Integrated Multi-Echelon Supply Chain Design with Inventories Under Uncertainty: MINLP Models, Computational Strategies

Fengqi You and Ignacio E. Grossmann
Dept. of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

DOI 10.1002/aic.12010
Published online September 15, 2009 in Wiley InterScience (www.interscience.wiley.com).

We address in this article a problem that is of significance to the chemical industry, namely, the optimal design of a multi-echelon supply chain and the associated inventory systems in the presence of uncertain customer demands. By using the guaranteed service approach to model the multi-echelon stochastic inventory system, we develop an optimization model to simultaneously determine the transportation, inventory, and network structure of a multi-echelon supply chain. The model is an MINLP with a nonconvex objective function including bilinear, trilinear, and square root terms. By exploiting the properties of the basic model, we reformulate this problem as a separable concave minimization program. A spatial decomposition algorithm based on the integration of Lagrangean relaxation and piecewise linear approximation is proposed to obtain near global optimal solutions with reasonable computational expense. Examples for specialty chemicals and industrial gas supply chains with up to 15 plants, 100 potential distribution centers, and 200 markets are presented. © 2009 American Institute of Chemical Engineers *AIChE J,* 56: 419–440, 2010
Keywords: supply chains, mixed-integer programming, inventory control, Lagrangean decomposition, industrial gases

Introduction

A recent report released by US Census Bureau shows that the total value of inventory in United States is around \$1.4 trillion (≈10% of US GDP),[1] of which around 24% are contributed by the chemical process industry. Another report by SmartOps shows that more than 50% of these inventories are inefficient.[2] Thus, there is a great economic incentive to optimize the inventories across process supply chains.[3,4] To accomplish this objective, the main challenge is how to effectively integrate inventory management with network design for multi-echelon process supply chains, so that decisions on locations to stock the inventory and the associated amount of inventories can be determined simultaneously to minimize costs. The integration is nontrivial for multi-echelon supply chains and their associated inventory systems in the presence of uncertain customer demands.[5] This problem has not been addressed before.

Although inventory management is a very important problem for the process industry, most of the models in the chemical engineering literature consider inventory management and supply chain network design separately. On the other hand, there are related works on supply chain optimization that take into account the inventory costs, but consider inventory issues coarsely without detailed inventory management policy.[6–10] In these models, the safety stock level is given as a parameter and usually treated as a lower bound of the total inventory level[11–19] or it is considered as an inventory target that would lead to some penalty costs if violated.[20] This approach cannot optimize the safety stock levels, especially when considering demand uncertainty. Thus, it can only provide an approximation of the inventory cost and may lead to suboptimal solutions.

Research on integrated supply chain network design and stochastic inventory management is relatively new. Most of the existing literature focuses on "single stage" inventory

Correspondence concerning this article should be addressed to I. E. Grossmann at grossmann@cmu.edu

© 2009 American Institute of Chemical Engineers

structure integrated with supply chain design, but without extending it to ''multi-echelon'' inventory systems addressed in this work. Daskin et al.[21] and Shen et al.[22] present a joint location-inventory model, which extends the classical uncapacitated facility location model to include nonlinear working inventory and safety stock costs for a two-stage supply chain network, so that decisions on the installation of distribution centers (DCs) and the detailed inventory replenishment decisions are jointly optimized. To simplify the problem, inventories in the retailers are neglected and they also assume that all the DCs have the same constant replenishment lead time, and the demand at each customer has the same variance-to-mean ratio. With the same assumptions, Ozsen et al.[23] have extended the model to consider capacitated limits in the DCs. Their work is further extended by Ozsen et al. (submitted for publication) to compare with the cases in which customers restrict to single sourcing and the case in which customers allow multi-sourcing. Another extension is given by Sourirajan et al.[24] in which the assumption of identical replenishment lead time is relaxed, whereas the assumption on demand uncertainty is still enforced. Recently, You and Grossmann[6] proposed a mixed-integer nonlinear programming (MINLP) approach to study a more general model based on the one developed by Daskin et al.[21] and Shen et al.[22] relaxing the assumption on identical variance-to-mean ratio for customer demands.

The objective of this work is to develop optimization models and solution algorithms to address the problem of joint multi-echelon supply chain network design and inventory management. By using the guaranteed service approach to model the time delays in the material flows in the multi-echelon inventory system,[7–10,25–30] we capture the stochastic nature of the product deliveries at each stage (echelon) of the supply chain and develop an equivalent deterministic optimization model. The model determines the supply chain design decisions (including the locations of DCs, assignments of customer demand zones to DCs, assignments of DCs to plants), shipment levels from plants to the DCs and from DCs to customers, and inventory decisions such as pipeline inventory and safety stock in each node of the supply chain network. The model also captures risk-pooling effects[31] by consolidating the safety stock inventory of downstream nodes to the upstream nodes in the multi-echelon supply chain. The model is first formulated as a mixed-integer nonlinear program (MINLP) with a nonconvex objective function and then reformulated as a separable concave minimization program after exploiting the properties of the basic model. To solve the problem efficiently, a tailored spatial decomposition algorithm based on Lagrangean relaxation and piece-wise linear approximation is developed to obtain near global optimal solutions within 1% optimality gap with modest CPU times. Several computational examples for industrial gases supply chains and specialty chemical supply chains are presented to illustrate the application of the model and the performance of the proposed algorithm.

This article includes several novel features. First, we explicitly model the multi-echelon inventory system of a process supply chain under demand uncertainty and use a holistic approach to simultaneously optimize the supply chain design decisions and multi-echelon stochastic inventory management decisions. To the best of our knowledge, the integration of multi-echelon stochastic inventory with supply chain design has not been addressed before, as most of the literature on joint supply chain design and inventory management considers only single-echelon inventory system. Capturing the multi-echelon inventory structure allows us to consider variable replenishment lead times and inventory allocation issues, both of which can significantly improve the decision-making across the process supply chains. Second, we develop a novel and efficient global optimization algorithm to obtain solutions within 1% global optimality gap for the resulting large scale nonconvex MINLP instances with thousands of discrete and continuous variables, and more than one million constraints. The proposed algorithm is based on effective integration of piece-wise linear approximation and Lagrangean relaxation that as far as we know has not been considered before. The entire solution process requires only a mixed-integer linear programming (MILP) solver without the need of an NLP solver. The efficient global optimization algorithm that allows large-scale solution and does not rely on nonlinear solvers is therefore another contribution of this work.

The outline of this article is as follows. Some basic concepts of inventory management with risk pooling and the guaranteed service model for multi-echelon inventory system are presented in Section Multi-Echelon Inventory. The problem statement is given in Section 3. In Section Model Formulation, we introduce the joint multi-echelon supply chain design and inventory management model. Two small illustrative examples are given in Section 5. To solve the large scale problem, an efficient decomposition algorithm based on Lagrangean relaxation and piecewise linear approximation is presented in Section Solution Algorithm. Our computational results and analysis are given in Section 7. We conclude this article in the last section.

Multi-Echelon Inventory Model

In this section, we briefly review some inventory management models that are related to the problem addressed in this work. Detailed discussion about single-stage and multi-echelon inventory management models are given by Zipkin[5] and Graves and Willem,[27] respectively.

Single-stage inventory model under base-stock policy

There are many control policies for single-stage inventory systems, such as base stock policy, (s,S) policy, (r, Q) policy, etc.[5] Among these policies, the periodic review base stock policy is the most widely used in the practice of inventory control. The reason is based on two facts. As shown by Federgruen and Zipkin,[32] the base stock policy is optimal for single-stage inventory system facing stationary demand. For multi-echelon inventory systems, the base stock policy, although not necessarily optimal, has the advantage of being simple to implement and close to the optimum.[33] Before introducing the multi-echelon inventory model, we first review the single-stage base stock policy, which is the common building block for most of the multi-echelon inventory models.

Figure 1 shows the inventory profile for a given product in a stocking facility operated under the periodic review base stock policy. As can be seen, the inventory level

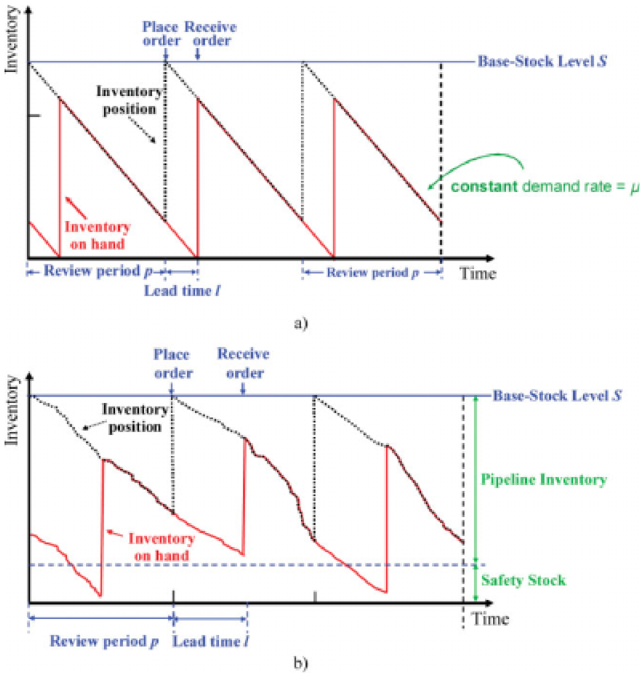

Figure 1. Inventory profile under base stock policy.

(a) Deterministic demand case. (b) Under demand uncertainty. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

decreases due to the customer demand and increases when replenishments arrive. Under the periodic review base stock policy, inventory is reviewed at the beginning of each review period and the difference between a specified base stock level and the actual inventory position (on-hand inventory plus in-process inventory minus backorders) is ordered for replenishment. It is interesting to note that the well-known continuous review (r, Q) policy can be treated as a special case of base stock policy with base stock level equal to $r + Q$, where r is the reorder point of inventory level and Q is the order quantity.

As under the base stock policy, the lengths of the review period and the replenishment lead time are determined exogenously, the only control variable is the base stock level. To determine the optimal base stock level for a single-stage inventory system, let us denote the review period as p, the replenishment lead time as l, and the average demand at each unit of time as μ. Recall that the inventory position is the total material in the system (on-hand plus on-order), and we start each review period with the same inventory position S, which is the base stock level. We must wait p units of time to review the inventory position again and place an order for replenishment, and then the order will take another l units of time to arrive (Figure 1). Therefore, the inventory in the system at the beginning of a review period should be large enough to cover the demand over review period p plus the replenishment lead time l, i.e., the optimal base stock level should be $\mu(p + l)$ if demand is deterministic as in Figure 1a.

Under demand uncertainty, we need more inventory (safety stock) to hedge against stockout before we get a chance to reorder as in Figure 1b. The accepted practice in this field is to assume a normal distribution of the demand, although of course other distribution functions can be specified. If the demand at each unit of time is normally distributed with mean μ and standard deviation σ, due to the property of normal distribution, the demand over review period p and the replenishment lead time l is also normally distributed with mean $\mu(p + l)$ and variance σ^2 $(p + l)$ (standard deviation $\sigma\sqrt{p+l}$). It is convenient to measure safety stock in terms of the number of standard deviations of demand denoted as safety stock factor, λ. Then the optimal base stock level is given by $S = \mu(p+l) + \lambda\sigma\sqrt{p+l}$, as shown in Figure 2.

We should note that if α is the *Type I service level* (the probability that the total inventory on hand is more than the demand) used to measure of service level, the safety stock factor λ corresponds to the α-quantile of the standard normal distribution, i.e., $\Pr(x \leq \lambda) = \alpha$.

Risk pooling effect

For single-echelon inventory system with multiple stocking locations, Eppen[31] proposed the "risk pooling effect," which states that significant safety stock cost can be saved by grouping in one central location the demand of multiple stocking locations. In particular, Eppen considers a single period problem with N retailers and one supplier. Each retailer i has uncorrelated normally distributed demand with mean μ_i and standard deviation σ_i. The review periods and replenishment lead times for all these retailers are the same and given as p and l, respectively. All the retailers guarantee the same *Type I service level* with the same safety stock factor λ. Eppen compared two operational modes of the N-retailer system: decentralized mode and centralized mode. In the decentralized mode, each retailer orders independently to minimize its cost. As in this mode, the optimal safety stock in retailer i corresponding to the safety stock factor λ is $\lambda\sqrt{p+l}\sigma_i$,[31] the total safety stock in the system is given by $\lambda\sqrt{p+l}\sum_{i=1}^{N}\sigma_i$. In the centralized mode, all the retailers are considered as a whole and a single quantity is ordered for replenishment, so as to minimize the total expected cost of the entire system. As in the centralized mode, all the

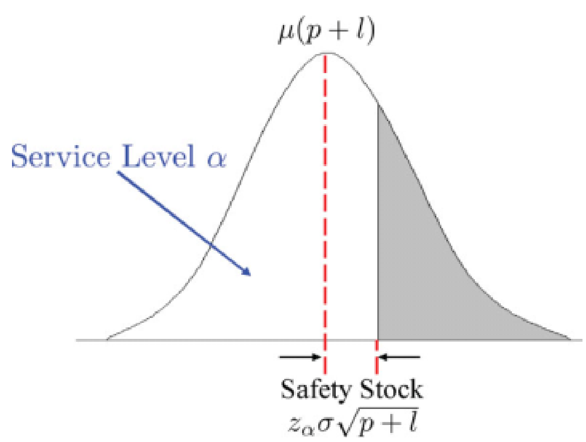

Figure 2. Safety stock level for normally distributed demand over lead time.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

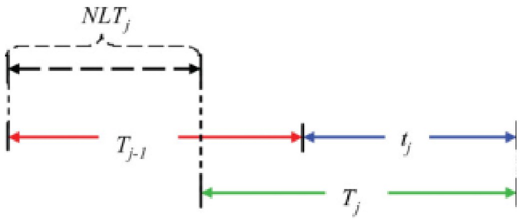

Figure 3. Timing relationship in guaranteed service approach.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

retailers are grouped and the demand at each retailer follows a normal distribution $N(\mu_i,\ \sigma_i^2)$, the total uncertain demand of the entire system during the order lead time will also follow a normal distribution with mean $(p\ +\ l)\sum_{i=1}^{N}\mu_i$ and standard deviation $\sqrt{p+l}\sqrt{\sum_{i=1}^{N}\sigma_i^2}$. Therefore, the total safety stock of the distribution centers in the centralized mode is given by $\lambda\sqrt{p+l}\sqrt{\sum_{i=1}^{N}\sigma_i^2}$, which is less than $\lambda\sqrt{p+l}\sum_{i=1}^{N}\sigma_i$. Eppen's simple model illustrates the potential saving in safety stock costs because of risk pooling. For example, consider a single-echelon inventory system with 100 retailers. Each retailer has uncorrelated normal distributed demand with identical mean μ and standard deviation σ. Thus, the total safety stock of this system is $100z\sigma\sqrt{p+l}$ under decentralized mode, and only $10z\sigma\sqrt{p+l}$ under centralized mode, i.e., 90% safety stocks in this system can be saved by risk pooling.

Guaranteed service model for multi-echelon inventory

One of the most significant differences between single-stage inventory system and multi-echelon inventory system is the lead time. For single-stage inventory system, lead time, which may include material handling time and transportation time, is exogenous and generally can be treated as a parameter. However, for a multi-echelon inventory system, lead time of a downstream node depends on the upstream node's inventory level and demand uncertainty, and thus the lead time and internal service level are stochastic. Based on this fact, simply propagating the single-stage inventory model to multi-echelon system leads to suboptimal solutions because the optimization of a multi-echelon inventory system is usually nontrivial.[5]

There are two major approaches to model the multi-echelon inventory system, the stochastic service approach and the guaranteed service approach.[27] For a detailed comparison of these two approaches, see Graves and Willems[27] and Humair and Willems.[29] In this work, we choose the guaranteed service approach to model the multi-echelon inventory system.

Definitions in guaranteed service approach

The main idea of the guaranteed service approach is that each node j in the multi-echelon inventory system quotes a guaranteed service time T_j, by which this node will satisfy the demands of material flow from its downstream customers. That is, the customer demand at time t must be ready to be shipped by time $t\ +\ T_j$. The guaranteed service times for internal customers are decision variables to be optimized, whereas the guaranteed service time for the nodes at the last echelon (facing external customers) is an exogenous input, which can be treated as a parameter. Besides the guaranteed service time, we consider that each node has a given deterministic order processing time, t_j, which is independent of the order size. The order processing time, t_j, which includes material handling time, transportation time, and review period, represents the time from all the inputs that are ready until the outputs are available to serve the demand. Therefore, the replenishment lead time, which represents the time from when we place an order to when all the goods are received, can be determined by the guaranteed service time of the direct predecessor T_{j-1} (the time that predecessor requires to have the chemicals ready to be shipped) plus the order processing time t_j. As we can see from Figure 3, the net lead time of node j, NLT_j, which is the required time span to cover demand variation using safety stocks at node j, is defined as the difference between the replenishment lead time of this node and its guaranteed service time to its successor.[10] The reason is that not all the customer demand at time t for node j should be satisfied immediately, but only needs to be ready by time $t\ +\ T_j$. Thus, the safety stock does not need to cover demand variations over the entire replenishment lead time, but just the difference between the replenishment lead time and the guaranteed service time to the successors, i.e., the net lead time. Therefore, we can calculate net lead time with the following formula:

$$\mathrm{NLT}_j = T_{j-1} + t_j - T_j$$

where node $j - 1$ is the direct predecessor of node j. Note that this formula implies that if the service time quoted by node j to its successor nodes equals the replenishment lead time, T_{j-1} + t_j, i.e., $\mathrm{NLT}_j\ =\ 0$ as shown in Figure 4a, no safety stock is required in node j because all products are received from the predecessors and processed within the guaranteed service time, i.e., this node is operating in ''pull'' mode. If the guaranteed service time T_j is 0, i.e., $\mathrm{NLT}_j = T_{j-1} + t_j$ as shown in Figure 4b, the node holds the most safety stock because all the orders, once they are placed, are fulfilled immediately, i.e., this node is operating in ''push'' mode.

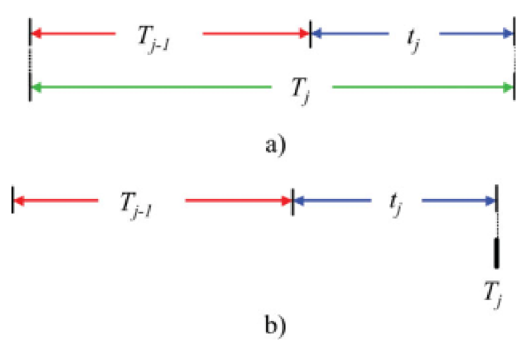

Figure 4. Extreme net lead times.

(a) Zero, no safety stock. (b) Largest, highest safety stock.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

$$\mathrm{PI}_j = t_j \mu_j$$

which is not affected by the coverage and guaranteed service time decisions.

Example of a series inventory system

To illustrate the main idea of guaranteed service approach, let us consider an example of a series inventory system with one plant, one DC, and one market as in Figure 5. In this example, the plant has a guaranteed service time of 2 days, which quantitatively represents the worst case supply uncertainty and production delay. It means that every order from the DC will be satisfied in at most 2 days, regardless of the uncertainty and delay from supply and production. The guaranteed service time of the market is zero, which means that every product in the market should be ready to be used immediately. Thus, the guaranteed service time of the plant is usually determined exogenously based on supply uncertainty and production delay, whereas the guaranteed service time of market is also exogenously determined by the market requirement.

In this way, the major decision variable of this system is the guaranteed service time of the DC. As shown in Figure 6a, the summation of plant's guaranteed service time and order processing time from plant to DC leads to the DC's worst case replenishment lead time, 5 days, which is the time span from the DC receives an order until the order is ready to be shipped out from the DC. Therefore, if the DC commits a guaranteed service time the same as the worst case replenishment lead time, 5 days, there is no need to hold safety stock in the DC because every order from the market requires at most 5 days to be fulfilled. In this case, the DC is operated under the "pull" mode and the net lead time is zero. Another extreme case is that the DC commits a zero guaranteed service time. In this case, the optimal safety stock should be able to cover the demand uncertainty over the worst case replenishment lead time because every order from the market should be satisfied immediately.

As the guaranteed service time increases from 0 to 5 days (the worst case replenishment lead time), the net lead time, which is the difference between worst case replenishment lead time and guaranteed service time, decreases from 5 days to 0, and consequently the optimal safety stock level decreases from the maximum level to zero, as in Figures 6b, c. Therefore, the guaranteed service time of the DC is an important decision variable that can adjust the optimal safety stock level in the DC.

Besides, the DC's guaranteed service time also affects the safety stock level of the market. As shown in Figure 7, the market's net lead time equals to the DC's guaranteed service time plus the order processing time from DC to customer. For instance, if the DC's guaranteed service time is 4 days, the market should hold sufficient safety stock to hedge the demand uncertainty over 6 days, which is the worst case replenishment lead time of the market. Thus, the major objective is to find out the optimal guaranteed service time, so as to reduce the total cost based on the adjusting the inventory allocation decisions between the DC and the market.

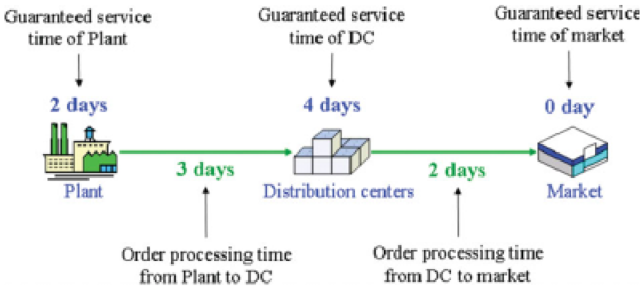

Figure 5. Motivating example of a series inventory system with a plant, a DC, and a market with fixed times to illustrate the timing relationships.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Optimal based stock level and inventory cost of guaranteed service approach

In the guaranteed service approach, each node in the multi-echelon inventory system is assumed to operate under a periodic review base stock policy with a common review period. Furthermore, demand over any time interval is also assumed to be normally distributed (mean μ_j and standard deviation σ_j for node j) and "bounded." This does not imply that demand can never exceed the bound, but this bound reflects the maximum amount of demand that a company wants to satisfy from safety stock. This interpretation is consistent with most practical applications that safety stocks are used to cope with regular demand variations that do not exceed a maximum reasonable bound. Larger demand variations are handled with extraordinary measures, such as spot-market purchases, rescheduling or expediting orders, other than using safety stocks. Thus, under these assumptions, each node sets its base stock so as to meet *all* the orders from its downstream customers within the guaranteed service time. That is, for node j there is an associated safety stock factor λ_j, which is given and corresponds to the maximum amount of demand that company wants to satisfy from safety stock. This yields the base stock level at node j:

$$S_j = \mu_j \mathrm{NLT}_j + \lambda_j \sigma_j \sqrt{\mathrm{NLT}_j}$$

This formula is similar but slightly different from the single-stage inventory model in terms of the expression for the lead time. Note that the review period has been taken into account as part of the order processing time and considered in the net lead time.

With the guaranteed service approach, the total inventory cost consists of safety stock cost and pipeline inventory cost. The safety stock of node j (SS_j) is given by the following formula as discussed earlier,

$$\mathrm{SS}_j = \lambda_j \sigma_j \sqrt{\mathrm{NLT}_j}$$

The expected pipeline inventory is the sum of expected on hand and on-order inventories. Based on Little's law,[34] the expected pipeline inventory PI_j of node j equals to the mean demand over the order processing time and is given by,

Problem Statement

We are given a potential supply chain (Figure 8) for a given chemical product that consists of a set of plants (or

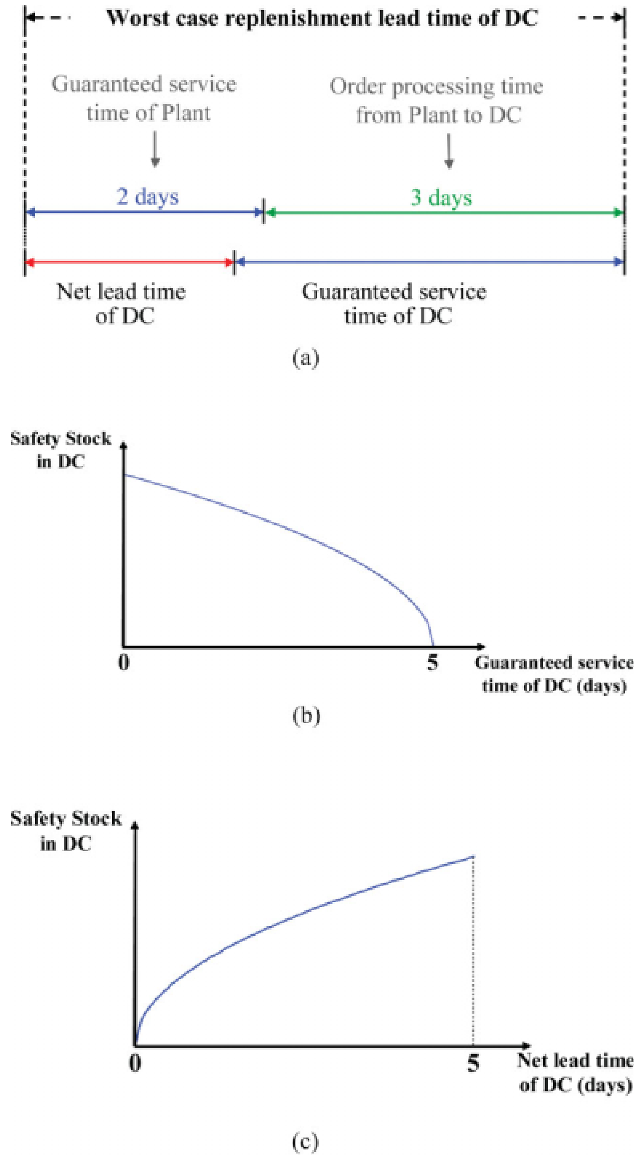

Figure 6. The relationship between times and safety stock level of the DC in the motivating example.

(a) Replenishment lead time, guaranteed service time and net lead time of the DC in the motivating example. (b) Conceptual relationship between safety stock level and guaranteed service time of DC. (c) Conceptual relationship between safety stock level and net lead time of DC. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

suppliers) $i \in I$, a number of candidate sites for distribution centers $j \in J$, and a set of customer demand zones $k \in K$, whose inventory costs should be taken into account. The customer demand zone can represent a distributor, a warehouse, a dealer, a retailer, or a wholesaler, which is usually a necessary component of the supply chain for specialty chemicals or advanced materials.[25,35] Alternatively, one might view the customer demand as the aggregation of a group of customers operated with vendor managed inventory (vendor takes care of customers' inventory), which is a common business model in the industrial gases industry and some chemical companies.[36,37]

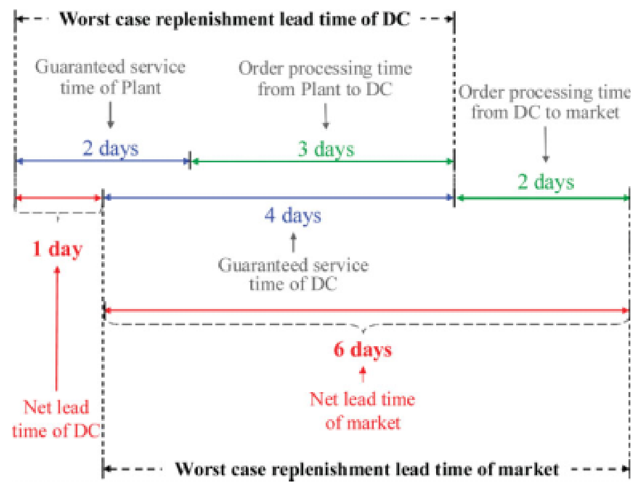

Figure 7. The timing relationship between plant, DC and market in the motivating example.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

In the given potential supply chain, the location of the plants, potential distribution centers, and customer demand zones are known and the distances between them are given. The investment costs for installing DCs are expressed by a cost function with fixed charges. Each retailer i has an uncorrelated normally distributed demand with mean μ_i and variance σ_i^2 in each unit of time. Single sourcing restriction, which is common in industrial gases supply chains[36,37] and specialty chemicals supply chains,[38] is employed for the distribution from plants to DCs and from DCs to customer demand zones. That is, each DC is only served by one plant, and each customer demand zone is only assigned to one DC to satisfy the demand. Linear transportation costs are incurred for shipments from plant i to distribution center j with unit cost $c1_{ij}$ and from distribution center j to customer demand zone k with unit cost $c2_{jk}$. The corresponding deterministic order processing times of DCs and customer demand zone, which includes the material handling time, transportation time, and inventory review period, are given by $t1_{ij}$ and $t2_{jk}$. The service time of each plant and the

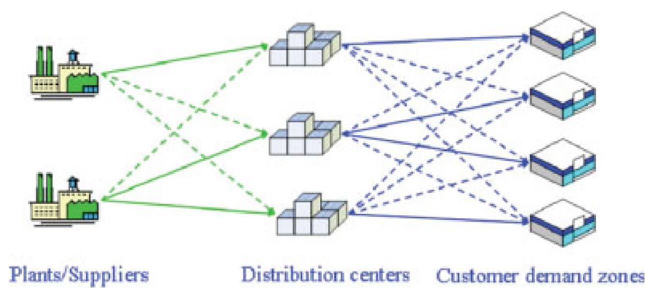

Figure 8. Supply chain network structure (three echelons).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

maximum service time of each customer demand zones are known. We are also given the safety stock factor for DCs and customer demand zones, $\lambda 1_j$ and $\lambda 2_k$, which correspond to the standard normal deviate of the maximum amount of demand that the node will satisfy from its safety stock. A common review period is used for the control of inventory in each node. Inventory costs are incurred at distribution centers and customers, and consist of pipeline inventory and safety stock, of which the unit costs are given.

The problem is then to determine how many distribution centers (DCs) to install, where to locate them, which plants to serve each DC, and which DCs to serve each customer demand zone, how long should each DC quote its service time, and what level of safety stock to maintain at each DC and customer demand zone so as to minimize the total installation, transportation, and inventory costs.

We should note that the guaranteed service time is only a variable for the DCs, and hence the net lead times are variables for the DCs and customer demand zones. Therefore, the problem corresponds effectively to a three-echelon supply chain with inventories under uncertainty.

Model Formulation

The joint multi-echelon supply chain design and inventory management model for a given chemical product can be formulated as a mixed-integer nonlinear program (MINLP). The definition of sets, parameters, and variables of the model are given latter.

Sets/Indices

I = set of plants (suppliers) indexed by i
J = set of candidate DC locations indexed by j
K = set of customer demand zones (wholesaler, regional distributor, dealer, retailer, or customers with vendor managed inventory) indexed by k

Parameters

$c1_{ij}$ = unit transportation cost from plant i to DC j
$c2_{jk}$ = unit transportation cost from DC j to customer demand zone k
f_j = fixed cost for installing a DC at candidate location j (annually)
g_j = variable cost coefficient for installing candidate DC j (annually)
$h1_j$ = unit inventory holding cost at DC j (annually)
$h2_k$ = unit inventory holding cost at customer demand zone k (annually)
R_k = maximum guaranteed service time to customers at customer demand zone k
SI_i = guaranteed service time of plant i
$t1_{ij}$ = order processing time of DC j if it is served by plant i, including material handling time of DC j, transportation time from plant i to DC j, and inventory review period
$t2_{jk}$ = order processing time of customer demand zone k if it is served by DC j, including material handling time of DC j, transportation time from DC j to customer demand zone k, and inventory review period
μ_k = mean demand at customer demand zone k (daily)
σ_k^2 = variance of demand at customer demand zone k (daily)
χ = days per year (to convert daily demand and variance values to annual costs)
$\theta 1_{ij}$ = unit cost of pipeline inventory from plant i to DC j (annual)
$\theta 2_{jk}$ = unit cost of pipeline inventory from DC j to customer demand zone k (annual)
$\lambda 1_j$ = safety stock factor of DC j
$\lambda 2_k$ = safety stock factor of customer demand zone k

Binary variables (0–1)

X_{ij} = 1 if DC j is served by plant i and 0 otherwise
Y_j = 1 if we install a DC in candidate site j and 0 otherwise
Z_{jk} = 1 if customer demand zone k is served by DC j and 0 otherwise

Continuous variables (0 to +∞)

L_k = net lead time of customer demand zone k
N_j = net lead time of DC j
S_j = guaranteed service time of DC j to its successive customer demand zones

Auxiliary variables (0 to +∞)

XZ_{ijk} = the product of X_{ij} and Z_{jk}
SZ_{jk} = the product of S_j and Z_{jk}
$SZ1_{jk}$ = auxiliary variable for linearization
NZ_{jk} = the product of N_j and Z_{jk}
$NZ1_{jk}$ = auxiliary variable for linearization
NZV_j = auxiliary variable for reformulation
$u1_{j,p_1}$ = auxiliary variable for piecewise linear approximation
$u2_{k,p_2}$ = auxiliary variable for piecewise linear approximation
$u3_{j,k,p_3}$ = auxiliary variable for piecewise linear approximation
W_j = objective function value of problem (APL$_j$)
$\widehat{W}_j$ = objective function value of problem (APLR$_j$)
$\overline{W}_j$ = objective function value of problem (APLP$_j$)

Auxiliary variables (0–1)

$v1_{j,p_1}$ = auxiliary variable for piecewise linear approximation
$v2_{k,p_2}$ = auxiliary variable for piecewise linear approximation
$v3_{j,k,p_3}$ = auxiliary variable for piecewise linear approximation

Objective Function

The objective of this model is to minimize the total supply chain design cost including the following items:

- Installation costs of DCs.
- Transportation costs from plants to DCs and from DCs to customer demand zones.
- Pipeline inventory costs in DCs and customer demand zones.
- Safety stock costs in DCs and customer demand zones.

The cost of installing a DC in candidate location j is expressed by a fixed-charge cost model that captures the economies of scale in the investment. The annual expected demand of DC j is $(\sum_{k\in K} \chi Z_{jk}\mu_k)$, which equals to the annual mean demand of all the customer demand zones $k \in K$ served by DC j. Hence, the cost of installing DC j consists of fixed cost f_j and variable cost $(g_j \sum_{k\in K} \chi Z_{jk}\mu_k)$, which is the product of variable cost coefficient and the expected demand of this DC in 1 year. Thus, the total installation costs of all the DCs is given by,

$$\sum_{j\in J} f_j Y_j + \sum_{j\in J}\left(g_j \sum_{k\in K} \chi Z_{jk}\mu_k\right) \tag{1}$$

The product of the annual mean demand of DC j $(\sum_{k\in K} \chi Z_{jk}\mu_k)$ and the unit transportation cost $(\sum_{i\in I} c1_{ij}X_{ij})$ between plant i and DC j yields the annual plant to DC transportation cost.

$$\sum_{i\in I}\sum_{j\in J}\left(c1_{ij}X_{ij}\sum_{k\in K}\chi Z_{jk}\mu_k\right) \tag{2}$$

Similarly, the product of yearly expected mean demand of customer demand zone k $(\chi\mu_k)$ and the unit transportation cost $(\sum_{j\in J} c2_{jk}Z_{jk})$ between customer demand zone k leads to the annual DC to customer demand zone transportation cost.

$$\sum_{j\in J}\sum_{k\in K} c2_{jk}\chi Z_{jk}\mu_k \tag{3}$$

Based on Little's law,[34] the pipeline inventory PI_j of DC j equals to the product of its daily mean demand $(\sum_{k\in K} Z_{jk}\mu_k)$ and its order processing time $(\sum_{i\in I} t1_{ij}X_{ij})$ which is given in terms of days. Thus, the annual total pipeline inventory cost of all the DCs is given by,

$$\sum_{i\in I}\sum_{j\in J}\left(\theta 1_j t1_{ij}X_{ij}\sum_{k\in K} Z_{jk}\mu_k\right) \tag{4}$$

where $\theta 1_j$ is the annual unit pipeline inventory holding cost of DC j.

Similarly, the total annual pipeline inventory cost of all the customer demand zones is given by,

$$\sum_{j\in J}\sum_{k\in K} \theta 2_k t2_{jk} Z_{jk}\mu_k \tag{5}$$

where $\theta 2_k$ is the annual unit pipeline inventory holding cost of customer demand zone k and $\sum_{k\in K} t2_{jk}Z_{jk}\mu_k$ is the pipeline inventory of customer demand zone k.

The demand at customer demand zone k follows a given normal distribution with mean μ_k and variance σ_k^2. Because of the risk-pooling effect,[31] the demand over the net lead time (N_j) at DC j is also normally distributed with a mean of $N_j\sum_{k\in J_k}\mu_k$ and a variance of $N_j\sum_{k\in J_k}\sigma_k^2$, in which J_k is the set of customer demand zones k assigned to DC j. Thus, the safety stock required in the DC at candidate location j with a safety stock factor $\lambda 1_j$ is $\lambda 1_j\sqrt{N_j}\cdot\sqrt{\sum_{k\in K}\sigma_k^2 Z_{jk}}$. Considering the annual inventory holding cost at DC j is $h1_j$, we have the total annual safety stock cost at all the DCs equals to

$$\sum_{j\in J}\lambda 1_j h1_j\sqrt{N_j}\cdot\sqrt{\sum_{k\in K}\sigma_k^2 Z_{jk}} \tag{6}$$

Similarly, the demand over the net lead time L_k of customer demand zones k is normally distributed with a mean of $L_k\mu_k$ and a variance of $L_k\sigma_k^2$. Thus, the annual safety stock cost at all the customer demand zones is given by,

$$\sum_{k\in K}\lambda 2_k h2_k\cdot\sigma_k\sqrt{L_k} \tag{7}$$

Therefore, the objective function of this model (total supply chain design cost) is given by

$$\begin{aligned}\min:\ &\sum_{j\in J} f_j Y_j + \sum_{j\in J}\left(g_j \sum_{k\in K} \chi Z_{jk}\mu_k\right)\\ &+\sum_{i\in I}\sum_{j\in J}\left(c1_{ij}X_{ij}\sum_{k\in K}\chi Z_{jk}\mu_k\right)+\sum_{j\in J}\sum_{k\in K} c2_{jk}\chi Z_{jk}\mu_k\\ &+\sum_{i\in I}\sum_{j\in J}\left(\theta 1_j t1_{ij}X_{ij}\sum_{k\in K} Z_{jk}\mu_k\right)+\sum_{j\in J}\sum_{k\in K} \theta 2_k t2_{jk} Z_{jk}\mu_k\\ &+\sum_{j\in J}\lambda 1_j h1_j\sqrt{N_j}\cdot\sqrt{\sum_{k\in K}\sigma_k^2 Z_{jk}}+\sum_{k\in K}\lambda 2_k h2_k\cdot\sigma_k\sqrt{L_k}\end{aligned} \tag{8}$$

where each term accounts for the DC installation cost, transportation costs of plants-DCs and DCs-customer demand zones, pipeline inventory costs of DCs and customer demand zones, and safety stock costs of DCs and customer demand zones.

Constraints

Three types of constraints are used to define the network structure. The first one is that if DC j is installed, it should be served by only one plant i. If the DC is not installed, it is not assigned to any plant. This can be modeled by,

$$\sum_{i\in I} X_{ij} = Y_j, \quad \forall j \tag{9}$$

The second constraint states that each customer demand zone k is served by only one DC,

$$\sum_{j\in J} Z_{jk} = 1, \quad \forall k \tag{10}$$

The third constraint is that if a customer demand zone k is served by the DC in candidate location j, the DC must exist,

$$Z_{jk} \le Y_j, \quad \forall j,k \tag{11}$$

Two constraints are used to define the net lead time of the DCs and customer demand zones. The replenishment lead time of DC j should be equal to the guaranteed service time (SI_i) of plant i, which serves DC j, plus the order processing time (t_{ij}). Note that the guaranteed service time (SI_i) of plant i is treated as parameters and represents the worst case supply uncertainty and production delay of plants. As each DC is served by only one plant, the replenishment lead time of DC j is calculated by $\sum_{i\in I}(SI_i + t_{ij}) \cdot X_{ij}$. Thus, the net lead time of DC j should be greater than its replenishment lead time minus its guaranteed service time to its successor customer demand zones, S_j, which is a variable.

$$N_j \geq \sum_{i\in I} (SI_i + t1_{ij}) \cdot X_{ij} - S_j, \quad \forall j \tag{12}$$

Similarly, the net lead time L_k of a customer demand zone k is greater than its replenishment lead time minus its maximum guaranteed service time, R_k, which is given by the nonlinear inequality

$$L_k \geq \sum_{j\in J} (S_j + t2_{jk}) \cdot Z_{jk} - R_k, \quad \forall k \tag{13}$$

Note that in this article we consider R_k as parameters, although this formulation can be extended to a bi-criterion optimization model by treating R_k as variables to simultaneously minimize the total cost and responsiveness.

Finally, all the decision variables for network structure are binary variables, and the variables for guaranteed service time and net lead time are non-negative variables.

$$X_{ij}, Y_j, Z_{jk} \in \{0,1\}, \quad \forall i,j,k \tag{14}$$

$$S_j \geq 0,\ N_j \geq 0, \quad \forall j \tag{15}$$

$$L_k \geq 0, \quad \forall k \tag{16}$$

MINLP Model

Grouping the parameters, we can rearrange the objective function and formulate the problem as the following mixed-integer nonlinear program (P0):

$$\text{Min}: \sum_{j\in J} f_j Y_j + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K} A_{ijk} X_{ij} Z_{jk} + \sum_{j\in J}\sum_{k\in K} B_{jk} Z_{jk} + \sum_{j\in J} q1_j \sqrt{N_j \sum_{k\in K} \sigma_k^2 Z_{jk}} + \sum_{k\in K} q2_k \sqrt{L_k} \tag{17}$$

$$\text{s.t.} \quad N_j \geq \sum_{i\in I} \bar{S}_{ij} \cdot X_{ij} - S_j, \quad \forall j \tag{12}$$

$$L_k \geq \sum_{j\in J} (S_j + t2_{jk}) \cdot Z_{jk} - R_k, \quad \forall k \tag{13}$$

$$\text{Constraints } (9)-(11),\ (14)-(16)$$

where

$$\bar{S}_{ij} = SI_i + t1_{ij}$$

$$A_{ijk} = \left(c1_{ij}\chi + \theta 1_j t1_{ij}\right) \cdot \mu_k$$

$$B_{jk} = \left(g_j\chi + c2_{jk}\chi + \theta 2_k t2_{jk}\right) \cdot \mu_k$$

$$q1_j = \lambda 1_j \cdot h1_j$$

$$q2_k = \lambda 2_k \cdot h2_k \cdot \sigma_k$$

Reformulated MINLP Model

The MINLP model (P0) has nonconvex terms, including bilinear and square root terms, in its objective function (17) and constraint (13). Because the bilinear terms are the products of two binary variables (such as $X_{ij} \cdot Z_{jk}$) or the products of a binary variable and a continuous variable (such as $N_j \cdot Z_{jk}$ and $S_j \cdot Z_{jk}$), they can be linearized by introducing additional variables.

To linearize the term ($X_{ij} \cdot Z_{jk}$) in the objective function (17), we first introduce a new continuous non-negative variable XZ_{ijk}. Then the product of X_{ij} and Z_{jk} can be replaced by this term XZ_{ijk} with the following constraints,[39]

$$XZ_{ijk} \leq X_{ij}, \quad \forall i,j,k \tag{19}$$

$$XZ_{ijk} \leq Z_{jk}, \quad \forall i,j,k \tag{20}$$

$$XZ_{ijk} \geq X_{ij} + Z_{jk} - 1, \quad \forall i,j,k \tag{21}$$

$$XZ_{ijk} \geq 0, \quad \forall i,j,k \tag{22}$$

where constraints (19), (20), and (22) ensure that if X_{ij} or Z_{jk} is zero, XZ_{ijk} should be zero, whereas constraint (21) ensures that if X_{ij} and Z_{jk} are both equal to one, XZ_{ijk} should be one.

The linearization of ($S_j \cdot Z_{jk}$) in constraint (13) requires two new continuous non-negative variable SZ_{jk} and $SZ1_{jk}$, and the following constraints,[39]

$$SZ_{jk} + SZ1_{jk} = S_j, \quad \forall j,k \tag{23}$$

$$SZ_{jk} \leq Z_{jk} \cdot S_j^U, \quad \forall j,k \tag{24}$$

$$SZ1_{jk} \leq (1 - Z_{jk}) \cdot S_j^U, \quad \forall j,k \tag{25}$$

$$SZ_{jk} \geq 0,\ SZ1_{jk} \geq 0, \quad \forall j,k \tag{26.1}$$

where constraints (24), (25), and (26.1) ensure that if Z_{jk} is zero, SZ_{jk} should be zero; if Z_{jk} is one, $SZ1_{jk}$ should be zero. Combining with constraint (23), we can have SZ_{jk} equivalent to the product of S_j and Z_{jk}.

Similarly, the product of N_j and Z_{jk} in the objective function (17) can be linearized as follows,

$$NZ_{jk} + NZ1_{jk} = N_j, \quad \forall j,k \tag{27}$$

$$NZ_{jk} \leq Z_{jk} \cdot N_j^U, \quad \forall j,k \tag{28}$$

$$NZ1_{jk} \leq (1 - Z_{jk}) \cdot N_j^U, \quad \forall j,k \tag{29}$$

$$NZ_{jk} \geq 0,\ \ NZ1_{jk} \geq 0, \quad \forall j,k \tag{26.2}$$

where NZ_{jk} and $NZ1_{jk}$ are two new continuous variables, and NZ_{jk} is equivalent to ($N_j \cdot Z_{jk}$).

The above linearizations introduce more constraints and variables, significantly reduce the number of nonlinear terms in the model (P0) and potentially reduce the computational effort. To further reduce the nonlinear terms in the objective function (17), the term ($N_j \cdot \sum_{k\in K} \sigma_k^2 \cdot Z_{jk}$) is replaced by a

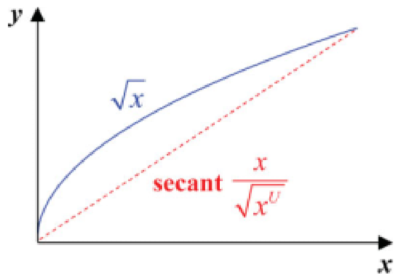

Figure 9. Secant of a univariate square root term.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

new nonnegative continuous variable NZV_j with the following constraint,

$$NZV_j = \sum_{k \in K} \sigma_k^2 \cdot NZ_{jk}, \quad \forall j \tag{30}$$

$$NZV_j \geq 0, \quad \forall j \tag{31}$$

Therefore, incorporating the above linearizations, we have the following reformulated MINLP model (P1).

$$\text{Min}: \sum_{j \in J} f_j Y_j + \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} A_{ijk} XZ_{ijk} + \sum_{j \in J} \sum_{k \in K} B_{jk} Z_{jk} + \sum_{j \in J} q1_j \sqrt{NZV_j} + \sum_{k \in K} q2_k \sqrt{L_k} \tag{17}$$

$$\text{s.t.} \ L_k \geq \sum_{j \in J} SZ_{jk} + \sum_{j \in J} t2_{jk} \cdot Z_{jk} - R_k, \quad \forall k \tag{32}$$

$$\text{Constraints } (9)-(11), \ (14)-(16), \ \text{and } (18)-(31)$$

Compared with model (P0), model (P1) is computationally more tractable because all the constraints in (P1) are linear and the only nonlinear terms are univariate concave terms in the objective function.

Initialization Process

The bounds of variables are quite important for nonlinear optimization problems. Tight upper bounds of the key continuous variables are given in the Appendix section. Besides variable bounds, a good initial solution is another key issue for numerical optimization of MINLP problems, as it provides a tight upper bound of the objective and consequently reduces the search space and computational time. As model (P1) has some special structure (linear constraints, univariate concave terms in the objective function), we can use a similar approach as in You and Grossmann[6] to obtain a "good" initial solution by solving a (MILP) approximation for initialization purposes.

As introduced by Falk and Soland[40] for a univariate square root term $\sqrt{x}$, in which the variable x has lower bound 0 and upper bound x^U, its secant $x/\sqrt{x^U}$ represents the convex envelope and provides a valid lower bound of the square root term as shown in Figure 9. As model (P1) is a minimization problem and all the constraints are linear, replacing all the univariate square root terms with their secants in the objective function (17) will lead to the MILP problem (P2) with a linear objective function,

$$\text{Min}: \sum_{j \in J} f_j Y_j + \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} A_{ijk} XZ_{ijk} + \sum_{j \in J} \sum_{k \in K} B_{jk} Z_{jk} + \sum_{j \in J} \frac{q1_j \cdot NZV_j}{\sqrt{NZV_j^U}} + \sum_{k \in K} \frac{q2_k \cdot L_k}{\sqrt{L_k^U}} \tag{34}$$

s.t. all the constraints of (P1)

As (P1) and (P2) have the same constraints, a feasible solution of (P2) is also a feasible solution of (P1). Furthermore, the optimal objective function value of (P2) is a valid lower bound of the optimal objective function value of (P1). Therefore, by solving the initialization MILP problem (P2), we can obtain a "good" initial solution for problem (P1) and a tight upper bound of the objective.

Note that based on this initialization procedure, a heuristic algorithm is to first solve problem (P2), then using its optimal solution as an initial point to solve problem (P1) with an MINLP solver that relies on convexity assumption (e.g., DICOPT, SBB). This algorithm can obtain "good" solutions quickly, although global optimality cannot be guaranteed (actually it provides a valid upper bound of the global optimal solution).

Illustrative Examples

To illustrate the application of the MINLP model (P1), we consider two small illustrative examples, an industrial liquid oxygen (LOX) supply chain example and an acetic acid supply chain example.

Example A: Industrial LOX supply chain

The first example is for an industrial LOX supply chain with two plants, three potential DCs, and six customers as given in Figure 10. In this supply chain, customer inventories are managed by the vendor, i.e., vendor-management-

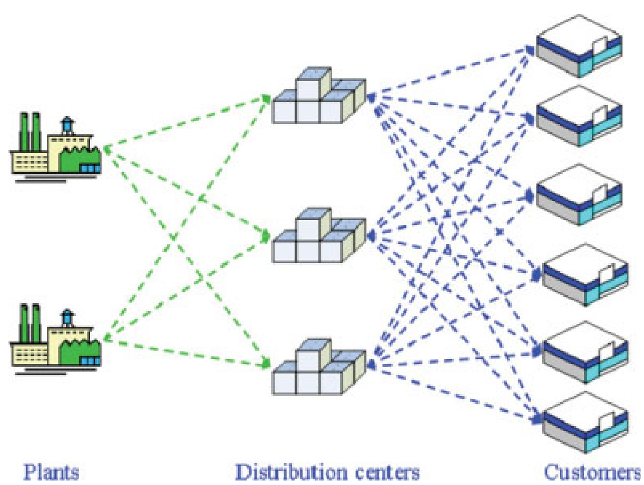

Figure 10. LOX supply chain network superstructure for Example A.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 1. Parameters for Demand Uncertainty for Example A

	Mean Demand μ_i (liters/day)	Standard Deviation σ_i (liters/day)
Customer 1	257	150
Customer 2	86	25
Customer 3	194	120
Customer 4	75	45
Customer 5	292	64
Customer 6	95	30

inventory (VMI), which is a common business model in the gas industry.[35–37] Thus, in this example inventory costs from DCs and customers should be taken into account in the total supply chain cost, and the joint multi-echelon supply chain design and inventory management model can be used to minimize total network design, transportation, and inventory costs.

In this instance, the annual fixed costs to install the three DCs (f_j) are \$10,000 per year, \$8,000 per year, and \$12,000 per year, respectively. The variable cost coefficient of installing a DC (g_j) at all the candidate locations is \$0.01 per liter/year. The safety stock factors for DCs ($\lambda 1_j$) and customers ($\lambda 2_k$) are the same and equal to 1.96, which corresponds to 97.5% service level if the demand is normally distributed. We consider 365 days in a year (χ). The guaranteed service times of the two plants (SI_i) are 2 days and 3 days, respectively. As the last echelon represents the customers, the guaranteed service time of customers (R_k) are set to 0. The data for demand uncertainty, order processing times, and transportation costs are given in Tables 1–5.

We consider two instances of this example. In the first instance, we consider zero holding cost for both pipeline inventory and safety stocks, i.e., the problem reduces to a supply chain network design problem without considering inventory cost. In the second instance, the pipeline inventory holding cost of LOX is \$6 per liter/year for all the DCs ($\theta 1_{ij}$) and customers ($\theta 2_{jk}$), and the safety stock holding cost is \$8 per liter/year for all the DCs ($h1_j$) and customers ($h2_k$).

Both the initialization MILP model (P2) and the MINLP model (P1) involve 27 binary variables, 124 continuous variables, and 256 constraints. As the problem sizes are relatively small, we solve model (P1) directly to obtain the global optimum with 0% optimality margin by using the BARON solver[41] with GAMS,[42] and the initialization MILP model (P2) is solved with GAMS/CPLEX. The resulting optimal supply chain networks with a minimum annual cost of \$108,329/year and the flow rate of each transportation link are given in Figure 11a. We can see that for the instance without considering inventory costs, all the three DCs are installed, and each of them serves two customers. The major trade-off is between DC installation costs and the transportation costs. For the instance taking into account the inventory costs, only two DCs are installed, and each serves three customers as seen in Figure 11b. The major trade-off is between DC installation costs, transportation costs, and inventory costs. The minimum annual cost in this instance is \$152,107/year with the variable cost items given in Figure 12. A comparison between these two instances suggests the importance of integrating inventory costs in the supply chain network design.

The optimal net lead times and guaranteed service times of DCs and customers for the design in Figure 11b are given in Figure 13. We can see in Figure 13a that in the path of material flow from Plant 2 to DC 1 and to Customers 1, 2, and 3, DC 1 has an optimal service time equal to the plant's service time plus the deterministic order processing time from Plant 2 to DC 1, i.e., 5 days. Note that this service time corresponds to the worst case response time from plant to DC, although the most optimistic case is that Plant 2 will fulfill the demand immediately and the minimum response time will be the deterministic order processing time, i.e., 2 days. We can observe similar phenomena in the path of material flow from Plant 1 to DC 3 and to Customers 4, 5, and 6 from Figure 13b. As both DC 1 and DC 3 have the optimal service times equal to the summation of plant's service time and the order processing time between plants and DCs, both of them have zero net lead time, i.e., do not hold any safety stock and work as pure "pull" systems. In Figures 14a, b, we can see that all the customers in the optimal solution have positive net lead times and hold sufficient safety stocks to ensure their guaranteed service times are zero, i.e., work as pure "push" system. This interesting phenomena shows that safety stock allocation of the multi-echelon LOX supply chain leads to a hybrid "push – pull" systems, in which each node either does not hold inventory or holds maximum required inventory, and the "pull" and "push" boundary is between the DCs and the customers. What we observed in this example is consistent and similar to the one discovered by Simpson[30] for multistage "series" inventory system. It is also interesting to note that all the optimal guaranteed service times and net lead times happen to be integer, although we do not restrict them to be integer values.

Example B: Acetic acid supply chain

The second example is for an acetic acid supply chain with three plants, three potential DCs, and four customer

Table 3. Order Processing Time ($t2_{jk}$) Between DCs and Customers (Days)

	Cust. 1	Cust. 2	Cust. 3	Cust. 4	Cust. 5	Cust. 6
DC1	2	2	3	3	4	4
DC2	4	4	1	1	4	4
DC3	4	4	3	3	2	2

Table 2. Order Processing Time ($t1_{ij}$) Between Plants and DCs (Days) for Example A

	DC1	DC2	DC3
Plant 1	7	4	2
Plant 2	2	4	7

Table 4. Unit Transportation Cost (Cl_{ij}) from Plants to DCs (\$/liter)

	DC1	DC2	DC3
Plant 1	0.24	0.20	0.20
Plant 2	0.18	0.19	0.23

Table 5. Unit Transportation Cost ($c2_{jk}$) from DCs to Customers ($/liter)

	Cust. 1	Cust. 2	Cust. 3	Cust. 4	Cust. 5	Cust. 6
DC1	0.01	0.03	0.10	0.44	1.60	2.30
DC2	1.50	0.25	0.01	0.02	0.25	1.50
DC3	2.27	1.73	0.51	0.10	0.01	0.03

demand zones, each of which has a wholesaler (Figure 15). On the basis of the conclusion of the previous example, we need to take into account the inventory costs from DCs and wholesalers when designing the optimal supply chain.

In this instance, the annual fixed DC installation cost (f_j) is $50,000 per year for all the DCs. The variable cost coefficient of installing a DC (g_j) at all the candidate locations is $0.5 per ton/year. The safety stock factors for DCs ($\lambda 1_j$) and wholesalers ($\lambda 2_k$) are the same and equal to 1.96. We consider 365 days in a year (χ). The guaranteed service times of the three plants (SI_i) are 3, 3, and 4 days, respectively. The pipeline inventory holding cost is $1 per ton/day for all the DCs ($\theta 1_{ij}/\chi$) and wholesalers ($\theta 2_{jk}/\chi$), and the safety stock holding cost is $1.5 per ton/day for all the DCs ($h1_j/\chi$) and wholesalers ($h2_k/\chi$). The guaranteed service time of wholesalers to end customers (R_k) is set to be 8 days. The data for

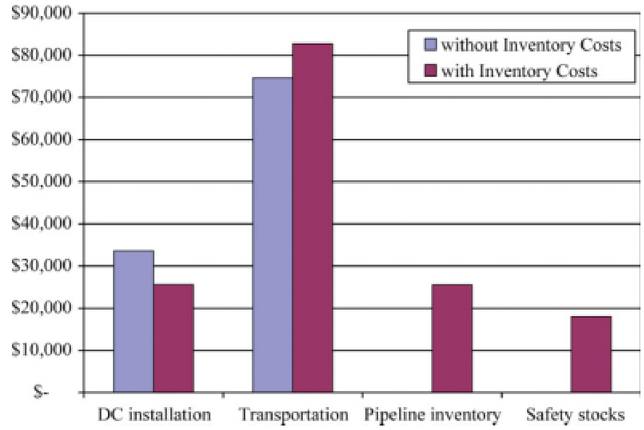

Figure 12. Cost items of the LOX supply chain in Example A.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

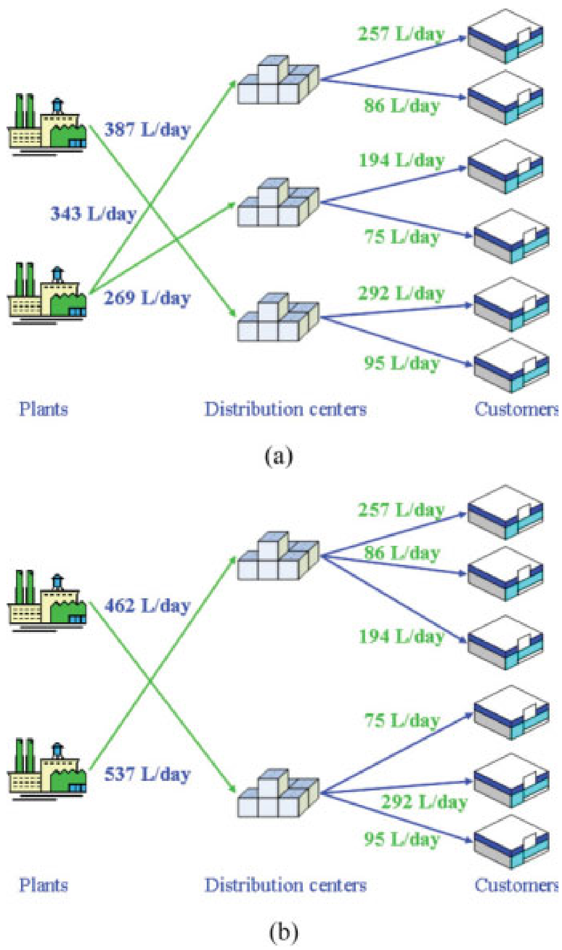

Figure 11. Optimal network structure for the LOX supply chain in Example A.

(a) Without considering inventory costs, total cost = $108,329. (b) Considering inventory costs, total cost = $152,107. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

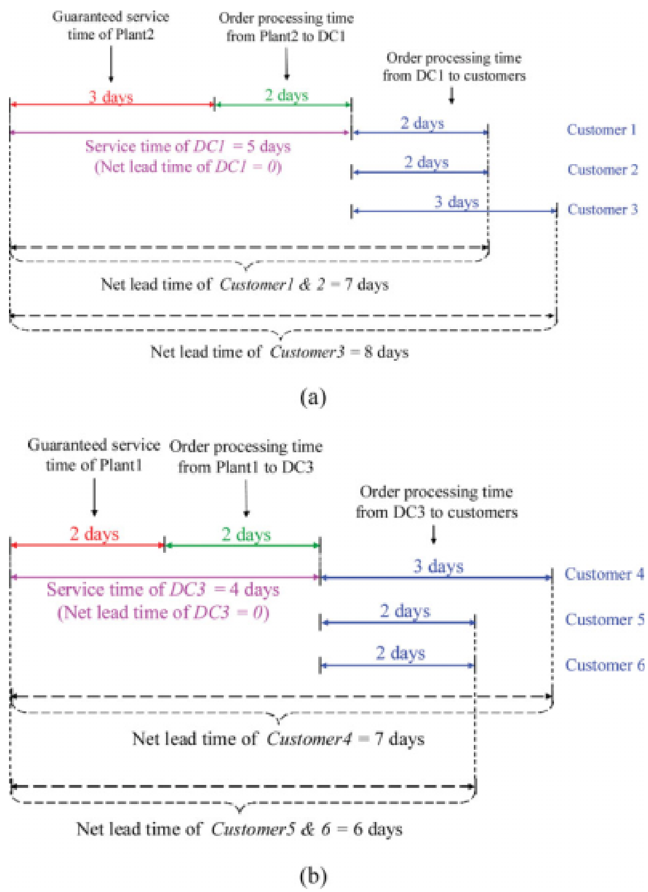

Figure 13. Optimal timing relationship of the LOX supply chain for the design in Figure 8b.

(a) Plant 2: DC 1 – Customers 1, 2, and 3. (b) Plant 1: DC 3 – Customers 4, 5, and 6. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

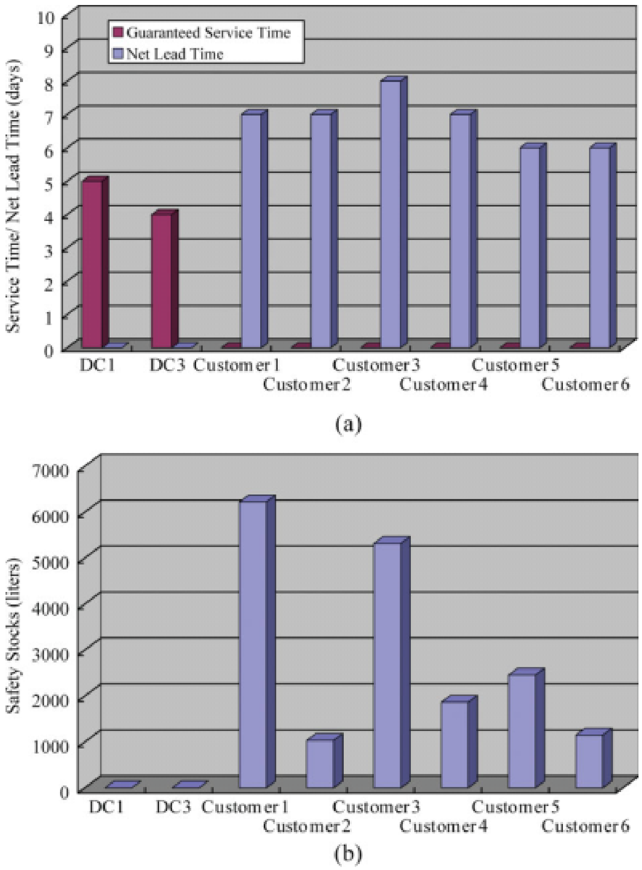

Figure 14. Optimal net lead times and safety stock levels of the LOX supply chain for the design in Figure 8b.

(a) Net lead times and guaranteed service times of DCs and customers. (b) Optimal safety stock levels of DCs and customers. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

demand uncertainty, order processing times, and transportation costs are given in Tables 6–10.

The computational study is carried out on an IBM T40 laptop with Intel 1.50GHz CPU and 512 MB RAM. The original model (P0) has 12 discrete variables, 22 continuous variables, and 27 constraints. Both the initialization MILP model (P2) and the MINLP model (P1) involve 24 binary variables, 185 continuous variables, and 210 constraints. We

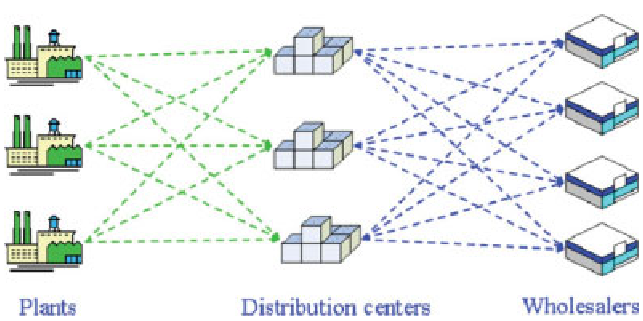

Figure 15. Acetic acid supply chain network superstructure for Example B.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 6. Parameters for Demand Uncertainty for Example B

	Mean Demand μ_i (ton/day)	Standard Deviation σ_i (ton/day)
Wholesaler 1	250	150
Wholesaler 2	180	75
Wholesaler 3	150	80
Wholesaler 4	160	45

solve the original model (P0) by using GAMS/BARON with 0% optimality margin, and it takes a total of 1376.9 CPU seconds. We then solve the reformulated model (P1) with the aforementioned initialization process, the CPU time reduces to only 7.6 s and the optimal solutions are the same as we obtained by solving model (P0). The possible reason is that initialization process helps BARON to find a "good" feasible solution during the preprocessing step. This solution provides a tighter upper bound that reduces the search space and speeds up the computation.

The resulting optimal supply chain network with a minimum annual cost of $1,986,148 per year and the flow rates of transportation links are given in Figure 16. We can see that two DCs are selected to install and one of them serves three wholesalers whereas the other one only serves one wholesaler. The major trade-off is between DC installation costs, transportation costs, and inventory costs. A detailed breakdown of the total cost is given in Figure 17.

Figure 18 shows the optimal net lead times and guaranteed service times of DCs and wholesalers. Because the maximum guaranteed service time of the wholesalers is set to be 8 days, whereas the total deterministic time from plant 2 to DC 1 and to Wholesaler 1 is only 7 days, it is reasonable that this path work as pure "pull" system to reduce the total inventory cost as in Figure 18a. In Figure 18b, we can see that for the path from Plant 1 to DC 2 and to Wholesalers 2, 3, and 4, Wholesaler 4 requires a significantly long order processing time (4 days) compared with others (1 day only). Therefore, it is understandable that most of the safety stocks are maintained in Wholesaler 1, although other nodes in this path also work as pure "pull" system. As seen in Figure 19, similarly as in the previous example, it can be observed that most of the nodes in this supply chain hold zero safety stock and work as pure "pull" system. The inventory allocation in this example also suggests the importance of incorporating multi-echelon stochastic inventory system in the supply chain design. Instead of focusing on the single-stage inventory only, the proposed approach can lead to an explicit improvement of the decision-making in supply chain design, and in turn reduce the total costs.

Solution Algorithm

Although small-scale problems can be solved to global optimality effectively by using a global optimizer, medium

Table 7. Order Processing Time ($t1_{ij}$) Between Plants and DCs (Days) for Example B

	DC1	DC2	DC3
Plant 1	4	4	2
Plant 2	2	4	3
Plant 3	3	4	4

Table 8. Order Processing Time ($t2_{jk}$) Between DCs and Wholesalers (Days)

	Wholesaler 1	Wholesaler 2	Wholesaler 3	Wholesaler 4
DC1	2	2	3	3
DC2	4	4	1	1
DC3	4	4	3	3

and large-scale joint multi-echelon supply chain design and inventory management problems are often computationally intractable with direct solution approaches because of the combinatorial nature and nonlinear nonconvex terms. In this section, we present an effective solution algorithm based on the integration of Lagrangean relaxation and piecewise linear approximation to obtain solutions within 1% of global optimality gap with reasonable computational expense.

Piecewise Linear Approximation

Instead of using the secants in the objective function of model (P1), a tighter lower convex envelope of the univariate square root terms is the piecewise linear function that uses a few additional continuous and discrete variables. Although this may require longer computational times to solve the initialization MILP problem, with the significant progress of MILP solvers piecewise linear approximations have recently been increasingly used for approximating different types of nonconvex nonlinear functions,[43,44] especially for univariate concave functions.[45–48]

There are several different approaches to model piecewise linear function for a concave term.[44,47,49] In this work, we use the "multiple-choice" formulation[47] to approximate the square root term $\sqrt{x}$. Let $P = \{1,2,3,\ldots,p\}$ denote the set of intervals in the piecewise linear function $\varphi(x)$, and M_1, $M_2,\ldots, M_p$, M_{p+1} be the lower and upper bounds of x for each interval p. The "multiple choice" formulation of $\varphi(x) = \sqrt{x}$ is given by,

$$\varphi(x) = \min \quad \sum_p (F_p v_p + C_p u_p)$$

$$\text{s.t.} \quad \sum_p v_p = 1$$

$$\sum_p u_p = x$$

$$M_p v_p \le u_p \le M_{p+1} v_p, \quad p \in P$$

$$v_p \in \{0,1\}, \quad p \in P$$

where $C_p = \dfrac{\sqrt{M_{p+1}} - \sqrt{M_p}}{M_{p+1} - M_p}$ and

$$F_p = \sqrt{M_p} - C_p M_p, \quad p \in P.$$

Table 9. Unit Transportation Cost ($c1_{ij}$) from Vendors to DCs ($/Ton)

	DC1	DC2	DC3
Vendor 1	1.8	1.6	2.0
Vendor 2	2.4	2.2	1.3
Vendor 3	2.0	1.3	2.5

Table 10. Unit Transportation Cost ($c2_{jk}$) from DCs to Wholesalers ($/Ton)

	Wholesaler 1	Wholesaler 2	Wholesaler 3	Wholesaler 4
DC1	1.0	3.3	4.0	7.4
DC2	1.0	0.5	0.1	2.0
DC3	7.7	7.3	5.1	0.1

As can be seen in Figure 20, the more intervals that are used, the better is the approximation of the nonlinear function, but more additional variables and constraints are required. Using this "multiple-choice" formulation to approximate the univariate concave terms $\sqrt{NZV_j}$ and $\sqrt{L_k}$ in the objective function of problem (P1) yields the following piecewise linear MILP problem (P3), which provides a tighter lower bounding problem for (P1).

(P3):

$$\begin{aligned}\text{Min}: & \sum_{j\in J} f_j Y_j + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K} A_{ijk} XZ_{ijk} + \sum_{j\in J}\sum_{k\in K} B_{jk} Z_{jk} \\ & + \sum_{j\in J} q1_j \sum_{p_1} \left(F1_{j,p_1} v1_{j,p_1} + C1_{j,p_1} u1_{j,p_1}\right) \\ & + \sum_{k\in K} q2_k \sum_{p_2} \left(F2_{k,p_2} v2_{k,p_2} + C2_{k,p_2} u2_{k,p_2}\right)\end{aligned} \quad (35)$$

$$\text{s.t.} \quad \sum_{p_1} v1_{j,p_1} = 1, \quad \forall j \quad (36.1)$$

$$\sum_{p_1} u1_{j,p_1} = NZV_j, \quad \forall j \quad (36.2)$$

$$M1_{j,p_1} v1_{j,p_1} \le u1_{j,p_1} \le M1_{j,p_1+1} v1_{j,p_1}, \quad \forall j, p_1 \quad (36.3)$$

$$v1_{j,p_1} \in \{0,1\}, \quad u1_{j,p_1} \ge 0, \quad \forall j, p_1 \quad (36.4)$$

$$\sum_{p_2} v2_{k,p_2} = 1, \quad \forall k \quad (37.1)$$

$$\sum_{p_2} u2_{k,p_2} = L_k, \quad \forall k \quad (37.2)$$

$$M2_{k,p_2} v2_{k,p_2} \le u2_{k,p_2} \le M2_{k,p_2+1} v2_{k,p_2}, \quad \forall k, p_2 \quad (37.3)$$

$$v2_{k,p_2} \in \{0,1\}, \quad u2_{k,p_2} \ge 0, \quad \forall k, p_2 \quad (37.4)$$

All the constraints of (P1).

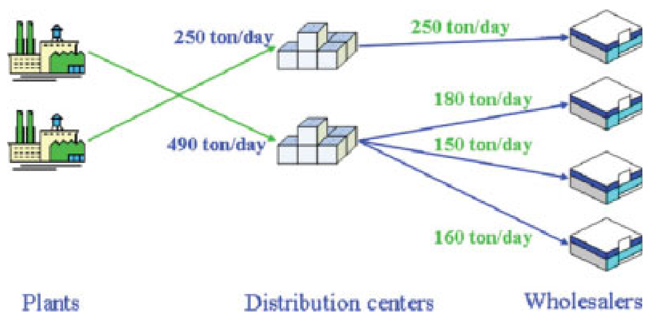

Figure 16. Optimal network structure for the acetic acid supply chain (Example B).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

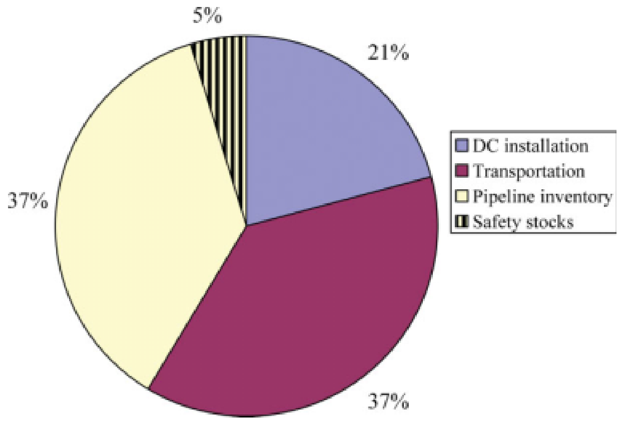

Figure 17. Cost items of the acetic acid supply chain in Example B.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Note that any feasible solution obtained from problem (P3) is also a feasible solution of (P1), and for each feasible solution the objective value of (P3) is always less than or equal to the objective value of (P1).

In summary, the solution obtained by solving (P3) can provide a "good" initial solution and tight upper bound of solving problem (P1). In addition, the optimal objective value of (P3) provides a valid global lower bound to the objective value of (P1).

To obtain an initial point "close" enough to the global solution, we can in principle use piecewise linear functions with sufficient large number of intervals to approximate the univariate concave terms in (P1). However, it is a nontrivial task to solve large-scale instances of (P3). To further improve the computational efficiency, we first exploit some properties of problem (P1).

Alternative Formulation

Let us first consider an alternative model formulation (AP) of problem (P1) in which (23)–(25) are excluded and represent the inequality (13) by the disjunction (39),

$$\text{Min}: \sum_{j\in J} f_j Y_j + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K} A_{ijk} XZ_{ijk} + \sum_{j\in J}\sum_{k\in K} B_{jk} Z_{jk} + \sum_{j\in J} q1_j \sqrt{NZV_j} + \sum_{j\in J}\sum_{k\in K} q2_k \sqrt{L_{jk}} \qquad (38)$$

s.t.

$$\begin{bmatrix} Z_{jk} \\ L_{jk} \ge S_j + t_{jk} - R_k \end{bmatrix} \vee \begin{bmatrix} \neg Z_{jk} \\ L_{jk} \le 0 \end{bmatrix}, \quad \forall j,k \qquad (39)$$

$$L_{jk} \ge 0, \quad \forall j,k \qquad (40)$$

Constraints (9)(11), (14)(15), (18)(22), and (27)(31)

Note that each disjunction in (39) is written in terms of the disaggregated variable L_{jk}, which is then substituted in the last term of the objective function (38). Applying the convex hull reformulation[50,51] to the above disjunctive constraint (39) in (AP) leads to:

$$L_{jk} \ge S1_{jk} + t_{jk}\cdot Z_{jk} - R_k \cdot Z_{jk}, \quad \forall j,k \qquad (41.1)$$

$$L_{jk} \le Z_{jk}\cdot L_{jk}^U, \quad \forall j,k \qquad (41.2)$$

$$S_j = S1_{jk} + S2_{jk}, \quad \forall j,k \qquad (41.3)$$

$$S1_{jk} \le Z_{jk}\cdot S_j^U, \quad \forall j,k \qquad (41.4)$$

$$S2_{jk} \le (1 - Z_{jk})\cdot S_j^U, \quad \forall j,k \qquad (41.5)$$

$$S1_{jk} \ge 0, \quad S2_{jk} \ge 0, \quad \forall j,k \qquad (41.6)$$

where $S1_{jk}$ and $S2_{jk}$ are two new auxiliary variables.

The value of the new continuous variable L_{jk} in (AP) is defined through two new constraints (39) and (40). It means that if customer demand zone k is assigned to DC j, i.e., $Z_{jk} = 1$, then the new variable L_{jk} represents the net lead time of customer demand zone k, otherwise $L_{jk} = 0$. Thus, any feasible solution of problem (AP) is also a feasible solution of problem (P1), and vice versa, as stated in the following property:

Property 1. *If (X^*_{ij}, Y^*_j, Z^*_{jk}, S^*_j, N^*_j, L^*_k) is a feasible solution of problem (P1) with objective function value W^*, then (X^*_{ij}, Y^*_j, Z^*_{jk}, S^*_j, N^*_j, L^*_{jk}), where $L^*_{jk} = 0$ if $Z^*_{jk} = 0$ and $L^*_{jk} = L^*_k$ if $Z^*_{jk} = 1$, is a feasible solution of problem (AP) and the associated objective function value of (AP) is W^*. If (X^*_{ij}, Y^*_j, Z^*_{jk}, S^*_j, N^*_j, L^*_{jk}) is a feasible solution of problem (AP) with objective function value W^*, then (X^*_{ij}, Y^*_j, Z^*_{jk}, S^*_j, N^*_j, L^*_k), where $L^*_k = \sum_{j\in J} L^*_{jk}$, is a feasible solution of problem (P1) and the associated objective function value of (P1) is W^*.*

The property can be easily proved and is omitted in this article. This property shows that we can solve either (P1) or (AP) to obtain a feasible or optimal solution, and then algebraically derive the solution to another problem with the

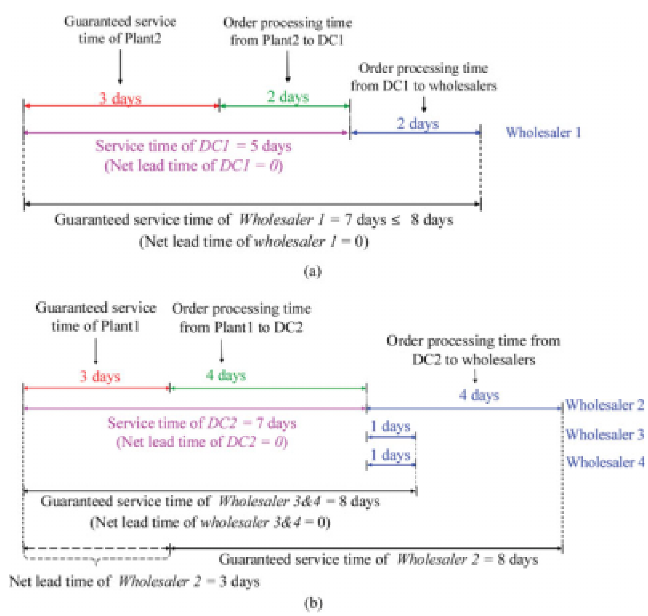

Figure 18. Optimal timing relationship for the acetic acid supply chain in Example B.

(a) Plant 2: DC 1 – Wholesaler 1. (b) Plant 1: DC 2 – Wholesalers 2, 3, and 4. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

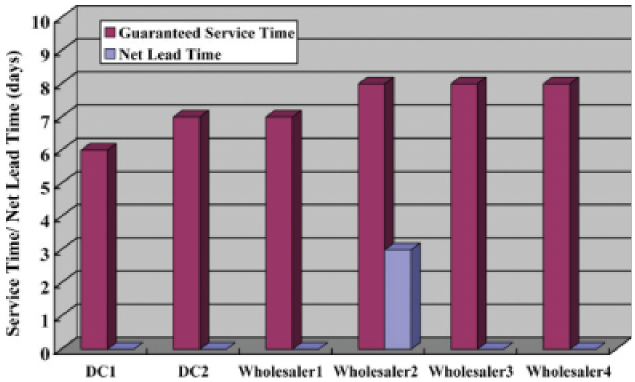

Figure 19. Net lead times and guaranteed service times of DCs and wholesalers for the acetic acid supply chain in Example B.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

same objective value. Compared with model (P1), model (AP) has more square root terms in the objective function due to the introduction of variable L_{jk}, although the linearization constraints (23), (24), and (25) for the product of S_j and Z_{jk} are not included in (AP).

Solving (AP) with the direct solution approach to obtain the optimal solution of (P1) can be expensive since there are more nonlinear terms in (AP). However, this model can be decomposed by using a Lagrangean relaxation algorithm. Furthermore, due to the equivalence of (P1) and (AP), we can solve (P1) instead of (AP) in the full or reduced variable space during the solution procedure of the Lagrangean algorithm. In the next section, we describe how to incorporate model (AP) and the piecewise linear approximation into a Lagrangean relaxation algorithm.

Lagrangean Relaxation Algorithm

To obtain near global optimal solutions to problems (P1) and (AP) with modest computational effort, we propose a decomposition algorithm based on Lagrangean relaxation.

The decomposition procedure

In this solution procedure, we use a decomposition scheme by dualizing the assignment constraints (10) in (AP) using the Lagrange multipliers λ_k. As a result, we obtain the following relaxed problem (APL),

$$W = \text{Min}: \sum_{j \in J}\left(f_j Y_j + \sum_{i \in I}\sum_{k \in K} A_{ijk} XZ_{ijk} + \sum_{k \in K}(B_{jk} - \lambda_k) Z_{jk} + q1_j\sqrt{NZV_j} + \sum_{k\in K} q2_k\sqrt{L_{jk}}\right) + \sum_{k \in K}\lambda_k \quad (42)$$

s.t. Constriants (9), (11), (14)–(16), (18)–(22), (27)–(31), and (41)

where W is the objective function value. Note that problem (APL) can be decomposed into $|J|$ subproblems, one for each candidate DC site $j \in J$, where each one is denoted by (APL$_j$) and is shown for a specific subproblem for candidate DC site j as follows:

$$W_j = \text{Min}: \; f_j Y_j + \sum_{i \in I}\sum_{k \in K} A_{ijk} XZ_{ijk} + \sum_{k \in K}(B_{jk} - \lambda_k) Z_{jk} + q1_j\sqrt{NZV_j} + \sum_{k\in K} q2_k\sqrt{L_{jk}} \quad (43)$$

$$\text{s.t.} \quad \sum_{i \in I} X_{ij} = Y_j,$$

$$Z_{jk} \le Y_j, \quad \forall k$$

$$N_j \ge \sum_{i \in I} \bar{S}_{ij} \cdot X_{ij} - S_j,$$

$$XZ_{ijk} \le X_{ij}, \quad \forall i, k$$

$$XZ_{ijk} \le Z_{jk}, \quad \forall i, k$$

$$XZ_{ijk} \ge X_{ij} + Z_{jk} - 1, \quad \forall i, k$$

$$NZ_{jk} + NZ1_{jk} = N_j, \quad \forall k$$

$$NZ_{jk} \le Z_{jk} \cdot N_j^U, \quad \forall k$$

$$NZ1_{jk} \le (1 - Z_{jk}) \cdot N_j^U, \quad \forall k$$

$$NZV_j = \sum_{k \in K} \sigma_k^2 \cdot NZ_{jk},$$

$$L_{jk} \le Z_{jk} \cdot L_{jk}^U, \quad \forall k$$

$$L_{jk} \ge S1_{jk} + t_{jk} \cdot Z_{jk} - R_k \cdot Z_{jk}, \quad \forall k$$

$$S_j = S1_{jk} + S2_{jk}, \quad \forall k$$

$$S1_{jk} \le Z_{jk} \cdot S_j^U, \quad \forall k$$

$$S2_{jk} \le (1 - Z_{jk}) \cdot S_j^U, \quad \forall k$$

$$S1_{jk} \ge 0, \;\; S2_{jk} \ge 0, \quad \forall k$$

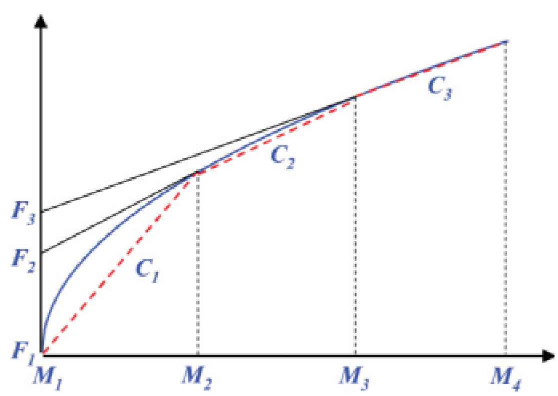

Figure 20. Piecewise linear function to approximate square root term.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

$$X_{ij}, Y_j, Z_{jk} \in \{0,1\}, \quad \forall i,k$$

$$S_j \geq 0, \; N_j \geq 0,$$

$$L_{jk} \geq 0, \quad \forall k$$

$$XZ_{ijk} \geq 0, \quad \forall i,k$$

$$NZV_j \geq 0,$$

Hence, (APL) can be decomposed into $|J|$ subproblems (APL$_j$) and one for each candidate DC site $j \in J$. Let W_j denote the globally optimal objective function value of problem (APL$_j$). As a result of the decomposition procedure, the global minimum of (APL), which corresponds to a lower bound of the global optimum of problem (AP), can be calculated by:

$$W = \sum_{j \in J} W_j + \sum_{k \in K} \lambda_k. \tag{44}$$

For each fixed value of the Lagrange multipliers λ_k, we solve problem (APL$_j$) for each candidate DC location j. Then, based on (44), the objective function value of problem (APL) can be calculated for each fixed value of λ_k. Using a standard subgradient method[52,53] to update the Lagrange multiplier λ_k, the algorithm iterates until a preset optimality tolerance is reached.

Lagrangean relaxation subproblems

At each iteration with fixed values of the Lagrange multipliers λ_k, the binary variables for installing a DC (Y_j) are optimized separately in each subproblem (APL$_j$) in the aforementioned decomposition procedure. For each subproblem (APL$_j$), we can observe that the objective function value of (APL$_j$) is 0 if and only if $Y_j = 0$ (i.e., we do not install DC j). In other words, there is a feasible solution that leads to the objective of subproblem (APL$_j$) equal to 0. Therefore, the global minimum of subproblem (APL$_j$) should be less than or equal to zero. Given this observation, it is possible that for some value of λ_k (such as $\lambda_k = 0$, $k \in K$) the optimal objective function values for all the subproblem (APL$_j$) are zero (i.e., $Y_j = 0$, $j \in J$, we do not select any DC). However, the original assignment constraint (8) implies that at least one DC should be selected to meet the demands, i.e.,

$$\sum_{j \in J} Y_j \geq 1. \tag{45}$$

Once constraint (10) is relaxed, constraint (45) becomes "non-redundant" and should be taken into account in the algorithm.[6,54] To satisfy constraint (45) in the Lagrangean relaxation procedure, we make the following modifications to the aforementioned solution step of problem (APL$_j$).

First, consider the problem (APLR$_j$), which is actually a special case of (APL$_j$) when $Y_j = 1$. The formulation for a specific j is given as:

$$\widehat{W}_j = \text{Min}: f_j + \sum_{i \in I}\sum_{k \in K} A_{ijk} XZ_{ijk} + \sum_{k \in K}(B_{jk} - \lambda_k) Z_{jk} + q1_j \sqrt{NZV_j} + \sum_{k \in K} q2_k \sqrt{L_{jk}} \tag{46}$$

s.t. the same constraints as (APL$_j$) with $Y_j = 1$.

where $\widehat{W}_j$ is denoted as the optimal objective function value of the problem (APLR$_j$).

With a similar scheme as introduced in our previous work,[6] we use the following solution procedure to deal with the implied constraint (45): for each fixed value of λ_k, we solve (APLR$_j$) for every candidate DC location j. Then select the DCs in candidate location j (i.e., let $Y_j = 1$), for which $\widehat{W}_j \leq 0$. For all the remaining DCs for which $\widehat{W}_j > 0$, we do not select them and set $Y_j = 0$. If all the $\widehat{W}_j > 0$, $\forall j \in J$, we select only one DC with the minimum $\widehat{W}_j$, i.e., $Y_{j^*} = 1$ for the j^* such that $\widehat{W}_{j^*} = \min_{j \in J}\{\widehat{W}_j\}$. By doing this at each iteration of the Lagrangean relaxation (for each value of the multiplier λ_i), we ensure that the optimal solution always satisfies $\sum_{j \in J} Y_j \geq 1$. Thus, the globally optimal objective function of (APL$_j$) can be recalculated as:

$$W = \sum_{j \in J, Y_j = 1} \widehat{W}_j + \sum_{k \in K} \lambda_k. \tag{57}$$

As (APLR$_j$) includes $|K| + 1$ univariate concave terms in the objective function, solving it to global optimality might be computationally expensive if $|K|$ is large. To improve the computational efficiency, piecewise linear approximations can be used. Similarly as in (P3), we use the "multiple-choice" formulation to approximate the concave terms in (APLR$_j$), and obtain the following MILP model (APLP$_j$),

$$\begin{aligned}\overline{W}_j = \text{Min}: f_j &+ \sum_{i \in I}\sum_{k \in K} A_{ijk} XZ_{ijk} \\ &+ q1_j \sum_{p_1}\left(F1_{j,p_1} v1_{j,p_1} + C1_{j,p_1} u1_{j,p_1}\right) \\ &+ \sum_{k \in K}(B_{jk} - \lambda_k) Z_{jk} \\ &+ \sum_{k \in K} q2_k \sum_{p_3}\left(F3_{j,k,p_3} v3_{j,k,p_3} + C3_{j,k,p_3} u3_{j,k,p_3}\right)\end{aligned} \tag{48}$$

$$\text{s.t.} \quad \sum_{p_1} v1_{j,p_1} = 1, \tag{36.1}$$

$$\sum_{p_1} u1_{j,p_1} = NZV_j, \tag{36.2}$$

$$M1_{j,p_1} v1_{j,p_1} \leq u1_{j,p_1} \leq M1_{j,p_1+1} v1_{j,p_1}, \quad \forall p_1 \tag{36.3}$$

$$v1_{j,p_1} \in \{0,1\}, \quad u1_{j,p_1} \geq 0, \quad \forall p_1 \tag{36.4}$$

$$\sum_{p_3} v3_{j,k,p_3} = 1, \quad \forall k \tag{49.1}$$

$$\sum_{p_3} u3_{j,k,p_3} = L_{jk}, \quad \forall k \tag{49.2}$$

$$M3_{j,k,p_3} v3_{j,k,p_3} \leq u3_{j,k,p_3} \leq M3_{j,k,p_3+1} v3_{j,k,p_3}, \quad \forall k, p_3 \tag{49.3}$$

$$v3_{j,k,p_3} \in \{0,1\}, \quad u3_{j,k,p_3} \geq 0, \quad \forall k, p_3 \tag{49.4}$$

All the constraints of (APLR$_j$).

where $\overline{W}_j$ is the optimal objective function value of problem (APLP$_j$).

As discussed earlier, $\overline{W}_j$ provides a valid "global" lower bound of $\widehat{W}_j$. Thus, the globally optimal objective function

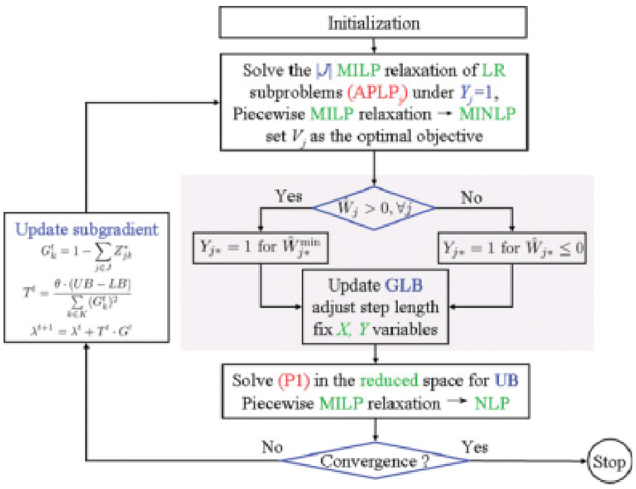

Figure 21. Algorithmic flowchart for the proposed algorithm based on the integration of Lagrangean decomposition and piecewise linear approximation.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

value of problem (AP) has a globally lower bound $\overline{W}$, which can be calculated by:

$$\overline{W} = \sum_{j \in J, Y_j = 1} \overline{W}_j + \sum_{k \in K} \lambda_k. \tag{50}$$

On the basis of this result, instead of solving each (APLR$_j$) to globally optimality, we can use the optimal solution of (APLP$_j$) as an initial point to solve (APLR$_j$) with an MINLP solver that relies on convexity assumptions, such as DICOPT, SBB, etc., so as to improve the computational efficiency. In this way, the optimal objective function of (APL$_j$), calculated by $W = \sum_{j \in J, Y_j = 1} \widehat{W}_j + \sum_{k \in K} \lambda_k$, is no longer a valid global lower bound of the optimal objective function value of problem (AP). However, the lower bound $\overline{W}$, obtained from piecewise linear lower bounding problem (APLP$_j$), still provides a "globally" lower bound to the optimal objective function value of problem (AP).

Upper bounding

A feasible solution of problem (AP) naturally provides a valid upper bound to its global optimal objective function value. To obtain a feasible solution, a common approach is to fix the values of the binary variables (X_{ij}, Y_j, and Z_{jk}) and solve (AP) in the reduced variable space with an NLP solver. However, solving a large-scale NLP problem can be computationally expensive.

Based on Proposition 1, we can solve problem (P1) instead of (AP) with fixed values of the binary variables (X_{ij}, Y_j, and Z_{jk}) to reduce the computational effort because (P1) has fewer nonlinear terms than (AP). To avoid solving the large scale NLP problem, we first solve the piecewise linear lower bounding MILP problem (P3) with fixed binary variables, and then substitute the optimal solution into the objective function of problem (P1) to calculate the associated objective function value. As discussed in Section 6.1, the optimal solution of (P3) is a feasible solution of (P1). Thus, the objective function value of (P1) obtained by function evaluation yields an upper bound of the global minimum of problem (AP). By using this approach, we avoid using an NLP solver and improve the computational efficiency and robustness without sacrificing the solution quality.

The solution algorithm

To summarize, the proposed solution algorithm based on the integration of Lagrangean relaxation and piece-wise linear approximation is as follows (also given in Figure 21):

Step 1: (Initialization).

Use an arbitrary guess as the initial vector of Lagrange multipliers λ', or else the dual of constraint (10) of a local optimum of the NLP relaxation of model (P1). Let the incumbent upper bound be UB $= +\infty$, incumbent lower bound be LB $= -\infty$, global lower bound be GLB $= -\infty$,

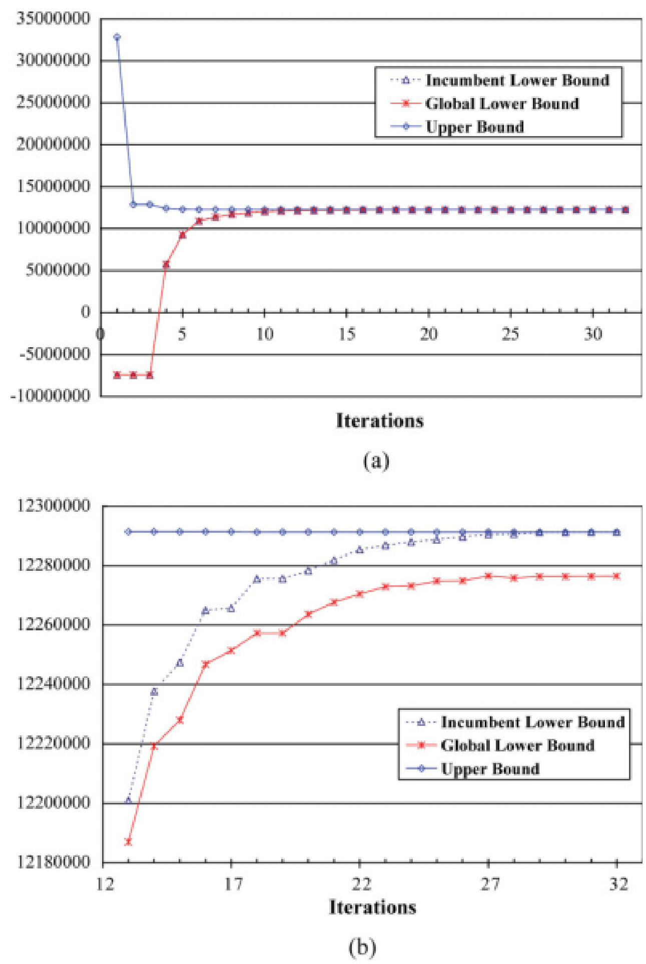

Figure 22. Bounds of each iteration of the proposed algorithm for instance with 3 plants, 50 potential DCs, and 150 customer demand zones.

(a) Bounds for Iteration 1 to Iteration 33 (the last iteration). (b) Bounds for Iteration 13 to Iteration 33 (the last iteration). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 11. Problem Sizes for the Medium and Large Scale Instances

			Original Formulation (P0) −MINLP			MINLP Reformulation (P1) & MILP Initialization Model (P2)			Decomposed Subproblem ($APLP_j$), MILP, 20 Intervals for Piecewise Linear Approx.		
Plants i	DCs j	Customer Demand Zones k	Dis. Var.	Con. Var.	Const.	Dis. Var.	Con. Var.	Const.	Dis. Var.	Con. Var.	Const.
2	20	20	460	60	480	460	2,480	5,300	442	563	1,165
5	30	50	1,680	110	1,660	1,680	13,640	33,190	1,105	2,658	3,355
10	50	100	5,550	200	5,300	5,550	70,250	185,350	2,210	5,813	8,205
20	50	100	6,050	200	5,300	6,050	120,250	335,350	2,220	6,823	11,205
3	50	150	7,700	250	7,900	7,700	52,800	120,450	3,300	4,353	9,155
15	100	200	21,600	400	20,500	21,600	300,400	1,040,700	11,603	42,103	11,245

Dis. Var., discrete variables; Con. Var., continuos variable; Const., constraints; NL-NZ, nonlinear nonzeros.

and iteration number be $t = 1$. Set the step length parameter $\theta = 2$.

Step 2

With fixed Lagrange multipliers λ^t, first solve the piecewise linear lower bounding problem of the modified Lagrangean relaxation subproblem ($APLP_j$) with an MILP solver. Then fix the value of the binary variables X_{ij} and Z_{jk}, use the optimal solution as the initial point, and solve the modified Lagrangean relaxation subproblem ($APLR_j$) with an NLP solver (not necessary global optimizer) or calculate the objective function value of ($APLR_j$) directly using the optimal solution of ($APLP_j$).

Denote the optimal objective function value of ($APLP_j$) as $\bar{W}_j^{(\lambda^t)}$, the optimal objective function value of ($APLR_j$) as $\widehat{W}_j^{(\lambda^t)}$, and the optimal solutions of ($APLR_j$) as $(\widehat{X}_{ij}^{(\lambda_t)}, \widehat{Z}_{jk}^{(\lambda_t)})$.

If all $\widehat{W}_j^{(\lambda^t)} > 0$, $\forall j \in J$, let $Y_{j*}^{(\lambda t)} = 1$, $X_{ij^*}^{(\lambda_t)} = \widehat{X}_{ij^*}^{(\lambda_t)}$, for j^* with $\widehat{W}_{j^*}^{(\lambda^t)} = \min_{j \in J} \left\{ \widehat{W}_j^{(\lambda^t)} \right\}$.

Else, let $y_j^{(\lambda_t)} = 0$, $X_{ij}^{(\lambda_t)} = \widehat{X}_{ij}^{(\lambda_t)}$ for all j with $\widehat{W}_j^{(\lambda^t)} \leq 0$, and $Y_j^{(\lambda t)} = 0$, $X_{ij}^{(\lambda t)} = 0$ for all j such that $\widehat{W}_j^{(\lambda^t)} > 0$.

Calculate $W^{(\lambda^t)} = \sum_{j \in J, Y_j^{(\lambda_t)} = 1} \widehat{W}_j + \sum_{k \in K} \lambda_k$, $\overline{W}^{(\lambda^t)} = \sum_{j \in J, Y_j^{(\lambda_t)} = 1} \overline{W}_j + \sum_{k \in K} \lambda_k$.

If $W^{(\lambda^t)} > \mathrm{LB}$, update lower bound by setting $\mathrm{LB} = W^{(\lambda^t)}$. If more than 2 iterations of the subgradient procedure[52] are performed without an increment of LB, then halve the step length parameter by setting $\theta = \frac{\theta}{2}$.

If $\overline{W}(\lambda^t) > \mathrm{GLB}$, update lower bound by setting $\mathrm{GLB} = \overline{W}(\lambda^t)$.

Note that the strictly global lower bound is given by GLB instead of LB, if ($APLR_j$) is not solved with a global optimizer.

Step 3

Fixing the design variable values as $X_{ij} = X_{ij}^{(\lambda_t)}$, $Y_j = Y_j^{(\lambda_t)}$. and using $\widehat{Z}_{jk}^{(\lambda_t)}$ as the initial values of the binary variables Z_{jk}, solve the piecewise linear lower bounding problem (P3) in the reduced space with fixed X_{ij}, Y_j, and λ^t using an MILP solver. Denote the optimal solution as $(Z_{jk}^{(\lambda_t)}, S_j^{(\lambda_t)}, N_j^{(\lambda_t)}, L_k^{(\lambda_t)}.)$. Substitute the optimal solution $(X_{ij}^{(\lambda_t)}, Y_j^{(\lambda_t)}, Z_{jk}^{(\lambda_t)}, S_j^{(\lambda_t)}, N_j^{(\lambda_t)}, L_k^{(\lambda_t)})$ into problem (P1) and calculate its objective function value, which is denoted as $\widehat{W}^{(\lambda^t)}$

If $\widehat{W}^{(\lambda^t)} < \mathrm{UB}$, update the upper bound by setting $\mathrm{UB} = \widehat{W}^{(\lambda^t)}$.

Step 4

Calculate the subgradient (G_k) using $G_j^t = 1 - \sum_{j \in J} \widehat{Z}_{jk}^{(\lambda_t)}$, $\forall k$. Compute the step size T,[52,53] $T^t = \frac{\theta \cdot (\mathrm{UB} - \mathrm{LB})}{\sum_{k \in K} (G_k^t)^2}$.

Update the multipliers, $\lambda^{t+1} = \lambda^t + T^t \cdot G^t$.

Step 5

If $gap = \frac{\mathrm{UB} - \mathrm{LB}}{\mathrm{UB}} < tol$ (e.g., 10^{-3}), or $\|\lambda^{t+1} - \lambda^t\|^2 < tol$ (e.g., 10^{-2}) or the maximum number of iterations has been reached, set UB as the optimal objective function value, and set $(X_{ij}^{(\lambda_t)}, Y_j^{(\lambda_t)}, Z_{jk}^{(\lambda_t)}, S_j^{(\lambda_t)}, N_j^{(\lambda_t)}, L_k^{(\lambda_t)})$ as the optimal solution.

Else, increment t as $t + 1$, go to Step 2.

We should note that the entire procedure requires only an MILP solver. An NLP solver can be used to solve problem ($APLR_j$) in the reduced variable space in Step 2, but it is not required. The reason is that the solution of the nonlinear optimization problem ($APLR_j$), which reduces to an NLP from MINLP after X_{ij} and Z_{jk} are fixed, can be substituted by simple function evaluation as stated in Step 2, although the solution quality may be sacrificed. Also, the above algorithm is guaranteed to provide rigorous global lower bounds

Table 12. Comparison of the Performance of the Algorithms for Medium and Large Scale Instances

			Solve (P0) Directly with BARON				Solve (P2) with CPLEX for at Most 1 h, Then Solve (P1) with DICOPT or SBB				Proposed Algorithm (Lagrangean Decomposition + Piecewise Linear Approximation)				
							DICOPT		SBB						
i	j	k	Solution	LB	Gap	Time (s)	Solution	Time (s)	Solution	Time (s)	Solution	Global LB	Global Gap	Time (s)	Iter.
2	20	20	1,889,577	1,159,841	63.2%	360,000	1,820,174	140.3	1,813,541	163.6	1,776,969	1,775,957	0.06%	175.0	11
5	30	50	–*	–*	–*	360,000	–†	360,000	–†	360,000	4,417,353	4,403,582	0.31%	3,279	24
10	50	100	–*	–*	–*	360,000	–†	360,000	–†	360,000	7,512,609	7,507,354	0.07%	27,719	42
20	50	100	–*	–*	–*	360,000	–†	360,000	–†	360,000	5,620,045	5,597,846	0.39%	27,748	53
3	50	150	–*	–*	–*	360,000	–†	360,000	–†	360,000	12,291,296	12,276,483	0.12%	16,112	32
15	100	200	–*	–*	–*	360,000	–†	360,000	–†	360,000	23,565,443	23,294,310	0.01%	35,612	69

i, number of plants; j, number of DCs; k, number of customer demand zones.
*No solution or bounds were returned due to solver failure.
†No solution was returned after 100 hours.

by GLB as discussed in Step 2. Thus, the gap for the global optimum is given by G_{gap} = (UB − GLB)/GLB. Because of the duality gap, this algorithm stops after a finite number of iterations. As will be shown in the computational results, the global gaps are quite small.

Computational Results

To illustrate the application of the proposed solution strategies, we present computational experiments for problems with up to 15 plants, 100 potential DCs and 200 customer demand zones on an IBM T40 laptop with Intel 1.50 GHz CPU and 512 MB RAM. The proposed solution procedure is coded in GAMS 22.8.1. The MILP problems are solved using CPLEX 11.0.1, the NLP problems in Step 2 of the Lagrangean solution approach are solved with solver CONOPT 3.3, and the global optimizer used in the computational experiments is BARON 8.1.4.

Input parameters

As we consider large scale problems, most of the input data are generated randomly. The safety stock factors for DCs ($\lambda 1_j$) and customers ($\lambda 2_k$) are the same and equal to 1.96, which corresponds to 97.5% service level is demand is normally distributed. We consider 365 days in a year (χ). The guaranteed service time of the last echelon customer demand zones (R_k) are set to 0. The annual fixed costs ($/year) to install the DCs (f_j) are generated uniformly on U[150,000,160,000] and the variable cost coefficient (g_j, $ per ton/year) are generated uniformly on U[0.01, 0.1]. The guaranteed service times of the plants (SI_i, days) are set as integers uniformly distributed on U[1,5]. The order processing time ($t1_{ij}$, days) between plants and DCs are generated as integers uniformly distributed on U[1,7], and the order processing time ($t2_{jk}$, days) between DCs and customer demand zones are generated as integers uniformly distributed on U[1,3]. The unit transportation cost from plants to DCs ($c1_{ij}$, $/ton) and from DCs to customer demand zones ($c2_{jk}$, $/ton) are set to

$$c1_{ij} = t1_{ij} \times U[0.05, 0.1], \quad c2_{jk} = t2_{jk} \times U[0.05, 0.1]$$

The expected demand (μ_i, ton/day) is generated uniformly distributed on U[75,150] and its standard deviation (σ_i, ton/day) is generated uniformly distributed on U[0,50]. The daily unit pipeline and safety stock inventory holding costs ($\theta 1_{ij}/\chi$, $\theta 2_{jk}/\chi$, $h1_j/\chi$, and $h2_k/\chi$) are generated uniformly distributed on U[0.1, 1].

In this instance, the annual fixed DC installation cost (f_j) is $50,000 per year for all the DCs. The variable cost coefficient of installing a DC (g_j) at all the candidate locations is $0.5 per ton/year. The safety stock factors for DCs ($\lambda 1_j$) and wholesalers ($\lambda 2_k$) are the same and equal to 1.96. We consider 365 days in a year (χ). The guaranteed service times of the three vendors (SI_i) are 3, 3, and 4 days, respectively. The pipeline inventory holding cost is $1 ton/day for all the DCs ($\theta 1_{ij}/\chi$) and wholesalers ($\theta 2_{jk}/\chi$), and the safety stock holding cost is $1.5 per ton/day for all the DCs ($h1_j/\chi$) and wholesalers ($h2_k/\chi$).

The number of intervals P_1, P_2, P_3, which are required to approximate each of the univariate concave terms $\sqrt{NZV_j}$, $\sqrt{L_k}$, and $\sqrt{L_{jk}}$ in (P3) and (APLP$_j$), are all set to 20. All the breakpoints of the piecewise linear function are evenly distributed between the lower and upper bound of the variables to be approximated.

Performance

The problem sizes for the six instances we considered in this example are given in Table 11. The numbers in i, j, k columns stand for the number of plants, potential DCs, and customer demand zones of each instance. The computational results of the six instances are given in Table 12. We solve each instance with three solution approaches. The first approach is solving the original MINLP problem (P0) directly with the global optimizer BARON. The second approach is to first solve the initialization MILP problem (P2) with at most 1 h computational time, then use the best solution obtained from (P2) as the initial solution to solve the MINLP problem (P1) with DICOPT and SBB solvers. The third approach is the proposed algorithm presented in Section 6.3.4. In the proposed algorithm, we solve the NLP relaxation of model (P1) to obtain the initial vector of Lagrange multipliers in Step 1 and solve the modified Lagrangean relaxation subproblem (APLR$_j$) with CONOPT 3.3 in Step 2. As we can see, global optimal solutions could not be obtained for 10 hours by solving these problems directly with BARON or solving the reformulated problem (P1) with the initialization scheme. However, optimal solutions within 1% global optimality are obtained for the instances by using the proposed algorithm.

Furthermore, we can see from Table 12 that only the medium size instance with 2 plants, 20 potential DCs, and 20 customer demand zones, can be solved directly with BARON to obtain a feasible solution (63% gap). For all the other instances, BARON fails to return any bounds or feasible solutions due to the large size of these problems. By solving the reformulated problem (P1) of the medium scale instance with convex MINLP solvers DICOPT and SBB after the initialization scheme, we can see that a better feasible solution can be obtained with smaller global optimality gap and much shorter computational times (140.3 and 163.6 CPU seconds, respectively). In contrast, the proposed Lagrangean relaxation and decomposition algorithm can obtain a much better solution ($1,776,969 vs. $1,889,577 and $1,820,174) within 0.06% of global optimality in much shorter time (175.0 CPU seconds). For the other large scale instances, BARON cannot finish preprocessing after 100 h when solving the reformulated problem (P1). In contrast the proposed Lagrangean relaxation algorithm can solve all the instances within 1% optimality in less than 10 h of computational time as shown in Table 12. Therefore, we can see the significant advantage of using the proposed Lagrangean relaxation algorithm for solving medium and large scale instances.

The change of bounds during the Lagrangean relaxation and decomposition algorithm for a large scale instance with 3 plants, 50 potential DCs, and 150 customer demand zones are given in Figure 22. We can see that as iterations proceed, the upper bound, which is the objective function value of a feasible solution, keeps decreasing, and both the incumbent lower bound and the global lower bound keep increasing, until the termination criterion is satisfied. We should note that the global optimality gap is given by the upper bound and the global lower bound instead of the incumbent lower bound.

Conclusion

In this article, we have presented an MINLP model that determines the optimal network structure, transportation, and inventory levels of a multi-echelon supply chain with the presence of customer demand uncertainty. The guaranteed service approach[7–10,25–30] is used to model the multi-echelon inventory system. The risk pooling effect is taken into account in the model by relating the demands in the downstream nodes to their upstream nodes. Examples on supply chains for industrial gases and specialty chemicals are presented to illustrate the applicability of the proposed model. To solve the resulting MINLP problem efficiently for large scale instances, a decomposition algorithm based on Lagrangean relaxation and piecewise linear approximation was proposed. Computational experiments on large scale problems show that the proposed algorithm can obtain global or near-global optimal solutions (typically within 1% of the global optimum) in modest computational expense without the need of a global optimizer.

Future work will address the integration of production planning and stochastic multi-echelon inventory management because production activities strongly influence the lead time, and the guaranteed service times of plants should be variables depending on the production planning, although in this work we consider the service time of plants as parameters with fixed values. Another possible future extension is to integrate stochastic inventory management, supply chain design, and production planning to further improve the decision-making across the process supply chains. This research can be also extended to address the responsiveness issue[55,56] of process supply chains and will be given in a further article.

Acknowledgments

The authors acknowledge the financial support from the National Science Foundation under Grant No. DMI-0556090 and No. OCI-0750826, and Pennsylvania Infrastructure Technology Alliance (PITA). Fengqi You is grateful to Prof. S. Willems (Boston University) and Prof. S. C. Graves (MIT) for helpful discussions.

Literature Cited

1. U.S. Census Bureau. Manufacturing and trade inventories and sales. Last accessed 8/30/2009.
2. Tayur S. *What is missing to enable optimization of inventory deployment and supply planning?* In: Grossmann IE, McDonald C, editors. *FOCAPO 2003*. Florida: Coral Springs, 2004, pp. 35.
3. Chopra S, Meindl P. *Supply Chain Management: Strategy, Planning and Operation*. Saddle River, NJ: Prentice Hall, 2003.
4. Grossmann IE. Enterprise-wide optimization: a new frontier in process systems engineering. *AIChE J*. 2005;51:1846–1857.
5. Zipkin PH. *Foundations of Inventory Management*. Boston, MA: McGraw-Hill, 2000.
6. You F, Grossmann IE. Mixed-integer nonlinear programming models and algorithms for large-scale supply chain design with stochastic inventory management. *Ind Eng Chem Res*. 2008;47:7802–7817.
7. Inderfurth K. Safety stock optimization in multistage inventory systems. *Int J Prod Econ*. 1991;24:103–113.
8. Inderfurth K. Valuation of leadtime reduction in multistage production systems. In: Fandel G, Gulledge T, Jones A, editors. *Operations Research in Production Planning and Inventory Control*. Berlin, Germany: Springer, 1993:413–427.
9. Inderfurth K, Minner S. Safety stocks in multistage inventory systems under different service measures. *Eur J Oper Res*. 1998;106: 57–73.
10. Minner S. Strategic safety stocks in reverse logistics supply chains. *Int J Prod Econ*. 2001;71:417–428.
11. Tsiakis P, Shah N, Pantelides CC. Design of multiechelon supply chain networks under demand uncertainty. *Ind Eng Chem Res*. 2001;40:3585–3604.
12. Bok J-K, Grossmann IE, Park S. Supply chain optimization in continuous flexible process networks. *Ind Eng Chem Res*. 2000; 39:1279–1290.
13. Verderame PM, Floudas CA. Integrated operational planning and medium-term scheduling of a large-scale industrial batch plants. *Ind Eng Chem Res*. 2008;47:4845–4860.
14. Relvas S, Matos HA, Barbosa Povoa APFD, Fialho J, Pinheiro AS. Pipeline scheduling and inventory management of a multiproduct distribution oil system. *Ind Eng Chem Res*. 2006;45:7841–7855.
15. Varma VA, Reklaitis GV, Blau GE, Pekny JF. Enterprise-wide modeling and optimization—an overview of emerging research challenges and opportunities. *Comput Chem Eng*. 2007;31:692–711.
16. Schulz EP, Diaz MS, Bandoni JA. Supply chain optimization of large-scale continuous process. *Comput Chem Eng*. 2006;29:1305–1316.
17. Gupta A, Maranas CD, McDonald CM. Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management. *Comput Chem Eng*. 2000;24:2613–2621.
18. Chen C-L, Lee W-C. Multi-objective optimization of multi-echelon supply chain networks with uncertain product demands and prices. *Comput Chem Eng*. 2004;28:1131–1144.
19. Lim M-F, Karimi IA. Resource-constrained scheduling of parallel production lines using asynchronous slots. *Ind Eng Chem Res*. 2003;42:6832–6842.
20. Jackson JR, Grossmann IE. Temporal decomposition scheme for nonlinear multisite production planning and distribution models. *Ind Eng Chem Res*. 2003;42:3045–3055.
21. Daskin MS, Coullard C, Shen Z-JM. An inventory-location model: formulation, solution algorithm and computational results. *Ann Oper Res*. 2002;110:83–106.
22. Shen Z-JM, Coullard C, Daskin MS. A joint location-inventory model. *Transp Sci*. 2003;37:40–55.
23. Ozsen L, Daskin MS, Coullard CR. Capacitated warehouse location model with risk pooling. *Nav Res Logist*, Vol. 5, pp. 295–312.
24. Sourirajan K, Ozsen L, Uzsoy R. A single-product network design model with lead time and safety stock considerations. *IIE Trans*. 2007;39:411–424.
25. Bossert JM, Willems SP. A periodic-review modeling approach for guaranteed service supply chains. *Interfaces*. 2007;37:420–435.
26. Graves SC, Willems SP. Optimizing strategic safety stock placement in supply chains. Manuf Serv Oper Manage. 2000;2:68–83.
27. Graves SC, Willems SP. Supply chain design: safety stock placement and supply chain configuration. In: de Kok AG, Graves SC, editors. *Handbooks in Operations Research and Management Science*, Vol. 11. North-Holland, Amsterdam: Elsevier, 2003:95–132.
28. Graves SC, Willems SP. Optimizing the supply chain configuration for new products. *Manage Sci*. 2005;51:1165–1180.
29. Humair S, Willems SP. Optimizing strategic safety stock placement in supply chains with clusters of commonality. *Oper Res*. 2006; 54:725–742.
30. Simpson KF. In-process inventories. *Oper Res*. 1958;6:863–873.
31. Eppen G. Effects of centralization on expected costs in a multi-echelon newsboy problem. *Manage Sci*. 1979;25:498–501.
32. Federgruen A, Zipkin PH. Computational issues in an infinite-horizon, Multiechelon Inventory Model. *Oper Res*. 1984;32:818–836.
33. Rao U, Scheller-Wolf A, Tayur S. Development of a rapid-response supply chain at caterpillar. *Oper Res*. 1999;48:189–204.
34. Little JDC. A proof of the queueing formula $L = \lambda W$. *Oper Res*. 1961;9:383–387.
35. Available at: http://www.effectiveinventory.com.
36. Campbell AM, Clarke LW, Savelsbergh MWP. Inventory routing in practice. In: Toth P, Vigo D, editors. *The Vehicle Routing Problem*. Philadelphia, PA: SIAM, 2001:309–330.
37. Savelsbergh M, Song J-H. Inventory routing with continuous moves. *Comput Oper Res*. 2007;34:1744–1763.
38. Hübner R. *Strategic Supply Chain Management in Process Industries: An Application to Specialty Chemicals Production Network Design*. Heidelberg: Springer-Verlag, 2007.

39. Glover F. Improved linear integer programming formulations of nonlinear integer problems. *Manage Sci*. 1975;22:455–460.
40. Falk JE, Soland RM. An algorithm for separable nonconvex programming problems. *Manage Sci*. 1969;15:550–569.
41. Sahinidis NV. BARON. A general purpose global optimization software package. *J Glob Optim*. 1996;8:201–205.
42. Brooke A, Kendrick D, Meeraus A, Raman R. *GAMS—A User's Manual*. Washington, DC: GAMS Development Corp., 1998.
43. Croxton KL, Gendron B, Magnanti TL. A comparison of mixed-integer programming models for nonconvex piecewise linear cost minimization problems. *Manage Sci*. 2003;49:1268–1273.
44. Wicaksono DS, Karimi IA. Piecewise MILP under- and overestimators for global optimization of bilinear programs. *AIChE J*. 2008;54:991–1008.
45. Bergamini ML, Grossmann IE, Scenna N, Aguirre P. An improved piecewise outer-approximation algorithm for the global optimization of MINLP models involving concave and bilinear terms. *Comput Chem Eng*. 2008;32:477–493.
46. Lin X, Floudas CA, Kallrath J. Global solution approach for a nonconvex MINLP problem in product portfolio optimization. *J Glob Optim*. 2005;32:417–431.
47. Magnanti TL, Shen Z-JM, Shu J, Simchi-Levi D, Teo C-P. Inventory placement in acyclic supply chain networks. *Oper Res Lett*. 2006;34:228–238.
48. Vidyarthi N, Celebi E, Elhedhli S, Jewkes E. Integrated production-inventory-distribution system design with risk pooling: model formulation and heuristic solution. *Transp Sci*. 2007;41:392–408.
49. Padberg MW. Approximating separable nonlinear functions via mixed zero-one programs. *Oper Res Lett*. 2000;27:1–5.
50. Raman R, Grossmann IE. Modeling and computational techniques for logic-based integer programming. *Comput Chem Eng*. 1994;18:563–578.
51. Balas E. Disjunctive programming and a hierarchy of relaxations for discrete continuous optimization problems. *SIAM J Algebr Discrete Methods*. 1985;6:466–486.
52. Fisher ML. The Lagrangean relaxation method for solving integer programming problems. *Manage Sci*. 1981;27:1–18.
53. Fisher ML. An application oriented guide to Lagrangian relaxation. *Interfaces*. 1985;15:2–21.
54. Beasley JE. Lagrangean heuristics for location problems. *Eur J Oper Res*. 1993;65:383–399.
55. You F, Grossmann IE. Optimal design and operational planning of responsive process supply chains. In: Papageorgiou L, Georgiadis M, editors. *Process System Engineering: Volume 3: Supply Chain Optimization*. Weinheim: Wiley-VCH, 2007:107–134.
56. You F, Grossmann IE. Design of responsive supply chains under demand uncertainty. *Comput Chem Eng*. 2008;32:2839–3274.

Appendix: Variable Bounds

In this Appendix, we derive the upper bounds of variables in the models. All the upper bounds of variables are denoted with the superscript U.

The maximum net lead time (N_j^U) of DC j is the maximum value of the sum of service time of plant i and the order processing time from plant i to DC j. It means that when the guaranteed service time of DC j is zero and it is served by plant i that has the maximum ($SI_i + t_{ij}$), the net lead time N_j equals to the maximum value,

$$N_j^U = \max_{i \in I}\{SI_i + t_{ij}\} = \max_{i \in I}\{\bar{S}_{ij}\}, \quad \forall j \tag{A1}$$

The maximum value of guaranteed service time of DC j is equal to the maximum value of its net lead time. It means that when the net lead time of DC j is zero and it is assigned to plant i that has the maximum ($SI_i + t_{ij}$), the net lead time S_j equals to the maximum value,

$$S_j^U = N_j^U = \max_{i \in I}\{\overline{S}_{ij}\}, \quad \forall j \tag{A2}$$

From constraints (13) and (16), it is easy to see that the maximum net lead time of customer demand zone k is as follows,

$$L_k^U = \max\left\{0, \max_{j \in J}\left\{S_j^U + t_{jk}\right\} - R_k\right\}, \quad \forall j \tag{A3}$$

The upper bounds of the auxiliary variables are easy to derive,

$$SZ_{jk}^U = S_j^U, SZ1_{jk}^U = S_j^U, \quad \forall j,k \tag{A4}$$

$$NZ_{jk}^U = N_j^U, NZ1_{jk}^U = N_j^U, \forall j,k \tag{A5}$$

$$NZV_j^U = \sum_{k \in K} \sigma_k^2 \cdot N_j^U, \quad \forall j \tag{A6}$$

Manuscript received Nov. 30, 2008, and revision received May 16, 2009.