Chemical Engineering Research and Design 186 (2022) 97–111

Available online at www.sciencedirect.com

Chemical Engineering Research and Design

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

Integrated production and inventory routing planning of oxygen supply chains

Yena Lee[a], Vassilis M. Charitopoulos[a], Karthik Thyagarajan[b], Ian Morris[c], Jose M. Pinto[b], Lazaros G. Papageorgiou[a,*]

[a] The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, UCL (University College London), Torrington Place, London WC1E 7JE, UK
[b] Linde Digital Americas, Linde PLC, 10 Riverview Drive, Danbury CT, 06810, US
[c] EMEA Production & Supply, BOC, Bawtry Road, Brinsworth, Rotherham S60 5NT, UK

A R T I C L E I N F O

Article history:
Received 15 February 2022
Received in revised form 18 July 2022
Accepted 19 July 2022
Available online 25 July 2022

Keywords:
Hybrid approach
Mixed-integer linear programming
Metaheuristics
Guided local search
Integrated oxygen supply chain planning

A B S T R A C T

In this work, we address a production and inventory routing problem for a liquid oxygen supply chain comprising production facilities, distribution network, and distribution resources. The key decisions of the problem involve production levels of production plants, delivery schedule and routing through heterogeneous vehicles, and inventory strategies for national stock-out prevention. Due to the problem complexity, we propose a two-level hybrid solution approach that solves the problem using both exact and metaheuristic methods. At the upper level, we develop a mixed-integer linear programming (MILP) model that determines production and inventory decisions and customer allocation. In the lower level, the original problem is reduced to several multi-trip heterogeneous vehicle routing problems by fixing the optimal production, inventory, and allocation decisions and clustering customers. A well-recognised metaheuristic, guided local search method, is adapted to solve the low-level routing problems. A real-world case study in the UK illustrates the applicability and effectiveness of the proposed optimisation framework.

1. Introduction

COVID-19 is a disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Lai et al., 2020). Some COVID-19 patients admit to the hospital where they are very likely to need breathing support due to severe respiratory symptoms (Higgins et al., 2021). Oxygen is a critical component in treating such patients. Thus, the COVID-19 pandemic is compelling oxygen suppliers to seek opportunities to improve operational efficiency and reduce costs of supply chains while maintaining a high quality of service. One widely acknowledged method to achieve this goal is the integrated supply chain optimisation, which simultaneously considers multiple activities in the supply chains such as production, inventory, and distribution (Fahimnia et al., 2013).

In a conventional approach, each element of the supply chain is often treated individually and optimised sequentially. For instance, production and inventory decisions are made in order to minimise the production cost; then, such decisions provide inputs for the distribution planning problem. Solving the supply chain planning problems in this way can avoid computational difficulties (Thomas and Griffin, 1996); however, it may lead to suboptimal or even infeasible solutions. Then again, as presented in some recent studies, an integrated approach to supply chain planning problems can result in a substantial amount of cost savings (Brown et al., 2001; Çetinkaya et al., 2009). The production, inventory, routing, and distribution problem (PIRP) is a type of integrated problem that involves simultaneously determining production, inventory, routing, and distribution decisions. In the PIRP, the decision-maker must decide on the

* Corresponding author.
E-mail address: l.papageorgiou@ucl.ac.uk (L.G. Papageorgiou).
https://doi.org/10.1016/j.cherd.2022.07.027

production level at each plant and each time period. The production decisions are made based on the production costs associated with the produced amount of product and the production capacity. The product deliveries from the plants to customers are being fulfilled by a set of capacitated vehicles considering the unit transportation costs. Regarding inventory at customer sites, the PIRP considers a vendor-managed inventory (VMI) paradigm. In the VMI policy, the product orders from customers are not given as they are in the traditional approach. In the VMI system, the supplier monitors the customer inventory levels and plans when and how much of the product to deliver to ensure that the inventory levels never fall under security levels (Waller et al., 1999). The PIRP also involves vehicle routing decisions in which the order of customer visits is determined. The first work on the PIRP dates back to the research conducted by Thomas and Griffin (1996). They considered the production at a single plant and distribution of multiple products via a fleet of vehicles. Two approaches were used, one in which the production planning and routing problems were decoupled and solved independently, and the other in which both problems were solved simultaneously. Solutions from these approaches were compared, and the coordinated approach showed a cost reduction ranging from 3 % to 20 %. A similar study was conducted by Fumero and Vercellis (1999), and an average improvement of 10 % was seen with the use of the integrated approach.

The PIRP has incurred high combinatorial complexities from routing decisions; therefore, many studies have focused on developing solution methods. Lei et al. (2006) developed a two-phase solution approach for the PIRP involving several plants and a fleet of vehicles with different capacities. In the first phase, the problem was solved without considering the consolidation of less than the transporter loads (LTL). The resulting inefficiencies in routing decisions were improved in the second phase. An extended optimal partitioning procedure was also proposed for an efficient routing heuristic. Boudia et al. (2006), (2007) developed a model for the PIRP consisting of a single facility, a set of identical vehicles, and customers. To handle the computational difficulty of the model, they proposed both a memetic algorithm with population management (MAPA) and a reactive greedy randomised adaptive search procedure (GRASP) improved by the path-relinking procedure. The MAPA was further improved by Boudia and Prins (2009), and the efficiency of these algorithms was addressed by randomly generated instances including 50, 100, and 200 customers. Bard and Nananukul (2009a), (2010) first investigated models previously developed and then introduced a heuristic based on a branch-and-price algorithm to solve a single-plant a homogeneous fleet PIRP. The algorithm iteratively solves the sub-pricing problems using a tabu search to generate feasible routing. Archetti et al. (2011) proposed a solution method for solving a problem with a facility and a set of customers based on a hybrid heuristic. The method first determines the delivery quantities and routes by assuming an unlimited production capacity at the facility. Then, the problem that considers only the production part is solved. Finally, the solution obtained from the previous steps is improved by iteratively removing and inserting two customers. Absi et al. (2015) worked on a single-item production routing problem. They introduced a two-phase iterative approach in which production and routing problems are solved in a sequential fashion. In phase I, a solution to a lot-sizing problem is established by simplifying the routing part. The routing decisions are determined in phase 2; then, the routing solution is used to update the corresponding costs for the next iteration. Cóccola et al. (2020) developed a two-stage solution procedure to solve the PIRP considering production facilities, multiple products, customers, and homogeneous multi-compartment trucks. In the first stage, multi-period routes are generated using a column generation method, and then production schedules, inventory profiles, and distribution routes are determined in the second stage.

When a problem involves vehicle routing decisions but not production decisions, it is referred to as an inventory routing problem (IRP), which is a relaxation of the PIRP. A review of the relevant literature can be found in Campbell et al. (1998); Bertazzi et al. (2008) and a more recent paper by Coelho et al. (2014). Concerning the solution method to the IRP, Campbell and Savelsbergh (2004) presented a two-phase solution approach that decouples a set of inventory and routing decisions. The first phase solves integer programming by clustering customers to reduce the number of routes, whereas the second phase focuses on finding routing schedules heuristically. In Savelsbergh and Song (2008), the authors studied an IRP with continuous moves where the repeated distribution of a product to a set of customers is considered. Etebari and Dabiri (2016) proposed a quadratic mixed-integer programming model for a multi-period IRP under the dynamic regional pricing problem. They developed a hybrid heuristic approach comprising five phases and being embedded in a simulated annealing framework. The first initialisation phase determines the time periods that customer inventories are potentially replenished; then, customer demands are generated and adjusted in the second and third phases. The fourth phase solves the reduced IRP by previous phases, and the solutions of the IRP are improved using the simulated annealing algorithm in the final phase. In addition, Dong et al. (2017) presented a solution technique that uses a preprocessing algorithm and a decomposition method for upper and lower-level subproblems. In the upper level, vehicle routing and scheduling subproblems. In the vehicle routing subproblem, routes are generated to fulfil the demands of customers selected by the preprocessing algorithm. On the other hand, detailed schedules for vehicles and drivers are determined in the lower level. More recently, Karakostas et al. (2019) studied an IRP in which location decisions, inventory profiles, vehicle routing, and distribution outsourcing decisions are integrated. They proposed a solution approach that uses a general variable neighbourhood search (GVNS) scheme to address the computational complexity of the problem. The proposed solution method consists of initialisation and improvement phases. In the initialisation phase, location and allocation decisions are constructed in addition to the inventory-routing decisions; then, the decisions are improved in the second phase. The efficiency of the proposed approach is illustrated with different-sized instances. This work was extended in Karakostas et al. (2020) by considering the environmental impact on the supply chain activities.

Despite the growing emphasis dedicated to the PIRP, such an integrated problem for the chemical process industry has received relatively little attention. An extensive survey of the literature regarding the integrated problem for the process industry was provided by Barbosa-Povoa and Pinto (2020); Ramaswamy et al. (2020). Glankwamdee et al. (2008) developed a linear programming (LP) model for coordinated optimisation of the production-distribution of industrial gases. In the model, routing decisions are disregarded; instead, they are approximated via constraints on the resources of trucks and drivers. You et al. (2011) addressed an industrial gas supply chain problem that considers distribution and inventory planning, except for production planning. An MILP model was formulated by combining short-term vehicle routing decisions with long-term inventory and customer-tank sizing decisions. Two computational strategies were also proposed. The first strategy relies on the decomposition of the full-space MILP model. By contrast, the second one is a continuous approximation that can predict the distribution cost without taking rigorous routing decisions into account. Marchetti et al. (2014) proposed a multi-period MILP model. The model was formulated by considering the production at multiple plants and distribution to multiple depots. A heuristic was applied to construct distribution routes; then, the routes were used in the MILP model. Zamarripa et al. (2016) extended the work by Marchetti et al. (2014). They introduced a rolling horizon decomposition method with two aggregation strategies to improve computational efficiency. In the first aggregation strategy, the discrete variables of the model were relaxed, whereas the distribution part was simplified in the second strategy. Different sizes of case studies examined the computational advantage of the developed method. However, when solving large-size industrial cases, both Marchetti et al. (2014) and Zamarripa et al. (2016) simplified the distribution model in which a detailed routing schedule could not be provided. In the paper presented by Zhang et al. (2017), a multigrid MILP model was proposed for a multiscale production routing problem involving plants, customers, multi-products, and homogeneous vehicles. In addition, an iterative heuristic-based method that dynamically updates the set of routes was introduced. Misra et al. (2018) studied a short-term planning problem for the cryogenic air separation industry. They developed an MILP model that can simultaneously establish the production and distribution plans. The model includes a simplistic distribution model that cannot provide detailed truck schedules but sufficiently rigorous solutions at the planning level. Furthermore, Lee et al. (2021) proposed an optimisation framework for the integrated decisions in industrial gas supply chains. They formulated the problem as an MILP. The model considered multiple features, including business contracts for the raw material and product, production, inventory control, and delivery through multimodal transportation. Rigorous truck scheduling was considered in the model; however, decisions on the tuck routing were disregarded. They also presented an efficient hierarchy-based solution approach to address the computational complexity of the model.

This work aims to develop an optimisation framework incorporating production, inventory, and routing decisions in oxygen supply chains during the COVID-19 pandemic. We propose a hybrid solution method that consists of two levels to efficiently solve the integrated problem. First, we develop an upper-level MILP model that solves the problem by disregarding the consideration of multiple customers on each route. The distribution part of the model considers complicating features, transportation by heterogeneous vehicles, multiple trips and multiple visits, restriction on the customer visit according to the vehicle type, and both fixed and quantity-plant- or quantity-vehicle-dependent loading/unloading times. At the lower level, we employ a metaheuristic approach to solve the detailed routing problem with the same complicating features. To the best of our knowledge, the production and inventory routing problem with this level of detail has not been investigated, especially for industrial gas applications. Furthermore, solution methods using metaheuristics have rarely been reported to solve this problem.

The main contributions of this work are the following:

1. Propose a mathematical framework that integrates production, inventory control, and distribution decisions. The framework considers a high level of detail on the distribution side, including a heterogeneous fleet of vehicles, customer-vehicle suitability, and fixed and quantity-plant- or quantity-vehicle-dependent loading/unloading times.
2. Develop a hybrid solution method that combines the exact method using mathematical programming with a metaheuristic approach for solving industrial-size problems.
3. Conduct a computational study on a real-world supply chain planning problem during the COVID-19 pandemic in the UK to test the applicability and capability of the proposed model and solution strategy.

The remainder of this paper is organised as follows. In Section 2, the problem statement is described. Then, Section 3 presents solution approach as well as the MILP formulation for this problem. In Section 4, the computational results of a case study in the UK and discussion are presented. Finally, the concluding remarks are provided in Section 5.

2. Problem statement

In this work, we consider a liquid oxygen supply chain network consisting of three echelons, production plants ($i \in I$), third-party suppliers ($m \in M$), and a number of customers ($k \in K$), as illustrated in Fig. 1.

The overall planning horizon of the problem is divided into multiple time periods ($t \in T$). Each plant can produce the product in a time period (P_{it}) to its capacity (P_i^{max}) and store it in the storage tank. The third-party suppliers can also be sources of the product.

Heterogeneous vehicles ($v \in V$) are used to transport the product from the plants to the customers. Each vehicle is designated to a specific plant or third-party supplier over the planning horizon (i.e., V_i or V_m) and defined with its capacity (CAP_v), fixed loading/unloading times (α^{LT}/α^{UT}), variable loading/unloading times (γ_i or γ_m/β_v), and variable cost per unit of travel distance (C_v^{veh}). The variable loading time (γ_i or γ_m) is dependent on both the loading quantity and loading location (i.e., plant i or third party m), while the variable unloading time (β_v) is dependent on the product quantity and vehicle v. The vehicles can perform multiple routes during a time period. We assume that each vehicle that leaves the plant or third-party supplier must return to the origin on completion of each route. In this problem, the routes are categorised into single- and multi-customer routes. Each vehicle visits one customer at most per trip in the single-customer route, while one or more customers can be visited in the multi-customer route. Further, there is a restriction on the customer visits according to the vehicle type, and this customer-vehicle suitability (K_v) is pre-specified. On the demand side, we consider two types of customers, industrial customers and hospitals. It should be noted that there is no particular distinction between industrial customers and

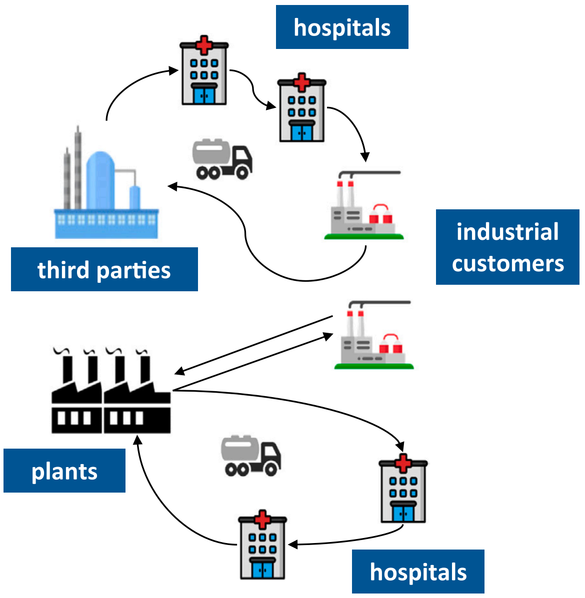

Fig. 1 – Liquid oxygen supply chain network.

hospitals. However, such a categorisation is common practice since the hospital demands might be more critical during the COVID-19 pandemic. The product being transported is stored in tanks at the industrial customer and hospital sites, and their inventories (I_{kt}) are controlled by the VMI methodology. Therefore, under the VMI, companies need to determine the delivery quantities and times of the product to satisfy customer demands (D_{kt}).

Overall, the considered optimisation problem is described as follows:

Given are:

- Plants: locations, maximum production capacity (P_i^{max}).
- Third-party suppliers: locations, limitations on product supplies (UB_m).
- Customers: locations and demand profiles (D_{kt}).
- Heterogeneous vehicles: vehicle-customer suitability (K_v), fixed loading/unloading times (α^{LT}, α^{UT}), variable loading/unloading times (γ_i, γ_m, β_v), loading capacities (CAP_v), available quantities (V_i, V_m).
- Inventories at plant and customer sites: initial (I_i^{ini}, I_k^{ini}), minimum (I_i^{min}, I_k^{min}), and maximum (I_i^{max}, I_k^{max}) levels.
- Cost data (i.e., production (C^{elec}, λ_i), third-party (C_m^{third}), unit transportation (C_v^{veh}) costs)

To determine:

- Liquid oxygen production levels.
- Inventory levels at customers and plants.
- Demand allocation to plants and third parties.
- Optimal routes and delivery schedules.

So as to:

Minimise total operating cost, including production, third-party, and transportation costs.

3. Solution strategy

The problem described in the previous section includes the decision on the vehicle routes, which is regarded as NP-hard. Given the NP-hardness, solving the full PIRP with exact methods might require expensive computational costs, especially when large-size problems are considered (Bard and Nananukul, 2009b). For instance, Bard and Nananukul (2010) attempted to solve small PIRP instances involving 30 customers and three vehicles for up to eight time periods in a monolithic way, and the reported average optimality gap after 1 h was 34 %. In addition, the instances that considered more than 40 customers and four vehicles were not solvable within 1 h. The computational difficulty of the PIRP arises from the vehicle routing decisions when accounting for the multi-customer routes. For example, the problem, including eight plants, 700 + customers, 30 + vehicles, and 30 time periods, involves more than 1 billion discrete variables. To address the computational challenge, we introduce a two-level solution strategy.

The upper level of the solution strategy focuses on finding optimal production and inventory decisions and customer allocation to plants and third-party suppliers. At this level, the decision on the multi-customer vehicle routing is disregarded. Instead, only single-customer routing is considered. By accounting for only the single-customer routes, not only can the computational burden arising from the combinatorial nature of the routing decisions be reduced, but efficient production and inventory plans can still be established. Section 3.1 describes the upper-level problem re-defined by the solution method. The lower level takes the multi-customer routing decisions into account. The upper-level production and inventory decisions are fixed, and the customers are clustered based on the optimal allocation obtained from the upper level. After this procedure, the original PIRP is transformed into several multi-trip heterogeneous vehicle routing problems (MTHVRPs), which can be defined for each cluster. The problem description for each MTHVRP is given in Section 3.2 then mathematical models are proposed subsequently.

3.1. *Upper-level problem*

The upper-level problem constitutes the same description as the original PIRP presented in Section 2. However, the following additional assumption needs to be introduced.

- On each route, only one customer can be visited at most.

3.2. *Lower-level problem*

At the lower level, the original PIRP is transformed into several MTHVRPs, and each MTHVRP is described as follows:

Given are:

- Location of a plant or third-party supplier.
- Location of customers.
- Customer demands need to be delivered.
- Vehicle capacities and available quantity at a plant or third-party supplier.
- Limitation on the vehicle operation times.
- Vehicle restriction on the customer visits.
- Unit transportation cost per distance.

To determine:

- Vehicle allocation to each customer.
- Optimal routes.

So as to:
Minimise total vehicle routing cost.

3.3. *Upper-level model*

First, we formulate the mathematical model used in the upper level as an MILP. The upper-level model is denoted as PIRP-L throughout the paper, and the notation used in the model is as follows:

Notation	
Indices	
i	production plant
k	customer
m	third-party supplier
t	time period
v	vehicle
Sets	
I	set of plants
K	set of customers
M	set of third-party suppliers
K_i	set of customers initially assigned to plant i
K_m	set of customers initially assigned to third party m
K_v	set of customers that vehicle v can visit
K_{iv}	set of customers initially assigned to plant i and can be served by vehicle v, $K_{iv} = \{k \mid k \in K_i \cap K_v\}$
K_{mv}	set of customers initially assigned to third party m and can be served by vehicle v, $K_{mv} = \{k \mid k \in K_m \cap K_v\}$
T	set of time periods
V_i	set of vehicles at plant i
V_m	set of vehicles at third party m
Parameters	
C^{elec}	electricity price (£/kWh)
C_m^{third}	unit purchasing cost of third party m (£/kg)
C_v^{veh}	unit transportation cost of vehicle v (£/mile)
CAP_v	capacity of vehicle v (kg)
D_{kt}	product demand of customer k in time period t (kg/day)
I_i^{ini}	initial inventory level at plant i (kg)
I_i^{max}	maximum inventory level at plant i (kg)
I_i^{min}	minimum inventory level at plant i (kg)
I_k^{ini}	initial inventory level at customer k (kg)
I_k^{max}	maximum inventory level at customer k (kg)
I_k^{min}	minimum inventory level at customer k (kg)
L_{ik}	distance of a round trip between plant i and customer k (mile)
L_{mk}	distance of a round trip between third party m and customer k (mile)
P_i^{max}	production capacity of plant i in each time period (kg)
UB_m	maximum product amount that can be sourced from third-party supplier m in each time period (kg)
α^{LT}	fixed loading time (day)
α^{UT}	fixed unloading time (day)
γ_i	variable loading time at plant i (day/kg)
γ_m	variable loading time at third party m (day/kg)
β_v	variable unloading time of vehicle v (day/kg)
Δ_t	length of time period t (day)
θ_{ik}	duration of a round trip between plant i and customer k (day)
θ_{mk}	duration of a round trip between third party m and customer k (day)
λ_i	unit specific power for plant i (kWh/kg)
Integer variables	
A_{ikvt}	number of deliveries from plant i to customer k executed by vehicle v during time period t
A_{mkvt}	number of deliveries from third-party supplier m to customer k executed by vehicle v during time period t
Continuous variables	
I_{it}	inventory level of plant i at the end of time period t (kg)
I_{kt}	inventory level of customer k at the end of time period t (kg)
P_{it}	production rate of plant i in time period t (kg/day)
Q_{ikvt}	amount of product transported from plant i to customer k by vehicle v during time period t (kg)
Q_{mkvt}	amount of product transported from third-party supplier m to customer k by vehicle v during time period t (kg)
TOC	total operating cost (£)
TP	production cost (£)
TS	third-party supplier cost (£)
TT	transportation cost (£)

3.3.1. *Objective function*

The objective of this problem is to minimise the total cost of the liquid oxygen supply chain, and the total cost comprises production, third party, and transportation costs.

$$\text{minimise } TOC \tag{1}$$

$$TOC = TP + TS + TT \tag{2}$$

where TOC denotes the total cost; TP is the production cost; TS is the third-party cost; and TT refers to the transportation cost of vehicles.

3.3.2. *Production and third party costs*

The production cost is calculated based on the electricity price (C^{elec}), unit specific power (λ_i), the production rate (P_{it}), and duration of each time period (Δ_t) as follows:

$$TP = \sum_{i \in I} \sum_{t \in T} C^{elec}\, \lambda_i\, P_{it}\, \Delta_t \tag{3}$$

The third-party cost is defined based on the unit purchasing cost from the third party (C_m^{third}) and the quantity of product transported from third party m to customer k.

$$TS = \sum_{m \in M} \sum_{k \in K_{mv}} \sum_{v \in V_m} \sum_{t \in T} C_m^{third}\, Q_{mkvt} \tag{4}$$

3.3.3. Transportation cost

$$TT = \sum_{i \in I} \sum_{k \in K_{iv}} \sum_{v \in V_i} \sum_{t \in T} C_v^{veh} L_{ik} A_{ikvt} + \sum_{m \in M} \sum_{k \in K_{mv}} \sum_{v \in V_m} \sum_{t \in T} C_v^{veh} L_{mk} A_{mkvt} \quad (5)$$

Eq. 5 calculates the total transportation cost of vehicles that leaves the plants or third-party suppliers. The transportation cost is defined based on the distance-dependent unit transfer cost (C_v^{veh}), the distance to make a round trip from plant i or third party m to customer k (L_{ik} and L_{mk}), and the number of trips each vehicle performs during time period t (A_{ikvt} and A_{mkvt}). where K_{iv} and K_{mv} are the sets of customers initially assigned to plant i and third party m based on their geographical locations and can be served by vehicle v, and V_i and V_m denote the sets of vehicles allocated to plant i and third-party supplier m, respectively.

3.3.4. Production constraints

The production capacity of each plant must be considered. Eq. 6 establishes the upper bound (P_i^{max}) for the production amount at plant i during time period t (P_{it}).

$$P_{it}\ \Delta_t \le P_i^{max} \quad \forall\ i \in I,\ t \in T \quad (6)$$

In addition, the product amount that can be purchased and delivered from third-party supplier m to any customers in time period t is limited by the upper bound (UB_m).

$$\sum_{k \in K_{mv}} \sum_{v \in V_m} Q_{mkvt} \le UB_m \quad \forall\ m \in M,\ t \in T \quad (7)$$

3.3.5. Vehicle constraints

The product amount delivered by vehicle v from plant i or third-party supplier m is limited by the vehicle capacity and the number of trips.

$$Q_{ikvt} \le CAP_v \quad A_{ikvt} \quad \forall\ i \in I,\ v \in V_i,\ k \in K_{iv},\ t \in T \quad (8)$$

$$Q_{mkvt} \le CAP_v \quad A_{mkvt} \quad \forall\ m \in M,\ v \in V_m,\ k \in K_{mv},\ t \in T \quad (9)$$

Eqs. 8-9 pose the delivery amount and it must be less than or equal to the number of deliveries (A_{ikvt} and A_{mkvt}) multiplied by the capacity of the vehicle (CAP_v).

$$\sum_{k \in K_{iv}} (\theta_{ik} + \alpha^{LT} + \alpha^{UT}) A_{ikvt} + \sum_{k \in K_{iv}} (\beta_v Q_{ikvt} + \gamma_i Q_{ikvt}) \le \Delta_t \quad \forall\ i \in I,\ v \in V_i,\ t \in T \quad (10)$$

$$\sum_{k \in K_{mv}} (\theta_{mk} + \alpha^{LT} + \alpha^{UT}) A_{mkvt} + \sum_{k \in K_{mv}} (\beta_v Q_{mkvt} + \gamma_m Q_{mkvt}) \le \Delta_t \quad \forall\ m \in M,\ v \in V_m,\ t \in T \quad (11)$$

Eqs. 10-11 pose the operation time that each vehicle can perform. The operation time of each vehicle includes the travelling time between locations (θ_{ik} and θ_{mk}), fixed loading time at the plant (α^{LT}), fixed unloading time at the customer (α^{UT}), and quantity-dependent variable loading and unloading times. Here, the quantity-dependent loading time (γ_i) is dependent on the plant, whereas the unit unloading time (β_v) depends on the vehicle. The total operation time of vehicle v during time period t cannot exceed the length of each time period (Δ_t).

3.3.6. Inventory constraints

The inventory level of each plant at the end of the time period is defined as the inventory level in the previous time period plus the production amount during the time period minus the amount delivered to customers.

$$I_{it} = I_{i,t-1}|_{t>1} + I_i^{ini}|_{t=1} + P_{it}\ \Delta_t - \sum_{k \in K_{iv}} \sum_{v \in V_i} Q_{ikvt} \quad \forall\ i \in I,\ t \in T \quad (12)$$

Similarly, the inventory level of each customer at the end of the time period is the inventory level in the previous time period plus the incoming product amount from any plants/third-party suppliers minus its demand.

$$I_{kt} = I_{k,t-1}|_{t>1} + I_k^{ini}|_{t=1} + \sum_{i:k \in K_{iv}} \sum_{v \in V_i} Q_{ikvt} + \sum_{m:k \in K_{mv}} \sum_{v \in V_m} Q_{mkvt} - D_{kt}\Delta_t\ \forall\ k \in K,\ t \in T \quad (13)$$

The inventory level in each time period must not fall below the minimum level and can not exceed the maximum capacity.

$$I_i^{min} \le I_{it} \le I_i^{max} \quad \forall\ i \in I,\ t \in T \quad (14)$$

$$I_k^{min} \le I_{kt} \le I_k^{max} \quad \forall\ k \in K,\ t \in T \quad (15)$$

Additionally, it is assumed that the inventory level at the end of the planning horizon is equal to the initial inventory level.

$$I_{i,t} = I_i^{ini} \quad \forall\ i \in I,\ t = |T| \quad (16)$$

$$I_{k,t} = I_k^{ini} \quad k \in K,\ t = |T| \quad (17)$$

3.4. Lower-level model

As aforementioned, at the lower level, the original PIRP is reduced to the MTHVRP after fixing the variables for production and inventory levels (P_{it}, I_{it}, and I_{kt}). In addition, the customers can be clustered to each plant and each time period based on the optimal customer allocation gained from the upper level. Each cluster can be defined based on the optimal values of the integer variables (denoted as A^*_{ikvt} and A^*_{mkvt}). For instance, if customer k received the product from plant i during time period t by any vehicle at the plant (i.e., $\sum_{v:k \in K_{iv}} A^*_{ikvt} \ge 1$), then the customer is clustered to the plant and time period. By clustering the customers, the MTHVRP can be solved for each plant and time period.

In the MTHVRP for each cluster, we set an additional assumption that each customer must be visited exactly once. Therefore, to allow only a single visit per customer, we fix the upper-level single-customer routing solution for customers who need multiple visits during a time period and for which the delivered amounts to the customers are the same as the vehicle capacities before solving the lower-level MTHVRPs. The upper-level single-customer routes, which are not fixed during this step, are part of the lower-level optimisation. Consequently, the upper-level solution guarantees a feasible solution to the lower-level problem. In addition, the worst solution (i.e., upper bound) that the lower-level problem can get is the single-customer routing solution obtained from the upper level. The detailed procedure and example for fixing the single-customer routes and calculating the values of the input parameters used for the MTHVRP are presented in Appendix A.

The mathematical framework for the MTHVRP is presented in the following section. Note that the indices for plant i, third-party suppliers m, and time period t are omitted from the model as the MTHVRP is solved for each plant/third party and each time period.

Notation	
Indices	
n, n'	echelon node (i, k, m), where 0 represents the plant or third-party node
r	route
Sets	
N^c	set of customer nodes
N_v	set of nodes that can be visited by vehicle v
N_v^c	set of customer nodes that can be visited by vehicle v
R	set of routes
Parameters	
γ	loading time at the plant or third-party node
$\theta_{nn'}$	travelling time from node n to node n' (day)
τ_v	operation time limit for vehicle v in each time period (day)
$L_{nn'}$	distance between node n and n' (mile)
D_n^{new}	product amount must be delivered to customer node n (kg)
TT^{fix}	transportation cost of fixed single-customer routes (£)
Binary variables	
E_{vrn}	1 if vehicle v visits node n on route r; 0 otherwise
$X_{vrnn'}$	1 if vehicle v travels from node n to node n' on route r; 0 otherwise
Continuous variables	
AQ_{vrn}	accumulation of delivered product amount before visiting node n on route r (kg)

$$\text{minimise } TT = TT^{fix} + \sum_{v\in V}\sum_{r\in R}\sum_{n\in N_v}\sum_{n'\in N_v} C_v^{veh} L_{nn'} X_{vrnn'} \quad (18)$$

$$\sum_{v:n\in N_v^c}\sum_{r\in R} E_{vrn} = 1 \quad \forall\, n \in K^c \quad (19)$$

$$\sum_{n'\in K_v} X_{vr0n'} \le 1 \quad \forall\, v \in V, r \in R \quad (20)$$

$$\sum_{n\in K_v} X_{vrn0} \le 1 \quad \forall\, v \in V, r \in R \quad (21)$$

$$\sum_{n'\in N_v} X_{vrnn'} = \sum_{n\in N_v} X_{vrn'n} = E_{vrn} \quad \forall\, v \in V, r \in R, n \in K_v^c \quad (22)$$

$$\sum_{r\in R}\sum_{n\in N_v}\sum_{n'\in N_v} \theta_{nn'} X_{vrnn'} + \sum_{r\in R}\sum_{n\in K_v^c} (\alpha^{UT} + \beta_v D_n^{new}) E_{vrn} + \sum_{r\in R}\sum_{n\in K_v^c} (\alpha^{LT} X_{vr0n} + \gamma D_n^{new} E_{vrn}) \le \tau_v \quad \forall\, v \in V \quad (23)$$

$$AQ_{vrn} + D_n^{new} \le AQ_{vrn'} + CAP_v(1 - X_{vrnn'}) \quad \forall\, v \in V, r \in R, n \in K_v^c, n' \in N_v \quad (24)$$

$$AQ_{vrn} \le CAP_v \quad \forall\, v \in V, r \in R, n \in N_v \quad (25)$$

Eq.18 is the objective function of the lower-level problem, and it is the single-customer routing cost fixed from the upper-level plus the routing cost. Eq.19 poses the number of visits for each customer node. Each customer node must be visited exactly once. Eqs.20-21 ensure at most one departure from a plant/third party and one arrival to a plant/third party. Eq.22 ensures the trip integrity in which there is exactly one precedent and following nodes if customer node n is visited by vehicle v on route r. Similar to Eqs.10-11, Eq.23 sets the limitation of the operation time for each vehicle. Eqs. 24-25 are the vehicle capacity constraints. where the AQ_{vrn} denotes the cumulative product amount delivered by vehicle v on route r before serving customer node n. These capacity constraints also eliminate the subtours.

Because the vehicle routing problem with heterogeneous vehicles is an NP-hard problem, solving such a problem with the MILP model is still a non-trivial task. Thus, we solve the problem by adapting the guided local search (GLS) method using the open-source tool, Google-OR tools. The GLS is a penalty-based metaheuristic algorithm that sits on top of other local search algorithms (Voudouris et al., 2010). The corresponding details are omitted for brevity, and the reader is referred to Voudouris and Tsang (1999); Beullens et al. (2003); Voudouris et al. (2010) for the details.

Finally, the overall structure of the two-level solution strategy is illustrated in Fig. 2.

4. Case study

4.1. Input data

To illustrate the applicability of the proposed model and approach, we solve a real-world case of oxygen supply chain planning during the COVID-19 pandemic in the UK. The case was provided by Linde/BOC, the largest oxygen provider in the UK. In the following, we present a brief introduction to the input data and assumptions. Note that detailed information regarding actual locations of production plants, third-party suppliers and customers, plant capabilities, and customer demands cannot be disclosed for confidentiality reasons. Therefore, in the remainder of this paper, costs are given in the monetary unit (MU), and numerical values are normalised, if necessary. In the UK oxygen supply chain network, there are eight production plants and third-party suppliers, over 700 customers, including hospitals and industrial customers, and dozens of heterogeneous vehicles. The vehicles are categorised into five types which take different unit transportation costs. The planning horizon is one month and the time discretisation is one day. In addition, we account for the oxygen demand during the first wave of the pandemic in the UK (i.e., April 2020). With respect to the hospital demands during the pandemic, it is mutually related to the number of COVID-19 inpatients. Thus, we calculate their daily demands based on historical data regarding delivery quantities and times provided by the industrial companies and the hospitalisation cases reported by the NHS England, Public Health Scotland, and Wales health boards. Graphs (A), (B), and (C) in Fig. 3 highlight the oxygen demand of hospitals in Scotland, Wales, and England, respectively. Each bar represents the cumulative oxygen demand of hospitals at each time period. The oxygen demand of each hospital is calculated based on the national demand profile so that the dynamic behaviour of the demand during the COVID-19 pandemic can be captured. The detailed method for calculating the hospital daily demands is presented in Appendix B. We use the estimated daily oxygen demand obtained from the method as the input parameters in the case study.

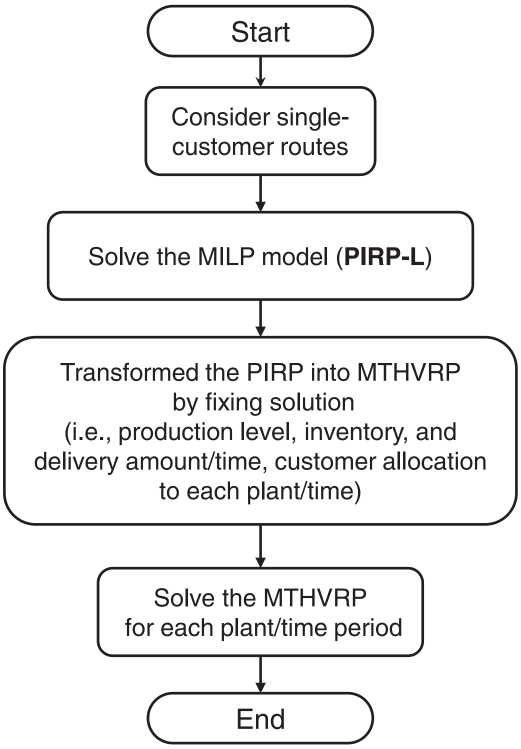

Fig. 2 – Structure of the two-level solution strategy.

4.2. *Results and discussion*

First, we solved the problem using the PIRP-L model. The upper-level model (PIRP-L) was implemented in GAMS 34.3.0 on an Intel 3.00 GHz, 128 GB RAM desktop and solved using Gurobi 9.1.1 solver. The optimality and time limit were set to 5 % and 3600 s, respectively. Table 1 reports the model statistics with the optimal solution. The upper-level model includes 177,663 equations, 176,674 continuous variables, and 155,040 discrete variables. The PIRP-L model can solve the integrated supply chain planning problem within a reasonable time of 1470 s. This is because the number of discrete variables is considerably reduced by considering only single-customer routes at this level. Fig. 4 provides the breakdown of the total cost gained from the PIRP-L model. The transportation cost accounts for almost half of the total cost (41 %), whereas the production and third-party costs account for 52 % and 7 %, respectively. Fig. 5 shows the inventory profiles and product flows at hospital $k17$ and industrial customer $k565$, respectively. In the figures, the left y-axis represents the inventory levels, while the right y-axis represents the product flows going in/out from the storage tank of the customer. The negative value of the right y-axis represents the product amount going out from the tank. The blue and green bars represent the amount of the product delivered from the plants to the customer. These bars can also be interpreted as the vehicle's load. For instance, the blue bar on day 2 in Fig. 5 for hospital $k17$ represents that a vehicle delivered the product from plant $i1$ to the hospital, and the vehicle load, i.e., delivered amount, is 6550 kg. In addition, the red bar shows the remaining load capacity in the vehicle, which is the vehicle capacity minus the load on the vehicle. As depicted, the frequency of the product deliveries for hospital $k17$ is relatively high, although its tank capacity is larger than that of industrial customer $k565$. This is because of the high security level of the hospital's tank. The security level, which is more than 75 % of the maximum capacity, limits the maximum amount of product to be delivered by vehicles. In addition, the delivered product amounts to hospital $k17$ are relatively small compared to the vehicle capacity. For example, a vehicle with the capacity of 13,004 kg delivered the product of 3670 kg from plant $i1$ to hospital $k17$ on day 12. The result indicates the necessity of re-optimising the single-customer routes obtained from the upper-level model.

After solving the upper-level problem, the customers were clustered for each plant/third-party supplier and each time period. Furthermore, the original PIRP was transformed into an independent MTHVRP for each cluster by fixing the production and inventory decisions. Each MTHVRP at the lower level was implemented in Python 3.8.10 and solved using the GLS in Google OR-Tools. The termination criterion of the GLS algorithm for each cluster was a time limit of 300 s Table 2 compares the transportation costs obtained from the upper and lower levels. The table also shows the vehicle capacity utilisation calculated based on the optimal values from the upper and lower levels. The values of the capacity utilisation are the average numbers. The capacity utilisation for each vehicle at each time period was calculated based on the number of trips, its capacity, and total product amount delivered to customers. As shown in Table 2, the transportation cost shows a 38 % improvement in the lower level when considering the multi-customer routes. In addition, the capacity utilisation can also be improved dramatically, and it shows an improvement between 3 % and 46 %. None of the vehicles in type 5 is used in both the upper and lower levels as the vehicles have the most expensive unit transportation cost. The improvement in the vehicle capacity utilisation of more clearly depicted in Fig. 6. In the figure, the bar graph represents the product amount delivered from plant $i1$ to customers and line graphs compare the vehicle capacity utilisation gained from upper and lower levels. As the delivery quantity and time are fixed after solving the upper-level problem, there is no difference in the delivered product amount (blue bars). On the other hand, the capacity utilisation of the vehicles at plant $i1$ is improved considerably at the lower level. Fig. 7 visualises the optimal routes of two clusters solved in the lower level. The figure also includes the information regarding the delivery amount and travelling distance. The single-customer routes are re-optimised, and detailed vehicle routing that considers multi-customer routing is determined. In addition, the vehicles utilise their capacities as much as possible.

To analyse the general performance of the GLS and examine the quality of their optimal solutions, we conducted a sensitivity analysis on the metaheuristic algorithm and termination criteria (i.e., CPU time limit). We applied three additional metaheuristic techniques, which are provided as solvers in the open-source tool (i.e., Google OR-Tools): tabu search (TA), simulated annealing (SA), and greedy descent method (GD). Note that the description of each metaheuristic solver can be found in Perron and Furnon (2022). In addition, we used four different CPU time limits for each cluster, including 30 s, 60 s, 300 s, and 3600 s

For the additional runs, we selected the clusters allocated to two different plants, plant $i1$ and $i4$, based on the solution from the upper-level model in the case study. We denoted the clusters related to plant $i1$ and $i4$ as C1 and C2. In C1, there are eight vehicles, whereas four vehicles are involved in C2. In each cluster for plant $i1$, different numbers of customers ranging from five to 27 are allocated. The average

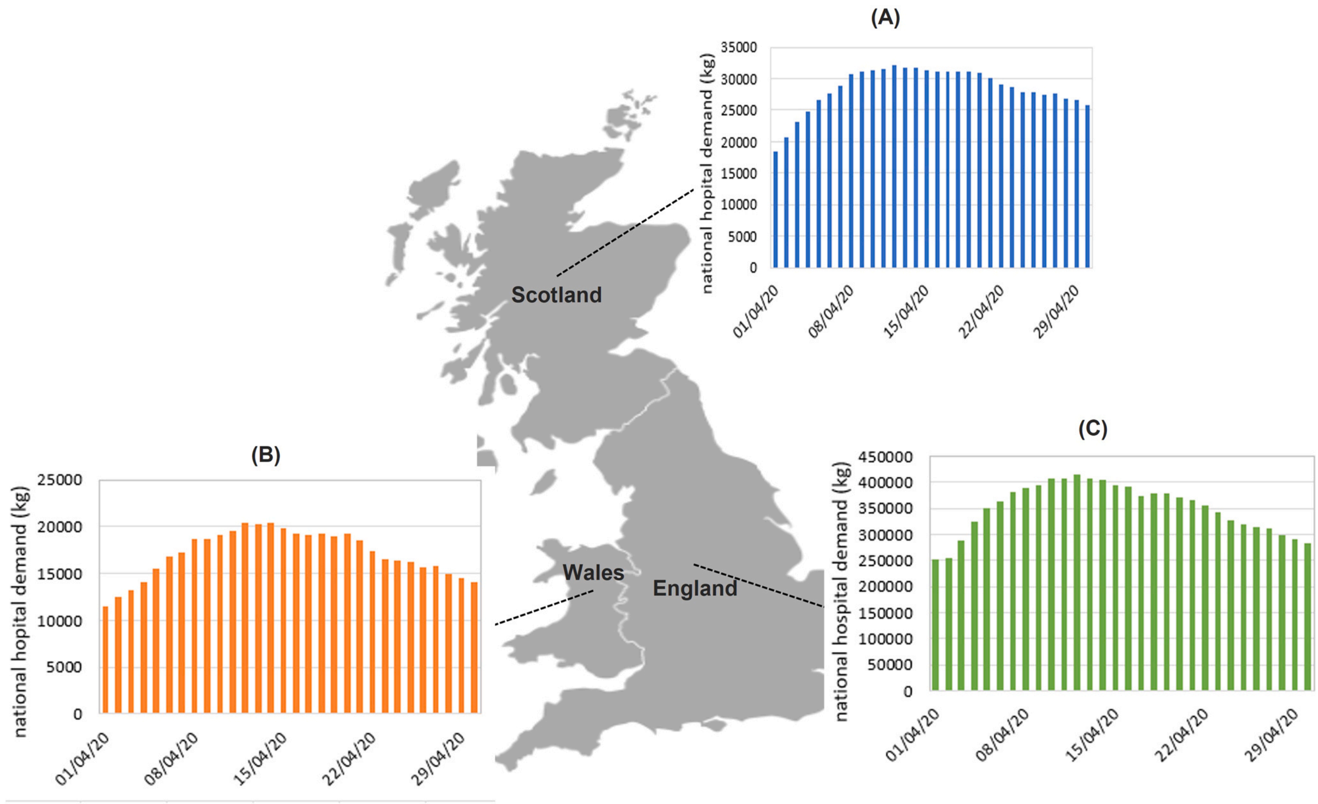

Fig. 3 – National oxygen demand profile of the hospitals (A) in Scotland; (B) in Wales; and (C) in England.

Table 1 – Model statistic of the PIRP-L model (upper level).

model	PIRP-L
no. of equations	177,663
no. of continuous variables	176,674
no. of discrete variables	155,040
optimal solution (MU)	1496,000
optimality gap (%)	3.6
CPU s	1470

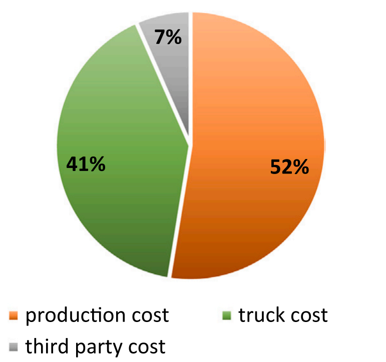

Fig. 4 – Breakdown of total cost obtained form the PIRP-L model (upper level).

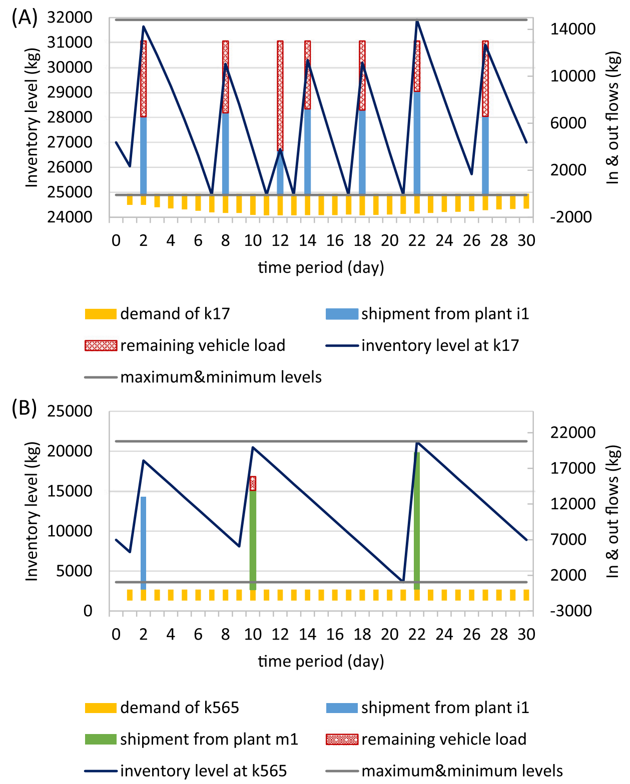

Fig. 5 – Inventory profile at (A) hospital $k17$ and (B) industrial customer $k565$ (upper level).

number of customers in each cluster for plant $i1$ is 21. On the other hand, customers from 11 to 21, on average 18 customers, are allocated to the clusters associated with plant $i4$. Note that we solved 60 independent problems, i.e., MTHVRPs for 60 clusters (2 plants × 30 days), as each cluster is defined by each plant each time period.

The results are presented in Fig. 8. In the figure, the y-axis shows the total transportation cost of the clusters related to each plant, while the x-axis indicates the CPU time limit that was set to solve the MTHVRP for each cluster. First, we compared the performance of the GLS with that of other metaheuristic algorithms. Focusing on the 300 s of CPU time limit that was used for the case study in the previous section, the GLS has a much better performance than other

Table 2 – Comparison of the transportation cost and capacity utilisation.

	vehicle type	average capacity (kg)	unit transportation cost (MU/mile)	upper level	lower level	improvement
transportation cost (MU)				611,198	378,033	38 %
capacity utilisation	type 1	12,800	2.4	35 %	81 %	46 %
	type 2	16,000	2.7	48 %	78 %	30 %
	type 3	20,200	3.0	85 %	88 %	3 %
	type 4	23,700	3.3	65 %	82 %	17 %
	type 5	25,700	3.6	–	–	–

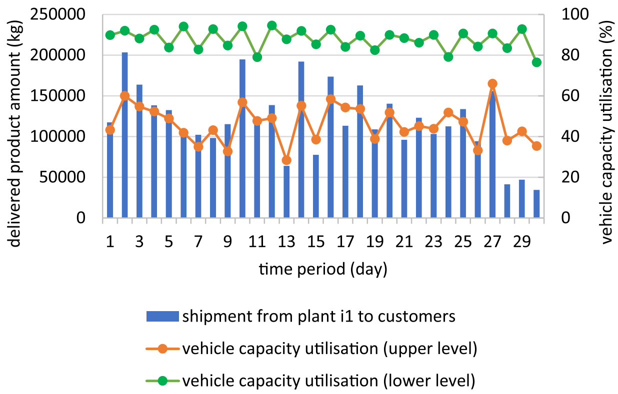

Fig. 6 – Comparison of the capacity utilisation of the vehicles at plant *i*1.

metaheuristics, whereas the performance of the SA and GD are not satisfactory. For C1 and C2, the total transportation costs from the GLS are 6.1 % and 10.8 % less than the costs from the SA and GD, on average. Although the TS has a higher transportation cost in both C1 and C2, the differences between the solutions are within 0.6 %. It is noticeable that the GLS shows a better performance in all cases, except for C1 with 3600 s time limit. The results show that the GLS outperforms other metaheuristic approaches, and it is a good option for solving the MTHVRPs.

Additionally, we investigated the trade-off between computational time and optimal solution. In case of C1, the GLS, TS, and SA improve the solutions by 4.8 %, 4,4 %, and 1.8 %, respectively, when the CPU time limit is increased from 30 s to 300 s. The metaheuristic approaches, except for the GD, with 3600 s of CPU time limit still show a small improvement of 0.9 %, on average, compared with the solutions gained after 300 s. For C2, the SA and GD do not show any noticeable improvement in the solutions with the increased CPU time. This might be due to existence of multiple local minima. Since the GD accepts only cost-reducing neighbours (i.e., better solutions than the current local solution) during a search process, it would be stuck at local minima and unable to improve the solution despite the increased CPU time limit. In case of the SA algorithm, there are a lot of tuneable parameter values, and these values are sensible (Weyland, 2008). These facts often make the SA fail to explore better solutions continuously (Voudouris and Tsang, 1999). By contrast, the optimal solutions gained from the GLS and TS after 300 s show 3.7 % and 3.5 % of improvement compared with the solutions gained in 30 s. In addition, the increased CPU time limit from 300 s to 3600 s results in only 0.3 % and 0.4 % fewer transportation costs for the GLS and TS, respectively. Fig. 9 depicts more detailed comparison of the solution quality across the clusters related to plant *i*1 and *i*4. Each row represents the metaheuristic methods with different CPU time limits and each column represents the time period, i.e.,

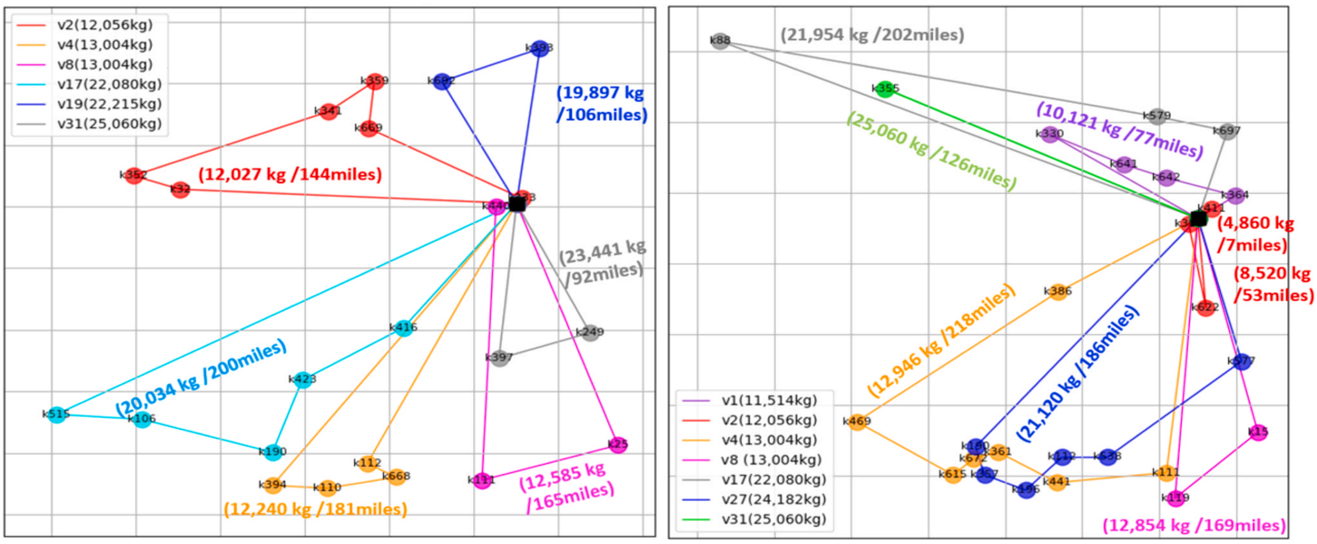

Fig. 7 – Optimal routing decisions gained from the lower level.

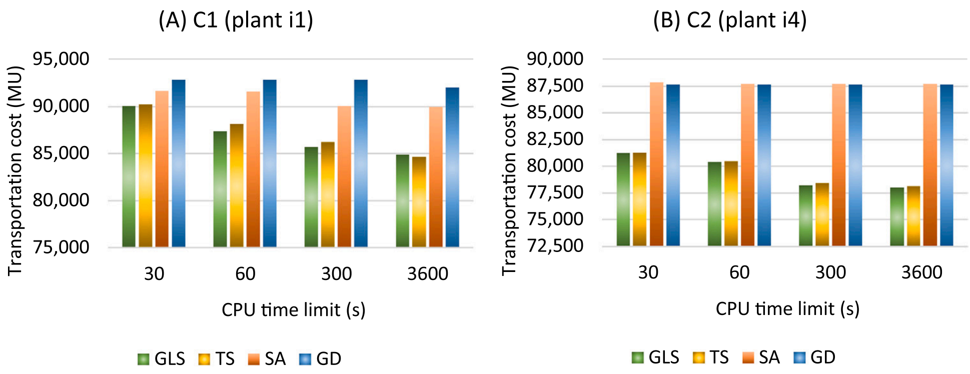

Fig. 8 – Comparison of the results from the sensitivity analysis.

(optimal value / best solution found) x 100 (%) : colour scale 100 – 95 – 90 – 85 – 70

(A) C1 (plant i1)

time period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GLS (30s)	94.8	96.2	94.9	94.1	90.4	88.0	86.9	92.9	92.9	91.1	86.9	81.6	99.6	98.3	98.3	94.4	97.4	88.5	91.5	97.3	96.6	85.7	97.0	97.1	88.1	92.9	96.3	99.5	99.6	100
GLS (60s)	94.8	96.2	99.0	94.1	98.7	97.4	87.3	98.1	95.4	99.2	95.6	81.6	99.6	98.7	98.3	94.4	97.4	95.8	91.8	97.8	99.6	94.1	100	98.3	96.7	92.9	96.3	99.5	99.6	100
GLS (300s)	99.1	98.0	99.0	95.6	99.7	98.0	96.8	97.5	98.6	100	95.6	85.4	100	99.0	100	98.5	97.7	99.3	92.3	100	100	99.6	99.6	98.5	99.4	94.9	97.6	100	100	100
GLS (1h)	100	100	100	98.0	100	100	100	97.5	100	100	97.5	85.4	100	100	100	100	99.5	100	92.3	100	100	100	100	100	99.4	94.9	100	100	100	100
TS (30s)	98.4	91.6	97.7	94.5	90.1	96.2	86.9	96.6	92.5	91.1	88.4	77.7	99.6	98.3	98.3	89.5	97.4	84.3	91.4	98.1	98.7	93.1	96.3	94.5	83.9	91.2	96.3	99.5	99.6	100
TS (60s)	98.4	96.9	97.7	94.5	90.1	96.3	93.0	97.6	93.6	98.2	91.4	77.7	99.6	98.3	99.6	93.1	97.4	97.0	91.7	98.1	98.7	93.3	96.3	94.5	92.1	94.7	96.3	99.5	99.6	100
TS (300s)	98.5	96.9	97.8	99.7	93.0	97.4	96.6	100	95.7	98.9	100	81.3	99.6	98.3	99.6	97.5	99.5	98.9	99.5	98.2	98.8	99.2	99.1	96.9	94.7	94.7	96.3	99.5	99.6	100
TS (1h)	99.1	100	100	100	100	98.4	98.4	100	99.2	99.6	100	84.3	99.6	99.6	99.6	99.3	100	99.7	100	100	99.6	100	100	99.7	99.4	100	98.8	99.5	99.6	100
SA (30s)	93.7	93.5	93.4	94.4	88.0	84.6	86.9	96.6	87.7	94.8	83.7	77.7	99.3	91.0	95.8	74.4	93.6	96.0	93.7	100	95.0	90.5	90.7	99.3	100	90.9	92.8	94.4	99.6	100
SA (60s)	93.7	93.5	95.2	94.4	88.0	84.6	86.9	96.6	87.7	94.8	83.7	77.7	99.3	91.0	95.8	74.4	93.6	96.0	93.7	100	95.0	90.5	90.7	99.3	100	90.9	92.8	94.4	99.6	100
SA (300s)	93.7	94.8	95.2	94.4	95.4	84.6	86.9	96.6	87.7	94.8	95.5	77.7	99.6	99.6	98.3	74.4	95.6	96.0	93.7	100	98.7	90.5	90.7	99.3	100	95.9	92.8	94.4	99.6	100
SA (1h)	93.7	94.8	95.2	98.0	95.4	83.8	86.9	96.6	92.5	94.8	95.5	77.7	99.6	99.6	98.3	74.4	95.6	87.6	87.4	100	98.7	93.7	96.3	94.5	100	95.9	96.4	99.5	99.6	100
GD (30s)	93.6	94.8	89.2	93.7	88.0	84.6	86.9	96.6	92.5	89.9	83.7	77.7	99.6	89.5	98.3	74.4	93.6	87.6	87.4	97.3	95.0	85.0	96.3	94.5	100	80.3	96.4	99.5	99.6	100
GD (60s)	93.6	94.8	89.2	93.7	88.0	84.6	86.9	96.6	92.5	89.9	83.7	77.7	99.6	89.5	98.3	74.4	93.6	87.6	87.4	97.3	95.0	85.0	96.3	94.5	100	80.3	96.4	99.5	99.6	100
GD (300s)	93.6	94.8	89.2	93.7	88.0	84.6	86.9	96.6	92.5	89.9	83.7	77.7	99.6	89.5	98.3	74.4	93.6	87.6	87.4	97.3	95.0	85.0	96.3	94.5	100	80.3	96.4	99.5	99.6	100
GD (1h)	93.6	94.8	89.2	93.7	88.0	84.6	86.9	96.6	92.5	89.9	83.7	100	99.6	89.5	98.3	74.4	93.6	87.6	87.4	97.3	95.0	85.0	96.3	94.5	100	80.3	96.4	99.5	99.6	100

(B) C2 (plant i4)

time period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GLS (30s)	99.5	95.3	92.9	96.2	95.2	96.0	97.6	98.7	94.3	96.3	98.5	93.0	95.1	94.8	93.2	98.1	95.5	100	94.3	86.4	97.5	93.3	94.1	97.1	97.8	99.2	95.0	98.2	96.2	99.7
GLS (60s)	99.5	99.2	98.8	96.2	100	99.5	97.6	98.7	94.6	96.3	98.5	95.3	95.1	94.8	93.2	98.1	95.6	100	94.3	86.6	97.5	93.5	97.7	97.1	98.3	99.2	96.0	99.1	96.2	99.7
GLS (300s)	100	100	98.2	100	100	100	100	100	100	96.5	100	98.2	97.8	100	100	98.5	100	100	98.0	100	99.6	100	100	99.6	100	99.6	100	100	100	100
GLS (1h)	100	100	98.8	100	100	100	100	100	100	100	100	98.2	100	100	100	98.5	100	100	100	100	100	100	100	100	100	99.6	100	100	100	100
TS (30s)	96.1	99.1	98.8	87.4	100	99.4	98.8	98.8	98.8	97.8	99.1	96.5	96.7	86.1	95.1	100	97.7	98.1	98.4	78.5	99.0	96.7	85.5	96.2	98.0	98.4	96.0	99.1	99.5	99.7
TS (60s)	96.3	99.1	99.7	87.4	100	99.4	99.6	98.9	98.8	99.8	99.1	96.5	96.7	86.1	96.5	100	98.7	98.1	98.4	78.7	100	97.4	100	96.2	100	99.2	96.0	99.1	99.5	99.7
TS (300s)	97.6	99.1	100	99.2	100	100	99.6	99.3	98.8	99.8	100	100	97.3	100	96.5	100	98.9	98.5	98.6	99.6	100	99.4	100	100	100	100	96.0	99.1	99.5	99.7
TS (1h)	99.5	99.1	100	99.2	100	100	99.6	99.4	99.9	100	100	100	99.7	100	99.6	100	99.6	100	99.9	99.6	100	99.5	100	100	100	100	96.0	99.6	99.5	99.7
SA (30s)	78.5	99.2	80.1	87.4	92.3	96.4	89.6	89.9	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.1	89.6	86.3	94.8	88.0	99.5	91.8	92.6	82.5
SA (60s)	78.5	99.2	80.1	87.4	92.3	96.4	89.6	89.9	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.1	93.6	86.3	94.8	88.0	99.5	91.8	92.6	82.5
SA (300s)	78.5	99.2	80.1	87.4	92.3	96.4	89.6	89.9	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.1	93.6	86.3	94.8	88.0	99.5	91.8	92.6	82.5
SA (1h)	78.5	99.2	80.1	87.4	92.3	96.4	89.6	89.9	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.1	93.6	86.3	94.8	88.0	99.5	91.8	92.6	82.5
GD (30s)	78.5	99.2	80.1	96.8	92.3	96.4	89.6	84.6	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.7	93.6	86.3	93.2	88.0	96.0	89.9	92.1	82.5
GD (60s)	78.5	99.2	80.1	96.8	92.3	96.4	89.6	84.6	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.7	93.6	86.3	93.2	88.0	96.0	89.9	92.1	82.5
GD (300s)	78.5	99.2	80.1	96.8	92.3	96.4	89.6	84.6	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.7	93.6	86.3	93.2	88.0	96.0	89.9	92.1	82.5
GD (1h)	78.5	99.2	80.1	96.8	92.3	96.4	89.6	84.6	71.5	92.2	97.1	87.1	91.4	94.8	85.5	81.8	91.6	84.4	76.9	81.4	96.0	89.7	93.6	86.3	93.2	88.0	96.0	89.9	92.1	82.5

Fig. 9 – Variation of the solution quality from the sensitivity analysis.

each cluster. The value in each cell was calculated based on the optimal solution gained from each metaheuristic approach with different CPU time limits and the best solution identified in each cluster. The cells also contain colours that is blended based on a colour scale given in the figure. By observing how cell colours change across the y-axis, we can see how the solution quality changes when using different metaheuristic and CPU time limits for each cluster. The pattern in cell colours across the time periods shows that time limit of 300 s with the GLS provides a reasonable balance between the computational time and the changes in solutions.

5. Concluding remarks

In this study, we addressed the integrated production and inventory routing planning of liquid oxygen supply chains during the COVID-19 pandemic. To overcome the computational difficulty of the problem, we proposed a two-level hybrid solution strategy that solves the problem via both the mathematical framework and metaheuristic method. At the upper level of the approach, we developed an MILP model in which only single-customer routes were considered. At the lower level, the problem was reduced to an MTHVRP by fixing the production and inventory decisions. Further, the MTHVRP was disaggregated into several independent problems by clustering customers based on the customer allocation results obtained from the upper level. We adapted the GLS approach to solve the MTHVRPs. The validity and efficiency of the proposed mathematical model and approach were tested in a case study on a real-world oxygen supply chain problem in the UK. The results showed that the developed MILP model and the proposed solution approach have a computational advantage. Further work could be directed at addressing the uncertainty of input parameters such as customer demand and transportation time and considering other solution techniques (e.g., rolling horizon and MILP decomposition methods) to enhance the computational efficiency of the proposed solution strategy.

CRediT authorship contribution statement

Yena Lee: Formal analysis, Investigation, Methodology, Visualisation, Validation, Software, Data curation, Writing – original draft, Writing – review & editing. **Vassilis M. Charitopoulos:** Investigation, Conceptualization, Writing – review & editing. **Karthik Thyagarajan:** Data curation, Validation, Writing – review & editing. **Jose M. Pinto:** Conceptualization, Validation, Writing – review & editing. **Ian Morris:** Data curation, Validation, Writing – review & editing. **Lazaros G. Papageorgiou:** Conceptualization, Investigation, Methodology, Validation, Project administration, Funding acquisition, Supervision, Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the financial support received from the UK Engineering and Physical Sciences Research Council (under the project EP/V050168/1).

Appendix A. Appendix: procedure for fixing the single-customer routes

In this section, we present the procedure for fixing single-customer routes obtained from the upper level. We fix the single-customer routes when the number of deliveries per customer in each time period is more than one.

First, we denote the optimal values of Q_{ikvt} and A_{ikvt}, as Q^*_{ikvt} and A^*_{ikvt}, respectively. Then, we define the binary and integer parameters as follows:

A^{fix}_{ikvt} fixed single-customer route in which vehicle v delivered the product from plant i to customer k in time period t
ND_{ikt} the number of deliveries from plant i to customer k in time period t
NT_{ikt} the number of vehicles travel from plant i to customer k in time period t
Y_{ikvt} 1 when customer k received the product from plant i by vehicle v in time period t

The binary and integer parameters are calculated based on Algorithm 1. As presented binary parameter Y_{ikvt} is equal to 1 when the vehicle v visits customer k in time period t (i.e., $A^*_{ikvt} \geq 1$). Then, the number of vehicles used for serving the product to customer k in time period t (NT_{ikt}) is calculated based on the binary parameter Y_{ikvt}. In addition, the number of deliveries to the customer is based on the optimal value of A^*_{ikvt} obtained from the previous level. Now, we fix the single-customer routes based on the parameters. For customer k who received the product from plant i more than once in time period t, there are two different types of deliveries: (i) served by the same vehicle; and (ii) served by different vehicles. The first case is that the number of deliveries is more than once, but the number of vehicles used is equal to one. In this case, we fix the single-customer routes except for the last delivery. (see lines 5–6 in Algorithm 1). On the other hand, the latter is the case in which the values of NQ_{ikt} and NT_{ikt} are greater than one. In this case, we fix the single-customer routes which the delivery quantity is equal to the vehicle capacity (see lines 7–8 in Algorithm 1).

Algorithm 1 Procedure of fixing the single-customer routes.

Input: A^*_{ikvt}, Q^*_{ikvt}
Output: A^{fix}_{ikvt}, Y_{ikvt}, Z_{ikvt}, NT_{ikt}, ND_{ikt}
1 **if** $A^*_{ikvt} \geq 1$ **then**
2 $\quad Y_{ikvt} = 1$
3 $\quad NT_{ikt} = \sum_{v:k \in K_{iv}} Y_{ikvt}$
4 $\quad ND_{ikt} = \sum_{v:k \in K_{iv}} A^*_{ikvt}$
5 **if** $ND_{ikt} > 1,\ NT_{ikt} = 1$ **then**
6 $\quad A^{fix}_{ikvt} = A^*_{ikvt} - 1$
7 **if** $Q^*_{ikvt} = CAP_v,\ ND_{ikt} > 1,\ NT_{ikt} > 1$ **then**
8 $\quad Z_{ikvt} = 1$

Next, we calculate the values of parameters that used for the MTHVRP at the lower level. The notation of the parameters is given below.

τ_{ivt} operation time limit of vehicle v at plant i in time period t (day)
D^{new}_{ikt} demand of customer k in a cluster associated with plant i and time period t (kg)
TT^{fix}_{it} transportation of fixed single-customer routes in a cluster associated with plant i and time period t (£)

Also, the equations for calculating values of the parameters are presented in.
Algorithm 2 .

Algorithm 2 Parameters used for the MTHVRP in the lower level.

Input: $Q^*_{ikvt}, A^{fix}_{ikvt}, Y_{ikvt}, Z_{ikvt}, NT_{ikt}, ND_{ikt}$
Output: $D^{new}_{ikt}, \tau_{ivt}, TT^{fix}_{it}$

if $ND_{ikt} = 1$ *and* $NT_{ikt} = 1$ **then**
  $D^{new}_{ikt} = \sum_{v:k\in K_{iv}} Q^*_{ikvt}$
if $ND_{ikt} > 1,\ NT_{ikt} = 1$ **then**
  $D^{new}_{ikt} = \sum_{v:k\in K_{iv}} (Q^*_{ikvt} - CAP_v A^{fix}_{ikvt})$
if $ND_{ikt} > 1,\ NT_{ikt} > 1$ **then**
  $D^{new}_{ikt} = \sum_{v:k\in K_{iv}} (Q^*_{ikvt} - CAP_v Z_{ikvt})$
$\tau_{ivt} = \Delta_t - \sum_{k\in K_{iv}} (\theta_{ik} + \alpha^{LT} + \alpha^{UT}) A^{fix}_{ikvt} - \sum_{k\in K_{iv}} (\beta_v CAP_v + \gamma_i CAP_v) A^{fix}_{ikvt} - \sum_{k\in K_{iv}} (\theta_{ik} + \alpha^{LT} + \alpha^{UT}) Z_{ikvt} - \sum_{k\in K_{iv}} (\beta_v CAP_v + \gamma_i CAP_v) Z_{ikvt}$
$TT^{fix}_{ivt} = \sum_{k\in K_{iv}} C^{veh}_v\ L_{ik}\ A^{fix}_{ikvt} + \sum_{k\in K_{iv}} C^{veh}_v\ L_{ik}\ Z_{ikvt}$

For a more comprehensive understanding of the procedure, assume that customer $k1$ received 10 kg of the product from plant $i1$ by vehicle $v1$, of which capacity is 10 kg, and 5 kg by vehicle $v2$, of which capacity is 8 kg, on a certain day. Also, customer $k2$ received 4 kg by vehicle $v2$ on the same day (see Fig. A.1). In case of customer $k1$, the delivery with 10 kg full load (red coloured route) is fixed, whereas the second delivery for 5 kg of the product (blue coloured route) will be reoptimised. Accordingly, in the lower-level MTHVRP, which is related to the plant and time period (i.e., plant i1 and the day), the demand of customer $k1$ will be 5 kg (D^{new}_{k1} = 5 kg). In addition, the remaining operation time for the vehicle will be Δ_t minus the operation time used for the fixed delivery of 10 kg of the product (i.e., τ_{v1} = 20 h). On the other hand, because customer $k2$ was visited only once, no route is fixed for the customer. In the MTHVRP, the available operation time for vehicle v2 (τ_{v2}) will be the same as the Δ_t and the demand of customer $k2$ (D^{new}_{k2}) will be 4 kg.

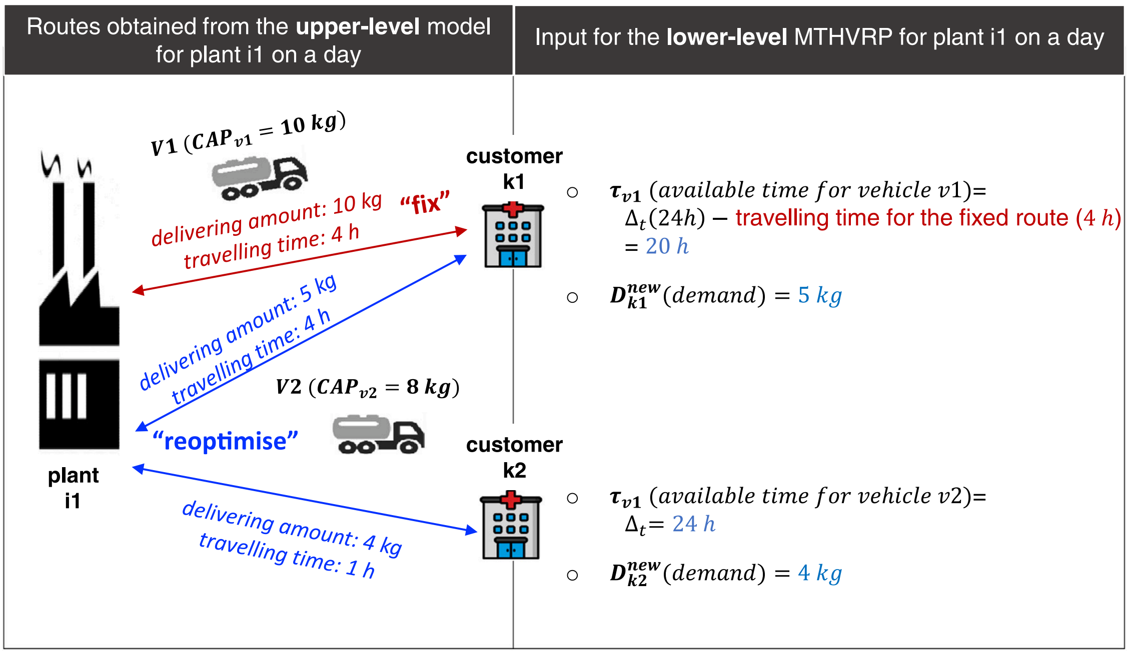

Fig. A.1 – Example for calculating input parameters used in the lower-level MTHVRP.

Table B.1 – Required oxygen demand per hospitalised patient.

	O2 flow rate (gas)		O2 flow rate (liquid)	
	per patient	total	total	per patient
severe 75 cases	10 L/min	10×60×75 = 45000 L/hr	45000/861[a]x24 = 1254 L/day	1254/75×1.14[b] = 19 kg/day
critical 25 cases	30 L/min	10×60×25 = 45000 L/hr	45000/861[a]x24 = 1254 L/day	1254/25×1.14[b] = 57 kg/day

[a] 1 L of liquid O2 at boiling point @ −183 C 101.31 kPa = 861 L of gas O2 @ 20 C 101.31 kPa
[b] 1 L of liquid O2 at boiling point @ −183 C 101.31 kPa = 1.14 kg

B. Appendix: hospital demand profile

The use of oxygen at hospitals is dominated by hospitalised coronavirus patients, who require on average 14 L/min oxygen gas for their survival based on the latest report by the World Health Organisation (WHO) (World Health Organization, 2020). In this section, we present the method used for estimating the hospital oxygen demand profile that is mainly related to the COVID-19 patients. First, we calculate the national oxygen demand based on the number of hospitalisations reported by the UK government (UK Health Security Agency, 2021) and WHO guideline that provides the required oxygen amount per COVID-19 patient (World Health Organization, 2020). According to the WHO guideline the COVID-19 patients requiring hospitalisation are classified into two cases, critical and severe cases. We assume that the hospitalisations at intensive care units (ICUs) are the critical cases, while the patients admitted to normal wards are the severe cases. The method for converting the number of hospitalisations to the oxygen demand is given in Table B.1. Next, we get the ratio of the total delivered amount to each hospital over the planning horizon to the national demand during the horizon. Note that the delivered amount is from historical data given by the industrial companies, Linde/BOC. Based on the ratio, we distribute the total delivered amount to follow the national demand profile. This procedure finds the hospital demand profile which is fitted to the given dataset regarding the national demand.

References

Absi, N., Archetti, C., Dauzère-Pérès, S., Feillet, D., 2015. A two-phase iterative heuristic approach for the production routing problem. Transp. Sci. 49, 784–795.

Archetti, C., Bertazzi, L., Paletta, G., Speranza, M.G., 2011. Analysis of the maximum level policy in a production-distribution system. Comput. Oper. Res. 38, 1731–1746.

Barbosa-Povoa, A.P., Pinto, J.M., 2020. Process supply chains: perspectives from academia and industry. Comput. Chem. Eng. 132, 106606.

Bard, J.F., Nananukul, N., 2009a. Heuristics for a multiperiod inventory routing problem with production decisions. Comput. Ind. Eng. 57, 713–723.

Bard, J.F., Nananukul, N., 2009b. The integrated production-inventory-distribution-routing problem. J. Sched. 12, 257–280.

Bard, J.F., Nananukul, N., 2010. A branch-and-price algorithm for an integrated production and inventory routing problem. Comput. Oper. Res. 37, 2202–2217.

Bertazzi, L., Savelsbergh, M., Speranza, M.G., 2008. Inventory routing. In: The Vehicle Routing Problem: Latest Advances and New Challenges. Springer, pp. 49–72.

Beullens, P., Muyldermans, L., Cattrysse, D., Van Oudheusden, D., 2003. A guided local search heuristic for the capacitated arc routing problem. Eur. J. Oper. Res. 147, 629–643.

Boudia, M., Prins, C., 2009. A memetic algorithm with dynamic population management for an integrated production-distribution problem. Eur. J. Oper. Res. 195, 703–715.

Boudia, M., Louly, M.A.O., Prins, C., 2006. A memetic algorithm with population management for a production-distribution problem. IFAC Proc. 39, 541–546.

Boudia, M., Louly, M.A.O., Prins, C., 2007. A reactive grasp and path relinking for a combined production-distribution problem. Comput. Oper. Res. 34, 3402–3419.

Brown, G., Keegan, J., Vigus, B., Wood, K., 2001. The kellogg company optimizes production, inventory, and distribution. Interfaces 31, 1–15.

Campbell, A., Clarke, L., Kleywegt, A., Savelsbergh, M., 1998. The Inventory Routing Problem. In: Fleet management and logistics. Springer, pp. 95–113.

Campbell, A.M., Savelsbergh, M.W., 2004. A decomposition approach for the inventory-routing problem. Transp. Sci. 38, 488–502.

Çetinkaya, S., Üster, H., Easwaran, G., Keskin, B.B., 2009. An integrated outbound logistics model for frito-lay: coordinating aggregate-level production and distribution decisions. Interfaces 39, 460–475.

Cóccola, M.E., Méndez, C.A., Dondo, R.G., 2020. A two-stage procedure for efficiently solving the integrated problem of production, inventory, and distribution of industrial products. Comput. Chem. Eng. 134, 106690.

Coelho, L.C., Cordeau, J.F., Laporte, G., 2014. Thirty years of inventory routing. Transp. Sci. 48, 1–19.

Dong, Y., Maravelias, C.T., Pinto, J.M., Sundaramoorthy, A., 2017. Solution methods for vehicle-based inventory routing problems. Comput. Chem. Eng. 101, 259–278.

Etebari, F., Dabiri, N., 2016. A hybrid heuristic for the inventory routing problem under dynamic regional pricing. Comput. Chem. Eng. 95, 231–239.

Fahimnia, B., Farahani, R.Z., Marian, R., Luong, L., 2013. A review and critique on integrated production-distribution planning models and techniques. J. Manuf. Syst. 32, 1–19.

Fumero, F., Vercellis, C., 1999. Synchronized development of production, inventory, and distribution schedules. Transp. Sci. 33, 330–340.

Glankwamdee, W., Linderoth, J., Shen, J., Connard, P., Hutton, J., 2008. Combining optimization and simulation for strategic and operational industrial gas production and distribution. Comput. Chem. Eng. 32, 2536–2546.

Higgins, V., Sohaei, D., Diamandis, E.P., Prassas, I., 2021. Covid-19: from an acute to chronic disease? potential long-term health consequences. Crit. Rev. Clin. Lab. Sci. 58, 297–310.

Karakostas, P., Sifaleras, A., Georgiadis, M.C., 2019. A general variable neighborhood search-based solution approach for the location-inventory-routing problem with distribution outsourcing. Comput. Chem. Eng. 126, 263–279.

Karakostas, P., Sifaleras, A., Georgiadis, M.C., 2020. Adaptive variable neighborhood search solution methods for the fleet size and mix pollution location-inventory-routing problem. Expert Syst. Appl. 153, 113444.

Lai, C.C., Shih, T.P., Ko, W.C., Tang, H.J., Hsueh, P.R., 2020. Severe acute respiratory syndrome coronavirus 2 (sars-cov-2) and coronavirus disease-2019 (covid-19): The epidemic and the challenges. Int. J. Antimicrob. Agents 55, 105924.

Lee, Y., Carrero-Parreno, A., Ramaswamy, S., Pinto, J.M., Papageorgiou, L.G., 2021. Hierarchical approach to integrated

planning of industrial gas supply chains. Ind. Eng. Chem. Res. 60, 5207–5219.
Lei, L., Liu, S., Ruszczynski, A., Park, S., 2006. On the integrated production, inventory, and distribution routing problem. IIE Trans. 38, 955–970.
Marchetti, P.A., Gupta, V., Grossmann, I.E., Cook, L., Valton, P.M., Singh, T., Li, T., André, J., 2014. Simultaneous production and distribution of industrial gas supply-chains. Comput. Chem. Eng. 69, 39–58.
Misra, S., Saxena, D., Kapadi, M., Gudi, R.D., Srihari, R., 2018. Short-term planning framework for enterprise-wide production and distribution network of a cryogenic air separation industry. Ind. Eng. Chem. Res. 57, 16841–16861.
Perron, L., Furnon, V., 2022.Google-Tools v7.2. Routing solver. ⟨https://developers.google.com/optimization/routing/routing_options⟩.
Ramaswamy, S., Madan, T., Thyagarajan, K., Pinto, J.M., Laínez-Aguirre, J.M., 2020. Advanced decision-support technologies for the design and management of industrial gas supply chains. In: Smart Manufacturing. Elsevier, pp. 387–421.
Savelsbergh, M., Song, J.H., 2008. An optimization algorithm for the inventory routing problem with continuous moves. Comput. Oper. Res. 35, 2266–2282.
Thomas, D.J., Griffin, P.M., 1996. Coordinated supply chain management. Eur. J. Oper. Res. 94, 1–15.
UK Health Security Agency, 2021.Healthcare in United Kingdom. ⟨https://coronavirus.data.gov.uk/details/healthcare⟩.
Voudouris, C., Tsang, E., 1999. Guided local search and its application to the traveling salesman problem. Eur. J. Oper. Res. 113, 469–499.
Voudouris, C., Tsang, E.P., Alsheddy, A., 2010. Guided local search. In: Handbook of Metaheuristics. Springer, pp. 321–361.
Waller, M., Johnson, M.E., Davis, T., 1999. Vendor-managed inventory in the retail supply chain. J. Bus. Logist. 20, 183.
Weyland, D., 2008. Simulated annealing, its parameter settings and the longest common subsequence problem. Proc. 10th Annu. Conf. Genet. Evolut. Comput. 803–810.
World Health Organization, 2020.Oxygen sources and distribution for COVID-19 treatment centres: interim guidance, 4 April 2020. Technical Report. World Health Organization.
You, F., Pinto, J.M., Capón, E., Grossmann, I.E., Arora, N., Megan, L., 2011. Optimal distribution-inventory planning of industrial gases. i. fast computational strategies for large-scale problems. Ind. Eng. Chem. Res. 50, 2910–2927.
Zamarripa, M., Marchetti, P.A., Grossmann, I.E., Singh, T., Lotero, I., Gopalakrishnan, A., Besancon, B., André, J., 2016. Rolling horizon approach for production-distribution coordination of industrial gases supply chains. Ind. Eng. Chem. Res. 55, 2646–2660.
Zhang, Q., Sundaramoorthy, A., Grossmann, I.E., Pinto, J.M., 2017. Multiscale production routing in multicommodity supply chains with complex production facilities. Comput. Oper. Res. 79, 207–222.