Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computers and Chemical Engineering 28 (2004) 2087–2106

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

A simulation based optimization approach to supply chain management under demand uncertainty☆

June Young Jung[a], Gary Blau[a], Joseph F. Pekny[a], Gintaras V. Reklaitis[a,*], David Eversdyk[b]

[a] School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA
[b] Dow Chemical, 2301 Brazosport Blvd., Freeport, TX 77541, USA

Received 9 April 2004; received in revised form 7 June 2004; accepted 8 June 2004

Available online 20 July 2004

Abstract

Cost effective supply chain management under various market, logistics and production uncertainties is a critical issue for companies in the chemical process industry. Uncertainties in the supply chain usually increase the variance of profits (or costs) to the company, increasing the likelihood of decreased profit. Demand uncertainty, in particular, is an important factor to be considered in the supply chain design and operations. To hedge against demand uncertainty, safety stock levels are commonly introduced in supply chain operations as well as in supply chain design. Although there exists a large body of literature on estimating safety stock levels based on traditional inventory theory, this literature does not provide an effective methodology that can address the complexity of real CPI supply chains and that can impact the current practice in their design, planning and scheduling. In this paper, we propose the use of deterministic planning and scheduling models which incorporate safety stock levels as a means of accommodating demand uncertainties in routine operation. The problem of determining the safety stock level to use to meet a desired level of customer satisfaction is addressed using a simulation based optimization approach. An industrial-scale case problem is presented to demonstrate the utility of the proposed approach.

Keywords: Supply chain management; Customer satisfaction level; Safety stock levels; Demand uncertainty; Stochastic optimization; Simulation based optimization

1. Introduction

The chemical process industry faces uncertainties in a range of factors such as demands for products, prices of raw materials and products, lead times for the supply of raw materials and in the production and the distribution of final products, process failures and quality failures (Subramanyam, Pekny, & Reklaitis, 1994). Among these, demand uncertainty may well have the dominant impact on profits and customer satisfaction. Demand uncertainty can result in over- or under-production, with resultant excess inventories or inability to meet customer needs, respectively. Excess inventory incurs unnecessary holding costs, while the inability to meet the customer needs results in both loss of profit and potentially, the long term loss of customers. The expected value of the ability to meet the product needs of a customer is traditionally called the customer satisfaction level. Under competitive market conditions, customer satisfaction level (or service level) is recognized as an important index that must be monitored and maintained at a high level. A high level of customer satisfaction can be achieved by maintaining increased inventory levels to hedge against demand uncertainties. Although additional inventory improves customer satisfaction, it entails increased inventory holding cost. This trade-off between customer satisfaction and inventory holding cost can be posed as a multi-stage stochastic optimization problem in which both the production and the inventory levels are the key optimization variables. Since even the full deterministic version of such problems can be challenging for routine solution, there is a need for practical alternatives.

One such approach entails the use of the concept of safety stock: a time independent lower bound on the inventory level which is chosen such as to absorb some level of the demand

☆ Supplementary data associated with this article can be found at doi: 10.1016/j.compchemeng.2004.06.006.

* Corresponding author. Tel.: +1 765 4943839; fax: +1 765 4940805.
E-mail addresses: reklaiti@ecn.purdue.edu (G.V. Reklaitis), daeversdyk@dow.com (D. Eversdyk).

0098-1354/$ – see front matter
doi:10.1016/j.compchemeng.2004.06.006

uncertainty. Indeed, there exists a large body of literature on estimating safety stock levels based on traditional inventory theory. While the analytical solutions developed in that work are elegant in their simplicity, they fail to address the key features of realistic supply chain problems, namely, multiple products sharing multiple production facilities with capacity constraints and demands originating from multiple customers. In addition, in real world supply chains, safety stock levels are dependent on factors such as the probabilistic distribution of the demands, the demand to capacity ratio, the dependence of the overall customer satisfaction level on meeting the demands for several different products produced in the same production facility and even the frequency with which production plans and schedules for the facility are updated. These factors introduce complexities which traditional inventory models simply do not accommodate.

The chemical process industry has a history of using deterministic linear programs (LP) and/or mixed integer linear programs (MILP) for supply chain planning and scheduling. In practice, demand and other uncertainties are accommodated in part by using such models in the so-called rolling horizon mode (Reklaitis, 1982). That is, the operating horizon is divided into a certain number of planning periods, the planning model with suitable deterministic demand forecasts is solved to yield production targets for each period, and the targets for the first or first few periods are used as inputs for deterministic scheduling models which yield detailed time-based production and sequencing solutions. At the end of the first period, the state of the system, including inventory levels, is updated and the planning and scheduling cycle is repeated with the horizon advanced by one period. The notion of safety stock can be imbedded in such planning models by including lower bounds on the inventory level levels of various products and/or sites. However, the incorporation of a measure of customer satisfaction is problematic since this is inherently a stochastic quantity. In this paper, we propose an approach for the determination of safety stock levels for use within the classical rolling horizon planning and scheduling strategy, which takes the stochastic nature of the customer satisfaction level into account.

In the following sections, a literature review will first be given. Then, we pose the rigorous problem of supply chain optimization and use this to motivate an approximate approach to the problem which can be implemented using existing tools. Then we describe the proposed computational approach for determining suitable safety stock levels. Finally the description of a case problem is given and illustrative results are reported.

2. Review of literature

Quantitative issues related to inventory management are addressed in a large body of literature known as stochastic inventory theory. For example, the literature survey by Porteus (1990) lists 122 references investigating these issues. Based on foundational work such as the classical EOQ (economic order quantity) model, Newsvendor model, and the optimal ordering policies, such as (S,s), (Q,r) and the base stock policy, research in inventory theory has been ongoing to extend these analytical results to the conditions experienced in the day-to-day operation and design of supply chains. An extensive discussion of the foundational work can be found in Porteus (2002) and Hopp and Spearman (2000). The works of DeCroix and Arreola-Risa (1998), Zipkin (1986) and Graves and Willems (2000) provide further examples of the ongoing extensions, where the consideration of production is incorporated in the inventory management problems. A key limitation of all of these studies, however, is that the effects that production planning and scheduling can and does have on inventories are not considered. Nonetheless the notion of safety stock level as simple target that can be used in routine operations has considerable merit and acceptance in practice.

In the process systems engineering literature, the problems associated with the optimal design and operation of process systems under uncertainty has been attracting increasing attention since the early 1990s. Subramanyam et al. (1994) proposed a hierarchical approach in which the design and scheduling stages are decomposed. The scheduling stage validates the feasibility of the design under various scenarios of different demand realizations. Ierapetritou and Pistikopoulos (1994) also proposed a two-stage stochastic optimization formulation for the planning and capacity expansion problem and reported a mathematical programming solution strategy. Shah and Pantelides (1992) reported a scenario-based approach to the design of multi-purpose batch plant under uncertain production requirements. However, the issues related with safety stock level or customer satisfaction were not addressed in these studies. Instead, hedged decisions on planning or design variables against demand uncertainty were sought. Gupta et al. published a series of studies on supply chain planning under demand uncertainty. Gupta and Maranas (2000a) formulated a two-stage stochastic program based on the MILP planning model proposed by McDonald and Karimi (1997). The two-stage model is composed of a here-and-now production model which constitutes the first stage and a wait-and-see inventory and distribution model which constitutes the second-stage problem. The uncertain demand parameter is included in the inventory and distribution model. The advantage of their approach lies in the fact that they derived a closed-form solution for the second-stage problem, such that the expectation of the second stage can be directly evaluated. However, the applicability of the model is limited to single period and single customer cases. In their second paper (Gupta et al., 2000b), which builds on the approach proposed in the first paper, chance constraints were also considered, allowing handling of customer satisfaction constraints. Gupta and Maranas (2003) further generalize their approach to the multi-period case. However, this approach may lead to a relatively optimistic measure of customer satisfaction level due to the inability to capture the dynamics of inventory changes within a planning period. In addition, as the optimization of their model is based on the trade-off between production cost and customer satisfaction, the production amount is adjusted to hedge against the demand uncertainty. This leaves unanswered the issues regarding to determination of the safety stock level.

Subramanian, Pekny, & Reklaitis (2001a) and Subramanian (2001b) proposed a computational architecture that addresses the combinatorial and stochastic aspects of R&D pipeline management. A simulation based optimization framework is developed for decision making with respect to project portfolio selection and project task scheduling. The computational framework, called the "Sim-Opt" architecture, combines deterministic mathematical programming for maximization of net present value of the R&D pipeline with discrete event simulation to incorporate the uncertainty in the successful progression of each project in the pipeline. The concept of timelines is used to study multiple unique realizations of the controlled evolution of the discrete-event pipeline system. Since Monte Carlo simulation provides flexibility in accommodating various types of uncertain parameters and their distributions, this framework has high potential for application to a wide range of large scale stochastic optimization problems. This "Sim-Opt" architecture is adopted as a part of the computational framework of this research.

Simulation-based optimization is an active area in the field of the stochastic optimization. Reviews of the current research on simulation-based optimization developments can be found in Carson and Maria (1997) and Andradottir (1998). Fu (2002) and Fu and Hu (1997) address the theoretical aspect of gradient based approaches to simulation optimization which will be applied in our work.

3. Problem description

Fig. 1 depicts the conceptual relation between the customer satisfaction level and the safety stock level of a product under uncertain demand. Each curve represents the relation between these two variables at a given level of demand variance. The curves associated with increasing demand variance are located in monotonically lower positions. Since the customer satisfaction level is a monotonically increasing function of the safety stock level, to achieve higher customer satisfaction, more safety stock is needed, resulting in increased inventory holding cost. The underlying assumption is that the safety stock and the production system operating at its normal production rate will be sufficient to meet the expected demand variations. In a sense, the safety stock provides a buffer that allows the production system to catch up to a surge in demand. In the presence of finite production capacity or rate, when the demand variation becomes sufficiently large, then for any given finite level of safety stock it will no longer be possible to meet customer demand or, in the expected value sense, to achieve the desired level of customer satisfaction. The trade-off between maximizing the customer satisfaction level and minimizing the safety stock level thus must also involve the production capacity constraint and thus the resulting problem is a constrained stochastic optimization problem.

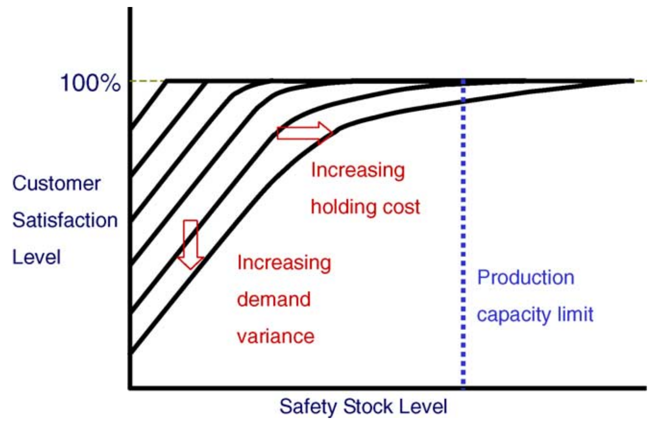


Fig. 1. Conceptual relation between customer satisfaction level and safety stock level.

Over the range of demand variance, the different ranges of the ratio of the expected value of the demand to the production capacity give rise to three different operational regimes. In regime I, when the demand to capacity ratio is sufficiently small, the production facility has sufficient spare capacity to cope with abrupt changes in the demand. Therefore, in this regime, relatively low or even zero safety stock level may be sufficient to achieve the desired customer satisfaction. In regime II, characterized by an intermediate range of expected demand to capacity ratio, the production capacity may be quite constrained when the demands for different products spike at some point in time. In this regime, even in the presence of safety stock, the customer satisfaction level for some products sharing the production facility may fail to reach their target values. Finally in regime III, the demand to capacity ratio is sufficiently large that the combined demands of different customers simply cannot be met and thus these demands will compete for the production resources. In this regime, the safety stock and production resources must be assigned strategically to meet the demands of some customers in preference to others. In this paper, we will address the safety stock levels appropriate under regimes I and II. The determination of safety stock levels for the case of regime III involves strategies for the assignment of quotas or the prioritization of customer demands. Such strategies, while of considerable interest, are beyond the scope of this work.

To formulate the problem that fully incorporates supply chain features relevant to the chemical process industry, the following problem context is assumed:

- A CPI company has multiple plant sites. Within each plant site, there are multiple production lines of processing units. Each production line has different production capacities and is assigned to a specific family of products based on their changeover costs and demand amounts. The units in different production lines may be connected to form a production network.

- The company produces multiple products and the products are categorized into several product families based on their similarity in physical or chemical properties.
- The actual customers are grouped into several sales regions and the demand forecast for each product is generated based on the sales regions.
- The demand for each product has seasonal and regional variation. The demand forecast can capture this variation over a one year horizon. The actual demand uncertainty for a product can be characterized by a certain probability distribution. The demand forecast is accurate enough to approximate the mean value of its actual realization. The standard deviation or probability distribution of the demand can be estimated from historical demand data.
- Each production facility can supply any sales region and the transportation resources are always available. Hence, in this study, logistical decisions are not explicitly considered.
- The Supply chain manager executes a multiperiod planning model and a scheduling model using the rolling horizon strategy. The planning model takes as input the current state of the system and the demand forecast for each product at each period over the planning horizon. The supply plan and/or production plan provides input to the scheduling model, which is used to generate actual production sequence, start time and amounts.

We consider the problem in which the company has set a target customer satisfaction level for each final product and wants to determine the safety stock level for each product at each of its production sites so as to minimize the expected value of the cost of supply chain.

4. The underlying stochastic problem

In this section, we first present a multi-stage stochastic program as a formalization of the above problem. In this formulation, the safety stock levels are not explicitly stated, rather they are implicit in the individual product inventory variables. The difficulties associated with the routine and regular solution of this formulation are discussed. Then, we propose a computational approach in which the safety stock levels are included as explicit variables and in which the determination of the safety levels is made so as to achieve desired customer satisfaction levels taking into account the stochastic demands. The production, inventory and supply variable values are generated using deterministic models in the rolling horizon mode.

4.1. Multi-stage stochastic program

The supply chain problem under consideration is inherently a stochastic multi-stage decision problem in the operating variables and involving several sets of operating and structural constraints. Each decision stage corresponds to a planning period (denoted by l) and each planning period is further divided into scheduling time periods (denoted by t).

Objective Function:

$$\begin{aligned}
&\mathrm{Min} \sum_{i,j,s,t\in\tau_1} c_{ijs}P_{ijst} + E_{\omega^1}\Bigg[\mathrm{Min} \sum_{i,s,c,t\in\tau_1} t_{sc}S_{isct} \\
&\quad + \sum_{i,s,t\in\tau_1} h_{is}I_{ist} + \sum_{i,s,t\in\tau_1} \xi_i\Gamma_{ist} + \sum_{i,j,s,t\in\tau_2} c_{ijs}P_{ijst} \\
&\quad + E_{\omega^2}\Bigg[\mathrm{Min} \sum_{i,s,c,t\in\tau_2} t_{sc}S_{isct} + \sum_{i,s,t\in\tau_2} h_{is}I_{ist} + \sum_{i,s,t\in\tau_2} \xi_i\Gamma_{ist} \\
&\quad + \sum_{i,j,s,t\in\tau_3} c_{ijs}P_{ijst} + E_{\omega^3}\Bigg[\dots E_{\omega^T}\Bigg[\mathrm{Min} \sum_{i,s,c,t\in\tau_T} t_{sc}S_{isct} \\
&\quad + \sum_{i,s,t\in\tau_T} h_{is}I_{ist} + \sum_{i,s,t\in\tau_T} \xi_i\Gamma_{ist}\Bigg]\cdots\Bigg]\Bigg]\Bigg] + \sum_i \mu_{is}J_i^{\Delta}
\end{aligned} \tag{1}$$

Subject to manufacturing constraints:

$$P_{ijst} = R_{ijs}RL_{ijst} \qquad \text{for all } i,\ j,\ s \text{ and } t \in \tau_l \tag{2}$$

$$\sum_i RL_{ijst} \leq H_{jst} \qquad \text{for all } i,\ j,\ s \text{ and } t \in \tau_l \tag{3}$$

Supply chain constraints:

$$I_{ist} = I_{is(t-1)} + \sum_j P_{ijst} - \sum_c S_{isct} \qquad \text{for all } i, s \text{ and } \tau_l \tag{4}$$

$$I_{ist_0} = I_{is}^0 \qquad \text{for all } i, s \text{ and } t_0 \in \tau_l \tag{5}$$

$$\Gamma_{ict} \geq \Gamma_{ic(t-1)} + \omega_{ict} - \sum_s S_{isct} \qquad \text{for all } i, c \text{ and } t \in \tau_l \tag{6}$$

Non-negativity constraints:

$$P_{ijst}, RL_{ijst}, S_{isct}, I_{ist}, \Gamma_{ict}, J_i^{\Delta} \geq 0 \tag{7}$$

Upper bound constraints:

$$P_{ijst} \leq R_{ijs}H_{jst}, \quad RL_{ijst} \leq H_{jst}, \quad S_{isct} \leq \sum_{t' \leq t} \omega_{ict'},$$

$$\Gamma_{ict} \leq \sum_{t' \leq t} \omega_{ict'} \tag{8}$$

$$\sum_{s,t'\leq t,l'\leq l} S_{isct'} \leq \sum_{t'\leq t,l'\leq l} \omega_{ict'} \qquad \text{for all } i, c \text{ and } t \tag{9}$$

Customer satisfaction constraints:

$$M_s(1 - Z_{ict}) \leq \Gamma_{ict} \leq M_b(1 - Z_{ict})$$
$$\text{for all } i, c \text{ and } t \in \tau_l \qquad (10)$$

$$J_i = E\left[\frac{\sum_{c,t} Z_{ict}}{N_c N_t}\right] \qquad \text{for all } i \qquad (11)$$

$$J_i + J_i^{\Delta} \geq J_i^{\text{target}} \qquad \text{for all } i \qquad (12)$$

where

Decision variables: P_{ijst}, production amount of product i on processing unit j at facility s in time period t; S_{isct}, supply of product i from facility s to sales region c in time period t.

State variables: J_i, customer satisfaction level for product i; I_{ist}, inventory level for product i at the end of time period t at facility s (I_{ist_0} is the initial inventory level for product i at facility s); Γ_{ict}, amount of shortage of product i in meeting demand from customer c at time period t; J_i^{Δ}, deviation below target customer satisfaction level for product i, J_i^{target}.

Other variables: RL_{ijst}, run length of product i on processing unit j at facility s in time period t; Z_{ict}, binary variable that takes value 1 when the realized demand ω_{ict} is met on time.

Uncertainty parameters: ω^l, set of uncertain demands up to planning period $l(\omega = \omega^T)$. ω_{ict}, uncertain demand for product i from customer c at time period t.

Cost coefficients: h_{ist}, inventory cost for holding a unit of product i in inventory facility s for the duration of time period t; c_{ijs}: unit production cost of product i on processing unit j at facility s; $\xi_{\iota\chi}$, unit penalty for inadequate supply of product i to region c; μ_i, unit penalty for missing target customer satisfaction level for product i; t_{sc}, unit transportation cost to move a unit of product from facility s to sales region c.

Other parameters: J_i^{target}, target customer satisfaction level for product i; R_{ijst}, effective rate for production i on processing unit j at facility s in time period t; H_{jst}, amount of time available for production on process j at s during time period t; I_{is}^0, initial inventory of product i at facility s; M_{b}, big constant (or maximum allowable size of an order); M_{s}, small constant (sufficiently smaller than minimum demand increment); τ_l, set of time periods that falls within planning period l; N_c, number of sales regions; N_t, number of scheduling periods; T, number of total planning periods considered in the problem.

This formulation is based on the deterministic linear programming model for supply chain planning proposed by McDonald and Karimi (1997) but extended to a multi-stage stochastic programming form. The objective function given in Eq. (1) is composed of three terms. The first term is the cost of production amount in the first planning period. The second term represents the cost of the T-stage recourse decisions involving the wait-and-see supply and inventory variables at each planning period and the here-and-now production variables of the adjacent planning period. It also includes a penalty cost for shortage of product in meeting demand which proportional to the amount of the shortage, Γ_{ict}. The nested expectations of $E_{\omega^1}[E_{\omega^2}[\ldots E_{\omega^T}[]\ldots]]$ denotes that the expectation is computed over the probability distribution of the cumulative demand, ω^l, up to each planning period l where the inner expectation is conditioned on the realization of the uncertain demand of the outer expectation. The first and the second term reflect the structure of a typical multi-stage stochastic program. The last term is the penalty for the deviation of the expected customer satisfaction levels from the target values set for each product. Note that this term effectively penalizes the cumulative frequency of failures to meet the demand whereas the Γ_{ict} terms in the objective function penalize the actual amounts. When the production capacity is limited it may be impossible to met the specified customer satisfaction level of all products, J_i^{target}. The role of slack variables J_i^{Δ} is to take on non-zero values so as to satisfy constraint (12).

Fig. 2 depicts the temporal alignment of decision variables, state variables and realizations of random demands along with the planning periods. In planning period l, the production decision variables,$P_{ijst\in\tau_l}$, are determined after the demand requirements up to $l - 1$ period have been realized but before the demand outcomes for period l and subsequent periods are known. Consequently, the decision on the production variables for period l should take into account the state at the beginning of planning period l and the possible demand outcomes in later periods. This is formalized through constraints (4 and 5) which link the decisions of two adjacent planning periods. The supply variables, $S_{isct\in\tau_l}$ take into account the demand outcomes for planning period l and serve to constrain the state variables $I_{ist\in\tau_l}$ and $\Gamma_{ict\in\tau_l}$.

If the demand distribution functions were discretized , then the evolution of random demands over time can be represented by the tree-like structure shown in Fig. 3. Starting from each node, a large number of possible demand realizations at the next planning period are expressed as branches stemming from that node. Assuming m possible next-period demand realizations at each node, the total number of scenarios will amount to m^T, where T is the total number of planning periods considered. At each period, each node is

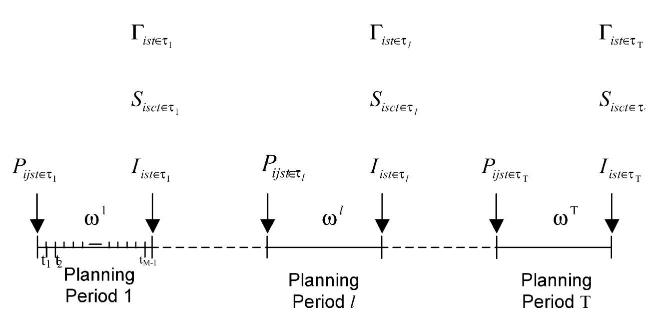


Fig. 2. Temporal alignment of planning periods and corresponding decisions for inner minimization.

associated with the realization of random demands, ω^l, the decision and state variables, $P_{ijst\in\tau_l}$, $S_{isct\in\tau_l}$, $I_{ist\in\tau_l}$, $\Gamma_{ict\in\tau_l}$.

The constraints, Eqs. (2)–(10), are generated for each demand sample path (or scenario) at each planning period in the deterministic equivalent formulation. Eq. (2) sets the production amounts and Eq. (3) constrains the production such that the capacity limit is not violated. Eq. (4) is a typical inventory balance. When the scheduling time period t is the first element of τ_l, $I_{is(t-1)}$ is determined at the parent node at Fig. 3, based on demands realized up to $l-1$ planning period, while I_{ist} is determined based on demands realized up to planning period l. For all other scheduling time periods in τ_l, $I_{is(t-1)}$ and I_{ist} are determined based on demands realized up to planning period l. Eq. (5) is the initial conditions of inventory variables. Eq. (6) involves the demand realization and calculates the backlogging amount. As $I_{is(t-1)}$ in Eq. (4), $\Gamma_{ic(t-1)}$ in Eq. (6) is dependent on demand realization up to planning period $l-1$, when the scheduling time period t is the first element in τ_l. Eq. (8) provides redundant but useful bounds for the variables. Eq. (9) insures that the backlogged demands are met later in time. Constraint (9) forces the binary variable Z_{ict} to take on the value 1 if the demand is met on time and 0, otherwise. The set of binary variables is required in order to compute the expected value of the customer satisfaction level for each product via Eq. (11). The inner summation of Z_{ict} over sales regions and time periods gives the number of demands met on time. Division of this value by the number of all demands (number of all sales regions multiplied by number of all time periods) results in an estimate of the customer satisfaction level. Thus the computation within the bracket in the equation provides an estimate of the customer satisfaction level for a given demand sample path. The expectation of these estimates gives the expected customer satisfaction level, J_i. Note that the customer satisfaction could readily be defined in terms of individual sales regions by dropping the summation over c and deleting the denominator factor N_c.

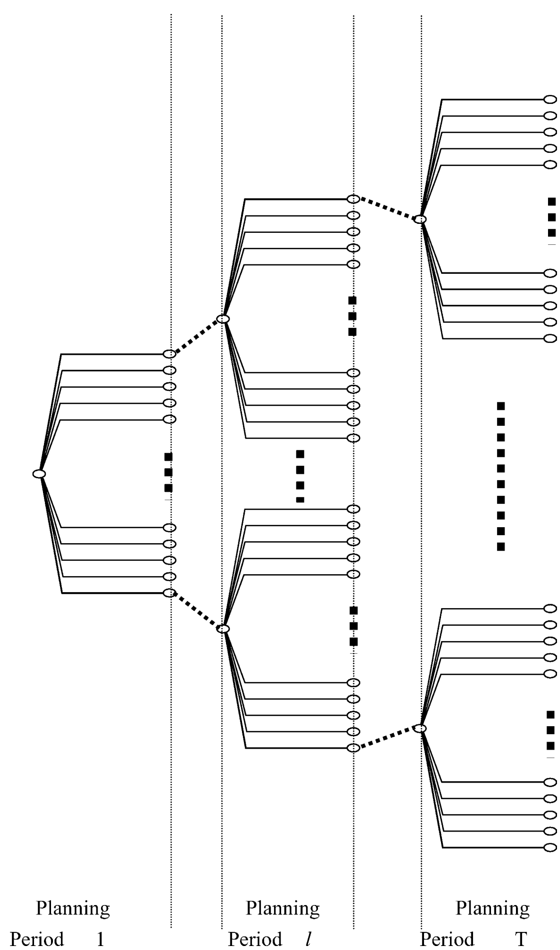


Fig. 3. Schematic diagram of evolution of sample paths.

As can be seen in Fig. 2, each planning period is subdivided into a number of scheduling time periods. The length of a time period is expected to be small enough for the solution of the model to represent a production schedule. Scheduling related constraints and variables are left out of the formulation, however, for simplicity of exposition. The formulation is also based on an assumption that the delivery from each plant site to each customer takes less time than the length of the time period. Thus, we avoid the consideration of transportation events that cover several time periods in this formulation. However, in principle such multiple-period transportation events and decisions can readily be considered by suitably modifying the formulation constraints along with the detailed scheduling constraints (which were not detailed here).

This multi-stage stochastic program insures that "here-and-now" decisions are taken using probabilistic information of uncertain future events, instead of just using their expected values. It is thus non-anticipatory in nature and provides the complete set of time dependent production and inventory variables required for execution. Under the rolling horizon strategy, once the first period has been executed, the model would be resolved and again a full set of production and inventory variables determined. Since the stochastically "best" set of these variables is generated, there is no explicit role for safety stocks. However, to achieve this non-anticipatory aspect, it is necessary to couple the "here-and-now" production decisions for each planning period with all the subsequent "wait-and-see" supply and inventory decisions and the production decisions of later planning periods up to and including the last planning period.

The direct solution of this multi-stage stochastic program using mathematical programming tools could in principle be undertaken. However, considering its scale and complexity, routine solution of the problem in rolling horizon fashion is at present impractical. For instance, the construction of a deterministic equivalent using the scenario approximation will create a problem much beyond the power of current LP or MILP solution techniques,—to say nothing of the additional dimensionality increase when the detailed scheduling variables and constraints are added to the formulation. Alternatively, Cheng, Subramanian, & Westerberg (2003) proposed the representation of multistage stochastic problems of this general type as discrete time Markov decision processes with recourse and suggested their possible solution using a dynamic programming strategy. While such a strategy is conceptually very attractive, it is limited computationally by the effects of state dimensionality and the presence of constraints which involve variables from different stages, as is the case with inventory balances.

4.2. Approximation strategy

To address this problem we propose an approximate strategy which relies on the use of deterministic supply chain planning and scheduling models employed in a rolling horizon mode. The deterministic model is built using expected values of future demands and incorporates safety stock levels for each product and site within the inventory balance constraints. The explicit safety stock level constraints reflect in an approximate matter the non-anticipatory aspects of the stochastic formulation. In order to determine the customer satisfaction level, demand scenarios extending over the entire planning horizon are generated in a Monte Carlo manner. A discrete event simulation of the supply chain is executed implementing the production and scheduling plans obtained via deterministic models for the full horizon for a given scenario. The results of multiple such simulations serve to provide an estimate of the customer satisfaction level. The safety stock level parameters are then adjusted as an outer loop optimization in which the weighted sum of the deviations from the target customer satisfaction levels are minimized. This strategy in essence employs two aspects of the "Sim-Opt" architecture, previously used to solve stochastic optimization problems involving dynamic discrete event systems. Fig. 4a represents the outer optimization on the safety stock levels while Fig. 4b represents the inner problem which involves a simulation with a series of imbedded planning/scheduling subproblems, solved in the rolling horizon mode. The set of Monte Carlo driven repetitions of the simulation with imbedded planning and scheduling optimizations essentially constitutes as function evaluation for the outer optimization loop of Fig. 4a.

We next define the two key subproblems, the outer optimization subproblem and the deterministic planning subproblem and show their connection to the stochastic problem of Eq. (1) through (12). In the subsequent section we discuss in more detail the various computational details which are needed to link these subproblems and to drive the computations to convergence on the desired safety stock levels.

4.3. Outer optimization subproblem

Objective function:

$$\mathrm{Min}_{\theta} f(\theta) = \sum_{i} \mu_{is} J_i^{\Delta} \tag{13}$$

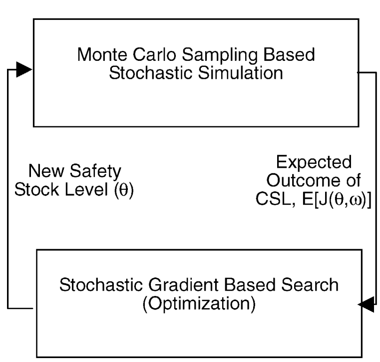
Subject to:

$$J_i + J_i^{\Delta} \geq J_i^{\text{target}} \quad \text{and} \quad J_i^{\Delta} \quad \text{for all } i \tag{14}$$

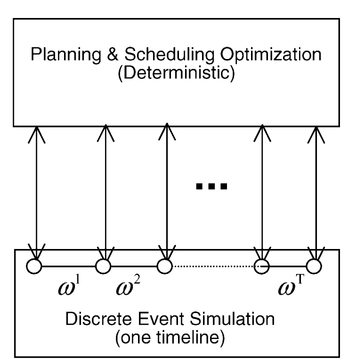
The objective function of the outer optimization, denoted by Eq. (13) involves a stochastic optimization in which the weighted deviations from the desired customer satisfaction level is minimized by choosing the safety stock levels θ, while ensuring the target for customer satisfaction J_i^{target} is met as closely as possible. In constraints (14), both J_i and J_i^{Δ} are functions of the safety stock level θ and are evaluated by using the results of multiple Monte-Carlo simulations with imbedded planning and scheduling problem solutions.

4.4. Deterministic planning model

The deterministic supply chain planning and scheduling models required for execution within the Monte-Carlo simulations are extracted directly from the original stochastic program formulation. In the following, only the objective function and the constraints of the deterministic planning model are shown: the details corresponding to the scheduling model are not presented for simplicity, as in the case of the stochastic formulation.



(a) Simulation Optimization



(b) Simulation with Embedded Optimization

Fig. 4. Configuration of simulation and optimization procedures.

Objective function:

$$\text{Min} \sum_{i,j,s,t} c_{ijs} P_{ijst} + \sum_{i,s,c,t_1} t_{sc} S_{isct} + \sum_{i,s,t} h_{is} I_{ist} + \sum_{i,s,t} \varsigma_{is} I^{\Delta}_{ist} + \sum_{i,c,t} \xi_i \Gamma_{ict} \quad (15)$$

Subject to manufacturing constraints:

$$P_{ijst} = R_{ijs} RL_{ijst} \qquad \text{for all } i,\ j,\ s \text{ and } t \quad (16)$$

$$\sum_i RL_{ijst} \le H_{jst} \qquad \text{for all } i,\ j,\ s \text{ and } t \quad (17)$$

Supply chain constraints:

$$I_{ist} = I_{is(t-1)} + \sum_j P_{ijst} - \sum_c S_{isct} \qquad \text{for all } i,\ s \text{ and } t \quad (18)$$

$$\Gamma'_{ict} \ge \Gamma'_{ic(t-1)} + E[\omega_{ict}] - \sum_s S'_{isct} \quad \text{for all } i,\ c \text{ and } t \quad (19)$$

$$I^{\Delta}_{ist} \ge \theta_{is} - I_{ist} \qquad \text{for all } i,\ s \text{ and } t \quad (20)$$

Lower bound constraints:

$$P_{ijst}, RL_{ijst}, S_{isct}, I_{ist}, \Gamma_{ict}, I^{\Delta}_{ist} \ge 0 \quad (21)$$

Upper bound constraints:

$$P_{ijst} \le R_{ijs} H_{jst}, \quad RL_{ijst} \le H_{jst}, \quad S_{isct} \le \sum_{t' \le t} E[\omega_{ict'}], \quad \Gamma_{ict} \le \sum_{t' \le t} E[\omega_{ict'}] \quad (22)$$

$$\sum_{s,t' \le t, l' \le l} S_{isct'} \le \sum_{t' \le t, l' \le l} E[\omega_{ict'}], \qquad \text{for all } i,\ c \text{ and } t \quad (23)$$

The notation for all the variables and constraints are consistent with those used in the formulation (1) through (12). Compared to the objective function (1), the variable J^{Δ}_i is not included in the objective function (15) of this deterministic model. In the deterministic planning model, the safety stock requirement is explicitly imposed through constraint (20) and violation of the safety stock is penalized by the term $\sum_{ist} \varsigma_{is} I^{\Delta}_{ist}$. The variable I^{Δ}_{ist} represents the amount of the inventory level in production facility s at time t which dips below the safety stock level θ_{is} determined by the outer optimization. Parameter ς_{is} is the unit penalty cost for this deviation from the safety stock level. It should be noted that the penalty term involving the variables I^{Δ}_{ist} does not constitute a penalty for failure to meet customer demands, rather it principally enforces the maintenance of safety stock levels. By virtue of the term $\sum_{ist} \xi_i \Gamma_{ist}$in the objective function (15), the optimum values of the production amounts P_{ijst} will include the production amounts that maximize meeting the expected values of the customer demands, as well as meeting the safety stock levels imposed by constraint (20).

It should be noted that the constraints for measuring the customer satisfaction levels, that is, Eqs. (10) and (11), do not appear in this deterministic model. This is due to the fact that in the deterministic planning model the value of the customer satisfaction levels which is by definition an expected value over realizations of a distribution of demands can not be defined. Instead, the customer satisfaction levels are determined as the result of multiple Monte Carlo simulations involving imbedded solutions of the planning model.It is evident that the proposed strategy constitutes a decomposition approach in that the objective function terms and several of the constraints are divided between the two subproblems. Moreover, the use of the safety stock level definitely introduces an additional approximation. Thus, it should be clear that the solutions obtained will be suboptimal relative to the full multi-stage stochastic formulation. Our claim is that it provides reasonable solutions using building blocks readily implementable in practice.

5. Computational framework

In this section, we describe the overall computational approach. We first summarize the computational logic for each decomposed problem and then detail the integration of these two computational elements.

5.1. Outer optimization

The outer optimization formulation, consisting of Eqs. (13) and (14), uses the customer satisfaction levels generated in the inner loop to generate improved values of the safety stock levels. From the discussion associated with Fig. 1, it can be expected that the expected customer satisfaction level is proportional to the safety stock level or more generally a monotonic function of θ. Thus, in regime I, the outer loop optimization would simply involve driving the safety stock levels until all customer satisfaction level targets are met, that is, until objective function (13) is driven to zero. In regime II, however, arbitrarily set levels of customer satisfaction levels may not be achievable. In this case, the outer loop should minimize the weighted deviations from the target values. This can readily be achieved via the following gradient-based stochastic search scheme:

Step 1: start with arbitrary set of safety stock levels, θ^n_{is}, where $n = 0$, for all i and s.

Step 2: run a sufficient number of Monte-Carlo sampling based simulations with their imbedded planning/scheduling optimizations to obtain a reliable estimate of the expected customer satisfaction level, $J_i(\theta^n_{is})$ and deviation $(J^{\Delta}_i)^n$.

Step 3: check for convergence, if $|(f(\theta^n) - (f(\theta^{n-1})| < \varepsilon_1$, for all i stop.

Otherwise, continue.

Step 4: calculate a new safety stock level using Eq. (24).

$$\theta_{is}^{n+1} = \theta_{is}^{n} + \alpha\beta_{is}\{\mu_i J_i^{\Delta}\} \tag{24}$$

Step 5: if $|\theta^{n+1} - \theta^{n}| < \varepsilon$, stop.
Otherwise, set $n = n + 1$ and go to step 2.

In Eq. (24), θ^{n+1} and θ^{n} are the vectors of safety stock levels in the $(n + 1)$th and nth stage of outer loop iterations, respectively. Parameter α is the step size factor, β_{is} is the distribution factor, which represents the ratio of product supply from each production sites and is μ_i is the weighting factor for product i in objective function (13). The value of α can be obtained experimentally or it can be recursively updated using the following equation.

$$\alpha_n = \frac{1}{\beta}\frac{\theta^{n} - \theta^{n-1}}{f(\theta^{n}) - f(\theta^{n-1})} \tag{25}$$

The distribution factor is of course optional and is used in the case study presented later in this paper by virtue of how the product demand forecasts are generated.

5.2. Simulation with embedded optimization

The computations inside the outer loop optimization, involve the repeated simulation of the supply chain operation over the planning horizon, each with a given Monte-Carlo sample of the demands. Within each such simulation, a series of planning and scheduling optimization problems are solved under the rolling horizon scheme. Each such complete simulation is called a timeline. The following summarizes the procedure for executing a timeline.

Step 1: run the deterministic planning and scheduling model with given state (in the first stage, deterministic initial inventory is used) to obtain the production decision, P_{ijst}, for stage l, (at the first iteration, $l = 1$).
Step 2: run the discrete event simulation for planning period l with demand outcomes for the planning period, ω^{l}.
Step 3: record the state (I_{ist}, Γ_{ict}, I_{ict}^{Δ}, WIP and amounts in transportation) at the end of planning period l from the simulation.
Step 4: set $l = l + 1$ and go to step 1 until $l = T$.
Step 5: estimate and record customer satisfaction estimate, $J_i(\theta, \omega)$.

By repeating the above procedure for a sufficient number of Monte Carlo samples, we can collect the performance results necessary to compute the expected value, $J_i(\theta) = E[J_i(\theta, \omega)]$. Note that the accuracy of the expected value will be governed by the number of replications of timelines.

5.3. Simulation based optimization framework

Fig. 5 presents a more detailed flow diagram which summarizes the overall computational strategy that combines the inner and outer loop components. At the beginning of the

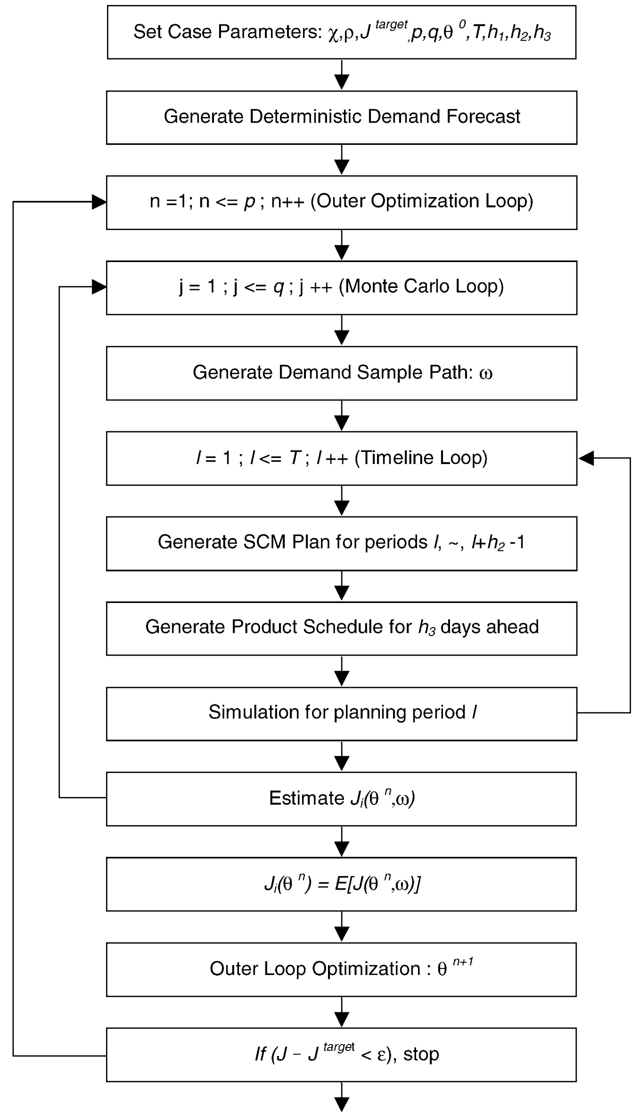


Fig. 5. Computational flow diagram.

computation, the case parameters, such as the coefficient of variation of the demand; χ, the demand to capacity ratio; ρ, target customer satisfaction level; J^{target}, initial safety stock level; θ^{0}, the number of iterations for the outer loop; p, and the Monte Carlo scheme; q, number of planning periods; T, and the lengths of the simulation time horizon; h_1 (days), the planning horizon, h_2 (months), and the scheduling horizon, h_3 (days) are set. Then the deterministic demand forecast is generated based on ρ. Before the simulation for one timeline starts, a sample path is generated using the deterministic demand forecast and χ. The simulation is built on an implicit loop over the embedded optimization. The number of the iteration on the implicit timeline loop is given as T. The interactions between the planning, scheduling and simulation are discussed in more detail at Section 6. Once the computations for a simulation timeline are completed, the customer satisfaction level, $J_i(\theta^{n}, \omega)$, is estimated. By iterating over the Monte-Carlo loop, the expected customer satisfaction level, $J_i(\theta^n) = E[J_i(\theta^n, \omega)]$ is obtained. Then the outer loop iteration is performed until the sum of the weighted deviations from the target customer satisfaction levels converges within the tolerance range specified.

5.4. Computational refinements

In the Monte-Carlo simulation, a sufficiently large number of simulation results, q in Fig. 5, are required to obtain a statistically significant estimate of the performance measure, $J_i(\theta^n)$. To increase the accuracy of that estimate, a computation load must be incurred which increases proportionally with the increase in q. To mitigate this computational load, we propose the use of antithetic variates. The antithetic variates technique is one of the commonly used variance reduction techniques in Monte-Carlo simulations studies. This technique builds on the assumption that the input random variable has positive or negative correlation with the output random variable. Thus, forcing the negative correlation between the outputs of a pair of simulation runs by using two inputs that are in opposite position from the mean of the input, the average of the outputs is anticipated to get closer to the actual value of the performance measure. For detailed discussion of the antithetic variates approach, the reader is referred to Law and Kelton (2000).

5.5. Discussion

As noted earlier, the proposed decomposition strategy can only provide suboptimal solutions. Particularly, the separation of the objective function of the underlying multistage stochastic program into the penalty portion which comprises the objective function for the outer loop and the remaining cost terms which are handled within the planning/scheduling optimizations is a key simplification. It is reasonable to question whether such separation does not force instability in the iterative scheme. While we can not give formal proof that this is not the case, we can offer a plausible argument. In the inner time line optimizations, a balance is achieved between production, transportation, inventory holding, safety stock shortage, and lost customer demands. We argue first that each of these components is proportional to the safety stock level. For instance, the inventory holding cost will be proportional to safety stock level as safety stock forces a net increase in the average inventory over time. The production cost will also increase with safety stock level as higher safety stock levels will lead to increased production to sustain such levels. While transport cost is not affected by safety stock levels, the safety stock shortage cost is likely to be a minor component at the optimal solution as it is solely intended to enforce the definition of safety stock. Finally, in the regime I and II domains which are the focus of this research, the amount of customer demand not met is expected to be very small and thus that term should be negligible. Thus, we conclude that the expected supply chain operational cost, as defined in the planning problem, will be roughly proportional to safety stock level. Since customer satisfaction is monotonic in safety stock level, the net effect of the outer optimization loop will be to drive down the safety stock levels and thus concurrently to drive down the supply chain cost as well. As a result, as the interaction progress we can expect to see a progressive general reduction in the supply chain cost.

6. Implementation of the computational framework

The computational framework for implementing the proposed strategy consists of the following components: the computation control module, the three models for planning, scheduling and simulation, and the static and dynamic databases, as shown in Fig. 6. The computational control module provides centralized computation management and control of interactions with the other computational components. Its various functional roles include:

- Issuing commands for retrieving/recording in the databases.
- Generating/updating and releasing commands for executing the planning, scheduling and simulation models.
- Generation of deterministic demand forecasts and demand scenarios.
- Re-processing the outputs of planning and scheduling models as inputs to scheduling and simulation models, respectively.
- Reporting the computation results.

The Static Database contains the information defining the case problem. The actual data is saved in a Microsoft Excel Workbook file. Within the file, the data is categorized based on their relevance and saved in its corresponding spreadsheets. The categories of the static data are as follows:

- SCM structure related: number of product types, number of grades, number of packages, number of processing units in each production facility, production capacities.
- Time related: length of simulation horizon, length of the horizon and of each period in the planning and scheduling models.

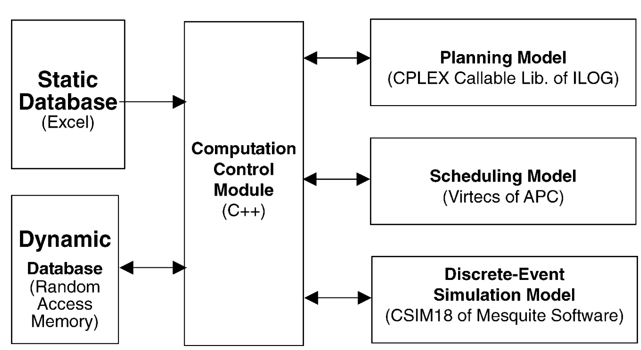


Fig. 6. Computational components and their interactions.

- Demand related: demand proportion of each product type, package ratio, proportion of regional demands, parameters for demand patterns of each product.
- Computation related: number of iterations for the outer optimization loop and the Monte-Carlo loop, step size factor, convergence tolerance.
- Case problem related: target customer satisfaction levels for each product, coefficient of the variation of demand, demand to capacity ratio.
- Others: inventory cost, backlogging cost, safety stock shortage cost, transportation cost, transportation time, initial safety stock levels.

Since the Computation Control Module is coded in C++, a publicly available class library is used to access each cell in the Excel sheets.

The Dynamic Database records the computational results and provides the three models with updated input data. This database is kept in random access memory by defining each data element as an object in the C++ program. The following summarizes the information kept in this database:

- Current time related indices: current iteration number of the outer loop, current iteration number of the Monte Carlo loop, current simulation clock (date and time of the current point at a simulation).
- State variables: on-hand inventory levels, WIP at each processing units at each plant, demands backlogged and products in transit, which are updated at each scheduling time period as the simulation proceeds.
- Demand data: the deterministic demand forecast and the demand scenarios.

The planning model is formulated as an LP based on McDonald and Karimi (1997) and its details can be readily derived from Eqs. (15)–(23). The planning model is generated/updated by CPLEX Callable Library functions used in the computation control module. For instance, in the case study described in Section 7, the horizon of the planning model is 3 months and at the beginning of each month, the planning model is updated in a rolling horizon manner. When the model is updated as the horizon rolls forward, the cost coefficients of the formulation remain unchanged, because they are independent of time. However, some of the parameters in the right hand sides of the constraints must be re-calculated to incorporate the current values of the state variables into the planning model. The WIP within the production facilities, which is recorded at the point of triggering of a new planning update, are added to the plant inventory when the simulation resumes. To incorporate this future addition in the inventory into the planning model, the WIP is added to the initial inventory, I_{is1}, in the planning model. Thus the initial inventory of the planning model will be the sum of the WIP and the inventory level (obtained from the simulation). The amount of product in-transit from each facility to each sales region that is recorded at the same time point is deducted from the total demand for the first period of the model. Thus the demand for each product from each sales region in the first period of the planning model, d_{ic1}, consists of the deterministic demand forecasts; with due dates the first period, minus the amounts in-transit. The incorporation of the state variables into the planning model update as described above assumes that the length of the planning period is significantly larger than the production and transportation lead times.

For the scheduling model, a number of different approaches have been tested and on that basis the VirtECS scheduling software (Advanced Process Combinatorics Inc., 2004) was chosen for formulating and solving the scheduling subproblems. A model for a 15-day schedule of two production facilities with a continuous time MILP formulation based on the model suggested by Karimi and McDonald (1997) entails approximately 156,000 constraints and 200,000 variables, among which 2400 variables are integer, even when product change-over cost is not considered. A feasible solution within 5% of the relaxed LP bound could be obtained with CPLEX in about 5 cpu/min on a 1 GHz Intel Pentium machine, for the case problem discussed in Section 7. Given that the Sim-Opt approach requires a large number of simulations and accordingly a larger number of scheduling executions, the MILP scheduling approach imposes too heavy of a computational load. By contrast, the VirtECS software provides 40-day feasible schedule for one plant in about 30 s.

The computation control module generates the XML-based input file for the VirtECS scheduling model. As in the update of the planning model, the state variables such as inventory levels, WIP and the amounts in-transit obtained from the simulation are incorporated accordingly when the scheduling model is triggered. In addition, the difference between the initial inventory level of the scheduling model and the safety stock target is added to the demand of the first date of the scheduling model as a way to implement the safety stock levels.

The discrete event simulation model is constructed with CSIM18 (Mesquite Software Inc., 2004), a process-oriented, general purpose simulation toolkit written with general C/C++ language functions. CSIM18 provides specific classes (simulation components) which include processes, facilities and events. In brief, a CSIM process is an independent thread of execution. The active entities of a simulation model are implemented as processes. Since many processes can be executing simultaneously, processes can mimic the activities of entities operating in parallel. Facilities are resources that are typically used by processes in the model. Facilities may have one or more servers. Processes queue up for access to a server. Events are used to synchronize actions of different processes. The implementation of this type of discrete event simulation model is complicated by the fact that the model must incorporate not only the representation of the active entities but also of various events and logic associated with these events. The reader unfamiliar with these typical features of discrete event simulation is directed to

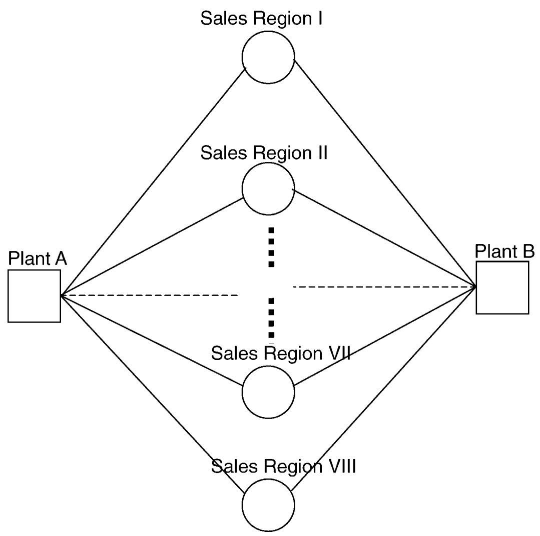


Fig. 7. Topology of the supply chain case problem.

Appendix A in which a simplified version of the simulation required for the case discussed in Section 7 is presented.

7. Case study

To demonstrate the performance of the proposed computational framework in solving realistic problems, a case study has been developed based on an application provided by a major US polyethylene producer.

7.1. Case problem

7.1.1. Products

It is assumed that the products are grouped into five different product types based on the proximity of the physical or chemical properties of the products. Each type is denoted by A, B, C, D, and E. Each type is assumed to have 10 different grades (grades 0–9) and each grade can be product in two different package types, bag and box. Thus there are a hundred final products.

Table 1
Production rates and yearly production capacities at each process unit

Process unit description	Production rates[a]	Yearly capacity[b]
Reactor A-1	102,740	900
Reactor A-2	68,493	600
Reactor B-1	102,740	900
Extruder A-1	102,740	900
Extruder A-2	68,493	600
Extruder B-1	102,740	900
Package Machine A-1	85,617	750
Package Machine A-2	85,617	750
Package Machine B-1	51,370	450
Package Machine B-2	51,370	450

[a] Unit: lbs/h.
[b] Unit: 1,000,000 lbs/year.

7.1.2. Production facilities

As can be seen in Fig. 7, two different production sites are available, plant A and plant B. Each plant has a different production layout and capacity. Fig. 8 shows the layout of the processing units. A serial production line is composed of a reactor, an extruder and one or more packaging machines. Plant A has two different serial production lines. In the first production line, product types A, B and C can be produced while in the second only types D and E are produced. However, both production trains share the packaging machines. Plant B has only one production line, which produces types A, B, and D. The total production capacity is assumed to be 2.4 billion lbs per year. Table 1 gives the production rates and yearly total production capacities of each processing unit. Table 2 provides the demand ratio (%) of each product type, where the demand ratio is defined to be the ratio of demand for each product type over the total demand for a year. The demand is distributed over types A, B, D, C and E in decreasing order. The table also gives the design production capacity assignment for each type. The assignments given indicate that the demand for type A and B is equally covered by production line A-1 and B-1. Type D is produced more in line A-2 than in line B-1, while the entire demand for type C and E is covered by the plant A lines.

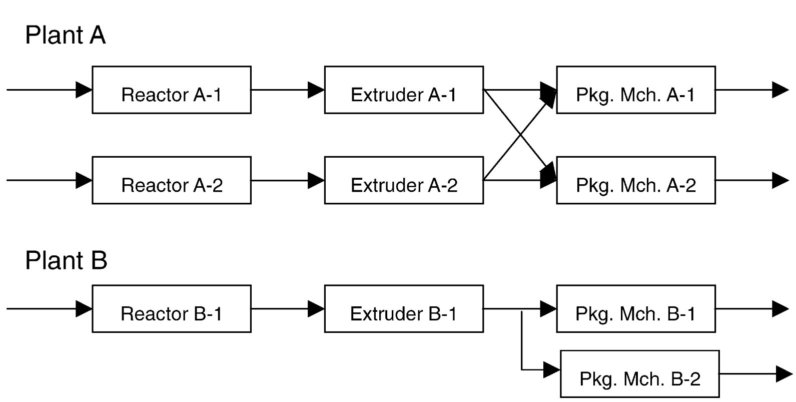


Fig. 8. Simplified process flow diagram of the production facilities.

Table 2
The demand ratio of each product type and the design capacity assignment for each product type at each reactor train

Product types	Demand ratio (%)	Design capacity assignment (%)		
		Production line A-1	Production line A-2	Production line B-1
A	31.25	42.00	–	42.00
B	25.00	33.00	–	33.00
C	9.38	25.00	–	–
D	24.38	–	60.00	25.00
E	10.00	–	40.00	–

Table 3
Demand proportion of each sales region

Sales region	Regional demand proportion (%)
I	23.11
II	10.56
III	31.82
IV	4.47
V	3.82
VI	12.48
VII	11.64
VIII	2.10

7.1.3. Sales regions

There are eight sales regions (sales regions I–VIII) which group the actual customers based on their regional proximity. Table 3 gives the proportions of total demand originating from each sales region. Table 4 provides the transportation time from each plant to each sales region. Transportation time is assumed to be linearly proportional to the distance. Generally, the transportation decision about which plant is to supply which customer is taken based on the distance/cost. However, when production capacity is constrained or the product on demand is unavailable in the nearer plant, then the other plant can supply the sales region.

Table 4
Transportation time

Sales region	Transportation time (h)		Transportation cost ($/lb)	
	Plant A	Plant B	Plant A	Plant B
I	22	47	0.04	0.09
II	11	38	0.02	0.07
III	21	35	0.04	0.06
IV	14	32	0.03	0.06
V	31	29	0.06	0.05
VI	22	26	0.04	0.05
VII	43	11	0.08	0.02
VIII	48	19	0.09	0.03

7.1.4. Demand

A demand generator is used to generate deterministic demand forecasts on a weekly basis for each type-grade-package-sales region combination. This deterministic demand is used directly in the planning and scheduling model. It is also used as the mean of the normally distributed random demand for the generation of demand scenarios for the discrete event simulation. For normally distributed demands, the standard deviation is given by the multiplication of χ (coefficient of variation of demand given as a case problem parameter) by the deterministic demand. Fig. 9 shows a sample of the demand scenario while varying the coefficient of variation of the demand. Increasing values of the coefficient, result in increasing fluctuations in the demand. Further details about the model used for deterministic demand generation are available as supplementary information to this paper. It should be noted that this particular form of demand generator is used mainly as a reasonably compact model for use in this and future case studies. In practice the generator would be constructed to fit historical product demand data.

Tables of costs used in the case study can be found in Appendix B.

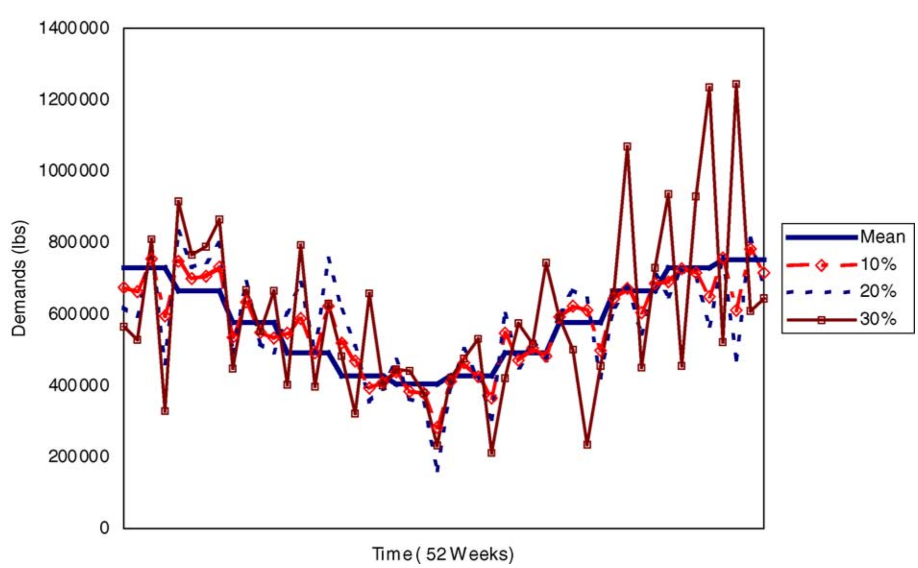


Fig. 9. Sample realizations for demand of product SR I-A0-Bag at different coefficient of variations.

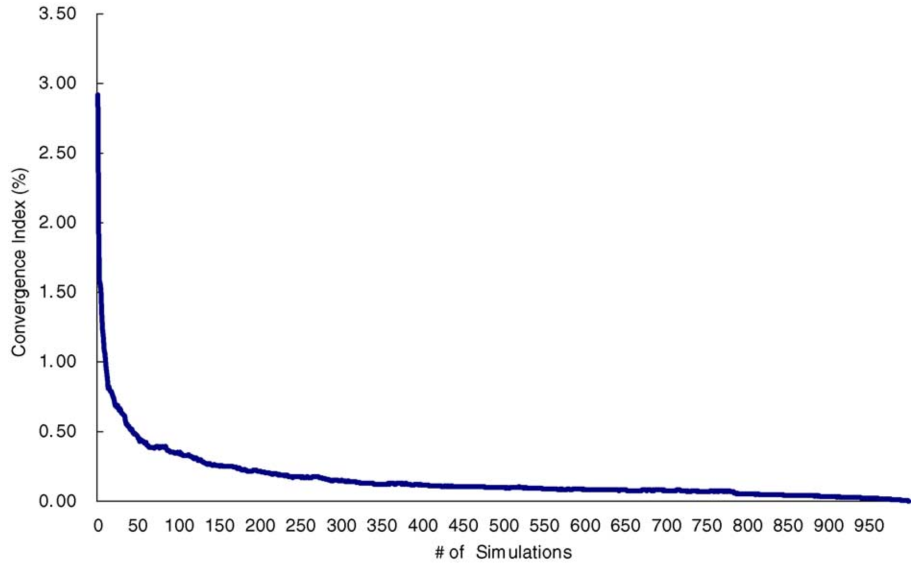


Fig. 10. Convergence of expected value of customer satisfaction level.

7.2. Results

Fig. 10 shows the averaged convergence of the sample mean of the customer satisfaction levels to the sample mean obtained with 1000 simulations. In the test, the coefficient of variation of demand is set at 30% and the demand to capacity ratio is set to 80%. The convergence index, C_i, defined in Eq. (26) is the averaged percent deviation of the sample means obtained after the *i*th simulation from the sample mean obtained after 1000 simulations.

$$C_i = \frac{\sum_{n=1}^{N} \left| S_{i,n} - S_{1000,n} \right| / S_{1000,n} \times 100}{N} \tag{26}$$

where $S_{i,n}$ is the sample mean of the customer satisfaction level for product n after *i*th simulation. N is the total number of products, (100, in this case problem). The lower the value of the convergence index, C_i, with smaller value i, the better the system convergence and thus the smaller the number of simulations needed to obtain statistically significant performance estimates. In the figure, after 50 simulations, the convergence index falls below 0.5%. Thus we can conclude that a relatively small number of simulations is sufficient to obtain statistically significant performance estimates.

Fig. 11 shows a comparison of the convergence of two cases. In the first case the antithetic variates technique is used in the Monte-Carlo simulation while the second case is the pure Monte-Carlo simulation. The antithetic variables technique is applied to the set of demands constituting a sample path. For 0.5% convergence, about 50 simulations were required for the pure Monte-Carlo case and about 20 simulations were required for the case when antithetic variates technique is used. The difference in the number of simula-

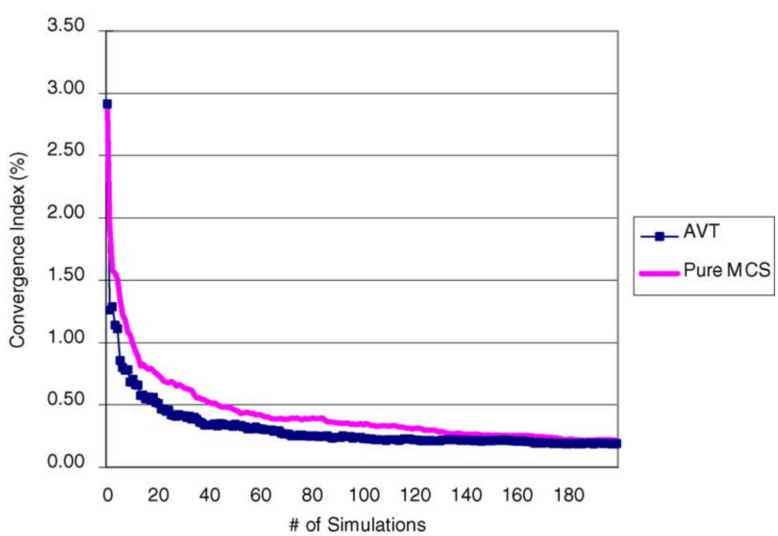


Fig. 11. Comparison of purely Monte-Carlo and Monte-Carlo with antithetic variate.

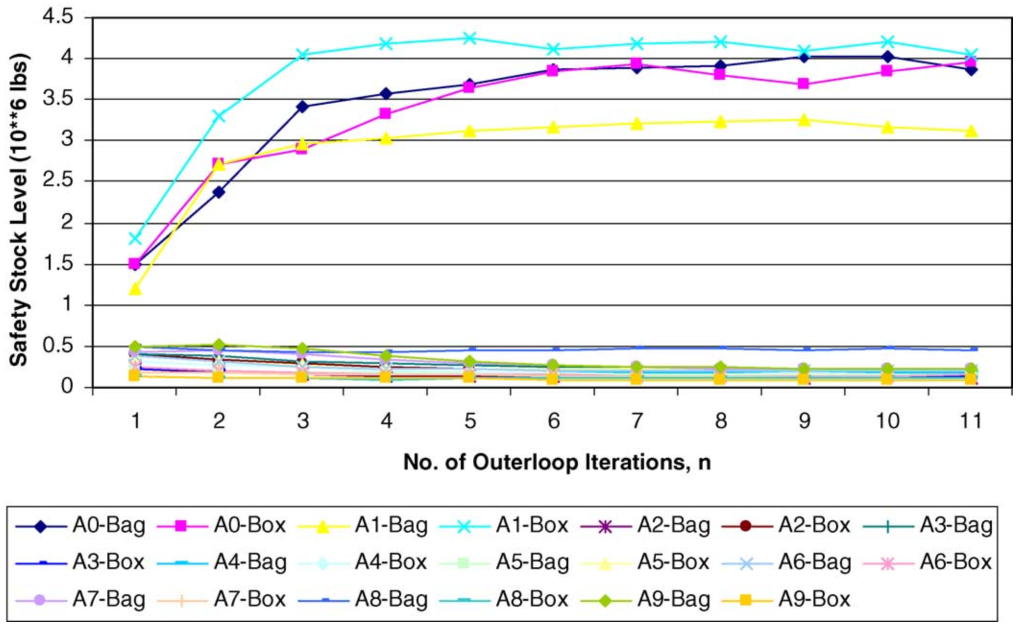


Fig. 12. Trajectory of safety stock levels of A type products at plant A as outer loop optimization proceeds.

tions required in each case increases as the convergence index decreases. Thus, we can conclude that the use of the antithetic variates technique can reduce the computation load quite effectively.

Next, the case problem was solved with the proposed decomposition strategy. The demand to capacity ratio is set to 80% and the coefficient of variation of demand is given 30%. The number of simulations within each Monte-Carlo

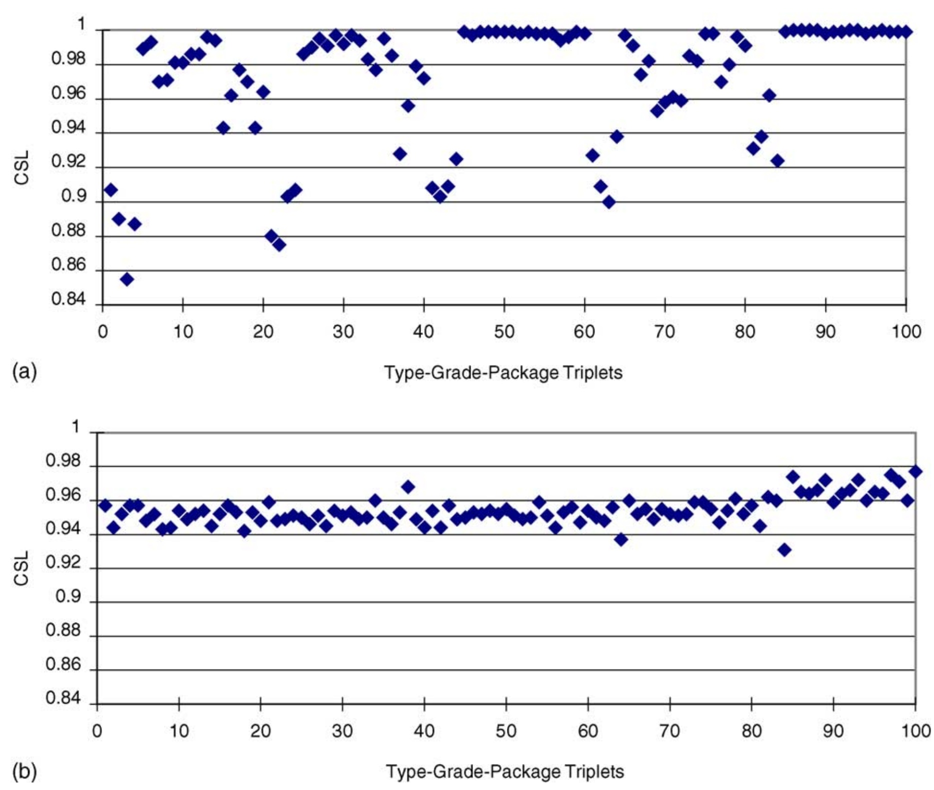


Fig. 13. (a) Sample means of customer satisfaction levels for each triplet before the first outer loop optimization. (b) Sample means of customer satisfaction levels for each triplet after the ninth outer loop optimization.

campaign is set to one hundred and the number of outer loop iterations is set to 10. The computation for the complete case was run on a 1.0 GHz Intel CPU and took about one hundred hours.

Fig. 12 illustrates the trajectory of the safety stock levels of type A as the outer loop iterations progressed. A1-bag denotes the final product of grade 1 in type A in bag package. The four products, A1-bag, A1-box, A2-bag and A2-box, make up 80% of the demand for type A and the other products, from A3-bag to A10-box, make up 20% of the demand. Likewise, for all the other types, the demand generation procedure is designed to have the account for 80% of the demand and rest of the 16 triplets (minor products) to share the remaining 20% of the demand assigned to each type. As might be expected, those products with larger demand need higher safety stock level and those with smaller demand need much lower safety stock level. This result reflects the fact that for larger demands, it is harder to cope with abrupt demand changes and for smaller demand products, it is relatively easy to make up the demand from safety stocks.

In Fig. 13a and b, the sample means of the customer satisfaction levels for each product after the ninth outer loop iteration, $E[J(\theta^9, \omega)]$, is compared with the corresponding initial values, $E[J(\theta^0, \omega)]$. The individual dots in the figure represent the sample mean of the customer satisfaction levels for each product. The elements in the vector $E[J(\theta^0, \omega)]$ in Fig. 13a are scattered over wide range of customer satisfaction levels, while those in the vector in Fig. 13b have

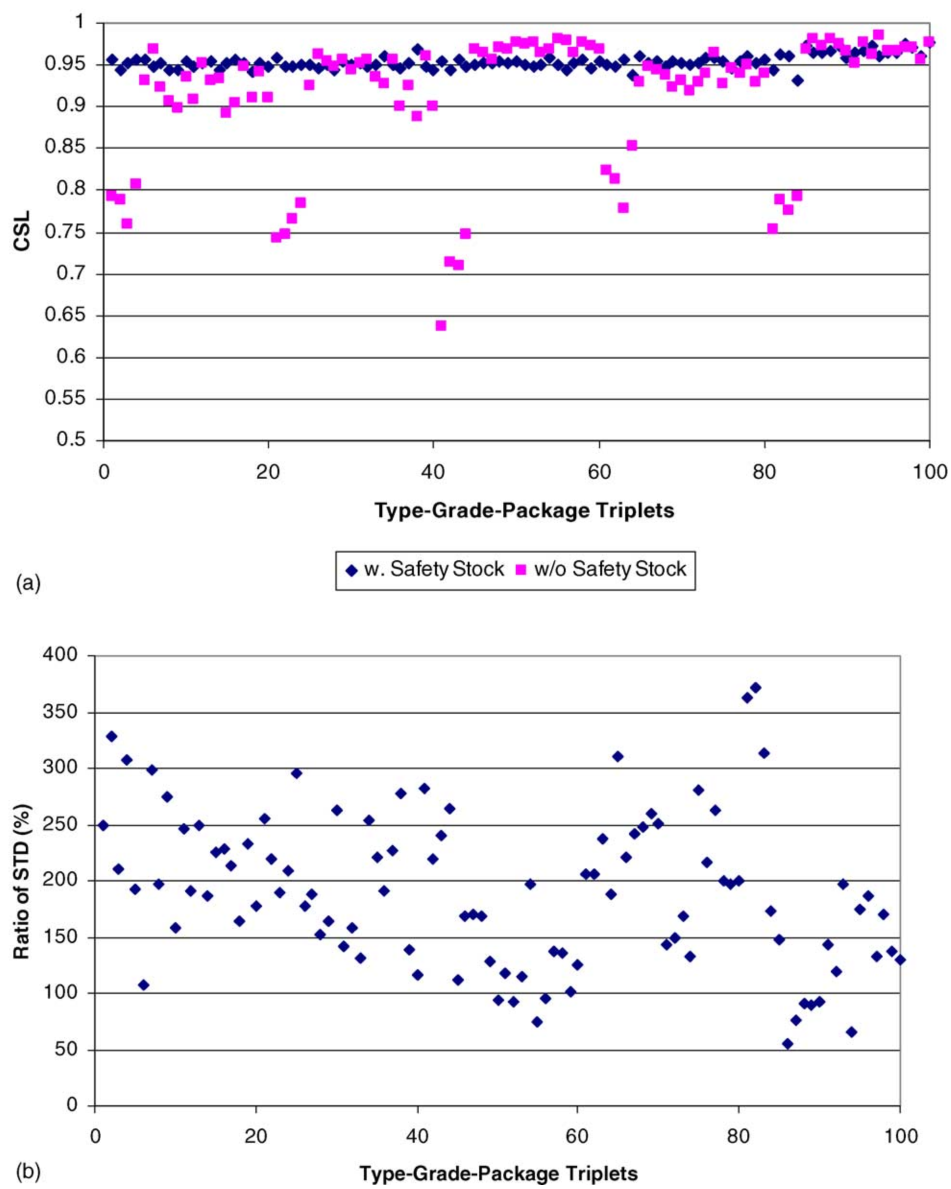


Fig. 14. (a) Comparison of expected customer satisfaction levels of the cases with and without safety stocks. (b) Ratios of standard deviation of customer satisfaction levels of the case without safety stocks to those of the case with safety stocks.

Table 5
Summary of safety stock levels and mean and standard deviation of customer satisfaction levels after each outer optimization

Number of outer optimizations	Total SSL	Mean of CSL	S.D. of CSL
1	0.53	0.97	0.036
2	0.57	0.97	0.027
3	0.59	0.97	0.021
4	0.59	0.96	0.017
5	0.59	0.96	0.014
6	0.59	0.96	0.012
7	0.58	0.96	0.011
8	0.58	0.96	0.010
9	0.58	0.96	0.009
10	0.57	0.95	0.008

converged near to the target value, 0.95. These results illustrate the effectiveness of the computational framework in optimizing the large number of safety stock levels from poor estimates in relatively few outer loop iterations.

Table 5 summarizes the changes in the total safety stock level (SSL) in the supply chain and the mean and the standard deviation of the customer satisfaction levels (CSL) averaged over all of the type-grade-package triplets. The total safety stock level is reported as the ratio of the total of the safety stock levels of all the products to the monthly total demand of all products. The total safety stock level resulting from the outer loop optimization corresponds to about 17 days of demand. The standard deviation of the customer satisfaction level decreases dramatically as the number of outer optimization increases, while the mean of the customer satisfaction level shows relatively small changes.

Finally, Fig. 14a and b compare the expectation and the standard deviation of the customer satisfaction level of each type-grade-package triplet when no safety stock is held within the supply chain, with those when the safety stock levels obtained by the full computation (nine outer optimizations) are used in the supply chain. The diamonds in each figure stand for the result obtained by the full computation just as is shown in Fig. 13b and the rectangles represent the results from the zero safety stock case. From Fig. 14a, the expected customer satisfaction levels with zero safety stock are lower for most of the type-grade-package triplets. It is noteworthy that the expected customer satisfaction levels of the major products (the first four type-grade-package triplets that take up 80% of the demand) in each type are significantly lower when no safety stock is held in the supply chain. This is due to the fact that without safety stocks the limited production capacity does not provide enough flexibility for the major products to cope with their uncertain demand. By contrast, the safety stock levels needed for the minor products to meet the target customer satisfaction levels are so small that they could achieve quite high customer satisfaction levels even without any safety stock level. Some of the products of types *C* (indexed as 41–60) and *E* (indexed as 81–100) show even higher customer satisfaction levels in the zero safety stock case than those obtained with safety stocks. The reason for this anomalous result can be attributed to the fact that as the safety stock levels of the major products are increased, the production of the minor products becomes more constrained and thus some demands can not be met in time. Fig. 14b presents the ratio of standard deviation of customer satisfaction level of each product in the zero safety stock case to that of the case with safety stocks. For most of the products, as can be expected, the standard deviations decrease significantly if safety stock is kept at its optimal level. However, the ratios of some of the minor products of type *C* and *D* decrease when safety stocks are present. This too is attributable to the constraints placed on the production capacity as the safety stock levels of major products increase.

8. Conclusion

In this work a computational framework was proposed for determining the safety stock levels for planning and scheduling applications involving realistically scaled supply chains which have the various complex characteristics arising in typical chemical process companies. These safety stock values allow the use of deterministic planning and scheduling tools in a rolling horizon manner for applications in which demand uncertainties would otherwise make such models less effective. While the approach has been demonstrated using demands as the uncertain parameters, the Monte-Carlo based timeline technique can be employed in the presence of other uncertain parameters, such as delivery times, production delays, and even change-orders. The only requirement is that empirical or fitted distribution functions be available for sampling purposes.

The key limitation of the overall approach lies in the large computing times required to address problems of increasing scope. While the long times required by the case reported are not prohibitive for periodic updating of safety stock level targets, these times are too long to allow the approach to be used to support various tactical analyses that are enabled by the abundant information provided by the simulations. Variance reduction techniques such as common random numbers and control variates and the parallelization of the computations by distributing the individual simulation for each sample path over different machines are viable means of achieving further computational time savings.

Appendix A. Implementation of the discrete event simulation

To illustrate the logic of the simulation model constructed for the case study of Section 7, consider a supply chain composed of one production facility, which has three processing

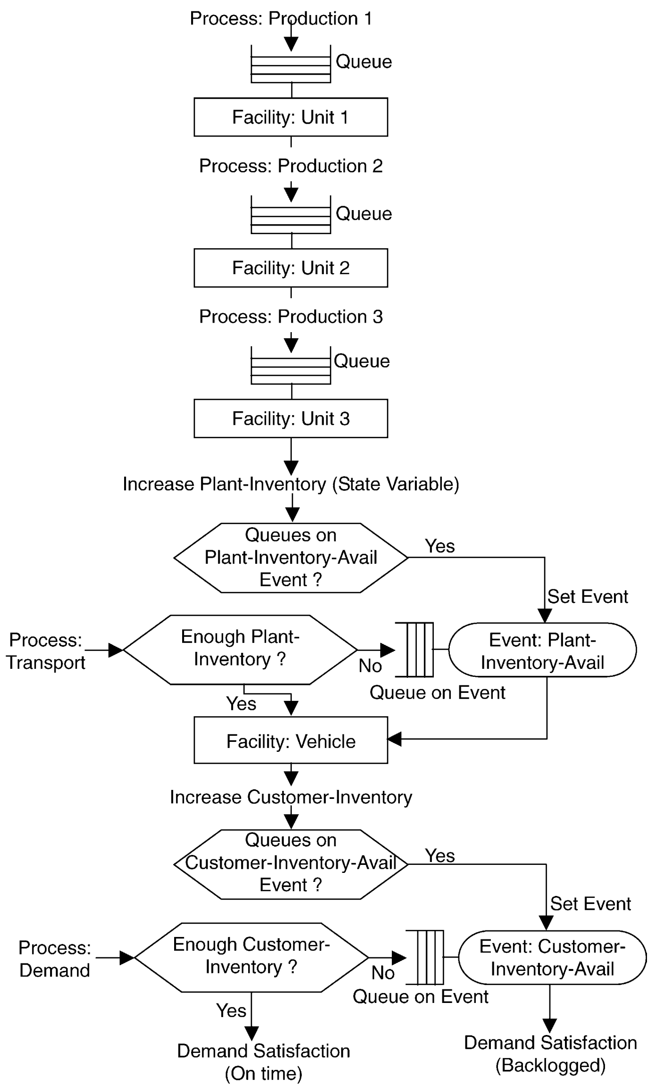


Fig. 15. Simulation flow diagram.

units in series for the production of several finished products, and one customer. Fig. 15 depicts the simulation flow of the supply chain simulation model under this assumption. The key simulation components are the following:

- Processes: production 1, 2 and 3, transportation, demand.
- Facilities: unit 1, 2 and 3, vehicle.
- Events: plant-inventory-avail, customer-inventory-avail.

Among these processes, the production 1, transportation and demand processes are initiating processes. The release of these initiating processes into the simulation is managed by the computation control module.

Once the production schedule is made but before the simulation begins or resumes, the computation control module prepares the schedule for the release time of the production 1 processes, the demand processes and the transportation processes. The scheduled activities of the production schedule are sorted in increasing order of time for the start of the activities. Each scheduled activity corresponds to a production 1 process. The production 1 process copies its information on its product id, production amount, production unit id (in case when multiple parallel units are available), and start time of production from the activity schedule. The production 1 processes are grouped based on the date of their production start times. As the simulation proceeds, at the beginning of each new day, the production 1 processes in the group whose start times fall within that new day, are released and stacked in the queue of the unit1. They are served by the unit 1 in the order of their scheduled start times.

To illustrate the logic of the simulation model constructed for the case study of Section 7, consider a supply chain composed of one production facility, which has three processing units in series for the production of several finished products, and one customer. Fig. 15 depicts the simulation flow of the supply chain simulation model under this assumption. The key simulation components are the following:

- Processes: production 1, 2 and 3, transportation, demand.
- Facilities: unit 1, 2 and 3, vehicle.
- Events: plant-inventory-avail, customer-inventory-avail.

Among these processes, the production 1, transportation and demand processes are initiating processes. The release of these initiating processes into the simulation is managed by the computation control module.

Once the production schedule is made but before the simulation begins or resumes, the computation control module prepares the schedule for the release time of the production 1 processes, the demand processes and the transportation processes. The scheduled activities of the production schedule are sorted in increasing order of time for the start of the activities. Each scheduled activity corresponds to a production 1 process. The production 1 process copies its information on its product id, production amount, production unit id (in case when multiple parallel units are available), and start time of production from the activity schedule. The production 1 processes are grouped based on the date of their production start times. As the simulation proceeds, at the beginning of each new day, the production 1 processes in the group whose start times fall within that new day, are released and stacked in the queue of the unit1. They are served by the unit 1 in the order of their scheduled start times.

The due dates and the deterministic amounts of the demands are set by the demand generator. The actual amounts of the demands involved in the simulation are given by the demand scenarios. The computation control unit maintains a list of the demand processes whose due dates are the same together with their corresponding deterministic demands. The list is sorted in the increasing order of the due dates of each demand process. The demand processes are released into the simulation exactly when the simulation clock reaches the due date of each demand process.

The computation control module generates transportation processes for each demand process. The transportation process requests the shipment of the amount corresponding to the demand process. The time when the transportation process is released is determined by subtracting the transportation time from the due date of the corresponding demand process.

In this illustrative example, the production 2 and 3 processes are dependent processes, which are released into the simulation as a result of the completion of other processes. The completion of a production 1 process generates a production 2 process and releases it into the simulation immediately. Likewise, production 3 processes are released right after the completion of production 2 processes.

Among the state variables, plant-inventory and customer-inventory are important variables that have direct effect on the customer satisfaction level. At the completion of the production 3 processes, plant-inventory is increased by the corresponding production amounts. Plant-inventory is decreased when a shipment to customer is initiated. Customer-inventory refers to the inventory kept at the customer sites. It is increased by the transportation amount when a transport process completes and decreased accordingly when a demand process claims its order at its due date.

When a transport process encounters insufficient plant-inventory, a queue entry is stacked in the plant-inventory-avail event queue. This event is executed and the transport process completed when there is a sufficient increase in the plant-inventory to meet the transportation amount. Likewise, a queue entry is generated in the customer-inventory-avail event queue, when a demand process meets with insufficient customer-inventory. This queued event is executed when the arrival of a new transportation results in sufficient customer-inventory level to meet the demand. If the demand process can be satisfied in full after its release, the demand process is counted as a demand met on time. Otherwise, it is counted as a backlogging case. These counts are used in the estimation of the customer satisfaction levels of each simulation timelines.

Appendix B. Case study data

See Tables 6–9

Table 6
Package ratio of each type-grade couple

Type	Grade									
	0	1	2	3	4	5	6	7	8	9
A	0.50	0.40	0.35	0.65	0.50	0.40	0.60	0.70	0.80	0.80
B	0.50	0.40	0.40	0.70	0.70	0.65	0.65	0.75	0.85	0.85
C	0.65	0.50	0.70	0.50	0.60	0.50	0.70	0.80	0.75	0.65
D	0.50	0.50	0.60	0.60	0.50	0.50	0.50	0.50	0.70	0.70
E	0.50	0.60	0.60	0.55	0.75	0.50	0.40	0.60	0.60	0.50

The given number is the fraction of bag package type for each type-grade couple. The fraction of box package type can be obtained as 1.0 minus given fraction.

Table 7
Inventory cost

Type	Grade									
	0	1	2	3	4	5	6	7	8	9
A	9.59	9.59	7.67	7.67	7.67	7.67	7.67	7.67	6.71	6.71
B	10.55	9.59	9.59	7.67	7.67	7.67	7.67	5.75	5.75	5.75
C	9.59	9.59	9.59	9.59	7.67	7.67	6.71	6.71	6.71	6.71
D	10.55	10.55	10.55	9.59	9.59	9.59	8.63	8.63	8.63	8.63
E	11.51	11.51	10.55	10.55	10.55	10.55	9.59	9.59	9.59	9.59

Unit: $10E-5$ \$/lbs per day.

Table 8
Backlogging cost

Type	Grade									
	0	1	2	3	4	5	6	7	8	9
A	47.95	47.95	38.36	38.36	38.36	38.36	38.36	38.36	33.56	33.56
B	52.74	47.95	47.95	38.36	38.36	38.36	38.36	28.77	28.77	28.77
C	47.95	47.95	47.95	47.95	38.36	38.36	33.56	33.56	33.56	33.56
D	52.74	52.74	52.74	47.95	47.95	47.95	43.15	43.15	43.15	43.15
E	57.53	57.53	52.74	52.74	52.74	52.74	47.95	47.95	47.95	47.95

Unit: $10E-5$ \$/lbs per day.

Table 9
Safety stock shortage cost

Type	Grade 0	1	2	3	4	5	6	7	8	9
A	23.97	23.97	19.18	19.18	19.18	19.18	19.18	19.18	16.78	16.78
B	26.37	23.97	23.97	19.18	19.18	19.18	19.18	14.38	14.38	14.38
C	23.97	23.97	23.97	23.97	19.18	19.18	16.78	16.78	16.78	16.78
D	26.37	26.37	26.37	23.97	23.97	23.97	21.58	21.58	21.58	21.58
E	28.77	28.77	26.37	26.37	26.37	26.37	23.97	23.97	23.97	23.97

Unit: $10E - 5$ \$/lbs per day.

References

Advanced Process Combinatorics Inc. (2004). Available on the web at http://www.combination.com/.

Andradottir, S. (1998). A review of simulation optimization techniques. *Proceedings of the 1998 Winter Simulation Conference*, 151–158.

Carson, Y., & Maria, A. 1997. Simulation optimization: methods and applications. *Proceedings of the 1997 Winter Simulation Conference*, 118–126.

Cheng, L., Subramanian, E., & Westerberg, A. W. (2003). Design and planning under uncertainty: Issues on problem formulation and solution. *Computers and Chemical Engineering, 27*, 781–802.

DeCroix, G. A., & Arreola-Risa, A. (1998). Optimal production and inventory policy for multiple products under resource constraints. *Management Science, 44*(7), 1201–1218.

Fu, M., & Hu, J. Q. 1997. *Conditional Monte-Carlo: Gradient Estimation and Optimization Applications*. Kluwer Academic Publishers.

Fu, M. (2002). Optimization for simulation: Theory vs. practice. *INFORMS Journal on Computing, 14*(3), 192–215.

Graves, S. C., & Willems, S. P. (2000). Optimizing strategic safety stock placement in supply chains. *Manufacturing and Service Operations Management, 2*(1), 68–83.

Gupta, A., & Maranas, C. D. (2000a). A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty. *Industrial and Engineering Chemistry Research, 39*(10), 3799–3813.

Gupta, A., & Maranas, C. D. (2003). Midterm supply chain planning under demand uncertainty: Customer demand satisfaction and inventory management. *Computers and Chemical Engineering, 27*, 1219–1227.

Gupta, A., Maranas, C. D., & McDonald, C. M. (2000b). Midterm supply chain planning under demand uncertainty: Customer demand satisfaction and inventory management. *Computers and Chemical Engineering, 24*, 2613–2621.

Hopp, J. W., & Spearman, M. L. (2000). *Factory Physics*. McGraw-Hill.

Ierapetritou, M. G., & Pistikopoulos, E. N. (1994). Novel optimization approach of stochastic planning models. *Industrial and Engineering Chemistry Research, 33*(8), 1930–1942.

Karimi, I. A., & McDonald, C. M. (1997). Planning and scheduling of parallel semicontinuous process: Short-term scheduling. *Industrial and Engineering Chemistry Research, 36*(7), 2701–2714.

Law, A. M., & Kelton, W. D. (2000). *Simulation Modeling and Analysis*. McGraw-Hill.

Mesquite Software, Inc. (2004). Available on the web at http://www.mesquite.com/.

McDonald, C. M., & Karimi, I. A. (1997). Planning and scheduling of parallel semicontinuous process: Production planning. *Industrial and Engineering Chemistry Research, 36*(7), 2691–2700.

Porteus, E. L. (1990). Stochastic inventory theory. In D. Heyman, & M. Sobel (Eds.), *Handbooks in Operations Research and Management Science* (vol 2., pp. 605–652). Amsterdam: Elservier Science Publishers.

Porteus, E. L. (2002). *Foundations of Stochastic Inventory Theory*. Stanford University Press.

Reklaitis, G. V. (1982). Review of scheduling of process operations. *AIChE Symposium Series, 78*(214), 119–133.

Shah, N., & Pantelides, N. N. (1992). Design of multipurpose batch plants with uncertain production requirements. *Industrial and Engineering Chemistry Research, 31*(5), 1325–1337.

Subramanyam, S., Pekny, J. F., & Reklaitis, G. V. (1994). Design of batch chemical plants under market uncertainty. *Industrial and Engineering Chemistry Research, 33*(11), 2688–2701.

Subramanian, D., Pekny, J. F., & Reklaitis, G. V. (2000a). A simulation-optimization framework for research and development pipeline management. *AIChE Journal, 47*(10), 2226–2242.

Subramanian, D. 2001b. *A Computational Architecture to Address Combinatorial and Stochastic Aspects of Process Management Problems*. Ph.D. Thesis, School of Chemical Engineering, Purdue University.

Zipkin, P. H. (1986). Models for design and control of stochastic, multi-item batch production systems. *Operations Research, 34*(1), 92–104.