ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing, Transportation and Logistics

Cyclic stochastic two-echelon inventory routing for an application in medical supply

Alexander Rave [a,b,*], Pirmin Fontaine [a,b], Heinrich Kuhn [a]

[a] Catholic University of Eichstätt-Ingolstadt, Ingolstadt School of Management, Auf der Schanz 49 85049 Ingolstadt, Germany
[b] Catholic University of Eichstätt-Ingolstadt, Mathematical Institute for Machine Learning and Data Science, Germany

ARTICLE INFO

Keywords:
Logistics
Inventory policies
Two-stage stochastic program
Adaptive large neighborhood search

ABSTRACT

Drug availability in clinics is essential for patient services, whose demand for medication is uncertain. Thus, clinics must have a variety of drugs available, leading to high inventory holding costs. In Germany, it is common for a larger central clinic to take over the procurement of drugs and distribute them to smaller surrounding clinics, which results in a two-echelon network structure. The clinics, however, operate according to their inventory policy as they plan independently. Additionally, the inventory policies include instant replenishment orders to avoid shortages, which can be executed by various vehicles, such as vans or aerial drones, because the orders only involve a few medications.

We present a two-stage stochastic program for a multi-product two-echelon inventory routing problem with stochastic demands. We decide on the cost-optimal cyclic delivery patterns and reorder points for the clinics with instant replenishment orders as recourse decision. Further, we introduce an adaptive large neighborhood search with problem-specific operators that modify the routing, delivery periods, and reorder points. We present a case study at a large German clinic that supplies multiple surrounding clinics and plans to integrate drone instead of van deliveries for emergency resupply. Our integrated approach leads to cost savings of 57% for the surrounding clinics and 18% for the central clinic. Using drone delivery compared to van delivery, the average stock of medication at surrounding clinics can be reduced, resulting in a total cost decrease of 29% while maintaining medication availability.

1. Introduction

To guarantee patient care, each clinic has a pharmacy that is appropriately stocked with medication. The pharmacy is under pressure to operate cost-effectively while providing full medical care to patients. This is particularly important since a clinic spends about 10 to 18% of its total expenses on medications (Nicholson et al., 2004; Volland et al., 2017). A way to reduce inventory costs is to make use of pooling effects. This means that inventory is stored more centrally to exploit the stochastic balancing effect between the independent demands in the individual clinics. As a consequence, additional transports are necessary. An aerial drone is an option to perform these transports since the drone has the particular advantage that it can carry a few medications to clinics faster and at lower costs than vans (Otto et al., 2018). In Ghana, for example, medicine is pooled in a central warehouse and then transported by drones to increase the overall availability of medication (Asadi et al., 2022). A similar situation exists in Germany; a larger central clinic is typically responsible for procuring medications and then distributing them to surrounding small clinics in a cyclic repeating manner (Quantum Systems, 2020).

For the latter example, two key questions arise. How many medications should be delivered to which clinic at what period, and what tours for delivery should be taken? This simultaneous inventory and supply optimization is known in the literature as the inventory routing problem (IRP) (e.g., Archetti & Ljubić, 2022). The delivery amount and delivery periods for the supply of clinics, however, depend on the inventory policy. In particular, clinics have economic advantages when choosing cyclic delivery patterns, e.g., they can improve the organization of their own clinic's supply chains. To avoid shortages, the additional supply of clinics is ensured via emergency deliveries. The standard deliveries include a high transport volume and thus are usually made with vans. The emergency deliveries, however, only include a few medications and can be performed by various vehicle types, for example, by drones (Quantum Systems, 2020). So far, the use of drones has not occurred in practice. Hence, this paper aims to assess the impact of drones as a new transportation option for instant replenishment orders.

* Corresponding author.
E-mail addresses: arave@ku.de (A. Rave), pirmin.fontaine@ku.de (P. Fontaine), heinrich.kuhn@ku.de (H. Kuhn).

https://doi.org/10.1016/j.ejor.2025.02.032
Received 22 July 2024; Accepted 24 February 2025
Available online 12 March 2025
0377-2217/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: Alexander Rave et al., *European Journal of Operational Research*, https://doi.org/10.1016/j.ejor.2025.02.032

Motivated by the previously introduced healthcare application, we develop a two-stage stochastic program for a cyclic multi-product two-echelon IRP with stochastic demand (2E-SIRP) that determines cost-optimal cyclic delivery periods and reorder points for both the central and the surrounding clinics. In addition, it determines the periods and routes for supplying the surrounding clinics. While these are decisions on the tactical plan of the first stage, in the second stage, instant replenishment orders performed by drones occur if the demand exceeds the inventory. Within this two-stage stochastic program, we take decisions that minimize the transport, expected inventory, and expected emergency delivery costs. To solve larger instances with a greater variety of products and scenarios, we develop a specialized adaptive large neighborhood search (ALNS) with problem-specific operators that modify the routing, delivery periods, or reorder points. Our ALNS further includes an algorithm that determines cost-optimal reorder points for each clinic individually if delivery periods or the routing are adjusted by operators. We benchmark the performance of the ALNS with optimally solved instances for the two-stage stochastic program and with instances of Archetti et al. (2007). Moreover, we generate managerial insights for a novel case study based on a large clinic in Germany. We present the benefits of our integrated planning approach as well as the influence of drones as new vehicle types for emergency deliveries on the inventory levels of different products and total cost reduction. Moreover, we show the value of a multi-product optimization on the resulting routing and reorder points.

We contribute to the literature as follows: Based on a real-life case study, we introduce a novel inventory routing problem variant that is applied to a (practical) medical supply problem for clinics. To tackle this problem, our paper combines an inventory routing problem with cyclic inventory policies on two echelons. These inventory policies further include instant replenishment orders to avoid shortages. Second, we formulate this problem setting as a two-stage stochastic program. Third, we develop an ALNS with problem-specific operators and an algorithm that determines cost-optimal reorder points for each clinic individually.

In the following, we describe the problem setting in detail (Section 2) and give a short literature review on inventory and routing planning (Section 3). We present the decision problem as a two-stage stochastic program in Section 4 and describe the ALNS in Section 5. In Section 6, we benchmark the performance of the ALNS and present the case study and managerial insights. Last, in Section 7, we summarize the results.

2. Problem setting

In this section, we first introduce the network structure of clinics in detail. In the second subsection, we present the inventory policy of clinics, and last, we describe the decision problem.

2.1. Network structure

Types and package sizes of products. The internal supply chain of the clinics handles and stores all kinds of medical goods like medicines, antibiotics, infusions, vaccines, essential medical aids, banked blood, etc., denoted as products. These products are generally sourced from the respective central distributor of these goods and then stored in the central clinic. The central clinic distributes these goods to the small surrounding clinics. The products are supplied, stored, and distributed in so-called packages, i.e., secondary packaging containing a certain quantity of individual units, pills, or blisters. The medical departments of the clinics, however, deliver these goods to their patients in individual consumer units (primary packaging). The package facilitates the handling of multiple consumer units in the supply chain and protects the individual products during picking and transportation (e.g., Broekmeulen et al., 2017; Wensing et al., 2018). One package contains a certain number of consumer units, denoted as package size. We assume that the regular deliveries contain complete packages of several products. The emergency deliveries, however, only contain consumer units of the products and in the quantity that cannot currently be fulfilled by their respective inventory level. Note that standard delivery is still the preferred delivery option due to the consolidation of clinics and products.

Central clinic (first echelon) supply. The central clinic – which has its own need for medications – procures the medical goods from the central distributor of these products, who is assumed to have sufficient inventory at all times. The demand for these medical goods for each clinic is independent of each other and uncertain but can be assumed to follow a known distribution (e.g., Johansson et al., 2020). Thus, shortages may arise at the central clinic if its inventory is too low. Due to the characteristics of the treatment of patients, the central clinic's missing inventory must be delivered in emergency deliveries ordered from a wholesaler. Note that emergency deliveries only have the purpose of satisfying urgent demand and do not additionally replenish inventories since standard deliveries are ordered from the central distributor. Emergency deliveries from the wholesaler only contain the currently missing products, which are quantified in consumer units. Standard deliveries from the distributor and emergency deliveries from the wholesaler are performed in commuting tours via vans. The standard delivery is executed at predefined cyclic periods, i.e., days, where different products are consolidated in a tour. The emergency delivery via van, on the other hand, can be executed whenever required.

Surrounding clinics (second echelon) supply. The central clinic supplies several surrounding clinics with the medical goods they need — both in standard deliveries by van and, if necessary, in emergency deliveries by drone. Due to long-term contracts, the central clinic is the only supplier of all small surrounding clinics. The surrounding clinics are supplied by multiple homogeneous vans on certain periods. Determining the number and exact days of the respective delivery periods is part of the optimization approach. The vans have a limited capacity and the supply is performed by a third-party logistics service provider. Therefore, it is not necessary to balance the workload of drivers between each period. In addition, the central clinic sets up an emergency delivery if shortages occur in the surrounding clinics. These deliveries are performed by vehicles specialized for fast deliveries. In our case, the fast deliveries are performed by drones (Quantum Systems, 2020). These drones are less expensive than delivery vans in terms of a single delivery trip, but they can only carry a few medications per flight due to limited capacity (Otto et al., 2018) — which is, however, mostly the case in emergency deliveries.

To sum up, the central clinic is supplied by a central distributor (standard delivery) and a wholesaler (emergency delivery). The surrounding clinics, on the other hand, are only supplied by the central clinic (standard and emergency delivery). There is no interaction between surrounding clinics and the central distributor or wholesaler.

The standard delivery for all clinics is only allowed on weekdays since a clinic's pharmacy is only staffed for emergency operations on the weekend. On the contrary, emergency deliveries can be conducted every period. Note that these are case-related assumptions, which can be easily relaxed and do not change our introduced algorithm.

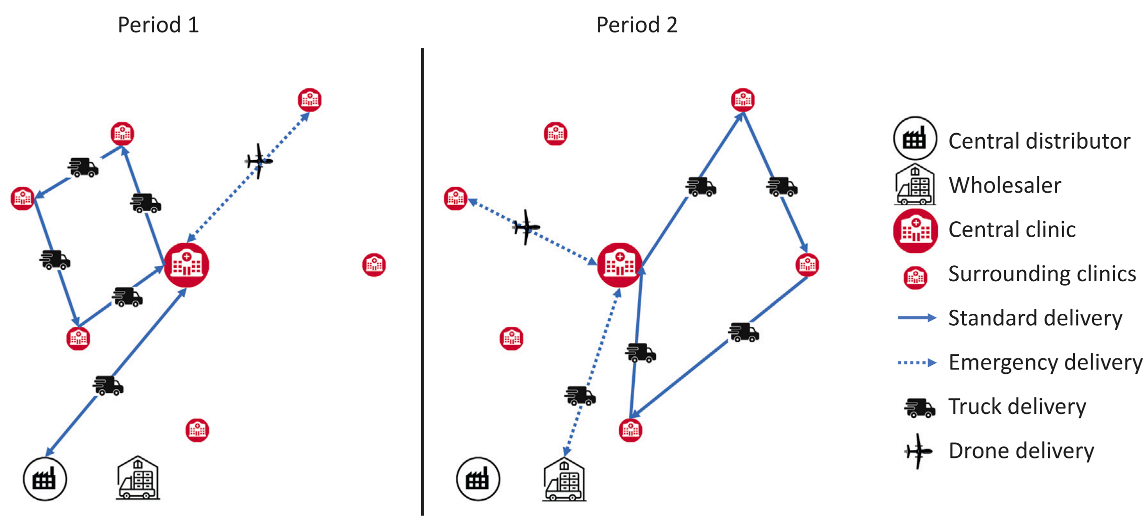

Fig. 1. Delivery system of medicine.

We assume an unlimited fleet of vans and drones. The vans and drones that visit the surrounding clinics have a limited capacity. On the other hand, the vans of the central distributor that supply the central clinic are assumed to have an unlimited capacity. If a drone's capacity would be exceeded in an emergency delivery, there are multiple drone flights within a period.

Fig. 1 presents an example network structure where the central clinic is fully responsible for delivering medication to six surrounding clinics. Three clinics are supplied in period 1, and the other three clinics in period 2 in standard delivery. Moreover, two emergency deliveries are carried out via drone due to shortages. The central clinic itself is supplied in period 1 via standard delivery by the central distributor and in period 2 via emergency delivery from the wholesaler.

2.2. Replenishment policy

The inventory of each product is replenished at each clinic using a periodic review inventory policy assuming a review period r and fixed package size q. This replenishment policy is, in general, denoted as (r, s, nq) inventory policy and frequently applied in the retailing industry (e.g., Broekmeulen et al., 2017; Wensing et al., 2018). Parameter s quantifies the reorder point that is used to trigger an order at a review moment, but only if the inventory is below the reorder point. No order is triggered in case of equality. nq indicates the order size, which must be an integer multiple n of the package size q. Note that the review period r can be defined as a time span with a fixed length or as a time span with different lengths within a cyclical, recurring planning period, e.g., a week. Assume a cyclic planning period of one week containing six working days (Monday to Saturday), then a review moment could possibly be, for example, each Tuesday, each Tuesday and Friday, or each Monday, Wednesday, and Friday, resulting in a fixed review period of six, three, and two days, respectively. However, this assumes a constant distance between reviews. The time intervals between review moments could also be different because of seasonal demand effects, uneven number of periods, or excluded review periods within the planning cycle. This results in so-called review or delivery patterns that are typically applied in retailing (e.g., Frank et al., 2021; Holzapfel et al., 2016). For example, review moments on Tuesday and Thursday of a six-day week result in a review interval of two and four days within the planning cycle. In our study, we assume this case, i.e., a fixed planning cycle and possibly several review moments within the cycle, which may have different time intervals.

Since we distinguish several products p, clinics c, and a planning cycle of several periods, t, we denote the considered replenishment policy as $(\dot{r}_{c,t}, s_{p,c}, n_{p,c} \cdot q_p)$, which includes additional emergency resupplies to prevent shortages. Typically, reorder points are set depending on the time. However, we have explicitly decided against this in our use case, i.e., the reorder points $s_{p,c}$ are independent of the period t, due to the practicability of the manual disposition that is applied in clinics.

The treatment of patients can only be rejected for very serious reasons and not if there is insufficient medication. Thus, if there is not enough medication in stock, this demand is satisfied by an emergency delivery. Note that the $(\dot{r}_{c,t}, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy with emergency resupply further includes the case of $s_{p,c} = 0$, where nothing is stored in general and delivered in standard delivery, and all demand is fulfilled via emergency deliveries. A characteristic of this problem with the focus on clinics is that the products are extremely varied, i.e., both daily and rarely needed (e.g., once every six months) drugs.

The demand for medication is any portion of a complete package, i.e., consumer units, but in a standard replenishment process, only complete packages are delivered. The emergency delivery, however, transports any package share since an emergency delivery is executed when the exact shortage is already known. In contrast, the target reorder points $s_{p,c}$ are quantified in complete packing units for practicability reasons.

Sequence of events for the replenishment cycle. The clinics' disposition and supply are differentiated according to whether a standard delivery arrives in the period. If there is a standard delivery planned in a period, the clinic first checks whether the inventory is below the reorder point $s_{p,c}$. If so, an order is placed for this product that includes whole packaging units and increases the inventory position at least to the level $s_{p,c}$. Note that this ordering process does not include any information on the demand of the current period. The further events are then identical for all periods. All clinics' demands become known at the same time, and the inventory position is adjusted. If there is not enough inventory available to meet the demand, an emergency order is placed (additionally to standard delivery, if necessary), which balances the inventory to zero. The resulting inventory level at the end of each period is used to quantify the clinic's inventory holding costs. The described scheduling and supply scheme of the clinics enables them to completely fulfill the respective demands of each clinic in all periods. Note that an efficient routing (for surrounding

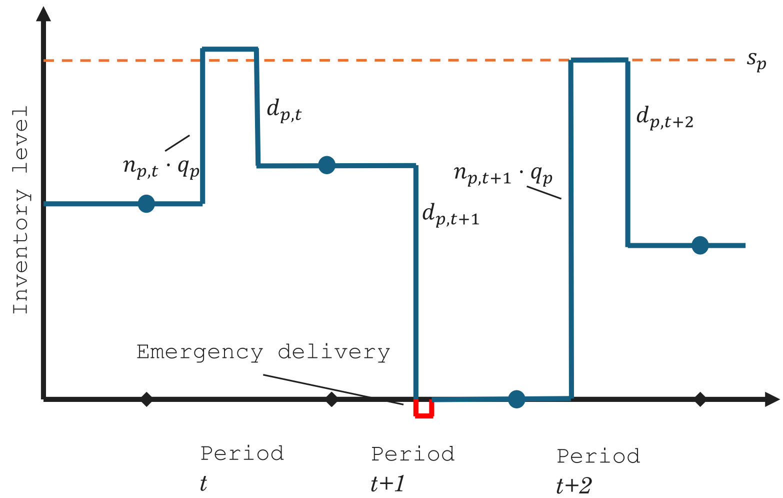

Fig. 2. Cyclic $(\dot{r}_{c,t}, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy of clinics.

clinics) also affects the inventory policies by reducing the transportation costs when the inventory is replenished, and thus, reorder points and routing must be optimized simultaneously. The only difference between the central clinic and the surrounding clinics – regarding their inventory policies – is that in the case of the central clinic, deliveries to the surrounding clinics also reduce inventory.

Fig. 2 presents an example inventory level of a drug for a single clinic for three periods. Standard delivery replenishes the inventory in period t and $t+2$ by the missing inventory at the end of the previous period compared to $s_{p,c}$ in complete packages. Afterwards, the inventory is reduced by the demand. In period $t+1$, there is no standard delivery, but demand exceeds the available inventory, and thus, an emergency order is placed. The emergency delivery only contains the missing inventory to fulfill the demand, and the inventory equals zero afterwards. Note that an emergency delivery might occur in the same period a standard delivery is conducted.

2.3. Decision problem

We determine cost-optimal cyclic delivery periods $\dot{r}_{c,t}$ and reorder points $s_{p,c}$ for all clinics and additionally the routing to serve all surrounding clinics. Then, emergency deliveries result from the delivery period and reorder point decisions to prevent unmet demand at one of the clinics. The inventory and supply of the central and surrounding clinics are optimized simultaneously as the inventory decisions of surrounding clinics affect the inventory of the central clinic.

We minimize routing costs as well as expected inventory (measured at the end of each period) and emergency delivery costs. Due to stochastic demand, we model this as a static stochastic program with a recourse decision in the form of emergency deliveries, which we formulate dependent on the scenario. We present a two-stage stochastic program that combines a two-echelon IRP with cyclic $(\dot{r}_{c,t}, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policies for all clinics.

The planning decision is made once for cyclic periods (in our case, one week) and then carried out over a longer planning horizon, e.g., three months, as long as there are no structural changes. Therefore, the cyclic inventory policy offers the advantage for clinics that the problem only has to be solved once, as long as no parameter values change structurally (e.g., seasonal fluctuations in demand). This is particularly advantageous for the smaller clinics, as there is a lack of personnel for extensive planning of, e.g., inventory. It follows that the considered problem setting is on a tactical level but also includes operational aspects. This is common for tactical planning problems (e.g., Arslan, 2021; Malicki & Minner, 2021; Rave et al., 2023).

We consider multiple products, as delivery patterns differ significantly from one another due to their differences in demand and holding costs. Moreover, different products are consolidated in standard and emergency delivery for both the surrounding and the central clinic, thus reducing the capacity of the respective vehicles. Products that have a low demand are only included in these tours, as we consider periodic review policies, where products are added if they fall below a reorder point. In addition, it has an impact on emergency deliveries. It follows that the simultaneous consideration of multiple products leads to a different solution than the consideration of each product separately. In Section 6.3.3, we show the impact of considering multiple products instead of each product separately. We neglect the limited shelf life of the products, as the inventory range of the given order quantity is well below the shelf life horizon of about two years. In addition, a strict first-in, first-out policy is followed in the clinics for inventory removal. The initial inventory of each product has a significant influence on the decisions to take in the cyclic planning and is thus dependent on the last period a clinic is supplied (e.g., Malicki & Minner, 2021).

3. Literature review

Originally introduced by Bell et al. (1983), there is a variety of publications on the IRP. For an extensive review on inventory routing, we recommend the paper of Coelho et al. (2014) and Roldán et al. (2017), and for inventory management in hospitals, the paper of Volland et al. (2017).

The problem considered also belongs to the class of periodic vehicle routing problems (PVRP). A wide range of publications are available concerning the PVRP (Campbell & Wilson, 2014). However, in our case, the clinics define clinic-specific cyclic delivery patterns, which constitute a defined combination of weekdays on which the individual clinics are supplied. Defining the delivery patterns also means determining the delivery frequency. The delivery frequency, however, impacts the volume per clinic delivery, affecting the associated cycle and safety stock in the central

ARTICLE IN PRESS

Table 1
Overview of inventory optimization literature.
[A] ALNS, * MILNP with additional linearization, ** depot has own stochastic demand, *** unmet demand is served via emergency delivery.

Paper	Assumptions		Inventory decisions			Methodology			
	multi-product	cyclic	limited inventory at depot	lost sales	back-orders	MILP	Markov chain	exact	heuristic
Higginson and Bookbinder (1995)							✓		
(Campbell et al., 1998), Campbell and Savelsbergh (2004)				✓					✓
Çetinkaya and Bookbinder (2003)					✓			✓	
Nolz et al. (2014)				✓		✓			✓[A]
Niakan and Rahimi (2015)	✓			✓		✓*			✓
Stenius et al. (2016)	✓		✓		✓			✓	
Yadollahi et al. (2017)		✓		✓		✓			
Rahimi et al. (2017)	✓				✓	✓*			✓
Jafarkhan and Yaghoubi (2018)	✓		✓			✓			✓
Rohmer et al. (2019)			✓			✓			✓[A]
Zheng et al. (2019)			✓			✓		✓	
Nikzad et al. (2019)	✓			✓		✓			✓
Johansson et al. (2020)	✓		✓		✓				✓
Malicki and Minner (2021)		✓		✓		✓			✓[A]
Raa and Aouam (2023)		✓		✓		✓			✓
Cui et al. (2023)					✓	✓		✓	
Sonntag et al. (2023)		✓			✓	✓		✓	✓
Our paper	✓	✓	✓**	***		✓			✓[A]

clinic and the surrounding clinics. Some recent publications quantify optimal delivery patterns in the area of grocery retailing. These approaches consider the supply of stores from a distribution center, taking inventory holding costs and joint delivery costs into account (e.g., Frank et al., 2021). However, these approaches neglect the two-echelon structure and only implicitly model stochastic demand. Taube and Minner (2018) also quantifies delivery patterns in the area of grocery retailing. They consider a classical joint replenishment problem with stochastic demand. In the following, we concentrate our literature review on contributions to inventory routing that focus, in particular, on problems with stochastic demand and time-based shipment consolidation.

Stochastic inventory routing — single product. When considering stochastic demands, there are typically lost sales or back-ordering, which state that unfulfilled demand can be met in a later period (Coelho et al., 2014). In our case, unfulfilled demand is served via emergency deliveries.

IRPs with stochastic demands (SIRP) are considered by Campbell et al. (1998), Campbell and Savelsbergh (2004). The authors focus on a heuristic solution method determining delivery periods and quantities in the first step and creating routes in the second step. Nolz et al. (2014) consider stochasticity in their SIRP through recourse decisions and propose an ALNS and two versions of an LNS for a bi-objective IRP. Cui et al. (2023) choose a robust optimization approach for a SIRP that is modeled with demand scenarios.

Yadollahi et al. (2017) and Malicki and Minner (2021) choose a representation with chance constraints considering certain service levels for safety stock calculation. To solve the problem, Malicki and Minner (2021) present a multi-start adaptive local search heuristic that runs multiple instances of an adaptive local search simultaneously and an ALNS for a cyclic IRP with stochastic demands. They show that setting the initial inventory endogenously in cyclic IRPs is sufficient. Raa and Aouam (2023) also consider a cyclic IRP with lost sales, but the delivery quantity depends on a base stock unless the capacity is insufficient. The authors develop a metaheuristic based on a first routing, second scheduling heuristic. Zheng et al. (2019) propose an exact algorithm based on the generalized benders decomposition for a SIRP, including an (T, S) inventory policy at the depot. Every cycle T, the inventory of the considered product is replenished to an amount of S.

Stochastic inventory routing — multiple products. Niakan and Rahimi (2015) and Rahimi et al. (2017) propose a formulation with fuzzy variables for the stochastic IRP in healthcare applications. Another healthcare application is considered by Jafarkhan and Yaghoubi (2018), who consider a SIRP for distributing blood cells, which are not in stock but can be substituted by other products. The authors propose a metaheuristic that makes use of two local searches in each iteration. Nikzad et al. (2019) determine the inventory and routing for serving medical centers with drugs. The authors present a two-stage stochastic program and a matheuristic for an IRP with application in medical supply. In contrast to our paper, among other things, the authors do not consider the inventory and supply of the depot, inventory replenishment policies, and emergency deliveries.

Time-based consolidation of customers. Time-based consolidation of customers, which means that orders are collected in a vehicle first until it departs, is considered first by Higginson and Bookbinder (1995). The vehicle departs if it is fully loaded, but it may leave earlier for cost reasons. This problem setting does not include a routing decision. Çetinkaya and Bookbinder (2003) extend this problem setting by deriving both a time-based and quantity-based consolidation policy. Stenius et al. (2016) derive an optimal solution framework for determining shipment intervals and reorder points for the time-based consolidation. Johansson et al. (2020) extend the article of Stenius et al. (2016) by deriving two approximation heuristics for large scale instances. While these publications did not include the routing to serve customers, Sonntag et al. (2023) combine the time-based shipment with an inventory routing problem, where additionally, each retailer considers an (T, S) inventory policy.

Delimitation to literature. Table 1 differentiates our assumptions, decisions, and methodology chronologically from the publications described before. Our paper adds to the available literature, particularly with respect to the following issues. First, we present a multi-product 2E-SIRP with the novel configuration of assumptions and decisions, especially by considering cyclic $(\dot{r}_{c,t}, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policies on both echelons. Further, we include instant replenishment orders as recourse decision. Second, we formulate a two-stage stochastic program and develop a specialized ALNS for solving larger instances. Third, we present a novel and real-life case study of a large German clinic.

Table 2
Index sets, parameters, and first and second stage variables.
* binary variables have a value of 1, if they are true, 0 else.

Index sets	
$\mathcal{C}$, $\hat{\mathcal{C}}$	Index set for clinics (including the central clinic $\hat{c}$)
$\mathcal{P}$	Index set for products
$\mathcal{T}$, $\mathcal{T}_0$	Index set for considered time periods (including period 0)
$\bar{\mathcal{T}}$	Index set for time periods in which standard delivery is permitted
Ω	Index set for scenarios
Parameters	
$k^S_{i,j}$	Costs traveling from clinic $i \in \hat{\mathcal{C}}$ to $j \in \hat{\mathcal{C}}$ in standard delivery
$\hat{k}^S$	Costs visiting the central clinic in standard delivery
$k^{em.}_c$	Emergency delivery costs visiting clinic $c \in \mathcal{C}$
k^I_p	Inventory holding costs for product $p \in \mathcal{P}$ per time period
$d_{p,c,t,\omega}$	Demand of product $p \in \mathcal{P}$ in clinic $c \in \hat{\mathcal{C}}$ at time $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
ε	Sufficiently small number
K^S	Maximum capacity for supply of surrounding clinics in standard delivery in volume equivalents
$K^{em.}$	Maximum capacity for the supply of surrounding clinics in emergency delivery in volume equivalents
M_l	Big M for $l = 1, \ldots, 5$ (detailed values for M_l can be found in Appendix A)
q^S_p	Consumer units, e.g. pills, of product $p \in \mathcal{P}$ that are included in a single package
v_p	Volume of a package of product $p \in \mathcal{P}$
First stage variables	
$\dot{r}_{c,t}$	Binary variable* indicating if clinic $c \in \hat{\mathcal{C}}$ is visited at period $t \in \bar{\mathcal{T}}$ with standard delivery
$s_{p,c}$	Integer variable, indicating the reorder point (entire packaging units) of each product $p \in P$ at clinic $c \in \hat{\mathcal{C}}$
$y_{i,j,t}$	Binary variable*, indicating, if a van travels from clinic $i \in \hat{\mathcal{C}}$ to clinic $j \in \hat{\mathcal{C}}$ for resupply at time $t \in \mathcal{T}$
$u_{i,t}$	Subtour-elimination variable for node $i \in \mathcal{C}$ at time $t \in \bar{\mathcal{T}}$ by Tucker and Zemlin (1960)
$I^{Init}_{p,c}$	Initial inventory of product $p \in \mathcal{P}$ at clinic $c \in \hat{\mathcal{C}}$
Second stage variables	
$I_{p,c,t,\omega}$	Positive real-value variable, indicating the inventory level of product $p \in \mathcal{P}$ for clinic $c \in \hat{\mathcal{C}}$ at the end of period $t \in \mathcal{T}_0$ and scenario $\omega \in \Omega$
$n^S_{p,c,t,\omega}$	Integer variable, indicating transported amount of packages of product $p \in \mathcal{P}$ to clinic $c \in \hat{\mathcal{C}}$ at time $t \in \bar{\mathcal{T}}$ and scenario $\omega \in \Omega$ in a standard delivery
$\phi^S_{i,j,t,\omega}$	Positive real-value variable, indicating accumulated transport volume from clinic $i \in \hat{\mathcal{C}}$ to clinic $j \in \hat{\mathcal{C}}$ at time $t \in \bar{\mathcal{T}}$ and scenario $\omega \in \Omega$ in a standard delivery
$q^{em.}_{p,c,t,\omega}$	Positive real-value variable, indicating transported amount of product $p \in \mathcal{P}$ to clinic $c \in \hat{\mathcal{C}}$ at time $t \in \mathcal{T}$ and scenario $\omega \in \Omega$ via emergency delivery
$w_{c,t,\omega}$	Integer variable, indicating the number of emergency resupplies for a clinic $c \in \hat{\mathcal{C}}$ at time $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$x_{p,c,t,\omega}$	Binary variable*, indicating, if there are emergency resupplies for product $p \in \mathcal{P}$ at clinic $c \in \hat{\mathcal{C}}$ at time $t \in \mathcal{T}$ and scenario $\omega \in \Omega$

4. Two-stage stochastic program

In this section, we introduce the mathematical model as a two-stage stochastic program for the 2E-SIRP. In Section 4.1, we describe the general problem setting, including notation and relevant sets, parameters, and variables, and in Section 4.2, we present the mathematical model.

4.1. Sets, parameters, and decision variables

We consider a set of $|\hat{\mathcal{C}}|$ nodes consisting both of the central clinic $\hat{c}$ as well as the set of surrounding clinics $\mathcal{C}$ with $\hat{\mathcal{C}} = \{\hat{c}\} \cup \mathcal{C}$. Moreover, we assume $|\mathcal{P}|$ products. $\mathcal{T} = \{1, \ldots, T\}$ is the index set for the periods with 1 and T being the first and last period and $\mathcal{T}_0 = \mathcal{T} \cup \{0\}$ further includes the initialization period 0. The standard delivery is permitted for all clinics in periods $\bar{\mathcal{T}} \subseteq \mathcal{T}$. Set Ω defines the demand scenarios.

General problem structure. We present the mathematical model in the form of a two-stage stochastic program. In the first stage, tactical decisions are taken on delivery days, routing to serve surrounding clinics and reorder points. In the second stage, emergency deliveries as recourse decisions are executed individually for each scenario. The emergency deliveries result from the tactical decisions in the first stage and are only executed to prevent a clinic from running out of stock.

Parameters. We consider travel costs $k^S_{i,j}$ per arc for standard delivery visiting the surrounding clinics, $\hat{k}^S$ for visiting the central clinic, $k^{em.}_c$ for emergency delivery as well as inventory holding costs k^I_p per period and product $p \in \mathcal{P}$.

For the supply of surrounding clinics, each van has a capacity of K^S and each drone of $K^{em.}$. For this, the volume v_p for each product is considered. q^S_p equals the consumer unit, e.g., pills, each package of product $p \in \mathcal{P}$ includes. Clinics $c \in \hat{\mathcal{C}}$ have a demand of $d_{p,c,t,\omega}$ for each product $p \in \mathcal{P}$, in period $t \in \mathcal{T}$, and scenario $\omega \in \Omega$.

Decision variables. The following decisions are taken for all demand scenarios consolidated and thus are first-stage decisions. The main decision variables determine the delivery periods $\dot{r}_{c,t}$ for clinics $c \in \hat{\mathcal{C}}$ and period $t \in \bar{\mathcal{T}}$ and the reorder points $s_{p,c}$ for product $p \in \mathcal{P}$ and clinic $c \in \hat{\mathcal{C}}$. Please note that $\dot{r}_{c,t}$ specifies the delivery period and not the delivery intervals. Moreover, variable $y_{i,j,t}$ decides on the routing in standard delivery, i.e., if a van travels from clinic $i \in \hat{\mathcal{C}}$ to clinic $j \in \hat{\mathcal{C}}$ in period $t \in \bar{\mathcal{T}}$.

The second stage defines the following directly related variables. For period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$, $I_{p,c,t,\omega}$ defines the inventory levels for each product $p \in \mathcal{P}$ in clinic $c \in \hat{\mathcal{C}}$, $n^S_{p,c,t,\omega}$ the delivery amount in packages in standard delivery, $w_{c,t,\omega}$ the number of emergency deliveries, and $q^{em.}_{p,c,t,\omega}$ the transported amount in emergency deliveries (see Table 2).

4.2. Mathematical model

Objective function.

$$\min Z^{total} = \sum_{t\in\bar{\mathcal{T}}} \hat{k}^S \cdot \dot{r}_{\hat{c},t} + \sum_{i,j\in\hat{C},t\in\bar{\mathcal{T}}} k^S_{i,j} \cdot y_{i,j,t} + \mathbb{E}_{\omega\in\Omega}\left\{ \sum_{c\in\hat{C},t\in\mathcal{T}} k^{em.}_c \cdot w_{c,t,\omega} + \sum_{p\in\mathcal{P},c\in\hat{C},t\in\mathcal{T}} k^I_p \cdot I_{p,c,t,\omega} \right\} \tag{1}$$

The objective function (1) minimizes the expected total costs in the time periods considered. It contains two major cost terms. The first applies to all scenarios and quantifies the delivery costs for standard deliveries for the central clinic (first sum) and routing costs to serve surrounding clinics (second sum). The second term includes the expected emergency delivery costs for supplying all clinics (third sum) and expected inventory holding costs for all clinics (fourth sum).

Routing constraints.

$$\sum_{i\in\hat{C}} y_{i,c,t} = \dot{r}_{c,t} \qquad \forall c\in C, t\in\bar{\mathcal{T}} \tag{2}$$

$$\sum_{i\in\hat{C}} y_{j,i,t} = \sum_{i\in\hat{C}} y_{i,j,t} \qquad \forall j\in\hat{C}, t\in\bar{\mathcal{T}} \tag{3}$$

$$\sum_{i,j\in\hat{C}} y_{i,j,t} \le M_1 \cdot \sum_{j\in C} y_{\hat{c},j,t} \qquad \forall t\in\bar{\mathcal{T}} \tag{4}$$

$$u_{i,t} + 1 \le u_{j,t} + M_1 \cdot (1 - y_{i,j,t}) \qquad \forall i,j\in C, t\in\bar{\mathcal{T}} \tag{5}$$

$$\sum_{j\in\hat{C}} (\phi^S_{j,c,t,\omega} - \phi^S_{c,j,t,\omega}) = \sum_{p\in\mathcal{P}} v_p \cdot n^S_{p,c,t,\omega} \cdot q^S_p \qquad \forall c\in C, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{6}$$

$$\phi^S_{i,j,t,\omega} \le K^S \cdot y_{i,j,t} \qquad \forall i,j\in\hat{C}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{7}$$

$$\sum_{t\in\bar{\mathcal{T}}} \dot{r}_{c,t} \ge 1 \qquad \forall c\in\hat{C} \tag{8}$$

$$y_{i,j,t} \in \{0,1\} \qquad \forall i,j\in\hat{C}, t\in\bar{\mathcal{T}} \tag{9}$$

$$\dot{r}_{c,t} \in \{0,1\} \qquad \forall c\in\hat{C}, t\in\bar{\mathcal{T}} \tag{10}$$

$$\phi^S_{i,j,t,\omega} \in \mathbb{R}^+ \qquad \forall i,j\in\hat{C}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{11}$$

$$n^S_{p,c,t,\omega} \in \mathbb{N} \qquad \forall p\in\mathcal{P}, c\in C, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{12}$$

$$u_{c,t} \in \mathbb{R} \qquad \forall c\in C, t\in\bar{\mathcal{T}} \tag{13}$$

Constraints (2) to (13) restrict the routing of the standard delivery in each period $t\in\bar{\mathcal{T}}$. Constraints (2) define the delivery periods when a surrounding clinic is visited. Further, these constraints ensure that each clinic is served a maximum of once per period. Next, Constraints (3) conserve flow, and Constraints (4) ensure that the van that visits surrounding clinics always starts its tour at the central clinic. Subtours are eliminated by Constraints (5) (Tucker & Zemlin, 1960). Constraints (6) define the accumulated transported volume of each van's tour. This accumulated transported volume must be less than or equal to the van's capacity (Constraints (7)). Constraints (8) ensure that each surrounding clinic is visited at least once in the considered periods. Last, variables are defined.

Reorder point constraints.

$$n^S_{p,c,t,\omega} \cdot q^S_p \le s_{p,c} - I_{p,c,t-1,\omega} + q^S_p - \varepsilon \qquad \forall p\in\mathcal{P}, c\in\hat{C}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{14}$$

$$n^S_{p,c,t,\omega} \cdot q^S_p \ge s_{p,c} - I_{p,c,t-1,\omega} - M_2 \cdot (1 - \dot{r}_{c,t}) \qquad \forall p\in\mathcal{P}, c\in\hat{C}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{15}$$

$$n^S_{p,c,t,\omega} \cdot q^S_p \le M_2 \cdot \dot{r}_{c,t} \qquad \forall p\in\mathcal{P}, c\in\hat{C}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{16}$$

$$s_{p,c} \in \mathbb{N} \qquad \forall p\in\mathcal{P}, c\in\hat{C} \tag{17}$$

$$I_{p,c,t,\omega} \in \mathbb{R}^+ \qquad \forall p\in\mathcal{P}, c\in\hat{C}, t\in\mathcal{T}_0, \omega\in\Omega \tag{18}$$

Constraints (14) to (18) define the reorder points $s_{p,c}$ and the resulting delivery quantities $n^S_{p,c,t,\omega} \cdot q^S_p$ for all clinics. Constraints (14) limit the delivery quantity to the reorder points, taking the inventory of the previous period into account such that the number of complete packages equals the reorder point. As the reorder points only measure the stock of complete packages, but the remaining stock of an open package, i.e., some consumer units, may still be present, the reorder points may be slightly exceeded by less than one package. This overlap is ensured by a term added on the right side ($q^S_p - \varepsilon$), representing all units of a package (q^S_p) with one consumer unit (ε) missing. Constraints (15) ensure that a certain number of packages are ordered to refill the inventory, which is only applied if a delivery occurs during this period. Further, drugs are only delivered in a period of standard delivery (Constraints (16)).

Emergency delivery constraints.

$$q^{em.}_{p,c,t,\omega} \le d_{p,c,t,\omega} - I_{p,c,t-1,\omega} + M_3 \cdot (1 - x_{p,c,t,\omega}) - \begin{cases} n^S_{p,c,t,\omega} \cdot q^S_p & ,t\in\bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p\in\mathcal{P}, c\in C, t\in\mathcal{T}, \omega\in\Omega \tag{19}$$

$$q^{em.}_{p,\hat{c},t,\omega} \le d_{p,\hat{c},t,\omega} - I_{p,\hat{c},t-1,\omega} + M_3 \cdot (1 - x_{p,\hat{c},t,\omega}) + \sum_{k\in C} q^{em.}_{p,k,t,\omega} + \begin{cases} \left(\sum_{k\in C} n^S_{p,k,t,\omega} - n^S_{p,0,t,\omega}\right) \cdot q^S_p & ,t\in\bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p\in\mathcal{P}, t\in\mathcal{T}, \omega\in\Omega \tag{20}$$

$$\frac{1}{M_4} \cdot q^{em.}_{p,c,t,\omega} \le x_{p,c,t,\omega} \le w_{c,t,\omega} \qquad \forall p\in\mathcal{P}, c\in\hat{C}, t\in\mathcal{T}, \omega\in\Omega \tag{21}$$

$$w_{c,t,\omega} \le M_4 \cdot \sum_{p\in\mathcal{P}} x_{p,c,t,\omega} \qquad \forall c \in \hat{C}, t \in \mathcal{T}, \omega \in \Omega \tag{22}$$

$$\sum_{p\in\mathcal{P}} q^{em.}_{p,\hat{c},t,\omega} \le M_4 \cdot w_{\hat{c},t,\omega} \qquad \forall t \in \mathcal{T}, \omega \in \Omega \tag{23}$$

$$\sum_{p\in\mathcal{P}} v_p \cdot q^{em.}_{p,c,t,\omega} \le K^{em.} \cdot w_{c,t,\omega} \qquad \forall c \in C, t \in \mathcal{T}, \omega \in \Omega \tag{24}$$

$$w_{c,t,\omega} \in \mathbb{N} \qquad \forall c \in \hat{C}, t \in \mathcal{T}, \omega \in \Omega \tag{25}$$

$$x_{p,c,t,\omega} \in \{0,1\}, q^{em.}_{p,c,t,\omega} \in \mathbb{R}^+ \qquad \forall p \in \mathcal{P}, c \in \hat{C}, t \in \mathcal{T}, \omega \in \Omega \tag{26}$$

Constraints (19) to (26) define the emergency delivery constraints for all clinics. Constraints (19) and (20) restrict the delivery quantity to the unsatisfied demand of the current period to avoid emergency deliveries enabling stockpiling. Constraints (21) and (22) link variables $q^{em.}_{p,c,t,\omega}$, $x_{p,c,t,\omega}$, and $w_{c,t,\omega}$, i.e., if there is a quantity delivered in emergency delivery, then there is an emergency delivery (Constraints (21)). Vice versa, Constraints (22) ensure that if there are emergency deliveries, then products are delivered. Please note that multiple products are consolidated in one emergency delivery if the capacity is sufficient. Constraints (23) determine if there is an emergency delivery for the central clinic. Constraints (24) restrict the capacity of emergency deliveries for surrounding clinics. Last, variables are defined.

Inventory balancing constraints.

$$I_{p,c,t,\omega} = I_{p,c,t-1,\omega} - d_{p,c,t,\omega} + q^{em.}_{p,c,t,\omega} + \begin{cases} n^S_{p,c,t,\omega} \cdot q^S_p & ,t \in \bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p \in \mathcal{P}, c \in C, t \in \mathcal{T}, \omega \in \Omega \tag{27}$$

$$I_{p,\hat{c},t,\omega} = I_{p,\hat{c},t-1,\omega} - d_{p,\hat{c},t,\omega} + q^{em.}_{p,\hat{c},t,\omega} - \sum_{k\in C} q^{em.}_{p,k,t,\omega} + \begin{cases} n^S_{p,\hat{c},t,\omega} \cdot q^S_p - \sum_{k\in C} n^S_{p,k,t,\omega} \cdot q^S_p & ,t \in \bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p \in \mathcal{P}, t \in \mathcal{T}, \omega \in \Omega \tag{28}$$

Constraints (27) to (28) are inventory balancing constraints for all clinics. Constraints (27) are the inventory balancing constraints for surrounding clinics for periods $t \in \mathcal{T}$ defining the inventory at the end of each period $t \in \mathcal{T}$ by reducing the inventory of the previous period by the demand and adding delivery quantities. The same is applied to the central clinic (Constraints (28)), but the inventory is additionally reduced by the total quantity delivered to surrounding clinics in standard and emergency deliveries.

Setting up the initial inventory.

$$I^{Init}_{p,c} = \max\left(\min_{\tau\in\bar{\mathcal{T}}}\left(s_{p,c} + M_5 \cdot \sum_{t=\tau+1}^{|\bar{\mathcal{T}}|} \dot{r}_{c,t} - \sum_{t=\tau}^{|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,c,t,\omega}\right\}\right), 0\right) \qquad \forall p \in \mathcal{P}, c \in C \tag{29}$$

$$I^{Init}_{p,\hat{c}} = \max\left(\min_{\tau\in\bar{\mathcal{T}}}\left(s_{p,\hat{c}} + M_5 \cdot \sum_{t=\tau+1}^{|\bar{\mathcal{T}}|} \dot{r}_{\hat{c},t} - \sum_{t=\tau}^{|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,\hat{c},t,\omega}\right\} - \sum_{t=\tau}^{|\bar{\mathcal{T}}|} \mathbb{E}_{\omega\in\Omega}\left\{\sum_{c\in C} n^S_{p,c,t,\omega} \cdot q^S_p\right\}\right), 0\right) \qquad \forall p \in \mathcal{P} \tag{30}$$

$$I_{p,c,0,\omega} = I^{Init}_{p,c} \qquad \forall p \in \mathcal{P}, c \in \hat{C}, \omega \in \Omega \tag{31}$$

$$I^{Init}_{p,c} \in \mathbb{R}^+ \qquad \forall p \in \mathcal{P}, c \in \hat{C} \tag{32}$$

Constraints (29) to (32) set up the initial inventory. Constraints (29) model the complete cycle on average for the surrounding clinics. The initial inventory is expected to be equal to the reorder point and is reduced by the expected average demand that arises after the last standard delivery arrives. To ensure that only the expected demand after the last delivery is taken into account, it is required to evaluate all periods $\tau \in \bar{\mathcal{T}}$. If a period is chosen before the last delivery arises, we add M_5 or multiples of M_5 to ensure that this period does not set $I^{Init}_{p,c}$. The initial inventory then equals the minimum inventory of all periods $\tau \in \bar{\mathcal{T}}$. However, it might be that the expected demand exceeds the reorder point, i.e., it is more cost-efficient to deliver the required medication via emergency delivery than to keep it in stock. For this case, the initial inventory is set to zero, which is the intuition behind the maximum function. The same is applied to the central clinic (Constraints (30)), but additionally, the expected transported medicine in standard delivery is considered that arises after the central clinic was supplied last.

These constraints are non-linear. Thus, we linearize them in Appendix B. We present the non-linear constraints here as these are easier to understand; however, in the numerical study, we implement the linearized constraints, as shown in the appendix.

5. Adaptive large neighborhood search

To solve the 2E-SIRP, we focus on the relevant decisions: delivery days, routing for surrounding clinics, and reorder points. The problem is that delivery days and reorder points are strongly dependent on each other, so a change in one also affects the other. Considering this dependency, we develop an ALNS with operators that modify delivery days or routing and adjust the reorder points in the same iteration. To adjust the reorder points, the ALNS includes an algorithm specifically developed to determine cost-optimal reorder points for each clinic and product individually. However, it is too time-consuming to consider dependencies between the clinics, and therefore, the globally optimal reorder points are not necessarily determined by this algorithm. Thus, the ALNS further includes an operator that slightly adjusts reorder points.

The structure of this section is as follows: First, we introduce the ALNS and describe the solution representation in Section 5.1, the pseudo-code of our ALNS in Section 5.2, our operators in Section 5.3, and the algorithm that determines optimal reorder points in Section 5.4.

5.1. Solution representation

A solution is represented by the clinics' cyclic delivery periods ($\dot{r}_{c,t}$), the routing of the standard delivery (R), including information on cyclic repeating delivery periods, and the reorder points for the clinics ($s_{p,c}$). In each iteration, operators are chosen that take changes on $\dot{r}_{c,t}$, R, or else $s_{p,c}$ (see Section 5.3). Additionally, if $\dot{r}_{c,t}$ or R are adjusted, the reorder points of the affected clinics are adjusted accordingly (see Section 5.4).

Within these changes, infeasible solutions are accepted but penalized in the objective function f dependent on the degree of infeasibility (e.g., Rave et al., 2023; Vidal et al., 2013; Voigt, 2024). Infeasibility may occur in two different ways: first, a van may exceed its capacity in a scenario, and second, not all clinics are visited at least once. Appendix C explains how we account for penalty costs in our objective function f.

5.2. ALNS algorithm

The general outline of the ALNS is presented in Algorithm 1. The algorithm begins with an initial solution (line 1) consisting of the delivery periods $\dot{r}_{c,t}$, the routing to serve surrounding clinics, R, and all clinics' reorder points $s_{p,c}$. This solution equals the global best solution ($\dot{r}^{global}_{c,t}$, R^{global}, $s^{global}_{p,c}$) in the beginning. The initial solution is generated by visiting all clinics once and only in the first period. The clinics are inserted one by one in the greedy best way in one tour.

Algorithm 1: ALNS framework for the cyclic inventory and routing planning

1 ($\dot{r}_{c,t}$,$\dot{r}^{global}_{c,t}$,R, R^{global}, $s_{p,c}$, $s^{global}_{p,c}$) ← Initial Solution;
2 $Temp$ ← GetInitialTemp();
3 **while** $Time < MaxTime$ **do**
4   Choose(Operator);
5   **if** Operator *modifies* $\dot{r}_{c,t}$ *or* R **then**
6     $\bar{r}_{c,t}$, $\bar{R}$ ← Operator($\dot{r}_{c,t}$, R);
7     $\bar{s}_{p,c}$ ← Optimize($s_{p,c}$); *//see Algorithm 2*
8   **else**
9     $\bar{s}_{p,c}$ ← Operator($s_{p,c}$);
10   **end**
11   Update($n^S_{p,c,t,\omega}$, $I_{p,c,t,\omega}$, $q^{em.}_{p,c,t,\omega}$);
12   **if** $f(\bar{r}_{c,t}, \bar{R}, \bar{s}_{p,c}) < f(\dot{r}^{global}_{c,t}, R^{global}, s^{global}_{p,c})$ **then**
13   ($\dot{r}_{c,t}$, $\dot{r}^{global}_{c,t}$, R, R^{global}, $s_{p,c}$, $s^{global}_{p,c}$) ← ($\bar{r}_{c,t}$, $\bar{R}$, $\bar{s}_{p,c}$);
14   **else**
15   **if** *accept*($f(\bar{r}_{c,t}, \bar{R}, \bar{s}_{p,c})$, $f(\dot{r}_{c,t}, R, s_{p,c})$, $Temp$) **then**
16   ($\dot{r}_{c,t}$, R,$s_{p,c}$) ← ($\bar{r}_{c,t}$, $\bar{R}$, $\bar{s}_{p,c}$);
17   **end**
18   **end**
19   Temp ← UpdateTemp($Temp$);
20   UpdateWeights();
21 **end**

In line 2, a temperature $Temp$ for the simulated annealing process is set. The while-loop in line 3 stops after a maximum time. At the beginning of each iteration, operators are chosen adaptively (line 4). If the selected operators modify the clinics' delivery periods or the routing to serve surrounding clinics (line 6), then in the next step, the reorder points of affected clinics are adjusted for each product accordingly (line 7). The algorithm to adjust the reorder points is explained in detail in Section 5.4. If an operator is chosen that modifies the reorder points, this reorder point algorithm is not executed (lines 8–10). Next, independent of the chosen operators, delivery quantities $n^S_{p,c,t,\omega}$, inventories $I_{p,c,t,\omega}$, and emergency delivery quantities $q^{em.}_{p,c,t,\omega}$ are updated (line 11). This is only applied to clinics that are affected by the operators' changes.

The simulated annealing process is described in lines 12–18. The costs f are based on changes made by operators and Algorithm 2. The solution is accepted if it is better than the previous global solution (lines 12–13). Depending on the temperature $Temp$, the new solution ($\bar{r}_{c,t}$, $\bar{R}$, $\bar{s}_{p,c}$) is accepted (lines 14–17) if it is worse than the global best solution. Last, the temperature and the weights of the operators are updated, dependent on the performance of this iteration (Sacramento et al., 2019).

5.3. Operators

We consider three different types of operators: Operators that modify the central clinics' delivery periods, operators that modify the routing of surrounding clinics, and operators that modify the reorder points. In each iteration, only one type of operator is chosen, i.e., in a single iteration, there are no operators chosen from both Sections 5.3.1 and 5.3.2, or 5.3.3.

5.3.1. Operators modifying the delivery period of the central clinic

If operators modify the delivery period of the central clinic, one of the following operators or both are chosen. A detailed description of both operators can be found in Appendix D.

Lowest standard quantity dependent delivery period removal - with noise (central clinic) - The delivery period is chosen in which there is the lowest total quantity delivered to the central clinic in standard delivery (including a certain noise), and this period's standard delivery to the central clinic is removed.

Largest emergency quantity dependent delivery period insertion - with noise (central clinic) - The delivery period is chosen in which there is the largest quantity delivered (based only on emergency deliveries) to the central clinic (including a certain noise), and the central clinic is served at this period.

5.3.2. Operators modifying the routing of surrounding clinics

The following operators modify the routing R for surrounding clinics' service and are extensions of the operators of Røpke and Pisinger (2006), Pisinger and Røpke (2007), Aksen et al. (2014), and Rave et al. (2023). We consider two operators that remove α clinics, three operators that insert these removed clinics, and one operator that shifts the delivery period. α is randomly drawn (uniform) with a limited number of clinics.

Random removal - This operator removes α clinics from any route and delivery period.

Closest removal - The closest removal operator removes a random clinic and, at maximum $\alpha - 1$ closest clinics of all or one specific route/s within the same period.

The following operators insert the removed clinics to the delivery period with the largest quantity of products transported in emergency delivery, taking the historical operators' performance for selecting each period and a noise into account. Appendix E presents the determination of the delivery periods in detail.

Greedy insertion - This operator adds clinics one by one to existing van tours at the selected period in a greedy best way, taking traveled distances into account. If there has not been a route on this period, a new route is created. Clinics might be inserted in any routes, as in the next step in the ALNS, the delivery quantity is adjusted.

Random insertion - This operator works in a similar way as the previous operator but adds clinics to a random place within existing van routes.

Greedy insertion - new route - Similar to the greedy insertion operator, this operator always inserts clinics in one new route at the selected period.

When applying a removal operator and an insertion operator, the number of clinic servings stays the same. To reduce the number of servings, we only apply a removal operator in an iteration. In order to increase the number of servings, we apply a removal operator that, however, only selects clinics and does not remove them. Afterwards, we apply an insertion operator to insert the selected clinics one by one in random order. Further, we consider an operator that shifts certain routes to the selected period. When this operator is selected, no additional removal or insertion operator is applied.

Delivery period shift - surrounding clinics - The delivery period of one randomly selected route is shifted to the selected period. If clinics have already been visited at that period, these clinics are removed from this shifted route.

5.3.3. Operator modifying the reorder points

Last, we consider an operator that modifies the reorder points without making any changes to the delivery periods or the routing. We need this operator as the reorder point algorithm described in the next section only gives us the cost-optimal reorder point for each clinic and product individually. The difficulty in reorder point computation arises especially from the dependencies between both echelons (central clinic vs. surrounding clinics) that make the problem significantly harder and more time-consuming to solve to (global) optimality.

Increase or decrease a clinic's reorder point - A random clinic's and product's reorder point is either increased or decreased by a random (integer) number of up to 10% of its value. The direction of the change depends on the number of emergency deliveries that the central clinic and the surrounding clinics have. A detailed description of this operator can be found in Appendix F.

5.4. Reorder point algorithm

Algorithm 2 describes the pseudo-code of the reorder point algorithm for calculating the reorder points $s_{p,c}$ when there was a change in a clinic's delivery periods and is executed for each affected clinic and product separately. For this, we use the integrality of $s_{p,c}$ to compute reorder points $s_{p,c}$ that are locally optimal for each clinic and product individually. This means that for the given delivery periods $\dot{r}_{c,t}$, $s_{p,c}$ is set cost-optimally without taking the two-echelon structure, i.e., the dependency of the central clinic on the surrounding clinics, and the van's payload into account. Determining globally optimal reorder points, on the other hand, would be very time-consuming as far as determinable and would significantly slow down the algorithm. To circumvent the dependencies between both echelons, the ALNS includes an operator that also modifies clinics' reorder points (Section 5.3.3).

Algorithm 2: Reorder point algorithm to determine $s_{p,c}$ on a given $\dot{r}_{c,t}$

```
1 Initialize (s^new_{p,c}, s_{p,c}, ṙ_{c,t});
2 while f(ṙ_{c,t}, s^new_{p,c}) < f(ṙ_{c,t}, s_{p,c}) do
3   | s_{p,c} ← s^new_{p,c};
4   | s^new_{p,c} ← adjust(s^new_{p,c});
5 end
```

In the beginning, a reorder point $s^{new}_{p,c}$, the best reorder point $s_{p,c}$ found so far, and the considered delivery periods $\dot{r}_{c,t}$ are initialized (line 1). As the initial reorder point, the reorder point of the delivery period before modification by operators is chosen. While there are still cost improvements (line 2) by adjusting $s^{new}_{p,c}$, $s_{p,c}$ is updated (line 3), and $s^{new}_{p,c}$ is either increased or decreased by one dependent on which change has a larger benefit (line 4). Please note that the direction of adjustment (increase or decrease) is only determined in the first iteration. In each further iteration, reorder points are adjusted in the same direction as long as it improves the solution. The sequence of clinics this algorithm is applied to is randomly selected before.

6. Numerical study

In this section, we show the performance of our ALNS and present a case study based on the project MEDinTime (Quantum Systems, 2020). First, in Section 6.1, we describe the numerical setup in detail. In Section 6.2, we show the performance of our ALNS compared to optimal results solving the two-stage stochastic program in CPLEX and to optimal solutions for IRP instances of Archetti et al. (2007). Further, we show the stability of the solution we found. In Section 6.3, we present results for the case study and perform multiple sensitivity analyses. The mathematical model is implemented in OPL and solved using CPLEX v12.10. and the ALNS is implemented in C++. All experiments are conducted on an AMD Ryzen 9 5950X with 128 GB RAM. The used instances are available at: https://github.com/FontainePirmin/2E-SIRP

6.1. Numerical setup

Clinics. The central clinic, Ingolstadt Hospital, supplies nine surrounding clinics ($|\hat{C}| = 10$) located between 3 and 75 km (euclidean distance) from the central clinic. The clinics are supplied in a weekly repeating manner ($|\mathcal{T}| = 7$), and all clinics are in the range of drone delivery.

Table 3
Parameter values for product categories.

$p \in \mathcal{P}$	Price ranges [€]	k_p^I [€]	v_p	Average mean of demand per period	Average variance of demand per period
1	< 12	0.0033	2	1.29	6.43
2	12–58	0.0193	2	0.21	1.07
3	59–129	0.0514	1	0.09	0.43
4	130–291	0.1153	1	0.06	0.32
5	292–778	0.2933	1	0.04	0.21
6	779–1758	0.6951	0.4	0.02	0.11
7	> 1759	2.4644	0.4	0.01	0.06

Cost structure. Various medications are needed for patient care, which we categorize by their sales price. The annual inventory holding costs k_p^I are approximated by 20% of the average sales price of the corresponding category. Based on the information given in our case study, standard delivery costs k^S are assumed to be 1.4€per km while drone delivery is 52% less costly and thus equals 0.672€per km. Each inventory replenishment at the central clinic is assumed to have delivery costs of 168€for a standard delivery ($\hat{k}^S$) and twice as high for an emergency delivery ($\hat{k}^{em.}$).

A standard supply for surrounding clinics includes approximately 200 different products. The costs for transporting the products are distributed evenly across all products, regardless of whether the products are fast or slow movers. Therefore, depending on the number of products considered, the transport costs are set in proportion to the 200 products. We consider only a portion of the products that is based on the products' cost prices. Considering only a share of these products leads to similar decisions but can save sufficient run-time. This proportional cost adjustment is applied to all standard delivery and emergency delivery costs, as different products might be consolidated for each delivery.

Representation of inventory. We consider a package size of 100 consumer units and count the stock from the whole pack view as this is done in practice, too. It follows that we assume a package size of $q_p^S = 1$ for all products $p \in \mathcal{P}$. If the inventory consists of, for example, three full packs and 27 remaining consumer units, then $I_{p,c,t,\omega} = 3.27$.

Demand. The (daily) demand for medication is assumed to be gamma-distributed with varying mean and variance for each product and clinic, as the demand is not negative and can be any share of a product. However, the demand is rounded to two decimal places as consumer units cannot be divided arbitrarily. Due to insufficient data available, the variance is taken from Johansson et al. (2020) and assumed to be five times as high as the mean of the distribution. Additionally, the demand for medications varies for each weekday, with its peak on Wednesday, and is especially lower at the weekend, as clinics have a 30% lower bed occupancy at the weekend. This is because patients in Germany are increasingly discharged at the weekend, and new patients mostly arrive at the beginning of the week.

Scenarios. Due to the uncertain demand of multiple products, the number of scenarios is growing exponentially with the number of products. We use Latin Hypercube Sampling (McKay et al., 1979) to sample a subset of these scenarios based on the multi-dimensional distribution of the demands. The distribution is divided into intervals with equal probability and then displayed in a hypercube. From this hypercube, $|\Omega|$ random demands are drawn for each product so that rows and columns do not overlap. This procedure is repeated for all hospitals and time periods. We choose a number of 100 scenarios and one product of each product category, i.e., the instances only vary in their demand scenarios. According to a sequential test (Wald, 1947), 100 scenarios are sufficient as a maximum desired deviation of 10% for a single criteria is not exceeded for a confidence interval of 95%. The tested criteria are the total costs, the number of emergency deliveries, and the average reorder points.

Products and capacities. Given the clinics' product range, we consider seven exemplary products, each in a different price category with varying demand. We assume that drugs that are high in demand have larger volume equivalents per package. The van's capacity is based on the Mercedes Sprinter (Mercedes-Benz, 2022) and reduced proportionally to the drugs under consideration so that the total average weekly demand of up to five clinics can be carried in each van's tour. This means that the van's capacity K^S depends on the considered drugs (e.g., $K^S = 138.7$ volume equivalents if all seven products are considered). The considered drone has a capacity of 3.765 volume equivalents ($K^{em.}$).

Table 3 presents the parameter settings for seven different products that we categorized by their price ranges. The table reports the resulting inventory holding costs per period k_p^I, their volume v_p, and their all clinics' average mean and variance of demand per period. To gather the demand for each clinic individually, the values need to be adjusted based on each clinic's size, i.e., the bed number. Please note that the products' volumes are only included due to the vans' and drones' limited capacities. Moreover, we found in the experiments that varying the volume only has minor influences.

For benchmark tests, we consider small and medium-sized instances with four, seven, or ten clinics, one to three products, and one to ten demand scenarios. These combinations result in a total of 520 instances. However, we consider only a few combinations with four and seven clinics, prioritizing customer sizes of ten as in our case study. For the detailed instance setting, we refer to the results in Table 4 where the number of instances per setting is shown in the column of optimally solved instances.

ALNS parameter. The ALNS parameters are based on pre-testing. In each iteration, a randomly drawn number of $\alpha \in \{1, \dots, |\mathcal{C}|\}$ surrounding clinics is removed if a removal operator is chosen. The reaction factors for the learning curve and the weight adjustment are set similarly to Rave et al. (2023). The initial temperature factor for the simulated annealing equals 1% of the costs of the initial solution and cools down proportionally by a cool rate dependent on the remaining run-time of the ALNS.

6.2. Performance analysis

6.2.1. 2E-SIRP instances

The performance of the ALNS is compared to optimal solutions if found and else to lower and upper bounds obtained using CPLEX for the two-stage stochastic program. As the instances in the numerical study are too extensive for solving them optimally, only subsets, i.e., fewer customers, products, and scenarios, of these instances are considered here. We further consider instances with three randomly generated clinic locations within a 100 km × 100 km map.

Table 4
Performance analysis for small instances.

$\|\hat{C}\|$	$\|\mathcal{P}\|$	$\|\Omega\|$	Cost objective				CPLEX		Run-time (s)	
			CPLEX	ALNS best	ALNS avg	ALNS σ [%]	Opt	GAP [%]	CPLEX	ALNS
4	1	5	4.08	4.08	4.09	0.2	12/12		2	60
4	2	5	7.80	7.80	7.81	0.1	4/4		7	60
4	3	5	14.32	14.31	14.40	0.7	3/4	6.9	1000	60
7	1	2	4.99	5.00	5.05	1.0	12/12		7	60
7	2	2	9.55	9.58	9.72	1.4	4/4		38	60
7	3	2	16.65	16.73	17.41	4.0	4/4		137	60
10	1	1	3.81	3.83	3.84	0.3	16/16		88	60
10	1	2	4.31	4.35	4.42	1.6	16/16		114	60
10	1	3	4.88	4.88	4.98	1.9	16/16		470	60
10	1	4	5.21	5.22	5.32	1.9	11/16	7.9	1627	60
10	1	5	5.55	5.52	5.63	1.9	10/16	10.9	2287	60
10	1	6	5.67	5.64	5.73	1.6	8/16	13.8	2607	60
10	1	7	5.91	5.91	6.01	1.6	2/16	10.3	3271	60
10	1	8	6.05	5.97	6.07	1.7	0/16	12.3	3600	60
10	1	9	6.23	6.06	6.16	1.5	0/16	17.0	3600	60
10	1	10	6.27	6.08	6.19	1.8	0/16	18.5	3600	60
10	2	1	13.89	14.08	14.54	3.1	9/16	10.1	2069	60
10	2	2	17.27	17.49	18.18	3.8	7/16	8.9	2635	60
10	2	3	20.45	20.36	21.07	3.3	0/16	18.5	3600	60
10	2	4	23.18	22.55	23.02	2.0	0/16	28.8	3600	60
10	2	5	23.51	22.81	23.26	1.9	0/16	32.7	3600	60
10	2	6	23.64	22.74	23.20	2.0	0/16	33.2	3600	60
10	2	7	23.95	22.49	22.96	2.0	0/16	36.1	3600	60
10	2	8	24.67	23.19	23.80	2.5	0/16	38.7	3600	60
10	2	9	25.28	23.31	23.95	2.6	0/16	39.7	3600	60
10	2	10	25.51	23.35	23.97	2.6	0/16	41.6	3600	60
10	3	1	32.73	34.92	36.45	4.2	0/16	14.1	3600	60
10	3	2	46.41	47.18	48.76	3.2	0/16	22.9	3600	60
10	3	3	51.91	51.98	53.50	2.8	0/16	37.1	3600	60
10	3	4	55.66	53.34	54.56	2.2	0/16	42.4	3600	60
10	3	5	58.10	55.29	56.50	2.1	0/16	45.4	3600	60
10	3	6	60.13	55.15	56.89	2.9	0/16	47.6	3600	60
10	3	7	59.91	55.34	56.45	2.0	0/16	50.3	3600	60
10	3	8	66.64	55.34	56.83	2.6	0/16	56.7	3600	60
10	3	9	70.65	55.25	56.76	2.6	0/16	56.4	3600	60
10	3	10	70.48	56.52	57.75	2.1	0/16	57.2	3600	60
Avg			26.67	24.77	25.41	2.2	97/520	28.2	2793	60

The run-time limit for the two-stage stochastic program was set to 60 min and for the ALNS to one minute to show that ALNS finds comparable or better results in a fraction of the time. The ALNS is executed five times. Table 4 presents aggregated results for 520 instances. Column $|\hat{C}|$ describes the number of clinics, column $|\mathcal{P}|$ the number of considered products, and column $|\Omega|$ the number of demand scenarios. The next four columns describe the best cost CPLEX found, the best costs the ALNS found, the average costs found, and the deviation between the best and average costs (σ). Columns eight and nine describe the number of optimal solutions found by CPLEX and the average gap of the not-to-optimality solved instances. Finally, the last two columns show the average run-time.

The results show that CPLEX can solve 97 out of the 520 instances to optimality. For the other 423 instances, the gaps are rather large. The ALNS, on the other hand, finds the same solutions or with only minimal deviations for clusters that are easy to solve, e.g., $|\hat{C}| = 4, |\mathcal{P}| = 1$. For clusters that are more difficult to solve (e.g., $|\hat{C}| = 4, |\mathcal{P}| = 3, |\Omega| = 5$), the ALNS finds good or slightly better in a small fraction of time compared to CPLEX, which might have run-time issues. Considering clusters that CPLEX cannot solve to optimality within one hour (e.g., $|\hat{C}| = 10, |\mathcal{P}| = 3, |\Omega| = 10$), the ALNS finds significantly better solutions. Thus, considering all clusters, the ALNS outperforms CPLEX in its best, and its average solution found, considering only a fraction of the run-time.

In Appendix G, we further benchmark our routing operators of Section 5.3.2 within the ALNS on the 30 IRP instances that were optimally solved by Archetti et al. (2007).

6.2.2. Stability tests

In this section, we demonstrate the stability of our ALNS solutions, i.e., how good the results are if demand fluctuates. For this, we apply the following procedure: First, we solve a single instance ($|\hat{C}| = 10, |\mathcal{P}| = 7, |\mathcal{T}| = 7, |\Omega| = 100$) and set the delivery periods $\dot{r}_{c,t}$, the routing R, and the reorder points $s_{p,c}$ for all clinics as in the found solution. In the second step, we set these decisions for a further 49 instances with the same demand distribution and determine the resulting costs. Please note that these 49 instances share the same parameters, except for different values of $d_{p,c,t,\omega}$. Third, we compare these costs with delivery periods $\dot{r}_{c,t}$, routing R, and reorder points $s_{p,c}$ for each of the 49 instances individually. Additionally, in the fourth step, we consider the case when there is an a-posteriori increase and decrease in demand by 15% on average. This means that for the 49 instances, the demand is, on average, significantly higher or lower compared to the instance delivery periods, routing, and reorder points are planned with.

Table 5 shows for all three cases the average cost increase and the standard deviation of the cost increase (σ). The ALNS is executed with a run-time limit of 60 min in each run, as the solution is used for a tactical planning problem (Rave et al., 2023).

Solving the single instance, each surrounding clinic is visited once a week with delivery periods on Thursday or Friday. Two tours are required for each delivery day (Thursday and Friday) in order to fulfill the scheduled deliveries. We now keep these delivery schedules and the respective

Table 5
Stability tests for unchanged, increased, and decreased average demand.

	Avg. cost increase [%]	σ [%]
Unchanged demand distribution	1.8%	1.5%
Increased demand by 15% on average	2.0%	1.1%
Decreased demand by 15% on average	4.0%	1.6%

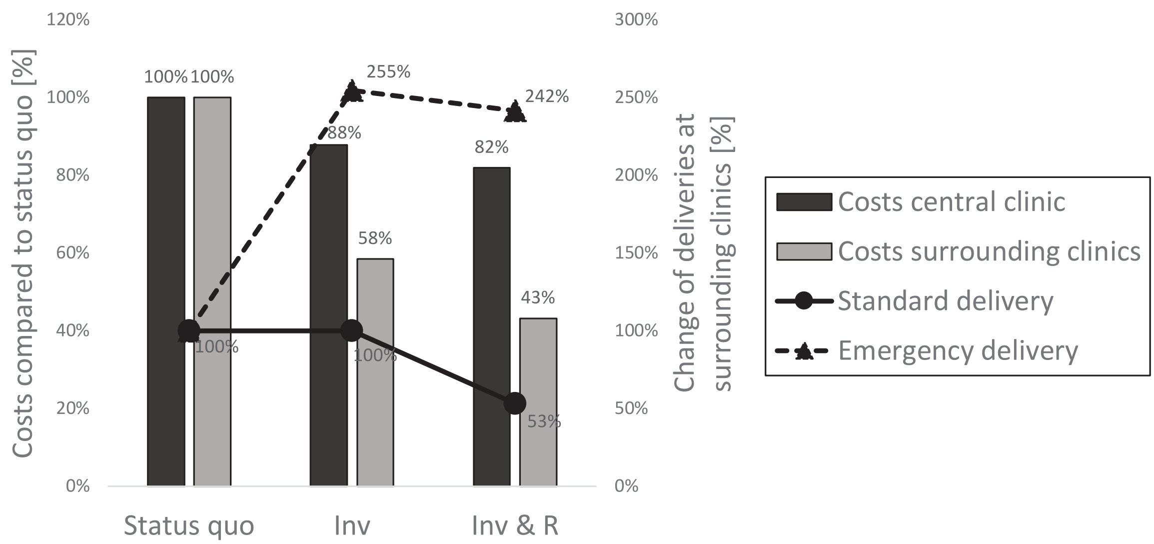

Fig. 3. Changes in cost and standard and emergency delivery for the central and surrounding clinics when optimizing inventory (Inv) and when additionally optimizing delivery periods and routes (Inv & R).

reorder points unchanged for the remaining 49 instances and determine the resulting costs in each case. On average, these costs are only 1.8% higher than the resulting costs, with an optimal solution for the respective instances. Moreover, the variations within the instances are rather low, with a standard deviation of 1.5%. When demand increases, the cost increase is only slightly higher with 2.0%. When average demand decreases, the average cost increase is higher at 4.0%, which is, however, still relatively low. Therefore, we can conclude that our approach leads to a stable solution for our instances.

6.3. Numerical results

6.3.1. Case study

We demonstrate the advantages of an optimization of delivery periods, routing to serve surrounding clinics, and reorder points compared to the status quo of clinic supply, which is as follows: Each surrounding clinic determines its delivery periods and reorder points on its own. There are, in total, about 250 emergency deliveries to surrounding clinics per year, as emergency deliveries are avoided if possible due to their extra costs. Subsequently, the central clinic plans the routing on the respective periods and its own inventory management based on the given information.

Our integrated approach not only optimizes the inventory of all clinics consolidated but also the routing simultaneously. Therefore, in the following, first, the reorder points for the central clinic and the surrounding clinics are optimized for the same routing as the status quo, and then the results are compared to the status quo. This is done by applying our reorder point algorithm to the routing of the status quo. Second, the routing is optimized additionally, and the results are again compared to the status quo. This is done by executing the ALNS. For both, we consider 50 instances with one product of each product category, and we set the run-time limit of the ALNS again to 60 min.

Fig. 3 shows the costs (central clinic: dark gray bar, surrounding clinics: light gray bar) compared to the status quo when optimizing the inventory (Inv) and when additionally optimizing delivery periods and routing (Inv & R) in percent. Moreover, the figure reports the change in standard (circles) and emergency deliveries (dashed line with triangles) in percent. Please note that the costs of the central clinic include the costs of supplying the central clinic via standard deliveries from the distributor and emergency supplies from the wholesaler, as well as the costs of inventory holding at the central clinic. The remaining costs are allocated to the surrounding clinics.

The inventory optimization leads to cost savings of 12% for the central clinic and 42% for the surrounding clinics, as demand peaks are eliminated. Additional cost savings of 15 percentage points can be reached for the surrounding clinics if delivery periods, routing, and reorder points are planned simultaneously. Moreover, this optimization leads to additional cost savings of 6 percentage points for the central clinic. Cost savings are mainly achieved because of the following: First, surrounding clinics decrease the number of standard deliveries to one visit per clinic per week but increase the number of emergency deliveries by 155% or 142%, respectively. Second, clinics make use of pooling effects, i.e., less inventory to be stored in surrounding clinics but more in the central clinic. This, however, does not increase the inventory of the central clinic in our case, as the central clinic itself uses emergency deliveries more often. Fig. 4 shows these pooling effects by reporting the reorder point change for the central and surrounding clinics in comparison to the status quo for inventory (Inv) and inventory and routing optimization (Inv & R).

The more expensive a product, the stronger the inventory pooling effect. Even after inventory optimization, the reorder points for product category 7 are set to zero for all surrounding clinics, which means that all demand is satisfied via emergency deliveries. Still, these products have an impact on emergency delivery and are, therefore, important to be considered in the integrated model. The central clinic, on the other hand, also

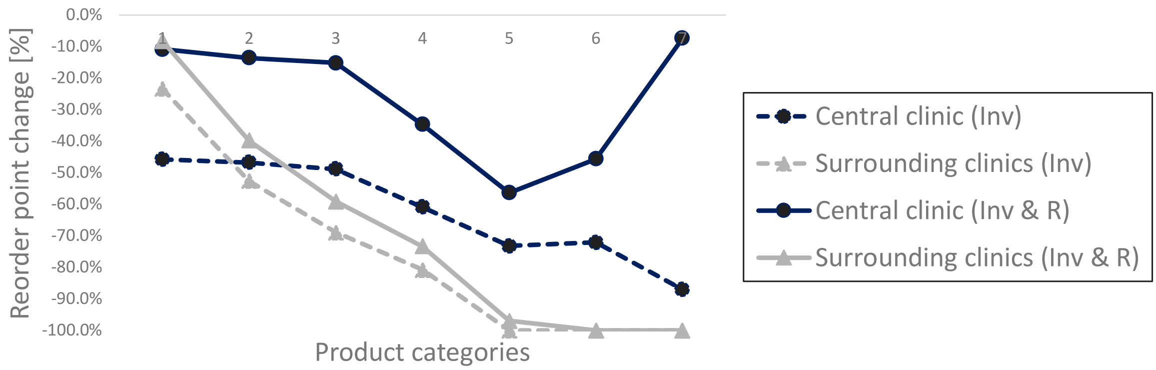

Fig. 4. Reorder point changes for the central and surrounding clinics when optimizing inventory (Inv) and when additionally optimizing delivery periods and routes (Inv & R).

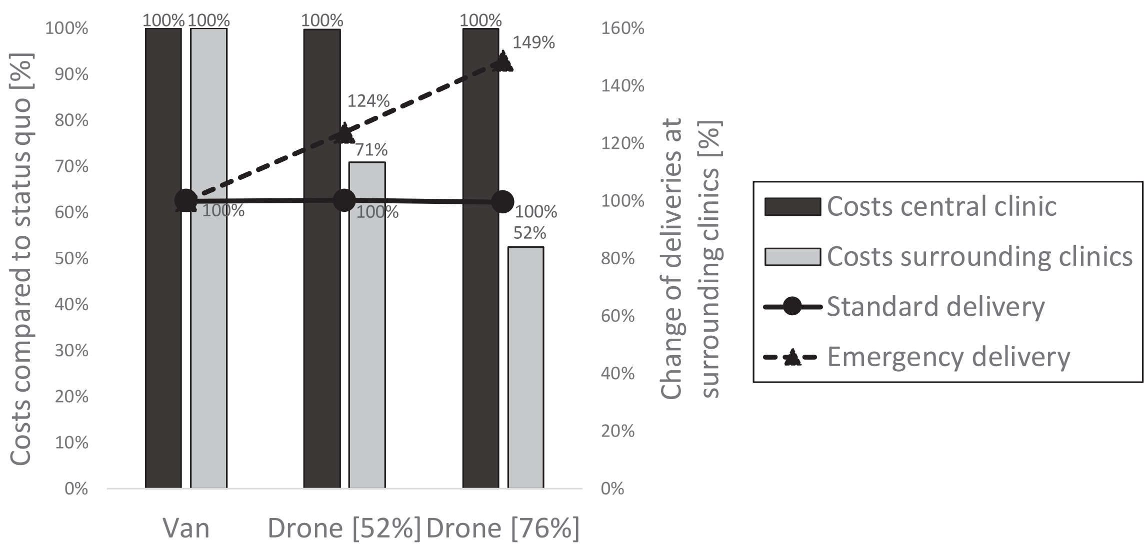

Fig. 5. Changes in cost and standard and emergency delivery for the central and surrounding clinics when considering vans and drones for emergency delivery.

has its reorder points reduced, as storing more inventory in the central clinic may not even be necessary when only split drugs are transported. Please note that we do not limit the number of emergency deliveries in our case study, which might be done in practice to prevent additional workload due to too often executed drone deliveries.

6.3.2. Impact of drone delivery

In this section, the influence of drone delivery as instant replenishment orders on total costs and inventory pooling effects is analyzed. Thus, Fig. 5 shows the change in costs and deliveries for the case if drones carry out the emergency delivery to the results for the case if the same vans carry out the emergency delivery as for standard delivery. Van delivery, on the one hand, has higher costs than drone delivery, but on the other hand, it has an unlimited capacity such that there is a maximum of one emergency delivery per clinic per period. We first consider costs for drone delivery as in our case study (Drone [52%]), i.e., drone delivery costs are 52% of van delivery costs in emergency delivery, and second, drone delivery costs are reduced by 76% (Drone [76%]).

Drones as emergency delivery options lead to cost savings of 29% for surrounding clinics but have nearly no impact on the costs for the central clinic. The cost savings for surrounding clinics increase to 48% when drone delivery is more cost-efficient. These cost savings can be achieved by reducing the stored inventory at the surrounding clinic if emergency deliveries are used more frequently instead. Fig. 6 shows the impact of drone delivery on the reorder points in comparison to van delivery.

For each product category, the pooling effects are stronger if the emergency delivery is more cost-efficient. The inventory of the central clinic also decreases slightly for some product categories, as there is an increase in the delivery of only individual consumer units to surrounding clinics so that whole packs are not stored, and the increased use of emergency deliveries leads to a distribution of delivery quantities over several periods, thus avoiding demand peaks. Please note that there are almost no changes for product categories 6 and 7, as these products already have a reorder point of 0 at the surrounding clinics when vans deliver the emergency supplies to the surrounding clinics.

6.3.3. Value of multi-product optimization

We now compare our multi-product approach, where reorder points for multiple products and routing and delivery periods are optimized simultaneously, to the following sequential procedure: first, we determine order intervals and reorder points for each product. For this, we solve our 2E-SIRP for each product individually, taking all relevant costs for inventory holding and transportation into account. Second, we use the

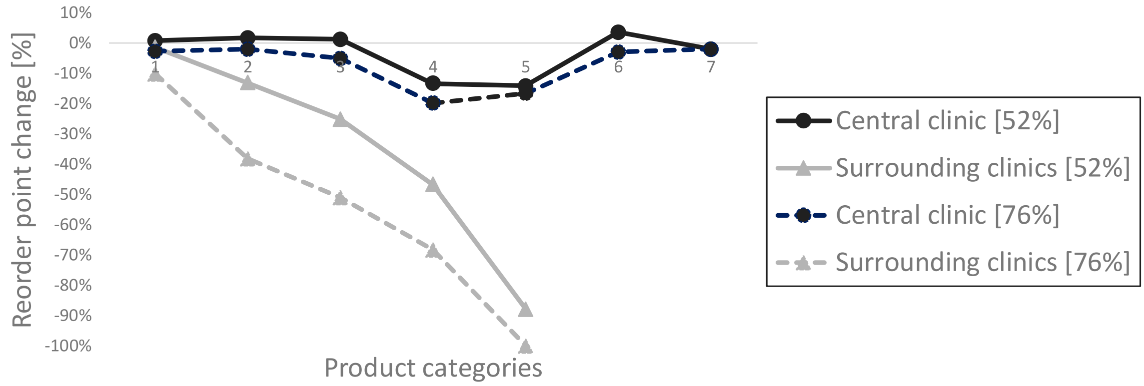

Fig. 6. Reorder point changes for the central and surrounding clinics when considering drones instead of vans for emergency deliveries.

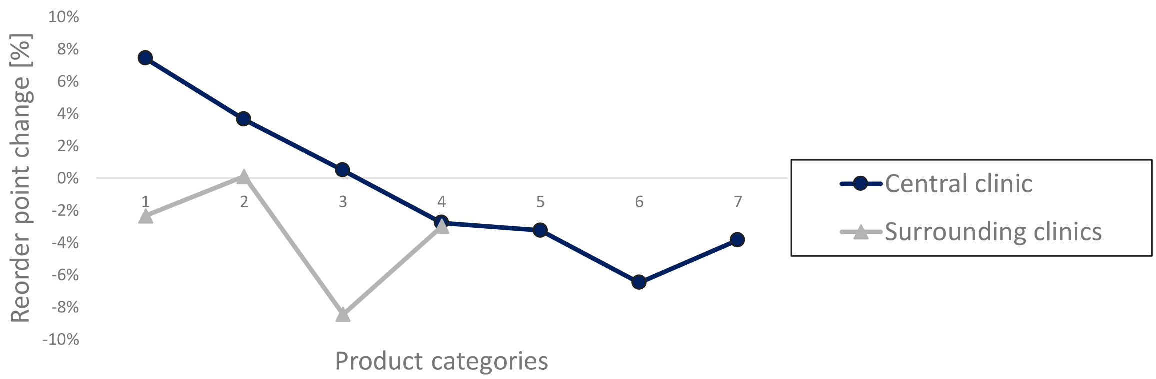

Fig. 7. Reorder point changes for each product category for the central clinic and for the surrounding clinics when considering all products simultaneously instead of solving the problem sequentially for each product.

resulting reorder points for all product categories as input parameters and optimize the routing by executing our ALNS with fixed reorder points. Note that with executing the first step, all product categories have the same order intervals at all clinics.

Fig. 7 presents for each product category the reorder point change when considering multiple products simultaneously in comparison to the individual consideration of products for the central clinic (circles) and the surrounding clinics on average (triangles). Note that the average reorder point changes are reported for the surrounding clinics.

We find that when considering multiple products simultaneously, the reorder points for the first product categories increase at the central clinic and decrease at the surrounding clinics, indicating an increase in pooling effects, i.e., more inventory is stored at the central clinic instead of at surrounding clinics. For product categories 5–7, it is vice versa. However, nothing is stored at all at surrounding clinics. The changes in reorder points can be predominantly attributed to the adjusted capacity ratios for each product in standard and emergency delivery, which have a direct impact on the transportation cost per delivery. The changes in reorder points lead to a cost decrease of 1% for surrounding clinics and 11% for the central clinic, so it is a decisive advantage to consider the multi-product case, especially for the central clinic on the first echelon. One major reason for these changes is that the vans' capacities are better balanced in the multi-product case. Van capacities for product categories 6 and 7 that are not fully utilized are also used for the delivery of product category 1.

7. Conclusion

This paper considers cyclic routing planning combined with inventory policies and drones as emergency delivery options for recourse decision for the medical supply of clinics. We introduce a two-echelon IRP that takes cyclic (r, s, nq) inventory policies for clinics in both echelons into account. We present the problem as a two-stage stochastic program and develop a specialized ALNS with problem-specific operators and an algorithm to calculate reorder points. In a performance analysis, we benchmark our approach to optimally solved instances and instances of Archetti et al. (2007) for the IRP. Further, we show the stability of our procedure and the value of multi-product consideration.

Summary of managerial implications: The case study shows that the use of pooling effects leads to total cost savings of 12% for the central clinic and 42% for surrounding clinics. The cost savings can even be increased by an additional 6% and 15% for the central clinic and the surrounding clinics, respectively if the route planning is also optimized and solved simultaneously with the inventory decision. We also found that a multi-product instead of a sequential single-product optimization approach significantly reduces costs, particularly for the central clinic. Furthermore, the study shows that drones as a more cost-efficient emergency delivery option reduce costs for surrounding clinics by 29% compared to an emergency delivery option via van. However, drone deliveries have nearly no impact on the inventory and supply costs of the central clinic.

Future research could include possible transshipments between the surrounding clinics. The transshipments are, for example, a pick-up of medications at one clinic during standard delivery and taken to the next or an emergency delivery launched from a closer surrounding clinic. Extensions of our problem setting, like consolidating multiple clinics in one tour in the emergency delivery, can be considered. This, however, leads to a vehicle routing problem (VRP) in the recourse, such that the ALNS presented in this paper needs to be adjusted to solve a VRP in the recourse as well. Further, shelf life can be considered. This is relevant when the routing and reorder points of, e.g., blood with an expiration date of some days instead of medication, need to be planned. Moreover, it can also be examined what the best fleet mix for emergency deliveries might look like, taking into account fixed costs and weather-related flight bans for drones.

CRediT authorship contribution statement

Alexander Rave: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Pirmin Fontaine:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization. **Heinrich Kuhn:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization.

Acknowledgments

The authors gratefully acknowledge the scientific and financial support of the Federal Ministry for Digital and Transport, Germany, for the research project MEDinTime reported in this paper. In addition, the authors would like to thank the anonymous reviewers and the editor for their valuable recommendations, which have significantly improved our paper.

Appendix A. Big M values

Setting big M values not arbitrarily large can reduce a standard solver's runtime. Thus, we recommend the following values for M_l with $l = 1, \dots, 6$. Please note that M_6 is introduced in Appendix B and is required for linearization.

- M_1 can be set as the number of all clinics considered ($|\hat{C}|$).
- M_2 can be set as the maximum reasonable delivery quantity a single product can have in standard delivery.
- M_3 needs to be greater than each clinic's possible maximum inventory capacities.
- M_4 can be set as a maximum number of transported packages in emergency deliveries.
- M_5 can be set as the sum of all clinics' expected demand for a product.
- M_6 can be set as twice the maximum reasonable inventory of a product.

Appendix B. Linearization of constraints (29) and (30)

In this section, we show how to linearize Constraints (29) and (30). For this, we introduce three additional variables, namely $\eta_{p,c,t} \in \{0,1\}$, $\theta_{p,c,t} \in \{0,1\}$, and $\zeta_{p,c,t} \in \mathbb{R}$ for $p \in \mathcal{P}$, $c \in \hat{C}$, and $t \in \bar{\mathcal{T}}$. $\eta_{p,c,t}$ determines the period with the last replenishment in standard delivery. If the expected demand since the last delivery exceeds the reorder point, $\theta_{p,c,t}$ ensures that the initial inventory is set to zero. $\zeta_{p,c,t}$ adjusts the inventory to linearize Constraints (29) and (30).

$$I^{Init}_{p,c} + \zeta_{p,c,\tau} = s_{p,c} + M_5 \cdot \sum_{t=\tau+1..|\bar{\mathcal{T}}|} r_{c,t} - \sum_{t=\tau..|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,c,t,\omega}\right\} \qquad \forall p \in \mathcal{P}, c \in C, \tau \in \bar{\mathcal{T}} \tag{33}$$

$$I^{Init}_{p,\hat{c}} + \zeta_{p,\hat{c},\tau} = s_{p,\hat{c}} + M_5 \cdot \sum_{t=\tau+1}^{|\bar{\mathcal{T}}|} r_{\hat{c},t} - \sum_{t=\tau}^{|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,\hat{c},t,\omega}\right\} - \sum_{t=\tau}^{|\bar{\mathcal{T}}|} \mathbb{E}_{\omega\in\Omega}\left\{\sum_{c\in C} n^S_{p,c,t,\omega\cdot q_p}\right\} \qquad \forall p \in \mathcal{P}, \tau \in \bar{\mathcal{T}} \tag{34}$$

$$\zeta_{p,c,t} \le M_5 \cdot \eta_{p,c,t} \qquad \forall p \in \mathcal{P}, c \in \hat{C}, t \in \bar{\mathcal{T}} \tag{35}$$

$$\sum_{t\in\mathcal{T}} \eta_{p,c,t} = |\bar{\mathcal{T}}| - 1 \qquad \forall p \in \mathcal{P}, c \in \hat{C} \tag{36}$$

$$\zeta_{p,c,t} \ge M_5 \cdot (\theta_{p,c,t} - 1) \qquad \forall p \in \mathcal{P}, c \in \hat{C}, t \in \bar{\mathcal{T}} \tag{37}$$

$$\zeta_{p,c,\tau} \ge s_{p,c} + M_5 \cdot \sum_{t=\tau+1..|\bar{\mathcal{T}}|} r_{c,t} - \sum_{t=\tau}^{|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,c,t,\omega}\right\} - M_6 \cdot \theta_{p,c,\tau} \qquad \forall p \in \mathcal{P}, c \in C, \tau \in \bar{\mathcal{T}} \tag{38}$$

$$\zeta_{p,\hat{c},\tau} \ge s_{p,\hat{c}} + M_5 \cdot \sum_{t=\tau+1..|\bar{\mathcal{T}}|} r_{\hat{c},t} - \sum_{t=\tau}^{|\mathcal{T}|} \mathbb{E}_{\omega\in\Omega}\left\{d_{p,\hat{c},t,\omega}\right\} - \sum_{t=\tau}^{|\bar{\mathcal{T}}|} \mathbb{E}_{\omega\in\Omega}\left\{\sum_{c\in C} n^S_{p,c,t,\omega\cdot q_p}\right\} - M_6 \cdot \theta_{p,c,\tau} \qquad \forall p \in \mathcal{P}, \tau \in \bar{\mathcal{T}} \tag{39}$$

$$\eta_{p,c,t}, \theta_{p,c,t} \in \{0,1\}, \zeta_{p,c,t} \in \mathbb{R} \qquad \forall p \in \mathcal{P}, c \in \hat{C}, t \in \bar{\mathcal{T}} \tag{40}$$

$$I^{Init}_{p,c} \in \mathbb{R}^+ \qquad \forall p \in \mathcal{P}, c \in \hat{C} \tag{41}$$

Constraints (33) and (34) equal Constraints (29) and Constraints (30) but without maximum and minimum function. For this, $\zeta_{p,c,t}$ is added on the left side, which adds or removes the remaining inventory to $I^{Init}_{p,c}$ to ensure equality. As $I^{Init}_{p,c}$ must be set to the lowest value or zero, $\zeta_{p,c,t}$ must be set to at maximum zero in a period $t \in \bar{\mathcal{T}}$ at least once. Constraints (35) ensure this, where binary variable $\eta_{p,c,t}$ is introduced, which is zero exactly for one period (Constraints (36)).

If the expected demand exceeds the reorder point, $I^{Init}_{p,c}$ must be set to zero. For this, $\theta_{p,c,t}$ is introduced to generate a lower bound for $\zeta_{p,c,t}$ that is zero if the reorder point is larger than the expected demand (Constraints (37)) or else to the reorder point minus the expected demand (Constraints (38) for surrounding clinics and Constraints (39) for the central clinic).

Appendix C. Objective function with penalty costs

To include penalty costs in our ALNS, we adjust our objective function as follows

$$f = Z^{total} \quad + k_P^{K^S} \cdot \sum_{t\in\bar{\mathcal{T}},\omega\in\Omega} \max\Big(\sum_{j\in C} \phi^S_{0,j,t,\omega} - K^S, 0\Big) + k_P^C \cdot \sum_{c\in\hat{C}} \max\Big(1 - \sum_{t\in\bar{\mathcal{T}}} \dot{r}_{c,t}, 0\Big)$$

where Z^{total} is our objective function, as in Section 4.2. The first term describes the costs that arise when van capacities are exceeded, with $k_P^{K^S}$ being the penalty factor for each capacity unit. If the accumulated transport volume each van carries when leaving the central clinic $\phi^S_{0,j,t,\omega}$ exceeds its capacity K^S, the difference is weighted with $k_P^{K^S}$. Otherwise, there are no penalty costs. The second term describes the costs if not all clinics are visited at least once, with k_P^C being the penalty factor for each clinic that is not supplied. If a clinic $c \in \hat{C}$ is not supplied, i.e., $\sum_{t\in\bar{\mathcal{T}}} \dot{r}_{c,t} = 0$, penalty costs arise. Otherwise, there are no penalty costs.

Appendix D. Detailed description of operators modifying the delivery period of the central clinic

Lowest standard quantity dependent delivery period removal - with noise (central clinic) - A delivery period T_i is chosen in iteration i, and the central clinic is no longer served on this period:

$$T_i = \min_{t\in\bar{\mathcal{T}}}\left(\sum_{c\in C,p\in\mathcal{P},\omega\in\Omega} n^S_{c,p,t,\omega} \cdot q^S \cdot \delta\right) \tag{42}$$

The period with the lowest total quantity delivered to the central clinic in standard delivery is chosen. To guarantee a certain randomness, a noise $\delta \in [0.8, 1.2]$ is added. Note that only periods are selected where there is a standard delivery to the central clinic.

Largest emergency quantity dependent delivery period insertion - with noise (central clinic) - A delivery period T_i is chosen in iteration i, and the central clinic is served at this period:

$$T_i = \max_{t\in\bar{\mathcal{T}}}\left(\sum_{p\in\mathcal{P},\omega\in\Omega} q^{em.}_{p,\hat{c},t,\omega} \cdot \delta\right) \tag{43}$$

The decision on which period to choose depends on the emergency delivery quantity of the corresponding periods supplied by the central clinic. We consider the number of emergency deliveries because multiple emergency deliveries in a period indicate that it may be reasonable to add a standard delivery instead. Again, a noise $\delta \in [0.8, 1.2]$ is added.

Appendix E. Detailed description of determining delivery periods for insertion operators

Insertion operators that modify the routing of surrounding clinics insert the removed clinics to a delivery period T_i that is determined as follows in an iteration i:

$$T_i = \max_{t\in\bar{\mathcal{T}}}\left(\sum_{c\in C,p\in\mathcal{P},\omega\in\Omega} q^{em.}_{p,c,t,\omega} \cdot \pi_{i-1,t} \cdot \delta\right) \tag{44}$$

This delivery period T_i is the period with the largest quantity of products transported in emergency delivery, i.e., via drone delivery, as it might be reasonable to serve the clinics at this period. Further, we consider the factor $\pi_{i,t}$ for the historical operators' performance of selecting period t and a noise $\delta \in [0.8, 1.2]$. $\pi_{i,t}$ is initialized and updated (only if selected) similar to the operator's weights.

Appendix F. Detailed description of the operator modifying the reorder points

Increase or decrease a clinic's reorder point - This operator chooses a random clinic $c \in \hat{C}$ and product p and increases or decreases its reorder point $s_{p,c}$ dependent on the number of emergency deliveries in the two-echelon network structure. If the selected clinic is a surrounding clinic, i.e., $c \in C$, then:

- If there are ρ times more emergency deliveries to the central clinic than to the considered surrounding clinic, $s_{p,c}$ is decreased by $\xi \in \{1, \dots, \xi^{max}\}$.
- If there are ρ times fewer emergency deliveries to the central clinic than to the considered surrounding clinic, $s_{p,c}$ is increased by $\xi \in \{1, \dots, \xi^{max}\}$.
- If none of the above occurs, $s_{p,c}$ is either increased or decreased by $\xi \in \{1, \dots, \xi^{max}\}$.

with $\rho \in \{1, \dots, 10\}$ and ξ being randomly chosen and ξ^{max} being set as 10% of the reorder point of product p and the selected clinic c: $\xi^{max} = \max(\lceil 10\% \cdot s_{p,c} \rceil, 1)$. The first two cases allow reorder points to be systematically adjusted in a more reasonable direction. The third case allows a change in both directions. We observed that the number of emergency deliveries has the strongest impact on the reorder points. However, it is not the only influence, as there are also inventory holding costs, for example. Thus, a certain randomness by choosing $\rho \in \{1, \dots, 10\}$ randomly and also by applying the third case is introduced so that not every influencing factor has to be taken into account. If the selected clinic is the central clinic, the cases are applied vice versa, e.g., $s_{p,c}$ is increased in the first case.

Table G.6
Performance analysis for benchmark instances of Archetti et al. (2007).

$\|\mathcal{C}\|$	Opt. costs	Our ALNS				
		Best costs	Avg. costs	Opt	Δ [%]	σ [%]
5	5538.02	5538.02	5572.80	5/5	0.00	0.67
10	8872.41	8903.47	9197.99	2/5	0.35	3.18
15	11721.83	11811.86	12114.40	0/5	0.78	2.47
20	14863.85	15006.32	15219.32	0/5	0.95	1.40
25	17170.81	17479.36	17768.22	0/5	1.81	1.63
30	20657.29	20923.60	21243.02	0/5	1.27	1.53
Avg	13137.37	13277.10	13519.29	7/30	0.86	1.81

Appendix G. Benchmark tests with IRP instances of Archetti et al. (2007)

In this section, we benchmark the performance of our routing operators of Section 5.3.2 within the ALNS on the 30 IRP instances that were optimally solved by Archetti et al. (2007) since they come closest to our problem setting of all IRP problems addressed in the literature so far. In comparison to our problem setting, Archetti et al. (2007) consider the following simplifications: The authors consider a single product whose demand is deterministic, constant, and known for the considered six periods, i.e., there is only a single scenario and no recourse decision. Moreover, the central clinic has no demand, and its inventory is replenished in a constant amount each period. All clinics' initial inventory is a given parameter as a fixed planning horizon is assumed. The inventory is always replenished to an order-up-to-level, which is a parameter. In addition, there is only one vehicle. Last, the inventory holding costs arise per clinic instead of per product h_c $(c \in \hat{\mathcal{C}})$. It follows that running the ALNS, neither the reorder point algorithm nor the operators that modify the central clinic's delivery periods or the clinics' reorder points are executed.

The 30 instances of Archetti et al. (2007) consider $|\mathcal{C}| \in \{5, 10, 15, 20, 25, 30\}$ clinics and $|\bar{\mathcal{T}}| = |\mathcal{T}| = 6$ delivery periods. Clinic-dependent inventory costs k_c^I $(c \in \mathcal{C})$ are randomly selected out of interval $[0.1, 0.5]$ for the respective instances. For the central clinic, it always holds $k_{\hat{c}}^I = 0.3$. In addition, the deterministic demand per period is randomly selected from interval [10,100].

Table G.6 shows the consolidated results of these instances with different numbers of clinics $|\mathcal{C}| \in \{5, 10, 15, 20, 25, 30\}$. The first column describes the number of considered surrounding clinics, and the next column describes the proven optimal solution by Archetti et al. (2007). Columns three and four present our best-found solution of five runs and the average found solution. Columns five and six give information on the number of solved instances to optimality and the average gap to optimality (Δ). Last, column seven shows the gap from best to average (σ). We chose a run-time of 60 min for running the ALNS.

Our ALNS solves 7 of the 30 instances to optimality with an average gap of 0.86%. Moreover, we observe that the solution is stable with an average σ of 1.81%. Especially the performance does not decrease monotonously with a larger number of considered clinics. These results show the performance, especially since our ALNS was not designed for this operational problem setting that has large differences but for a tactical problem setting with multiple products and stochastic demands that also includes the optimization of reorder points for cyclic delivery patterns.

References

Aksen, D., Kaya, O., Salman, F. S., & Tüncel, Ö. (2014). An adaptive large neighborhood search algorithm for a selective and periodic inventory routing problem. *European Journal of Operational Research*, *239*(2), 413–426.

Archetti, C., Bertazzi, L., Laporte, G., & Speranza, M. G. (2007). A branch-and-cut algorithm for a vendor-managed inventory-routing problem. *Transportation Science*, *41*(3), 382–391.

Archetti, C., & Ljubić, I. (2022). Comparison of formulations for the inventory routing problem. *European Journal of Operational Research*, *303*(3), 997–1008.

Arslan, O. (2021). The location-or-routing problem. *Transportation Research Part B: Methodological*, *147*, 1–21.

Asadi, A., Nurre Pinkley, S., & Mes, M. (2022). A Markov decision process approach for managing medical drone deliveries. *Expert Systems with Applications*, *204*, Article 117490.

Bell, W. J., Dalberto, L. M., Fisher, M. L., Greenfield, A. J., Jaikumar, R., Kedia, P., Mack, R. G., & Prutzman, P. J. (1983). Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *INFORMS Journal on Applied Analytics*, *13*(6), 4–23.

Broekmeulen, R. A., Sternbeck, M. G., van Donselaar, K. H., & Kuhn, H. (2017). Decision support for selecting the optimal product unpacking location in a retail supply chain. *European Journal of Operational Research*, *259*(1), 84–99.

Campbell, A., Clarke, L., Kleywegt, A., & Savelsbergh, M. (1998). The inventory routing problem. In *Fleet management and logistics* (pp. 95–113). Kluwer Academic Publishers.

Campbell, A. M., & Savelsbergh, M. W. P. (2004). A decomposition approach for the inventory-routing problem. *Transportation Science*, *38*(4), 488–502.

Campbell, A. M., & Wilson, J. (2014). Forty years of periodic vehicle routing. *Networks*, *63*(1), 2–15.

Çetinkaya, S., & Bookbinder, J. H. (2003). Stochastic models for the dispatch of consolidated shipments. *Transportation Research Part B: Methodological*, *37*(8), 747–768.

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

Cui, Z., Long, D. Z., Qi, J., & Zhang, L. (2023). The inventory routing problem under uncertainty. *Operations Research*, *71*(1), 378–395.

Frank, M., Ostermeier, M., Holzapfel, A., Hübner, A., & Kuhn, H. (2021). Optimizing routing and delivery patterns with multi-compartment vehicles. *European Journal of Operational Research*, *293*(2), 495–510.

Higginson, J. K., & Bookbinder, J. H. (1995). Markovian decision processes in shipment consolidation. *Transportation Science*, *29*(3), 242–255.

Holzapfel, A., Hübner, A., Kuhn, H., & Sternbeck, M. (2016). Delivery pattern and transportation planning in grocery retailing. *European Journal of Operational Research*, *252*(1), 54–68.

Jafarkhan, F., & Yaghoubi, S. (2018). An efficient solution method for the flexible and robust inventory-routing of red blood cells. *Computers & Industrial Engineering*, *117*, 191–206.

Johansson, L., Sonntag, D. R., Marklund, J., & Kiesmüller, G. P. (2020). Controlling distribution inventory systems with shipment consolidation and compound Poisson demand. *European Journal of Operational Research*, *280*(1), 90–101.

Malicki, S., & Minner, S. (2021). Cyclic inventory routing with dynamic safety stocks under recurring non-stationary interdependent demands. *Computers & Operations Research*, *131*, Article 105247.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245.

Mercedes-Benz, A. (2022). Technical data, weights and measures of the sprinter panel van. *World Wide Web*, URL https://www.mercedes-benz.ie/vans/en/sprinter/panel-van/technical-data, last access: December 01, 2022.

Niakan, F., & Rahimi, M. (2015). A multi-objective healthcare inventory routing problem; a fuzzy possibilistic approach. *Transportation Research Part E: Logistics and Transportation Review*, *80*, 74–94.

Nicholson, L., Vakharia, A. J., & Erenguc, S. S. (2004). Outsourcing inventory management decisions in healthcare: Models and application. *European Journal of Operational Research*, *154*(1), 271–290.

Nikzad, E., Bashiri, M., & Oliveira, F. (2019). Two-stage stochastic programming approach for the medical drug inventory routing problem under uncertainty. *Computers & Industrial Engineering*, *128*, 358–370.

Nolz, P. C., Absi, N., & Feillet, D. (2014). A bi-objective inventory routing problem for sustainable waste management under uncertainty. *Journal of Multi-Criteria Decision Analysis*, *21*(5–6), 299–314.

Otto, A., Agatz, N., Campbell, J., Golden, B., & Pesch, E. (2018). Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, *72*(4), 411–458.

Pisinger, D., & Røpke, S. (2007). A general heuristic for vehicle routing problems. *Computers & Operations Research*, *34*(8), 2403–2435.

Quantum Systems (2020). Long-range drones to deliver emergency medication to hospitals – project MEDinTime. *World Wide Web*, URL https://www.quantum-systems.com/category/medintime/, last access: December 01, 2022.

Raa, B., & Aouam, T. (2023). A shortfall modelling-based solution approach for stochastic cyclic inventory routing. *European Journal of Operational Research*, *305*(2), 674–684.

Rahimi, M., Baboli, A., & Rekik, Y. (2017). Multi-objective inventory routing problem: A stochastic model to consider profit, service level and green criteria. *Transportation Research Part E: Logistics and Transportation Review*, *101*, 59–83.

Rave, A., Fontaine, P., & Kuhn, H. (2023). Drone location and vehicle fleet planning with trucks and aerial drones. *European Journal of Operational Research*, *308*(1), 113–130.

Rohmer, S., Claassen, G., & Laporte, G. (2019). A two-echelon inventory routing problem for perishable products. *Computers & Operations Research*, *107*, 156–172.

Roldán, R. F., Basagoiti, R., & Coelho, L. C. (2017). A survey on the inventory-routing problem with stochastic lead times and demands. *Journal of Applied Logic*, *24*, 15–24.

Røpke, S., & Pisinger, D. (2006). An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, *40*(4), 455–472.

Sacramento, D., Pisinger, D., & Røpke, S. (2019). An adaptive large neighborhood search metaheuristic for the vehicle routing problem with drones. *Transportation Research. Part C: Emerging Technologies*, *102*, 289–315.

Sonntag, D. R., Schrotenboer, A. H., & Kiesmüller, G. P. (2023). Stochastic inventory routing with time-based shipment consolidation. *European Journal of Operational Research*, *306*(3), 1186–1201.

Stenius, O., Karaarslan, A., Marklund, J., & de Kok, A. G. (2016). Exact analysis of divergent inventory systems with time-based shipment consolidation and compound Poisson demand. *Operations Research*, *64*(4), 906–921.

Taube, F., & Minner, S. (2018). Data-driven assignment of delivery patterns with handling effort considerations in retail. *Computers & Operations Research*, *100*, 379–393.

Tucker, C. E. M. A. W., & Zemlin, R. A. (1960). Integer programming formulation of traveling salesman problem. *Journal of the ACM*, *7*(4), 326–329.

Vidal, T., Crainic, T. G., Gendreau, M., & Prins, C. (2013). A hybrid genetic algorithm with adaptive diversity management for a large class of vehicle routing problems with time-windows. *Computers & Operations Research*, *40*(1), 475–489.

Voigt, S. (2024). A review and ranking of operators in adaptive large neighborhood search for vehicle routing problems. *European Journal of Operational Research*, available online 18 May 2024.

Volland, J., Fügener, A., Schoenfelder, J., & Brunner, J. O. (2017). Material logistics in hospitals: A literature review. *Omega*, *69*, 82–101.

Wald, A. (1947). *Sequential analysis*. John Wiley & Sons.

Wensing, T., Sternbeck, M. G., & Kuhn, H. (2018). Optimizing case-pack sizes in the bricks-and-mortar retail trade. *OR Spectrum*, *40*(4), 913–944.

Yadollahi, E., Aghezzaf, E.-H., & Raa, B. (2017). Managing inventory and service levels in a safety stock-based inventory routing system with stochastic retailer demands. *Applied Stochastic Models in Business and Industry*, *33*(4), 369–381.

Zheng, X., Yin, M., & Zhang, Y. (2019). Integrated optimization of location, inventory and routing in supply chain network design. *Transportation Research Part B: Methodological*, *121*, 1–20.