Optimal Distribution-Inventory Planning of Industrial Gases: I. Fast Computational Strategies for Large-Scale Problems

Fengqi You,[1,2] Jose M. Pinto,[3] Elisabet Capón,[1] Ignacio E. Grossmann,[1*]
Nikhil Arora,[3] Larry Megan[3]

[1]Dept.of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213
[2] Argonne National Laboratory, Argonne, IL 60439
[3]Praxair Inc., 39 Old Ridgebury Road, Danbury, CT 06810

November, 2010

To be submitted to *Industrial & Engineering Chemistry Research*

Abstract

In this paper, we address the optimization of industrial gas distribution systems, which consist of plants and customers, as well as storage tanks, trucks and trailers. A mixed-integer linear programming (MILP) model is presented to minimize the total capital and operating cost, and to integrate short-term distribution planning decisions for the vehicle routing with long-term inventory decisions for sizing storage tanks at customer locations. In order to optimize asset allocation in the industrial gas distribution network by incorporating operating decisions, the model also takes into account the synergies among delivery schedule, tank sizes, customer locations and inventory profiles. To effectively solve large scale instances, we propose two fast computational strategies. The first approach is a two-level solution strategy based on the decomposition of the full scale MILP model into an upper level route selection - tank sizing model and a lower level reduced routing model. The second approach is based on a continuous approximation method, which estimates the operational cost at the strategic level and determines the tradeoff with the capital cost from tank sizing. Three cases studies including instances with up to 200 customers are presented to illustrate the applications of the models and the performance of the proposed solution methods.

Keywords: planning & scheduling, industrial gas supply chain, MINLP, vehicle routing, continuous approximation, tank sizing

[*] To whom all correspondence should be addressed. E-mail: grossmann@cmu.edu

1. Introduction

A distribution network of industrial merchant liquid products (Nitrogen, Oxygen, Argon, Carbon Dioxide, Helium and Hydrogen) consists of plants and customers, as well as storage facilities, trucks and trailers. In particular, customer inventories in this distribution network are managed by the vendor of industrial gases, i.e. the vendor installs storage tanks in customer locations with proper sizes and manages their replenishments to satisfy customer demands by coordinating the deliveries. Short-term distribution planning decisions involve deciding which customers receive deliveries each day, when to deliver, how much to deliver, how to combine deliveries into routes, how to combine routes into the drivers' daily schedules, determining which truck or trailer for each delivery and the capacity of each truck for delivery. The long-term inventory decisions involve deciding how many tanks to install in each customer location, the size of each tank, and when and how to install new tanks at customer locations, as well as when and how to upgrade and downgrade existing tanks. To minimize the total capital and operating costs, the short-term distribution planning decisions should be integrated with the long-term inventory decisions. This integration requires accounting for the synergies between the customers in terms of locations and tank sizes, and to consider the interactions of tank sizes and inventories between customers. The challenge is how to effectively solve the resulting large-scale mixed-integer programming model in order to optimize the capital asset allocation in the industrial gas distribution network by incorporating operating decisions.

In this paper, we present an integrated mixed-integer linear programming (MILP) model for industrial gas distribution-inventory planning using a slot-based scheduling model for vehicle routing. While effective for short-term problems, the model becomes computationally expensive to solve for long planning horizons, which is necessary for the integration of strategic tank sizing decisions and operational vehicle routing decisions. Hence, two solution strategies are proposed to reduce the computational effort.

The first approach given in the appendix, consists of a two-level strategy. In the upper level, we solve a simultaneous route selection and tank sizing model, which is a relaxation of the integrated MILP model by neglecting the decisions on delivery schedules and considering the "worst case" working inventory for tank sizing. The solution of the upper level problem yields the optimal tank sizes and the possible routes for delivery. Next, we fix the previous tank sizes and solve a reduced routing problem which only considers those routes determined by the upper level problem. Since the reduced routing problem only considers a subset of all possible routes, it is computationally more efficient than the original routing problem. Therefore, the detailed schedule and quantity of each delivery and the inventory profile of each customer are determined by the lower level problem.

The second solution strategy given in Section 5, is based on a continuous approximation method.[1-3] This approach consists of two phases: in the first phase we solve an upper level continuous approximation model and in the second phase we solve a lower level detailed routing model based on the results obtained in the first phase. The continuous approximation model predicts the optimal sizes of tanks to be installed, downgraded or upgraded in each customer location over the given planning horizon. Using the continuous approximation method for capacitated vehicle routing to estimate the total distribution cost, the upper level model trades off the capital cost from tank sizing with the operating cost from continuous approximation without considering the routing details. The resulting upper level problem is a non-convex mixed-integer nonlinear programming (MINLP) model with nonlinear terms from the continuous approximation constraints. After introducing additional variables and constraints to exactly linearize the nonlinear terms, the model is reformulated as an MILP, which can be globally optimized very effectively even for large-scale instances. In the second phase, we fix the previously determined tank sizing decisions and solve the detailed routing problem in the reduced variable space in. This model predicts the detailed vehicle routing decisions including the sizes of deliveries and the inventory levels of each customer over the planning horizon, as well as the detailed timing and sequence of deliveries with trucks of different capacities.

We present the aforementioned model formulations and computational strategies in this paper. Three case studies with up to 200 customers are solved to illustrate the application of the proposed models and solution approaches. The results show that the proposed solution strategies, especially the continuous approximation method, can obtain global optimal or near-optimal solutions very quickly even for large-scale problems.

The rest of this paper is organized as follows. We first review the related literature in Section 2. The general problem statement is provided in Section 3, which is followed by the integrated MILP model formulation for simultaneous tank sizing and vehicle routing, in Section 4. The proposed continuous approximation approach is presented in Sections 5. Computational results for three case studies and the conclusions of this work are then given at the end of this paper. The simultaneous route selection and tank sizing approach is given in the Appendix.

2. Literature Review

Although distribution-inventory planning is an important problem for the industrial gas industry, there is relatively little chemical engineering literature on this topic. Glankwamdeea et al.[4] studied the production and distribution planning of an industrial gas supply chain. However, they did not consider tank sizing issues or detailed distribution planning (i.e. vehicle routing) as are addressed in this paper. On the other hand, there are a number of articles addressing the vehicle routing problems for the process industry, but none of them has considered the tank sizing issue. Choi et al.[5] developed an approximate stochastic dynamic programming approach for the traveling salesman problem under uncertainty. Their algorithmic framework is shown to be computationally very efficient compared to the stochastic dynamic programming in the full space, without significant loss in the solution quality. Using a mathematical programming approach, Jetlund and Karimi[6] proposed an MILP model based on variable-length slots for the maximum-profit scheduling of a fleet of multi-parcel tankers engaged in shipping bulk liquid chemicals. The authors propose a heuristic decomposition algorithm that obtains the fleet schedule by repeatedly solving the base formulation for a single ship. An MILP model is proposed by Huang and Chung[7] that integrates the routing and scheduling decisions for production planning of pipeless plants with different layouts. Recently, Dondo et al.[8] developed an exact MILP mathematical formulation for the multiple vehicle time-window-constrained pickup and delivery problem. Their approach is able to account for many-to-many transportation requests, pure pickup and delivery tasks, heterogeneous vehicles and multiple depots. Based on this work, Dondo and Cerda[9] proposed an MILP model for large-scale multi-depot vehicle routing problems with time windows. To further reduce the problem size, the authors also developed a spatial decomposition scheme, such that large problems with up to 200 customers, multiple depots and different vehicle-types were solved with quite reasonable computational effort. A novel MILP mathematical framework for the short-term vehicle routing problem of multiechelon multiproduct transportation networks was recently proposed by Dondo et al.[10] Their model relies on a continuous-time representation and applies the general precedence notion to model the sequencing constraints establishing the ordering of vehicle stops on every route.

Related routing problems have also been extensively studied by the operations research community in the past decades, but most of the existing works only consider operational planning of inventory and distribution, without integration with the strategic tank sizing decisions. A general review of vehicle routing problem is given by Laporte.[11] Specific reviews for inventory-routing problems are given by Baita et al.[12] and Moin and Salhi,[13] and a specific review for strategic location-routing problem is given by Nagy and Salhi.[14]

Few works closely related to this paper are reviewed below. Webb and Larson[15] addressed a similar problem that integrates strategic fleet sizing decisions with distribution and inventory planning. A decomposition approach for solving the large scale MILP models for inventory-routing problem was proposed by Campbell and Savelsbergh.[16] Lei et al.[17] proposed a two-phase solution approach to the integrated production, inventory, and routing problem. The main advantage of their approach is that the two-phase approach is able to simultaneously coordinate the production, inventory, and transportation operations of the entire planning horizon, without the need to aggregate the demand or relax the constraints on transportation capacities. The authors also reported real-world case studies to illustrate the performance of their computational framework.

The distribution-inventory planning of supply chains usually leads to large-scale optimization problems that are difficult to solve. To address the computational challenge, various decomposition methods, such as Benders decomposition,[18] Lagrangean decomposition,[19-20] bilevel decomposition[21] and hierarchical decomposition[22] are proposed. Benders decomposition[18] is usually used to solve complex mixed-integer programming models arising from process planning and scheduling (e.g. see the work by Pinto & Grossmann[23]) or large-scale stochastic programming problems in supply chain operations under uncertainty.[24] Lagrangean decomposition[19-20] can effectively solve large-scale supply chain planning problems with “decomposable” structures.[25-29] If the supply chain planning problem includes both strategic and operational decisions, bilevel decomposition[21] can be implemented to iteratively solve an upper level aggregated model and a lower level detailed model.[30-32] Hierarchical decomposition[22] is effective for supply chain optimizatoin problems with multiple levels of decision-making.[33-36]. In addition to these standard decomposition methods, some problem-dependent solution algorithms, such as those based on successive MILP decomposition[37] and successive piece-wise linear approximation[38], have been proposed to handle the computational challenges arising in industrial-sized supply chain planning problems.

3. Problem Statement

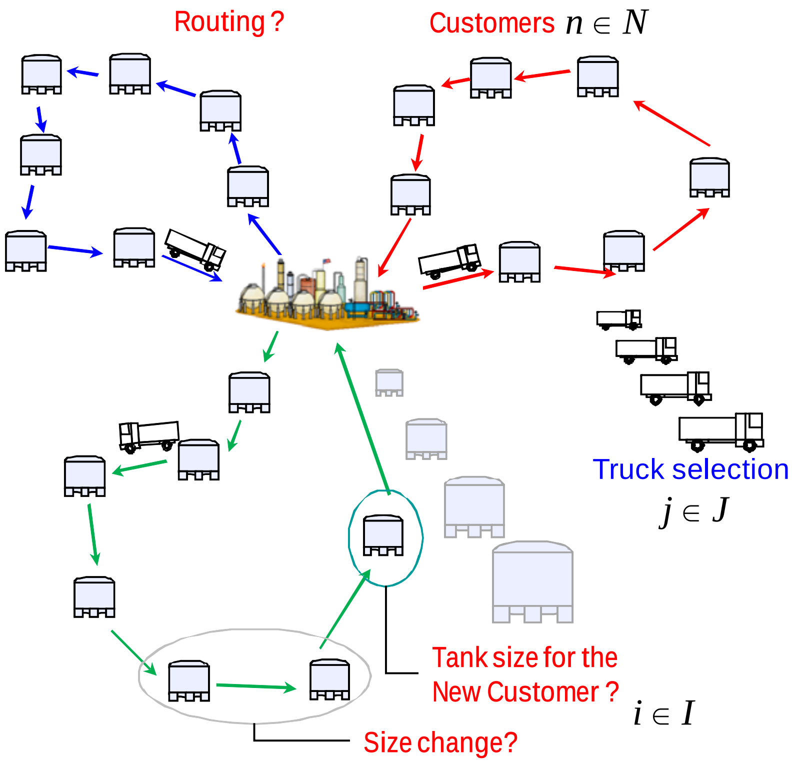


Figure 1. Tank sizing and vehicle routing of industrial gas supply chains

We are given an industrial gas distribution network consisting of a production plant and a set of customers $n \in N$ as shown in Figure 1. The locations of the plant and customers, as well as the distances between them are given. Each customer has a deterministic and constant demand rate $dem_{n,y}$ and safety stock level $safety_{n,y}$ in each year $y \in Y$ within the planning horizon. We are also given a set of possible tanks with different sizes $i \in I$. The lower and upper bounds for tank with size *i* are defined as T_i^L and T_i^U.

If customer *n* has an existing tank with size *i*, we set the parameter $ot_{i,n} = 1$, otherwise $ot_{i,n} = 0$. Similarly, if customer *n* is a new customer without any existing tank, we set the parameter $new_n = 1$, otherwise it is equal to zero. Each newly installed tank is full of merchant liquid at the beginning, i.e. initial inventory $Vzero_{n,y}$ is assumed to be equal to the tank capacity. For customers with existing tanks, their specific initial inventory levels are given. New customers need to determine the size of tank to be installed in their location. Existing customers can upgrade or downgrade the existing tanks, or add a second tank if extra space is available.

There are nt_j trucks of size $j \in J$ with discrete capacities are defined by $Vtruck_j$. The delivery cost per distance traveled for every truck $j \in J$ is ck_j. All the trucks are assumed to have the same average traveling speed ($speed$) and a maximum number of working hours per day (hpd). For each delivery, there is a fixed percent of product loss, denoted as $loss$, and a minimum unloaded fraction given as $frac$.

The capital cost of tank with size $i \in I$ is given as $Ccap_i$, and the service cost of installing, upgrading or downgrading a tank with size $i \in I$ is $Cser_i$. Both capital and service costs are depreciated with a working capital discount factor, $wacc$, and a depreciation period in years (dep). In addition, there is a unit outage cost $Cout_{n,y}$ for unsatisfied demand of customer *n* in year *y*.

The problem is to simultaneously determine the tank sizing and modification decisions at each new and existing customer, as well as the schedule and quantity of each delivery in order to minimize the total capital, service, distribution and outage costs.

Based on the aforementioned discussion, the major assumptions of this problem are listed as follows:

- Only one type of industrial gas is considered
- When a new tank is installed, the initial inventory is the full tank capacity
- All the trucks have the same traveling speed

4. Integrated MILP Model

We first formulate the aforementioned problem as an integrated MILP model, which simultaneously considers tank sizing and vehicle routing, and predicts the optimal delivery schedule, delivery quantity, truck selection decisions, new tank installation decisions, upgrade and downgrade decisions of existing tanks at customer location, and the detailed inventory profile of each customer. A slot-based scheduling representation is used to model the vehicle routing decisions. In this model, the set $t \in T$ is introduced for the events occurring in each year or each time period. The number of time events is selected such that the optimal solution does not change if the number increases. We also introduce the set $r \in R$ for all the possible routes that start at the plant, by way of at least one customer and ending at the plant. If the number of customers involved in a route is not constrained, there are $2^{|N|}-1$ possible routes for a distribution network with $|N|$ customers and one plant. We further define two subsets N_r, which includes all the customers that are served by route *r*, and R_n, which includes all the routes involving customer *n*. In addition, the parameter $dist_r$ represents the distance of route *r*. A list of indices, sets, parameters and variables are given in the Nomenclature section.

4.1 Objective function

The objective function of this MILP is to minimize the total cost, including capital cost, service cost, distribution cost and outage cost, as given in Equation (1),

Min: $Cost = capcost + servcost + distcost + outcost$ (1)

where the detailed cost components are listed in constraints (2) – (6).

$$capcost = \frac{1}{dep}\left(\sum_i Ccap_i \sum_y \left[\begin{array}{l} \displaystyle\sum_{n|new_n=0 \wedge tsize_n=0} \frac{ot_{i,n}}{(1+wacc)^{y-1}} + \sum_{\substack{n|new_n=0 \wedge tsize_n=1 \\ \wedge espace_n=0}} \frac{yt_{i,n}}{(1+wacc)^{y-1}} + \\ \displaystyle\sum_{n|new_n=0 \wedge tsize_n=1 \wedge espace_n=1} \frac{ot_{i,n}+et_{i,n}}{(1+wacc)^{y-1}} + \sum_{n|new_n=1} \frac{yt_{i,n}}{(1+wacc)^{y-1}} \end{array}\right]\right) \quad (2)$$

$$servcost = \frac{1}{dep}\left(\sum_i \sum_n Cser_i \sum_y \frac{tins_{i,n}+et_{i,n}}{(1+wacc)^{y-1}}\right) \quad (3)$$

where $tins_{i,n} \geq yt_{i,n} + ot_{i,n}$ (4)

$$outcost = \sum_n \sum_t \sum_y Cout_{n,y} out_{n,t,y} \quad (5)$$

$$distcost = \sum_j \sum_r \sum_t \sum_y \frac{dist_r \cdot ck_j \cdot z_{j,r,t,y}}{(1+wacc)^y} \quad (6)$$

The four terms in equation (2) for capital cost correspond to rating (no tank sizing), replacing an existing tank, adding a tank to the extra space an existing customer and sizing the tank at a new customer, respectively. Any tank change or addition leads to a service cost, as shown in Equations (3) and (4). Equation (5) accounts for outage cost, where $out_{n,t,y}$ is the outage amount of customer n at event point t in year y. The outage cost guarantees feasibility of the model in case of large demands and represents a penalty for not satisfying them. The total distribution cost is given as the summation of all the delivery costs as in Equation (6). $z_{j,r,t,y}$ is a binary variable that equals to 1 if truck j is used in time event t for delivery with route r in year y. If truck j delivers product using route r, the corresponding distribution cost of this trip equals to the product of route distance ($dist_r$) and the delivery cost per unit distance traveled of this truck (ck_j). Note that the capital, service and distribution costs are all depreciated with working capital discount factor $wacc$.

4.2 Tank selection constraints

Two parameters $tsize_n$ and $espace_n$ are introduced to define the conditions for tank sizing. If there is extra space for installing another tank at customer n, we have the parameter $espace_n = 1$, otherwise we set it to zero. If the tank at customer n needs to be sized or changed, we have the parameter $tsize_n = 1$, otherwise $tsize_n = 0$.

If the tank of customer n needs to be sized or changed, we would either install a new tank in the extra space or change the current tank. If no tank is sized, no tank modification action will be taken. This relationship can be modeled by the following constraint:

$$\sum_i \left(yt_{i,n} + et_{i,n} \right) \le tsize_n, \qquad \forall n \tag{7}$$

where $et_{i,n}$ is a binary variable that equals to 1 if customer n has tank of size i installed in extra space, and $yt_{i,n}$ is also a binary variable that equals to 1 if the customer n will be installed with a tank of size i.

If the storage tanks at customer n should be changed ($tsize_n = 1$) and there is extra space ($espace_n$=1), then at most one type of tank should be selected to be installed at the extra space, i.e. one binary variable $et_{i,n}$ must be selected, and at the same time there is no change to the existing tank, i.e. all $yt_{i,n}$ variables are zero. On the other hand, if there is no extra space, then no tank should be installed at the extra space, and at least one type of tank should be selected to replace the existing tank. Constraints (8) and (9) model these logic relationships.

$$\sum_i yt_{i,n} = 1 - espace_n \qquad \forall n \mid tsize_n = 1 \tag{8}$$

$$\sum_i et_{i,n} \le espace_n \qquad \forall n \tag{9}$$

We note that although a potentially better approach for capital investment planning would be to allow the tank size to change every year using a multi-period formulation for tank selection, we assume in this work that the tank sizes will not change in the planning horizon after installation in the first year. Given the dynamic nature of the market and that uncertainty customer demand and in its set of neighbors grows in the future, capital investment decisions are made in the present and the model is optimized on a periodic basis to assess potential changes in the capacity of the network

4.3 Tank balance constraints

The inventory level of a customer decreases due to product demand, and increases due to replenishments. The mass balance of the customer inventory implies that the inventory at the beginning of any time event *t* plus the replenishment amount, should be equal to the inventory level at end of this time event plus the satisfied demand. Thus, eq. (10) shows that the initial inventory ($Vzero_n$) is the same as the inventory at the first time point minus the demand over the duration of the first time event ($\Delta t_{t=0,y=1}$).

$$Vo_{n,t=1,y=1} = Vzero_n - dem_{n,y=1}\Delta t_{t=0,y=1} \qquad \forall n \tag{10}$$

where $Vo_{n,t,y}$ is the inventory level of customer *n* at time event *t* of year *y* and $dem_{n,y}$ is the demand rate of customer *n* in year *y*.

The inventory balance of a customer at other time points is given by constraints (11) and (12).

$$Vo_{n,t=1,y} = Vo_{n,t=|T|,y-1} + p_{n,t=|T|,y-1} - dem_{n,y}\Delta t_{t=0,y} \qquad \forall n, y > 1 \tag{11}$$

$$Vo_{n,t,y} = Vo_{n,t-1,y} + p_{n,t-1,y} - dem_{n,y}\Delta t_{t-1,y} \qquad \forall n, t > 1, \forall y \tag{12}$$

where $p_{n,t,y}$ is the delivery (replenishment) amount to customer *n* at time event *t* of year *y*. The volumetric balance constraints above represent the tank levels at different time slots over the horizon.

At any time, the inventory level of a customer should not fall below the minimum inventory (Vl_n), which is determined by the property of the product and the storage tank. Thus, we have the following constraint.

$$Vl_n \leq Vo_{n,t,y} \qquad \forall n,t,y \qquad (13)$$

If the inventory level falls below the safety stock level, there will be unsatisfied demand $out_{n,y}$.

$$Vo_{n,t,y} + out_{n,t,y} \geq Vl_n + safety_{n,y} \qquad \forall n,t,y \qquad (14)$$

where $safety_{n,y}$ is the safety stock level of customer *n* in year *y*.

The inventory level plus the replenishment amount ($p_{n,t,y}$) of a customer should not exceed the maximum inventory level (Vu_n), which is the customer's tank size at any time.

$$Vo_{n,t,y} + p_{n,t,y} \leq Vu_n \qquad \forall n,t,y \qquad (15)$$

The minimum and maximum inventory levels of customer *n* depend on the storage tank(s) installed for this customer. Thus, they are modeled through the following two equations,

$$Vl_n = \sum_i T_i^L \left(ot_{i,n} + yt_{i,n} + et_{i,n} \right) \qquad \forall n \qquad (16)$$

$$Vu_n = \sum_i T_i^U \left(ot_{i,n} + yt_{i,n} + et_{i,n} \right) \qquad \forall n \qquad (17)$$

where T_i^U is the discrete tank size and T_i^L is the corresponding inventory lower bound for tank with size *i*.

If a new customer *n* joins the distribution network, at least a new tank is selected and assumed to be at full level; otherwise initial inventory level parameters are inputted.

$$Vzero_n = Vu_n \qquad \forall n \mid new_n = 1 \qquad (18)$$

4.4 Truck delivery constraints

Constraint (19) enforces that if a delivery is made through a route *r*, then it has to satisfy a minimum fraction of the truck load,

$$\sum_{n \in N_r} pr_{n,r,t,y} \geq frac \sum_j z_{j,r,t,y} Vtruck_j \left(1 - loss\right) \qquad \forall r,t,y \qquad (19)$$

where $Vtruck_j$ is the capacity of truck j, *loss* is the product loss percentage per delivery, *frac* is the minimum tanker fraction unloaded, $z_{j,r,t,y}$ is a binary variable that equals to 1 if truck j delivers in time event t of year y, and $pr_{n,r,t,y}$ is the delivery amount to customer n in route r at time event t of year y.

The total replenishment amount per delivery should not exceed the truck capacity after accounting for the product loss, although some product is allowed to return to the source. Thus, we have the following constraint,

$$\sum_{n\in N_r} pr_{n,r,t,y} \le \sum_j z_{j,r,t,y} Vtruck_j \left(1-loss\right) \qquad \forall r,t,y \qquad (20)$$

Constraint (21) represents the fact that the total replenishment amount that customer n receives at time event t of year y is the summation of the through all the possible routes involving this customer.

$$p_{n,t,y} = \sum_{r\in R_n} pr_{n,r,t,y} \qquad \forall n,t,y \qquad (21)$$

The number of deliveries is bound by the demands and truck sizes. It yields the following constraint.

$$\frac{dem_{n,y}}{\max_j \{Vtruck_j\}} \le \sum_j \sum_{r\in R_n} \sum_t z_{j,r,t,y} \le \frac{dem_{n,y}}{\min_j \{Vtruck_j\}} \qquad \forall n,y \qquad (22)$$

4.5 Timing constraints

The time interval cannot be less than the period to deliver to the customers. This is composed by the travel time, the time to lad the tank of the customer, and another time period to set up the truck at the source,

$$\Delta t_{t,y} \ge z_{j,r,t,y}\left(\frac{dis_r}{speed\cdot hpd} + FT\_load\cdot\left|N_r\right| + FT\_del\right), \qquad t>1, \forall y \qquad (23)$$

where $\Delta t_{t,y}$ is the time interval in time event t of year y, hpd is the maximum number of working hours per day, dis_r is the total traveling distance of route r, *FT_load* is the loading time for each customer, and *FT_del* is the loading time for each delivery at the plant.

The time interval in event t-1 of year y ($\Delta t_{t-1,y}$) should be equal to the initial time of event t in year y ($ti_{t,y}$) minus the initial time of the previous time event.

$$\Delta t_{t-1,y} = ti_{t,y} - ti_{t-1,y} \qquad \forall t,y \qquad (24)$$

For every year y, all the events start at time zero and end at day 365. Thus, we have the following initial conditions.

$$\Delta t_{t=1,y} = ti_{t=1,y} \qquad \forall y \qquad (25)$$

$$ti_{t=|T|,y} = 365 \qquad \forall y \qquad (26)$$

The following constraint is introduced to restrict that earlier time slots are selected first.

$$\sum_j \sum_r z_{j,r,t,y} \geq \sum_j \sum_r z_{j,r,t-1,y} \qquad \forall t, y \qquad (27)$$

4.6 Computational complexity

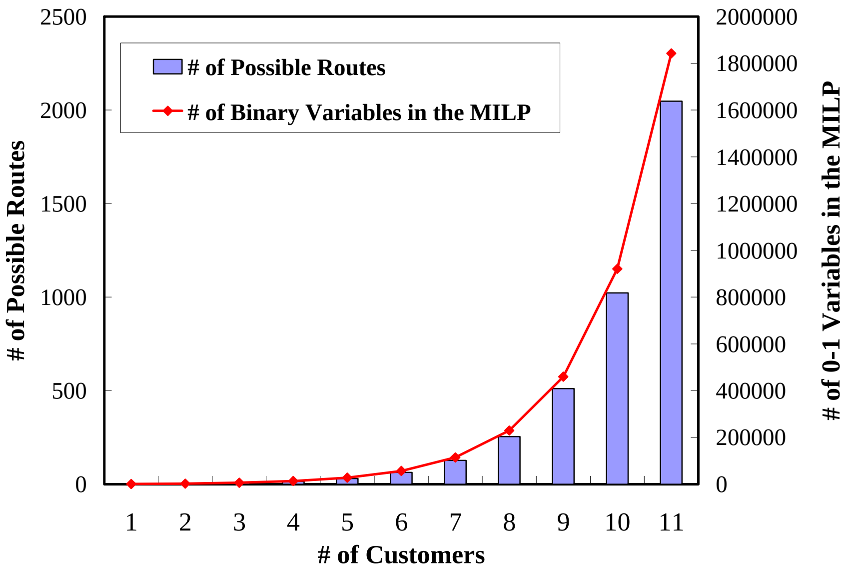


Figure 2. Computational complexity of the integrated MILP model

The simultaneous tank sizing and vehicle routing problem discussed above yields an MILP model, including constraints (2)-(27) and the objective function (1). Although the model is quite comprehensive as it accounts for both strategic decisions on tank modification and the operational decisions for vehicle routing, it can be computationally very expensive for practical applications. For example, an industrial gas distribution network with only 5 customers, 4 possible truck capacities and 6 potential tank sizes under 3-year planning horizon leads to an MILP problem with more than 100,000 binary variables. The significant computational complexity arises from the binary variable $z_{j,r,t,y}$ for the detailed vehicle routing. Since the number of possible routes increases exponentially as the number of customers increases, the problem size of the corresponding MILP model can greatly increase when more customers are considered simultaneously to account for synergy effects (see Figure 2). To address the computational challenge, we propose two solution strategies: the continuous approximation approach given in the following section and the simultaneous route selection and tanks sizing approach discussed in the Appendix.

5. Continuous Approximation Approach

This strategy employs a continuous approximation approach to estimate the annual delivery cost without considering the detailed schedules of the routing problem. By accounting for the capacitated vehicle routing cost at the strategic level, the trade-off between the capital cost and operational cost is established. After the strategic tank sizing decisions are determined, detailed vehicle routing is considered for operational decisions. The major advantage is that both the upper level continuous approximation model and the lower level detailed routing problem can be solved effectively without sacrificing too much solution quality. The major drawback is that the optimality gap cannot be estimated because a theoretical lower bound is not available.

The detailed vehicle routing model in the lower level can be considered as a reduced model of the integrated MILP presented in Section 4, after fixing the binary variables for tank sizing $et_{i,n}$ and $yt_{i,n}$. Therefore, in this section we only present the formulation of the upper level continuous approximation model for tank sizing.

We first formulate the continuous approximation model as an MINLP with the following objective function and constraints. After exact linearization, the model is then reformulated as an MILP, of which the formulation is presented at the end of this section.

5.1 Objective function

The objective function of this continuous approximation model is to minimize total cost, including capital investment cost, service cost and distribution cost.

Min: $Cost = capcost + servcost + distcost$ (28)

The detailed cost components are given by constraints (2) – (4) and (29).

$$distcost = \sum_{y} \frac{crot_y}{(1+wacc)^y} \quad (29)$$

Note that constraints (2) - (4) for capital and service costs are the same as those given in Section 4.1. The total distribution cost equals to the summation of discounted annual routing cost as in equation (29), where $crot_y$ is the annual routing cost calculated from the continuous approximation.

5.2 Routing cost approximation

Since the vehicle routing problem is an NP-hard problem, solving such a problem for a long time horizon (e.g. years) is a non-trivial task. As this work focuses on strategic tank sizing decisions, we can employ a continuous approximation method to estimate the optimal routing cost as a result of different tank sizing decisions. General reviews of various continuous approximation models for routing problems are given by Daganzo,[1] Langevin et al.[3] and Dasci and Verter.[2] As pointed out by the authors, mathematical programming and continuous approximations are two important approaches for routing problems. Continuous approximation models can be used to supplement mathematical programming models, and are very useful for strategic decision-making, e.g. location-routing problem[39] and strategic transportation-inventory problem.[40] In this problem, tank sizing decisions are strategic decisions made on a yearly basis. Thus, a continuous approximation model can be used to simplify the detailed routing problem, while still capturing the trade-off between capital costs and routing costs at the strategic level.

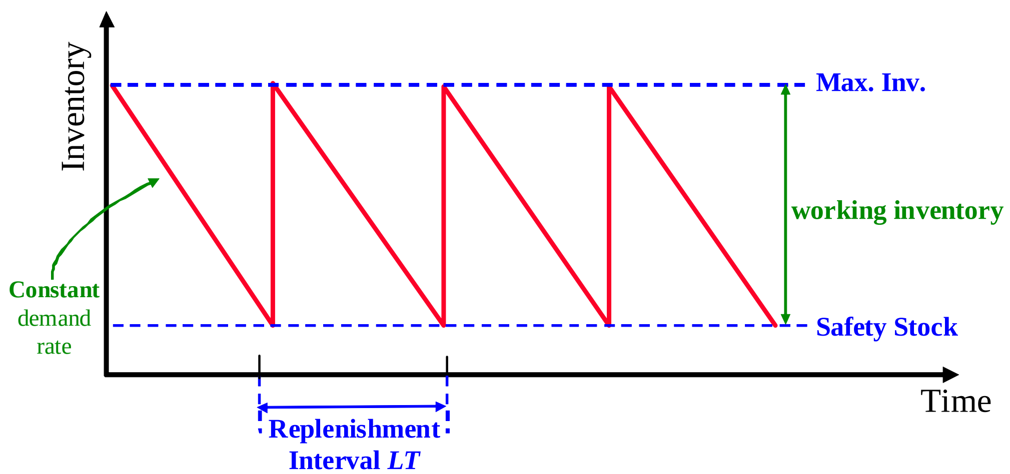


Figure 3. Inventory profile of a customer under cyclic inventory routing

In the continuous approximation model, we approximate discrete variables and parameters associated with vehicle routing using continuous functions, which represent distributions of customer locations and demands. We assume in this continuous approximation model that customers are replenished at a fixed frequency each year (i.e. cyclic inventory-routing for each year) and only one type of truck is used for delivery each year. Following these assumptions, we have the inventory profile of a customer tank for a given year as in Figure 3.[41] As we can see, the inventory level of a customer should generally lie between the lower and upper bounds, and these bounds of depend on the size of tank installed. We consider the difference between the current inventory position and the inventory lower bound as working inventory. Clearly, the larger the tank in a customer location, the larger the corresponding maximum working inventory is. From Figure 3, we can see that each time after replenishment, the working inventory level first goes up quickly and then decreases gradually due to product demand. The major decision of tank sizing is to determine what type of tank should be installed or changed in a customer. Therefore, if we could estimate the maximum working inventory of each customer, we can then determine the tank sizes. Since the customers are replenished at a fixed interval, the maximum working inventory should be the same for all the replenishments and equal to the demand rate times the replenishment interval. If the replenishment frequency were high, but the maximum working inventory level would be low at the expense of a high distribution cost, and we only need a small tank; and vice versa. With this assumption, we capture the trade-off between the routing and capital costs and consider the routing problem in a "cyclic" way. Note that a similar approach is also used for inventory-routing problems as discussed by Viswanathan and Mathur,[42] Jung and Mathur,[43] and Sindhuchao et al.[44]

As can be seen from Figure 3, the working inventory level equals to the demand rate times the replenishment interval. Thus, the required tank size should be no less than the maximum inventory level, which is the summation of the working inventory and the safety stock level. Because the demand rate of customer n in year y is given ($dem_{n,y}$), if we use x_y to denote the number of replenishment cycles in year y, then the replenishment interval or the worst case replenishment lead time of year y (LT_y) should satisfy the equation below,

$$LT_y \cdot x_y = Hz_y, \qquad \forall y \qquad (30)$$

where Hz_y is the time duration of year *y*.

Let $Trp_{n,y}$ be the total amount of product delivered from plant to customer *n* in year *y*, $ccapic_y$ be the effective truck capacity (truck capacity after accounting for product loss), rr_n be the distance between the plant and customer *n*, and *TSP* be the length of the optimal travelling salesman tour that all the customers are visited once. Haimovich and Rinnooy-Kan[45] show that the minimum routing distance for each replenishment cycle in year *y* (mrt_y) can be approximated with the following formula:

$$mrt_y \approx 2 \cdot \left(\frac{\sum_n Trp_{n,y} \cdot rr_n}{x_y \cdot ccapic_y} \right) + \left(1 - \frac{1}{ccapic_y} \right) \cdot TSP, \quad \forall y \qquad (31)$$

The detailed derivation of this formula from the original one proposed by Haimovich and Rinnooy Kan[45] is given as follows. Haimovich and Rinooy-Kan[45] proposed the following formula to determine the minimum routing distance (mrt_y) of the capacitated vehicle routing problem for a distribution system consisting of one plant and multiple customers,

$$\max\{2\frac{N}{q}\overline{r}, TSP\} \le mrt_y \le 2\left\lceil \frac{N}{q} \right\rceil \overline{r} + (1 - \frac{1}{q}) \cdot TSP \qquad (32)$$

where *N* denotes the total number of customers, *q* denotes the maximum number of customers that a truck can visit in one trip, i.e. capacity in terms of the number of customers, $\overline{r}$ denotes the *average* distance between the plant and a customer, and *TSP* is the shortest traveling salesman tour visiting each customer exactly once. The left hand side and right hand side of the above equation provide the lower and upper bounds of the minimum routing distance, respectively. Please refer to Haimovich and Kan[45] for a proof of (32).

There are a few questions that must be addressed in order to tailor this formula as a continuous approximation model for this work. The first question is how to derive an equation from the inequalities given in (32). The second question is how to measure vehicle capacity in terms of quantity, instead of the number of customers as in (32). The third question is how to incorporate customer demand information into this formula, although (32) only considers customer locations and assumes no differences among them. The last question is how to improve the accuracy of the continuous approximation formula.

To address these questions, we use a similar approach as Shen and Qi.[39] First, we take the upper bound for continuous approximation by dropping off the ceiling.

$$mrt \approx 2\frac{N}{q}\cdot \overline{r} + (1-\frac{1}{q})\cdot TSP \tag{33}$$

Computational studies by Shen and Qi[39] show that the approximation error can be bounded to 2% when the number of customers increases to more than 50. Of course, the more customers we have, the more accurate (33) will be.

In the next step, for an industrial gas supply chain with multiple customers, we "disaggregate" the customers into a number of "unit demand" customers. For example, if there is a customer with demand of 5,000L within the replenishment cycle, we disaggregate this customer into 5,000 customers, each of who has unit demand of 1L per replenishment cycle. Note that these 5,000 customers after disaggregation are still in the same location as the original customer. With this approach, the truck capacity measured by quantity is the same as the one measured by the maximum number of "unit demand" customers that can be replenished by one truck visit. In addition, the total number of "unit demand" customers is much larger than the total number of the original customers, so the accuracy of (33) can be improved.

In our problem, x_y denotes replenishment cycles in year y and $Trp_{n,y}$ is the total amount of product delivered from plant to customer n in year y. Thus, each customer n is disaggregated into $Trp_{n,y}/x_y$ "unit demand" customers. Note that the total traveling salesman tour of all the original customers (*TSP*) should be equal to the traveling salesman tour of all the "unit demand" customers, because the disaggregation process does not change the customer locations. Thus, the average distance between the plant and the customers, $\overline{r}$, can be obtained with the following formula:

$$\overline{r} = \frac{\sum_n \left(\frac{Trp_{n,y}}{x_y}\cdot rr_n\right)}{\sum_n \left(\frac{Trp_{n,y}}{x_y}\right)} = \frac{\sum_n \left(Trp_{n,y}\cdot rr_n\right)}{\sum_n Trp_{n,y}} \tag{34}$$

where rr_n is the distance between the plant and customer n. Note that after disaggregation, there are $Trp_{n,y}/x_y$ "unit demand" customers at the same location as the original customer *n*.

Therefore, substituting N in (33) with $\sum_n \left(\frac{Trp_{n,y}}{x_y} \right)$, q with $ccapic_y$, and $\bar{r}$ with (34), we can reformulate (33) to the following equation:

$$
\begin{aligned}
mrt_y &\approx 2\frac{N}{q}\cdot \bar{r} + (1-\frac{1}{q})\cdot TSP \\
&= 2\cdot \left(\frac{\sum_n Trp_{n,y}}{x_y \cdot ccapic_y} \right)\cdot \frac{\sum_n \left(Trp_{n,y}\cdot rr_n \right)}{\sum_n Trp_{n,y}} + \left(1-\frac{1}{ccapic_y} \right)\cdot TSP \\
&= 2\cdot \frac{\sum_n \left(Trp_{n,y}\cdot rr_n \right)}{x_y \cdot ccapic_y} + \left(1-\frac{1}{ccapic_y} \right)\cdot TSP
\end{aligned}
\qquad (35)
$$

which is the constraint in (31).

Shen and Qi[39] conducted computational tests for the above continuous approximation and showed that it is quite accurate if the number of customers is sufficiently large - the error of using continuous approximation is bounded by less than 2% in general cases.[46]

To reduce the nonlinearities, we introduce a new positive variable seg_y such that

$$seg_y = \frac{\sum_n Trp_{n,y}\cdot rr_n}{x_y \cdot ccapic_y}, \qquad \forall y \qquad (36)$$

Thus, the continuous approximation of the minimum routing distance for each replenishment cycle is given as follows:

$$mrt_y = 2\cdot seg_y + \left(1-\frac{1}{ccapic_y} \right)\cdot TSP, \qquad \forall y \qquad (37)$$

If we know the unit distance transportation cost of year y ($cunit_y$), then the total delivery cost of this year ($crot_y$) is the product of the unit transportation cost, the number of replenishment cycles and the minimum routing distance of each replenishment cycle.

$$crot_y = cunit_y \cdot mrt_y \cdot x_y \qquad \forall y \qquad (38)$$

Note that constraints (30) and (36) - (38) are nonlinear constraints with nonconvex terms, but they can all be exactly linearized as discussed in Section 5.6.

5.3 Tank selection and sizing constraints

In this model, we have the same tank selection and sizing constraints (7) - (9) and (16) - (18) as in Section 4.

5.4 Mass balance constraints

Let $Vm_{n,y}$ be the maximum inventory level of customer *n* in year *y*. From Figure 3, we know that the maximum inventory level should be no less than the summation of working inventory ($winv_{n,y}$), safety stock ($safety_{n,y}$) and minimum volume of the tank in customer *n* in year *y* (Vl_n) defined in constraint (16). Thus, we have the following constraint,

$$Vm_{n,y} \geq winv_{n,y} + Vl_n + safety_{n,y} \qquad \forall n, y \qquad (39)$$

The maximum inventory level should not exceed the maximum volume of the tank defined by the tank size of customer *n* in year *y* in constraint (17).

$$Vm_{n,y} \leq Vu_n \qquad \forall n, y \qquad (40)$$

Based on mass balance, the total amount of product delivered from plant to customer *n* in year *y* ($Trp_{n,y}$) is given by the following constraints,

$$Trp_{n,y} = dem_{n,y} + Vend_{n,y} - Vzero_n + Vl_n + safety_{n,y} \qquad \forall n,\ y = 1 \qquad (41)$$

$$Trp_{n,y} = dem_{n,y} + Vend_{n,y} - Vend_{n,y-1} \qquad \forall n, y \mid y \geq 2 \qquad (42)$$

$$Vend_{n,y} \leq winv_{n,y} \qquad \forall n, y \qquad (43)$$

where $Vend_{n,y}$ is the inventory level of customer *n* at the end of year *y* after adjustment for minimum tank volume and safety stocks, and it should be less than the working inventory level. Note that in the first year we need to account for the initial inventory level and adjust for minimum tank volume and safety stock.

For customer *n*, its working inventory ($winv_{n,y}$) is the replenishment that it received in a replenishment cycle. Thus, the working inventory times the number of replenishment cycles should be equal to the annual delivery amount to this customer.

$$winv_{n,y} \cdot x_y = Trp_{n,y} \qquad \forall n \qquad (44)$$

Constraint (44) is also nonconvex due to the bilinear term on the left hand side, but it can be exactly linearized by introducing additional variables and constraints. Details are discussed in Section 5.6.

5.5 Truck constraints

Following the assumption that only one type of truck is selected for delivery in each year, we have the following constraints,

$$\sum_{j} tru_{j,y} = 1 \qquad \forall y \qquad (45)$$

$$cunit_y = \sum_{j} ck_j \cdot tru_{j,y} \qquad \forall y \qquad (46)$$

$$ccapic_y = \sum_{j} tru_{j,y} \cdot Vtruck_j \cdot (1 - loss) \qquad \forall y \qquad (47)$$

Constraint (45) shows that only one type of truck is selected per year. Constraints (46) and (47) further define the unit transportation cost of a year, and the effective delivery capacity (after adjustment of loss) of truck for a year, respectively.

The lead time of a replenishment cycle should not be less than the total travel and loading time. The total traveling time is given by the minimum routing distance divided by the traveling speed and the working hours per day (*hpd*). The total loading time includes the loading times at the customer locations and at the plant. Because each customer will be visited at least once in a replenishment cycle, there will be at least $|N|$ times of loading in the customers. The loading time at the plant should be greater than the loading time of deliveries from the plant (*FT_del*). These relations are modeled through the following constraint,

$$LT_y \cdot hpd \geq \frac{mrt_y}{speed} + FT\_load \cdot |N| + FT\_del\,, \ \forall y \qquad (48)$$

where the lead time (LT_y) is measured in days, the loading times (*FT_load* and *FT_del*) are measured in hours, and the unit of traveling speed is km/hour.

5.6 MILP Reformulation

The continuous approximation model is a non-convex MINLP with the objective function given in (28) and constraints (2) – (4), (7) – (9), (16) – (18), (29) – (31) and (36) – (48). In particular, the nonlinear nonconvex terms in this model appear in constraints (30), (36) - (38) and (44). In this section, we perform exact linearizations to reformulate the MINLP model into an MILP by introducing additional variables and constraints.

First, we introduce binary variables $Ix_{k,y} \in \{0,1\}$ to represent the integer variable x_y as follows:

$$x_y = \sum_k 2^{|k|-1} \cdot Ix_{k,y} \qquad \forall y \qquad (49)$$

where $Ix_{k,y}$ determines the value of the kth digit of the binary representation of x_y. Note that the elements in set K depend on the upper bound of x_y. For example, if $x_y^U = 63$, we can set K = 1, 2, 3, 4, 5 or 6.

Further, the reciprocal of $ccapic_y$ ($Tccapic_y$) can be modeled through the following equation,

$$Tccapic_y = \sum_j \frac{tru_j}{Vtruck_j \cdot (1-loss)} \qquad \forall y \qquad (50)$$

which comes directly from equation (47).

With equations (49) and (50), we can linearize the nonlinear constraints. The bilinear term in (30) can be reformulated as $LT_y \cdot x_y = \sum_k 2^{|k|-1} \cdot LT_y \cdot Ix_{k,y}$. By introducing a nonnegative continuous variable $LTIx_{k,y} = LT_y \cdot Ix_{k,y}$, constraint (30) can be reformulated as:

$$\sum_k 2^{|k|-1} \cdot LTIx_{k,y} = Hz_y \qquad \forall y \qquad (51)$$

which is a linear constraint.

We also need the following linearization constraints to define the new variable:[47]

$$LTIx_{k,y} + LTIx1_{k,y} = LT_y \qquad \forall k, y \qquad (52.1)$$

$$LTIx_{k,y} \le LT_y^U \cdot Ix_{k,y} \qquad \forall k, y \qquad (52.2)$$

$$LTIx1_{k,y} \le LT_y^U \cdot (1 - Ix_{k,y}) \qquad \forall k, y \qquad (52.3)$$

$$LTIx_{k,y} \ge 0, \ LTIx1_{k,y} \ge 0 \qquad \forall k, y \qquad (52.4)$$

where $LTIx1_{k,y}$ is an auxiliary variable and the upper bound of LT_y is given by the time duration of year y (Hz_y).

Substituting equation (44) into the right hand side of constraint (36), we have:

$$seg_y = \frac{\sum_n winv_{n,y} \cdot x_y \cdot rr_n}{x_y \cdot ccapic_y} = \frac{\sum_n winv_{n,y} \cdot x_y \cdot rr_n}{x_y \cdot ccapic_y}, \qquad \forall y$$

Because x_y is a positive integer variable and $ccapic_y$ is also positive, the above equation implies that constraint (44) is equivalent to the following constraint:

$$seg_y \cdot ccapic_y = \sum_n winv_{n,y} \cdot rr_n \,, \qquad \forall y \tag{53}$$

Based on equation (47), we can reformulate the bilinear term in (53) as $seg_y \cdot ccapic_y = \sum_j seg_y \cdot tru_{j,y} \cdot Vtruck_j \cdot (1 - loss)$. By introducing a nonnegative continuous variable $TruSeg_{j,y} = seg_y \cdot tru_{j,y}$, constraint (53) can be reformulated as the following linear constraint:

$$\sum_j TruSeg_{j,y} \cdot Vtruck_j \cdot (1 - loss) = \sum_n winv_{n,y} \cdot rr_n \,, \qquad \forall y \tag{54}$$

We also need the following linearization constraints:

$$TruSeg_{j,y} + TruSeg1_{j,y} = Seg_y \qquad \forall j, y \tag{55.1}$$

$$TruSeg_{j,y} \le seg_y^U \cdot tru_{j,y} \qquad \forall j, y \tag{55.2}$$

$$TruSeg1_{j,y} \le seg_y^U \cdot (1 - tru_{j,y}) \qquad \forall j, y \tag{55.3}$$

$$TruSeg_{j,y} \ge 0 \,, \; TruSeg1_{j,y} \ge 0 \qquad \forall j, y \tag{55.4}$$

where $TruSeg1_{j,y}$ is an auxiliary variable and the upper bound of Seg_y is given by a sufficient large number, e.g. ($|N| \cdot \max\{rr_n\}$).

Based on equation (50), constraint (37) can be easily reformulated as the following linear constraint:

$$mrt_y = 2 \cdot seg_y + (1 - Tccapic_y) \cdot TSP \qquad \forall y \tag{56}$$

Based on equations (46) and (49), we can reformulate the tri-linear term in the right hand side of constraint (38) as $cunit_y \cdot mrt_y \cdot x_y = \sum_j \sum_k 2^{|k|-1} \cdot ck_j \cdot tru_{j,y} \cdot Ix_{k,y} \cdot mrt_y$. By introducing nonnegative continuous variables $mrIx_{k,y} = Ix_{k,y} \cdot mrt_y$ and $mrItru_{j,k,y} = tru_{j,y} \cdot mrIx_{k,y}$, constraint (38) can be reformulated as the following linear constraint:

$$crot_y = \sum_j \sum_k 2^{|k|-1} \cdot ck_j \cdot mrItru_{j,k,y} \qquad \forall y \tag{57}$$

We also need the following linearization constraints:

$$mrIx_{k,y} + mrIx1_{k,y} = mrt_y \qquad \forall k, y \tag{58.1}$$

$$mrIx_{k,y} \le mrt_y^U \cdot Ix_{k,y} \qquad \forall k, y \tag{58.2}$$

$$mrIx1_{k,y} \le mrt_y^U \cdot \left(1 - Ix_{k,y}\right) \qquad \forall k, y \qquad (58.3)$$

$$mrIx_{k,y} \ge 0,\ mrIx1_{k,y} \ge 0 \qquad \forall k, y \qquad (58.4)$$

$$mrItru_{j,k,y} + mrItru1_{j,k,y} = mrIx_{k,y} \qquad \forall j, k, y \qquad (59.1)$$

$$mrItru_{j,k,y} \le mrt_y^U \cdot tru_{j,y} \qquad \forall j, k, y \qquad (59.2)$$

$$mrItru1_{j,k,y} \le mrt_y^U \cdot \left(1 - tru_{j,y}\right) \qquad \forall j, k, y \qquad (59.3)$$

$$mrItru_{j,k,y} \ge 0,\ mrItru1_{j,k,y} \ge 0 \qquad \forall j, k, y \qquad (59.4)$$

where $mrIx1_{k,y}$ and $mrItru1_{j,k,y}$ are auxiliary variables

Based on equation (49), we can reformulate the bilinear term on the left hand side of constraint (44) as $winv_{n,y} \cdot x_y = \sum_k 2^{|k|-1} \cdot winv_{n,y} \cdot Ix_{k,y}$. By introducing a nonnegative continuous variable $wIx_{k,n.y} = winv_{n,y} \cdot Ix_{k,y}$, constraint (44) can be reformulated as the following linear constraint:

$$\sum_k 2^{|k|-1} \cdot wIx_{k,n,y} = Trp_{n,y} \qquad \forall n \qquad (60)$$

We also need the following linearization constraints:

$$wIx_{k,n.y} + wIx1_{k,n.y} = winv_{n,y} \qquad \forall j, y \qquad (61.1)$$

$$wIx_{k,n.y} \le winv_{n,y}^U \cdot Ix_{k,y} \qquad \forall j, y \qquad (61.2)$$

$$wIx1_{k,n.y} \le winv_{n,y}^U \cdot \left(1 - Ix_{k,y}\right) \qquad \forall j, y \qquad (61.3)$$

$$wIx_{k,n.y} \ge 0,\ wIx1_{k,n.y} \ge 0 \qquad \forall j, y \qquad (61.4)$$

where $wIx1_{k,n.y}$ is an auxiliary variable.

With exact linearization, we reformulate the continuous approximation model as an MILP, with the objective function (28) and constraints (2) – (4), (7) – (9), (16) – (18), (29), (39) – (43), (45) – (52), (54) – (61).

6. Case Studies

In this section, we present computational results for three examples to illustrate the application of the proposed models and the performance of the proposed solution strategies. Each example includes a number of new customers, whose tanks need to be sized. We do not consider the changes of existing customers' tanks, although this issue can be easily addressed by our computational framework. All the computational studies are performed on an IBM T400 laptop with Intel 2.53GHz CPU and 2 GB RAM. The proposed solution procedure is coded in GAMS 23.2.1.[48] The MILP problems are solved using CPLEX 12. The optimality tolerances are all set to 10^{-9}.

Case study 1: a network with two customers

The first case study is illustrative and represents a small isolated cluster of an industrial gas supply chain, with one production plant and two customers, N15 and N16. The supply chain network structure and the monthly demand rates of the first year for both customers are given in Figure 4. Note that all the data are scaled with volume unit (vu) due to confidential agreement. Other data for this case study are given in Tables 1-2. From Figure 4, it is easy to figure out that there are three possible routes in this example:

- Route 1: plant-N15-N16. Total round trip distance is 2,225km.
- Route 2: plant-N15. The total round trip distance is 2,200km.
- Route 3: plant-N16. The total distance is 2,200km.

In addition, it is easy to see that the TSP distance to visit all the customers once is 50km for this case study

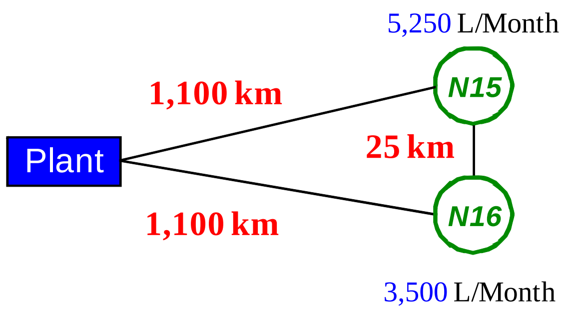


Figure 4. Case study 1 – two customer industrial gas supply chain

We consider two instances of this example. In the first instance, the planning horizon is one year, customer *N16* has an existing tank of 13,000 L and customer *N15* is a new customer whose tank should be sized. The planning horizon in the second instance is three years, and both *N15* and *N16* are new customers without existing tanks. In the three year horizon, we consider a 15% demand growth rate for both customers. Both instances are solved using three approaches: a) the integrated MILP approach as discussed in Section 4, b) the simultaneous route selection and tank sizing approach as presented in Section 5, including an upper level route selection – tank sizing model and a lower level reduced routing model with a subset of possible routes, c) the continuous approximation approach as discussed in Section 6, which includes an upper level continuous approximation model for tank sizing and a lower level detailed routing model.

Table 1 General parameters used in the models

Number of truck types	4
Number of tank sizes	6
Depreciation period (*dep*)	15 years
Time duration of year *y* (Hz_y)	365 days
Maximum number of working hours per day (*hpd*)	15 hours/day
Average truck speed in km per hr (*speed*)	22 km/hour
Minimum tanker fraction unloaded (*frac*)	10%
Product loss percentage per delivery (*loss*)	5%
Safety stock as a percentage of the tank size	15%

Table 2 Available tank sizes and the corresponding capital and service costs

Tank sizes: T_i^U (L)	Service cost in terms of percentage of capital cost ($Cser_i$/ $Ccap_i$)
1,000	26.04%
6,000	16.34%
10,000	12.95%
13,000	13.44%
16,000	12.66%
20,000	11.92%

Table 3 Optimal solution of the first instance of case study 1 (one year planning horizon, N15 is a new customer and N16 has an existing tank of 13,000L)

	Simultaneous Approach (Sec.4)	Route Selection – tank sizing Approach (Appendix)		Continuous Approximation Approach (Sec. 5)	
	Integrated MILP model	Route selection – tank sizing model	Reduced routing model	Continuous approximation model	Detailed routing model
Dis. Var.	738	636	408	34	732
Cont. Var.	1,360	852	896	134	1,358
Constraints	2,115	1,383	2,018	197	2,110
CPU (s)	69	12	15	0.1	29
Total cost	$18,128	$18,128		$18,128	
Proposed tank size for N15	13,000 L	13,000 L		13,000 L	

For the first instance, the problem sizes and computational times of all the models in the three approaches, as well as their optimal solutions are shown in Table 3. We can see that all the approaches yield the same optimal solution (minimum total cost of

$18,128, and installing a 13,000L tank for customer N15), but the CPU times are different. The simultaneous approach discussed in Section 4 solves an integrated MILP model, which has the largest size and requires the most computational time, 69s. The second approach solves an upper level route selection – tank sizing model and a lower level routing model in the reduced space of possible routes. Both models are MILP, and they have fewer variables and constraints than the integrated MILP model. Solving the two MILP models in the route selection – tank sizing approach takes less CPU times (12s and 15s) than solving the integrated model, and the upper level and lower level models require similar computational effort. In the continuous approximation approach, the upper level approximation model includes very few variables and constraints, and was solved almost instantaneously (0.1s). The lower level detailed routing model is equivalent to the integrated MILP model after fixing the integer variables for tank sizing. Thus, the detailed routing model has slightly fewer discrete variables than the integrated MILP model, and was solved much faster (29s) than the integrated MILP model (69s). Presumably, the reason is that each integer variable for tank sizing represents long-term decisions that may affect multiple integer variables for routing due to the time scale and decision hierarchy, and a pure routing can be solved very effectively once tank sizing decisions are fixed.

Since all the three approaches lead to the same optimal solution, the optimal inventory profiles of customers N15 and N16 are given in Figure 5. The inventory profiles include the information regarding the tank sizes, routing/deliveries and customer demands. We can see that the maximum inventory level corresponds to the tank of 13,000L, and the minimum inventory level is the safety stock level, which is 15% of the tank size, i.e. 1,950L. Inventory levels decrease following a constant demand rate and "jump" up once replenishments arrive. Because customers N15 and N16 have different demand rates, N15 needs 5 replenishments per year while N16 needs only 4 replenishments. Although the optimal inventory-routing decisions do not exactly follow the "cyclic" pattern as we assumed in the continuous approximation approach, the difference is relatively small. Thus, the continuous approximation approach predicts the same optimal tank sizing decisions as the simultaneous approach.

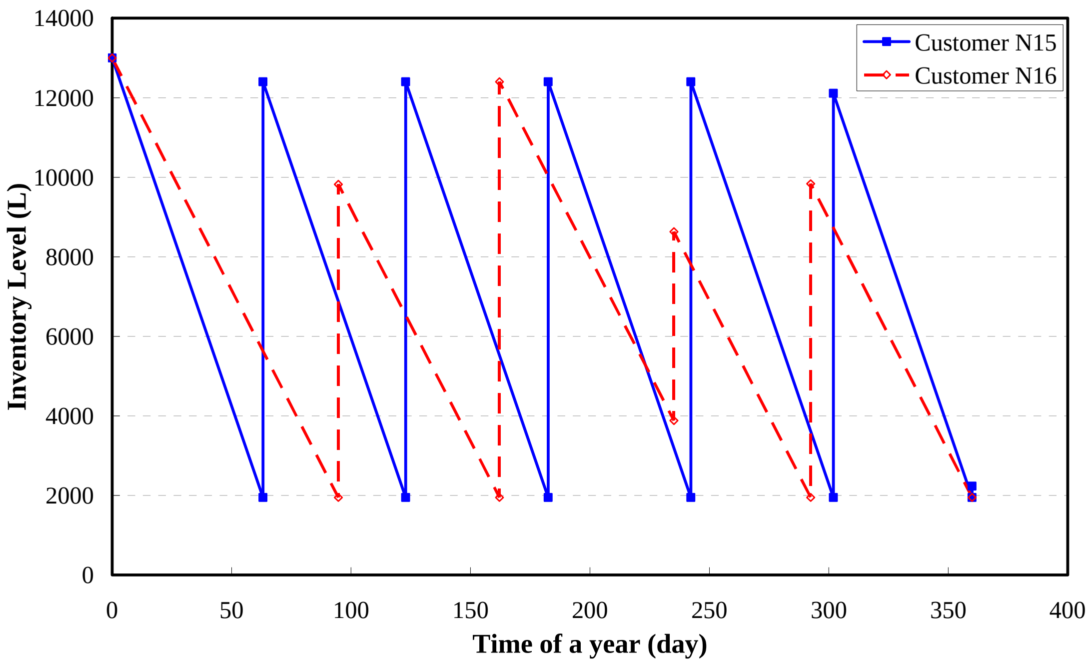


Figure 5. Optimal inventory profiles of the two customers in the first instance of case study 1 (one year planning horizon, N15 is a new customer and N16 has an existing tank of 13,000L)

Table 4 Optimal solution of the second instance of case study 1 (three year planning horizon, and both N15 and N16 are new customers)

	Simultaneous Approach	Route Selection – tank sizing Approach		Continuous Approximation Approach	
	Integrated MILP model	Route selection – tank sizing model	Reduced routing model	Continuous approximation model	Detailed routing model
Dis. Var.	2,182	1,836	1,206	54	2,170
Cont. Var.	3,996	2,452	2,666	232	3,992
Constraints	6,302	4,082	6,188	253	6,292
CPU (s)	21,159 (>memory)	128.9	35	9	59
Total cost	$53,789 * (1.23% gap)	$53,329		$53,329	
Proposed tank size for N15	13,000 L*	10,000 L		10,000 L	
Proposed tank size for N16	6,000L*	6,000 L		6,000 L	

>memory: computation was terminated due to running out of memory
***: best found solution with 1.23% gap**

In the second instance, we consider a planning horizon of 3 years, and treat both N15 and N16 as new customers without any existing tanks. The problem sizes, computational times and optimal solutions of the three approaches are given in Table 4. We can see that the problem sizes for this instance are significantly larger than the ones for the first instance, because we have a longer planning horizon and one more new customer. The large problem size makes the simultaneous approach fail to solve the problem to global optimum: CPLEX ran out of memory after around 6 hours and the best known solution has a gap of 1.23%. The simultaneous approach yields a total cost of $53,789 and the optimal tank sizes for customers N15 and N16 are 13,000L and 6,000L, respectively. Note that this is a suboptimal solution and the global optimal solution may have less total cost and different tank sizing decisions.

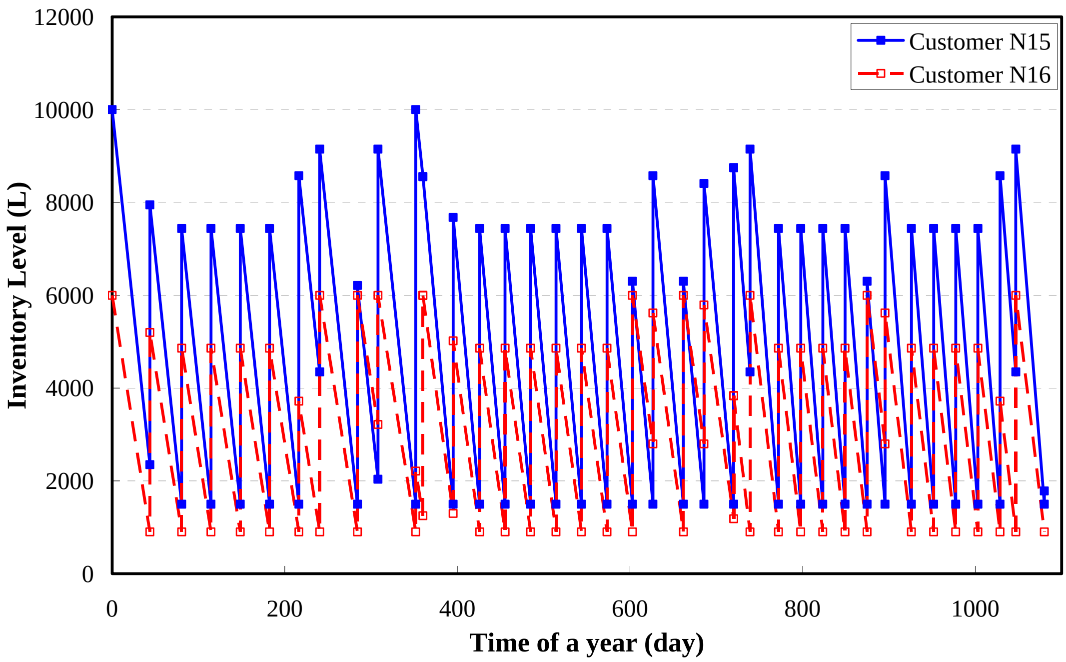


Figure 6. Optimal inventory profiles of the two customers in the second instance of case study 1 (three year planning horizon, and both N15 and N16 are new customers)

The routing selection – tank sizing approach needs 129s for solving the upper level problem and 35s for solving the reduced routing problem. Compared to the simultaneous approach, this method requires much less computational time, and the major computational effort is in solving the upper level problem for route selection and tank sizing. This approach yields a total cost of $53,329, which is lower than the suboptimal cost predicted by the simultaneous approach. The optimal tank sizing selection for customer N15 is 10,000L, which is a lower volume than the one predicted by the simultaneous approach. Since this approach has a lower total cost, sizing a 10,000L tank to customer N15 might be a better decision. The continuous approximation approach leads to the same optimal solution as the route selection – tank sizing approach, but requires slightly less CPU times. It took only 9s for the upper level continuous approximation model, but the lower level detailed routing model requires 59s, which is longer than the one for the reduced routing model in the second approach. Note that the reduced routing model only considers those routes selected by its upper level problem, and thus has fewer variables and constraints and requires less CPU time. The optimal inventory profiles of the customers predicted by the continuous approximation approach are given in Figure 6, where we can see the tradeoffs between tank sizes, deliveries and customer demands. The two customers have different maximum and minimum inventory levels due to their different tank sizes. Although they have different demand rates, it turns out that both customers have 37 replenishments during the three-year planning horizon.

Case study 2: a network with four customers

In the second case study we consider a four-customer industrial gas cluster, of which the network structure and the demand rates of the first year are given in Figure 7. We also use the data provided in Tables 1-2 for this case study. Based on the network structure in Figure 7, there are 15 possible routes for this case study. The set of possible routes and the total round trip distance for each route are listed in details in Part II.[49] From the network structure, it is easy to see that the TSP distance to visit all the customers once is 4507.47km for this case study.

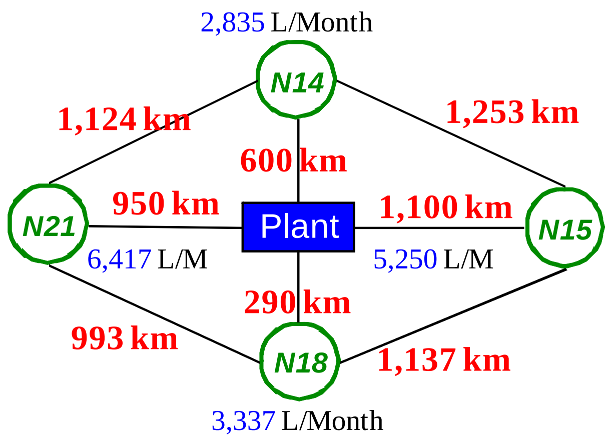


Figure 7. Case study 2 – four customer industrial gas supply chain

Two instances are considered for this case study. The first instance has a planning horizon of one year, and considers N14 as a new customer and each of the other three customers has an existing tank of 16,000L. Similar to the previous case study, we solve this instance with the three approaches. The computational results are listed in

Table 5. Since the number of customers increases, the problem size of the integrated MILP also increases. Before running out of memory, the best found solution of the simultaneous approach has an optimality gap of 14.75%. The suboptimal solution indicates a total cost of $37,559 and a tank size of 13,000L for customer N21. With the route selection – tank sizing approach and continuous approximation approach, we obtained a better solution with a lower total cost of $33,824. Thus, 10,000L, which is the tank size for customer N14 determined by the last two approaches, should be a better selection for this instance. The detailed inventory profiles as predicted by the last two approaches for the four customers are given in Figure 8. Due to different locations, tank sizes and demand rates, the four customers have different numbers of replenishments in the one year planning horizon. Note that the maximum and minimum inventory levels for customer N14 is 10,000L and 1,500L, respectively, while for the other three customers they are 16,000L and 2,400L, respectively. This in turn reveals the trade-offs between the strategic tank sizing decisions and the operational routing decisions. We note that the number of replenishment cycles may be different from the number of replenishments of a customer, because a customer can receive more than one replenishment in a replenishment cycle. The number of replenishment cycles is determined from the continuous approximation model, and the detailed number of replenishments of each customer comes from the solution of the detailed routing model.

Table 5 Optimal solution of the first instance of case study 2 (one year planning horizon, N14 is a new customer, N15, N18 and N21 all have an existing tank of 16,000L)

	Simultaneous Approach	Route Selection – tank sizing Approach		Continuous Approximation Approach	
	Integrated MILP model	Route selection – tank sizing model	Reduced routing model	Continuous approximation model	Detailed routing model
Dis. Var.	3,630	3,108	2,648	58	3,618
Cont. Var.	6,190	4,760	5,220	232	6,194
Constraints	7,965	8,676	6,804	705	7,960
CPU (s)	16,390 (>memory)	9,492	4,113	2	6,420
Total cost	$37,559* (14.75% gap)	$33,824		$33,824	
Proposed tank size for N14	13,000L*	10,000 L		10,000 L	

>memory: computation was terminated due to running out of memory
***: best found solution with 14.75% gap**

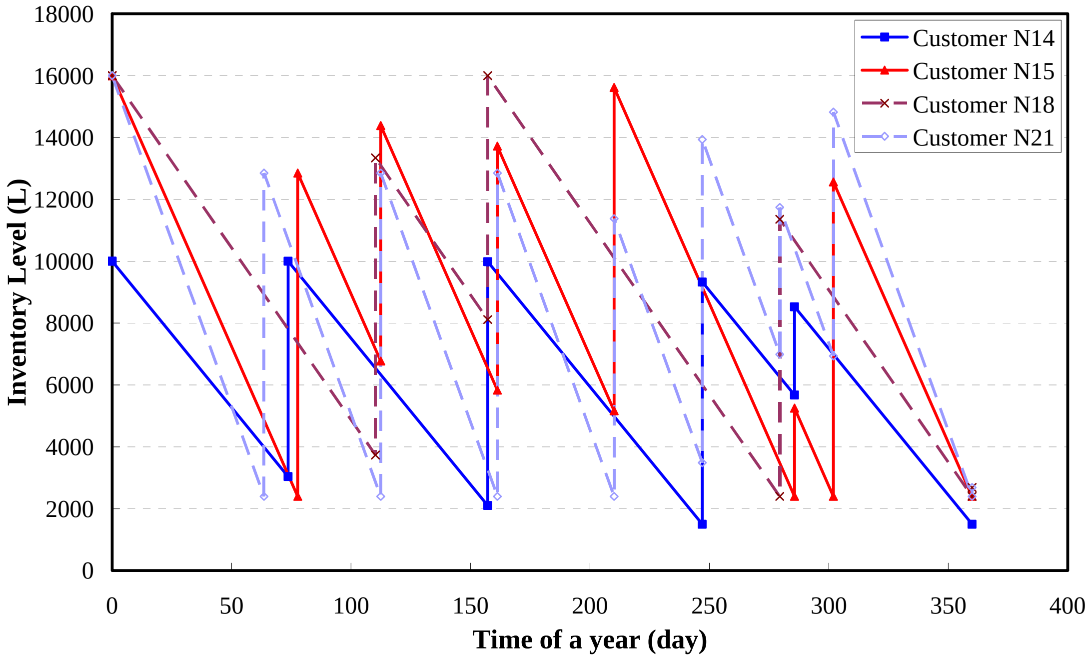


Figure 8. Optimal inventory profiles of the four customers in the first instance of case study 2 (one year planning horizon, N14 is a new customers, N15, N18 and N21 all have an existing tank of 16,000L)

Table 6 Optimal solution of the second instance of case study 2 (three year planning horizon, and N14, N15, N18 and N21 are all new customers)

	Simultaneous Approach (Sec.4)	Route Selection – tank sizing Approach (Appendix)		Continuous Approximation Approach (Sec. 5)	
	Integrated MILP model	Route selection – tank sizing model	Reduced routing model	Continuous approximation model	Detailed routing model
Dis. Var.	10,848	9,108	N/A	78	10,800
Cont. Var.	18,440	13,960	N/A	576	18,428
Constraints	23,834	25,776	N/A	726	23,814
CPU (s)	51,406 (>memory)	54,718 (>memory)	N/A	5.3	31,420
Total cost	$156,774 * (74.54% gap)	N/A		$101,402	
Proposed tank size for N14	16,000L*	N/A		10,000 L	
Proposed tank size for N15	20,000L*	N/A		16,000 L	
Proposed tank size for N18	10,000L*	N/A		10,000 L	
Proposed tank size for N21	20,000L*	N/A		20,000 L	

>memory: computation was terminated due to running out of memory
***: best found solution with 74.54% gap**

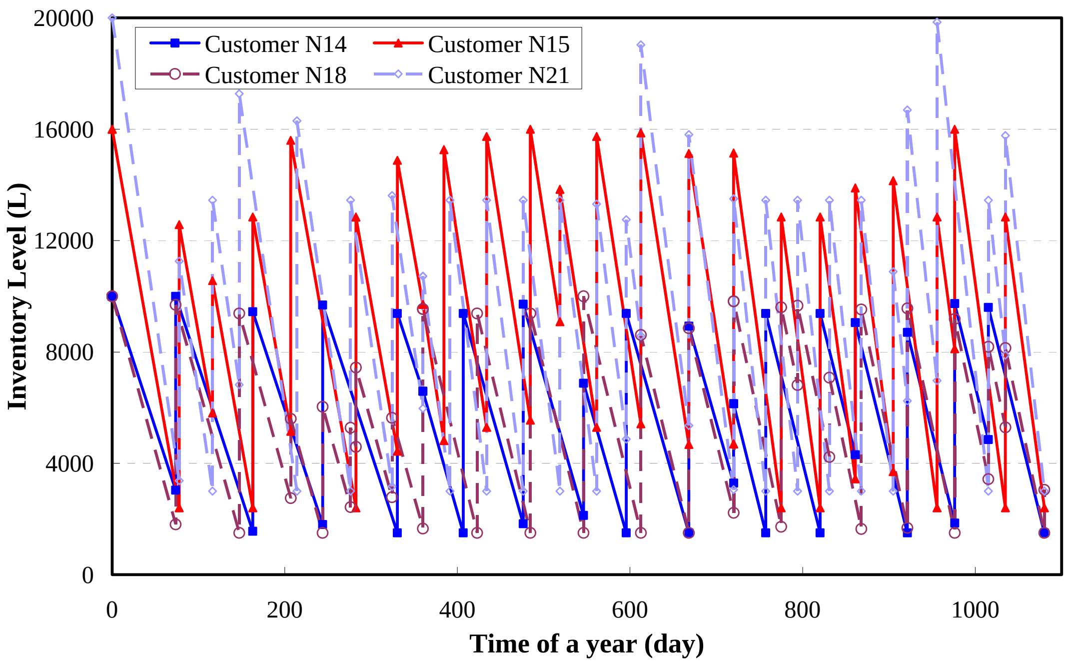


Figure 9. Optimal inventory profiles of the four customers in the second instance of case study 2 (three year planning horizon, and N14, N15, N18 and N21 are all new customers)

In the second instance of case study 2, we consider a three-year planning horizon and treat all the customers as new customers without any existing tanks. In the three year horizon, we consider a 15% demand growth rate for all customers. The computational results of solving this instance with three approaches are given in Table 6. We can clearly see that the problem sizes increase significantly and some problems include more than 10,000 binary variables. The simultaneous approach ran out of memory after around 4 hours, and the best found solution ($156,774) has a gap as large as 74.54%. Because there are 4 customers and 15 possible routes in this instance, the problem size of the route selection – tank sizing model also becomes computationally intractable when considering the three-year planning horizon. Because the route selection – tank sizing problem ran out of memory, we were not able to obtain the selected routes and solve the reduced routing model in the lower level. With the continuous approximation approach, the problem size of the upper level approximation model is still rather small and can be solved very efficiently (only 5.3s for the global optimum), although this instance is relatively large. The detailed routing model, despite its large size, was solved to global optimality in about 9 hours. The solution predicted by this approach has a lower optimal total cost of $101,402, and the optimal tank sizes for customers N14, N15, N18 and N21 are 10,000L, 16,000L, 10,000L and 20,000L, respectively. The detailed inventory profiles are given in Figure 9, where we can see similar trade-offs between tank sizes, demand rates and deliveries.

Case study 3: large scale instances with 30, 60, 100 and 200 customers

In the last case study, we consider four large-scale industrial gas supply chains with 30, 60, 100 and 200 customers, respectively. In all these four instances, a 3-year planning horizon is considered and all the customers are treated as new customers without any existing tanks. As we can see from case study 2, the simultaneous approach and the route selection – tank sizing approach can be computationally intractable for such a large scale instance. Thus, we only use the continuous approximation approach for this case study.

The data provided in Tables 1-2 are used for the four instances in this case study. Due to the large number of customers, we generate randomly their locations and demand rates. All the customer locations are generated in a $400\text{km} \times 400\text{km}$ square following uniform distribution, and the plant is located in the center of this square. The detailed locations of the customers and plant for these four instances are given in Figures 10(a) – 10(d). The TSP distances to visit all the customers once (not including the plant) for different scenarios and years are obtained with Concorde TSP Solver[50] through its NEOS interface[51] with CPLEX 12. The resulting TSP distances for the four instances are 1,600km, 2,402km, 3,005km and 4,335km, respectively. We note that the Concorde TSP solver is computationally very effective – for a 200 customer case that will be solved later, it took less than 2 seconds to obtain the global optimal solution for the TSP values.

The monthly demand rates of customers in the first year ($demc_{n,y}$, L/month) are generated using normal distributions as follows:

For the 30 customer instance: $demc_{n,y} = 100 + 100 \times \left| N[0,\ 40] \right|$

For the 60 customer instance: $demc_{n,y} = 100 + 100 \times \left| N[0,\ 30] \right|$

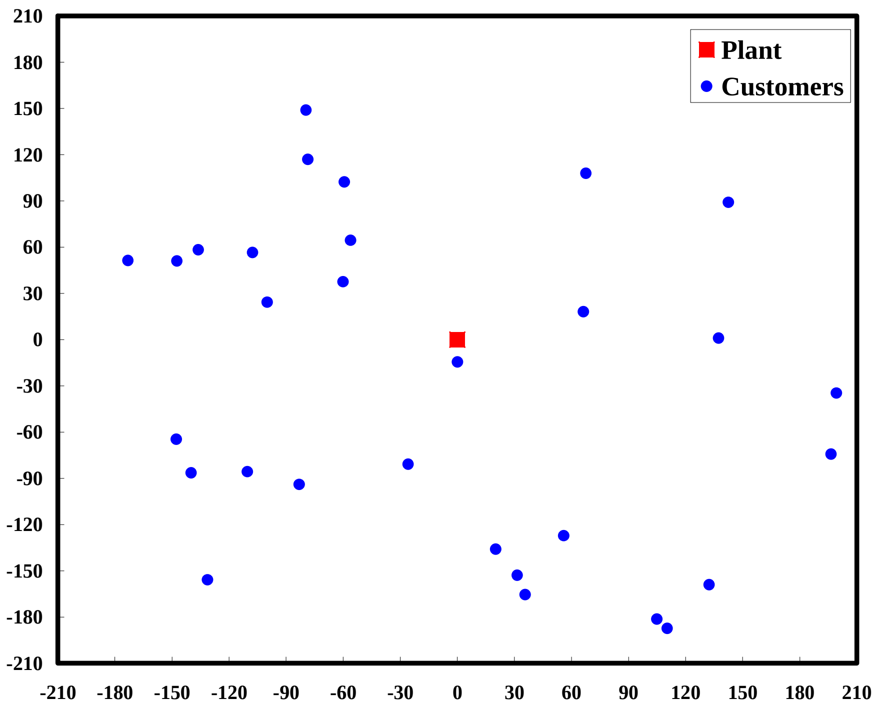
For the 100 customer instance: $demc_{n,y} = 100 + 100 \times \left| N[0,\ 15] \right|$

For the 200 customer instance: $demc_{n,y} = 100 + 100 \times \left| N[0,\ 5] \right|$

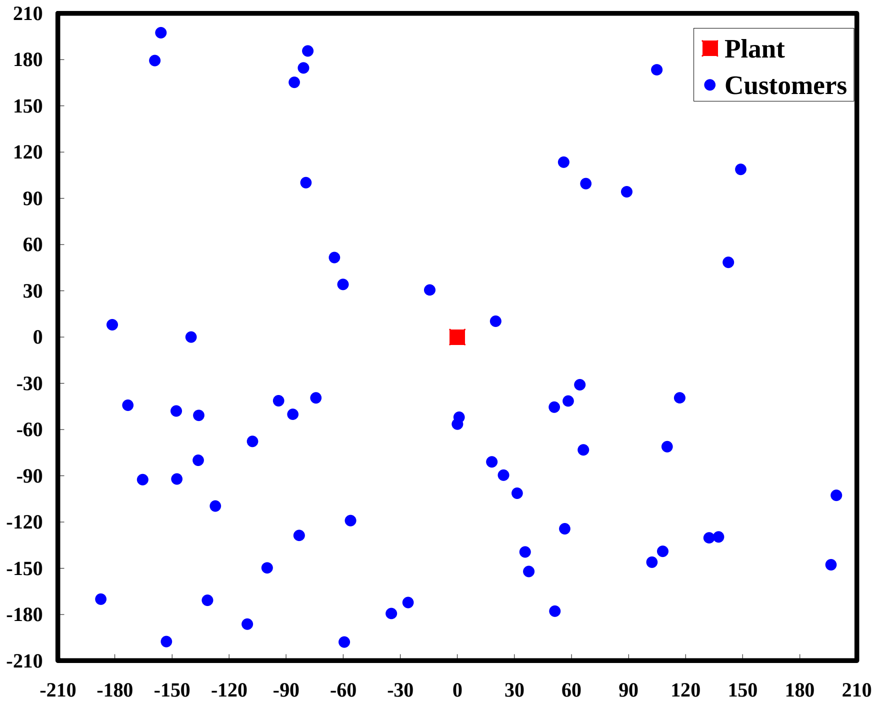
Note that we take the absolute values of the normal distribution so that the monthly demand rates are always higher than 100L/month. Although the normal distribution is unbounded, the maximum monthly demand rate we obtained from the sampling is 16,966.61L/month. In the three year horizon, we consider a 15% demand growth rate for all customers.

Table 7   Optimal solution of case study 3 using continuous approximation

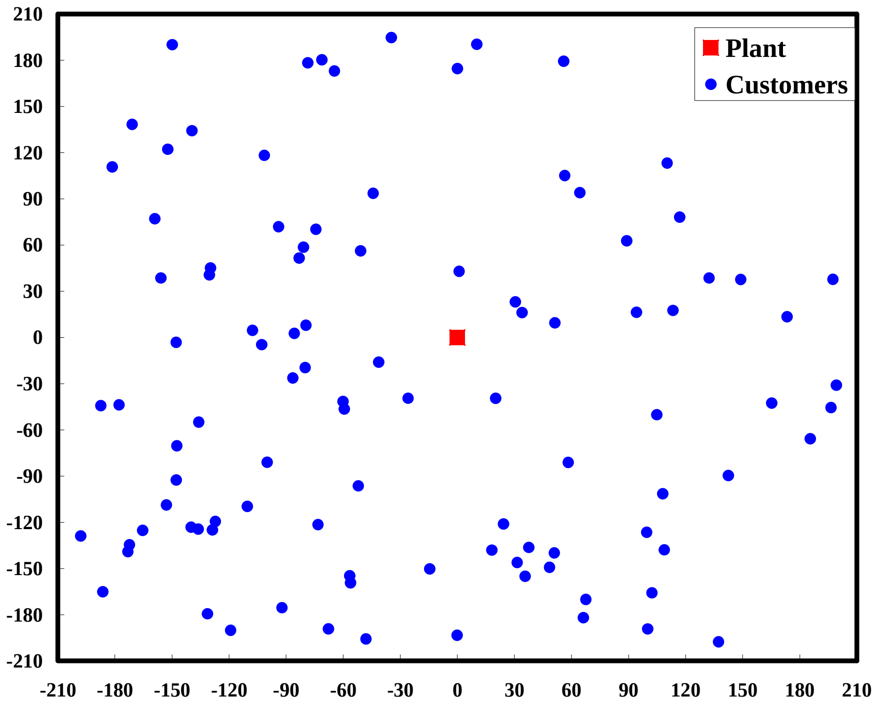
	Dis. Var.	Cont. Var.	Constraints	CPU (s)	Total cost ($)
30 customer instance	1,260	1,062	2,944	30.5	337,195.75
60 customer instance	4,290	3,862	5,494	21.3	721,413.51
100 customer instance	11,130	6,262	8,894	350.8	965,453.32
200 customer instance	42,230	12,262	17,394	299.4	1,664,093.11



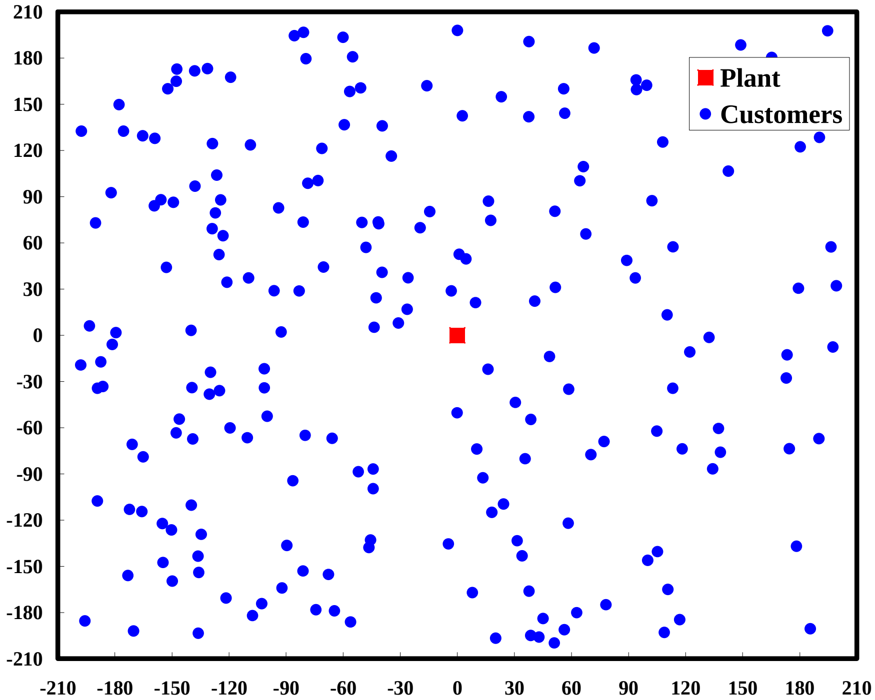
10(a) Location map of the 30 customers



10(b) Location map of the 60 customers



10(c) Location map of the 100 customers



10(d) Location map of the 200 customers

Figure 10. Case study 3 – industrial gas supply chains consisting of 30, 60, 100 and 200 customers (three year planning horizon, and all the customers are new customers without any existing tank)

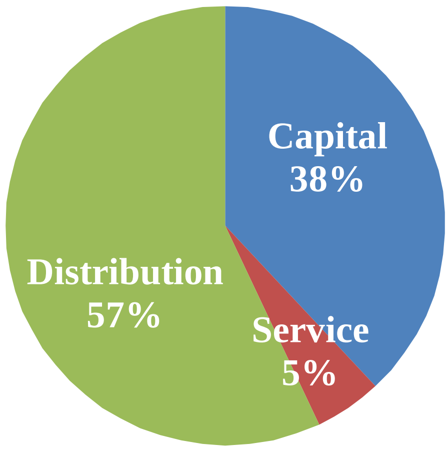


Figure 11. The breakdown of the total cost for the 200 customer case

In Table 7, we report the problem sizes, computational times and optimal solutions of the continuous approximation model for the four instances. Note that the computational costs for solving the TSP problems are not included in the CPU times reported in this Table. Although the problem sizes increase exponentially as the number of customers increases, we can still solve the 200 customer instance, which includes 42,230 binary variables, 12,262 continuous variables and 17,394 constraints, in less than 5 minutes. The results clearly show that the continuous approximation model has very high computational efficiency and is capable of handling the strategic tank sizing problems for large-scale industrial gas supply chains. The breakdown of the total cost for the 200 customer instance is given in Figure 11. We can see that the total distribution cost is close to the capital cost, which again reveals the tradeoff between vehicle routing and tank sizing.

In this case study, we do not solve the detailed routing model due to the large problem size. Although solving an integrated MILP for a 200-customer routing problem is a nontrivial task, there are many existing heuristics and decomposition methods that can help to obtain a “good” near-optimal solution for the pure routing problem within reasonable computational time. One possible approach is to employ an integrated clustering method and location-based heuristics to group the customers into a number of small clusters and solve the routing problem within each cluster independently. By iteratively changing the customers in the clusters, we can obtain a near-optimal solution within the required computational time. The details of this method will be introduced in the second part of this paper.[49] The key point is that once we can determine the strategic tank sizing decisions for large-scale industrial gas supply chains with the proposed approaches (e.g. continuous approximation method), the lower level detailed routing problem is very similar to the many vehicle routing problems that have been well studied in the past decades.

This case study illustrates the application of the proposed continuous approximation method and the effectiveness of this approach for large-scale problems. After all, solving an integrated MILP with the simultaneous approach for the tank sizing decisions of a 200 customer industrial gas supply chain is most likely beyond the capability of the current state-of-the-art computational architecture and software.

7. Conclusion

In this paper, we have proposed an MILP model to simultaneously optimize the tank sizing and vehicle routing decisions in the distribution-inventory planning of industrial gas supply chains. To effectively integrate the strategic and operational decisions and to handle long planning horizon, we have also proposed two computational strategies. The first approach includes an upper level route selection - tank sizing problem and a lower level reduced routing problem. The upper level problem predicts the optimal tank sizing decisions and the routes for deliveries, and the lower level problem is then solved with fixed tank sizes and those selected routes for higher computational efficiency. The second approach is based on a continuous approximation method, which can provide a rough estimation of the total distribution cost without considering detailed routing decisions. We have proposed an MINLP model in the upper level based on continuous approximation to determine the tank sizing decisions. The model is then reformulated as an MILP, which can be globally optimized very effectively even for large-scale instances. The lower level of this approach is to solve the detailed routing problem in the reduced variable space after fixing the tank sizing decisions. These models and computational strategies were applied to three case studies for industrial gas supply chains with up to 200 customers. The results clearly show that the proposed computational strategies, especially the continuous approximation approach, are very effective for solving the distribution-inventory planning problem of large-scale industrial gas supply chains

Acknowledgment

The authors acknowledge the financial support from Praxair, Inc., Pennsylvania Infrastructure Technology Alliance (PITA) and National Science Foundation under Grant No. DMI-0556090 and No. OCI-0750826. Elisabet Capón also wishes to thank the Spanish Ministerio de Educación y Ciencia for the FPU program. Fengqi You is partially supported by the U.S. Department of Energy under contract DE-AC02-06CH11357.

Appendix

Simultaneous Route Selection and Tank Sizing Approach

The simultaneous route selection and tank sizing approach consists of two steps. First, the tank sizes and the potential routes to be used in the second step are selected by solving an aggregated MILP model. The MILP model accounts for the “worst case” working inventory, and simultaneously predicts the optimal routes to supply the different customers and the optimal tank sizes for satisfying the specified demand. In the second step, a vehicle routing problem is solved in the reduced space with the selected routes and the installed tank capacities.

The detailed vehicle routing model in the second step can be considered as a reduced model of the integrated MILP presented in Section 4, after limiting the elements in the set $r \in R$ to the selected routes and fixing the binary variables for tank sizing $et_{i,n}$ and $yt_{i,n}$. Therefore, in this section we only present the formulation of the simultaneous route selection and tank sizing model that is solved in the first step.

The simultaneous route selection and tank sizing model aims at determining the optimal routes and tank sizes so as to minimize the total cost, regardless of the synergies of truck deliveries. The main assumption in this model is that the tank size should be greater than the sum of the minimum inventory, safety stocks and the maximum delivered amount among all the replenishments, i.e. the "worst case" working inventory. Additionally, the set $d \in D$ is introduced for considering the deliveries, that is, the occasions in which a given route is covered in a year. The detailed formulation of this aggregated model is given as follows.

A.1 Objective function

The objective function of simultaneous route selection and tank sizing model as given in (28) is to minimize the summation of the total capital investment, service and distribution costs. Note that the outage cost is neglected in this model.

Min: $Cost = capcost + servcost + distcost$ (A1)

The detailed cost components are given in constraints (2) – (4) and (A2).

$$distcost = \sum_j \sum_r \sum_d \sum_y \frac{dist_r \cdot ck_j \cdot Dz_{j,r,d,y}}{(1+wacc)^y} \quad \text{(A2)}$$

Note that constraints (2) - (4) are the same as those given in Section 4.1.

Capital and service costs have the same formulation as in Section 4. The total distribution cost represents an aggregated formulation as given in (A2), where $Dz_{j,r,d,y}$ is a binary variable that equals to 1 if truck *j* is used for route *r* in trip *d* in year *y*. Therefore, the sum over the set *d* of $Dz_{j,r,d,y}$ represents the number of times that route *r* is covered with a given truck *j* in year *y*.

A.2 Tank selection constraints

In this model, we have the same tank selection constraints (7) – (9) as in Section

4.2.

A.3Truck delivery constraints

Constraint (A3) defines that at most one truck can be assigned to a selected route. If a given route is selected ($Dzz_{j,r,y}$ = 1), then it can be replenished by the selected truck as stated in Constraint (A4),

$$\sum_j Dzz_{j,r,y} \leq 1 \qquad \forall r, y \qquad \text{(A3)}$$

$$Dz_{j,r,d,y} \leq Dzz_{j,r,y} \qquad \forall j, d, y,\ r \in R_n \qquad \text{(A4)}$$

where $Dzz_{j,r,y}$ is a binary variable that equals to 1 if truck *j* is used for delivery with route *r* in year *y*, and the binary variable $Dz_{j,r,d,y}$ allows to account for the trips using a given route *r* is covered with truck *j* in year *y*.

A new continuous variable $Dp_{n,r,y}$ is introduced to represent the replenishment amount of a trip to customer *n* with route *r* in year *y*. This variable is defined through the following constraint,

$$Dpr_{n,r,d,y} \leq Dp_{n,r,y} \quad \forall n \in N_r, r \in R_n, d \in D, y \in Y \qquad \text{(A5)}$$

where $Dpr_{n,r,d,y}$ is the replenishment amount to customer *n* with route *r* in trip *d* of year *y*.

Constraint (A6) enforces that the quantity delivered to a given customer ($Dpr_{n,r,d,y}$) has to be the same as ($Dp_{n,r,y}$) in every replenishment to the customer through route r. Additionally, constraint (A7) imposes the amount to be replenished ($Dpr_{n,r,d,y}$) is different from 0, if the binary variable $Dz_{j,r,d,y}$ is 1.

$$Dp_{n,r,y} - \max_j \left\{Vtruck_j\right\} \cdot \left(1 - \sum_j Dz_{j,r,d,y}\right) \leq Dpr_{n,r,d,y} \quad \forall n \in N_r, r \in R_n, d \in D, y \in Y \quad \text{(A6)}$$

$$Dpr_{n,r,d,y} \leq \max_j \left\{Vtruck_j\right\} \cdot \sum_j Dz_{j,r,d,y} \qquad \forall n \in N_r, r \in R_n, d \in D, y \in Y \quad \text{(A7)}$$

Constraint (A8) ensures that the total amount received by all the customers served by a given route *r* cannot exceed the capacity of the truck *j* that delivers through that route. In addition, a minimum quantity of the truck must be delivered to the customers included in route r, which is imposed as a fraction of the total truck capacity through constraint (A9).

$$\sum_{n \in NC_{r,n}} Dp_{n,r,y} \leq \sum_j Dzz_{j,r,y} \cdot vtru_j \cdot \left(1 - loss \cdot \left|N_r\right|\right) \qquad \forall r \in R_n, y \in Y \qquad \text{(A8)}$$

$$Dp_{n,r,y} \geq fraction \cdot \sum_j Dzz_{j,r,y} \cdot vtru_j \cdot \left(1 - loss \cdot \left|N_r\right|\right) \quad \forall n, r \in R_n, y \in Y \qquad \text{(A9)}$$

The minimum number of visits to a customer can be estimated as a function of its demand and the capacity of the largest truck as in constraint (A10),

$$\sum_j \sum_{r \in R_n} \sum_d Dz_{j,r,d,y} \geq \frac{dem_{n,y}}{\max_j \left\{Vtruck_j\right\}} \quad \forall n, y \qquad \text{(A10)}$$

The above equation represents a lower bound of the overall number of trips that must be done to a customer *n* in a given year *y*, considering that the largest truck supplies all routes, and that the whole capacity is delivered to that customer *n*.

A.4Mass balance constraints

Constraint (A11) ensures that the total amount received by a customer plus its initial inventory must satisfy the demand for that year plus the minimum level of the tank and the safety stock.

$$\sum_{r \in R_n} \sum_d Dpr_{n,r,d,y} \geq dem_{n,y} Hz_y - Vzero_n + Vl_{\mathrm{n}} + safety_{n,y} \quad \forall n, y \qquad \text{(A11)}$$

In addition, the maximum level of the tank installed in a customer location cannot be exceeded by any delivery to that customer in any of the routes, as stated in constraint (A12). The existing level is calculated as the sum of the lower level plus the delivered quantity and the safety stock.

$$vu_n \geq vl_n + Dp_{n,r,y} + safety_{n,y} \quad \forall n \in N,\ r \in R_n,\ y \in Y \qquad \text{(A12)}$$

Note that in this aggregated model, we neglect the detailed timing issues of the deliveries, and constraints (10)-(14) of the integrated model are aggregated into constraints (A11) and (A12).

The constraints for modeling the minimum, maximum and initial inventory level of customer *n* are the same as Constraints (16) – (18) given in Section 4.3.

The following constraint avoids the degeneracy in the set *D* so that the convergence to the solution is faster. It imposes that the first elements of the set *D* are assigned first.

$$\sum_j Dz_{j,r,d,y} \leq \sum_j Dz_{j,r,d-1,y} \quad \forall r \in R_n,\ y \in Y \qquad \text{(A13)}$$

A.5 Computational complexity

The simultaneous route selection and tank sizing problem corresponds to an MILP model with the objective function in (A1) and constraints (2)-(4), (7)-(9), (16)-(18) and (A2)-(A13). By approximating the maximum inventory levels and underestimating the distance costs, this first step selects the delivery routes without going into the details of the timing. A reduced vehicle routing problem is then solved by fixing tank sizing and route selection decisions. This approach can significantly reduce the computational time of the routing problem by considering a subset of all the possible routes. Since the upper level aggregated model has potentially smaller size than the original integrated model, it should be computationally more efficient than the integrated MILP model presented in Section 4. However, as the problem size increases, solving the aggregated model can also become intractable due to the combinatorial complexity of route enumeration.

Nomenclature

Sets/Indices

d	set of trips
i	set of tank sizes
j	set of truck sizes
k	set for binary representation of integers
n	set of customers
r	set of all possible routes
t	set of events
y	set of years

Subsets

$n \in N_r$	subset of customers that are served by route r
$r \in R_n$	subset of routes used by customer n

Parameters

$Ccap_i$	capital cost of tank of size i
$Cout_{n,y}$	outage cost for customer n in year y
$Cser_i$	service cost of tank of size i

ck_j	delivery cost per distance traveled of truck *j*
ca_r	1 if route r is considered for detailed routing
$dem_{n,y}$	demand of customer *n* in year *y*
$demc_{n,y}$	monthly demand rate of customer *n* in year *y*
dep	depreciation period in years
$dist_r$	distance of route *r*
$espace_n$	1 if there is extra space for installing another tank at customer *n*
frac	minimum tanker fraction unloaded
FT_load	loading time for each customer
FT_del	loading time for each delivery from the plant
hpd	maximum number of working hours per day
Hz_y	time duration of year *y*
loss	product loss percentage per delivery
new_n	1 if customer *n* is new
nt_j	number of tankers of size *j*
$ot_{i,n}$	1 if tank size *i* originally installed at customer *n*
rr_n	distance between the plant and customer *n*
$safety_{n,y}$	safety level in year *y* for customer *n*
speed	average truck traveling speed in km per hr
T_i^L	lower bound for discrete tank size *i*
T_i^U	upper bound for discrete size of tank *i*
tm	maximum number of slots
$tsize_n$	1 if tank of customer *n* is sized
TSP	traveling sales man distance of all the customers (exclude the plant)
$Vtruck_j$	full transportation capacity of truck *j*
$Vzero_n$	initial volume at customer *n*
wacc	working capital discount factor

Binary Variables (0-1)

$et_{i,n}$	1 if customer *n* has tank of size *i* installed in extra space; 0 otherwise
$yt_{i,n}$	1 if customer *n* has tank of size *i* installed; 0 otherwise
$z_{j,r,t,y}$	1 if truck *j* is used for delivery with route *r* in time event *t* of year *y*
$Dz_{j,r,d,y}$	1 if truck *j* is used for route *r* in trip *d* in year *y*

$Dzz_{j,r,y}$	1 if truck j is used for delivery with route r in year y
$tru_{,j,y}$	1 if truck j is selected for replenishmsnt in year y
$Ix_{k,y}$	0-1 variable for the binary representation of the number of replenishments in year y (x_y)

Continuous Variables (0 to $+\infty$)

$Cost$	total cost
$capcost$	capital cost
$ccapic_y$	effective capacity of truck for the replenishments in year y
$cunit_y$	unit transportation cost of year y
$crot_y$	approximated routing cost of year y
$distcost$	distribution cost
$Dp_{n,r,y}$	maximum single delivery amount to customer n with route r in year y
$Dpr_{n,r,d,y}$	replenishment amount to customer n with route r in trip d of year y
LT_y	worst case replenishment lead time of year y
mrt_y	minimum routing distance to visit all the customers once in year y
$out_{n,t,y}$	outage for customer n in time t in year y
$outcost$	outage cost
$p_{n,t,y}$	delivery to customer n in time event t of year y
$pr_{n,r,t,y}$	delivery to customer n in route r at time event t of year y
seg_y	auxiliary variable, for groups of all the customers
$servcost$	service cost
$Tccapic_y$	reciprocal of $ccapic_y$
$ti_{t,y}$	initial time in time event t of year y
$\Delta t_{t,y}$	time interval in time event t of year y
$Trp_{n,y}$	total replenishment amount from the plant to customer n in year y
$Vend_{n,y}$	inventory level of customer n at the end of year y
Vl_n, Vu_n	minimum and maximum volume of tank at customer n
$Vo_{n,t,y}$	volume at customer n in time event t of year y
$Vm_{n,y}$	maximum inventory level of customer n in year y
$winv_{n,y}$	maximum working inventory of customer n in year y
x_y	number of replenishment in year y in scenario s

Auxiliary Variables (0 to $+\infty$)

$tins_{i,n}$	auxiliary variable
$LTIx_{k,y}$	auxiliary variable for the product of $Ix_{k,y}$ and LT_y
$LTIx1_{k,y}$	auxiliary variable for linearization
$TruSeg_{j,y}$	auxiliary variable for the product of $tru_{,j,y}$ and seg_y
$TruSeg1_{j,y}$	auxiliary variable for linearization
$mrIx_{k,y}$	auxiliary variable for the product of $Ix_{k,y}$ and mrt_y
$mrIx1_{k,y}$	auxiliary variable for linearization
$mrItru_{j,k,y}$	auxiliary variable for the product of $tru_{,j,y}$ and $mrIx_{k,y}$
$mrItru1_{j,k,y}$	auxiliary variable for linearization
$wIx_{n,k,y}$	auxiliary variable for the product of $Ix_{k,y}$ and $winv_{n,y}$
$wIx1_{n,k,y}$	auxiliary variable for linearization

References

1. Daganzo, C., Logistics Systems Analysis. In *Lecture Notes in Economics and Mathematical Systems*, Beckmann, M.; Krelle, W., Eds. Springer Verlag: Berlin, 1996.
2. Dasci, A.; Verter, V., A continuous model for production-distribution system design. *European Journal of Operational Research* **2001,** 129, 287-298.
3. Langevin, A.; Mbaraga, P.; Campbell, J. F., Continuous Approximation Models for Freight Distribution: An Overview. *Transportation Research B* **1996,** 30, 163-188.
4. Glankwamdeea, W.; Linderoth, J.; Shen, J.; Connard, P.; Hutton, J., Combining optimization and simulation for strategic and operational industrial gas production and distribution. *Computers & Chemical Engineering* **2009,** 32, 2536-2546.
5. Choi, J.; Lee, J. H.; Realff, M. J., An algorithmic framework for improving heuristic solutions Part II. A new version of the stochastic traveling salesman problem. *Computers & Chemical Engineering* **2004,** 28, 1297-1307.
6. Jetlund, A. S.; Karimi, I. A., Improving the logistics of multi-compartment chemical tankers. *Computers & Chemical Engineering* **2004,** 28, 1267-1283.
7. Huang, W.; Chung, P. W. H., Integrating routing and scheduling for pipeless plants in different layouts. *Computers & Chemical Engineering* **2005,** 29, 1069-1081.
8. Dondo, R.; Mendez, C. A.; Cerda, J., Optimal management of logistic activities in multi-site environments. *Computers & Chemical Engineering* **2008,** 32, 2547-2569.
9. Dondo, R., G.; Cerdá, J., A hybrid local improvement algorithm for large-scale multi- depot vehicle routing problems with time windows. *Computers & Chemical Engineering* **2009,** 33, 513-530.
10. Dondo, R.; Mendez, C. A.; Cerda, J., Managing Distribution in Supply Chain Networks. *Industrial & Engineering Chemistry Research* **2009,** 48, 9961–9978.
11. Laporte, G., What You Should Know about the Vehicle Routing Problem. *Naval Research Logistics* **2007,** 54, 811-819.

12. Baita, F.; Ukovich, W.; Pesentib, R.; Favaretto, D., Dynamic routing-and-inventory problems: a review. *Transportation. Research-A* **1998,** 32, 585-598.
13. Moin, N. H.; Salhi, S., Inventory routing problems: a logistical overview. *Journal of Operations Research Society* **2007,** 58, 1185-1194.
14. Nagy, G.; S., S., Location-routing: Issues, models and methods. *European Journal of Operational Research* **2007,** 177, 649-672.
15. Webb, I. R.; R.C., L., Period and phase of customer replenishment: A new approach to the Strategic Inventory/Routing Problem. *European Journal of Operational Research* **1995,** 85, 132-148.
16. Campbell, M.; Savelsbergh, M. W. P., A Decomposition Approach for the Inventory-Routing Problem. *Transportation Science* **2004,** 38, 488-502.
17. Lei, L.; Liu, S.; Ruszczynski, A.; Park, S., On the integrated production, inventory and distribution routing problem. *IIE Transactions* **2006,** 38, 955-970.
18. Benders, J. F., Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* **1962,** 4, 238-252.
19. Fisher, M. L., The Lagrangean relaxation method for solving integer programming problems. *Management Science* **1981,** 27, 18.
20. Guignard, M.; Kim, S., Lagrangean decomposition: A model yielding stronger lagrangean bounds. *Mathematical Programming* **1987,** 39, 215-228.
21. Iyer, R. R.; Grossmann, I. E., A Bilevel Decomposition Algorithm for Long-Range Planning of Process Networks. *Industrial & Engineering Chemistry Research* **1998,** 37, 474-481.
22. Daichendt, M. M.; Grossmann, I. E., Integration of hierarchical decomposition and mathematical programming for the synthesis of process flowsheets. *Computers & Chemical Engineering* **1995,** 22, 147-175.
23. Pinto, J. M.; Grossmann, I. E., Optimal cyclic scheduling of multistage continuous multiproduct plants. *Computers & Chemical Engineering* **1994,** 18, 797-816.
24. You, F.; Wassick, J. M.; Grossmann, I. E., Risk Management for Global Supply Chain Planning under Uncertainty: Models and Algorithms. *AIChE Journal* **2009,** 55, 931-946.
25. You, F.; Grossmann, I. E., Balancing Responsiveness and Economics in the Design of Process Supply Chains with Multi-Echelon Stochastic Inventory. *AIChE Journal* **2009,** In press, DOI: 10.1002/aic.12244.
26. Chen, P.; Pinto, J. M., Lagrangean-based techniques for the supply chain management of flexible process networks. *Computers & Chemical Engineering* **2008,** 32, 2505-2528.
27. Neiro, S. M. S.; Pinto, J. M., A general modeling framework for the operational planning of petroleum supply chains. *Computers & Chemical Engineering* **2004,** 28, 871-896.
28. You, F.; Grossmann, I. E., Mixed-Integer Nonlinear Programming Models and Algorithms for Large-Scale Supply Chain Design with Stochastic Inventory Management. *Industrial & Engineering Chemistry Research* **2008,** 47, (20), 7802-7817.
29. You, F.; Grossmann, I. E., Integrated Multi-Echelon Supply Chain Design with Inventories under Uncertainty: MINLP Models, Computational Strategies. *AIChE Journal* **2010,** 56, (2), 419 - 440.
30. Bok, J.-K.; Grossmann, I. E.; Park, S., Supply Chain Optimization in Continuous Flexible Process Networks. *Industrial & Engineering Chemistry Research* **2000,** 39, (5), 1279-1290.

31. Carvalho, M. C. A.; Pinto, J. M., An MILP model and solution technique for the planning of infrastructure in offshore oilfields. *Journal Of Petroleum Science And Engineering* **2006,** 51, 97-110.
32. You, F.; Grossmann, I. E.; Wassick, J. M., Multisite Capacity, Production and Distribution Planning with Reactor Transformation: MILP Model, Bilevel Decomposition Algorithm vs. Lagrangean Decomposition Scheme. *Industrial & Engineering Chemistry Research* **2010,** In press, DOI: 10.1021/ie100559y.
33. Lainez, J. M.; Puigjaner, L.; Reklaitis, G. V., Financial and financial engineering considerations in supply chain and product development pipeline management. *Computers & Chemical Engineering* **2009,** 33, 1999-2011.
34. Papageorgiou, L. G.; Rotstein, G. E.; Shah, N., Strategic Supply Chain Optimization for the Pharmaceutical Industries. *Industrial & Engineering Chemistry Research* **2001,** 40, 275-286.
35. Tsiakis, P.; Shah, N.; Pantelides, C. C., Design of multi-echelon supply chain networks under demand uncertainty. *Industrial & Engineering Chemistry Research* **2001,** 40, 3585-3604.
36. You, F.; Grossmann, I. E., Design of Responsive Supply Chains under Demand Uncertainty. *Computers & Chemical Engineering* **2008,** 32, (12), 2839-3274.
37. You, F.; P. M. Castro; Grossmann, I. E., Dinkelbach's algorithm as an efficient method to solve a class of MINLP models for large-scale cyclic scheduling problems. *Computers & Chemical Engineering* **2009,** 33, 1879-1889.
38. You, F.; Grossmann, I. E., Stochastic Inventory Management for Tactical Process Planning under Uncertainties: MINLP Model and Algorithms. *AIChE Journal* **2010,** In press, DOI: 10.1002/aic.12338.
39. Shen, Z.-J. M.; Qi, L., Incorporating inventory and routing costs in strategic location models. *European Journal of Operations Research* **2007,** 179, 372-389.
40. Burns, L. D.; Hall, R. W.; Blumenfeld, D. E., Distribution Strategies that Minimize Transportation and Inventory Costs. *Operations Research* **1985,** 33, 469-489.
41. Zipkin, P. H., *Foundations of Inventory Management*. McGraw-Hill: Boston, MA, 2000.
42. Viswanathan, S.; Mathur, K., Integrating Routing and Inventory Decisions in One-Warehouse Multiretailer Multiproduct Distribution Systems. *Management Science* **1997,** 43, 294.
43. Jung, J.; Mathur, K., An Efficient Heuristic Algorithm for a Two-Echelon Joint Inventory and Routing Problem. *Transportation Science* **2007,** 41, 55-73.
44. Sindhuchao, S.; Romeijn, H. E.; Akcali, E.; Boondiskulchok, R., An Integrated Inventory- Routing System for Multi-item Joint Replenishment with Limited Vehicle Capacity. *Journal of Global Optimization* **2005,** 32, 93-118.
45. Haimovich, M.; Rinnooy-Kan, A. H. G., Bounds and Heuristics for Capacitated Routing Problems. *Mathematics of Operations Research* **1985,** 10, 527-541.
46. Agarwal, R.; Ahuja, R. K.; Laporte, G.; Shen, Z. J., A composite very large-scale neighborhood search algorithm for the vehicle routing problem. In *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*, Leung, J. Y.-T., Ed. Chapman Hall / CRC: 2004.
47. Balas, E., Disjunctive Programming and a Hierarchy of Relaxations for Discrete Continuous Optimization Problems. *SIAM Journal on Algebraic and Discrete Methods* **1985,** 6, (3), 466-486.
48. Rosenthal, R. E., GAMS - A User's Manual. In GAMS Development Corp.:

Washington, DC, 2008.
49. You, F.; Grossmann, I. E.; Pinto, J. M.; Megan, L., Optimal Distribution-Inventory Planning of Industrial Gases: II. MINLP Models and Algorithms for Stochastic Cases, Argonne National Laboratory Preprint ANL/MCS-P1812-1110, November 2010.
50. Applegate, D.; Bixby, R. E.; Chvátal, V.; Cook., W. J. *Concorde TSP Solver (http://www.tsp.gatech.edu/concorde)*, 2005.
51. Mittelmann, H. http://www-neos.mcs.anl.gov/neos/solvers/co:concorde/TSP.html. (Access on Feb. 8, 2010),