Optimal Distribution-Inventory Planning of Industrial Gases: II. MINLP Models and Algorithms for Stochastic Cases

Fengqi You,[1,2] Jose M. Pinto,[3] Ignacio E. Grossmann,[1*] 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

Submitted to *Industrial & Engineering Chemistry Research*

Abstract

In this paper we consider the inventory-distribution planning under uncertainty for industrial gas supply chains by extending the continuous approximation solution strategy proposed in part I. A stochastic inventory approach is proposed and it is incorporated into a multi-period two-stage stochastic mixed-integer nonlinear programming (MINLP) model to handle uncertainty of demand and loss or addition of customers. This nonconvex MINLP formulation takes into account customer synergies and simultaneously predicts the optimal sizes of customers' storage tanks, the safety stock levels and the estimated delivery cost for replenishments. To globally optimize this stochastic MINLP problem with modest computational time, we develop a tailored branch-and-refine algorithm based on successive piece-wise linear approximation. The solution from the stochastic MINLP is fed into a detailed routing model with shorter planning horizon to determine the optimal deliveries, replenishments and inventory. A clustering-based heuristic is proposed for solving the routing model with reasonable computational effort. Three case studies including instances with up to 200 customers are presented to demonstrate the effectiveness of the proposed stochastic models and solution algorithms.

Key words: planning & scheduling, industrial gas supply chain, inventory-routing, stochastic programming, global optimization, MINLP

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

1. Introduction

The design and operations of industrial gas inventory-distribution systems involve uncertainties at the strategic and at the operational levels. In the short term, the most common uncertainty is concerned with the product consumption and demands of the customers. Besides demand fluctuations, it is common to have losses or additions of customers due to contract termination or new contracts that are signed. A deterministic planning model[1] is a useful tool to help reduce costs by taking into account customer synergies in the inventory-distribution planning, and integrating strategic tank sizing decisions with operational vehicle routing decisions. However, uncertain demand fluctuations and uncertain additions and losses of customers may significantly affect the decision-making across the industrial gas supply chain. Thus, it is necessary to extend the deterministic planning models to address these uncertainties, and develop effective optimization algorithms for these problems.

A key component of decision-making under uncertainty is the representation of the stochastic parameters. There are two distinct ways of representing uncertainties. The scenario-based approach[2-6] attempts to capture the uncertainties by representing them in terms of a number of discrete realizations of the stochastic parameters, where each complete realization of all uncertain parameters gives rise to a scenario. In this way all the possible future outcomes are taken into account through the use of scenarios, in which recourse actions are anticipated for each scenario realization. The objective is to find a solution that on average performs well under all scenarios. This approach provides a straightforward way of formulating the problem, but its major drawback is that it typically relies on either a priori forecasting of all possible outcomes, or the explicit/implicit discretizations of continuous probability distributions. Thus, the problem size increases exponentially as the number of uncertain parameters and scenarios increases. This is particularly true when using continuous multivariate probability distributions with Gaussian quadrature integration schemes. Alternatively, Monte Carlo sampling could be used, but it also requires a rather large number of samples to achieve a desired level of accuracy.

The uncertainty can also be addressed through a chance constraint approach, which considers the uncertainty by treating one or more parameters as random variables with known probability distributions. Through multivariate integration over the continuous probability distribution functions, this approach can lead to a reasonable size deterministic equivalent representation of the probabilistic model. It circumvents the exponentially growing number of scenarios from explicit/implicit discretization or sampling, at the expense of introducing certain number of nonlinear terms into the model.[7-9] Although this approach has the limitation of not explicitly integrating recourse actions, it can effectively handle demand uncertainty at the operational level where the probabilistic description renders some operational planning variables to be stochastic. Based on the chance constraint method and some features of demand uncertainty, You & Grossmann[10-12] have recently proposed stochastic inventory models to deal with demand uncertainty in the design and operation of process systems. In their works, the uncertain demand is hedged by holding a certain amount of safety stocks before demand realization. The safety stock level is estimated by using a chance constraint that links service level with a demand probability distribution. The need of allowing recourse actions as in stochastic programming,[13] which can significantly increase problem size, is obviated by taking proactive action with the safety stocks. The stochastic attributes of the problem are translated into a deterministic optimization problem at the expense of adding nonlinear safety stock terms into the model.

As argued by Zimmermann,[14] the choice of the appropriate method to model the stochastic parameters is context-dependent, with no single theory being sufficient to model all kinds of uncertainty. The problem addressed in this paper includes two types of uncertainty: customer demand fluctuation at the operational level and uncertain changes in the distribution network due to loss or addition of customers at the strategic level. Based on the nature of these nonlinearities, we employ the stochastic inventory approach to deal with demand uncertainty, and use the scenario-based stochastic programming approach to handle the uncertain loss or addition of customers in the distribution network. We extend the continuous approximation approach, which is an efficient computational strategy for large-scale inventory-distribution planning of industrial gases, to address the two aforementioned types of uncertainty. The stochastic version of the continuous approximation solution strategy includes a two-stage stochastic mixed-integer nonlinear programming (MINLP) problem in the upper level and a decomposable mixed-integer linear programming (MILP) problem in the lower level. A global optimization method based on successive piecewise linear approximation is proposed to effectively solve the stochastic MINLP at the upper level, and a clustering-based heuristic is proposed for solving the routing model at the lower level with reasonable computational effort. We present the 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 the performance of the solution algorithms.

The rest of this paper is organized as follows. The general problem statement is provided in Section 2, which is followed by the proposed stochastic continuous approximation method in Section 3. The model formulation of the stochastic continuous approximation model and the global optimization algorithm for solving it effectively are given in Sections 4 and 5, respectively. In Section 6, we present computational results for three case studies. The last section concludes this paper and the clustering-based heuristic method for solving the routing problem is given in the appendix.

2. Problem Statement

We are given an industrial gas distribution network consisting of a production plant and a set of customers $n \in N$. The locations of the plant and customers, as well as the distances between them are given. We are also given a set of tanks with different sizes $i \in I$. The lower and upper bounds for tank with size *i* are given as T_i^L and T_i^U. We set the parameter $ot_{i,n} = 1$ if customer *n* has an existing tank with size *i*, and set it to zero otherwise. Similarly, the parameter $new_n = 1$ if customer n is a new customer without any existing tank, otherwise it is set to be zero. Each newly installed tank is full of merchant liquid product at the beginning. That is, the initial inventory $Vzero_{n,y}$ is assumed to be the same as the tank capacity for the new customers. For customers with existing tanks, their initial inventory levels are given. There is a set of trucks $j \in J$ with maximum capacity $Vtruck_j$. The delivery cost per distance traveled is given as ck_j. All the trucks are assumed to have an average traveling speed ($speed$) and a maximum number of working hours per day, hpd. For each delivery by truck, 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 cost and service costs are discounted with a working capital discount factor $wacc$ and a depreciation period in years given as dep.

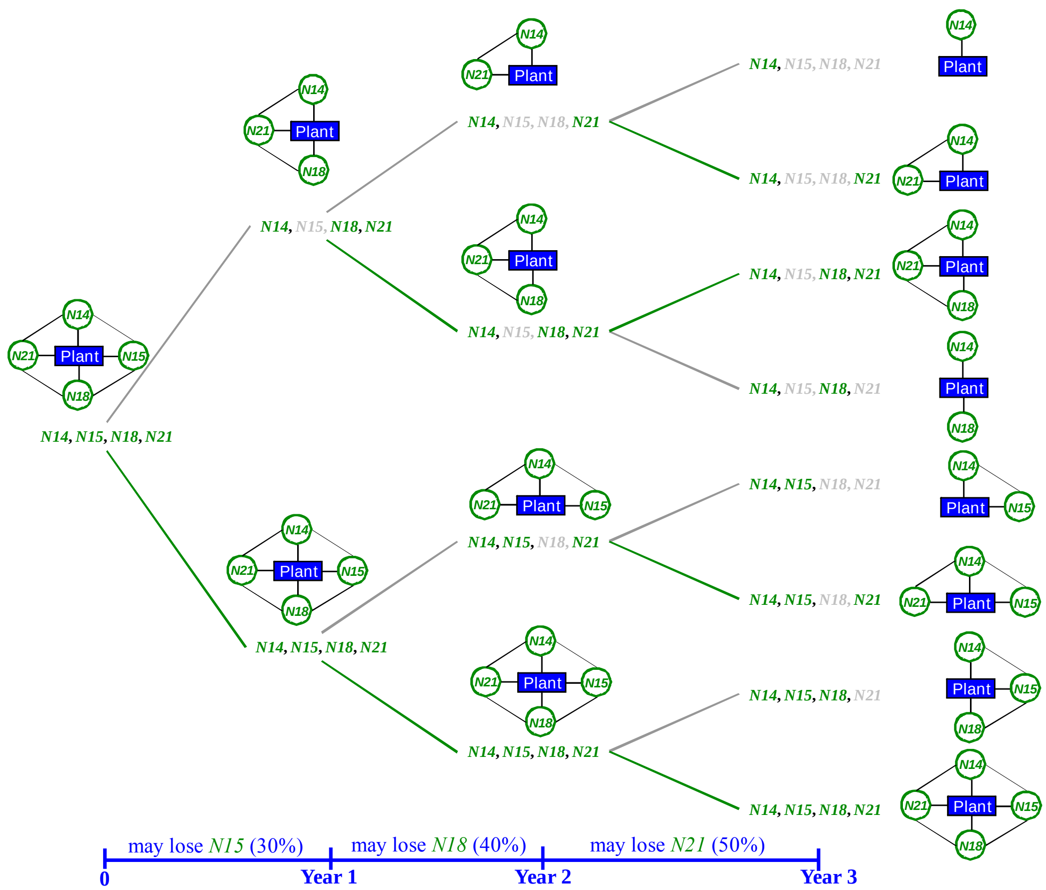


Figure 1. Industrial gas distribution network under uncertain loss of customers over time

The uncertainties arise from demand fluctuations and the losses or gains of customers in the distribution network. We assume that customer demands follow normal distributions, while the uncertain losses/additions of customers are represented through a set of scenarios $s \in S$. We assume that all the contracts are terminated or signed at the beginning of each year. Thus, the losses or additions of customers happen at the beginning of each year, which then defines a network structure for that year in the given scenario (see Figure 1 for an example). The probability of each scenario is given or can be derived from the probabilities of gaining or losing customers.

The problem is then to simultaneously determine the tank sizes and modification decisions at each customer location, and the schedule and quantity of each delivery for all the possible realizations of uncertainty. The objective is to minimize the total expected cost.

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

- Only one type of industrial gas is considered
- Scenarios with different network structures and the associated probabilities can be obtained
- Customers are added/lost at the beginning of the time period (year)
- Customer demand fluctuations follow normal distribution

For the continuous approximation approach, the major assumptions are:

- Cyclic inventory-routing
- All the customers have the same replenishment lead time
- The replenishment lead times are all the same in a year of a scenario
- Only one type of truck is used in a given year of a scenario

3. Solution Strategy

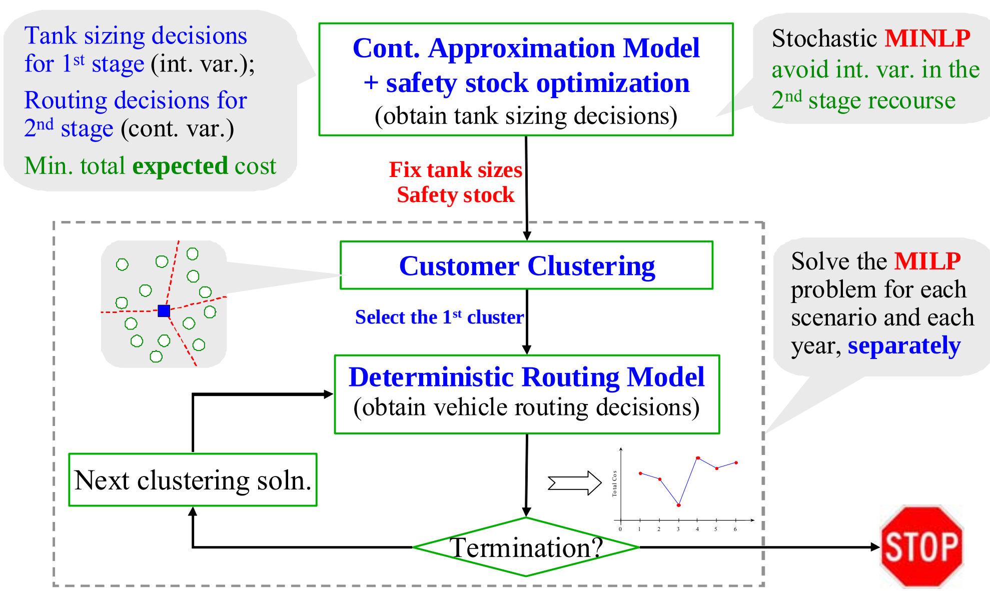


Figure 2. Inventory profile of a customer under cyclic inventory routing

We use the continuous approximation approach of Part I to deal with industrial gas inventory-distribution planning under uncertainty due to two major reasons. The first one is its computational efficiency. As shown through the case studies in our previous work,[1] this approach is much more efficient than the integrated MILP approach and the route selection-tank sizing approach for solving large scale problems, with little sacrifice of the solution quality. By decoupling the decision-making at the strategic and operational levels and using continuous approximation to estimate the routing cost, the proposed model can establish the optimal tradeoff between the strategic tank sizing cost and the approximate routing cost to predict optimal tank sizing decisions in the upper level. The lower level problem is then decomposed and solved as a vehicle routing problem with fixed tank sizes and shorter time horizon. The second major reason for using the continuous approximation approach is that the first stage decisions (here-and-now) of this problem are the tank sizing decisions and the second stage decisions (wait-and-see) are the routing decisions. These decisions are usually modeled with binary variables, and their presence in the second stage of the stochastic programming problem gives rise to significant computational challenge. However, the continuous approximation model can handle this challenge because it relies on very few binary variables and most discrete decisions are approximated using "continuous" functions. Therefore, the continuous approximation model is expected to have higher computational efficiency than other approaches that also incorporate stochastic programming.

The detailed algorithmic framework of the stochastic continuous approximation approach is given in Figure 2. At the upper level, we solve a stochastic version of the continuous approximation model by incorporating a stochastic inventory model for safety stock optimization. Although the deterministic continuous approximation model can be reformulated as an MILP, the inclusion of stochastic inventory approach introduces additional nonlinear terms (square root terms) and renders the model as an MINLP. The stochastic continuous approximation model also includes a two-stage stochastic programming element by treating all the tank sizing decisions as the first stage decisions and the remaining ones as the second stage. In this way, we avoid a large number of integer variables from detailed routing in the second stage of the stochastic programming model. Note that in principle, the problem could be formulated as a multi-stage stochastic programming model, but we only consider a two-stage approach in order to reduce the computational effort. Section 4 presents the detailed model formulation of the stochastic continuous approximation model and Section 5 introduces an efficient global optimization algorithm to solve this upper level problem.

After solving the upper level stochastic continuous approximation model to obtain the optimal tank sizes and safety stocks, these decisions are passed to the lower level for detailed routing. Since both kinds of uncertainty are taken into account in the stochastic continuous approximation model, we only need to solve a deterministic version of the detailed routing problem for each year and each scenario, which includes a specific network structure. The detailed routing model is the same as the integrated MILP model presented in Part I,[1] except that all the decisions for tank sizing and safety stocks are fixed using the values from the optimal solution of the stochastic continuous approximation model. Thus, the model formulation of the detailed routing problem is omitted in this paper. Vehicle routing problems have been well studied in the past decades and there are many heuristic methods that can be used to improve the computational efficiency.[15] In this work we use a clustering-based heuristic that is described in the Appendix.

4. Stochastic Continuous Approximation Model

The stochastic continuous approximation model is a stochastic MINLP that simultaneously considers tank sizing, safety stock optimization and approximated vehicle routing. The detailed model formulation is given below, and a list of indices, sets, parameters and variables are given in the Appendix.

4.1 Objective function

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

Min: $$E[Cost] = \sum_s prob_s \cdot \left(capcost_s + servcost_s + distcost_s \right) \tag{1}$$

where $prob_s$ is the probability of scenario s and the detailed cost components are listed in constraints (2) – (5).

$$capcost_s = \frac{1}{dep} \left(\sum_i Ccap_i \sum_y \left[\begin{array}{l} \sum\limits_{n | new_n = 0 \wedge tsize_n = 0 \wedge n \in N_{y,s}} \dfrac{ot_{i,n}}{(1+wacc)^{y-1}} + \sum\limits_{\substack{n | new_n = 0 \wedge tsize_n = 1 \\ \wedge espace_n = 0 \wedge n \in N_{y,s}}} \dfrac{yt_{i,n}}{(1+wacc)^{y-1}} + \\ \sum\limits_{\substack{n | new_n = 0 \wedge tsize_n = 1 \\ \wedge espace_n = 1 \wedge n \in N_{y,s}}} \dfrac{ot_{i,n} + et_{i,n}}{(1+wacc)^{y-1}} + \sum\limits_{n | new_n = 1 \wedge n \in N_{y,s}} \dfrac{yt_{i,n}}{(1+wacc)^{y-1}} \end{array} \right] \right) \tag{2}$$

Equation (2) is for the capital cost of scenario s. The four terms 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 location, respectively. $N_{y,s}$ is a subset of customers included in the distribution network of scenario *s* in year *y*.

Any tank change or addition leads to a service cost, as shown in Equations (3) and (4). The total distribution cost equals to the summation of discounted annual routing cost as in equation (5), where $crot_{y,s}$ is the estimated routing cost of scenario *s* in year *y* coming from the continuous approximation. Note that the capital, service and distribution costs are all discounted with the working capital discount factor $wacc$.

$$servcost_s = \frac{1}{dep}\left(\sum_i \sum_{n \in N_{y,s}} Cser_i \sum_y \frac{tins_{i,n} + et_{i,n}}{(1+wacc)^{y-1}}\right) \tag{3}$$

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

$$distcost_s = \sum_y \frac{crot_{y,s}}{(1+wacc)^y} \tag{5}$$

We note that the distribution cost given in objective function (5) is the estimation cost of vehicle routing from the "continuous approximation method", and it is not the exact routing cost that can be derived from solving the detailed routing problem.

4.2 Tank sizing constraints

Tank sizing decisions are the first stage decisions in the stochastic programming model, and they are independent of the scenarios. 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 as in Part I. 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

First, two parameters $tsize_n$ and $espace_n$ are introduced to define the conditions for tank sizing. We set $espace_n = 1$ if there is extra space for installing another tank at customer *n*, otherwise it is set it to be zero. Similarly, $tsize_n = 1$ if the tanks at customer *n* need to be sized or changed, otherwise it is set to be zero.

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 modification will be made. This relationship can be modeled by the following constraint.

$$\sum_i \left(yt_{i,n} + et_{i,n} \right) \le tsize_n, \qquad \forall n \in N_{y,s} \qquad (6)$$

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

If there is no extra space ($espace_n$=0), 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. On the other hand, if the storage tanks at customer n should be upgraded ($tsize_n = 1$) and there is extra space 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. These logic relationships can be modeled by:

$$\sum_i yt_{i,n} = 1 - espace_n \qquad \forall n \in N_{y,s} \text{ such that } tsize_n = 1 \quad (7)$$

$$\sum_i et_{i,n} \le espace_n \qquad \forall n \in N_{y,s} \qquad (8)$$

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,

$$Vl_n = \sum_i T_i^L \left(ot_{i,n} + yt_{i,n} + et_{i,n} \right) \qquad \forall n \in N_{y,s} \qquad (9)$$

$$Vu_n = \sum_i T_i^U \left(ot_{i,n} + yt_{i,n} + et_{i,n} \right) \qquad \forall n \in N_{y,s} \qquad (10)$$

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

$$Vzero_n = Vu_n \qquad \forall n \in N_{y,s} \mid new_n = 1 \qquad (11)$$

4.3 Truck constraints

Similarly to the deterministic version of this model, we assume that only one type of truck is selected for delivery in each year and each scenario. This assumption implies that all trucks have the same capacity and unit cost in each year and each scenario, although we do not explicitly account for truck availability in the continuous approximation. However, it should be noted that the "one type of truck" assumption is associated with the continuous approximation, which is used in the strategic level for tank sizing decisions. The availability of the trucks is considered in the lower level detailed routing model (please refer to Part I for details).

We note that different scenarios in different years may have different network structures, and thus different selection of trucks. The following constraints are used to define the truck selection.

$$\sum_j tru_{j,y,s} = 1 \qquad \forall y,s \tag{12}$$

$$cunit_{y,s} = \sum_j ck_j \cdot tru_{j,y,s} \qquad \forall y,s \tag{13}$$

$$ccapic_{y,s} = \sum_j tru_{j,y,s} \cdot Vtruck_j \cdot (1 - loss) \qquad \forall y,s \tag{14}$$

Constraint (12) shows that only one type of truck is selected per year and $tru_{j,y,s}$ is a binary variable that is equal to 1 if truck j is selected for delivery in year y and scenario s, otherwise it is equal to zero. Constraints (13) and (14) are for the unit transportation cost ($cunit_{y,s}$) in year y of scenario s, and the effective delivery capacity ($ccapic_{y,s}$) after adjustment of loss of the truck for a scenario in a specific year.

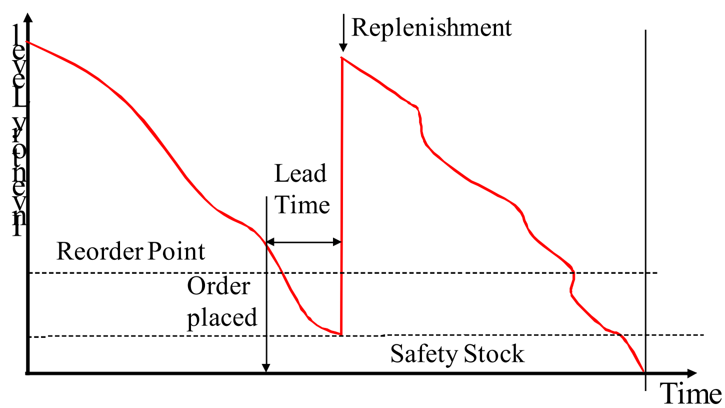
The lead time of a replenishment cycle ($LT_{y,s}$) in year y of scenario s should not be less than the sum of the total travel time, which is given by the minimum routing distance ($mrt_{y,s}$) divided by the traveling speed and the working hours per day (*hpd*), and the total loading time, which includes the loading times at the customer locations and at the plant. There will be at least $|N_{y,s}|$ times of loading in the customers as each customer will be visited at least once in a replenishment cycle. The loading time at the plant should be greater than the loading time of deliveries from the plant (*FT_del*). Thus, the lead time constraint is given by,

$$LT_{y,s} \geq \frac{mrt_{y,s}}{speed \cdot hpd} + FT\_load \cdot |N_{y,s}| + FT\_del\,, \ \forall y,s \tag{15}$$

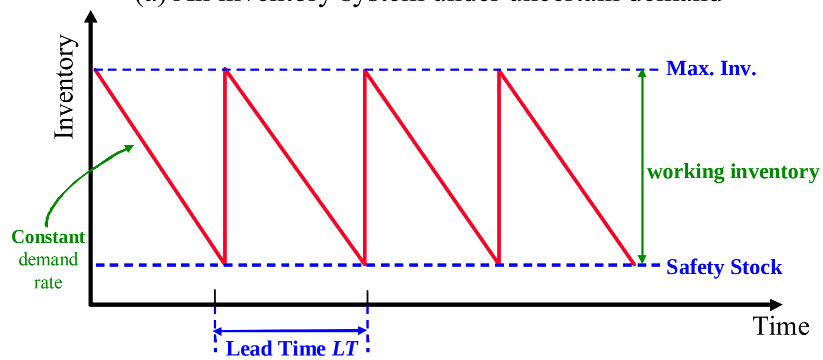
4.4 Cyclic inventory-routing under uncertainty based on continuous approximation

Because the focus of this work is on the strategic tank sizing decisions, we employed a continuous approximation method to estimate the optimal routing cost as a result of different tank sizing decisions. By using the continuous approximation we simplify the detailed routing problem, while still capturing the tradeoff between capital costs and routing costs at the strategic level.

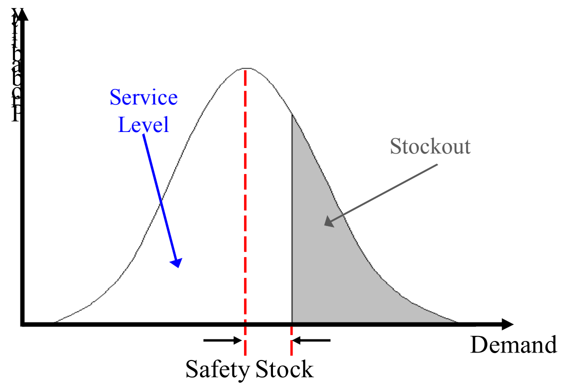
With the proposed approach, we approximate the discrete variables and parameters associated with vehicle routing using continuous functions, which represent distributions of customer locations and demands. The profile of an inventory system under uncertain demand is given in Figure 3(a). Similarly to the deterministic case, we assume cyclic inventory-routing for each scenario in each year due to their potentially different network structures and demands. Thus, the inventory profile of a customer for a scenario in a specific year is given in Figures 3(b). From this figure, we have the tank size no less than the maximum inventory level, which is the summation of working inventory and safety stock.



(a) An inventory system under uncertain demand



(b) Inventory profile under cyclic inventory-routing



(c) Safety stock and service level under normally distributed demand

Figure 3. Stochastic inventory system under continuous approximation

The working inventory equals to the demand rate times the replenishment lead time. The safety stock level can be determined by using the stochastic inventory approach as follows. A storage facility under demand uncertainty may not always have sufficient stock to handle the changing demand. If the inventory level is less than the demand during the replenishment lead time, stockout may happen. *Type I service level* is defined as the probability that the total inventory on hand is more than the demand (as shown Figure 3b). If the demand is normally distributed with mean μ and standard deviation σ, the optimal safety stock level to guarantee a service level α is $z_\alpha \sigma$, where z_α is a standard normal deviate such that $\Pr(z \le z_\alpha) = \alpha$.[16] If the demand rate is normally distributed with mean μ and standard deviation σ, then the uncertain demand over the replenishment lead time LT also follows a normal distribution with mean $\mu \cdot LT$ and variance $\sigma^2 \cdot LT$. Thus, the optimal safety stock level to guarantee a service level α is $z_\alpha \sqrt{\sigma^2 \cdot LT} = z_\alpha \cdot \sigma \cdot \sqrt{LT}$.[16] We should note that the acceptable practice in this field is to assume a normal distribution of the demand, although of course other distribution functions can be specified. The stochastic inventory model has been proved to provide very good approximations for inventory system under demand uncertainty.[17] In this way, the maximum inventory level is modeled as a nonlinear function of the replenishment lead time and demand probability distribution. A tradeoff between the inventory and routing costs is also established. If the replenishment frequency is high, the routing cost could also be high, but the working inventory level maybe low, so we only need a small tank and vice versa.

If we use $x_{y,s}$ to denote the number of replenishment cycles in year y of scenario s, then the corresponding replenishment lead time ($LT_{y,s}$) should satisfy the equation below:

$$LT_{y,s} \cdot x_{y,s} = Hz_y \qquad \forall y,s$$

where Hz_y is the time duration of year y. Note that both Hz_y and $LT_{y,s}$ are in terms of physical days instead of working day, as customers consume demand continuously. The above equation has a bilinear term on the left hand side. To linearize it, similar to Part I, the integer variable $x_{y,s}$ can be reformulated with binary variables $Ix_{k,y,s} \in \{0,1\}$ as follows:

$$x_{y,s} = \sum_k 2^{|k|-1} \cdot Ix_{k,y,s} \qquad \forall y,s \qquad (16)$$

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

With equations (16), we can linearize the nonlinear constraint $LT_{y,s} \cdot x_{y,s} = Hz_y$ as $LT_{y,s} \cdot x_{y,s} = \sum_k 2^{|k|-1} \cdot LT_{y,s} \cdot Ix_{k,y,s}$. By introducing a nonnegative continuous variable $LTIx_{k,y,s} = LT_{y,s} \cdot Ix_{k,y,s}$, we have the following reformulated constraint:

$$\sum_k 2^{|k|-1} \cdot LTIx_{k,y,s} = Hz_y \qquad \forall y,s \qquad (17)$$

We also need the following linearization constraints to define the new variable $LTIx_{k,y,s}$:[18-19]

$$LTIx_{k,y,s} + LTIx1_{k,y,s} = LT_{y,s} \qquad \forall k,y,s \qquad (18.1)$$

$$LTIx_{k,y,s} \le LT^U_{y,s} \cdot Ix_{k,y,s} \qquad \forall k,y,s \qquad (18.2)$$

$$LTIx1_{k,y,s} \le LT^U_{y,s} \cdot \left(1 - Ix_{k,y,s}\right) \qquad \forall k,y,s \qquad (18.3)$$

$$Ix_{k,y,s} \in \{0,1\},\ LTIx_{k,y,s} \ge 0,\ LTIx1_{k,y,s} \ge 0 \qquad \forall k,y,s \qquad (18.4)$$

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

Let $Vm_{n,y,s}$ be the maximum inventory level of customer n in scenario s and year y. From Figure 3, we know that the maximum inventory level should be no less than the summation of working inventory ($winv_{n,y,s}$), safety stock ($safety_{n,y,s}$) and minimum volume of the tank in customer *n* in year *y* (Vl_n) defined in constraint (9). Thus, we have the following constraint.

$$Vm_{n,y,s} \geq winv_{n,y,s} + safety_{n,y,s} + Vl_n \qquad \forall n,y,s \qquad (19)$$

The maximum inventory level should not exceed the maximum volume of the tank defined by the tank size of customer *n* in constraint (10).

$$Vm_{n,y,s} \leq Vu_n \qquad \forall n,y,s \qquad (20)$$

As discussed above, the safety stock level should be equal to the product of the service level parameter z_α, the standard deviation of daily demand $\sigma_{n,y}$ and the square root of replenishment lead time $LT_{y,s}$, which is measured in days.

$$safety_{n,y,s} = z_\alpha \cdot \sigma_{n,y} \cdot \sqrt{LT_{y,s}} \qquad \forall n,y,s \qquad (21)$$

For customer *n*, its working inventory ($winv_{n,y,s}$) is the replenishment that it receives in a replenishment cycle. Thus, the working inventory times the number of replenishment cycles should be equal to the annual amount of product delivered from plant to customer *n* in scenario *s* and year *y* ($Trp_{n,y,s}$).

$$winv_{n,y,s} \cdot x_{y,s} = Trp_{n,y,s} \qquad \forall n,y,s$$

Based on equation (16), we can reformulate the bilinear term on the left hand side of the above equatioin as $winv_{n,y,s} \cdot x_{y,s} = \sum_k 2^{|k|-1} \cdot winv_{n,y,s} \cdot Ix_{k,y,s}$. By introducing a nonnegative continuous variable $wIx_{k,n.y,s} = winv_{n,y,s} \cdot Ix_{k,y,s}$, we can obtain the following reformulated linear constraint:

$$\sum_k 2^{|k|-1} \cdot wIx_{k,n,y,s} = Trp_{n,y,s} \qquad \forall n,y,s \qquad (22)$$

We also need the following linearization constraints to define $wIx_{k,n.y,s}$:

$$wIx_{k,n.y,s} + wIx1_{k,n.y,s} = winv_{n,y,s} \qquad \forall k,n,y,s \qquad (23.1)$$

$$wIx_{k,n.y,s} \leq winv^U_{n,y,s} \cdot Ix_{k,y,s} \qquad \forall k,n,y,s \qquad (23.2)$$

$$wIx1_{k,n.y,s} \leq winv^U_{n,y,s} \cdot \left(1 - Ix_{k,y,s}\right) \qquad \forall k,n,y,s \qquad (23.3)$$

$$Ix_{k,y,s} \in \{0,1\},\ wIx_{k,n.y,s} \geq 0,\ wIx1_{k,n.y,s} \geq 0 \qquad \forall k,n,y,s \qquad (23.4)$$

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

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

$$Trp_{n,y,s} = dem_{n,y} + Vend_{n,y,s} - Vzero_n + Vl_n + safety_{n,y,s} \qquad \forall n,s,\ y=1 \tag{24}$$

$$Trp_{n,y,s} = dem_{n,y} + Vend_{n,y,s} - Vend_{n,y-1,s} \qquad \forall n,s,y \mid y \geq 2 \tag{25}$$

$$Vend_{n,y,s} \leq winv_{n,y,s} \qquad \forall n,s,y \tag{26}$$

where $dem_{n,y}$ is the demand rate of customer n in year y, $Vend_{n,y,s}$ is the inventory level of customer n in scenario s 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 stocks.

4.5 Continuous approximation for routing cost

The capacitated vehicle routing cost is estimated via a continuous approximation approach. Let $Trp_{n,y,s}$ be the total amount of product delivered from plant to customer n in scenario s and year y, $ccapic_{y,s}$ be the effective truck capacity (truck capacity after accounting for product loss), rr_n be the distance between the plant and customer n, and $TSP_{y,s}$ be the length of the optimal traveling salesman tour that all the customers included in scenario s and year y are visited once. Based on a continuous approximation, the minimum routing distance for each replenishment cycle in scenario s and year y ($mrt_{y,s}$) can be approximated with the following formula:[1]

$$mrt_{y,s} \approx 2 \cdot \left(\frac{\sum_{n \in N_{y,s}} Trp_{n,y,s} \cdot rr_n}{x_{y,s} \cdot ccapic_{y,s}} \right) + \left(1 - \frac{1}{ccapic_{y,s}} \right) \cdot TSP_{y,s}, \quad \forall y,s$$

To reduce the nonlinearities, we introduce a new positive variable $seg_{y,s}$ such that

$$seg_{y,s} = \frac{\sum_{n \in N_{y,s}} Trp_{n,y,s} \cdot rr_n}{x_{y,s} \cdot ccapic_{y,s}}, \qquad \forall y,s$$

which is equivalent to the following constraint,

$$seg_{y,s} \cdot x_{y,s} \cdot ccapic_{y,s} = \sum_{n \in N_{y,s}} Trp_{n,y,s} \cdot rr_n = \sum_{n \in N_{y,s}} winv_{n,y,s} \cdot x_{y,s} \cdot rr_n = x_{y,s} \cdot \sum_{n \in N_{y,s}} winv_{n,y,s} \cdot rr_n, \qquad \forall y,s$$

Because $x_{y,s}$ is a positive integer variable, the above equation is equivalent to the following one,

$$seg_{y,s} \cdot ccapic_{y,s} = \sum_{n \in N_{y,s}} \left(winv_{n,y,s} \cdot rr_n \right), \qquad \forall y,s$$

Based on equation (14), we can reformulate the bilinear term on the left hand side of the above equation as $seg_{y,s} \cdot ccapic_{y,s} = \sum_j seg_{y,s} \cdot tru_{j,y,s} \cdot Vtruck_j \cdot (1 - loss)$. By introducing a nonnegative continuous variable $TruSeg_{j,y,s} = seg_{y,s} \cdot tru_{j,y,s}$, we have the following reformulated linear constraint:

$$\sum_j TruSeg_{j,y,s} \cdot Vtruck_j \cdot (1 - loss) = \sum_{n \in N_{y,s}} \left(winv_{n,y,s} \cdot rr_n \right), \quad \forall y,s \tag{27}$$

We also need the following linearization constraints:

$$TruSeg_{j,y,s} + TruSeg1_{j,y,s} = Seg_{y,s} \qquad \forall j,y,s \tag{28.1}$$

$$TruSeg_{j,y,s} \le seg^U_{y,s} \cdot tru_{j,y,s} \qquad \forall j,y,s \tag{28.2}$$

$$TruSeg1_{j,y,s} \le seg^U_{y,s} \cdot \left(1 - tru_{j,y,s}\right) \qquad \forall j,y,s \tag{28.3}$$

$$tru_{j,y,s} \in \{0,1\},\ TruSeg_{j,y,s} \ge 0,\ TruSeg1_{j,y,s} \ge 0 \qquad \forall j,y,s \tag{28.4}$$

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

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

$$mrt_{y,s} = 2 \cdot seg_{y,s} + \left(1 - \frac{1}{ccapic_{y,s}}\right) \cdot TSP_{y,s}, \qquad \forall y,s$$

Further, the reciprocal of $ccapic_{y,s}$ ($Tccapic_{y,s}$) can be modeled through the following linear equation,

$$Tccapic_{y,s} = \sum_j \frac{tru_{j,y,s}}{Vtruck_j \cdot (1 - loss)} \qquad \forall y,s \tag{29}$$

which comes directly from equation (14).

Based on equation (29), we can easily reformulate the minimum routing distance constraint as:

$$mrt_{y,s} = 2 \cdot seg_{y,s} + \left(1 - Tccapic_{y,s}\right) \cdot TSP_{y,s}, \qquad \forall y,s \tag{30}$$

If we know the unit distance transportation cost of scenario s in year y ($cunit_{y,s}$), then the total delivery cost of this scenario in this year ($crot_{y,s}$) is the product of the unit transportation cost, the number of replenishment cycles and the minimum routing distance of each replenishment cycle.

$$crot_{y,s} = cunit_{y,s} \cdot mrt_{y,s} \cdot x_{y,s} \qquad \forall y,s$$

Based on equations (13) and (16), similarly as in Part I, we have $cunit_{y,s} \cdot mrt_{y,s} \cdot x_{y,s} = \sum_j \sum_k \left(2^{|k|-1} \cdot ck_j \cdot tru_{j,y,s} \cdot Ix_{k,y,s} \cdot mrt_{y,s} \right)$. If we set $mrIx_{k,y,s} = Ix_{k,y,s} \cdot mrt_{y,s}$ and $mrItru_{j,k,y,s} = tru_{j,y,s} \cdot mrIx_{k,y,s}$, then the above nonlinear constraint can be linearized as follows:

$$crot_{y,s} = \sum_j \sum_k 2^{|k|-1} \cdot ck_j \cdot mrItru_{j,k,y,s} \qquad \forall y,s \qquad (31)$$

We also need the following linear constraints:

$$mrIx_{k,y,s} + mrIx1_{k,y,s} = mrt_{y,s} \qquad \forall k,y,s \qquad (32.1)$$

$$mrIx_{k,y,s} \le mrt^U_{y,s} \cdot Ix_{k,y,s} \qquad \forall k,y,s \qquad (32.2)$$

$$mrIx1_{k,y,s} \le mrt^U_{y,s} \cdot \left(1 - Ix_{k,y,s}\right) \qquad \forall k,y,s \qquad (32.3)$$

$$Ix_{k,y,s} \in \{0,1\},\ mrIx_{k,y,s} \ge 0,\ mrIx1_{k,y,s} \ge 0 \qquad \forall k,y,s \qquad (32.4)$$

$$mrItru_{j,k,y,s} + mrItru1_{j,k,y,s} = mrIx_{k,y,s} \qquad \forall j,k,y,s \qquad (33.1)$$

$$mrItru_{j,k,y,s} \le mrt^U_{y,s} \cdot tru_{j,y,s} \qquad \forall j,k,y,s \qquad (33.2)$$

$$mrItru1_{j,k,y,s} \le mrt^U_{y,s} \cdot \left(1 - tru_{j,y,s}\right) \qquad \forall j,k,y,s \qquad (33.3)$$

$$tru_{j,y,s} \in \{0,1\},\ mrItru_{j,k,y,s} \ge 0,\ mrItru1_{j,k,y,s} \ge 0 \qquad \forall j,k,y,s \qquad (33.4)$$

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

4.6 Stochastic MINLP Reformulation

After reformulation and linearization, the stochastic continuous approximation model is a non-convex MINLP with the objective function given in (1) and constraints (2) – (33). The remaining nonlinear nonconvex term in this model is the square root term in safety stock constraint (21).

5. Global Optimization Algorithm

Although small scale instances of the stochastic continuous approximation model can be solved to global optimality by using a global optimizer, medium and large-scale problems are often computationally intractable with a direct solution approach due to the combinatorial nature and nonlinear nonconvex terms. In this section, we introduce an efficient global optimization algorithm based on a special property of the model and on successive piecewise linear approximations to tackle this nonconvex MINLP problem.

An important property of this model is given as follows.

Property 1. *If we replace the square root terms in the safety stock constraints (21) with piecewise linear under-estimators, the solution of the resulting MILP model provides a global lower bound of the global minimum solution of the original MINLP.*

We omit the proof because Property 1 is straightforward and easy to prove. Based on Property 1, we can first construct a lower bounding MILP problem based on piecewise linear approximations, and then employ a branch-and-refine method to globally optimize the nonconvex stochastic continuous approximation problem.

5.1 Piece-wise Linear Approximation

The only nonlinear terms of the stochastic continuous approximation model are the univariate square root terms, $\sqrt{LT_{y,s}}$, in the safety stock constraint (21). To improve computational efficiency, we consider piece-wise linear approximations for the concave square root terms. There are several different approaches to model piecewise linear functions for a concave term. In this work, we use the "multiple-choice" formulation[20-21] to approximate the square root term $\sqrt{LT_{y,s}}$. Let $P_{y,s} = \{1, 2, 3, \cdots, p\}$ denote the set of intervals in the piecewise linear function $\varphi\left(LT_{y,s}\right)$, and $u_{y,s,0}$, $u_{y,s,1}$, $u_{y,s,2}$,... $u_{y,s,p}$, be the lower and upper bounds of $LT_{y,s}$ for each interval. The "multiple choice" formulation of $\varphi\left(LT_{y,s}\right) = \sqrt{LT_{y,s}}$ for a given year and scenario is given by,

$$\varphi\left(LT_{y,s}\right) = \min \quad \sum_{p}\left(\beta_{y,s,p}E_{y,s,p} + \alpha_{y,s,p}F_{y,s,p}\right) \tag{34}$$

s.t.

$$\sum_{p} E_{y,s,p} = 1 \tag{35}$$

$$\sum_p F_{y,s,p} = LT_{y,s} \tag{36}$$

$$u_{y,s,p-1}E_{y,s,p} \le F_{y,s,p} \le u_{y,s,p}E_{y,s,p}\,,\ \forall p \tag{37}$$

$$E_{y,s,p} \in \{0,1\}\,,\ F_{y,s,p} \ge 0\,,\ \forall p \tag{38}$$

where $\alpha_{y,s,p} = \dfrac{\sqrt{u_{y,s,p}} - \sqrt{u_{y,s,p-1}}}{u_{y,s,p} - u_{y,s,p-1}}$ and $\beta_{y,s,p} = \sqrt{u_{y,s,p}} - \alpha_{y,s,p}u_{y,s,p}$, $p \in P$.

Substituting (34) - (38) into the safety stock constraint (21) yields a mixed-integer linear programming (MILP) model, which is a piece-wise linear under-estimator of the stochastic MINLP problem. The MILP lower bounding model formulation is given as follows.

Min: $E[Cost] = \sum_s prob_s \cdot \left(capcost_s + servcost_s + distcost_s\right)$ (1)

s.t. Constraints (2) – (20), (22) – (33)

$$safety_{n,y,s} = z_\alpha \cdot \sigma_{n,y} \cdot \sum_p \left(\beta_{y,s,p}E_{y,s,p} + \alpha_{y,s,p}F_{y,s,p}\right), \quad \forall n \in N_{y,s}, y, s \tag{39}$$

$$\sum_p E_{y,s,p} = 1 \qquad \forall y,s \tag{40}$$

$$\sum_p F_{y,s,p} = LT_{y,s} \qquad \forall y,s \tag{41}$$

$$u_{y,s,p-1}E_{y,s,p} \le F_{y,s,p} \le u_{y,s,p}E_{y,s,p} \qquad \forall y,s,p \tag{42}$$

$$E_{y,s,p} \in \{0,1\}\,,\ F_{y,s,p} \ge 0\,, \qquad \forall y,s,p \tag{43}$$

5.2 Branch-and-Refine Algorithm

In order to globally optimize the non-convex MINLP problem, we can first solve the MILP lower bounding problem, whose solution provides a valid lower bound to the global optimal solution, and then solve a reduced MINLP problem by fixing the binary variables $tru_{j,y,s}$ and $Ix_{k,y,s}$. Note that we do not fix the tank sizing decisions, because in the MILP lower bound problem, the safety stock levels are underestimated and thus the optimal tank sizes might be underestimated. To avoid infeasibility, we only fix $tru_{j,y,s}$ and $Ix_{k,y,s}$ to solve the reduced MINLP.

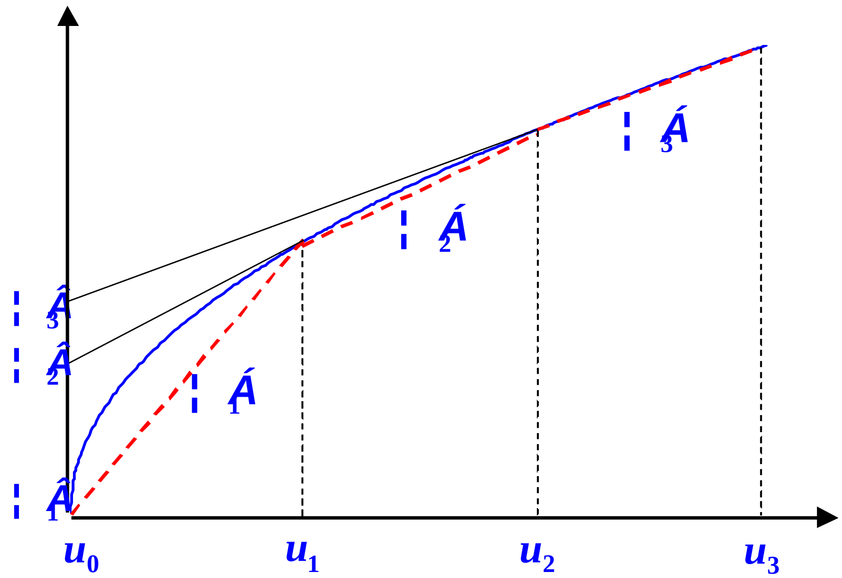


Figure 4 Piece-wise linear under-estimator of a concave term

The optimal solution of the reduced MINLP is also a feasible solution of the original stochastic continuous approximation model, so its objective value provides a valid global upper bound of the original MINLP problem. The remaining challenge is how to iteratively refine and improve the solution so that the global optimal solution can be obtained after a finite number of iterations. If we use a sufficiently large number of intervals in the piecewise linear lower bonding MILP problem, we obtain the global solution within sufficiently small optimality margin. The reason is that the more intervals that are used, the better is the approximation of the square root function (see Figure 4). However, more intervals require additional variables and constraints in the lower bounding MILP model. Similarly to our previous work,[22] we use an iterative branch-and-refine strategy based on successive piece-wise linear approximation to control the size of the problem.

In the first step of this algorithm, we consider a single linear approximation in the lower bounding MILP, i.e. replacing all the square root terms in the stochastic continuous approximation MINLP model with their corresponding secants as shown in Figure 5a. Thus, the optimal solution of the MILP problem provides the first lower bound LB1. An upper bound can be obtained by fixing the values of the binary variables $tru_{j,y,s}$ and $Ix_{k,y,s}$ and then solve the stochastic continuous approximation model in the reduced variable space. Due to the presence of tank sizing variables, the reduced problem is still an MINLP. As the lower bounding MILP underestimates the tank sizes and the maximum tank size is fixed, it is possible that the reduced MINLP can be infeasible. In that case, we move on to the next iteration and the upper bound is still the best upper bound provided in the previous iterations. Of course, the lower bounding MILP can be infeasible as well, but it implies that the original problem is infeasible, because the MILP is a relaxation of the original MINLP.

In the next step, we use the optimal solution of variable $LT_{y,s}$ in the upper bounding problem as the lower bound of a new interval, and consider a two-interval linear approximation of the square root terms as shown in Figure 5b. If the optimal solution of the upper bounding problem in the previous iteration lies at the bounds of some intervals, we do not add any new interval for the corresponding square root term $\sqrt{LT_{y,s}}$. After we construct the two-interval linear approximation MILP model, we can similarly obtain a lower bound, and then an upper bound by solving the reduced MINLP.

As shown in Figure 5c, the number of intervals in the piecewise linear model increases as the iteration number increases. Meanwhile, the best lower bound increases while the best upper bond decreases. The algorithm keeps iterating until the lower bound and upper bound are close enough to reach an optimality tolerance, e.g. 10^{-6}. Note that the number of intervals does not always equal to the number of iterations, because the optimal solutions in some iterations may lie at the bounds of the intervals and in that case we do not increase new intervals for the corresponding square root terms.

To summarize, the proposed branch-and-refine algorithm based on successive piece-wise linear approximation is as follows:

<u>Step 1:</u> (Initialization)

Initialize $iter = 1$, $LB = 0$, $UB = +\infty$. Set $NP_{y,s}^{iter} = 1$, $p \in P_{y,s} = \{0,1,...,NP_{y,s}^{iter}\}$. Use a single linear approximation, i.e. the secant, for the square root terms. To achieve this, set $u_{y,s,0} = 0$ and $u_{y,s,1} = LT_{y,s}^{U}$, as well as $\alpha_{y,s,1} = \dfrac{\sqrt{u_{y,s,1}} - \sqrt{u_{y,s,0}}}{u_{y,s,1} - u_{y,s,0}} = \dfrac{1}{\sqrt{LT_{y,s}}}$ and $\beta_{y,s,1} = \sqrt{u_{y,s,1}} - \alpha_{y,s,1} \cdot u_{y,s,1} = 0$.

<u>Step 2:</u>

At iteration $iter$, solve the piece-wise linear approximation MILP model. If the MILP problem is infeasible, then the original problem is also infeasible and the algorithm stops. If not, denote the optimal objective function value as φ^{iter} and the optimal solution of variables $tru_{j,y,s}$ and $Ix_{k,y,s}$ as ($tru_{j,y,s}^{*iter}$, $Ix_{k,y,s}^{*iter}$). If $\varphi^{iter} > LB$, then set $LB = \varphi^{iter}$.

Step 3:

Fix the values of binary variables $tru_{j,y,s} = tru_{j,y,s}^{*iter}$ and $Ix_{k,y,s} = Ix_{k,y,s}^{*iter}$, and solve the original stochastic continuous approximation MINLP model in the reduced space, which is an MINLP with fewer binary variables to obtain the *local* optimal solution.

If the reduced MINLP is feasible, denote the optimal value of the objective function as Φ^{iter} and the optimal solution as $LT_{y,s}^{iter}$. If $\Phi^{iter} < UB$, then set $UB = \Phi^{iter}$, store the current optimal solution.

If the reduced MINLP is infeasible, denote the optimal solution of the lead times in the lower bounding MILP as $LT_{y,s}^{iter}$.

Find $n_{y,s}^{iter}$ such that the optimal solution $LT_{y,s}^{iter}$ lies in the $n_{y,s}^{iter}$ interval. i.e. $u_{y,s,n_{ci,y,s}^{iter}-1} \le LT_{y,s} \le u_{y,s,n_{ci,y,s}^{iter}}$. One approach to find the proper n_j^{iter} is to compute the product $\left(LT_{y,s}^{iter} - u_{y,s,p-1}\right)\cdot\left(LT_{y,s}^{iter} - u_{y,s,p}\right)$ for all the p = 2, 3, …, $NP_{y,s}^{iter}$, and then denote the first p that leads to a non-positive value (zero or negative value) of the product as $n_{y,s}^{iter}$, i.e. if $\left(LT_{y,s}^{iter} - u_{y,s,p-1}\right)\cdot\left(LT_{y,s}^{iter} - u_{y,s,p}\right) \le 0$, set $n_{y,s}^{iter} = p$.

If $UB - LB \le \varepsilon$ (e.g. 10^{-9}), stop and output the optimal solution; otherwise, go to the next step.

Step 4:

For those year y and scenario s such that $\sqrt{LT_{y,s}^{iter}} = \sum_p (\beta_{y,s,p}^{iter} E_{y,s,p}^{iter} + \alpha_{y,s,p}^{iter} F_{y,s,p}^{iter})$, set $NP_{y,s}^{iter+1} = NP_{y,s}^{iter}$, $\alpha_{y,s,p}^{iter+1} = \alpha_{y,s,p}^{iter}$ and $\beta_{y,s,p}^{iter+1} = \beta_{y,s,p}^{iter}$.

For other year y and scenario s, set $NP_{y,s}^{iter+1} = NP_{y,s}^{iter} + 1$ and update set $P_{y,s}$, i.e. $p \in P_{y,s} = \left\{0, 1, ..., NP_{y,s}^{iter+1}\right\}$ and update $u_{y,s,p}$, $\alpha_{y,s,p}$ and $\beta_{y,s,p}$ as follows:

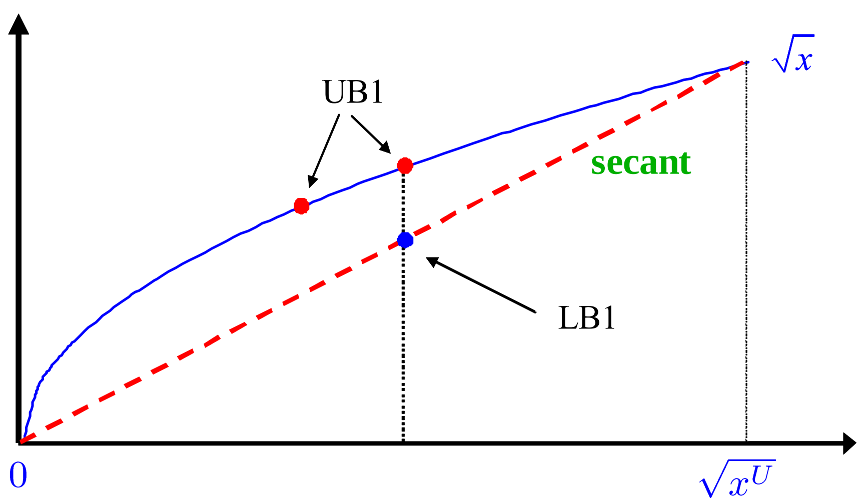
- For $p < n_{y,s}^{iter}$ (i.e. p = 0, 1, 2 … , $n_{y,s}^{iter} - 1$), set $u_{y,s,p}^{iter+1} = u_{y,s,p}^{iter}$, $\alpha_{y,s,p}^{iter+1} = \alpha_{y,s,p}^{iter}$ and $\beta_{y,s,p}^{iter+1} = \beta_{y,s,p}^{iter}$.

- For $p = n_{y,s}^{iter}$, set $u_{y,s,n_j^{iter}}^{iter+1} = LT_{y,s}^{iter}$, $\alpha_{y,s,n_{y,s}^{iter}}^{iter+1} = \dfrac{\sqrt{LT_{y,s}^{iter}} - \sqrt{u_{y,s,n_{y,s}^{iter}-1}}}{LT_{y,s}^{iter} - u_{y,s,n_{y,s}^{iter}-1}}$ and $\beta_{y,s,n_{y,s}^{iter}}^{iter+1} = \sqrt{LT_{y,s}^{iter}} - \alpha_{y,s,n_{y,s}^{iter}}^{iter+1} \cdot LT_{y,s}^{iter}$

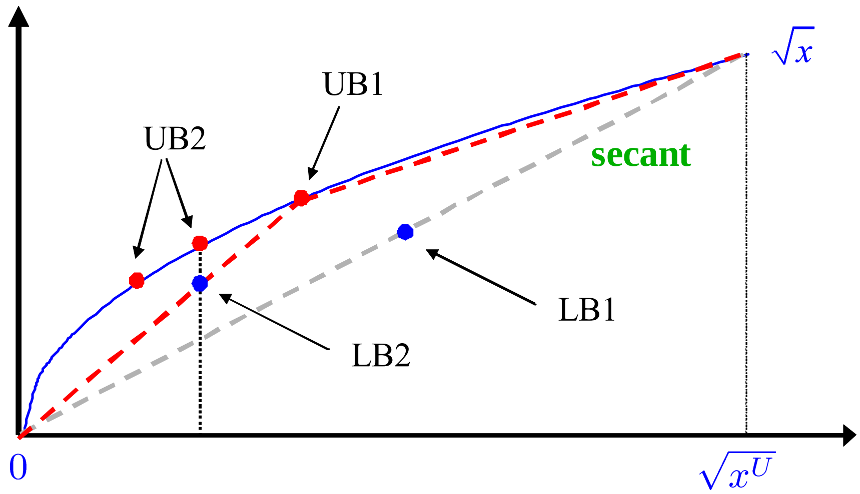
- For $p = n_{y,s}^{iter} + 1$, set $u_{y,s,n_{y,s}^{iter}+1}^{iter+1} = u_{y,s,n_{y,s}^{iter}}^{iter}$, $\alpha_{y,s,n_{y,s}^{iter}+1}^{iter+1} = \dfrac{\sqrt{u_{y,s,n_{y,s}^{iter}}} - \sqrt{LT_{y,s}^{iter}}}{u_{y,s,n_{y,s}^{iter}} - LT_{y,s}^{iter}}$ and $\beta_{y,s,n_{y,s}^{iter}+1}^{iter+1} = \sqrt{u_{y,s,n_{y,s}^{iter}}} - \alpha_{y,s,n_{y,s}^{iter}}^{iter+1} \cdot u_{y,s,n_{y,s}^{iter}}$

- For $p > n_{y,s}^{iter} + 1$ (i.e. p = $n_{y,s}^{iter} + 2$, $n_{y,s}^{iter} + 3$, …, $NP_{y,s}^{iter}$), set $u_{y,s,p}^{iter+1} = u_{y,s,p-1}^{iter}$, $\alpha_{y,s,p}^{iter+1} = \alpha_{y,s,p-1}^{iter}$ and $\beta_{y,s,p}^{iter+1} = \beta_{y,s,p-1}^{iter}$.

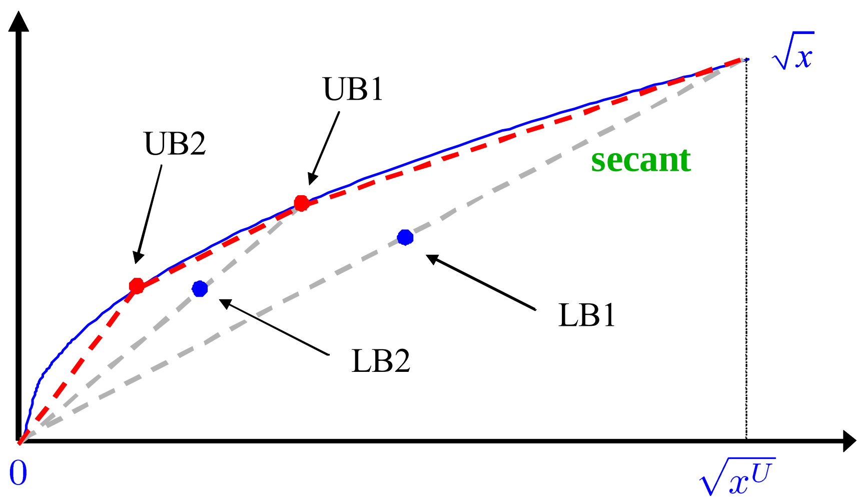
then, set $iter = iter + 1$ and go to Step 2.



(a) Iteration 1 – replace the square root terms with their secant, the optimal solution of the MILP provides a lower bound LB1, and the upper bound can be obtained from function evaluation or from solving the reduced NLP.



(b) Iteration 2 – add a new interval based on the optimal solution of the upper bounding problem in Iteration 1, and consider a two-piece linear approximation of the square root terms



(c) Iteration 3 – add another new interval based on the optimal solution of the upper bounding problem in the previous iteration

Figure 5 Branch-and-refine algorithm based on successive piece-wise linear approximations

6. Case Studies

In order to illustrate the application of the proposed model and the performance of the proposed solution strategies we consider three case studies. All the computational experiments 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.[23] The MILP problems are solved using CPLEX 12 and the reduced MINLP problems in Step 2 of the branch-and-refine algorithm are solved with MINLP solver DICOPT. We use

DICOPT as the convex MINLP solver and the global optimizer used in the computational experiments is BARON 8.1.5. The optimality tolerances of DICOPT, BARON and the proposed branch-and-refine algorithm are all set to 10^{-3} and optimality margins of solving the piecewise linear approximation MILP model (P4) and the reduced NLP model (P3) are both 10^{-6}.

Case study 1: four customer case

In the first case study, we consider a four-customer cluster of an industrial gas supply chain, of which the network structure and the monthly mean demand rates of the first year are given in Figure 6. All the customers need to size their tanks. In the three year horizon we consider a 15% annual demand growth rate for all customers, and the standard deviation of uncertain daily demand is considered as 1/3 of the daily mean demand. Other major input data for the case study are given in Tables 1-2.

Although all the four customers are included in the network at time zero, some of them may terminate the contract in a certain future year:

- N14 will not terminate the contract by the end of Year 3
- N15 has a 30% chance of terminating the contract in Year 1
- N18 has a 40% chance of terminating the contract in Year 2
- N21 has a 50% chance of terminating the contract in Year 3

We assume that each event is independent from the others, so eight scenarios are generated for this case study. The detailed network structure and TSP distances to visit all the customers once for each scenario in each year and the probability of each scenario are given in Figure 1 and Table 3.

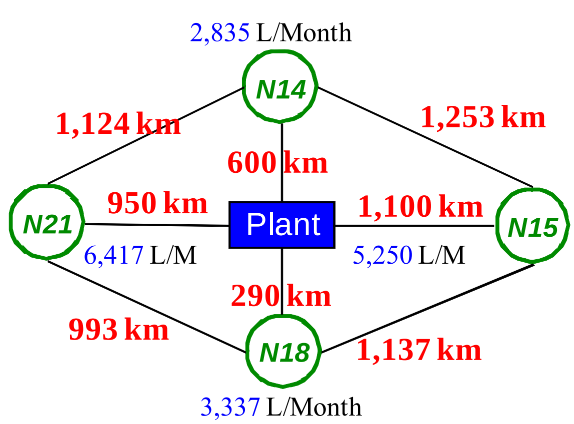


Figure 6. Case study 1 – four customer industrial gas supply chain

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 Network structure of each scenario in each year for case study 1

Scenario/Year		Customers in the network	TSP distance ($TSP_{y,s}$)	Scenario probability ($prob_s$)
S1	Year 1	N14, N15, N18, N21	4,507km	21%
	Year 2	N14, N15, N18, N21	4,507km	
	Year 3	N14, N15, N18, N21	4,507km	
S2	Year 1	N14, N18, N21	3,007km	9%
	Year 2	N14, N18, N21	3,007km	
	Year 3	N14, N18, N21	3,007km	
S3	Year 1	N14, N15, N18, N21	4,507km	14%
	Year 2	N14, N15, N21	4,427km	
	Year 3	N14, N15, N21	4,427km	
S4	Year 1	N14, N15, N18, N21	4,507km	21%
	Year 2	N14, N15, N18, N21	4,507km	
	Year 3	N14, N15, N18	3,281km	
S5	Year 1	N14, N15, N18, N21	4,507km	14%
	Year 2	N14, N15, N21	4,427km	
	Year 3	N14, N15	2,506km	
S6	Year 1	N14, N18, N21	3,007km	9%
	Year 2	N14, N18, N21	3,007km	
	Year 3	N14, N18	1,780km	
S7	Year 1	N14, N18, N21	3,007km	6%
	Year 2	N14, N21	2,247km	
	Year 3	N14, N21	2,247km	
S8	Year 1	N14, N18, N21	3,007km	6%
	Year 2	N14, N21	2,247km	
	Year 3	N14	0	

We consider 17 instances with service level α ranging from 50% to 98% for all customers and the corresponding service level parameter z_α ranging from 0 to 2.07, where z_α is a standard normal deviate such that $\Pr(z \le z_\alpha) = \alpha$. We note that all the 17 instances have similar problem structure – the only difference among them is the service level parameter. As these instances are solved sequentially with service level ranging from 50% to 98%, an instance's solution provides a good starting point for solving the next instance. We solved these 17 instances with DICOPT, BARON and the proposed branch-and-refine algorithm (CPLEX+DICOPT). The original MINLP model for stochastic continuous approximation includes 264 binary variables, 3,300 continuous variables and 4,344 constraints. DICOPT returned "infeasible" for all the instances, presumably due to the numerical difficulty arising from the square root terms. One way to overcome this issue is to add a sufficiently small number ε to each square root term. Please refer to our earlier work[22] for details. The global optimizer BARON took 70,354 CPU seconds for solving all the 17 instances (in average 4,138 CPUs/instance) to an optimality gap of 10^{-6}. With the proposed branch-and-refine algorithm using CPLEX and DICOPT, we solved all the 17 instances to global optimality with a total of 285 CPU seconds (in average 17 CPUs/instance). Clearly, the proposed global optimization algorithm demonstrated much better performance than the general purpose commercial MINLP solvers.

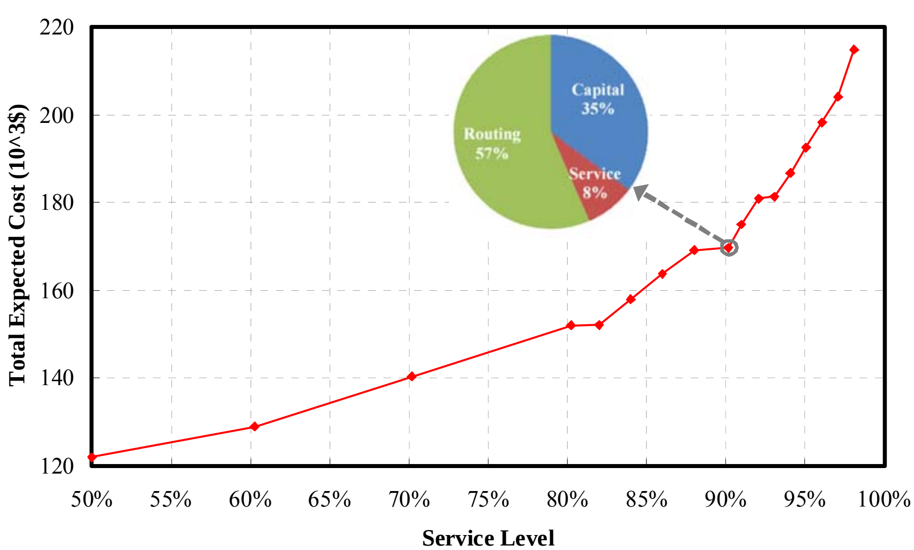


Figure 7. Pareto curve for Case study 1 (total expected cost vs. service level)

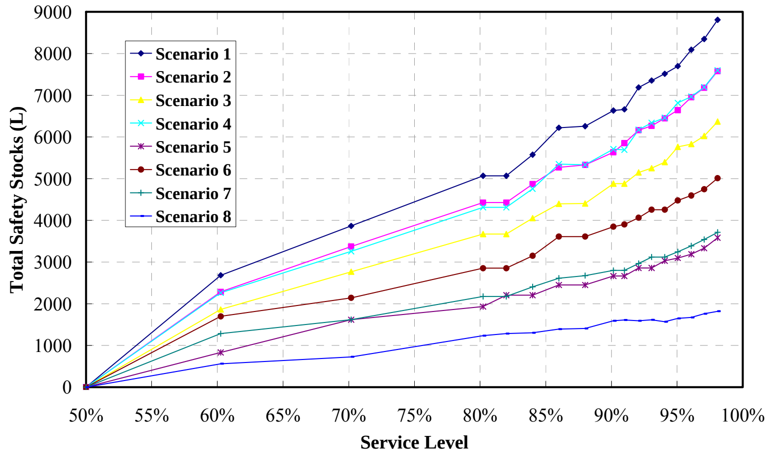


Figure 8. Scenario total safety stocks of the 3rd year under different service levels for Case study 1

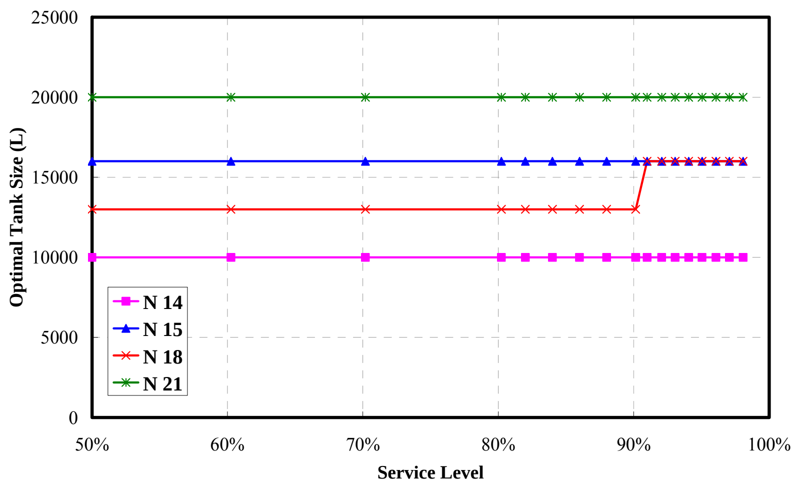


Figure 9. Optimal tank sizing decisions under different service levels for Case study 1

The optimal solutions of these 17 instances are given in Figures 7-9. Figure 7 shows the Pareto optimal curve between the total expected cost and the service level. As the service level increases from 50% to 98%, the total expected cost increases from \$121,965 to \$214,866. Thus, a higher service level implies higher total expected cost. In particular, when the service level increases from 80% to 82% and from 88% to 90%, the total expected cost only increases by \$40 and \$170, respectively. Therefore, setting the service level to 82% or 90% might be good choices in terms of balancing economics and service level. The cost breakdown for the 90% service level case is also given in Figure 7. We can see that the summation of capital and service costs is close to the estimated routing cost. The pie chart reveals the tradeoff between distribution cost and costs for installing and maintaining the storage tanks.

Figure 8 shows the scenario total safety stock levels (summation of the safety stocks of all the customers) for the third year under different specifications of the service level. We can see a similar trend that the higher service level we maintain, the more cost it will be for each scenario. Due to different network structures and different numbers of customers, different scenarios have different costs and they can be generally classified into four groups: Scenario 1 includes all the four customers in all the three years, so it has the highest scenario cost; Scenario 8 has the fewest number of customers throughout the planning horizon, so it has the lowest scenario cost; Scenario 2, 3 and 4 have more customers than Scenario 5, 6 and 7, so their costs are between the ones for Scenario 1 and Scenario 8.

Figure 9 depicts how the optimal tank sizing decisions, which are first stage decisions independent of scenarios, change as the service level increases from 50% to 98%. It is interesting to see that the optimal tank sizing decisions for all the customers do not change when the service level increases from 50% to 90%. The result in turn suggests that optimal tank sizing decisions are relatively robust in terms of service level. When the service level increases above 90%, the optimal tank size for customer N18 increases from 13,000L to 16,000L. This is because higher service level requires more safety stocks, and thus a larger tank is needed.

Case study 2: eight customer case

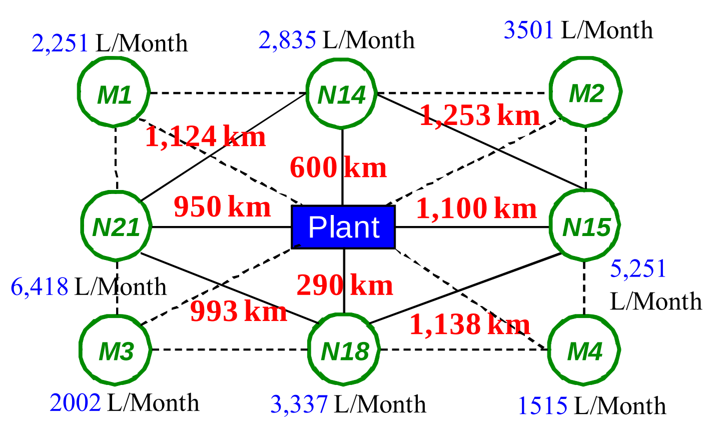


Figure 10. Case study 2 – eight customer industrial gas supply chain

In the second case study, we consider an eight customer cluster of an industrial gas supply chain as shown in Figure 10. Customers N14, N15, N18 and N21 are existing customers, and M1-M4 are potentially new customers that may join the network in

Year 2. All the customers need to size their tanks. The monthly demand rates of the first year for all customers are also given in Figure 10. We consider a 15% annual demand growth rate for all customers throughout the planning horizon, and the standard deviation of uncertain daily demand is considered as 1/3 of the daily mean demand. Other major input data for this case study are the same as those given in Tables 1-2.

Although all the four customers are included in the network at time zero, some of them may terminate the contract in a future year:

- N14 will not terminate the contract by the end of Year 2
- N15 has a 30% chance of terminating the contract in Year 1
- N18 has a 10% chance of terminating the contract in Year 2
- N21 has a 20% chance of terminating the contract in Year 2
- M1 has a 25% chance of signing a new contract starting in Year 2
- M2 has a 35% chance of signing a new contract starting in Year 2
- M3 has a 20% chance of signing a new contract starting in Year 2
- M4 has a 50% chance of signing a new contract starting in Year 2
- At most one of M1, M2, M3 and M4 will sign the contract

We assume each event is independent of others, so 40 scenarios are generated for this case study (note that at most only one new customer will join the network). The detailed network structure for each scenario in each year and the probabilities of scenarios are given in Table 4.

Table 4 Network structure of each scenario in each year for case study 2

Scenario	Year	Customers in the network	TSP distance ($TSP_{y,s}$)	Scenario probability ($prob_s$)
S1	Year 1	N14, N15, N18, N21	4,507km	16.1446%
	Year 2	N14, N15, N18, N21	4,507km	
S2	Year 1	N14, N15, N18, N21	4,507km	6.9191%
	Year 2	N14, N15, N18	3,281km	
S3	Year 1	N14, N15, N18, N21	4,507km	1.7938%
	Year 2	N14, N15, N21	4,427km	
S4	Year 1	N14, N15, N18, N21	4,507km	4.0361%
	Year 2	N14, N15	2,505km	
S5	Year 1	N14, N18, N21	3,007km	0.4485%
	Year 2	N14, N18, N21	3,007km	
S6	Year 1	N14, N18, N21	3,007km	1.7298%
	Year 2	N14, N18	1,780km	
S7	Year 1	N14, N18, N21	3,007km	0.7688%
	Year 2	N14, N21	2,247km	
S8	Year 1	N14, N18, N21	3,007km	0.1922%

	Year 2	N14	0	
S9	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, N21, M1	4,507km 4,934km	5.3815%
S10	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, M1	4,507km 4,642km	2.3064%
S11	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N21, M1	4,507km 4,934km	0.5979%
S12	Year 1 Year 2	N14, N15, N18, N21 N14, N15, M1	4,507km 4,339km	1.3454%
S13	Year 1 Year 2	N14, N18, N21 N14, N18, N21, M1	3,007km 3,433km	0.1495%
S14	Year 1 Year 2	N14, N18, N21 N14, N18, M1	3,007km 3,142km	0.5766%
S15	Year 1 Year 2	N14, N18, N21 N14, N21, M1	3,007km 2,674km	0.2563%
S16	Year 1 Year 2	N14, N18, N21 N14, M1	3,007km 1,900km	0.0641%
S17	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, N21, M2	4,507km 4,954km	8.6932%
S18	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, M2	4,507km 3,728km	3.7257%
S19	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N21, M2	4,507km 4,874km	0.9659%
S20	Year 1 Year 2	N14, N15, N18, N21 N14, N15, M2	4,507km 2,953km	2.1733%
S21	Year 1 Year 2	N14, N18, N21 N14, N18, N21, M2	3,007km 4,632km	0.2415%
S22	Year 1 Year 2	N14, N18, N21 N14, N18, M2	3,007km 3,405km	0.9314%
S23	Year 1 Year 2	N14, N18, N21 N14, N21, M2	3,007km 4,360km	0.4140%
S24	Year 1 Year 2	N14, N18, N21 N14, M2	3,007km 2,200km	0.1035%
S25	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, N21, M3	4,507km 4,754km	4.0361%
S26	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, M3	4,507km 4,642km	1.7298%
S27	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N21, M3	4,507km 4,737km	0.4485%
S28	Year 1 Year 2	N14, N15, N18, N21 N14, N15, M3	4,507km 4,625km	1.0090%
S29	Year 1 Year 2	N14, N18, N21 N14, N18, N21, M3	3,007km 3,254km	0.1121%
S30	Year 1 Year 2	N14, N18, N21 N14, N18, M3	3,007km 3,142km	0.4324%
S31	Year 1 Year 2	N14, N18, N21 N14, N21, M3	3,007km 2,715km	0.1922%
S32	Year 1	N14, N18, N21	3,007km	0.0480%

	Year 2	N14, M3	2,604km	
S33	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, N21, M4	4,507km 4,760km	16.1446%
S34	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N18, M4	4,507km 3,533 km	6.9191%
S35	Year 1 Year 2	N14, N15, N18, N21 N14, N15, N21, M4	4,507km 4,737 km	1.7938%
S36	Year 1 Year 2	N14, N15, N18, N21 N14, N15, M4	4,507km 2,958 km	4.0361%
S37	Year 1 Year 2	N14, N18, N21 N14, N18, N21, M4	3,007km 4,632km	0.4485%
S38	Year 1 Year 2	N14, N18, N21 N14, N18, M4	3,007km 3,405km	1.7298%
S39	Year 1 Year 2	N14, N18, N21 N14, N21, M4	3,007km 4,609km	0.7688%
S40	Year 1 Year 2	N14, N18, N21 N14, M4	3,007km 2,830km	0.1922%

Similarly, we consider 17 instances with service level α ranging from 50% to 98% and the corresponding service level parameter z_α ranging from 0 to 2.07. All the 17 instances are solved with DICOPT, BARON and the proposed branch-and-refine algorithm (CPLEX+DICOPT). The original MINLP model for stochastic continuous approximation includes 816 binary variables, 12,516 continuous variables and 16,236 constraints. DICOPT returned "infeasible" for all instances, and the global optimizer BARON cannot return any feasible solution after running for 7 days. With the proposed branch-and-refine algorithm using CPLEX and DICOPT, we solved all the 17 instances to global optima with a total of 1,770 CPU seconds (on average 104 CPUs/instance) under an optimality tolerance of 10^{-6}. It requires at most 6 iterations to solve each instance. As an example, the instance under 82% service level needs 6 iterations to reach the optimality tolerance. The upper and lower bounds of each iteration for this instance are shown in Figure 11. As the iteration number increases, the upper bound decreases and the lower bound increases until the optimality margin is reached. The lower bounding MILP problem in the last iteration has the maximum problem size, including 1,296 binary variables, 12,996 continuous variables and 167,356 constraints. The computational results show that for this problem the proposed global optimization algorithm is much more computationally efficient than the general purpose commercial MINLP solvers.

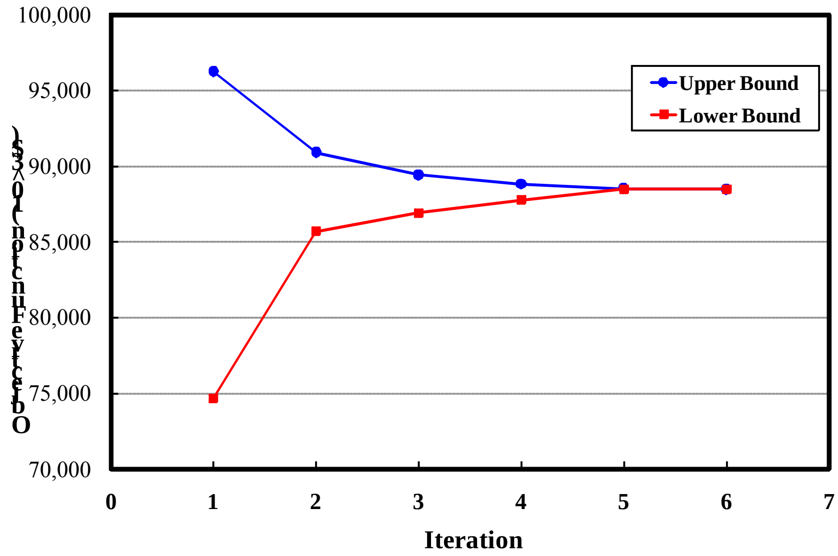


Figure 11. Bounds of each iteration of the proposed branch-and-refine algorithm for case study 2 under 82% service level

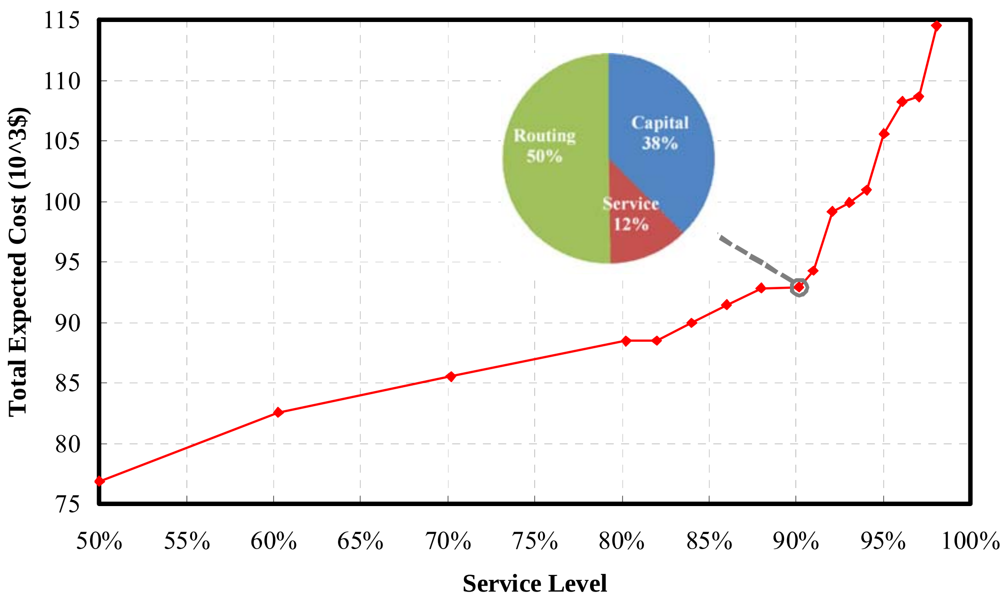


Figure 12. Pareto curve for Case study 2 (total expected cost vs. service level)

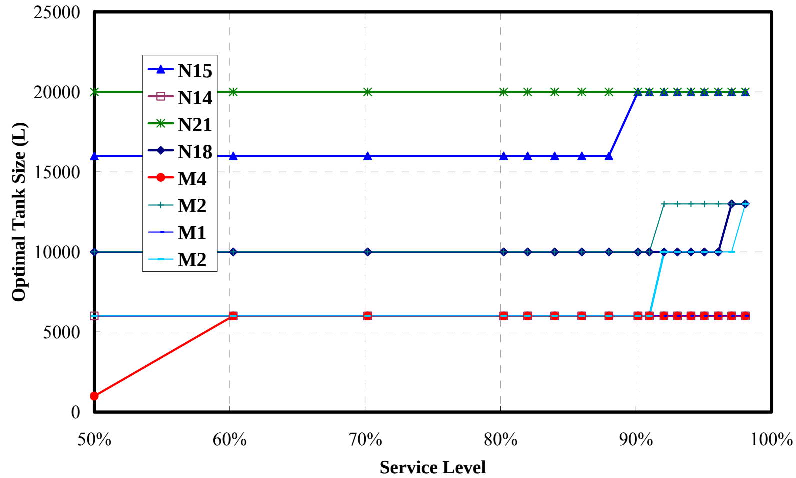


Figure 13. Optimal tank sizing decisions under different service levels for Case study 2

The optimal solutions of these 17 instances are given in Figures 12 and 13. Figure 12 shows the Pareto optimal curve between the total expected cost and the service level. We can see that as the service level increases from 50% to 98%, the total expected cost increases from $76,861 to $114,515. Thus, a higher service level implies more total expected cost. Figure 13 depicts how the optimal tank sizing decisions, which are the first stage decisions independent of scenarios, change as the service level increases from 50% to 98%. We can see that the optimal tank sizes mainly depend on the customer demands and locations. As the service level increases, the required amount of safety stocks also increases, and thus, the optimal tank sizes may stay unchanged or increase. The cost breakdown for the 90% service level case is also given in Figure 12. The pie chart also reveals the tradeoff between distribution cost and costs for installing and maintaining the storage tanks.

Case study 3: large scale instances with 200 customers

In the last case study, we consider a large-scale industrial gas supply chain with 200 customers. A one year planning horizon is considered and all the customers are new and need to size their tanks. As we can see from the previous two case studies, the proposed branch-and-refine algorithm is more efficient for this problem than the commercial general purpose MINLP solvers. Thus, we only use the proposed global optimization algorithm for this case study.

The data provided in Tables 1-2 are used for the three instances in this case study. Due to the large number of customers, we randomly generate their locations and demand rates. All the customer locations are generated in a $400\text{km}\times 400\text{km}$ square following a uniform distribution, and the plant is located in the center of this square. The detailed locations of the customers and plant are given in Figure 14. The TSP distances to visit all the customers once (not including the plant) for different scenarios and years are obtained with Concorde TSP Solver[24] through its NEOS interface[25] with CPLEX 12. The Concorde TSP solver is quite computationally 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:

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

Note that we take the absolute values of the normal distribution so that the monthly demand rates are always higher than 100 L/month. Although the normal distribution is unbounded, the maximum monthly demand rate we obtained from the sampling is 5,783 L/month. We consider a 15% annual demand growth rate for all customers throughout the planning horizon, and the standard deviation of uncertain daily demand is considered as 1/3 of the daily mean demand.

This example was designed for the case that a group of customers belonging to the same organization, who will decide to terminate the existing contract or signing a contract for the entire group, i.e. a group of customers might join or leave the network at the same time in some situation. We consider three organizations that have master contracts with the vendor and have branches in different locations. If any of the organizations decides to terminate the contract with the vendor, all its locations will have its supply finished and tanks pulled out. Hence, we consider three scenarios in this case study. Their probability and network structure are given in Table 5. Specifically, we consider 160 customers are "existing customers" that will not terminate the contract in the coming year, another 20 customers are "potentially lost customers" that may terminate the contract in Year 1, and the last 20 customers are the "potentially new customers", who might sign the contract and join the distribution network in Year 1. Scenario 1, with 40% chance, includes a network consisting of "existing customers" and "potentially lost customers"; Scenario 2, with 30%, is for a network consisting of only "existing customers", i.e. the "potentially lost customers" terminate their contract in Year 1; and Scenario 3, with 30% chance, has a distribution network with all the customers, i.e. those "potentially new customers" sign the contract and join the distribution network in Year 1.

Table 5 Network structure of each scenario in each year for case study 1

Scenario	Customers in the network	Total number of customers	TSP distance ($TSP_{y,s}$)	probability ($prob_s$)
S1	"existing customers" & "potentially lost customers"	180	3,914km	40%
S2	only "existing customers"	160	3,083km	30%
S3	"existing customers", "potentially lost customers" & "potentially new customers"	200	4,335km	30%

We consider 17 instances with service level α ranging from 50% to 97% and the corresponding service level parameter z_α ranging from 0 to 1.89. All the 12 instances are solved with the proposed branch-and-refine algorithm (CPLEX+DICOPT). It took 19,783 CPUs to solve all the 12 instances (in average 1,649 CPUs/instance). Each instance requires 4 to 6 iterations depending on the specified service level. The maximum size of the lower bounding MILP problem (at the 6th iteration) includes 2,439 discrete variables, 12,623 continuous variables and 14,150 constraints.

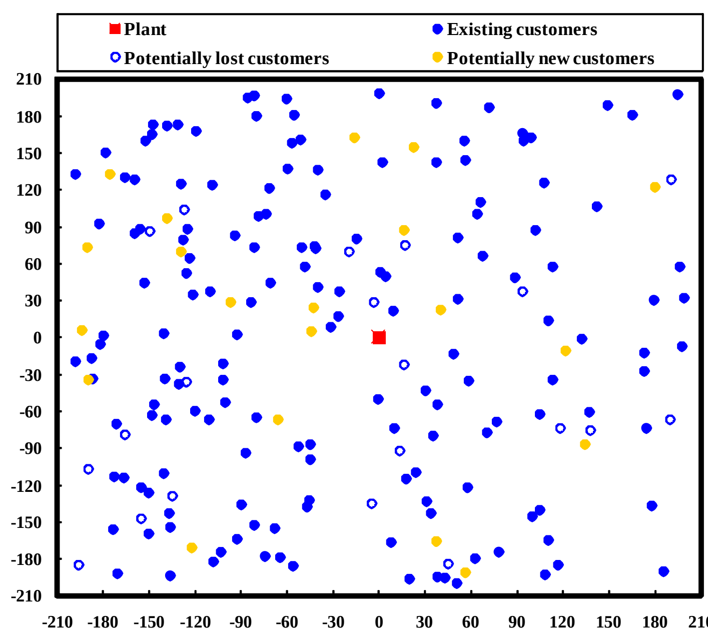


Figure 14. Location map of the 200 customers in the industrial gas supply chain

of Case study 3

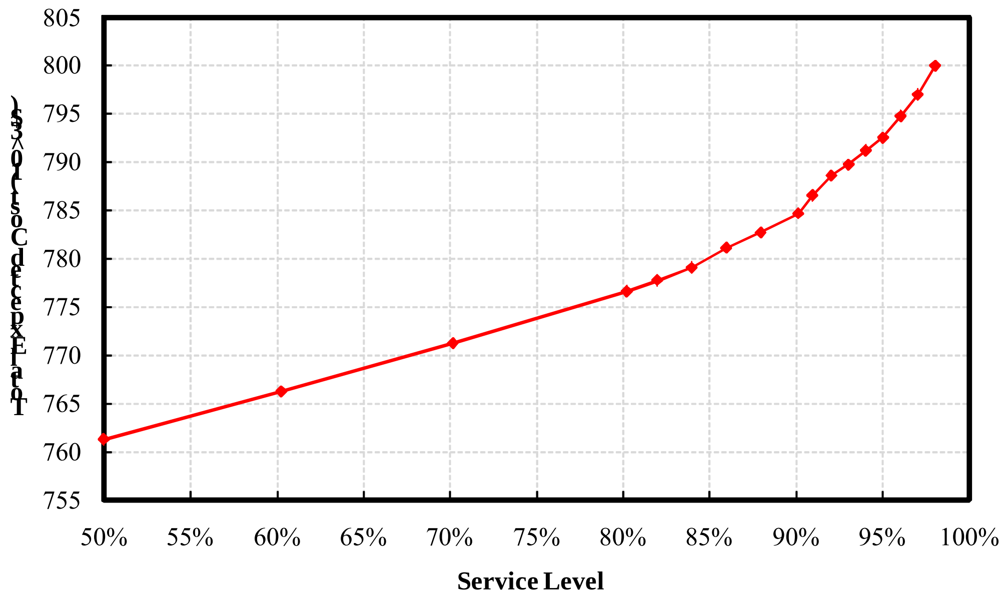


Figure 15. Pareto optimal curves for the 200 customer industrial gas supply chain in Case study 3

The Pareto curves for this case is given in Figures 15. Due to the one year planning horizon and the relatively small customer demand standard deviations, the total expected costs only increases by around $40,000 as the service level increases from 50% to 97%. We can see a similar trend as in the previous case studies that higher service level requires higher total expected cost.

This case study illustrates the application of the proposed stochastic continuous approximation method and the effectiveness of the branch-and-refine algorithm for solving large-scale problems.

Conclusion

In this paper, we have developed a computational framework to deal with industrial gas inventory-distribution planning under uncertain demand and customer presence. The framework consists of an upper level stochastic continuous approximation model and a lower level detailed routing model. The stochastic continuous approximation model is an MINLP problem that captures uncertainty of demand fluctuation and customer presence by incorporating a stochastic inventory model and a two-stage stochastic programming framework. This model simultaneously optimizes the tank sizing decisions, safety stock levels and estimated vehicle routing costs. To efficiently solve this stochastic MINLP problem for large-scale problems, we have proposed an efficient branch-and-refine algorithm based on successive piece-wise linear approximations. Detailed operational decisions in the inventory-distribution planning can be obtained by solving the lower level routing model after fixing the solutions from the stochastic continuous approximation model. Three case studies were presented to demonstrate the applicability of the proposed model. Computational experiments on large-scale industrial gas supply chains with up to 200 customer show that the proposed algorithm is much more efficient than commercial MINLP solvers for solving the stochastic continuous approximation problems.

Acknowledgment

The authors acknowledge the financial support from Praxair, Inc., Pennsylvania Infrastructure Technology Alliance and National Science Foundation under Grant No. DMI-0556090 and No. OCI-0750826. Fengqi You is partially supported by the U.S. Department of Energy under contract DE-AC02-06CH11357.

Appendix: clustering-based heuristic for detailed routing problem

There are a number of methods to improve the computational efficiency of solving the detailed vehicle routing problems.[15] In this appendix, we present a clustering-based heuristic that is easy to implement in practice.

In the first step, we create all the possible clusters such that the maximum number of customers allowed in a cluster is 4. Note that the maximum number of customers included in a cluster is problem-dependent. For the problem addressed in this work, we find out that each truck trip visits at most 4 customers in most cases. Thus, we consider each cluster has at most 4 customers.

For example if we have 7 customers {1, 2, 3, 4, 5, 6, 7} and customers 1–5 cannot be clustered with customers 6–7, then we have the following clusters: {1}, {2}, {3}, {4}, {5}, {6}, {7}, {1, 2}, {1, 3}, {1,4}, {1, 5}, {2, 3}, {2, 4}, {2, 5}, {3, 4}, {3, 5}, {4, 5}, {6, 7}, {1, 2, 3}, {1, 2, 4}, {1, 2, 5}, {1, 3, 4}, {1, 3, 5}, {1, 4, 5}, {2, 3, 4}, {2, 3, 5}, {2, 4, 5}, {3, 4, 5}, {1, 2, 3, 4}, {1, 2, 3, 5}, {1, 2, 4, 5}, {1, 3, 4, 5}, {2, 3, 4, 5}.

The round-trip distances (RTD) for each cluster can be calculated by the traveling salesman distance between the plant and all the customers included in the cluster. For instance, assume that a delivery is made to all customers in the cluster {1, 2, 3}. Figure 16 shows all the possible ways to travel:

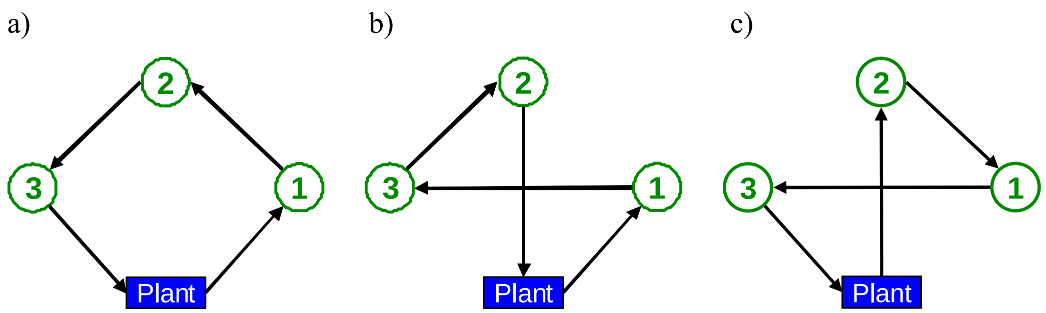


Figure 16. All the possible routes for delivery for a three customer cluster

If the shortest RTD is obtained by the routing b), then for the cluster {1, 2, 3}, we can assume that the RTD is the one calculated by the method b).

From the set of all possible clusters formed, we select the clusters that include all the customers and lead to the shortest RTDs from the plants. Therefore, for the customers 1 – 7 described above, a set of clusters that includes all the customers and is the shortest RTD could be: {1, 4, 5}, {2, 3}, {6, 7}.

After the clustering, we solve the detailed routing problem for each of the clusters formed above, i.e., consider the routing for some or all the customers in {1, 4, 5}, {2, 3}, and {6, 7} separately. In this way, we are able to capture the location and volume synergies for large-scale routing problems.

The clusters were selected on the basis of shortest total RTD. However, note that shortest RTD does not necessarily mean that we would have the lower total distribution cost than any other clusters. Hence, we have to solve the routing problem for another combination of clusters that includes all the 7 customers. We have to continue this analysis until we have found the clusters that lead to the lowest total costs. For example, if we have the total cost versus the combination of clusters as Figure 17, where the third set of clusters leads to the smallest total cost. It is also clear that other clusters lead to larger costs and hence we do not need to evaluate more than six clusters in this example. Figure 18 summarizes the entire heuristic method.

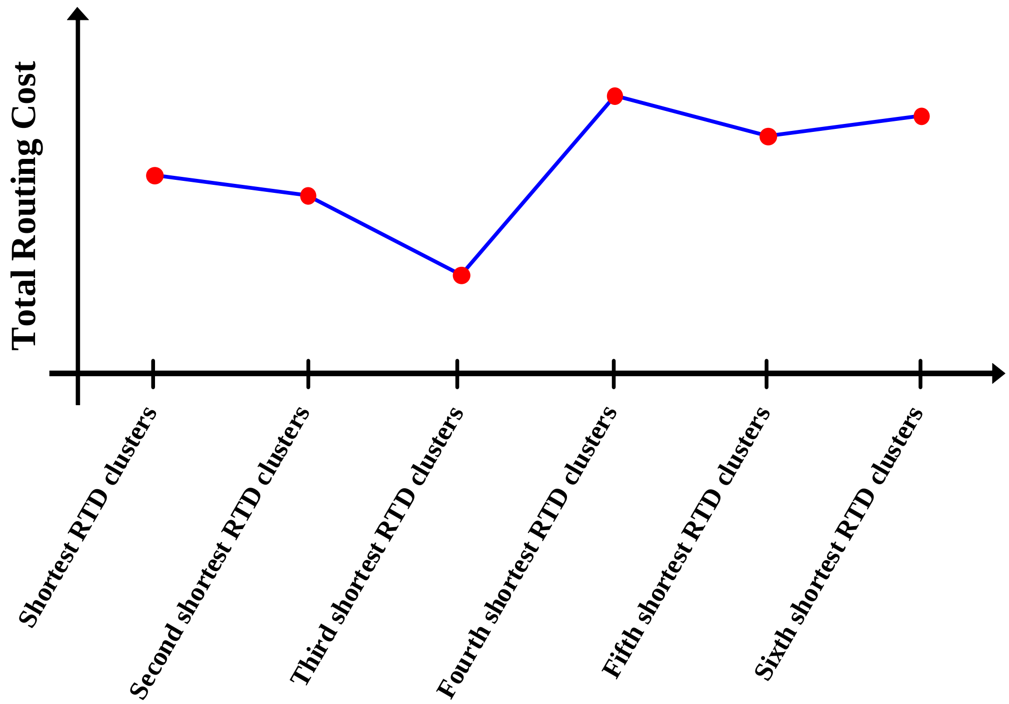


Figure 17. Total cost changes as the combination of clusters changes

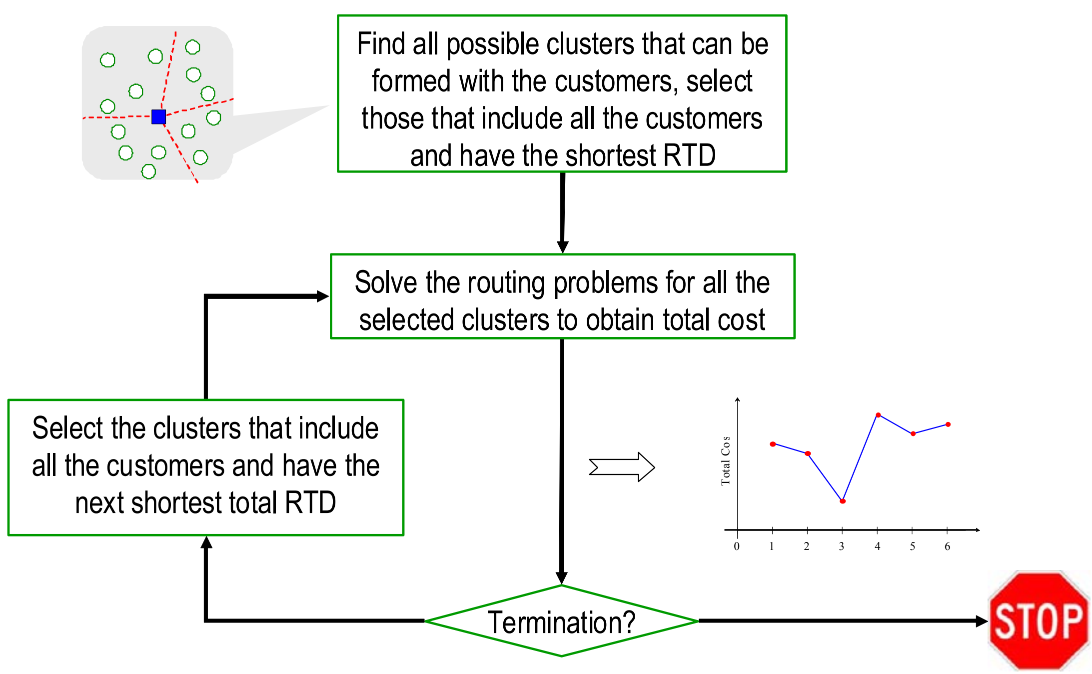


Figure 18. Algorithmic framework of the clustering-based heuristic for solving the detailed routing problem

Nomenclature

Sets/Indices

i	set of tank sizes
j	set of truck sizes
k	set for binary representation of integers
n	set of customers
s	set of scenarios
y	set of years
$n \in N_{y,s}$	subset of customers included in the network of scenario s and year y

Parameters

$Ccap_i$	capital cost of tank of size i
Cd_j	delivery cost of truck j
$Cout_{n,y}$	outage cost for customer n in year y
$Cser_i$	service cost of tank of size i
ck_j	Delivery cost per kilometer traveled of truck j
$dem_{n,y}$	demand customer n in year y
$demc_{n,y}$	monthly demand rate of customer n in year y
dep	depreciation period in years
$espace_n$	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
$ot_{i,n}$	1 if tank size i originally installed at customer n
rr_n	distance between the plant and the customer n
$speed$	average truck traveling speed in km per hr
T_i^L, T_i^U	lower and upper bounds for tank of size i
$tsize_n$	1 if tank of customer n is sized
$TSP_{y,s}$	traveling sales man distance of all the customers (exclude the plant)

	included in the distribution network of scenario s in year y
$Vtruck_j$	full transportation capacity of truck j
$wacc$	working capital discount factor
z_α	service level parameter, the standard normal deviate of α

Binary Variables (0-1)

$et_{i,n}$	1 if customer n has tank of size i installed in extra space; 0 otherwise
$Ix_{k,y,s}$	0-1 variable for the binary representation of the number of replenishments in scenario s and year y ($x_{y,s}$)
$tru_{j,y,s}$	1 if truck j is selected for replenishmsnt in scenario s and year y
$yt_{i,n}$	1 if customer n has tank of size i installed; 0 otherwise

Continuous Variables (0 to $+\infty$)

$capcost_s$	capital cost of scenario s
$ccapic_{y,}$	effective capacity of truck for the replenishments in scenario s, year y
$crot_{y,s}$	approximated routing cost of scenario s, year y
$cunit_{y,s}$	unit transportation cost of scenario s, year y
$distcost_s$	distribution cost of scenario s
$E[Cost]$	total expected cost
$LT_{y,s}$	Replenishment lead time of scenario s, year y
$mrt_{y,s}$	minimum routing distance to visit all customers in scenario s, year y
$servcost_s$	service cost of scenario s
$safety_{n,y,s}$	safety stock level in scenario s and year y for customer n
$seg_{y,s}$	auxiliary variable, for groups of all the customers
$Tccapic_{y,s}$	reciprocal of $ccapic_{y,s}$
$Trp_{n,y,s}$	total delivery amount from plant to customer n in scenario s and year y
$Vend_{n,y,s}$	inventory level of customer n at the end of scenario s, year y
Vl_n, Vu_n	min and max volume of tank at customer n
$Vm_{n,y,s}$	maximum inventory level of customer n in scenario s, year y
$Vzero_{n,y,s}$	initial volume at customer n in scenario s and year y
$winv_{n,y,s}$	maximum working inventory of customer n in scenario s, year y
$x_{y,s}$	number of replenishment in year y in scenario s

Auxiliary Variables (0 to $+\infty$)

$LTIx_{k,y,s}$	auxiliary variable for the product of $Ix_{k,y,s}$ and $LT_{y,s}$
$LTIx1_{k,y,s}$	auxiliary variable for linearization
$mrIx_{k,y,s}$	auxiliary variable for the product of $Ix_{k,y,s}$ and $mrt_{y,s}$
$mrIx1_{k,y,s}$	auxiliary variable for linearization
$mrItru_{j,k,y,s}$	auxiliary variable for the product of $tru_{,j,y,s}$ and $mrIx_{k,y,s}$
$mrItru1_{j,k,y,s}$	auxiliary variable for linearization
$TruSeg_{j,y,s}$	auxiliary variable for the product of $tru_{,j,y,s}$ and $seg_{y,s}$
$TruSeg1_{j,y,s}$	auxiliary variable for linearization
$wIx_{n,k,y,s}$	auxiliary variable for the product of $Ix_{k,y,s}$ and $winv_{n,y,s}$
$wIx1_{n,k,y,s}$	auxiliary variable for linearization

References

1. You, F.; Pinto, J. M.; Capón, E.; Grossmann, I. E.; Megan, L., Optimal Distribution-Inventory Planning of Industrial Gases: I. Fast Computational Strategies for Large-Scale Problems. *Industrial & Engineering Chemistry Research* **2010**, Submitted.

2. Liu, M. L.; Sahinidis, N. V., Optimization of Process Planning Under Uncertainty. *Industrial & Engineering Chemistry Research* **1996,** 35, 4154.

3. Neiro, S. M. S.; J.M., P., Lagrangean Decomposition applied to multiperiod planning of petroleum refineries under uncertainty. *Latin American Applied Research* **2006,** 36, 213-220.

4. Neiro, S. M. S.; Pinto, J. M., Multiperiod optimization for production planning of petroleum refineries. *Chemical Engineering Communications* **2005,** 192, 62-88.

5. Subrahmanyam, S.; Pekny, J. F.; Reklaitis, G. V., Design of Batch Chemical Plants under Market Uncertainty. *Industrial & Engineering Chemistry Research* **1994,** 33, 2688-2701.

6. 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.

7. Petkov, S. B.; Maranas, C. D., Design of Single Product Campaign Batch Plants under Demand Uncertainty. *AIChE Journal* **1998,** 44, 896.

8. Wellons, H. S.; Reklaitis, G. V., The design of multiproduct batch plants under uncertainty with staged expansion. *Computers & Chemical Engineering* **1989,** 13, 115.

9. You, F.; Grossmann, I. E., Design of Responsive Supply Chains under Demand Uncertainty. *Computers & Chemical Engineering* **2008,** 32, (12), 2839-3274.

10. 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.

11. 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.

12. 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.
13. Birge, J. R.; Louveaux, F., *Introduction to Stochastic Programming*. Springer-Verlag: New York, 1997.
14. Zimmermann, H. J., An application-oriented view of modeling uncertainty. *European Journal of Operational Research* **2000,** 122, 190-198.
15. Laporte, G., What You Should Know about the Vehicle Routing Problem. *Naval Research Logistics* **2007,** 54, 811-819.
16. Zipkin, P. H., *Foundations of Inventory Management*. McGraw-Hill: Boston, MA, 2000.
17. Zheng, Y. S., On properties of stochastic inventory systems. *Management Science* **1992,** 38, 87-103.
18. 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.
19. Glover, F., Improved Linear Integer Programming Formulations of Nonlinear Integer Problems. *Management Science* **1975,** 22, (4), 455-460.
20. Bergamini, M. L.; Grossmann, I. E.; Scenna, N.; Aguirre, P., An improved piecewise outer-approximation algorithm for the global optimization of MINLP models involving concave and bilinear terms. *Computers & Chemical Engineering* **2008,** 32, 477-493.
21. Padberg, M. W., Approximating separable nonlinear functions via mixed zero-one programs. *Operations Research Letters* **2000,** 27, 1-5.
22. 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.
23. Rosenthal, R. E., GAMS - A User's Manual. In GAMS Development Corp.: Washington, DC, 2008.
24. Applegate, D.; Bixby, R. E.; Chvátal, V.; Cook., W. J. *Concorde TSP Solver (http://www.tsp.gatech.edu/concorde)*, 2005.
25. http://www-neos.mcs.anl.gov/neos/solvers/co:concorde/TSP.html. (Access on Feb. 8, 2010),