Cyclic Stochastic Inventory Routing Planning with Reorder Points and Recourse Decision for an Application in Medical Supply

Alexander Rave, Pirmin Fontaine, Heinrich Kuhn

Catholic University Eichstätt-Ingolstadt, Ingolstadt School of Management, Department of Operations, Auf der Schanz 49, 85049 Ingolstadt, Germany
e-mails: arave@ku.de, pirmin.fontaine@ku.de, heinrich.kuhn@ku.de

Abstract

Drug availability in clinics is essential for patient services, whose demand for medication on the other hand is uncertain. Thus, clinics need to have a variety of medication available, which lead 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 own inventory policy as they plan independent of each other. Additionally, these inventory policies include instant replenishment orders to avoid shortages. As these instant replenishment orders only include few medication, these can be performed by a variety of vehicles, including vans, cabs, and aerial drones.

We present a two-stage stochastic program for a multi-product two-echelon inventory routing problem with stochastic demands under consideration of inventory policies. We decide on the cost-optimal cyclic delivery patterns and reorder points for the clinics with instant replenishment orders performed by drones as recourse decision in case of shortages. Due to the problem's complexity caused by the interdependencies of inventory policies in the two echelons, we develop a specialized adaptive large neighborhood search that includes an algorithm for determining local optimal reorder points for the routing in each iteration. We present a case study at a large German clinic, which is responsible for serving surrounding clinics and plans to integrate drone instead of truck deliveries as emergency resupply decision. Our integrated approach not only leads to cost savings of 58.4% for the surrounding clinics but also 15.9% for the central clinic. By using drone delivery compared to van delivery, average stock of medication at surrounding clinics can be reduced by at least 1.7% to 100% for expensive and rarely needed medication, which results in a total cost decrease of 29.8% while maintaining medication availability.

Keywords: Inventory policies, Drone delivery, Two-stage stochastic program, Adaptive large neighborhood search

1. Introduction

Clinics are obligated to care for patients. To guarantee this treatment, each clinic has a pharmacy that is appropriately equipped with medication. The pharmacy is under pressure to operate cost-effectively, while still providing a full range of medical care to patients. About 10 to 18% of a clinic's total expenses is spent on medication (Nicholson et al., 2004;

Volland et al., 2017). A way to reduce inventory costs is to make use of pooling effects. This means that safety stock is stored more centrally in order to exploit the stochastic balancing effect between the possibly independent demands in the individual clinics. As a consequence, however, additional transports are necessary. An aerial drone is an option to perform these transports, since the drone has the particular advantage that they can carry few medication to clinics faster and at lower costs than vans (Otto et al., 2018). In Ghana, for example, medication is pooled in a central warehouse and then transported by drones in order to increase 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 cycling 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 and Ljubić, 2022). The delivery amount for the supply of clinics, however, is dependent on the inventory policy, which is planned individually by each clinic. The clinics make use of inventory policies as inventory can thus be planned over a longer planning horizon. In order to avoid shortages, 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 few medication and can be performed by a variety of vehicles, for example, by drones (Quantum Systems, 2020).

Motivated by the previous introduced healthcare application, we develop a two-stage stochastic program for a cyclic multi-product two-echelon IRP with stochastic demand and inventory policies (2E-SIRP) that determines cost-optimal reorder points for both the central and the surrounding clinics. In addition, it determines the delivery periods and routes for supplying the surrounding clinics. While these are decisions on the tactical plan of the first stage, the second stage decides on the resulting instant replenishment orders performed by drones, 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 greater variety of products and scenarios, we develop a specialized adaptive large neighborhood search (ALNS). In each iteration of the ALNS, the locally optimal reorder points are determined for the routing that is modified by removal or insertion operators. We benchmark the performance of the ALNS with optimally solved instances for the two-stage stochastic program and generate managerial insights for a novel case study based on a large clinic in Germany. We present both the benefits of our integrated planning approach as well as the influence of drones as new types of vehicles for emergency deliveries on the inventory levels of different products and total costs reduction.

We contribute to the literature as follows: First, our paper is the first that combines two different inventory policies with routing decisions. Second, we formulate this problem setting as a two-stage stochastic program. Third, we derive an ALNS that determines local optimal reorder points in each iteration after taking changes on the routing to clinics. Fourth, we present a novel case study based on a large clinic in Germany.

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

The central clinic - which has its own demand - procures all kind of medical goods like medicines, antibiotics, infusions, etc. from the central distributor of these products who is assumed to have sufficient inventory at all times. The demand of 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, if the inventory at the central clinic is too low, shortages may arise. Due to the characteristics of the treatment of patients, these have to be satisfied through emergency deliveries ordered from and supplied by 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. Standard deliveries from the distributor and emergency deliveries from the wholesaler are performed in commuting tours. However, different products are consolidated, if they are ordered at the same day.

The central clinic supplies several surrounding clinics with the medical goods they need. The clinics are supplied by multiple homogeneous vans on predefined days within a recurring supply period. The vans are considered to have a limited payload and the supply is performed by a third party logistics. Therefore, the number of driver utilization is not balanced per period. The determination of the respective delivery pattern for each clinic is part of the planning problem under consideration. In addition, an emergency delivery is performed by the central clinic if shortages might 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 more cost-effective in terms of delivery than vans, but, they can only carry few medication per flight due to limited payload (Otto et al., 2018) - which is, however, mostly the case in emergency deliveries.

The standard delivery for all clinics is only allowed during weekdays, since the pharmacy of clinics is only staffed for emergency operation at the weekend. Contrary, emergency deliveries can be conducted on every day, but are costlier at the weekend. Note that these are case related assumptions, which can be easily relaxed and do not change our introduced algorithm.

We assume an unlimited fleet of vans and drones. The vans and drones, which serve the surrounding clinics have a limited payload. The vans of the central distributor that serve the central clinic, on the other hand, are assumed to have an unlimited payload. If a surrounding clinic's demand exceeds the payload in emergency delivery, the drone might serve the same clinic multiple times within a period.

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

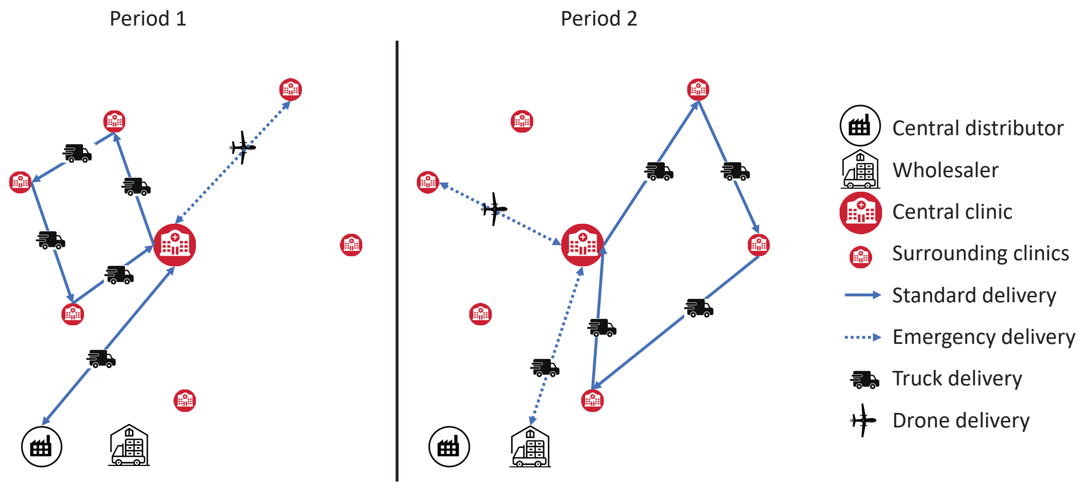


Figure 1: Delivery system of medicine

2.2. Inventory policy

For each product $p \in \mathcal{P}$ the central clinic $\hat{c}$ operates an $(s_{p,\hat{c}}, Q_p)$ inventory policy with additional emergency resupply to avoid shortages. An order is placed at the central distributor of the product $p \in \mathcal{P}$ with an amount of Q_p, if the inventory level is below the reorder point $s_{p,\hat{c}}$.

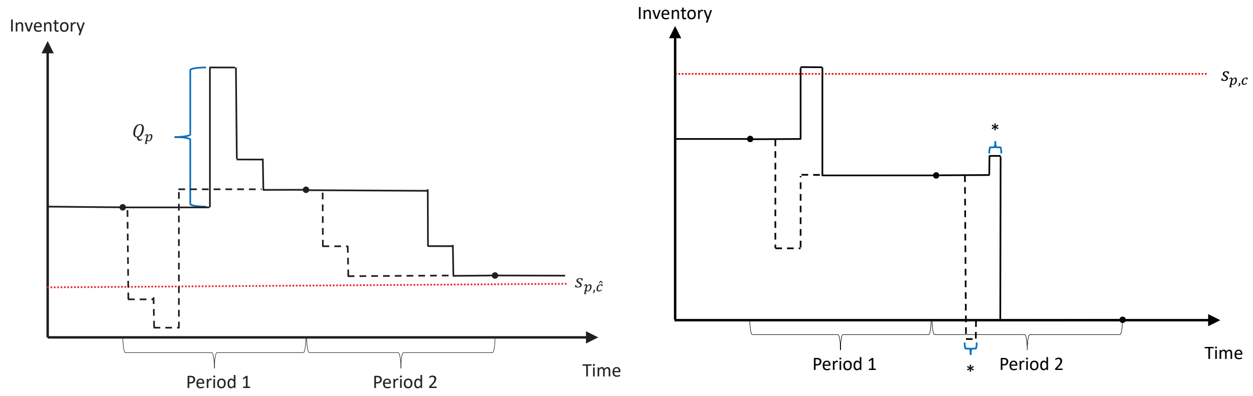
The replenishment quantity Q_p is assumed to be given. This quantity is typically based on long-term contracts between central clinic and distributor that includes total demand forecasts, minimum order quantities, quantity discounts as well as ordering and delivery costs. If the demand would exceed the current inventory level, additional emergency delivery costs arise, as fast delivery is conducted. In this case, the delivery quantity is equal to the resulting shortages, i.e., only the missing quantity.

Surrounding small clinics $c \in \mathcal{C}$ operate an $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy with additional emergency resupply to prevent shortages. In periods of $\mathcal{T}_c$ in a cyclic repeating pattern the inventory is refilled by an order of $n_{p,c}$ packages á q_p consumer units, e.g., pills, if the inventory is below the reorder point $s_{p,c}$. The inventory is then replenished so that the inventory exceeds the reorder points $s_{p,c}$, if no additional demand occurs between ordering and delivery. Note that all subsets $\mathcal{T}_c$ of periods of $\mathcal{T}$ can be selected (e.g., delivery on Monday and Thursday). If there is not enough medication on stock, this demand is satisfied by an emergency delivery. This is a special case of back-ordering (Coelho et al., 2014), because the corresponding demand is served directly and not at a later stage. Note that the $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ inventory with emergency resupply 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 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 of 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 share of a package, 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 - central clinic. At the beginning of each period, the disposable inventory equals the physical inventory. In the first step, the demand of the surrounding clinics and the demand of the central clinic are known, which change the disposable inventory. If the disposable inventory falls below the reorder point $s_{p,\hat{c}}$ during weekdays (Monday to Friday), an order of Q_p is triggered in the second step, which includes whole packing units. The disposable inventory is adjusted accordingly. If this results in total in a negative disposable inventory, the third step is to additionally place an emergency order to prevent a negative physical inventory. Then, in the fourth step, the regular order quantity Q_p and emergency orders, if triggered, arrive in the central clinic. Next, in the fifth step, the surrounding and central clinics' demands are physically fulfilled. Disposable and physical inventory are then equivalent again and denote the inventory level at the end of the period, which is then used to quantify the inventory holding cost in the central clinic. The described disposition, supply and delivery scheme of the central clinic allows to fulfill the demand of the surrounding clinics and central clinic itself in all periods under consideration, without a situation-related delivery time.

Sequence of events for the replenishment cycle - surrounding clinics. At the beginning of each period, the physical and disposable inventory are identical. Subsequently, the disposition and the supply (delivery) of the surrounding clinics is differentiated according to whether a standard delivery is planned in the period. In periods with a standard delivery, it is first checked whether the disposable inventory is below the reorder point $s_{p,c}$ and, if so, an order is placed for this product that includes whole packaging units and increases the disposable inventory at least to the level $s_{p,c}$. The further events are then identical for all periods. The



(a) $(s_{p,\hat{c}}, Q_p)$ policy at the central clinic

(b) $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ policy at surrounding clinics
* emergency resupply

Figure 2: Replenishment policies of clinics
The solid line represents the physical stock, the dashed line the disposable stock.

demand of the clinic becomes known and the disposable inventory is adjusted. If this results in a negative disposable inventory, an emergency order is placed (additionally to a standard delivery, if necessary), which balances the disposable inventory to zero. Subsequently, any triggered deliveries (standard delivery and/or emergency delivery) arrive at the considered clinic. Next, the clinic's demand is fulfilled. Disposable and physical inventory are equal again. The resulting inventory level is used to quantify the clinic's inventory holding costs. The described scheduling and supply scheme of the clinics enables to completely fulfill the respective demands of a clinic in all periods.

Example inventory level. Figure 2 presents an example inventory level of a drug for the central clinic (Figure 2a) and one surrounding clinic (Figure 2b) for two periods. For the central clinic, orders are set as soon as the disposable inventory is below the reorder point $s_{p,\hat{c}}$, which is the case in period 1. The physical inventory is then increased by the delivery quantity Q_p and reduced by both the demand of clinics and the central clinic's own demand. In period 2 the disposal and the physical demands are decreased, but no order is placed, as the remaining disposal inventory is above the reorder point. The large points on the lines mark the time when inventory is assessed for the subsequent period.

For the surrounding clinics, standard delivery replenishes the inventory in period 1 by the missing physical inventory at the end of the previous period compared to $s_{p,c}$ in complete packages. Afterwards the physical inventory is reduced by the demand. In period 2 there is no standard delivery, but demand exceeds the disposal inventory level and thus an emergency order is placed ($*$). The emergency delivery only contains the missing inventory to fulfil the demand. Thus, the physical inventory equals 0 afterwards. Note that an emergency delivery might occur in the same delivery period a standard delivery is conducted.

2.3. Decision problem

We determine cost-optimal reorder points $s_{p,\hat{c}}$ for the inventory policy of the central clinic and reorder points $s_{p,c}$, delivery days T for the inventory policy of the surrounding clinics. For this, we make use of the two-echelon IRP structure of the present problem and thus also determine the routing for service of the surrounding clinics. We minimize routing costs as well as expected inventory and emergency delivery costs. Due to stochastic demand, this methodically leads to a static stochastic program with a recourse decision in 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 a $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy for surrounding clinics and a $(s_{p,\hat{c}}, Q_p)$ inventory policy for the central clinic.

The planning decision is made once for cyclic planning periods (in our case: one week) and then carried out over a longer planning horizon, e.g., three month, as long as there are no structural changes. The inventory policy therefore 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).

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. Therefore, the simultaneous consideration of multiple products leads to a different solution than consideration 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 cylcic planning, and is thus dependent on the last delivery for surrounding clinics (e.g., Malicki and Minner, 2021). Since the central clinic has no cyclical supply patterns, the initial stock is estimated as the average inventory level.

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 an overview on 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). There is a wide range of publications available concerning the PVRP (Campbell and Wilson, 2014). However, in our case the clinics define clinic-specific delivery patterns (DP) and the central clinic cyclically supplies clinics with their requirements of medical goods. The DPs constitute a defined combination of weekdays on which the individual clinics are supplied. Defining the DPs also means determining the delivery frequency. The delivery frequency, however, impacts the volume per clinic delivery, which in turn affects the associated cycle and safety stock in the central clinic as well in 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 DPs 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 also consider time-based shipment consolidation.

Stochastic inventory routing - single product. When considering stochastic demands, there are either lost sales or back-ordering, which state that unfulfilled demand can be met in a later period (Coelho et al., 2014). In our problem setting, additional emergency deliveries in form of back-ordering occur. The difference between this emergency back-ordering and lost sales is that shortages from one clinic lead to inventory changes in the central clinic, since in this case the latter releases an emergency delivery. Additionally, lost sales lead to penalty costs per patient and back-ordering to penalty costs per delivery.

IRPs with stochastic demands (SIRP) are considered by Campbell et al. (1998) and Campbell and Savelsbergh (2004). The authors focus on a heuristic solution method and decompose the decisions in two phases: first the delivery period and quantities are determined, second routes are created by an insertion heuristics. Nolz et al. (2014) consider stochasticity in their SIRP through recourse decisions and propose an ALNS and two versions of a LNS for a bi-objective IRP. Cui et al. (2022) choose a robust optimization approach for a SIRP that is modelled 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. For solving 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 it is sufficient to set the initial inventory endogenously in cyclic IRPs. Raa and Aouam (2023) also consider a cyclic IRP with lost sales, but the delivery quantity is dependent on a base-stock unless the capacity is not sufficient (shortfall). The authors develop a metaheuristic that is 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 cycles T the inventory of the one 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 application. Another healthcare application is considered by Jafarkhan and Yaghoubi (2018) who consider an SIRP for the distribution of 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.

Time-based consolidation of customers. Time-based consolidation of customers, which means that orders are collected in a vehicle first until it departures, is considered first by Higginson and Bookbinder (1995). The vehicle departures, if its 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.

	assumptions		inventory decisions				methodology			
Paper	multiple products	cyclic	limited inventory at depot	lost sales	back-orders	substitution by different products	MILP	Markov chain	exact solution method	heuristics
Higginson and Bookbinder (1995)								✓		
Campbell et al. (1998), Campbell and Savelsbergh (2004)				✓						✓
Çetinkaya and Bookbinder (2003)					✓				✓	
Nolz et al. (2014)				✓			✓			✓
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)			✓				✓		✓	
Johansson et al. (2020)	✓		✓		✓					✓
Malicki and Minner (2021)		✓		✓			✓			✓[A]
Raa and Aouam (2023)		✓		✓			✓			✓
Cui et al. (2022)					✓		✓		✓	
Our paper	✓	✓	✓**		✓***		✓			✓[A]

Table 1: Overview of inventory optimization literature with shipment consolidation

[A] ALNS, * MILNP with additional linearization, ** depot has own stochastic demand, *** unmet demand is served via emergency delivery

Table 1 differentiates our assumptions, decisions, and methodology from the publications described before chronologically. Our paper adds to the available literature, particularly with respect to the following issues.

1. Presenting a 2E-SIRP with novel configuration of assumptions and decisions, especially:
   a) Combining both an $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ and an $(s_{p,\hat{c}}, Q_p)$ inventory policy in a two-echelon IRP.
   b) Depot has its own demand.
   c) Instant replenishment orders as recourse decision.
2. Formulating a two-stage stochastic program and developing a specialized ALNS for solving larger instances.
3. Presenting a novel and real-life case study of a large German clinic.

4. Two-stage stochastic program

In this section, we introduce the mathematical model in form of a two-stage stochastic program for the 2E-SIRP. In Section 4.1, we introduce 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 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. $\mathcal{T}_0 = 0...T$ is the index set for the periods including the initialization period 0. The standard delivery is permitted for all clinics in time periods $\bar{\mathcal{T}} \subseteq \mathcal{T}$. Set Ω defines the demand scenarios.

General problem structure. We present the mathematical model in form of a two-stage stochastic program. In the first stage, tactical decisions are taken on routing and reorder points. In the second stage, emergency deliveries as recourse decisions are executed for each scenario, if an inventory of a clinic is running out of stock.

Parameters. The total costs are composed out of travel costs $k^S_{i,j}$ per arc for standard delivery serving the surrounding clinics, $\hat{k}^S$ for serving the central clinic, $k^{em.}_{c,t}$ for emergency delivery as well as inventory holding costs k^I_p per period and product $p \in \mathcal{P}$.

For service of surrounding clinics, each van has a payload of K^S and each drone of $K^{em.}$, respectively. For this the volume v_p for each product is considered. Q_p equals the fixed delivery quantity for product $p \in P$ at the central clinic. 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 reorder points $s_{p,c}$ for product $p \in \mathcal{P}$ and clinic $c \in \hat{\mathcal{C}}$ and the routing $y_{i,j,t}$, i.e., if a van travels from clinic $i \in \hat{\mathcal{C}}$ to clinic $j \in \hat{\mathcal{C}}$ in period $t \in \bar{\mathcal{T}}$. Thus, $y_{i,j,t}$ gives information on the delivery days a clinic $j \in \mathcal{C}$ is served.

The following decisions are second stage decisions. 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 $c \in \hat{\mathcal{C}}$, $q^S_{i,j,t,\omega}$ the delivery amount by standard delivery, $w_{c,t,\omega}$ the number of emergency deliveries, and $q^{em.}_{p,c,t,\omega}$ the transported amount in emergency deliveries. Moreover, $z_{p,t,\omega}$ determines, if the central clinic is served in period $t \in \bar{\mathcal{T}}$ with product $p \in \mathcal{P}$, as the delivery days of the central clinic are not fixed and dependent on the reorder points.

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 serving the central clinic in standard delivery
$k^{em.}_{c,t}$	emergency delivery costs serving clinic $c \in \mathcal{C}$ at time $t \in \mathcal{T}$
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	sufficient small number
K^S	Maximum payload for service of surrounding clinics in standard delivery in volume equivalents

Continued on next page

$K^{em.}$	Maximum payload for service of surrounding clinics in emergency delivery in volume equivalents
M_l	Big M for $l = 1, ..., 11$
Q_p	Delivery quantity of product $p \in P$ for the central clinic in the standard delivery, quantified in entire packaging units
v_p	Volume of a package of product $p \in \mathcal{P}$
First stage decision variables	
$s_{p,c}$	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 $t \in \mathcal{T}$
First stage auxiliary variables	
$u_{i,t}$	Subtour-elimination variable for node $i \in \mathcal{C}$ at time $t \in \bar{\mathcal{T}}$ by Miller et al. (1960)
$I_{p,c}^{Init}$	Initial inventory of product $p \in \mathcal{P}$ at clinic $c \in \hat{\mathcal{C}}$
$\eta_{p,c,t}$, $\theta_{p,c,t}$, $\zeta_{p,c,t}$	Binary*, binary*, and real-value variables to ensure cyclic planning
Second stage auxiliary variables	
$I_{p,c,t,\omega}$	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 $\omega \in \Omega$
$q_{p,c,t,\omega}^{S}$	Integer variable, indicating transported amount of packages of product $p \in \mathcal{P}$ to clinic $c \in \mathcal{C}$ at time $t \in \bar{\mathcal{T}}$ and scenario $\omega \in \Omega$ in a standard delivery
$\phi_{i,j,t,\omega}^{S}$	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
$q_{p,c,t,\omega}^{em.}$	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
$r_{t,\omega}$	Binary variable*, indicating, if there is a standard resupply of the central clinic at time $t \in \bar{\mathcal{T}}$ and scenario $\omega \in \Omega$
$w_{c,t,\omega}$	Integer variable, indicating, the number of emergency resupplies for 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 $\omega \in \Omega$
$z_{p,t,\omega}$	Binary variable*, indicating, if there is a standard resupply of the central clinic at time $t \in \bar{\mathcal{T}}$ for product $p \in \mathcal{P}$ and scenario $\omega \in \Omega$
$\xi_{p,c,t,\omega}$, $\kappa_{p,c,t,\omega}$, $\iota_{p,c,t,\omega}$	Binary*, real-value, and real-value variable for soft constraints

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

4.2. Mathematical model

In this subsection, we present our two-stage stochastic program for the 2E-SIRP as a mixed-integer linear program. For a better reading we decomposed the two-stage stochastic program into the objective function and different constraints.

Objective function.

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

The objective function (1) minimizes the total expected logistics 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 in the second stage. The second term approximates expected costs, which are the average of additional delivery and inventory costs for all scenarios evaluated, i.e., standard delivery costs in the first echelon, emergency delivery costs in the second echelon, and inventory holding costs for all clinics.

Routing constraints.

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

$$\sum_{i,j\in\hat{\mathcal{C}}} y_{i,j,t} \leq M_1\cdot\sum_{j\in\mathcal{C}} y_{\hat{c},j,t} \qquad \forall t\in\bar{\mathcal{T}} \tag{3}$$

$$\sum_{i\in\hat{\mathcal{C}}} y_{i,j,t} \leq 1 \qquad \forall j\in\mathcal{C}, t\in\bar{\mathcal{T}} \tag{4}$$

$$u_{i,t}+1 \leq u_{j,t} + M_2\cdot(1-y_{i,j,t}) \qquad \forall i,j\in\mathcal{C}, t\in\bar{\mathcal{T}} \tag{5}$$

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

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

$$\sum_{i\in\hat{\mathcal{C}}t\in\bar{\mathcal{T}}} y_{i,j,t} \geq 1 \qquad \forall j\in\mathcal{C} \tag{8}$$

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

$$\phi^S_{i,j,t,\omega}\in\mathbb{R} \qquad \forall i,j\in\hat{\mathcal{C}}, t\in\bar{\mathcal{T}}, \omega\in\Omega \tag{10}$$

$$q^S_{p,c,t,\omega} \in \mathbb{N} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{11}$$

$$u_{i,t} \in \mathbb{R} \qquad \forall i \in \mathcal{C}, t \in \bar{\mathcal{T}} \tag{12}$$

Constraints (2) to (12) restrict the routing of the standard delivery in each period $t \in \bar{\mathcal{T}}$. Constraints (2) conserve the flow and Constraints (3) ensure that the van that serves the clinics in the second echelon always starts at the central clinic. M_1 can be set as $|\hat{\mathcal{C}}|$, since each clinic is served maximum once per period (Constraints (4)). Subtours are eliminated by Constraints (5) (Miller et al., 1960). M_2 can be set as $|\hat{\mathcal{C}}|$. Constraints (6) defines the accumulated transported volume of each van's tour. This accumulated transported volume must be less than the van's payload (Constraints (7)). Constraints (8) ensure that each surrounding clinic is served at least once in the considered periods. Last, variables are defined.

Reorder point constraints for the central clinic.

$$I_{p,\hat{c},t,\omega} - Q_p \cdot z_{p,t,\omega} \geq 0 \qquad \forall p \in \mathcal{P}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{13}$$

$$I_{p,\hat{c},t,\omega} - Q_p \leq s_{p,\hat{c}} \leq I_{p,\hat{c},t,\omega} - \varepsilon \qquad \forall p \in \mathcal{P}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{14}$$

$$\sum_{p \in \mathcal{P}} z_{p,t,\omega} \leq |\mathcal{P}| \cdot r_{t,\omega} \qquad \forall t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{15}$$

$$r_{t,\omega} \in \{0, 1\} \qquad \forall t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{16}$$

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

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

$$z_{p,t,\omega} \in \{0, 1\} \qquad \forall p \in \mathcal{P}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{19}$$

Constraints (13) to (19) define the reorder points for the central clinic. There are emergency deliveries only if the inventory at the end of the previous period is not sufficient high to meet all demand (Constraints (13)). Constraints (14) (right side) ensure that there is only a delivery, if inventory is below the reorder point. A standard delivery must be carried out, if inventory is below the reorder point (left side of Constraints (14)). Constraints (15) ensure that the products are consolidated in delivery. Last, variables are defined.

Reorder point constraints for surrounding clinics.

$$q^S_{p,c,t,\omega} \leq s_{p,c} - I_{p,c,t-1,\omega} + 1 - \varepsilon \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{20}$$

$$q^S_{p,c,t,\omega} \geq s_{p,c} - I_{p,c,t-1,\omega} - M_3 \cdot \left(1 - \sum_{i \in \hat{\mathcal{C}}} y_{i,c,t}\right) \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{21}$$

$$q^S_{p,c,t,\omega} \leq M_3 \cdot \sum_{i \in \hat{\mathcal{C}}} y_{i,c,t} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \bar{\mathcal{T}}, \omega \in \Omega \tag{22}$$

Constraints (20) to (22) define the reorder points $s_{p,c}$ and the resulting delivery quantities $q^S_{p,c,t,\omega}$ for the surrounding clinics. Constraints (20) and (21) ensure that the delivery quantities - which have an integer value - fill up the inventory to the reorder point $s_{p,c}$. M_3 can be set as the fraction of the payload of the van and the smallest volume of any product, e.g., one pill. Medicine is only delivered in a period to a clinic, if there is a standard delivery for this clinic in the considered period (Constraints (22)).

Emergency delivery constraints.

$$\sum_{p \in \mathcal{P}} q^{em.}_{p,\hat{c},t,\omega} \leq M_6 \cdot w_{\hat{c},t,\omega} \qquad \forall t \in \mathcal{T}, \omega \in \Omega \tag{23}$$

$$\begin{aligned} q^{em.}_{p,c,t,\omega} \leq\ & d_{p,c,t,\omega} - I_{p,c,t-1,\omega} + M_7 \cdot \big(1 - x_{p,c,t,\omega}\big) \\ & - \begin{cases} q^S_{p,c,t,\omega} & , t \in \bar{\mathcal{T}} \\ 0 & , else \end{cases} \end{aligned} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \mathcal{T}, \omega \in \Omega \tag{24}$$

$$\begin{aligned} q^{em.}_{p,\hat{c},t,\omega} \leq\ & d_{p,\hat{c},t,\omega} - I_{p,\hat{c},t-1,\omega} + M_7 \cdot \big(1 - x_{p,\hat{c},t,\omega}\big) \\ & + \sum_{k \in \mathcal{C}} q^{em.}_{p,k,t,\omega} + \begin{cases} \sum_{k \in \mathcal{C}} q^S_{p,k,t,\omega} & , t \in \bar{\mathcal{T}} \\ 0 & , else \end{cases} \end{aligned} \qquad \forall p \in \mathcal{P}, t \in \mathcal{T}, \omega \in \Omega \tag{25}$$

$$w_{c,t,\omega} \leq M_8 \cdot \sum_{p \in \mathcal{P}} x_{p,c,t,\omega} \qquad \forall c \in \hat{\mathcal{C}}, t \in \mathcal{T}, \omega \in \Omega \tag{26}$$

$$q^{em.}_{p,c,t,\omega} \leq M_8 \cdot x_{p,c,t,\omega} \qquad \forall p \in \mathcal{P}, c \in \hat{\mathcal{C}}, t \in \mathcal{T}, \omega \in \Omega \tag{27}$$

$$\sum_{p \in \mathcal{P}} v_p \cdot q^{em.}_{p,c,t,\omega} \leq K^{em.} \cdot w_{c,t,\omega} \qquad \forall c \in \mathcal{C}, t \in \mathcal{T}, \omega \in \Omega \tag{28}$$

$$w_{c,t,\omega} \in \mathbb{N} \qquad \forall c \in \mathcal{C}, t \in \mathcal{T}, \omega \in \Omega \tag{29}$$

$$x_{p,c,t,\omega} \in \{0,1\}, q^{em.}_{p,c,t,\omega} \in \mathbb{R}^+ \qquad \forall p \in \mathcal{P}, c \in \hat{\mathcal{C}}, t \in \mathcal{T}, \omega \in \Omega \tag{30}$$

Constraints (23) to (30) define the emergency delivery constraints for all clinics. Constraints (23) determine, if there is an emergency delivery for the central clinic. M_6 must be set sufficient high, as the payload of the emergency delivery to the central clinic is unlimited. Constraints (24) and (25) restrict the delivery quantity to the unsatisfied demand of the current period to avoid emergency deliveries enabling stockpiling. M_7 needs to be greater than all possible maximum inventory capacities of each clinic. Constraints (26) consolidate multiple products in one emergency delivery. Constraints (27) ensure that medicine is only delivered, if there is an emergency delivery. M_8 can be set as a maximum reasonable amount of transported volume in emergency deliveries. Constraints (28) restrict the payload of the emergency deliveries for surrounding clinics. Last, auxiliary variables are defined.

Inventory constraints.

$$I_{p,c,t,\omega} = I_{p,c,t-1,\omega} - d_{p,c,t,\omega} + q^{em.}_{p,c,t,\omega} + \begin{cases} q^{S}_{p,c,t,\omega} & ,t \in \bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \mathcal{T}, \omega \in \Omega \quad (31)$$

$$\begin{aligned} I_{p,\hat{c},t,\omega} = {} & I_{p,\hat{c},t-1,\omega} - d_{p,\hat{c},t,\omega} + q^{em.}_{p,\hat{c},t,\omega} - \sum_{k \in \mathcal{C}} q^{em.}_{p,k,t,\omega} \\ & + \begin{cases} Q_p \cdot z_{p,t,\omega} - \sum_{k \in \mathcal{C}} q^{S}_{p,k,t,\omega} & ,t \in \bar{\mathcal{T}} \\ 0 & ,else \end{cases} \qquad \forall p \in \mathcal{P}, t \in \mathcal{T}, \omega \in \Omega \quad (32) \end{aligned}$$

$$\begin{aligned} I^{Init}_{p,c} + \zeta_{p,c,\tau} = {} & s_{p,c} + M_9 \cdot \sum_{t=\tau+1..|\bar{\mathcal{T}}|, i \in \hat{\mathcal{C}}} y_{i,c,t} \\ & - \sum_{t=\tau..|\mathcal{T}|} \mathbb{E}_{\omega \in \Omega} \Big\{ d_{p,c,t,\omega} \Big\} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, \tau \in \mathcal{T} \quad (33) \end{aligned}$$

$$M_{10} \cdot (\theta_{p,c,t} - 1) \leq \zeta_{p,c,t} \leq M_{10} \cdot \eta_{p,c,t} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \mathcal{T} \quad (34)$$

$$\sum_{t \in \mathcal{T}} \eta_{p,c,t} = |\mathcal{T}| - 1 \qquad \forall p \in \mathcal{P}, c \in \mathcal{C} \quad (35)$$

$$\begin{aligned} \zeta_{p,c,\tau} \geq {} & s_{p,c} + M_9 \cdot \sum_{t=\tau+1..|\bar{\mathcal{T}}|, i \in \hat{\mathcal{C}}} y_{i,c,t} \\ & - \sum_{t=\tau..|\mathcal{T}|} \mathbb{E}_{\omega \in \Omega} \Big\{ d_{p,c,t,\omega} \Big\} - M_{11} \cdot \theta_{p,c,\tau} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, \tau \in \mathcal{T} \quad (36) \end{aligned}$$

$$I^{Init}_{p,\hat{c}} = \frac{s_{p,\hat{c}} + Q_p}{2} \qquad \forall p \in \mathcal{P} \quad (37)$$

$$I_{p,c,0,\omega} = I^{Init}_{p,c} \qquad \forall p \in \mathcal{P}, c \in \hat{\mathcal{C}}, \omega \in \Omega \quad (38)$$

$$I^{Init}_{p,c} \in \mathbb{R}^+ \qquad \forall p \in \mathcal{P}, c \in \hat{\mathcal{C}} \quad (39)$$

$$\eta_{p,c,t}, \theta_{p,c,t} \in \{0,1\}, \zeta_{p,c,t} \in \mathbb{R} \qquad \forall p \in \mathcal{P}, c \in \mathcal{C}, t \in \mathcal{T} \quad (40)$$

Constraints (31) to (40) are inventory constraints for all clinics. Constraints (31) 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 for the central clinic, but the inventory is additionally reduced by the total quantity delivered to clinics in the second echelon in standard and emergency deliveries (Constraints (32)). Constraints (33) are the cyclic formulations for the surrounding clinics. The initial inventory is dependent on the last replenishment in the cyclic planning periods, i.e. the reorder points reduced by the expected demand from the last delivery on. M_9 can be set as the maximum reasonable inventory of a product. $\eta_{p,c,t}$ and $\zeta_{p,c,t}$ are auxiliary variables that ensure equality holds. Only for one period $\zeta_{p,c,t}$ is set to zero (Constraints (34) and (35)). Constraints (36) ensure that $\zeta_{p,c,t}$ might have negative values, if the reorder points are lower than the expected demand. Constraints (37) set the initial inventory for the central clinic as the half of both the reorder point plus the order quantity. Constraints (38) defines the inventory at period 0 for all scenarios. Last, auxiliary variables are defined.

5. Adaptive large neighborhood search

To solve the 2E-SIRP, we decompose the problem in its two echelons. We develop a problem specific ALNS to generate routes and delivery days. Using this routing, we determine optimal reorder points in each iteration. Thus, this approach leads to a focus of the ALNS on the second echelon of the problem.

The structure of this section is as follows: First, we introduce the ALNS and describe the solution representation in Subsection 5.1, the pseudocode of our ALNS in Subsection 5.2, our routing specific operators in Subsection 5.3, and the algorithm to find local optimal reorder points in Subsection 5.4.

5.1. Solution representation

The ALNS was first introduced in Røpke and Pisinger (2006) and Pisinger and Røpke (2007) for the class of vehicle routing problems and has been extended for inventory routing problems (e.g., Aksen et al., 2014). Our ALNS is specialized for the $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy of clinics on the second echelon and the $(s_{p,\hat{c}}, Q_p)$ inventory policy of the central clinic on the first echelon, which means that the delivery quantities are dependent on the inventory policies.

A solution is represented by the routing of the standard delivery (R) including information on delivery days within the cyclic planning horizon and the reorder points for the clinics ($s_{p,c}$). In each iteration operators are chosen that take changes on just the routing R (see Section 5.3. Afterwards, the reorder points $s_{p,c}$ are optimized depending on the created routing R (see Section 5.4).

Within these changes, infeasible solutions are accepted, but penalized in the objective function dependent on the degree of infeasibility (e.g., Vidal et al., 2013; Rave et al., 2023). Infeasibility may occur in two different ways: first, a van may exceed its payload in a scenario and second, not all clinics are served at least once.

5.2. ALNS algorithm

The general outline of the ALNS is presented in Algorithm 1. The algorithm begins with an initial solution (line 1) and a temperature $Temp$ for the simulated annealing process (line 2). The initial solution is generated by serving all clinics once and only in the first period of the cyclic planning periods. The clinics are inserted one by one in the greedy best way in one tour. This initial solution is not necessarily feasible. The while-loop in line 3 stops after a maximum time. At the beginning of each iteration, operators and a delivery day T for the repair operator are chosen adaptively (line 4). In line 5 operators modify the routing R. Local optimal reorder points $\bar{s}_{p,c}$ for this routing $\bar{R}$ are computed afterwards in our reorder point algorithm (lines 6), see Algorithm 2, and delivery quantities, inventories and emergency deliveries are updated.

The simulated annealing process is described in lines 7 - 10. The costs f are based on changes made by operators by Algorithm 2. The solution is always accepted, if it is better than the previous global solution (lines 7 - 8). Depending on the temperature $Temp$ the new solution ($\bar{R}$, $\bar{s}_{p,c}$) is accepted (lines 9 - 10), if it is worse than the global best solution. Last, the temperature and the weights of the operators and delivery days are updated, dependent on the performance of this iteration.

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

$(R,\ R^{global},\ s_{p,c},\ s_{p,c}^{global}) \leftarrow$ Initial Solution;
$Temp \leftarrow$ GetInitialTemp();
while $Time < MaxTime$ **do**
    Choose(Operator, T);
    $\bar{R} \leftarrow$ Operator(R);
    $\bar{s}_{p,c} \leftarrow$ Optimize($s_{p,c}$);   *//see Algorithm 2*
    if $f(\bar{R}, \bar{s}_{p,c}) < f(R^{global}, s_{p,c}^{global})$ **then**
        $(R,\ R^{global},\ s_{p,c},\ s_{p,c}^{global}) \leftarrow (\bar{R},\ \bar{s}_{p,c})$;
    else if $accept(f(\bar{R},\ \bar{s}_{p,c}),\ f(R, s_{p,c}),\ Temp)$ **then**
        $(R, s_{p,c}) \leftarrow (\bar{R},\ \bar{s}_{p,c})$;
    Temp $\leftarrow$ UpdateTemp($Temp$);
    UpdateWeights();
end

5.3. Operators

The used removal, selection, and insertion operators are extensions of the operators of Røpke and Pisinger (2006), Pisinger and Røpke (2007), Aksen et al. (2014), and Rave et al. (2023). A selection operator only chooses clinics for insertion, in contrast to a removal operator that removes these clinics. In each iteration operators are chosen in the following way, as the removal of customers decreases and an insertion of clinics increases the number of servings per clinic, which, however, is variable.

1. Remove clinic, insert clinic *(two operators chosen)*
2. Select clinic, insert clinic *(two operators chosen)*
3. Remove clinic *(one operator chosen)*
4. Shift delivery day of one route *(one operator chosen)*

We consider two operators that remove or select α clinics, three operators that insert these removed or selected clinics, and one operator that shifts the delivery day. α is randomly drawn (uniform) with a limited number of clinics.

Random removal/selection - The random removal/selection operator removes/selects α clinics from any route and delivery day.

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

Greedy insertion - This operator adds clinics one by one to existing van tours at day T in a greedy best way taking traveled distances into account. If there has not been a route at this day, 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 the route.

Greedy insertion - new route - Similar to the previous operator, this operator inserts clinics always in one new route at the selected day T.

Delivery day shift - The delivery day of one route is shifted to the selected day T. If clinics have already been served at that day, these clinics are removed from this shifted route.

5.4. Reorder point algorithm

Algorithm 2 describes the pseudo code of the reorder point algorithm for calculating the reorder points $s_{p,c}$ and is executed for each clinic, separately. For this, we use the integrality of $s_{p,c}$ to compute local optimal reorder points $s_{p,c}$ in each iteration of the ALNS. In the beginning a reorder point $s_{p,c}^{new}$, the best reorder point $s_{p,c}$ found so far, and the considered routing R are initialized (line 1). As initial reorder point the reorder point of the routing before modification by operators is chosen. While there are still cost improvements (line 2) by adjusting $s_{p,c}^{new}$, $s_{p,c}$ is updated (line 3) and $s_{p,c}^{new}$ is either increased or decreased by 1 dependent on which change has a larger benefit (line 4).

Algorithm 2: Reorder point algorithm to determine $s_{p,c}$ on a given R

1 Initialize ($s_{p,c}^{new}$, $s_{p,c}$, R);
2 **while** $f(R, s_{p,c}^{new}) < f(R, s_{p,c})$ **do**
3 $\quad s_{p,c} \leftarrow s_{p,c}^{new}$;
4 $\quad s_{p,c}^{new} \leftarrow \text{adjust}(s_{p,c}^{new})$;
5 **end**

The calculation of $s_{p,c}$ for a clinic and product requires a complex evaluation of the objective function and also has dependencies with the reorder points of both the central clinic and even other surrounding clinics, because the payload in standard delivery is limited. These dependencies are circumvented by randomly selecting the sequence of clinics and products the algorithm is applied to. Thus, the found reorder points are local optimal and not necessarily global optimal.

6. Numerical Study

In this section, we show the performance of our ALNS and present a case study, which is based on the project MEDinTime (Quantum Systems, 2020). First in Section 6.1, we describe the numerical setup in detail and in Section 6.2, we show the performance of our ALNS compared to optimal results solving the two-stage stochastic program in CPLEX. In Section 6.3, we present results for the case study and perform a sensitivity analysis where we analyze the effect of drone delivery as instant replenishment orders on inventory pooling effects. 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 3950X with 32 GB.

6.1. Numerical setup

Location and number of clinics. The central clinic, which is Ingolstadt Hospital, serves nine surrounding clinics, which are located between 3 and 75 km (euclidean distance) from the central clinic. The surrounding clinics are served in a weekly repeating manner and all clinics are in range of a drone delivery.

For a performance analysis, we also consider three random locations of clinics within a 100km $\times$ 100km map and the five locations as in Archetti et al. (2007) for 15 clinic locations, which are also conducted by, e.g., Cui et al. (2022). For the instances by Archetti et al. (2007) we assume similar demands as in Archetti et al. (2007) and Cui et al. (2022), i.e. daily demands with randomly chosen integer values between one and five for each clinic that are constant over time

Cost structure. There is a variety of medication that is needed for patient's 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% cheaper and thus equals 0.672€ per km.

Each inventory replenishment at the central clinic is assumed to have delivery costs of 168€ for standard delivery ($\hat{k}^S$) and twice as high for emergency delivery ($\hat{k}^{em.}$). Due to weekend emergency staffing at all clinics, emergency delivery costs are always increased by a factor of three at this time.

A standard supply for surrounding clinics includes approximately 200 different products, for which the transportation costs are total. Thus, these costs are "fixed" costs and are reduced to the considered number of products, as we only consider a portion of the products. Considering only a share of these products leads to similar decisions but can save sufficient running time (e.g., Rahimi et al., 2017). This proportional cost adjustment is applied to all standard delivery and emergency delivery costs, as in each delivery, different products might be consolidated.

Demand. The demand of 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. The variance is assumed to be five times as high as the mean of the distribution (Johansson et al., 2020). 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.

Due to multiple products, the distribution of the scenarios is potentiated. Thus, to reduce this probability space, we use the latin hypercube sampling by McKay et al. (1979) to generate demands of each product, within each period for each hospital.

Delivery quantities and capacities. We assume that drugs that are high in demand are delivered in larger packages, which results in larger volume equivalents per package.

The payload of a van is based on the Mercedes Sprinter (Mercedes-Benz AG, 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 payload K^S is dependent on the considered drugs (e.g., $K^S = 138.7$ volume equivalents, if one product of each category is considered). The considered drone has a payload of 3.765 volume equivalents ($K^{em.}$).

Table 3 presents the parameter settings for the seven different products categorized by their price ranges, the resulting inventory holding costs per day k_p^I, the delivery quantity for the central clinic Q_p, their volume v_p, and their average mean and variance, which need to be adjusted based on the size, i.e., the bed number, for each clinic.

$p \in \mathcal{P}$	**Price ranges [€]**	k_p^I [€]	Q_p	v_p	**Average mean of demand**	**Average variance of demand**
1	< 12	0.0033	100	2	1.29	6.43
2	12 - 58	0.0193	30	2	0.21	1.07
3	59 - 129	0.0514	20	1	0.09	0.43
4	130 - 291	0.1153	15	1	0.06	0.32
5	292 - 778	0.2933	10	1	0.04	0.21
6	779 - 1758	0.6951	5	0.4	0.02	0.11
7	> 1759	2.4644	3	0.4	0.01	0.06

Table 3: Product categories

ALNS parameter. The ALNS parameters are based on pre-testing. In each iteration a randomly drawn number of one to all clinics is selected/removed, if a selection/removal operator is chosen. The reaction factor for the learning curve and the weight adjustment are set similar 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 by a cool rate dependent on the remaining run time of the ALNS.

6.2. Performance analysis

In this section we first benchmark our ALNS to solutions generated by CPLEX (Section 6.2.1) and next we demonstrate the stability of the found solution (Section 6.2.2).

6.2.1. Small instances

Benchmark instances in the literature cannot be considered, as there is no publication with a similar problem setting in the literature to our knowledge. The performance of the ALNS is compared to the solutions obtained using CPLEX for the two-stage stochastic program. The run time limit for the two-stage stochastic program was set to 60 minutes and for the ALNS to five minutes. The ALNS is executed five times. Table 4 presents aggregated results for 78 instances. Detailed results are provided in Table 8 in the Appendix. Column $|\hat{\mathcal{C}}|$

describes the number of surrounding clinics, column $|\mathcal{P}|$ number of considered products, and column $|\Omega|$ the number of demand scenarios. The next four columns describe the best cost, CPLEX found, the best cost the ALNS found, the average cost found, and the deviation between best and average costs. Column 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 run time.

			cost objective				CPLEX		run time (s)	
$\lvert\hat{\mathcal{C}}\rvert$	$\lvert\mathcal{P}\rvert$	$\lvert\Omega\rvert$	CPLEX	ALNS best	ALNS avg	ALNS σ [%]	Opt	GAP [%]	CPLEX	ALNS
4	1	5	4.45	4.45	4.46	0.59	12 / 12	-	82	78
4	2	5	8.37	8.27	8.28	0.23	3 / 4	28	1465	183
4	3	5	16.70	16.50	16.51	0.12	1 / 4	21	2794	225
7	1	2	6.27	6.22	6.24	1.02	6 / 12	21	1940	207
7	2	2	14.47	12.49	12.51	0.12	0 / 4	42	3600	300
7	3	2	40.27	22.85	22.93	0.34	0 / 4	63	3600	300
10	1	1	13.51	13.46	13.49	1.01	17 / 28	24	1192	136
15	1	1	23.36	20.93	21.21	1.34	2 / 10	27	3083	278
avg			13.59	12.24	12.30	0.38	43 / 78	23	1722	180

Table 4: Performance analysis for small instances

Findings: The results show that CPLEX can only solve 43 out of the 78 small instances to optimality. For the other 35 instances, the gaps are rather large. The ALNS, on the other hand, finds for all clusters the same or better costs on average with up to 56% less costs by using only 10% of the run time. The results of the ALNS are stable as σ has a value of 0.38% on average for all instances.

6.2.2. Large instances

In this section, we demonstrate the stability of the solutions of our ALNS. For this, we compare the routing for ten instances for the location of clinics as in our case study. We choose a number of 100 varying scenarios and one product of each product category. Thus, the instances only vary in their demand scenarios. The number of scenarios is sufficient as a maximum desired deviation of 10% is not exceeded for a confidence interval of 95%. It is tested for the total costs, number of emergency deliveries, and average amount of reorder points. The ALNS is executed five times with a run time limit of 60 minutes in each run, as the solution is used for a tactical planning problem (Rave et al., 2023). Table 5 presents for each instance the best routing found at the corresponding day, the best costs, the average costs, and the deviation from average to best costs (σ). The routing always starts and ends with the central clinic. The driving order equals the order of the numbered clinics in the squared brackets. If there is a second route at the same weekday, the routing is presented in a second row.

	Best found routing at weekdays							
Instance	Mo.	Tu.	We.	Th.	Fr.	Best costs	Avg. costs	σ [%]
1				[8]-[3]-[5] [7]	[1]-[6]-[2]-[4]-[9]	167.70	168.52	0.49
2				[8]-[3]-[5]-[7]	[1]-[6]-[2]-[4]-[9]	167.78	168.53	0.44
3				[8]-[3]-[5] [4]-[9]	[7]-[1]-[6]-[2]	171.58	172.58	0.58
4				[8]-[3]-[5]	[7]-[1]-[6]-[2]-[4]-[9]	166.48	167.81	0.79
5				[8]-[3]-[5] [4]-[9]	[7]-[1]-[6]-[2]	172.77	173.58	0.46
6				[8]-[3]-[5]-[7]	[1]-[6]-[2]-[4]-[9]	171.84	173.08	0.72
7				[9]-[8]-[3]-[5]-[7]	[1]-[6]-[2]-[4]	170.43	171.51	0.63
8				[8]-[3]-[5]-[7]	[1]-[6]-[2]-[4]-[9]	166.63	167.74	0.67
9				[8]-[3]-[5] [4]-[9]	[7]-[1]-[6]-[2]	173.63	174.69	0.61
10				[8]-[3]-[5]-[7]	[1]-[6]-[2]-[4]-[9]	167.43	168.19	0.45

Table 5: Performance analysis for large instances

Findings: The results are stable in costs as σ has always values below 1% and each clinic is served once, either on Thursday, or on Friday. The serving is just after the weekly demand peak in clinics to prevent the additional emergency delivery costs at the weekend. Clinics [8],[3], and [5] are always served in the same order and day. The same holds for clinics [1], [6], and [2]. Clinics [4], [7], and [9] are added to different routes, or even new routes, in the instances under consideration. However, this variation does not lead to large cost differences. Clinic [7] is mainly served on Thursday and [4] and [9] mainly on Friday. To sum up, the results appear very satisfactory as the costs are stable and the best routing has only little variation.

6.3. Case study

In this section, we present the results of the case study, i.e., we demonstrate advantages of an optimization of reorder points and routing compared to the status quo of clinic supply. The status quo is as follows: Each surrounding clinic determines its delivery days and reorder points on its own. There are in total about 250 emergency deliveries per year, as emergency deliveries are avoided if possible due to their extra costs. Subsequently, the central clinic plans the routing on the respective days 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 2 to the routing of the status quo. Second, additionally the routing is optimized and the results are again compared to the status quo. This is done by executing the ALNS. For both, we consider 50 instances á one product of each product category, 100 scenarios, and the clinic location as in our case study and we set the run time limit of the ALNS again to 60 minutes. We obtain the status quo results by running the reorder point algorithm that chooses reorder points in that way that there are 250 emergency deliveries for surrounding clinics per year.

The following KPIs are compared in Table 6 all in comparison to the status quo in percent: total costs for the central clinic and for surrounding clinics (1. row), total number of standard delivery servings per week for the surrounding clinics (2. row), the average number of emergency deliveries per week for surrounding clinics (3. row), and last the reorder points for each product category. Each KPI consolidates all nine surrounding clinics.

	Inventory optimization [%]		Inventory & routing optimization [%]	
	Central clinic	Surrounding clinics	Central clinic	Surrounding clinics
Costs	− 16.0	− 45.7	− 15.9	− 58.4
# Standard Del.		± 0.0		− 47.1
# Emergency Del.		+ 185.9		+ 160.9
Reorder points				
product category 1	+ 3.7	− 9.7	+ 46.6	− 1.9
product category 2	+ 4.6	− 42.4	+ 19.8	− 32.4
product category 3	+ 3.8	− 64.4	+ 16.8	− 56.7
product category 4	− 14.0	− 76.6	+ 4.5	− 64.1
product category 5*	− 60.0	− 99.2	− 59.3	− 88.9
product category 6*	− 53.6	− 100.0	− 41.7	− 100.0
product category 7*	− 100.0	− 100.0	− 100.0	− 100.0

Table 6: Percentage change from status quo with cost-optimal inventory (2nd and 3rd column) and with simultaneous route and inventory optimization (4th and 5th column)
* Smaller clinics do not have any inventory of this product category already in the status quo

Findings: The inventory optimization leads to cost savings of 16.0% for the central clinic and 45.7% for the surrounding clinics, as demand peaks are eliminated. Additional cost savings of 12.7 percentage points can be reached for the surrounding clinics, if the routing is planned simultaneous. However, this additional routing optimization does have nearly no an impact on costs for the central clinic.

The cost savings are mainly achieved by applying the following: First, clinics make use of pooling effects, i.e., less inventory to be stored in surrounding clinics, but more in the central clinic. The following applies in particular: The more expensive a product, the stronger the pooling effect. Even after inventory optimization, the reorder points for product category 7 are set to 0 for all clinics which means that all demand is satisfied via emergency deliveries. It may not even be necessary to store more in the central clinic, when only split drugs are transported. Second, surrounding clinics decrease number of standard deliveries to one service per clinic per week, but increase the number of emergency deliveries by 185.9% or 160.9%, respectively.

Impact of drone delivery. Last, the influence of drone delivery as instant replenishment orders on total costs and inventory pooling effects is analyzed. Thus, in Table 7, we compare results for the case, if the emergency delivery is carried out by drones, to the results for the case, if the emergency delivery is carried out by the same vans as for standard delivery. Van delivery on the one hand has higher costs compared to drone delivery, but on the other hand has an unlimited payload such that there is at maximum one emergency delivery per clinic per day. We first consider costs for drone delivery as in our case study (first two columns) and second drone delivery costs reduced by a factor of two (last two columns).

Findings: Drones as emergency delivery option lead to cost savings of 29.8% for surrounding clinics, but have nearly no impact on the costs for the central clinic. The cost savings for surrounding clinics can be increased to 49.8%, 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. Once again, the more expensive a product, the stronger the pooling effects. This has no effect on product category 7, as this already has a reorder point of 0 at van delivery for all clinics. 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 days, thus avoiding demand peaks.

	Cost advantage of emergency delivery by drones			
	52% [%]		76% [%]	
	Central clinic	Surrounding clinics	Central clinic	Surrounding clinics
Costs	− 0.6	− 29.8	− 0.6	− 49.8
# Standard Del.		− 0.7		− 0.7
# Emergency Del.		+ 22.5		+ 53.5
Reorder points				
product category 1	− 0.2	− 1.7	− 3.0	− 9.3
product category 2	− 0.1	− 12.0	− 4.5	− 36.1
product category 3	− 0.5	− 30.7	− 1.8	− 48.2
product category 4	− 4.9	− 33.3	− 22.2	− 67.2
product category 5	− 10.9	− 63.2	− 4.7	− 100.0
product category 6	+ 2.1	− 100.0	+ 6.3	− 100.0
product category 7*	./.	./.	./.	./.

Table 7: Percentage change from van delivery as emergency delivery option with drone delivery when drone delivery is 52% (2nd and 3rd column) and 76% (4th and 5th column) more cost efficient
* All clinics have a reorder point equal to 0 for this product category

7. Conclusion

This paper considers a cyclic routing planning combined with inventory policies and drones as emergency delivery options for recourse decision for medical supply of clinics. We introduce a two-echelon IRP that takes a cyclic $(\mathcal{T}_c, s_{p,c}, n_{p,c} \cdot q_p)$ inventory policy for surrounding clinics in the second echelon and a $(s_{p,\hat{c}}, Q_p)$ policy for the central clinic in the first echelon into account. We represent the problem as a two-stage stochastic program and develop a specialized ALNS that determines local-optimal reorder points $s_{p,c}$ in each iteration. In a performance analysis we show the solution quality and stability of the ALNS.

The case study shows that the use of pooling effects leads to total cost savings of 16.0% for the central clinic and 45.7% for surrounding clinics. The cost savings for surrounding clinics can even be increased by additional 12.7 percentage points, if the routing is optimized simultaneously. The study also shows that drones as more cost-efficient emergency delivery option reduces costs for surrounding clinics by 29.8% compared to an emergency delivery option via van. However, drones have nearly no impact on the 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. 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.

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.

References

Aksen, D., Kaya, O., Sibel Salman, F., 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.
URL https://doi.org/10.1287/trsc.1060.0188

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

Asadi, A., Nurre Pinkley, S., Mes, M., 2022. A markov decision process approach for managing medical drone deliveries. Expert Systems with Applications 204, 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.

Campbell, A., Clarke, L., Kleywegt, A., Savelsbergh, M., 1998. The inventory routing problem. In: Fleet Management and Logistics. Kluwer Academic Publishers, pp. 95–113.

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., 2022. The inventory routing problem under uncertainty. Operations Research.

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.

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, 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 AG, 2022. Technical data, weights and measures of the Sprinter Panel Van. World Wide Web, last access: December 01, 2022.
URL https://www.mercedes-benz.ie/vans/en/sprinter/panel-van/technical-data

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

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., Selcuk Erenguc, S., 2004. Outsourcing inventory management decisions in healthcare: Models and application. European Journal of Operational Research 154 (1), 271–290.

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, last access: December 01, 2022.
URL https://www.quantum-systems.com/category/medintime/

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.

Stenius, O., Karaarslan, A. G., 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.

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.

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

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.

				CPLEX			ALNS				
Inst.	$\lvert\hat{\mathcal{C}}\rvert$	$p \in \mathcal{P}$	$\lvert\Omega\rvert$	Costs	Gap [%]	Time [s]	Best costs	Avg. costs	σ [%]	Time [s]	Δ [%]
M-1	4	1	5	2.4330	-	37	2.4330	2.4330	-	27	-
R1-2	4	1	5	2.9737	-	43	2.9737	2.9737	-	1	-
R2-3	4	1	5	2.1938	-	1	2.1938	2.1938	-	0	-
R3-4	4	1	5	2.7103	-	1	2.7103	2.7103	-	28	-
M-5	4	2	5	5.9270	-	841	5.9271	5.9564	0.49	300	-
R1-6	4	2	5	3.8962	-	41	3.8962	3.8962	-	35	-
R2-7	4	2	5	4.5830	-	3	4.5830	4.6248	0.90	120	-
R3-8	4	2	5	4.8799	-	6	4.8799	4.8869	0.14	173	-
M-9	4	3	5	4.6614	-	3	4.6614	4.6621	0.01	63	-
R1-10	4	3	5	7.0386	-	2	7.0386	7.0386	-	7	-
R2-11	4	3	5	5.6938	-	3	5.6938	5.7741	1.39	181	-
R3-12	4	3	5	6.3950	-	5	6.3950	6.3950	-	0	-
M-13	4	1,2	5	10.2422	28	3600	9.8354	9.8354	-	300	-3.97
R1-14	4	1,2	5	7.8025	-	2052	7.8025	7.8117	0.12	179	-
R2-15	4	1,2	5	7.2629	-	118	7.2629	7.2629	-	0	-
R3-16	4	1,2	5	8.1623	-	91	8.1623	8.1900	0.34	253	-
M-17	4	1,2,3	5	17.2051	21	3600	16.2412	16.2412	-	300	-5.6
R1-18	4	1,2,3	5	16.9167	27	3600	16.9167	16.9430	0.15	300	-
R2-19	4	1,2,3	5	14.6688	-	377	14.6688	14.6688	-	0	-
R3-20	4	1,2,3	5	18.0214	14	3600	18.1574	18.1712	0.08	300	0.75
M-21	7	1	2	3.9107	29	3600	3.7736	3.7736	-	300	-3.51
R1-22	7	1	2	5.5495	12	3600	5.3368	5.3368	-	300	-3.83
R2-23	7	1	2	4.7262	20	3600	4.6513	4.6523	0.02	300	-1.58
R3-24	7	1	2	6.4221	30	3600	6.1125	6.1125	-	300	-4.82
M-25	7	2	2	6.3795	-	657	6.3795	6.3795	-	0	-
R1-26	7	2	2	7.6421	19	3600	7.6421	7.6421	-	300	-
R2-27	7	2	2	4.7228	-	24	4.7228	4.7240	0.03	142	-
R3-28	7	2	2	7.9691	15	3600	7.9691	7.9691	-	300	-
M-29	7	3	2	7.2791	-	397	7.2791	7.2791	-	0	-
R1-30	7	3	2	4.3390	-	2	4.3390	4.3390	-	0	-
R2-31	7	3	2	6.8301	-	127	6.8301	7.0485	3.10	240	-
R3-32	7	3	2	9.5096	-	476	9.5628	9.6539	0.94	300	0.56
M-33	7	1,2	2	16.3952	57	3600	12.2557	12.2603	0.04	300	-25.25
R1-34	7	1,2	2	15.1775	55	3600	12.9558	12.9743	0.14	300	-14.64
R2-35	7	1,2	2	10.9020	37	3600	10.3046	10.3046	-	300	-5.48
R3-36	7	1,2	2	15.3906	20	3600	14.4562	14.4833	0.19	300	-6.07
M-37	7	1,2,3	2	33.9267	64	3600	22.4928	22.5572	0.29	300	-33.7
R1-38	7	1,2,3	2	51.5159	78	3600	21.7242	21.8438	0.55	300	-57.83
R2-39	7	1,2,3	2	22.0797	40	3600	19.7454	19.7859	0.20	300	-10.57
R3-40	7	1,2,3	2	53.5479	70	3600	27.4347	27.5259	0.33	300	-48.77
M-41	10	1	1	5.1196	-	342	5.1196	5.1196	-	3	-
R1-42	10	1	1	5.7673	44	3600	4.8426	4.8524	0.20	300	-16.03
R2-43	10	1	1	4.5026	16	3600	4.5026	4.5127	0.22	300	-
R3-44	10	1	1	5.7289	20	3600	5.7289	5.7289	-	300	-
R1-45	10	2	1	7.0747	17	3600	7.0747	7.0747	-	300	-
M-46	10	2	1	7.6590	22	3600	7.6201	7.6244	0.06	300	-0.51
R2-47	10	2	1	5.3971	-	33	5.3971	5.3971	-	0	-
R3-48	10	2	1	8.1021	38	3600	7.7737	7.7737	-	300	-4.05
M-49	10	3	1	9.5964	20	3600	9.5964	9.6423	0.48	300	-

Continued on next page

R1-50	10	3	1	4.8449	-	5	4.8449	4.8449	-	22	-
R2-51	10	3	1	5.4927	-	7	5.4927	5.4927	-	0	-
R3-52	10	3	1	10.3821	-	128	10.3821	10.3821	-	51	-
M-53	10	4	1	7.4996	-	6	7.4996	7.6642	2.15	198	-
R1-54	10	4	1	13.2423	-	325	13.2423	13.2423	-	112	-
R2-55	10	4	1	12.4876	8	3600	12.4876	12.4876	-	300	-
R3-56	10	4	1	16.1072	-	63	16.1071	16.4444	2.05	218	-
M-57	10	5	1	13.7504	-	5	13.7504	13.7504	-	25	-
R1-58	10	5	1	11.7721	-	2	11.7721	11.7721	-	84	-
R2-59	10	5	1	11.6002	-	2	11.6002	11.6002	-	39	-
R3-60	10	5	1	13.6624	-	3	13.6624	13.6624	-	0	-
M-61	10	6	1	17.6657	-	23	17.6657	18.0118	1.92	184	-
R1-62	10	6	1	24.8828	-	23	24.8828	24.8828	-	27	-
R2-63	10	6	1	9.5174	-	1	9.5174	9.5174	-	0	-
R3-64	10	6	1	21.4008	-	4	21.4008	21.4008	-	49	-
M-65	10	7	1	17.1729	-	1	17.1729	17.1729	-	47	-
R1-66	10	7	1	41.5341	-	4	41.5341	41.5341	-	29	-
R2-67	10	7	1	38.1420	28	3600	38.1420	38.1420	-	300	-
R3-68	10	7	1	28.1036	-	1	28.1036	28.1036	-	14	-
A-69	15	1	1	15.9818	-	1265	15.9818	15.9919	0.06	107	-
A-70	15	1	1	16.4826	-	761	16.4826	16.608	0.75	269	-
A-71	15	1	1	16.2164	6	3600	16.2164	16.4751	1.57	300	-
A-72	15	1	1	16.7131	4	3600	16.7131	17.3854	3.87	300	-
A-73	15	1	1	17.6030	4	3600	17.6030	17.6117	0.05	300	-
A-74	15	1	1	27.8470	42	3600	27.4362	27.6236	0.68	300	-1.50
A-75	15	1	1	27.3037	42	3600	26.6746	27.2452	2.09	300	-2.36
A-76	15	1	1	16.2097	5	3600	16.0320	16.3224	1.78	300	-1.11
A-77	15	1	1	27.9306	42	3600	27.4960	27.8916	1.42	300	-1.58
A-78	15	1	1	51.2633	70	3600	28.6530	28.9684	1.09	300	-78.91

Table 8: Performance analysis for small instances (* optimal solution)
The clinics have the locations as in our case study (“M“), or three random locations (“R1“,“R2“,“R3“), or as in Archetti et al. (2007) (“A“) for all clinics all with a hyphen separated consecutive numbering.