Design of Responsive Process Supply Chains under Demand Uncertainty

Fengqi You and Ignacio E. Grossmann[*]

Center for Advanced Process Decision-making, Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

ABSTRACT

This paper addresses the optimization of supply chain design and planning under the criteria of responsiveness and economics with the presence of demand uncertainty. The supply chain consists of multi-site processing facilities and corresponds to a multi-echelon production network with both dedicated and multiproduct plants. The economic criterion is measured in terms of net present value, while the criterion for responsiveness accounts for transportation times, residence times, cyclic schedules in multiproduct plants, and inventory management. By using a probabilistic model for stockout, the expected lead time is proposed as the quantitative measure of supply chain responsiveness. The probabilistic model can also predict the safety stock levels by integrating stockout probability with demand uncertainty. These are all incorporated into a multi-period mixed-integer nonlinear programming (MINLP) model, which takes into account the selection of manufacturing sites and distribution centers, process technology, production levels, scheduling and inventory levels. The problem is formulated as a bi-criterion optimization model that maximizes the net present value and minimizes the expected lead time. The model is solved with the ε-constraint method and produces a Pareto-optimal curve that reveals how the optimal net present value, supply chain network structure and safety stock levels, change with different values of the expected lead time. A hierarchical algorithm is also proposed based on the decoupling of different decision-making levels (strategic and operational) in the problem. The application of this model and the proposed algorithm are illustrated with two examples of polystyrene supply chains.

Keywords: Supply Chain Management, Responsiveness, Lead Time, Demand Uncertainty, Safety Stock, MINLP

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

1. INTRODUCTION

There is a growing recognition that individual businesses no longer compete as stand-alone entities, but rather as supply chains whose success or failure is ultimately determined in the marketplaces by the end consumers (Christopher and Towill, 2001). For better customer satisfaction and market understanding, companies are striving to achieve the best performance from their supply chains by different measures, of which accurate demand forecasting, inventory and responsive supply chain are three key components (Fisher, 1997).

Quick response enables supply chains to meet the customer demands for ever-shorter lead times, and to synchronize the supply to meet the peaks and troughs of demand (Sabath, 1998). Nowadays responsive supply chains have become keys to competitive success and survival (Fisher, 1997; Christopher, 2000, 2005) due to the increasing pressure to reduce costs and inventories for competitions in the global marketplace (Grossmann, 2005). Although sophisticated methods such as time series have been developed to improve the forecasting accuracy, uncertainties in demand are unavoidable due to ever changing market conditions. In supply chains, inventory improves the service by helping deal with demand uncertainty and providing flexibility, although it can be costly (Chase and Aquilano, 1995).

In this work, we consider the design of a responsive supply chain with integration of inventory and safety stock under demand uncertainty. The supply chain consists of multi-site processing facilities and corresponds to a multi-echelon production network with both dedicated and multiproduct facilities. The major goal is to determine the processes that are to be integrated in the supply chain network with their corresponding suppliers, distribution centers and the associated transport links between them. The major considerations in the design are the supply chain responsiveness and profitability. Profitability is expressed in terms of net present value, while responsiveness accounts for transportation times, residence times, cyclic schedules in multiproduct plants, and inventory management. By using a probabilistic model for stockout, a quantitative characterization of responsiveness for supply chain networks is presented, which measures the expected response time or expected lead time to changes under uncertain demands with integration of inventory and safety stock level. The probabilistic model can also predict the safety stock levels by integrating stockout probability with demand uncertainty. These are incorporated into a multi-period mixed-integer non-linear programming (MINLP) model, which takes into account the selections of manufacturing sites and distribution centers, process technology, production levels, scheduling and inventory levels. The problem is formulated as a bi-criterion optimization model in which the objectives are to maximize the net present value and to minimize the expected lead time. Aside from relying on the ε-constraint method to generate the Pareto-optimal curve, a hierarchical algorithm is also proposed for the solution of the resulting large-scale nonconvex MINLP model based on the decoupling of the different decision-making levels (strategic and operational) identified in our problem. The application of this model is illustrated through two examples of polystyrene supply chains.

The rest of the paper is organized as follows. We briefly review related literature in the next section, and the main quantitative characteristics of responsiveness of process supply chain networks are discussed in Section 3. A formal problem statement along with the key assumptions is given in Section 4, while the proposed mathematical model is described in Section 5. Section 6 presents a hierarchical solution approach, and its applicability is demonstrated by two illustrative examples in Section 7. Finally, concluding remarks are presented in Section 8.

2. LITERATURE REVIEW

Most of the “responsiveness” literature for supply chains tends to be qualitative and conceptual, and has not been subjected to the kind of quantitative analysis that this paper intends to address. There are, however, several related works that offer relevant insights. Forrester (1961) first illustrated in a series of case studies the effect of dynamics in industrial systems, which gives rise to the “bullwhip effect”. Lee et al (1997) further demonstrated that the “bullwhip” effect is a consequence of the information delay due to the structure of supply chains, and the severity of this effect is positively related to lead times. Responsiveness in the wider supply chain context has been discussed by Fisher (1997), who argues that the product characteristics (innovative or functional) and life cycles need to be linked to the layout and functions (conversion and market mediation) of the supply chain. He also pointed out the need of reducing the lead time, which enables quick response to unpredictable demand to minimize stockouts, markdowns and obsolete inventory. Matson et al (1999) discussed concepts and issues associated with responsiveness in production and illustrate the audit tools they proposed from a case study in the steel industry. Recently, several conceptual models on supply chain responsiveness have been proposed. Christopher and Towill (2001) integrate lead time and agility to highlight the differences in their approach, and combined them to propose an integrated hybrid strategy for designing cost-effective responsive supply chains with seamless connection between manufacturing and logistics. In a later work, Yusuf et al (2004) have reviewed emerging patterns for creating responsive supply chains based on survey research driven by a conceptual model. Holweg (2005) proposed in his paper that product, process and volume are three key factors that determine the responsiveness of a supply chain system, and provided guidelines on how to align the supply chain strategy to these three factors in order to balance responsiveness to customer demand and supply chain efficiency. An examination of supply chain systems in process industries from a responsiveness view point was carried out by Shaw et al (2005). These authors also proposed a conceptual management strategy to improve responsiveness.

Another group of relevant papers to be considered are on supply chain design and operation. A general review of this area is given in Kok and Graves (2003), and a specific review for supply chains in process industries is presented by Shah (2005). Some recent works include the following. Tsiakis et al (2001) presented a supply chain design model for the steady-state continuous processes. Their supply chain model was developed based on determining the connection between multiple markets and multiple plants with fixed locations. Jackson and Grossmann (2003) presented a temporal decomposition scheme based on Lagrangean decomposition for a nonlinear programming problem model for multi-site production planning and distribution, where nonlinear terms arise from the relationship between production and physical properties or blending ratios. Schulz et al (2005) described two multi-period MINLP models for short term planning of petrochemical complexes. Linearization techniques were applied to reformulate the nonconvex bilinear constraints as MILP models. Recently, Sousa et al (2006) presented a two stage procedure for supply chain design with responsiveness testing. In the first stage, they design the supply chain network and optimize the production and distribution planning over a long time horizon. In the second stage, responsiveness of the first stage decisions are assessed using the service level to the customers (i.e. delay in the order fulfillment). However, all these models consider supply chain networks with only dedicated processes. Multi-product batch plants or flexible processes were not taken into account, and hence no scheduling models were included.

There are works on supply chain optimization with consideration of flexible processes in the production network, but most of them are restricted to planning and scheduling for a given facility in a fixed location without extension to the multisite supply chain network design problems. Bok et al (2000) proposed a multiperiod supply chain optimization model for operational planning of continuous flexible process networks, where sales, intermittent deliveries, production shortfalls, delivery delays, inventory profiles and job changeovers are taken into account. A bilevel decomposition algorithm was proposed, which reduced the computational time significantly. Kallrath (2002) described a tool for simultaneous strategic and operational planning in a multi-site production network, where key decisions include operating modes of equipment in each time period, production and supply of products, minor changes to the infrastructure and raw material purchases and contracts. A multi-period model is formulated where equipment may undergo one change of operation mode per period. The standard material balance equations are adjusted to account for the fact that transportation times are much shorter than the period durations. Chen et al (2003) presented a multi-product, multistage and multiperiod production and distribution planning model. They also proposed a two-phase fuzzy decision making method to obtain a compromise solution among all participants of the multi-enterprise supply chain.

To account for product demand fluctuation and to obtain a better understanding of how uncertainty affects the supply chain performance, a number of approaches have been proposed in the chemical engineering literature for the quantitative treatment of uncertainty in the design, planning and scheduling problems. A classification of different areas of uncertainty for batch chemical plant design is suggested by Subrahmanyam et al (1994), where uncertainty in prices and demand, equipment reliability and manufacturing are taken into account. The authors used the popular scenario-based approach, which attempts to capture uncertainty by representing it in terms of a number of discrete realizations of the stochastic quantities, constituting distinct scenarios. Each complete realization of all uncertain parameters gives rise to a scenario and all the possible future outcomes are forecasted and taken into account through the use of scenarios. The objective is to find a robust solution which performs well under all scenarios. The scenario-based approach provides a straightforward way to implicitly account for uncertainty (Liu and Sahinidis, 1996). Its main drawback is that it typically relies on either a priori forecasting of all possible outcomes, or the explicit/implicit discretizations of a continuous multivariate probability distribution by methods of Gaussian quadrature integration or Monte Carlo sampling, which can result in an exponential number of scenarios.

Another popular method to address the uncertainty is using probabilistic approaches, which consider the uncertainty aspect of the supply chain by treating one or more parameters as random variables with known probability distribution. By introducing a certain number of nonlinear terms from continuous distribution, this approach can lead to a reasonable size of the deterministic equivalent representation of the probabilistic model, circumventing the need for explicit/implicit discretization or sampling. As argued by Zimmermann (2000), the choice of the appropriate method is context-dependent, with no single theory being sufficient to model all kinds of uncertainty. Recently, Shen et al (2003) proposed a novel approach to deal with the demand uncertainty for facility location problems. In their work, demand uncertainty is hedged by holding certain amount of safety stocks, and the safety stock level is decided by the demand variance and a specific service level. By adding the safety stock term in the model, the recourse problem for uncertain parameters is avoided.

Thus, these papers either focus only on the long-term strategic design models, or else are restricted to short-term planning and scheduling models. Hence, no quantitative analyses are available for responsive supply chains under demand uncertainty.

3. SUPPLY CHAIN RESPONSIVENESS

A major goal of this paper is to develop a quantitative definition of supply chain responsiveness with integration of inventory under demand uncertainty. Responsiveness is defined as the ability of a supply chain to respond rapidly to the changes in demand, both in terms of volume and mix of products (Christopher, 2000; Holweg, 2005). Since the definition is qualitative, we need to find a quantitative measure for supply chain responsiveness.

Lead time is the time of a supply chain network to respond to customer demands. We will consider have the lead time to the one corresponding to the longest time for all paths. Furthermore, in the worst case lead time corresponds to the response time when there are zero inventories. This was used as a measure of responsiveness in our previous work (You and Grossmann, 2007). As shown in Figure 1, a supply chain network with long lead time implies that its responsiveness is low, and vice versa. In this work, since we consider uncertain demands and safety stocks, expected lead time will be used as the measure of supply chain responsiveness. Thus, the challenge is to quantitatively define the expected lead time with integration of the supply chain network structure, inventory and operation details under demand uncertainty. In the following sections we will first review some key definitions, and then present our new proposal for expected lead time.

3.1. Time Delays in Simple Linear Supply Chains

Consider first the case of a simple linear supply chain as given in Figure 2 that consists of one supplier, several manufacturing plants, one distribution center and one customer. The material flow starts from the supplier by way of manufacturing plants and distribution center(s)[1] and ends at the customer. The information transfers in the reverse direction. In this work, we assume that information transfers instantaneously, thus the time delay for the entire supply chain comes from the time delay incurred by the transfer of materials.

From Figure 2, we can see that the transfer of material flow is delayed by both transportation and production. The transportation delays between supplier, plants, distribution center and customer are equal to the associated transportation times between them (Figure 3). The production delay by each single-product plant is equal to the residence time of the products. The production delay in a multiproduct plant is more complicated as it needs to account for scheduling details. Therefore, we will leave this to the discussion in Section 5.3 cyclic scheduling.

The time delay of the entire linear supply chain can be partitioned into two parts, delivery lead time and production lead time. The delivery lead time is defined as the time to transfer the product from distribution center to the customer, and the production lead time is the time that the material flow takes to transfer from supplier to the distribution center (Figure 4). Thus, the delivery lead time is equal to the transportation time from the distribution center to the customer; the production lead time is equal to the summation of all the time delays incurred by transportation and production from the supplier to the distribution center. Note that this characterization of the time delay of product activities is similar to the lean tool "value stream mapping" (Voekel and Chapman, 2003).

3.2. Expected Lead Time of Simple Linear Supply Chain

If there is sufficient inventory in the distribution center to handle the demand changes, the lead time should be equal to the time to transfer products from distribution center to customers, which is the delivery lead time. If there is no sufficient inventory in the distribution center to handle the demand changes (i.e. the product is out of stock), the worst case is when there is no extra stock for raw materials or intermediate products, and the only way is to go back to reorder the raw materials from the supplier. In this way, after a series of transportation and production steps, the product will be finally shipped to the

[1] A linear supply chain can have more than one distribution centers. For a supply chain network, the distribution system can be multi-echelon. For simplicity, we only consider one distribution center for the linear supply chain in this work.

customer. Therefore, in this case the lead time would be equal to the production lead time plus the delivery lead time. Because the demand is uncertain, there is a probability that the product will be out of stock. We denote $Prob$ as the probability of stock out, L_D as the delivery lead time, and L_P as the production lead time. Thus, if the product is out of stock, the lead time is the production lead time plus delivery lead time ($L_P + L_D$) with the stock out probability $Prob$. If the product is not stocked out, the lead time is delivery lead time L_D with the probability of $(1 - Prob)$. Therefore, we have that the expected lead time $E(L)$ of this simple linear supply chain is given by:

$$E(L) = Prob \cdot (L_P + L_D) + (1 - Prob) \cdot L_D$$

which can be arranged as,

$$E(L) = Prob \cdot L_P + L_D$$

This implies that the expected lead time is equal to the delivery lead time plus the expected production lead time (the stock out probability times production lead time).

3.3. Expected Lead Time of Process Supply Chain Network

Although a general process supply chain network is more complex than a simple linear supply chain, we can still "decompose" the supply chain network into paths of material flows that start from a supplier, and end at a customer, by way of several plants and distribution centers (as shown in Figure 5). For simple supply chain networks we can determine all the paths by inspection. For complex supply chain networks, various pathfinding algorithms (such as the one by Lengauer and Tarjan, 1979) that can be used to figure out all the possible paths in the supply chain network. It is worth mentioning that value stream mapping (Voekel and Chapman, 2003), a lean manufacturing tool widely used in industries, can also analyze the process and find all the possible paths in a supply chain network automatically.

Thus, each path is equivalent to a simple linear supply chain for which the expected lead time can be easily determined. We define the maximum expected lead time of all possible paths as the total expected lead time of the entire supply chain. One could also consider a weighted expected lead time according to the importance of customers. But for simplicity, we consider the former definition in this work.

4. PROBLEM STATEMENT

An integrated approach is needed in order to consider simultaneously supply chain network design, production planning and scheduling, demand uncertainty and inventory management to resolve the trade-offs between economics and responsiveness in an optimal manner. The problem of responsive supply chain design under demand uncertainty can be formally stated as follows.

Given is a potential process supply chain network (PSCN) that includes possible suppliers, manufacturing sites, distribution centers and customers as shown in Figure 6. Also, a set of processes and a time horizon consisting of a number of time periods are given. The processes may be dedicated or flexible. Flexible processes are multi-product processes that operate under different production schemes using different raw materials and/or producing different products. Furthermore, changeovers are incurred between products (Figure 7). For all the production schemes, mass balances are expressed linearly in terms of the main product's production. The investment costs for installing the plants and distribution centers are expressed by a cost function with fixed charges. There could be different transportation modes, continuous (e.g. pipelines) or discrete (e.g. barges, rail cars, vessel), for each path that connects the suppliers, plant sites, distribution centers and customers. For simplicity, we will assume that the transportation of materials in this problem is continuous. Thus, fixed charge cost functions provide good estimations of transportation costs. The transportation time of each route and the residence time of each product are assumed to be known. The PSCN involves a set of chemicals, which includes raw materials, intermediate products and final products. Prices for raw materials and final products are assumed to be known over the entire time horizon. Raw materials are subject to availability constraints (i.e., within lower and upper bounds). Demands in each time period are uncertain and are described by a specified continuous probability distribution function. Most of the inventories and all the safety stocks are hold in distribution centers, while plant sites also maintain a certain amount of inventory. Unit inventory cost for raw materials, intermediate products and final products are also given.

In order to design a responsive supply chain, one objective is to minimize the expected lead time of the entire supply chain network. From the economic aspect, the other objective function is to maximize the net present value (NPV) over the specified time horizon. The income from sales, along with the investment, operating, transportation and purchase costs are taken into account in the NPV objective function.

Since the two conflicting objectives need to be optimized simultaneously, the corresponding problem yields an infinite set of trade-off solutions. These solutions are Pareto-optimal in the sense that it is impossible to improve both objective functions simultaneously (Halsall and Thibault, 2006). This implies that any designs for which the expected net present value and the expected lead time can be improved simultaneously are "inferior" solutions that do not belong to the Pareto-optimal curve. The aim of this problem is to determine the supply chain network configurations and operational decisions that define the Pareto optimal curve by maximizing NPV and minimizing the expected lead time.

5. MODEL

The model will be formulated as a multi-period MINLP problem, which will predict the detailed design decisions, production and inventory profiles and schedules of the PSCN with different specifications of the expected lead time. A list of indices, sets, parameters and variables are given in the Appendix. Four types of constraints are included in this model. They are network structure constraints, operational planning constraints, cyclic scheduling constraints and probabilistic constraints.

Constraints (1) to (12) determine the network structure, constraints (13) to (23) refer to the operational planning constraints, constraints (24) to (36) are used for the cyclic scheduling of multi-product plants, constraints (37) to (40) are probabilistic constraints. Finally, inequalities (41) to (44) define the expected lead time and equation (46) defines the net present value, both of which are objective functions to be optimized.

5.1. Network Structure Constraints

To determine the topology of the network structure and model the selection of suppliers, plant sites, distribution centers, together with the transportation links between them, the binary variables ($Y^P_{k,i}$, Y^m, $Y^I_{k,ls}$, $Y^N_{k',k}$, $Y^O_{k,m}$, $Y^S_{m,ld}$) for plants, distribution centers and transportation links are introduced for the design decisions. Four types of network structure constraints are applied to represent the relationships between each node in the supply chain network.

5.1.1. Supplier – Plant (Site)

The first type of relationship is between suppliers and plant sites (Figure 8). A transportation link for raw material j from supplier ls to plant site k exists ($Y^I_{k,ls}$), only if at least one plant that consumes raw material j exists in site k ($Y_{k,i}^{P}$). The relationships discussed above can be expressed by the following logic proposition:

$$Y_{k,ls}^{I} \Rightarrow \vee_{i\in I_j} Y_{k,i}^{P}, \qquad \forall k\in K, ls\in LS_j, j\in J \qquad (1.a)$$

These logic propositions can be further transformed into inequalities as described in Raman and Grossmann (1993).

$$Y_{k,ls}^{I} \le \sum_{i\in I_j} Y_{k,i}^{P}, \qquad \forall k\in K, ls\in LS_j, j\in J \qquad (1)$$

On the plant side, if a plant that consumes raw material j is set up ($Y_{k,i}^{P}$), at least one transportation link from the supplier ls to site k must be selected ($Y_{k,ls}^{I}$). The logic propositions are:

$$Y_{k,i}^{P} \Rightarrow \vee_{ls\in LS_j} Y_{k,ls}^{I}, \qquad \forall k\in K_i, i\in I_j, j\in J \qquad (2.a)$$

which can be transformed into inequalities:

$$Y_{k,i}^{P} \le \sum_{ls\in LS_j} Y_{k,ls}^{I}, \qquad \forall k\in K_i, i\in I_j, j\in J \qquad (2)$$

5.1.2. Input and output relationship of a plant

The second type of network structure relationship is the input and output relationship of a plant (Figure 9). This type of relationship is somewhat more complicated than the previous one because the inter-site transportation must be taken into account. If an inter-site transportation link from site k' to site k is installed for chemical j ($Y_{k',k}^{N}$), it implies that at least one plant i' in site k' is installed that produces chemical j ($Y_{k',i'}^{P}$), and also at least one plant i in site k is installed that consumes chemical j ($Y_{k,i}^{P}$),

$$Y_{k',k}^{N} \le \sum_{i'\in O_j} Y_{k',i'}^{P}, \qquad \forall k,k'\in K, j\in (JP_{k',i'} \cap JM_{k,i}) \qquad (3)$$

$$Y_{k',k}^{N} \le \sum_{i\in I_j} Y_{k,i}^{P}, \qquad \forall k,k'\in K, j\in (JP_{k',i'} \cap JM_{k,i}) \qquad (4)$$

If a plant i in site k is installed ($Y_{k,i}^{P}$), that consumes chemical j, then site k is connected to one of the suppliers of chemical j denoted as ls ($Y_{k,ls}^{I}$), or connected to another site k' that produces chemical j ($Y_{k',k}^{N}$), or there is another plant i' in site k that produces chemical j ($Y_{k,i'}^{P}$). The logic propositions can be written as follows:

$$Y_{k,i}^{P} \Rightarrow \vee_{ls\in LS_j} Y_{k,ls}^{I} \vee_{k'\in K_{i'}} Y_{k',k}^{N} \vee_{i'\in O_j} Y_{k,i'}^{P}, \qquad \forall k\in K_i, i\in I_j, j\in JM_{k,i} \qquad (5.a)$$

which can be transformed into linear inequalities as,

$$Y_{k,i}^{P} \le \sum_{ls \in LS_j} Y_{k,ls}^{I} + \sum_{k' \in K_{i'}} Y_{k',k}^{N} + \sum_{i' \in O_j} Y_{k,i'}^{P}, \qquad \forall k \in K_i, i \in I_j, j \in JM_{k,i} \tag{5}$$

Similarly, if the chemical j is produced by plant i in site k, then at least one other plant i' in the same site is installed that consumes chemical j ($Y_{k,i'}^{P}$), or there is at least one transportation link to a distribution center ($Y_{k,m}^{O}$), or to another site ($Y_{k,k'}^{N}$) that consumes chemical j:

$$Y_{k,i}^{P} \le \sum_{m \in M} Y_{k,m}^{O} + \sum_{k' \in K_{i'}} Y_{k,k'}^{N} + \sum_{i' \in I_j} Y_{k,i'}^{P}, \qquad \forall k \in K_i, i \in O_j, j \in JP_{k,i} \tag{6}$$

Constraints (5) are defined for raw materials and intermediate products, and constraints (6) are defined for intermediate products and final products.

5.1.3. Plant (site) – Distribution Center

The third type of relationship is between plant sites and distribution centers as shown in Figure 10. A transportation link for product j from plant site k to distribution center m exist ($Y_{k,m}^{O}$), only if at least one plant that consumes raw material j exists in site k ($Y_{k,i}^{P}$). On the plant side, if a plant that consumes raw material j is set up, there should be at least one link from the distribution center m to site k exists. Similarly, transforming the corresponding logic propositions, leads to the following inequalities:

$$Y_{k,m}^{O} \le \sum_{i \in O_j} Y_{k,i}^{P}, \qquad \forall k \in K, m \in M, j \in JP_{k,i} \tag{7}$$

$$Y_{k,i}^{P} \le \sum_{m \in M} Y_{k,m}^{O}, \qquad \forall k \in K, i \in O_j, j \in JP_{k,i} \tag{8}$$

5.1.4. Input and output relationship of distribution center

The last type of network structure relationship is the input and output relationship of a distribution center as in Figure 11. A transportation link from plant site k to distribution center m exists, only if the distribution center m exists (Y^{m}). A transportation link from distribution center m to customer ld exists ($Y_{m,ld}^{S}$), only if the distribution center m exists. The relationships can be expressed as,

$$Y_{k,m}^{O} \le Y^{m}, \qquad \forall k \in K, m \in M \tag{9}$$

$$Y_{m,ld}^{S} \le Y^{m}, \qquad \forall m \in M, ld \in LD \tag{10}$$

On the other hand, if a distribution center m is set up, at least one transportation link from the plant site k to distribution center m ($Y^O_{k,m}$) and at least one transportation link from distribution center m to customer ld ($Y^S_{m,ld}$) must exist.

$$Y^m \le \sum_{k \in K} Y^O_{k,m}, \qquad \forall m \in M \tag{11}$$

$$Y^m \le \sum_{ld \in LD} Y^S_{m,ld}, \qquad \forall m \in M \tag{12}$$

5.2. Operational Planning Constraints

In the operational planning model investment in plant capacity, and purchases, sales, production, transportation and mass balance relationships are considered together with the constraints of these activities due to the supply chain structure.

5.2.1. Production Constraints

All the chemical flows $W_{k,i,j,s,t}$ of chemical j associated with production scheme s in plant i of site k, other than the main product j', are given by the mass balance coefficient $\mu_{i,j,s}$. The following equation relate input or inlet flow of chemical j ($W_{k,i,j,s,t}$) with the output of a main product j' ($W_{k,i,j',s,t}$) of each process,

$$W_{k,i,j,s,t} = \mu_{i,j,s} W_{k,i,j',s,t}, \qquad \forall k \in K_i, i \in I_j, j \in J_{i',s}, j' \in JP_{i,s}, s \in S_i, t \in T \tag{13}$$

The production amount should not exceed the design capacity $Q_{k,i}$ defined by the main product j for each process:

$$W_{k,i,j,s,t} \le \rho^1_{i,s} Q_{k,i} Lenp_t, \qquad \forall k \in K_i, i \in O_j, j \in JP_{k,i}, s \in S_i, t \in T \tag{14}$$

where $Lenp_t$ is the length of time period t and $\rho^1_{i,s}$ is the relative production amount of main product j of production scheme s in plant i in turns of capacity

The formulation is based on the assumption that there are no capacity expansions over the entire time horizon. However, multi-period capacity planning events and decisions can readily be considered by suitably modifying the formulation constraints along with the detailed capacity investment constraints (Sahinidis et al, 1989) which are not detailed here.

For flexible processes, the maximum production rate $r_{k,i,s}$ of the each product s in plant i of site k is proportional to the capacity of the plant (see Norton and Grossmann, 1994),

$$r_{k,i,s} = \rho^2_{i,s} Q_{k,i}, \qquad \forall k \in K_i, i \in O_j, j \in JP_{k,i}, s \in S_i \tag{15}$$

where $\rho_{i,s}^2$ is the relative maximum production rate of main product of production scheme s in turns of the capacity of plant i

5.2.2. Mass Balance Constraints

The mass balance for chemical j in manufacturing site k at time period t is given as follows:

$$\sum_{ls \in LS} P_{k,j,ls,t} + \sum_{k' \in K} TR_{k,k',j,t} + \sum_{i \in O_j} \sum_{s \in S} W_{k,i,j,s,t} = \sum_{m \in M} F_{k,j,m,t} + \sum_{k \in K} TR_{k,k',j,t} + \sum_{i' \in I_j} \sum_{s \in S} W_{k,i',j,s,t} ,$$

$$\forall k \in K_i, j \in J, t \in T \qquad (16)$$

where $P_{k,j,ls,t}$ is purchase amount, $TR_{k,k',j,t}$ is the inter-site shipping amount and $W_{k,i,j,t}$ is the production amount.

The mass balance for chemical j in distribution center m at time period t is given as follows:

$$\sum_{k \in K} F_{k,j,m,t} = \sum_{ld \in LD} S_{j,m,ld,t} , \qquad \forall j \in J, m \in M, t \in T \qquad (17)$$

where $F_{k,j,m,t}$ is the shipping amount from production site to distribution center and $S_{j,m,ld,t}$ is the sale amount.

5.2.3. Inventory Constraints

The total available amount of chemical j for customer ld ($QS_{j,ld,t}$) should be equal to the safety stocks ($I_{j,m,ld,t}$) committed to this customer in all distribution centers plus the sale amount.

$$QS_{j,ld,t} = \sum_{m \in M} I_{j,m,ld,t} + \sum_{m \in M} S_{j,m,ld,t} , \qquad \forall j \in J, ld \in LD, t \in T \qquad (18)$$

The sale amount of chemical j to each customer ld at time period t should be equal to the associated target demand $d_{j,ld,t}^m$ (usually the target is equal to the mean value of uncertain demand),

$$\sum_{m \in M} S_{j,m,ld,t} = d_{j,ld,t}^m , \qquad \forall j \in J, ld \in LD, t \in T \qquad (19)$$

The total available amount of chemical j will be less than the upper bound of the demand,

$$QS_{j,ld,t} \le d_{j,ld,t}^U , \qquad \forall j \in J, ld \in LD, t \in T \qquad (20)$$

The working inventories of the plant sites ($WI^{PS}_{k,j,t}$) and the distribution centers ($WI^{DC}_{m,j,t}$) are related linearly to the inlet flows of materials that they handle (Tsiakis et al, 2001). This is expressed via the constraints,

$$WI^{PS}_{k,j,t} \geq \alpha_{k,j,t} \sum_{ls \in LS} P_{k,j,ls,t} , \qquad \forall k \in K, j \in JM_k, t \in T \qquad (21)$$

$$WI^{DC}_{m,j,t} \geq \beta_{m,j,t} \sum_{k \in K} F_{k,j,m,t} , \qquad \forall m \in M, j \in JP_k, t \in T \qquad (22)$$

where $\alpha_{k,j,t}$ and $\beta_{m,j,t}$ are given coefficients coming from empirical studies (such as the probability of machines broken down or supply limitation).

5.2.4. Upper Bound Constraints

Purchases $P_{k,j,ls,t}$ from supplier ls to plant site k take place only if the transportation link between them is set up,

$$P_{k,j,ls,t} \leq P^U_{k,j,ls,t} Y^I_{k,ls} , \qquad \forall k, j, ls, t \qquad (23.1)$$

Inter-site transportation $TR_{k,k',j,t}$ from site k to site k' take place only if the transportation link between them is set up,

$$TR_{k,k',j,t} \leq TR^U_{k,k',j,t} Y^N_{k,k'} , \qquad \forall k, k', j, t \qquad (23.2)$$

Sales $S_{j,m,ld,t}$ from distribution center m to customer ld take place only if the transportation link between them is selected,

$$S_{j,m,ld,t} \leq S^U_{j,m,ld,t} Y^S_{m,ld} , \qquad \forall j, m, ld, t \qquad (23.3)$$

Nonzero production flows $W_{k,i,j,s,t}$ are allowed in plant i of site k only if the plant is installed,

$$W_{k,i,j,s,t} \leq W^U_{k,i,j,s,t} Y^P_{k,i} , \qquad \forall k, i, j, s, t \qquad (23.4)$$

$$Q_{k,i} \leq Q^U_{k,i} Y^P_{k,i} , \qquad \forall k, i \qquad (23.5)$$

The transportation amount $F_{k,j,m,t}$ from plant site k to distribution center m takes place only if the transportation link between them is set up,

$$F_{k,j,m,t} \leq F^U_{k,j,m,t} Y^O_{k,m} , \qquad \forall k, j, m, t \qquad (23.6)$$

5.3. Cyclic Scheduling Constraints

To address detailed operations of the multi-product plants, we have considered a cyclic scheduling policy (Pinto and Grossmann, 1994). Under this policy, the sequences to produce each product are decided together with the cycle time, and then the identical schedule is repeated over each time period (Figure 12). The trade-offs between inventories and transitions are established by optimizing the cycle times.

Important decisions in cyclic scheduling including the sequence of production ($SY_{k,i,s,sl,t}$) and precedence relationship for changeovers between pairs of products ($Z_{k,i,s,s',sl,t}$), are determined through the following assignment and sequencing constraints.

5.3.1. Assignment Constraints

The assignment constraints state that exactly one time slot must be assigned to one product and vice versa. The total number of time slots will be exactly equal to the total number of products,

$$\sum_{sl \in SL_i} SY_{k,i,s,sl,t} = 1, \qquad \forall k \in K_i, i \in I_j, s \in S_i, t \in T \tag{24}$$

$$\sum_{s \in S_i} SY_{k,i,s,sl,t} = 1, \qquad \forall k \in K_i, i \in I_j, sl \in SL_i, t \in T \tag{25}$$

5.3.2. Sequence Constraints

The sequence constraints state that exactly one transition from product s occurs in the beginning of any time slot if and only if s was being processing during the previous time slot. On the other hand, exactly one transition to product s occurs in the time slot if and only if product s is being processed during that time slot. As suggested in Wolsey (1997), the transition variables $Z_{k,i,s,s',sl,t}$ can be replaced by continuous variables between 0 and 1, instead of binary variables. This significantly reduces the number of discrete variables and improves the computational efficiency.

$$\sum_{s \in S_i} Z_{k,i,s,s',sl,t} = SY_{k,i,s',sl-1,t}, \qquad \forall k \in K_i, i \in I_j, s' \in S_i, sl \in SL_i, t \in T \tag{26}$$

$$\sum_{s' \in S'_i} Z_{k,i,s,s',sl,t} = SY_{k,i,s,sl,t}, \qquad \forall k \in K_i, i \in I_j, s \in S_i, sl \in SL_i, t \in T \tag{27}$$

$$0 \le Z_{k,i,s,s',sl,t} \le 1, \qquad \forall k, i, s, s', sl, t \tag{28}$$

5.3.3. Production Constraints

The production amount of product s in a cycle ($W^S_{k,i,s,t}$) is equal to the processing rate $r_{k,i,s}$ times the processing time $\delta_{k,i,s,sl,t}$:

$$W^S_{k,i,s,t} = \sum_{sl \in SL_i} r_{k,i,s} \delta_{k,i,s,sl,t}, \qquad \forall k \in K_i, i \in I_j, s \in S_i, t \in T \tag{29}$$

The amount produced for each product in time period t ($N_{k,i,t}$ cycles in the time period) should be no less than the total production predicted from operational planning in this time period:

$$W^{S}_{k,i,s,t} N_{k,i,t} = W_{k,i,j,s,t}, \qquad \forall k \in K_i, i \in O_j, j \in JP_{k,i}, s \in S_i, t \in T \tag{30}$$

5.3.4. Timing Constraints

Constraints (31) to (34) are used to restrict the timing issues in the cyclic scheduling.

The processing time $\delta_{k,i,s,sl,t}$ in a certain time slot is equal to the summation of the processing times assigned to all the products in this time slot,

$$\delta_{k,i,sl,t} = \sum_{s \in S_i} \delta_{k,i,s,sl,t}, \qquad \forall k \in K_i, i \in I_j, sl \in SL_i, t \in T \tag{31}$$

The cycle time $TC_{k,i,t}$ is equal to the summation of all the processing times in each time slot plus the summation of transition times in this cycle,

$$TC_{k,i,t} = \sum_{sl \in SL_i} \delta_{k,i,sl,t} + \sum_{s \in S_i} \sum_{s' \in S_i} \sum_{sl \in SL_i} Z_{k,i,s,s',sl+1,t} \tau_{i,s,s'}, \qquad \forall k \in K_i, i \in I_j, t \in T \tag{32}$$

The total production time should not exceed the duration of each time period $H_{k,i,t}$,

$$TC_{k,i,t} N_{k,i,t} \le H_{k,i,t}, \qquad \forall k \in K_i, i \in I_j, t \in T \tag{33}$$

The production for scheme s in time slot sl can take place only if the time slot is assigned to the production scheme:

$$\delta_{k,i,s,sl,t} \le \delta^{U}_{k,i,s,sl,t} SY_{k,i,s,sl,t}, \qquad \forall k \in K_i, i \in I_j, s \in S_i, sl \in SL_i, t \in T \tag{34}$$

5.3.5. Cost Constraints

To integrate the cyclic scheduling with the strategic planning, the inventory and transition costs from cyclic scheduling are considered as part of the operating cost. Constraint (35) represents that cost from scheduling in a time period for a certain plant. The first term on the right hand side of the equation stands for the total transition cost in a time period. The second term is the inventory cost for all the chemicals involved in the production. The change of inventory level in a time period is given in Figure 13. In the work by Pinto and Grossmann (1994), these authors consider inventory only for final products, as their model is for single plant. In our case, each manufacturing site may have more than one production plant, and inventory for materials of multi-product plants must be also taken into account. Since we assume that material balances are expressed linearly in terms of the main product's production, the cumulative inventory levels for raw materials are also related linearly to the cumulative inventory level of main product in each production scheme and the coefficients of the linear relationships are equal to the mass balance coefficients. This leads to the second term on the right hand side of the following constraint. Thus, the inventory and transition costs of multiproduct processes are given by,

$$COST^S_{k,i,t} = \sum_{s\in S_i}\sum_{s'\in S_i}\sum_{sl\in SL_i} CTR_{i,s,s'} Z_{k,i,s,s',sl,t} N_{k,i,t} + \sum_{s\in S_i}\sum_{j\in J_i}\sum_{sl\in SL_i} (\mu_{i,j,s}\varepsilon_{k,j,t})(r_{k,i,s} H_{k,i,t} - W^S_{k,i,s,t} N_{k,i,t})\delta_{k,i,sl,t} / 2$$
$$\forall k \in K_i, i \in I_j, t \in T \qquad (35)$$

This constraint is nonlinear and nonconvex, with bilinear and trilinear terms. If all the processes in the production network are dedicated, cyclic scheduling need not be taken into account, and this constraint can be discarded.

5.3.6. Upper Bound Constraints

As a multi-site problem, we need to make sure that if a plant i in site k is not installed, there are no production cycles. To model this, we introduce the upper bound constraint (36) for the number of cycles $N_{k,i,t}$ in each time period for each multiproduct plant in each manufacturing site:

$$N_{k,i,t} \le N^U_{k,i,t} Y^P_{k,i}, \qquad \forall k \in K_i, i \in I_j, t \in T \qquad (36)$$

Also the assignment constraints are reformulated to account for the fact that all the scheduling activities can take place only if the plant is installed:

$$\sum_{sl\in SL_i} SY_{k,i,s,sl,t} = Y^P_{k,i}, \qquad \forall k \in K_i, i \in I_j, s \in S_i, t \in T \qquad (24)$$

$$\sum_{s\in S_i} SY_{k,i,s,sl,t} = Y^P_{k,i}, \qquad \forall k \in K_i, i \in I_j, sl \in SL_i, t \in T \qquad (25)$$

5.4. Probabilistic Constraints for Demand Uncertainty

A key component of decision making under uncertainty is the representation of the stochastic parameters. There are two distinct ways of representing uncertainties. The scenario-based approach (Subrahmanyam et al, 1994; Liu and Sahinidis, 1996) attempts to capture the uncertainties by representing them in terms of a number of discrete realizations of the stochastic parameters where each complete realization of all uncertain parameters gives rise to a scenario. In this way all the possible future outcomes are taken into account through the use of scenarios. This approach provides a straightforward way to formulate the problem, but its major drawback is that the problem size increases exponentially as the number of scenarios increases. This is particularly true when using continuous multivariate probability distribution with Gaussian quadrature integration schemes.

Alternatively, Monte Carlo sampling could be used, although it also requires a rather large number of points to achieve a desired level of accuracy. These difficulties can be somewhat circumvented by analytically integrating continuous probability distribution functions for the random parameters (Wellons et al, 1989; Petkov et al, 1998). This approach can lead to a reasonable size deterministic equivalent representation of the probabilistic model, at the expense of introducing certain amount of nonlinearities into the model through multivariate integration over the continuous probability space. In this work, this approach is used for describing the demand uncertainty.

The probabilistic description of demand uncertainty renders some operational planning variables (amount of sale, shortfall and salvage) to be stochastic. A traditional way to deal with this is to allow “corrective actions” by adding recourse actions in the model. Stochastic programming problems with recourse are usually complicated in its nature and difficult to solve. Based on the special characteristics of demand uncertainty, Shen et al (2003) recently proposed a novel approach to hedge unexpected demand by holding a certain amount of safety stock before demand realization. The amount of safety stock is estimated by using a chance constraint linking service level to a demand probability distribution. The need for recourse is obviated by taking the proactive action with safety stock, and the stochastic attributes of the problem are translated into a deterministic optimization problem at the expense of introducing nonlinear terms into the model. In this work, we use a similar approach. Instead of specifying the service level (or over-stocked probability), we treat the stock-out probability as a variable, and integrate it with supply chain responsiveness. The detailed model formulations are given below.

5.4.1. Stock-out Probability

If a particular product demand realization $d_{j,ld,t}$ is higher than its total available amount $QS_{j,ld,t}$, we are under-stocked, i.e. stock-out will happen. If a particular product demand realization $d_{j,ld,t}$ is less than its total available amount $QS_{j,ld,t}$, we are over-stocked. The probability of over-stock is defined as β-service level in manufacturing literatures. The probability of stock-out plus the service level should be always equal to 1 (as Figure 14). Thus the stock-out probability can be expressed as,

$$Prob_{j,ld,t} = \Pr(QS_{j,ld,t} \le d_{j,ld,t}) = 1 - \Pr(QS_{j,ld,t} \ge d_{j,ld,t}), \qquad \forall j \in J, ld \in LD, t \in T \quad (37)$$

The form of this constraint is very similar to a chance constraint (Charnes and Cooper, 1963), which suggests that the above equation can be transformed into a deterministic form. It is worth mentioning that this general definition for estimating safety stock level can be applied for all types of demand distributions. In real world cases, demand is often assumed to be normally distributed when there are sufficient samples, or triangularly distributed when limited sample applied. In the following sections we consider the cases when the demand follows a normal distribution and a triangular distribution.

5.4.2. Normal Distribution

Due to the central limit theorem, the product demands are often modeled as random parameters that are normally distributed (Wellons et al, 1989; Petkov et al, 1998). For a normal distributed demand, we are given the mean ($\mu_{j,ld,t}$) and standard deviation ($\sigma_{j,ld,t}$) of the demand.

To facilitate the calculation of the stock-out probability $Prob_{j,ld,t}$, standardization of the variables is first performed. Normal random variables can be recast into the standardized normal form, with a mean of zero and a variance of 1, by subtracting their mean and dividing by their standard deviation. For the deterministic variables $QS_{j,ld,t}$, the standardized normal variables $K_{j,ld,t}$ are given as:

$$K_{j,ld,t} = \frac{QS_{j,ld,t} - \mu_{j,ld,t}}{\sigma_{j,ld,t}} \tag{38}$$

Substituting (38) into the general definition of stock-out probability defined in (37), we have the stock-out probability of normally distributed demand for product j customer ld at time period t is,

$$Prob_{j,ld,t} = 1 - \Phi(K_{j,ld,t}), \qquad \forall j \in J, ld \in LD, t \in T \tag{39.a}$$

where Φ(x) denotes the cumulative probability function of standard normal random variable in the form of,

$$\Phi(\text{x}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{x^2}{2}) \,\mathrm{d}\, x \tag{39.b}$$

The stock-out probability for the entire planning horizon is considered as the worst case for all the individual stock-out probabilities of all the time periods. Hence we have that the stock-out probability ($Prob_{j,ld}$) for product j and customer ld is:

$$Prob_{j,ld} \geq Prob_{j,ld,t} = 1 - \Phi(K_{j,ld,t}), \qquad \forall j \in J, ld \in LD, t \in T \tag{39}$$

5.4.3. Triangular Distribution

For triangular distribution, we are given the demand's lower bound d^L, upper bound d^U, and the most likely value d^M (as in Figure 15). Due to the non-differentiable characteristics of triangular distribution, the stock-out probability has different representations when the total available amount QS is less or more than the most likely demand d^M. Their relationship can be represented by the following disjunction (for simplicity, the subscripts j, ld, t are omitted)

$$\begin{bmatrix} y \\ d^L \le QS \le d^M \\ Prob = 1 - \dfrac{(QS-d^L)^2}{(d^U-d^L)(d^M-d^L)} \end{bmatrix} \vee \begin{bmatrix} \neg y \\ d^U \ge QS \ge d^M \\ Prob = \dfrac{(d^U-QS)^2}{(d^U-d^L)(d^U-d^M)} \end{bmatrix} \quad (40)$$

Using the convex hull reformulation, the disjunction can be transformed into MINLP constraints as discussed by Lee and Grossmann (2000):

$$QS_{j,ld,t} = QS^1_{j,ld,t} + QS^2_{j,ld,t} \quad (40.1)$$

$$d^L_{j,ld,t} Y^d_{j,ld,t} \le QS^1_{j,ld,t} \le d^M_{j,ld,t} Y^d_{j,ld,t} \quad (40.2)$$

$$d^M_{j,ld,t}(1-Y^d_{j,ld,t}) \le QS^2_{j,ld,t} \le d^U_{j,ld,t}(1-Y^d_{j,ld,t}) \quad (40.3)$$

$$Prob_{j,ld,t} = Prob^1_{j,ld,t} + Prob^2_{j,ld,t} \quad (40.4)$$

$$Prob^1_{j,ld,t} \le Y^d_{j,ld,t} \quad (40.5)$$

$$Prob^2_{j,ld,t} \le 1 - Y^d_{j,ld,t} \quad (40.6)$$

$$Prob^1_{j,ld,t} \ge Y^d_{j,ld,t} - \frac{(QS^1_{j,ld,t} - d^L_{j,ld,t})^2}{(d^U_{j,ld,t} - d^L_{j,ld,t})(d^M_{j,ld,t} - d^L_{j,ld,t})} \quad (40.7)$$

$$Prob^2_{j,ld,t} \ge \frac{(ydu_{j,ld,t} - QS^2_{j,ld,t})^2}{(d^U_{j,ld,t} - d^L_{j,ld,t})(d^U_{j,ld,t} - d^M_{j,ld,t})} \quad (40.8)$$

where $Y^d_{j,ld,t}$ is a binary variable that equal to 1 if $QS_{j,ld,t}$ is less than $d^M_{j,ld,t}$.

5.5 Lead Time Definition

As discussed in Section 3, for each path the expected lead time is equal to the delivery lead time, plus the production lead time, times the stock out probability. The delivery lead time and the production lead time are in turn equal to the summation of all the production delays and transportation delays incurred in the corresponding path. The expected lead time of the whole supply chain network is equal to the maximum expected lead time of each path. As a supply chain design problem, we need to consider the case that if a plant or a transportation link is not selected, the associated delay is 0. Binary variables are used to model the lead time TP with the following inequalities:

$$TP \geq Prob_{j,ld} Y^{I}_{k,ls} \lambda^{I}_{k,ls} + \sum_{x=1}^{n} Prob_{j,ld} Y^{P}_{k_x,i_x} \theta_{k_x,i_x} + \sum_{x=1}^{n-1} Prob_{j,ld} Y^{N}_{k_x,k_{x+1}} \lambda^{N}_{k_x,k_{x+1}} + Prob_{j,ld} Y^{O}_{k_n,m} \lambda^{O}_{k_n,m} + Y^{S}_{m,ld} \lambda^{S}_{m,ld}$$
$$\forall (ls, k_1, k_2 ... k_n, m, ld) \in Path_{ls,k,m,ld} \qquad (41)$$

where $Prob_{j,ld}$ is the stock-out probability, all the Y are binary variables for design decisions, λ denotes transportation delays and θ represents production delay. The superscript $(*)^I$ denotes the transport link from supplier to plant site, the superscript $(*)^P$ denotes the plant, $(*)^N$ is for inter-site transport link, $(*)^O$ represents the transport link from plant site to distribution center, and $(*)^S$ is for the transport link from distribution center to customer.

The set $Path_{ls,k,m,ld}$ includes all the possible paths in a given potential PSCN network. All the elements in the set $Path_{ls,k,m,ld}$ are in the form of $(ls, k_1, k_2 ... k_n, m, ld)$, where supplier ls is the start of the path, $k_1, k_2 ... k_n$ are the manufacturing sites and m is the distribution center that the associated stream goes through, and customer ld is at the end of the path.

In equation (41) the transportation delay of each route and the production delay of each single product plant are given parameters. The production time delay for a multiproduct plant is not so obvious. Before introducing our definitions, consider the motivating example shown in Figure 7, 12 and 13. A multi-product plant produces three chemicals *A*, *B* and *C*. Assume there is a demand change of chemical *A*. The worst case is when we just finished producing *A*, and there is no extra inventory of *A* besides the one committed to the former demand. There are two operating policies that can be implemented to deal with this situation.

If the demand of chemical *A* has a large change, one would usually stop the current production for chemical *C* as soon as possible and skip all the other products (Chemical *B*) to produce chemical *A* directly. In this case, the production delay is equal to the residence time of chemical *A* (Figure 16). So we have the production delay $\theta_{k,i}$ for multiproduct plant i in site k is equal to the maximum residence time ($\theta^{R}_{i,s}$) of all the products produced by this plant,

$$\theta_{k,i} \geq \theta^{R}_{i,s}, \qquad \forall k \in K_i, i \in I_j \qquad (42.a)$$

If the demand change of chemical A is not very significant, one will wait until the plant produces A again, so that we can adjust the production to meet the demand change. This takes some time which is given by the processing time of chemical B and C, plus residence time of A. In this way we define for multiproduct plant, the time delay for each product as cycle time plus residence time minus its processing time (Figure 17). Similarly, the production delay for a multiproduct plant is equal to the maximum time delay for each product:

$$\theta_{k,i} \geq TC_{k,i,t} + \theta^{R}_{i,s} - \sum_{sl \in SL_i} \delta_{k,i,s,sl,t}, \qquad \forall k \in K_i, i \in I_j \qquad (42.b)$$

In this definition cycle times of each plant are taken into account as part of the delay due to production, so that we have integrated the production details into the quantitative definition of responsiveness.

The terms $Prob_{j,ld}Y$ (stock-out probability times binary variable omitting the subscripts for simplicity) in the lead time definition can be linearized. We use a continuous variable PY to replace the $Prob_{j,ld}Y$ term in the lead time constraint:

$$Prob_{j,ld}Y = PY, \qquad \forall j, ld \qquad (43.1)$$

$$TP \geq PY^{I}_{k,j,ls,ld}\lambda^{I}_{k,ls} + \sum_{x=1}^{n} PY^{P}_{k_x,i_x,j,ld}\theta_{k_x,i_x} + \sum_{x=1}^{n-1} PY^{N}_{k_x,k_{x+1}j,ld}\lambda^{N}_{k_x,k_{x+1}} + PY^{O}_{k_n,j,m,ld}\lambda^{O}_{k_n,m} + Y^{S}_{m,ld}\lambda^{S}_{m,ld}$$

$$\forall (ls, k_1, k_2 ... k_n, m, ld) \in Path_{ls,k,m,ld} \qquad (43)$$

The equation (43.1) is equivalent to the following disjunction,

$$\begin{bmatrix} Y \\ PY = Prob_{j,ld} \end{bmatrix} \vee \begin{bmatrix} \neg Y \\ PY = 0 \end{bmatrix}, \qquad \forall j, ld \qquad (44.1)$$

Applying the convex hull reformulation (Balas, 1985) to the above disjunctive constraint leads to:

$$PY + PY2 = Prob_{j,ld} \qquad \forall j, ld \qquad (44.2)$$

$$PY \leq Y \qquad \forall j, ld \qquad (44.3)$$

$$PY2 \leq 1 - Y \qquad \forall j, ld \qquad (44.4)$$

where $PY2$ is a new continuous variable introduced as a slack variable. The constraints (33) are applied for all the terms with superscript $(*)^{I}$, $(*)^{P}$, $(*)^{N}$, $(*)^{O}$, $(*)^{S}$ in the expected lead time definition.

5.6. Nonnegative Constraints

All continuous variables must be nonnegative and the binary variables should be integer:

$$Q_{k,i}, W_{k,i,j,t}, P_{k,j,ls,t}, TR_{k,k',j,t}, F_{k,j,m,t}, S_{j,m,ld,t}, I_{j,m,ld,t}, QS_{j,ld,t}, TP, \theta \geq 0 \quad (45.1)$$

$$Z_{k,i,s,s',sl,t}, W^S_{k,i,s,t}, r_{k,i,s}, \delta_{k,i,sl,t}, TC_{k,i,t}, Te_{k,i,sl,t}, Ts_{k,i,sl,t}, N_{k,i,t}, \theta_{k,i}, COST^S_{k,i,t} \geq 0 \quad (45.2)$$

$$Y^I_{k,ls}, Y^P_{k,i}, Y^O_{k,m}, Y^m, Y^S_{m,ld}, SY_{k,i,s,sl,t} \in \{0,1\} \quad (45.3)$$

5.7. Net Present Value

The NPV of the supply chain network is given by the following equations,

$$NPV = Income - C_{purch} - C_{oper} - C_{tranp} - C_{invest} - C_{inventory} \quad (46)$$

$$Income = \sum_j \sum_{ld} \sum_t \varphi_{j,ld,t} Sa_{j,ld,t}$$

$$C_{purch} = \sum_k \sum_j \sum_{ls} \sum_t \varphi_{j,ls,t} P_{k,j,ls,t}$$

$$C_{operate} = \sum_k \sum_i \sum_s \sum_{j \in JP_{i,s}} \sum_t \sigma_{i,s,t} W_{k,i,j,s,t} + \sum_k \sum_i \sum_t COST^S_{k,i,t}$$

$$C_{tranp} = \sum_k \sum_j \sum_{ls} \sum_t \omega^I_{k,j,ls,t} P_{k,j,ls,t} + \sum_k \sum_{k'} \sum_j \sum_t \omega^N_{k,k',j,t} TR_{k,k',j,t} + \sum_k \sum_j \sum_m \sum_t \omega^O_{k,j,m,t} F_{k,j,m,t} + \sum_j \sum_m \sum_{ld} \sum_t \omega^S_{j,m,ld,t} S_{j,m,ld,t}$$

$$C_{invest} = \sum_k \sum_i \omega^P_{k,i} Q_{k,i} + \sum_k \sum_i \gamma^P_{k,i} Y^P_{k,i} + \sum_k \sum_{ls} \gamma^I_{k,ls} Y^I_{k,ls} + \sum_k \sum_m \gamma^O_{k,m} Y^O_{k,m} + \sum_k \sum_{k'} \gamma^N_{k,k'} Y^N_{k,k'} + \sum_m \gamma^m Y^m + \sum_m \sum_{ld} \gamma^S_{k,ld} Y^S_{k,ld}$$

$$C_{inventory} = \sum_j \sum_m \sum_{ld} \sum_t \varepsilon_{j,m,t} I_{j,m,ld,t} + \sum_k \sum_j \sum_t \varepsilon_{k,j,t} (WI^{PS}_{k,j,t} + WI^{DC}_{m,j,t})$$

All the parameters in the above formulation are discounted at a specified interest rate and include the effect of taxes and interest rate on the net present value.

6. SOLUTION PROCEDURE

6.1. Solution Procedure for Multi-objective Optimization

In order to obtain the Pareto-optimal curve for the bi-criterion optimization problem[2], one of the objectives is specified as an inequality with a fixed value for the bound which is treated as a parameter. There are two major approaches to solve the problem in terms of

[2] Two objectives are given by (43) and (46), constraints are given by (1)-(42), (44)-(45)

this parameter. One is to simply solve it for a specified number of points to obtain an approximation of the Pareto optimal curve. The other is to solve it as a parametric programming problem (Dua and Pistikopoulos, 2004), which yields the exact solution for the Pareto optimal curve. While the latter provides a rigorous solution approach, the former is simpler to implement for nonlinear models. For this reason we have selected this approach. The procedure includes the following three steps. The first one is to minimize the expected lead time TP to obtain the shortest expected lead time TP_S, which in turn yields the lowest Pareto optimal NPV. The second step is to maximize NPV that in turn yields the longest Pareto optimal expected lead time TPL. In this case the objective function is set as,

$$NPV - \varepsilon \cdot TP \tag{47}$$

where ε is a very small value (e.g., it is on the order of 0.001). The last step is to fix the expected lead time TP to discrete values between TP_S and TP_L, and optimize the model by maximizing NPV at each selected point. In this way we can obtain an approximation to the Pareto-optimal curve, together with the optimal configurations of PSCN for different values of lead time.

6.2. Shortest Optimal Expected Lead Time

In the aforementioned solution procedure for this bi-criterion optimization problem, one of the important steps is to find the shortest optimal expected lead time. Instead of minimizing the expected lead time by solving the entire problem directly, we use the following solution strategy to improve the computational efficiency.

The expected lead time of each path of chemical flow as defined in section 3.2 is given by $E(L) = Prob \cdot L_P + L_D$, which equals to the delivery lead time (L_D) plus stock-out probability ($Prob$) times production lead time (L_P). From the above equation, we can see that as the stock-out probability decreases, the expected lead time will decrease. In the step for determining the shortest optimal expected lead time, we do not account for the economic objective. Therefore, if there are sufficiently high safety stock levels in all the distribution centers to hedge the uncertain demands, the stock-out probability will be 0. Then the expected lead time of the supply chain network will be reduced to the maximum delivery lead time of each path of chemical flow. The delivery lead time is equal to the transportation delay from a distribution center to a customer. Because we can select what distribution centers to install and what transportation links to set up, we are able to choose a subset of all the possible transport links between the customers and distribution centers, such that each customer has at least one transport link connected to, and the maximum transportation delay of all these transport links is minimized. In summary, the minimum expected lead time can be calculated by:

$$TP_S = \max_{ld \in LD}\{\min_{m \in M}\{\lambda_{m,ld}^S\}\} \tag{48}$$

where $\lambda_{m,ld}^S$ is the shipping time from distribution center m to customer ld.

Equation (48) defines the shortest optimal expected lead time. It can be interpreted as follows. For each customer ld, we only choose one distribution center that have minimum shipping time to this customer, and set up the transport link between them. In this way, we set up as many transport links as the number of the customers (number of elements in set ld). Since the expected lead time of the supply chain network is equal to the longest expected lead time for each path, the minimum expected lead time of the supply chain network will be equal to the maximum shipping time for those transport links that have been set up.

Instead of solving the large scale nonconvex MINLP problem, we can obtain the optimal solution with much less computational effort by using the proposed method. However, this method works only for the shortest expected lead time case. To obtain other points in the Pareto curve, we need to solve a series of large scale nonconvex MINLP problems. To reduce the computational time, a heuristic hierarchical solution approached is proposed.

6.3. Heuristic Hierarchical Approach

The solution of each problem for the points in the Pareto curve can be very computationally demanding tasks due to the large number of discrete decisions and the highly nonlinear nonconvex terms3. In this section we propose a heuristic hierarchical solution approach that is able to handle the combinatorial and nonconvex nature of the responsive supply chain design problem and to reduce the computational effort needed. The basic idea is to exploit the fact that the operational cost arising form scheduling (inventory costs and transition costs) makes up only a small part of the total NPV. Therefore, we can use a two-stage decision approach as follows. We first determine the supply chain network structure and strategic planning decisions (production and shipping amounts), neglecting changeovers and transitions for multiproduct plants. In the second

[3] The nonlinear nonconvex terms arise from the cyclic scheduling constraints (29), (30), (35), stock-out probability definition (37)-(40) and expected lead time definition (41), (43).

stage, we then consider the detailed production scheduling after the strategic design and planning decisions are made. The proposed methodology employs a simplified model (first stage) to determine the strategic design and planning decisions, which are then fed into the detailed model (second stage) in order to derive the operational and scheduling decision variables.

It is important to note that this solution approach can be applied when one defines the production delay with constraint (42.a), which assumes that the demand is undergoing large changes and the production delay is irrelevant to the cycle time (please refer to section 5.5 for details). Clearly this heuristic approach may not yield optimal or near-optimal solutions if the model includes constraint (42.b) for the production delays. The reason is given in section 6.3.1.

6.3.1. Simplified Model

The simplified model formulation is an approximation of the detailed model formulation. The main advantage of the simplified model is that it does not include the large amount of the nonconvex terms and binary variables for cyclic scheduling. In this way, the problem can be solved more efficiently and still capture the main trends in the supply chain design and planning. Furthermore, the model provides a valid upper bound by overestimating the objective function of the original problem.

As transitions and changeovers for multiproduct plants are not taken into account in the simplified model, all the constraints for cyclic scheduling are not included. The objective functions are the same for each step as in the aforementioned solution procedure for multi-objective optimization problem. In summary, the simplified model includes the following constraints: (1)-(23), (37)-(40) (depends on the associated demand probability distribution), (42.a), (44)-(45). The objective functions are (43) and (46).

In the simplified model we define the production delay with constraint (42.a), which considers production delay equal to the maximum residence time of all the products. Therefore, the expected lead time of a supply chain depends on the design and planning decisions only, and unrelated to production scheduling. Thus, the expected lead time obtained from the simplified model is exactly the same as what we can get by solving the original model. Based on this, in the second step of ε-constraint method, which is to calculate the longest optimal expected lead time, we only need to solve the simplified model for the solution.

On the other hand, the simplified model might not provide a good approximation to the original model, if the model includes constraint (42.b) to define the production delays instead of (42.a). Because constraint (42.b) considers production delay equal to the cycle time minus processing time plus residence time. However, cycle times and processing times are decisions for cyclic scheduling, which have been neglected in the simplified model. Missing information of scheduling may lead to inaccuracies when optimizing the expected lead time. Thus, this heuristic approach may not yield near-optimal solutions if the model includes constraint (42.b) for the production delays

6.3.2. Solving Detailed Model

In the detailed model, the design variables (binary variables for design decisions) and planning variables (production and transportation amounts, inventory levels) are fixed to their values as determined from the simplified model. The original detailed model is then solved in the reduced variable space in order to determine the optimal levels for the scheduling decision variables. However, the detailed model is still a large scale nonconvex MINLP model. To obtain a "good" starting point for the near-optimal solutions we formulate a heuristic subproblem and a convexified subproblem to select the initial values of the scheduling variables.

The heuristic subproblem is used for the production sequence of multiproduct plants. It is based on the fact that transition costs are usually proportional to the transition times. Thus, a production sequence that can minimize the total transition time in a production cycle is often the optimal production sequence (Sahinidis and Grossmann, 1991). Therefore, the heuristic subproblem is to minimize the total transition times in a production cycle (constraint (32)), subject to the constraints (24) – (28).

In the convexified subproblem, all the bilinear terms in the detailed model are replaced by convex envelopes (McCormick, 1976; Quesada and Grossmann, 1995). For example, the equation $W^S_{k,i,s,t} N_{k,i,t} = W_{k,i,j,s,t}$ (30) was replaced by the following constraints,

$$W^{S\ \ LO}_{k,i,s,t} N_{k,i,t} + W^S_{k,i,s,t} N_{k,i,t}{}^{LO} - W^{S\ \ LO}_{k,i,s,t} N_{k,i,t}{}^{LO} \le W_{k,i,j,s,t} \quad (49.1)$$

$$W^{S\ \ UP}_{k,i,s,t} N_{k,i,t} + W^S_{k,i,s,t} N_{k,i,t}{}^{UP} - W^{S\ \ UP}_{k,i,s,t} N_{k,i,t}{}^{UP} \le W_{k,i,j,s,t} \quad (49.2)$$

$$W^{S\ \ LO}_{k,i,s,t} N_{k,i,t} + W^S_{k,i,s,t} N_{k,i,t}{}^{UP} - W^{S\ \ LO}_{k,i,s,t} N_{k,i,t}{}^{UP} \ge W_{k,i,j,s,t} \quad (49.3)$$

$$W^{S\ \ UP}_{k,i,s,t} N_{k,i,t} + W^S_{k,i,s,t} N_{k,i,t}{}^{LO} - W^{S\ \ UP}_{k,i,s,t} N_{k,i,t}{}^{LO} \ge W_{k,i,j,s,t} \quad (49.4)$$

where LO and UP represent lower and upper bounds on the variables, respectively.

The solutions of the two subproblems are then used as the starting point to solve the detailed model. By doing this, we increase the chance to obtain a near-optimal solution.

6.3.3. Algorithm

In summary, the hierarchical algorithm comprises of the following steps:

Step 1: Solve the simplified nonconvex MINLP model (objective functions are (43) and (46), subject to constraints (1)-(23), (37)-(40), (42), (44)-(45)) by neglecting transitions and changeovers of multiproduct plants and then fix the design and planning decisions.

Step 2: For each multiproduct plant, solve the MILP heuristic subproblem (objective function is (32), constraints are (24)-(28)) to minimize the total transition time in a production cycle to obtain the initial values of the scheduling variables.

Step 3: Solve the convexified model that uses convex envelopes to replace the bilinear constraints (30), (33) and (35)

Step 4: Use the solutions from the subproblems in Step 2 and 3 as the starting points, solve the detailed MINLP model in the reduced variables space to obtain the scheduling and operational decisions.

7. NUMERICAL EXAMPLES

In order to illustrate the application of the proposed model and its corresponding solution strategy, we consider two examples for the design of polystyrene supply chains. The first example is a medium size problem that is solved with different production delay definitions under different demand uncertainties (triangular distribution and normal distribution), and using different solution approaches (direct approach and hierarchical approach). The second example is a large scale problem motivated by a real world application of which two instances with and without considering safety stocks are solved. The summary of all the instances is given in Table 1.

In both examples the time horizon is 10 years, and three time periods are considered with lengths of 2 years, 3 years and 5 years, respectively. An annual interest rate of 10% and a tax rate of 45% have been considered for the calculation of the net present value. All the other input data are available upon request.

All the instances are modeled with GAMS (Brooke et al, 1998) and solved on an IBM T60 laptop with an Intel Core Duo 1.83 GHz CPU and 1GB RAM. Due to the non-convexity of the MINLP problems, the global optimization solver BARON (Sahinidis, 1996) was used for the instances where the demands are triangular distributed, and the

SBB solver was used for the instances where the demands are normal distributed (because BARON does not support error function). All the instances are first solved with the outer-approximation algorithm (Duran and Grossmann, 1986) solver in DICOPT for obtaining a lower bound before solving with BARON or SBB solvers.

7.1. Example 1

This example has a production network with three types of candidate plants (Figure 18). Plant I is used to produce styrene monomers from ethylene and benzene; Plant II is a multiproduct plant for the production of three different types of solid polystyrene (SPS) resins; Plant III is also a multiproduct plant for the production of two different types of expandable polystyrene (EPS) resins. The entire supply chain network includes three potential suppliers, three potential production sites, two potential distribution centers, three customers and the associated potential transport links between them. The superstructure of the potential process supply chain network for Example 1 is given in Figure 19. Two raw materials (benzene and ethylene), one intermediate (styrene monomer) and seven products (three types of SPS resins and two types of EPS resins) are included in the supply chain network.

7.1.1. Solutions for Different Production Delay Definitions

We first consider two instances for Example 1 where the demand follows triangular distributions, but different production delay definitions are needed for multi-product plants. Instance 1 considers that the demand has large changes, and thus the production delay is irrelevant to the cycle time (Constraint 42.a); Instance 2 considers that the demand changes are small and that the production delay is closely related to cycle time (Constraint 42.b). Both instances consist of 126 binary variables, 2,970 continuous variables and 3,438 constraints and they are solved directly by using GAMS/BARON with 0% optimality tolerance. Six points in the Pareto optimal curve require 4,063 CPU seconds for Instance 1 and 16,893 seconds for Instance2.

The Pareto curve is shown in Figure 20. We can see that for Instance 1 the Pareto curve ranges from 1.5 to 4.8 days in the expected lead time, and from $532 million to $667 million for the NPV. For Instance 2, the Pareto curve ranges from 1.5 to 9.32 days in the expected lead time, and from $491 million to $667 million for the NPV.

For both curves the NPV increases as the expected lead time increases. This means that the price to reduce the expected lead time and to improve the supply chain responsiveness is to decrease the NPV. From the trends of both curves, we can see that the rate of increase of NPV decreases as the expected lead time increases. This means that the cost to reduce the expected lead time increases when the expected lead time becomes smaller.

Although these two instances use different production delay definitions, they still have the same maximum NPV, because different definitions of production delay have only an impact on the expected lead times and the NPV for intermediate points.

Because the lead time definition (42.b) not only takes into account the residence time, but also considers the cycle times and processing times, the production delay we can obtain from constraint (42.b) will always be greater than the one we can get from (42.a). Thus, the longest optimal expected lead time for Instance 2 is greater than the one for Instance 1.

It is reasonable that the two instances have the same minimum expected lead time, because as long as there are sufficient safety stocks in the distribution centers, the expected lead time will be equal to the expected delivery lead time regardless of the production delay. However, for the minimum expected lead time the optimal NPV from Instance 2 is less than the NPV form Instance 1. The reason is that we define the expected lead time of the supply chain as the longest expected lead time for all the possible paths. Although these two cases have the same expected lead time for the entire supply chain, the lead time for each individual path is not the same. As Instance 2 always has a longer production lead time than Instance 1 under the same circumstances, more inventories need to be held to reduce the stock out probability so that the longer production lead time of the paths except the longest one could be traded off.

Figure 21 shows the change of cumulative inventories for all the EPS resins in distribution center "DC1" for different expected lead time specifications for the two production delay definitions. The safety stock levels are both zero at the longest expected lead time case. As the expected lead time decreases, the inventory level increases. More inventories are required for Instance 2 to obtain the same expected lead time as Instance 1. This is also because the production delay definition allows Instance 1 to have shorter production lead time.

The optimal network structure (Figure 22) is the same for all the points in the Pareto curves in both instances. All the three manufacturing sites are selected, but only site 2 installs plant I to produce styrene monomers, and then the monomers are shipped to plant sites II and III as the raw materials for different polystyrene resins. Plant site 1 connects to both distribution centers, but plant site 2 only connects to the first distribution center. And each customer is served by only one distribution center.

Instance 2 requires much longer computational time than Instance 1 due to the large number of bilinear terms in the lead time definition (because the production delays are variables for instance 2). For computational simplicity, we will use Constraint 42.a as the production delay definition in the following examples, i.e. production delay irrelevant to cycle time.

7.1.2. Normal Distributed Demand

For Instance 3, we solve Example 1 with demand following a normal distribution. This instance consists of 105 binary variables, 2,466 continuous variables and 2,963 constraints, and is solved with GAMS/SBB and 5% margin of optimality. The reason we used the SBB solver instead of BARON is that the later one cannot handle the nonlinear term (error function) arising from the cumulative probability distribution of normal distribution (Constraint 39). Six points in the Pareto optimal curve require 753.57 CPU seconds.

The Pareto curve is shown in Figure 23. It has the same range as Instance 1 in the expected lead time, but a wider range of NPV from $190 million to $591 million. This curve shows a similar trend as the Pareto curve for Instance 1, but the first point (the optimal NPV for the shortest lead time case) is much lower than expected. A possible reason is that the normal distribution has a “bell” shape curve. To ensure that the stock–out probability is zero for the longest path, a sufficiently large amount of safety stock should be held due to the long tail of the normal distribution, which leads to significant inventory cost.

Figure 24 represents the safety stock level for three types of SPS and two types of EPS in the two distribution centers. Similar to the case of the triangular distribution, the inventory level goes down as the expected lead time increases.

The optimal network structure (Figure 22) is the same as we obtained from Instance 1.

7.1.3. Hierarchical Solution Approach

To test the performance of our proposed solution algorithm, we solve Instance 1 with the proposed hierarchical solution approach. The solver we used is GAMS/BARON, and the optimality margin is set to be 0% for the simplified model and 5% for the detailed model.

The computational results are showed in Table 2. We can see that for all the points, the computational times are smaller than the times required for the direct approach. Since Instance 1 is solved with a 0% optimality tolerance, the results are globally optimized. After comparing with the results from hierarchical approach, the differences for the solutions of these two instances are small (around 3%).

As we can see from this instance, the proposed hierarchical algorithm can obtain a good near-optimal solution in short computational times. Another important advantage of using the hierarchical algorithm is that we can obtain near optimal solutions in reasonable time for very large scale problems that are unable to be solved directly with GAMS/BARON. In the next example, we are going to solve an industrial size problem that GAMS/BARON failed to converge for more than two weeks if solving it directly. However, with our hierarchical approach, solutions for six points of the Pareto curves are found within 5 hours. The details will be discussed next.

7.2. Example 2

Example 2 is a large scale problem which is motivated by a real world application. It has the same production network as in Figure 18. The potential supply chain network (Figure 25) includes three possible ethylene suppliers located in Illinois, Texas and Mississippi, and three potential benzene suppliers located in Texas, Louisiana, Alabama. Four potential manufacturing sites can be located in Michigan, Texas, California and Louisiana. The Michigan manufacturing site can set up all the three types of plants, the Texas manufacturing site can only install Plant I, the California manufacturing site can only set up Plants II and III, and the Louisiana manufacturing site can only set up Plants I and II. The supply chain can have five distribution centers, located in Nevada, Texas, Georgia, Pennsylvania and Iowa. Customers are pooled into nine sale regions across the country based on their geographical proximity. The corresponding superstructure of the supply chain network is given in Figure 26.

We assume that the demands follow triangular distributions, and solve this problem with consideration of safety stock (Example2, Instance 1) and without holding any safety stock (Example2, Instance 2). Both instances consist of 215 binary variables, 8,216 continuous variables and 14,617 constraints and they are solved with the proposed hierarchical algorithm by using GAMS/BARON with 0% optimality tolerance for the simplified model and 5% for the detailed model. Six points in the Pareto optimal curve require 15,396 CPU seconds for Instance 1 and 16,927 seconds for Instance 2. The reason it takes longer to solve Instance 2 is because for the step for calculating the shortest expected lead time we need to solve the problem by minimizing the expected lead time instead of using Equation 48 directly, because safety stocks are not considered in Instance 2, and thus Equation 48 is not applicable for this case.

The Pareto optimal curves are given in Figure 27. The Pareto curve for Example2, Instance 1 (with safety stock) has a similar trend as the Instance 1 of Example 1, and ranges from 1.6 to 5 days in the expected lead time and from $409 million to $683 million for NPV. The Pareto curve of Example 2, Instance 2 (without safety stock) ranges from 4.3 to 5 days and is very similar to the curve obtained for the deterministic supply chain design reported in You and Grossmann (2007).

Although with different ranges of NPV, these two curves has the same optimal longest expected lead time and the associated NPV (due to the optimality margin, there is a small difference between their optimal NPV). This is because in the longest expected lead time case, the supply chain needs to reduce cost by setting the safety stock levels to zero for Instance 1, which is equivalent to the case of no safety stock in Instance 2. Since there is no inventory held in Instance 2, the only factor that can change the expected lead time is the network structure. No matter how the network structure changes, the expected lead time in Instance 2 always includes the production lead time. Due to this reason, the range of expected lead time for Instance 2 is much smaller than that of Instance 1. On the other hand, changing the supply chain network structure is always much more expensive than holding a certain amount of safety stock. Thus, the Pareto curve for Instance 2 is below the curve for Instance 1.

The optimal network structures under different expected lead times for Instance 1 are shown in Figure 28-30. It is interesting to see that all the four sites are selected, and that different types of plants are installed in the network structures. With the shortest expected lead time, 1.6 days (NPV = $489.39 MM), (Figure 28) eight plants in the four sites are installed, and all the four suppliers are selected and connected to the associated nearest plant sites. The CA site is only supplied by the TX site for styrene monomer. As the expected lead time increases to 2.96 days (NPV = $644.46 MM), the supplier in LA is selected to provide benzene to the TX site, which leads to cheaper raw material, in turn increasing the NPV (Figure 29). As shown in Figure 30 and 31, a new inter-site transportation link from LA site to CA site for the shipping of styrene monomer is added. The change of network structure increases the expected lead time, and leads to the highest NPV up to $690 MM. These examples shows the importance of establishing trade-offs between responsiveness and economics in the design and planning of a PSCN for the improvement of overall earning and performance of a company.

8. CONCLUSIONS

In this paper we have presented a quantitative approach for designing responsive supply chains under demand uncertainty. The expected lead time was proposed as a measure of process supply chain responsiveness, and defined quantitatively with integration of supply chain network structure and inventory level. A multiperiod mixed integer nonlinear programming (MINLP) mathematical model was developed for the bi-criterion optimization of economics and responsiveness, while considering customer demand uncertainty. The model integrates the long-term strategic decisions (e.g. installation of plants, selection of suppliers, manufacturing sites, distribution centers and transport links) with the short-term operational decisions (e.g. product transitions and changeovers) for the multi-site multi-echelon process supply chain network. The model also includes a novel approach to predict the safety stock levels with consideration of responsiveness, demand uncertainty and economic objectives.

A bi-criterion optimization model was implemented to obtain the trade-offs between responsiveness and economics using the ε-constraint method. A hierarchical algorithm was further presented for the solution of the resulting large-scale MINLP problem based on decoupling of the decision-making levels (strategic and operational). Without compromising the solution quality, significant savings in computational effort was achieved by employing the proposed algorithm in the illustrative examples.

Two examples related to styrene production were solved to illustrate the industrial application of this model. The results show that small changes in expected lead time can lead to significant changes in the net present value and the network structure, which in turn suggests the importance of integrating responsiveness into the design and operations of process supply chain network.

Acknowledgements

The authors gratefully acknowledge financial support from National Science Foundation under Grant No. DMI-0556090.

Nomenclature

Indices/Sets

k	Manufacturing Sites
i	Plants
j	Chemicals
m	Distribution centers
ls	Suppliers
ld	Customers
t	Time periods
s, s'	Production schemes
K_i	Set of sites that can set up plant i
$J_{i,s}$	Set of chemicals involved in scheme s of plant i
$JP_{i,s}$	Set of main products for production scheme s of plant i
$JP_{k,i}$	Set of main products for plant i in site k
$JM_{k,i}$	Set of materials of plant i in site k
S_i	Set of production schemes for plant i
SL_i	Set of time slots for plant i in the production scheduling
LS_j	Set of suppliers that supply chemical j
LD_j	Set of customers that need chemical j
I_j	Set of plants that consume chemical j
O_j	Set of plants that produce chemical j
$Path_{ls,k,m,ld}$	Set of possible path of chemical flow from a supplier to some sites and distribution center, finally ends at a customer. Elements are in the form of $(ls, k_1, k_2 ... k_n, m, ld)$

Parameters

$Lenp_t$	Length of each time period t
$d^L_{j,ld,t}$	Lower bound of demand of chemical j in market ld during time period t
$d^U_{j,ld,t}$	Upper bound of demand of chemical j in market ld during time period t
$d^M_{j,ld,t}$	Most likely demand of chemical j in market ld during time period t
$d^m_{j,ld,t}$	Target demand of chemical j in market ld during time period t
$\alpha_{k,j,t}$	Coefficient for throughput working inventory amount of chemical j for site k
$\beta_{m,j,t}$	Coefficient for throughput working inventory of chemical j for distribution center m
$\varphi_{j,ld,t}$	Selling price of chemical j in market ld during time period t
$\varphi_{j,ls,t}$	Purchase price of chemical j in market ls during time period t
$\varepsilon_{j,m,t}$	Inventory cost of chemical j in distribution center m in time period t
$\varepsilon_{k,j,t}$	Inventory cost of chemical j in plant sites k in time period t
$\theta^R_{i,s}$	Residence time of the main product for production scheme s of plant i

$\gamma^I_{k,ls}$	Fixed cost of transport link from suppliers ls to plant sites k
$\gamma^O_{k,m}$	Fixed cost of transport link from plant sites k to distribution center m
$\gamma^S_{m,ld}$	Fixed cost of transport link from distribution center m to customer ld
$\gamma^N_{k,k'}$	Fixed cost of inter-plant site transportation
$\gamma^P_{k,i}$	Fixed cost of installation of plant i in site k
$\omega^I_{k,j,ls,t}$	Unit shipping cost for chemical j from suppliers ls to plant sites k in time period t
$\omega^O_{k,j,m,t}$	Unit shipping cost for chemical j from plant sites k to DC m in period t
$\omega^S_{j,m,ld,t}$	Unit shipping cost for chemical j from distribution center m to customer ld
$\omega^N_{k,k',j,t}$	Unit shipping cost for chemical j for inter-plant site transportation
$\omega^P_{k,i}$	Variable cost of installation of plant i in site k
$\lambda^I_{k,ls}$	Transportation time from supplier ls to plant site k
$\lambda^N_{k,k'}$	Transportation time from plant site k to k'
$\lambda^O_{k,m}$	Transportation time from plant site k to distribution center m
$\lambda^S_{m,ld}$	Transportation time from distribution center m to customer ld
$\rho^1_{i,s}$	relative production amount of main product j of production scheme s in plant i
$\rho^2_{i,s}$	relative maximum production rate of main product of production scheme s of plant i
$\mu_{i,j,s}$	Mass balance coefficients of chemical j in scheme s of plant i
$\sigma_{i,s,t}$	Unit operating cost of scheme s of plant i during period t
$\tau_{i,s,s'}$	Transition time from product s' to s in plant i
$CTR_{i,s,s'}$	Transition cost from product s to s' in plant i
$H_{k,i,t}$	Total available production time in plant i of site k in period t
$P^U_{k,j,ls,t}$	Upper bound of purchase of chemical j from supplier ls to site k during period t
$TR^U_{k,k',j,t}$	Upper bound of shipment of product j from site k to k' in period t
$S^U_{j,m,ld,t}$	Upper bound of sales of product j to market ld from distribution center m in time period t
$F^U_{k,j,m,t}$	Upper bound of shipment of product j from site k to distribution center m in time period t
$W^U_{k,i,j,s,t}$	Upper bound of production of chemical j in plant i of site k in period t
$Q^U_{k,i}$	Upper bound of capacity of each plant
$\delta^U_{k,i,s,sl,t}$	Upper bound of processing time for product s in slot sl of plant i in site k

Continuous Variables:

$Q_{k,i}$	Capacity of plant i in site k
$r_{k,i,s}$	Production rate of product s in plant i of site k
$W_{k,i,j,s,t}$	Amount of chemical j produced in plant i of site k in period t

$P_{k,j,ls,t}$	Purchase of chemical j from supplier ls to site k during period t
$S_{j,m,ld,t}$	Sales of product j to market ld from distribution center m during time period t
$F_{k,j,m,t}$	Shipping amount of chemical j from site k to distribution center m in time period t
$TR_{k,k',j,t}$	Shipping amount of chemical j from site k to k' in period t
$QS_{j,ld,t}$	Total available amount of chemical j for customer ld in time period t
$I_{j,m,ld,t}$	Safety stock of chemical j for market ld in distribution center m during period t
$WI_{k,j,t}^{PS}$	Working inventory of chemical j in site k during time period t
$WI_{m,j,t}^{DC}$	Working inventory of chemical j in distribution center m during time period t
TP	Expected lead time of the whole supply chain network
NPV	Net present value of the supply chain network
$\theta_{k,i}$	Time delay by production of plant i in site k
$Ts_{k,i,sl,t}$	Starting time of slot sl in plant i of site k in period t
$Te_{k,i,sl,t}$	End time of slot sl in plant i of site k in period t
$\delta_{k,i,s,sl,t}$	Processing time of scheme s in slot sl of plant i in site k
$\delta_{k,i,sl,t}$	Processing time of the time slot sl of plant i in site k
$TC_{k,i,t}$	Cycle time of plant i in site k in period t
$W_{k,i,s,t}^{S}$	Amount produced of main product in scheme s of plant i of site k in period t
$N_{k,i,t}$	Number of cycle in plant i of site k in period t
$COST_{k,i,t}^{S}$	Total cost for inventories and transitions of plant i in site k in period t
$K_{j,ld,t}$	Standardized normal variables of product j for customer ld in time period t
$Prob_{j,ld,t}$	stock-out probability for product j customer ld at time period t

Binary Variables

$Y_{k,ls}^{I}$	1 if a transportation link from supplier ls to plant site k is set up
$Y_{k,i}^{P}$	1 if plant i in site k is installed
$Y_{k',k}^{N}$	1 if an inter-site transportation link from site k' to site k is set up
$Y_{k,m}^{O}$	1 if transportation link from site k to distribution center m is set up
Y^{m}	1 if distribution center m is installed
$Y_{m,ld}^{S}$	1 if a transportation link from distribution center m to customer ld is set up
$SY_{k,i,s,sl,t}$	1 if the slot sl is assigned to the product s in plant i site k in period t
$Z_{k,i,s,s',sl,t}$	1 if product s is preceded by s' in time slot sl of plant i site k time period t

REFERENCES

1) Balas E., "Disjunctive Programming and a Hierarchy of Relaxations for Discrete Continuous Optimization Problems", *SIAM Journal on Algebraic and Discrete Methods*, **1985**, 6(3), 466-486
2) Bok, J. K.; Grossmann, I. E.; Park, S., "Supply Chain Optimization in Continuous Flexible Process Networks", *Industrial and Engineering Chemistry Research,* **2000**, 39, 1279-1290
3) Brooke, A.; Kendrick, D.; Meeraus, A.; Raman, R., "GAMS- A User's Manual", GAMS Development Corp., **1998**
4) Charnes, A.; Cooper W. W., "Deterministic Equivalents for Optimizing and Satisfying under Chance Constraints", *Operations Research,* **1963**, 11(1), 18-39
5) Chase, R. B.; Aquilano, N. J., "Production and Operations Management: Manufacturing and Service", Seventh Edition, Irwin, Chicago, **1995**
6) Chen, C. L.; Wang, B. W.; Lee, W. C., "Multiobjective Optimization for a Multienterprise Supply Chain Network", *Industrial and Engineering Chemistry Research*, **2003**, 42, 1879-1889
7) Christopher, M; Towill, D., "An integrated model for the design of agile supply chains", *International Journal of Physical Distribution & Logistics Management*, **2001**, 31, 235-246
8) Christopher, M., "The Agile Supply Chain Competing in Volatile Markets", *Industrial Marketing Management*, **2000**, 29, 37-44
9) Christopher, M., "Logistics and Supply Chain Management", Publisher: Prentice Hall, **2005**
10) Dua V.; Pistikopoulos, E.N., "Parametric Optimization in Process Systems Engineering: Theory and Algorithms", *Proceeding of Indian National Science Academy*, **2004**, 69A, 429-444
11) Duran, M. A.; Grossmann, I. E., "An outer approximation algorithm for a class of mixed integer nonlinear programs", *Mathematical Programming*, **1986**, 36, 307
12) Fisher, M.L., "What is the right supply chain for your product?", *Harvard Business Review*, **1997**, 75(2), 105-116
13) Forrester, J., "Industrial Dynamics", MIT Press, and John Wiley & Sons, Inc., New York **1961**
14) Grossmann I. E., "Enterprise-wide Optimization: A New Frontier in Process Systems Engineering", *AICHE Journal*, **2005**, 51, 1846-1857
15) Halshall-Whitney, H.; Thibault, J., "Multi-objective optimization for chemical processes and controller design: Approximating and classifying the Pareto domain", *Computers and Chemical Engineering*, **2006**, 30, 1155-1168
16) Holweg, M., "The three dimensions of responsiveness", *International Journal of Operations & Production Management*, **2005**, 25, 603-622
17) Jackson, J. R.; Grossmann, I. E., "Temporal Decomposition Scheme for Nonlinear Multisite Production Planning and Distribution Models", *Industrial and Engineering Chemistry Research*, **2003**, 42, 3045-3055
18) Kallrath, J., "Combined strategic and operational planning: An MILP success story in chemical industry", *OR Spectrum*, **2002**, 24, 315
19) Kok, A. G.; Graves, S. C., "Supply Chain Management: Design, Coordination and Operation, Handbooks in Operations Research and Management Science", Publisher: Elsevier, **2003**
20) Lee, S.; Grossmann, I. E., "New algorithms for nonlinear generalized disjunctive programming", *Computers and Chemical Engineering*, **2000**, 24(9-10), 2125-2141
21) Lee, H. L.; Padmanabhan, V.; Whang, S., "Information distortion in a supply chain: the bullwhip effect", *Management Science*, **1997**, 43(4), 546-558
22) Lengauer, T.; Tarjan, R. E., "A Fast Algorithm for Finding Dominators in a Flowgraph", *ACM Transactions on Programming Language and Systems*, **1979**, 1(1), 121-141
23) Liu, M. L.; Sahinidis, N. V., "Optimization of Process Planning Under Uncertainty", *Industrial and Engineering Chemistry Research*, **1996**, 35, 4154
24) Matson, J. B.; McFarlane, D. C., "Assessing the responsiveness of existing production operations", *International Journal of Operations & Production Management*, **1999**, 19(8), 765-784

25) McCormick, G. P., "Computability of Global Solutions to Factorable Nonconvex Programs – Part I – Convex Underestimating Problems", *Mathematical Programming*, **1976**, 10, 146
26) Norton, L. C.; Grossmann, I. E., "Strategic Planning Model for Complete Process Flexibility", *Industrial and Engineering Chemistry Research*, **1994**, 33, 69-76
27) Petkov, S. B.; Maranas, C. D., "Design of Single Product Campaign Batch Plants under Demand Uncertainty", *AICHE Journal*, **1998**, 44, 896
28) Pinto, J. M.; Grossmann, I. E., "Optimal cyclic scheduling of multistage continuous multiproduct plants", *Computers and Chemical Engineering*, **1994**, 18, 797-816
29) Quesada, I..; Grossmann, I. E., "Global Optimization of Bilinear Process Networks with Multicomponenet Flows", *Computers and Chemical Engineering*, **1995**, 12, 1219
30) Raman, R.; Grossmann, I. E., "Symbolic integration of logic in mixed integer linear programming techniques for process synthesis", *Computers and Chemical Engineering*, **1993**, 27, 909
31) Sabath, R., "Volatile demand calls for quick response", *International Journal of Physical Distribution & Logistics Management*, **1998**, 28(9/10), 698-704
32) Sahinidis, N. V., "BARON: A general purpose global optimization software package", *Journal of Global Optimization*, **1996**, 8(2), 201-205
33) Sahinidis, N. V.; Grossmann, I. E., "MINLP model for cyclic multiproduct scheduling on continuous parallel lines", *Computers and Chemical Engineering*, **1991**, 15(2), 85-103
34) Sahinidis, N. V.; Grossmann, I. E.; Fornari, R. E.; Chathrathi, M., "Optimization model for long range planning in the chemical industry", *Computers and Chemical Engineering*, **1989**, 13(9), 1049-1063
35) Schulz, E. P.; Diaz, M. S.; Bandoni, J. A., "Supply chain optimization of large-scale continuous processes", *Computers and Chemical Engineering*, **2005**, 29, 1305-1316
36) Shah, N., "Process industry supply chains: Advances and challenges", *Computers and Chemical Engineering*, **2005**, 29, 1225-1235
37) Shaw, N. E.; Burgess, T. F.; Mattos, C.; Stec, L. Z., "Supply chain agility: the influence of industry culture on asset capabilities within capital intensive industries", *International Journal of Production Research*, **2005**, 43(16), 3497-3516
38) Shen, Z. M.; Coullard, C.; Daskin, M. S., "A Joint Location-Inventory Model", *Transportation Science*, **2003**, 37(1), 40-55
39) Sousa, R. T.; Shah, N.; Papageorgiou, L. G., "Supply Chain Design and Planning with Responsiveness Testing – a Two-level Holistic Approach to an Industrial Case", *AICHE Annual Meeting*, San Francisco, CA, **2006**
40) Subrahmanyam, S.; Pekny, J. F.; Reklaitis, G. V., "Design of Batch Chemical Plants under Market Uncertainty", *Industrial and Engineering Chemistry Research*, **1994**, 33, 2688-2701
41) Tsiakis, P.; Shah, N.; Pantelides, C. C., "Design of Multi-echelon Supply Chain Networks under Demand Uncertainty", *Industrial and Engineering Chemistry Research*, **2001**, 40, 3585-3604
42) Voelkel, J. G.; Chapman, C., "Value Stream Mapping", *Quality Progress*, **2003**, 36(5), 65-69
43) Wellons, H. S.; Reklaitis, G. V., "The design of multiproduct batch plants under uncertainty with staged expansion", *Computers and Chemical Engineering*, **1989**, 13, 115
44) Wolsey, L. A., "MIP modeling of changeovers in production planning and scheduling problems", *European Journal of Operational Research*, **1997**, 99, 154-165
45) You, F.; Grossmann, I. E., "Optimal Design and Operational Planning of Responsive Process Supply Chains", Supply Chain Optimization (Papageorgiou and Georgiadis, editors), Wiley-VCH, Weinheim, **2007**
46) Yusuf, Y. Y.; Gunasekaran, A.; Adeleye, E. O.; Sivayoganathan, K., "Agile supply chain capabilities: Determinants of competitive objectives", *European Journal of Operational Research*, **2004**, 159, 379-392
47) Zimmermann, H. –J., "An application-oriented view of modeling uncertainty", *European Journal of Operational Research*, **2000**, 122, 190-198

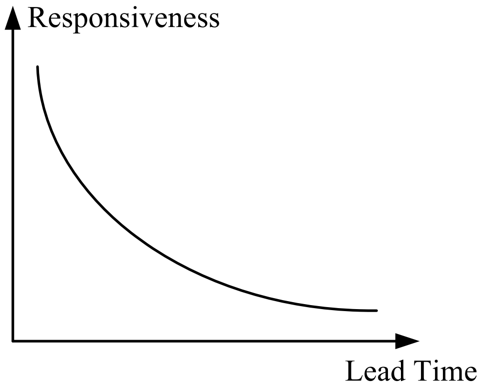

Figure 1. Conceptual relationship between lead time and responsiveness

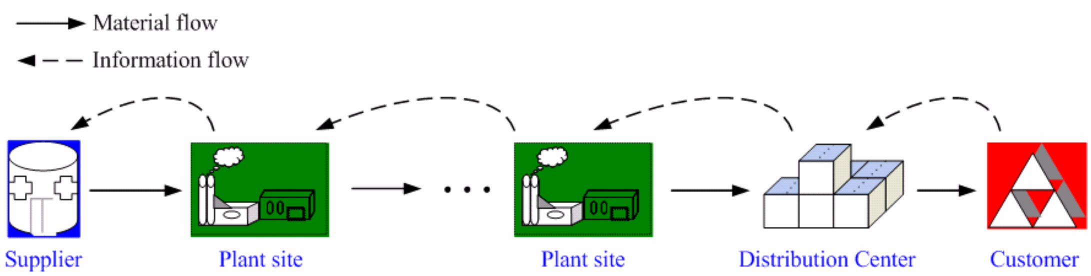

Figure 2. A simple linear supply chain

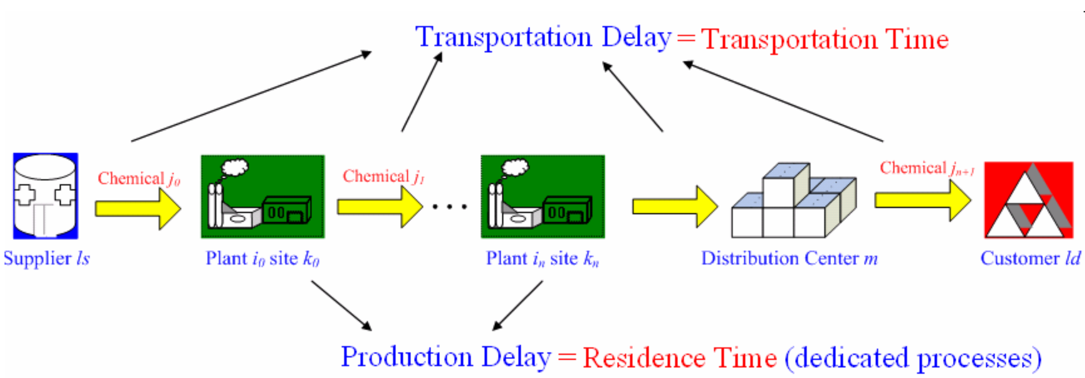

Figure 3. Time delays of a simple linear supply chain

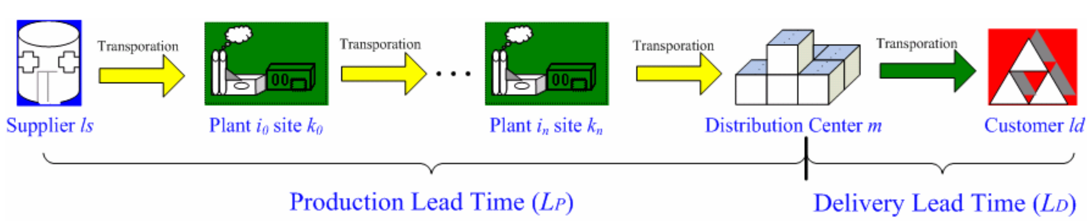

Figure 4. Production lead time and delivery lead time

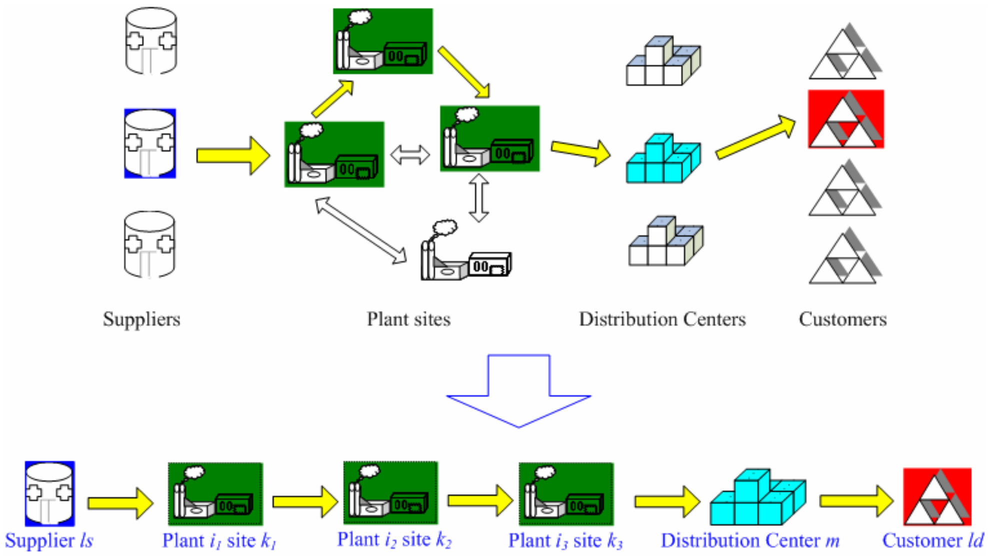

Figure 5. A path of chemical flow in a PSCN

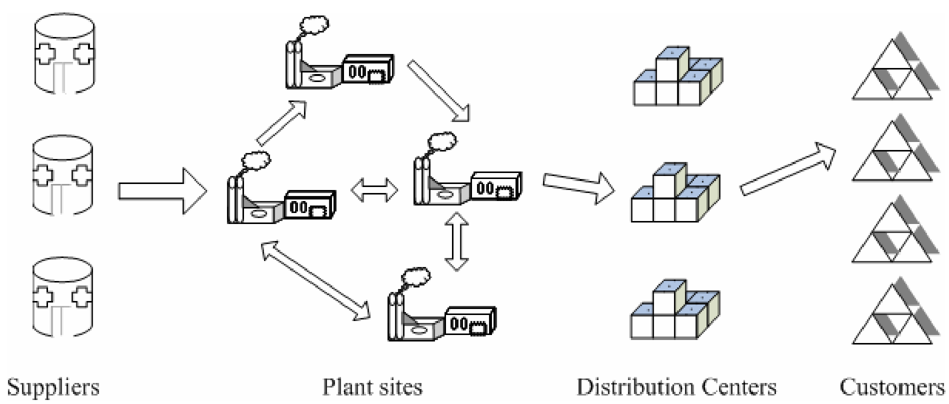

Figure 6. Process supply chain network

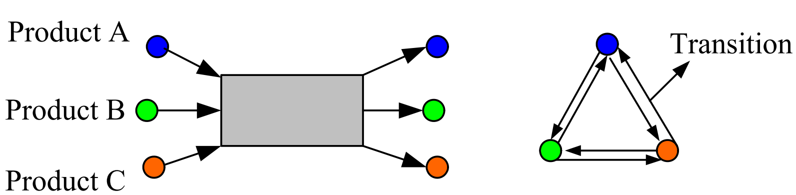

Figure 7. Changeovers of flexible processes

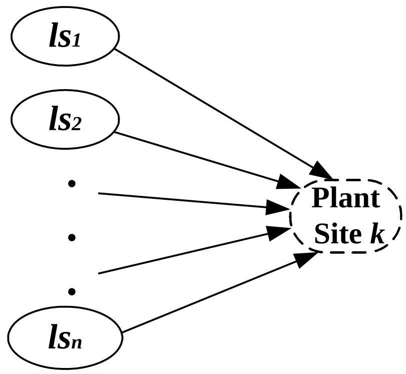

Figure 8. Relationship between suppliers and manufacturing sites

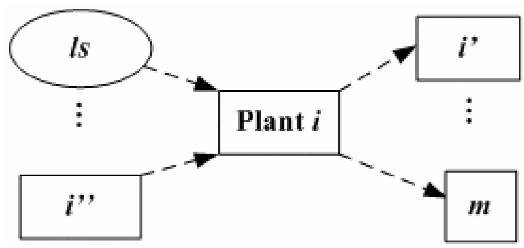

Figure 9. Input and output relationship of a plant

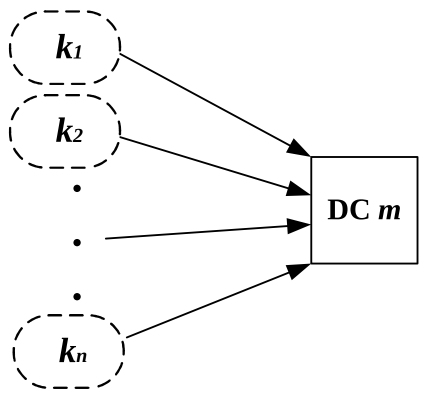

Figure 10. Relationship between manufacturing sites and distribution centers

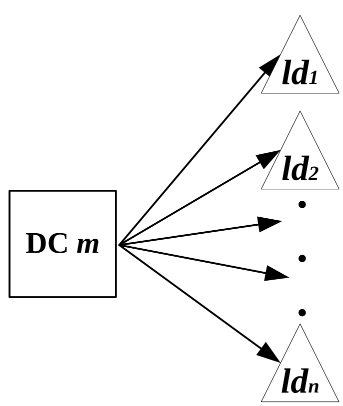

Figure 11. Input and output relationship of a distribution center

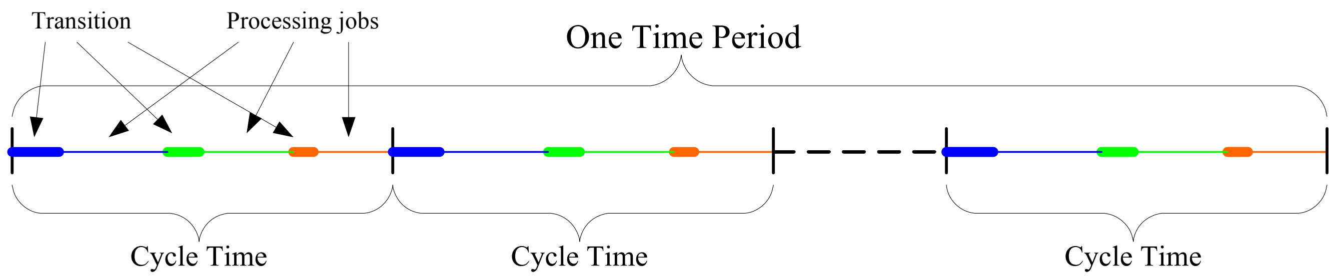

Figure 12. Cyclic scheduling of each time period

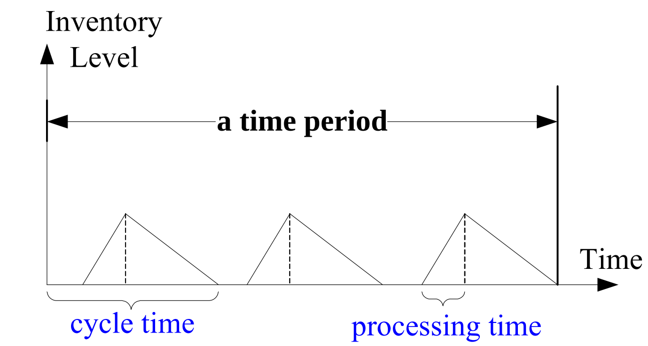

Figure 13. Inventory level change in cyclic scheduling

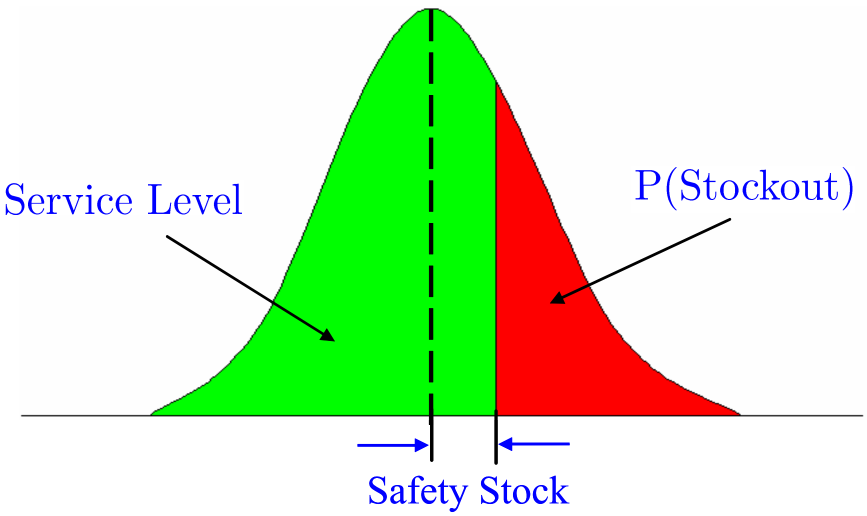

Figure 14. Service level and stock-out probability

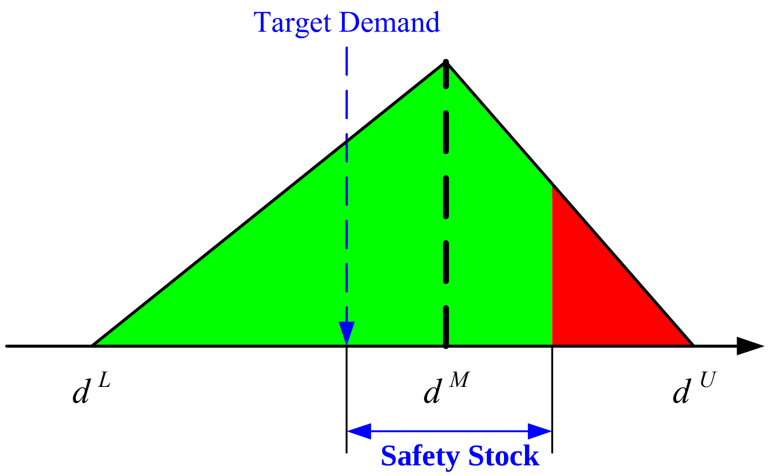

Figure 15. Safety stock for triangular distribution

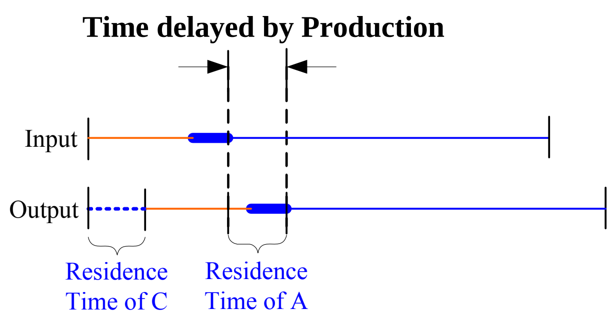

Figure 16. Time delay by production (large change of demand)

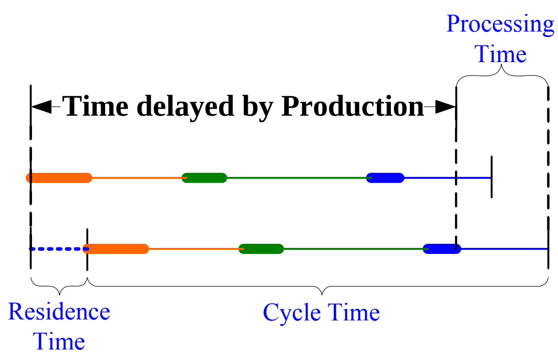

Figure 17. Time delay by production (small change of demand)

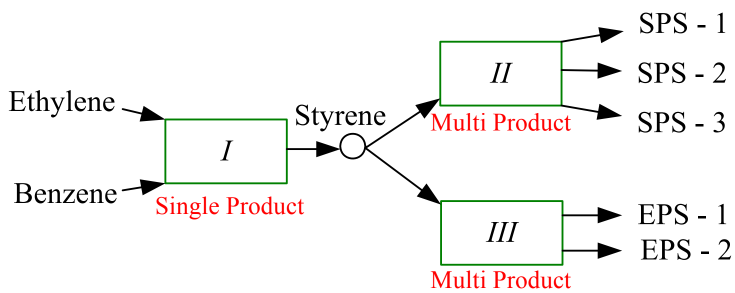

Figure 18. Production network for polystyrene supply chains

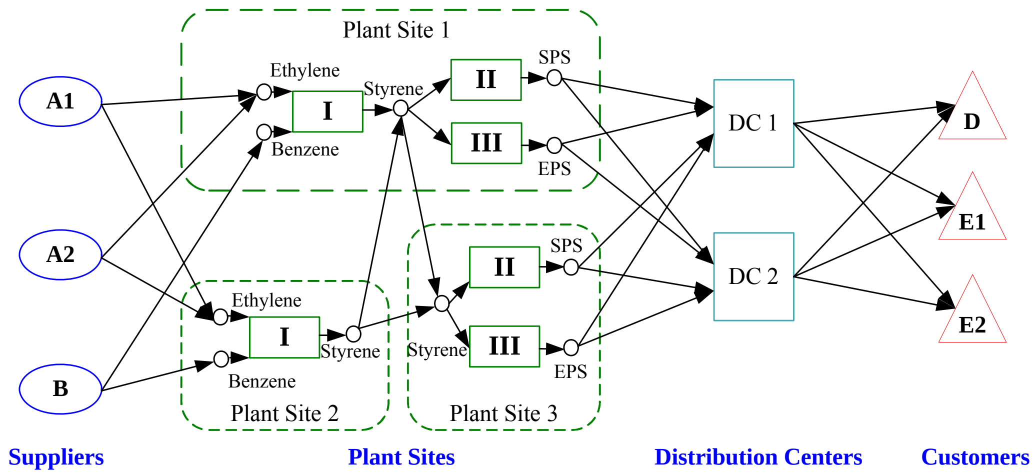

Figure 19. Potential process supply chain network superstructure for Example 1

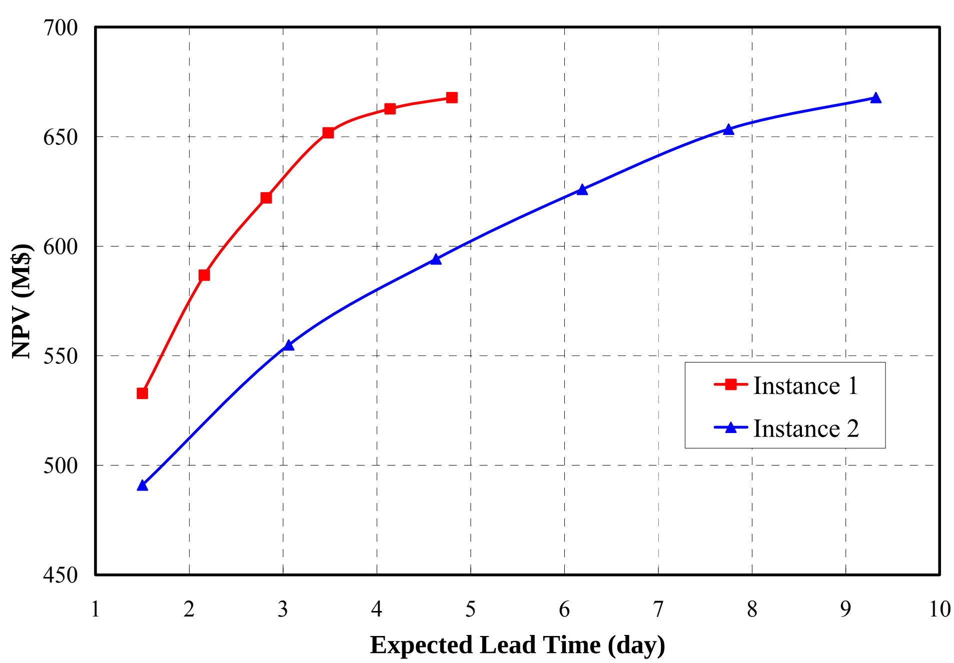

Figure 20. Pareto optimal curve for different production delay definitions in Instance 1 and 2 of Example 1 (demand uncertainty both follow triangular distribution)

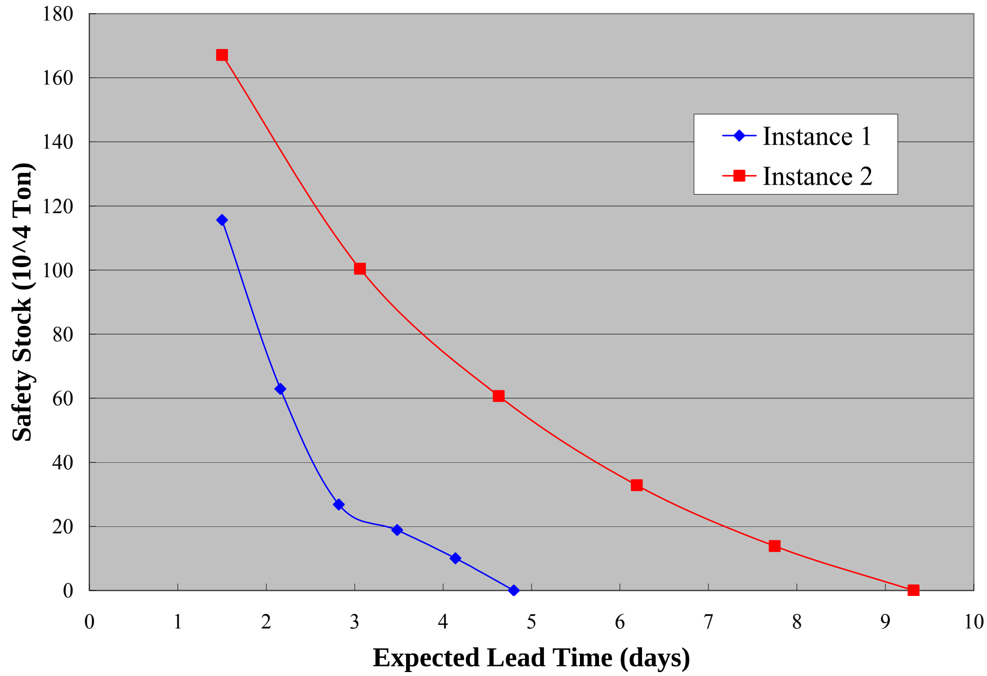

Figure 21. Safety stock levels for EPS resins in DC1 for Instance 1 and 2 of Example 1 (triangular distributed demand for both instances)

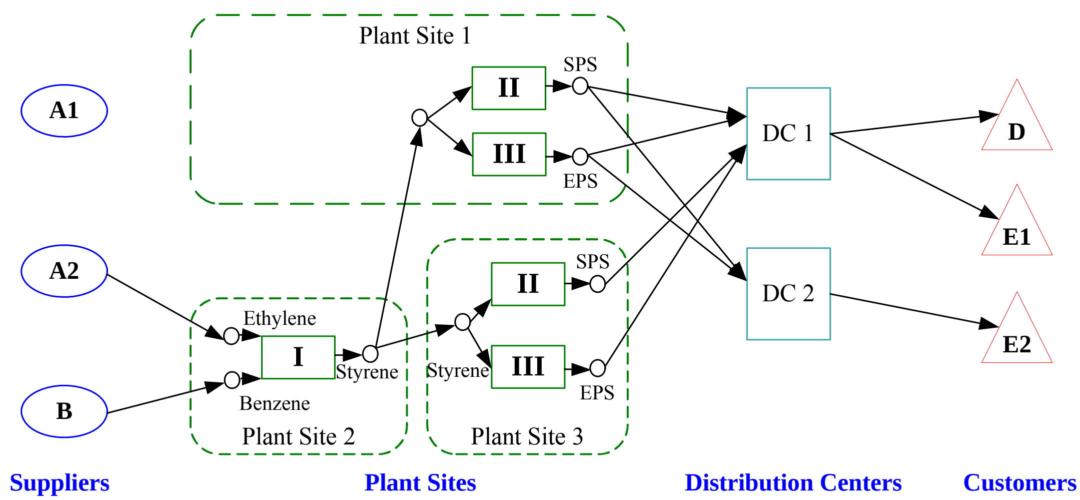

Figure 22. Optimal network structure for Instance 1 and 2 of Example 1

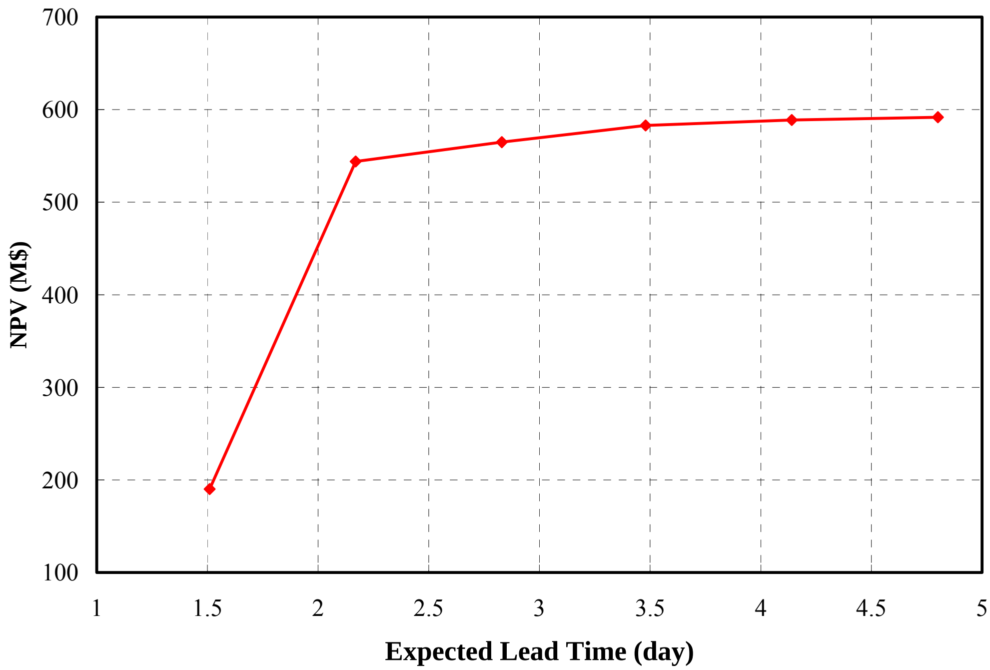

Figure 23. Pareto optimal curve for Instance 3 of Example 1 (normal distributed demand)

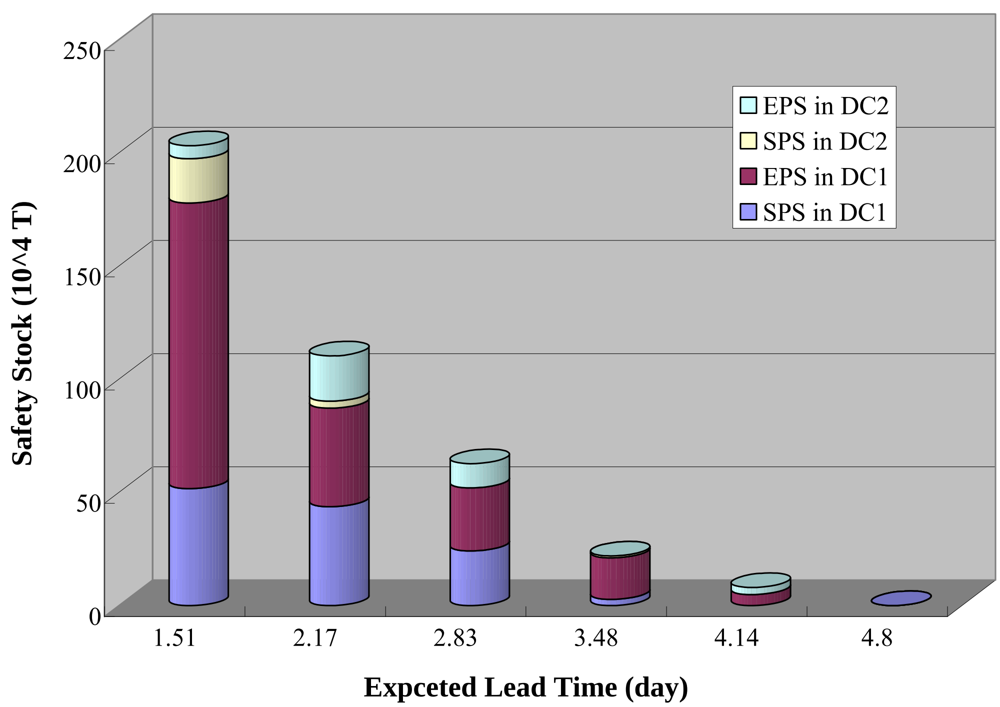

Figure 24. Safety stock level for normal distributed demand in Instance 3 of Example 1

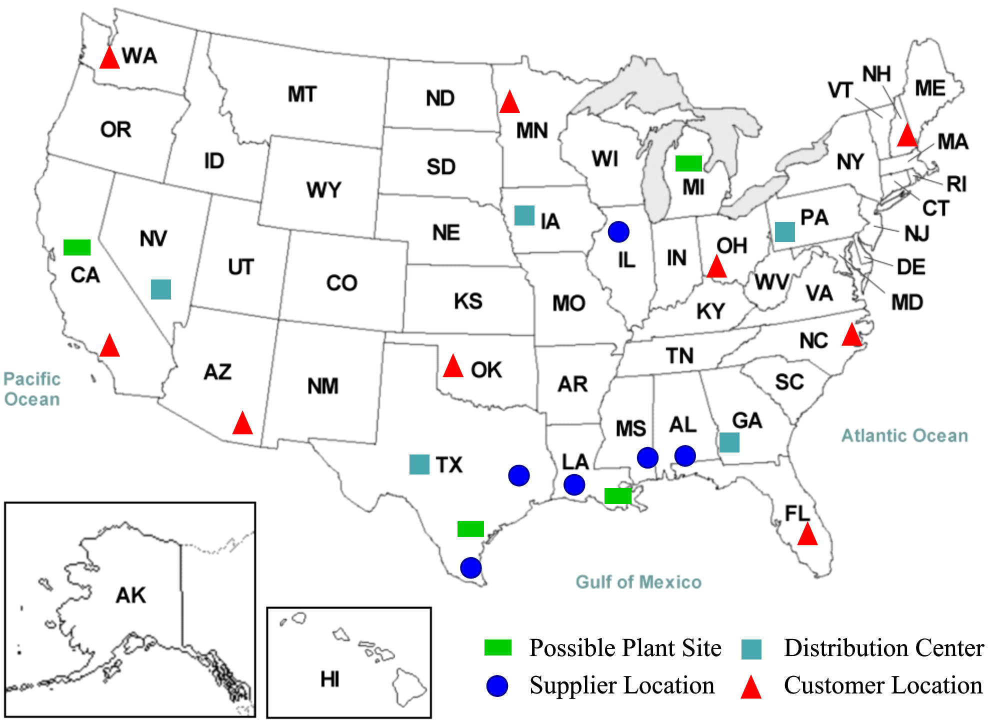

Figure 25. Location map for Example 2

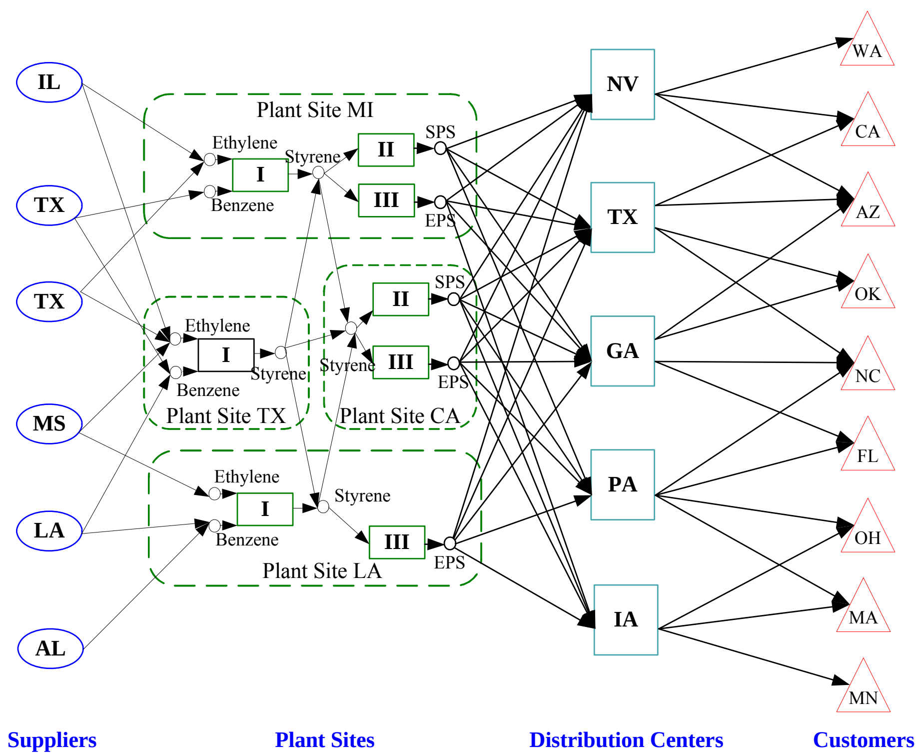

Figure 26. Potential process supply chain network superstructure for Example 2

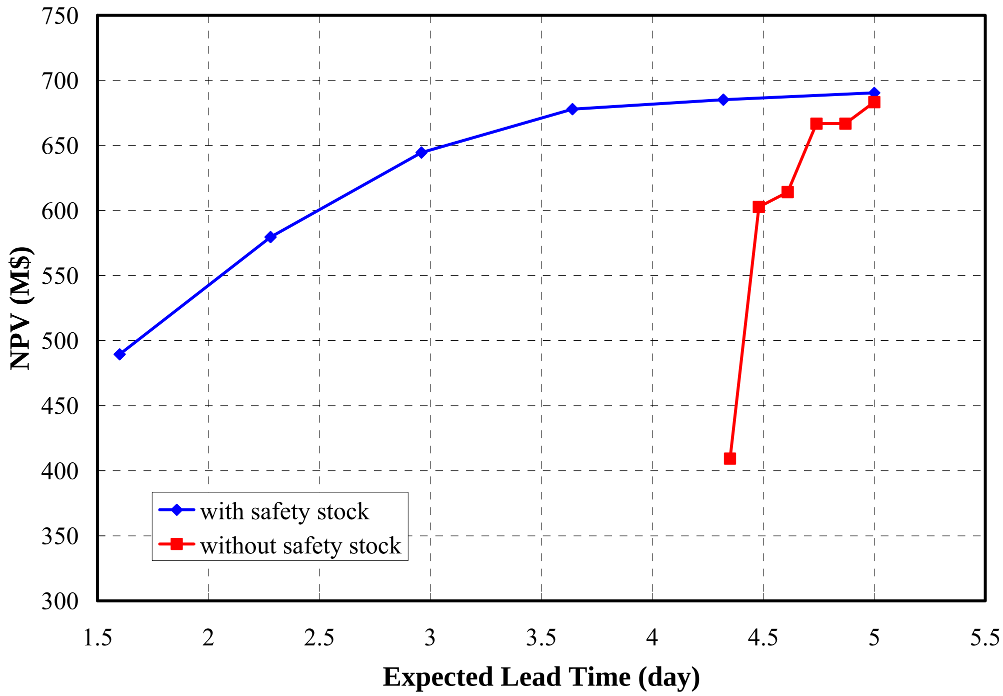

Figure 27. Pareto curve for Example 2 with safety stock and without safety stock (Instance 1 and 2 of Example 2)

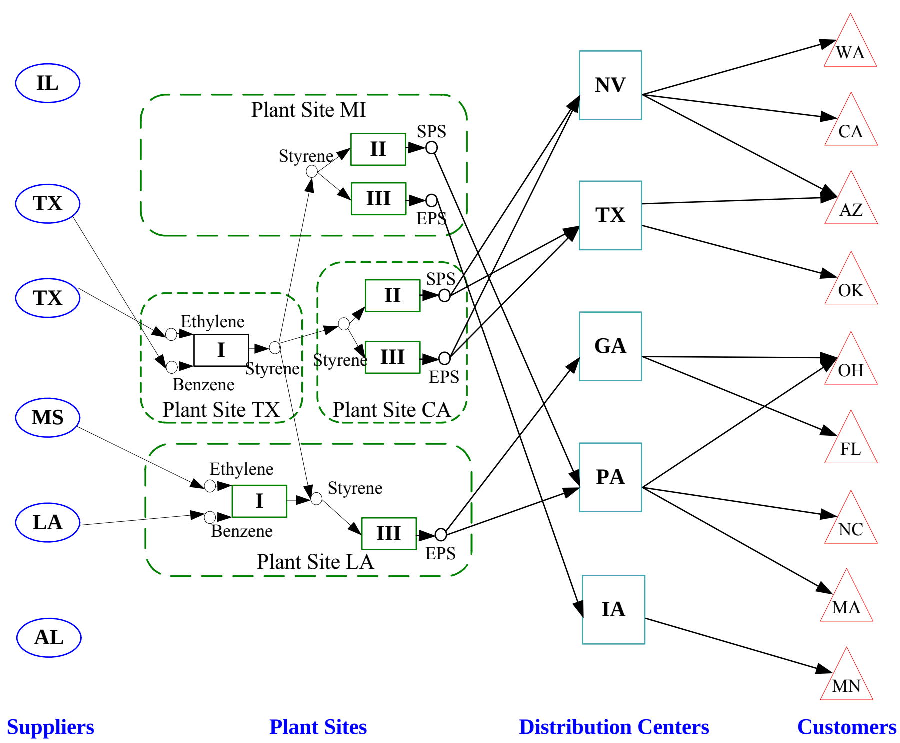

Figure 28. Optimal network structure of Example 2 at expected lead time = 1.5 days, NPV = $489.39 MM

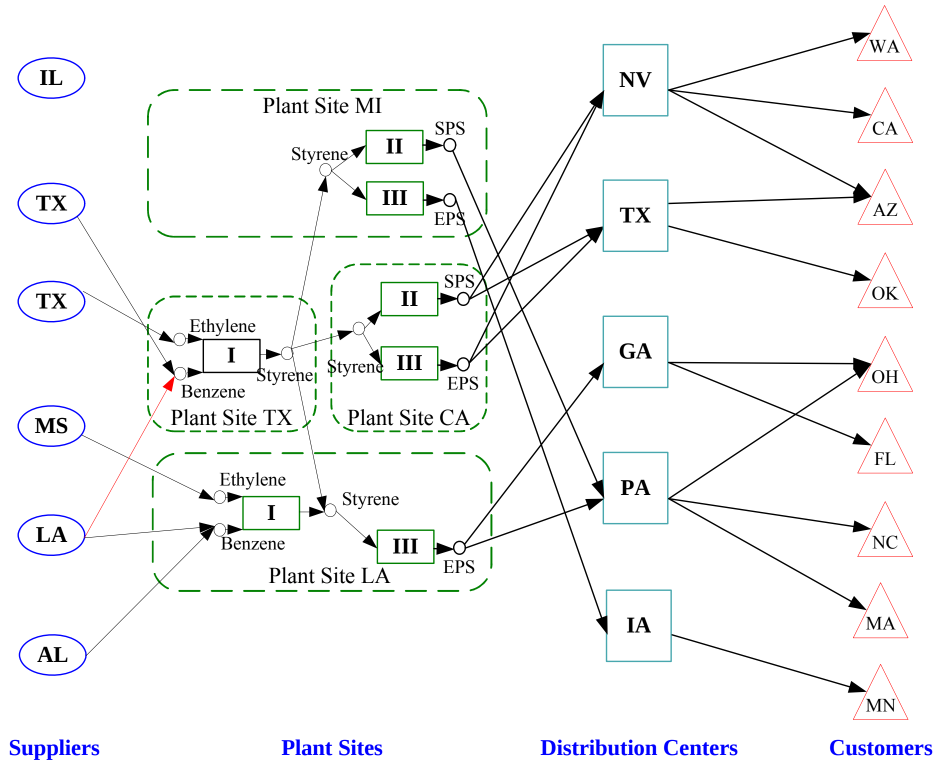

Figure 29. Optimal network structure of Example 2 at expected lead time = 2.96 days, NPV = $644.46 MM

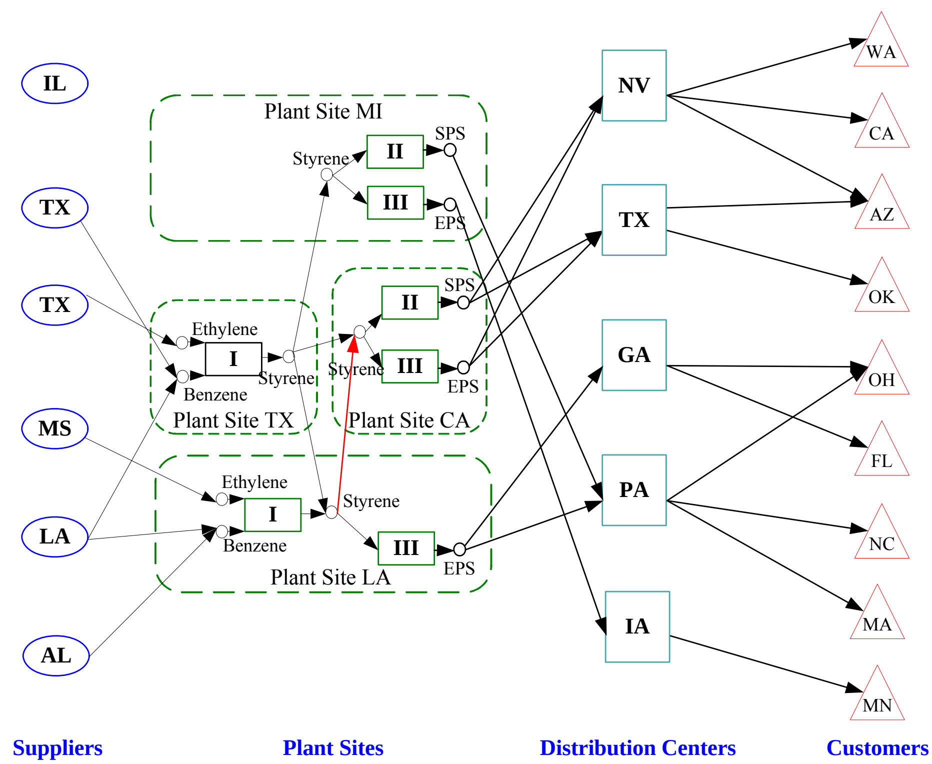

Figure 30. Optimal network structure of Example 2 at expected lead time = 5.0 days, NPV = $690 MM

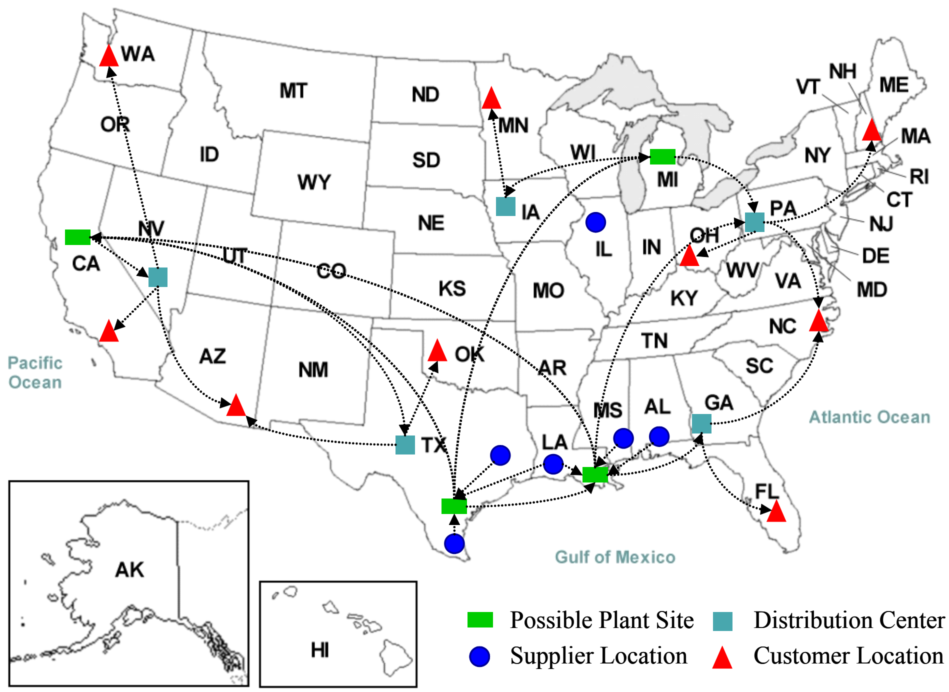

Figure 31. Material flows in the location map for the longest expected lead time (5.0 days) case of Example 2

Instances	Demand Uncertainty	Production Delay	Solution Approach	Safety Stock
Example 1, Instance 1	Triangular	Constraint 42.a	Direct	Yes
Example 1, Instance 2	Triangular	Constraint 42.b	Direct	Yes
Example 1, Instance 3	Triangular	Constraint 42.a	Hierarchical	Yes
Example 1, Instance 4	Normal	Constraint 42.a	Hierarchical	Yes
Example 2, Instance 1	Triangular	Constraint 42.a	Hierarchical	Yes
Example 2, Instance 2	Triangular	Constraint 42.a	Hierarchical	No

Table 1. Summary of all the numerical examples and instances for case studies

Points on Pareto Curve	Expected Lead Time (days)	Direct Approach (Example 1, Instance 1)		Hierarchical Approach (Example 1, Instance 3)		Difference from Optimum
		NPV (M$)	CPU(s)	NPV (M$)	CPU(s)	
I	1.5	532.82	316.48	514.68	104.25	3.40%
II	2.16	586.78	484.08	568.65	70.52	3.09%
III	2.82	622.07	649.85	603.94	106.49	2.91%
IV	3.48	651.73	1292.35	633.6	168.33	2.78%
V	4.14	662.69	960.83	644.56	138.99	2.74%
VI	4.8	667.79	359.65	649.65	101.34	2.72%

Table 2. Comparison for six points in the Pareto curve with different solution approaches (Instance 1 and 3 of Example 1). In the direct approach (Instance 1), the optimality margin is set to be 0%, which is the global optimum; in the hierarchical approach (Instance 3), the optimality margin is set to 0% for simplified model and 5% for the detail model.