Inventory Policies and Safety Stock Optimization for Supply Chain Planning

Braulio Brunaud[1], Jose M. Laínez-Aguirre[2], Jose M. Pinto[2], and Ignacio E. Grossmann[1]

[1]Carnegie Mellon University, Department of Chemical Engineering
[2]Praxair Inc.

May 31, 2018

Abstract

In this paper, traditional supply chain planning models are extended to simultaneously optimize inventory policies. The inventory policies considered are the (r,Q) and (s,S) policies. In the (r,Q) inventory policy and order for Q units is placed every time the inventory level reaches level r. While in the s,S policy the inventory is reviewed in predefined intervals. If the inventory is found to be below level s, an order is placed to bring the level back to level S. Additionally, to address demand uncertainty four safety stock formulations are presented: 1) proportional to throughput, 2) proportional to throughput with risk-pooling effect, 3) explicit risk-pooling, and 4) guaranteed service time. The models proposed allow simultaneous optimization of safety stock, reserve and base stock levels in tandem with material flows in supply chain planning. The formulations are evaluated using simulation.

1 Introduction

Supply chain management is a demand propagation problem. The last stage in the chain is the distribution of finished goods to end customers. Most operations upstream from that stage is driven by an action taken by the customer, either walking into a store to purchase a product or placing an order to have the product delivered. Since the expected service time is normally much smaller than the production lead time, the demand of a customer must be anticipated through a forecast. This is the origin of most decisions involved in a supply chain. The optimization of a supply chain plan becomes the estimation of the optimal decisions to respond to a given demand forecast. The main difficulty is that a forecast is an estimation, and like every estimation, it is prone to error. In supply chain optimization, this error is referred to demand uncertainty. Furthermore, to address the mismatch between lead times and required service times, inventory is held at different stages of the process.

An optimal supply chain plan defines the amount of material transported between facilities at any given time period within the planning horizon. When determining these flows, the inventory levels at the storage facilities are simultaneously determined, because the multiperiod planning models employed include inventory balance constraints (Bradley and Arntzen, 1999). This indicates that given a demand forecast it is possible to determine the exact timing and amount of inventory replenishments. However, in practice warehouses are actually managed in terms of policies, which are simple rules that dictate when to replenish an inventory and the corresponding replenishment amount. The definition of the policy parameters is typically done using average demand and lead time as input, using defined mathematical expressions, historical data, or using simulation (Kapuscinski and Tayur, 1999). In this paper we resolve the discrepancy between the inventory curves obtained from a planning model and the implementation of inventory policies by proposing a mixed-integer programming models capable of simultaneously determining the optimal flows and the inventory policy parameters. Garcia-Herreros et al. (2016) propose logic formulations to implement inventory policies in production systems with arrangements of inventories in series and in parallel. The inventory policy parameters are optimized using stochastic programming. The policy considered is a simple basestock policy to approximate multistage-stochastic programming formulations. In this paper we consider a general derivation for traditional inventory policies commonly used in practice.

The first policy considered in this paper is the continuous review (r,Q) (Galliher et al., 1959). When the inventory reaches level r a replenishment order for Q units is placed. Methods proposed to determine the (r,Q) policy parameters include heuristics (Platt et al., 1997), and unconstrained optimization (Federgruen and Zheng, 1992). The second policy considered is the periodic review (s,S). The inventory is reviewed in defined periods. If the level is below s, an order to bring back the inventory position to S is placed. To determine the policy parameters several methods including heuristics (Zheng and Federgruen, 1991), and simulation-based optimization (Bashyam and Fu, 1998) have been proposed. For both policies, previous works based on constrained optimization have found it difficult to solve the resulting models. In this work we propose a mixed-integer linear programming model, which together with advances in MILP solvers (Linderoth, 2017), provides feasible alternatives for practical applications.

The uncertainty in the demand must also be addressed to prevent stockouts. Uncertainty can be considered either using a stochastic programming framework or considering a safety stock. Stochastic inventory optimization problems are still very challenging to model and solve. Thus, the problem size they can address is limited. On the other hand, safety stock (Enke, 1958) is a very old and intuitive concept, although its incorporation in supply chain planning models is quite recent. In this paper, we present and analyze four alternatives to estimate the optimal amount of safety stock amount in a supply chain planning context. The safety stock formulations considered are: 1) proportional to throughput, 2) proportional to throughput with risk-pooling effect, 3) explicit risk-pooling, and 4) guaranteed service time.

The literature on inventory models is extensive, from the work of Arrow, Karlin and Scarf (Arrow et al., 1958) to the study of optimal inventory management in a variety of situations. However, the inclusion of inventory models for safety stock and inventory policies in a mathematical programming framework has been limited. In previous approaches, the safety stock is considered as a fixed parameter that acts as a lower bound for the inventory (Relvas et al., 2006; Varma et al., 2007). Jackson and Grossmann (2003) and Lim and Karimi (2003) also consider the safety stock as fixed lower bound for inventory, but they also include a penalty term in the objective function to penalize the violation of this bound.

Shen et al. (2003) propose a mixed-integer nonlinear programming (MINLP) formulation for the location of facilities, that explicitly includes the risk-pooling effect (Eppen, 1979). The model was used by Miranda and Garrido (2009), and extended by You and Grossmann (2008) to incorporate variable coefficient of variation (variance to mean ratio) between customers. Because of the nonlinearities, these models are difficult to solve, and the size of problem they can address is limited. Diabat and Theodorou (2015) proposed a general linearization for the model to formulate a mixed-integer linear programming model (MILP). Recently, Brunaud et al. (2017) proposed a piecewise-linear formulation for the problem and showed that the approximation yields similar results to the MINLP formulation. You and Grossmann (2010) integrated the guaranteed service level concept proposed by Graves and Willems (2000) in MINLP models.

The formulations proposed in this paper provide a wide array of options to model a wide range of applications. The problem is described in Section 2, the formulations for inventory policies are presented in Section 3. Safety stock is considered in Section 4. A case study with optimization and simulation results is presented in Sections 5 and 6, respectively.

2 Problem Description

A supply chain network structure is given, including suppliers, warehouses, retailers, and a number of customers (Fig. 1). It is required to determine the optimal material flows and inventory levels to satisfy the demand forecast. The objective is to minimize the transportation and inventory costs.

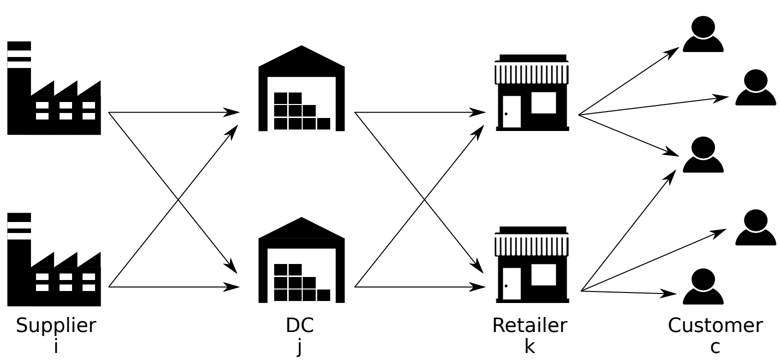

Figure 1. Supply chain network structure

To address this problem a multiperiod linear programming model (LP) is formulated as defined by Eqs. 1–7.

$$
\begin{aligned}
Max \quad & \sum_i \sum_j \sum_t TC_{ij} x_{ijt} + \sum_j \sum_k \sum_t TC_{jk} x_{jkt} + \sum_k \sum_c \sum_t TC_{kc} x_{kct} \\
& + \sum_j \sum_t HC_j inv_{jt} + \sum_k \sum_t HC_k inv_{kt} && (1) \\
s.t. \quad & \sum_k x_{kct} = D_{ct} \qquad \forall k, t && (2) \\
& inv_{jt} = inv_{jt-1} + \sum_i x_{ijt} - \sum_k x_{jkt} - \sum_c x_{jct} \qquad \forall j, t && (3) \\
& inv_{kt} = inv_{kt-1} + \sum_j x_{jkt} - \sum_c x_{kct} \qquad \forall i, t && (4) \\
& inv_{j0} = InitInv_j \qquad \forall j && (5) \\
& inv_{jT} = InitInv_j \qquad \forall j && (6) \\
& x_{ijt}, x_{jkt}, x_{kct} \geq 0, inv_{jt} \leq MaxInv_j, inv_{kt} \leq MaxInv_k && (7)
\end{aligned}
$$

In the model, i is a supplier, j is a distribution center, k is a retailer, c is a customer, and t is the time period. The parameter TC is the transportation cost, HC is the inventory holding cost, $MaxInv$ is the warehouse capacity, and D is the demand. The variables inv and x represent inventory and flow, respectively. In order to reduce the impact of the initial inventory, both initial and final inventories (period 0 and period T) have been set equal to the initial inventory parameter $InitInv$ (Eqs. 5 and 6). It is possible to hold inventory at the distribution centers (DC) and also at the retailers. Although it is possible to consider a single type of warehouse we consider DCs separated from retailers to explicitly illustrate the effects of safety stock in multi-echelon systems.

Provided there is enough production capacity and the total lead time is less than the time required by the customer, the above model yields a solution with zero inventory. This would be the situation of a just-in-time operation. Every order would need to be moved throughout the supply chain in a single day. For most supply chains, this is neither feasible nor practical. It would require frequent replenishment orders to the warehouses which would not be accepted by inventory managers. A partial solution would be to incorporate the cost of placing an order. This cost does not depend on the amount ordered. Binary variables y_{ijt} and y_{jkt} are added to the model (LP) to indicate when an order is placed to replenish the inventory. For now, we assume there is no replenishment lead time. To consider ordering cost Eqs. 8–10 are added to the model.

$$x_{ijt} \leq MaxInv_j y_{ijt} \qquad \forall i, j, t \tag{8}$$

$$x_{jkt} \leq MaxInv_k y_{jkt} \qquad \forall j, k, t \tag{9}$$

$$y_{ijt}, y_{jkt} \in \{0, 1\} \tag{10}$$

Also, the order cost is added to the objective

$$OrderCost = \sum_t \left(\sum_i \sum_j CO_{ij} y_{ijt} + \sum_j \sum_k CO_{jk} y_{jkt} \right) \tag{11}$$

where CO is cost for placing a replenishment order. For a small-sized problem like the one from Fig. 1, the model is solved in under a second. The inventory curves obtained are presented in Fig. 2. The model solved to obtain the curves from Fig. 2 is formulated with Eqs. 2–10. The objective function is the summation of Eqs. 1 and 11.

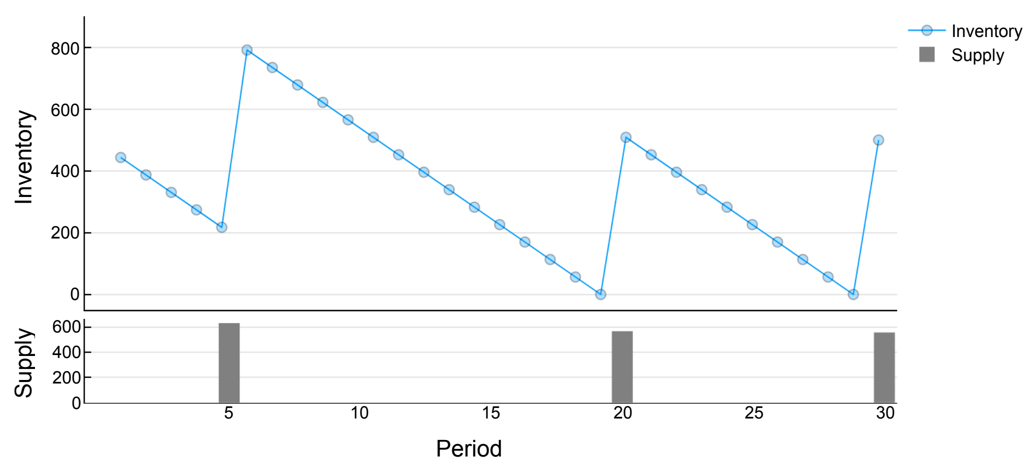

Figure 2. Inventory curves for model including fixed ordering cost

The curves from Fig. 2 resemble the inventories observed in practice. However, they are obtained for a specific, usually average, demand forecast. Deviations from the demand forecast are likely to occur in practice. The output of the model does not indicate how to react if this happens. Because of the dynamic behavior of demand, inventory is usually managed in terms of policies, which are simple rules that dictate when to replenish an inventory and the replenishment amount. To our knowledge, the main inventory policies have not been incorporated in supply chain planning models despite being used in practice. Garcia-Herreros et al. (2016) incorporated a basestock inventory policy and optimized its parameters using stochastic programming.

In the following sections, the base model will be extended by including blocks of constraints to model inventory policies (Section 3) and to handle safety stock (Section 4). Figure 3 summarizes the formulations that are developed. A complete model is generated by the combination of the base model with an inventory policy and a safety stock formulation. It is also possible to not consider inventory policies or safety stock. The combination of these elements yields 15 possible models, using one inventory policy option from the set $P = \{None, (r, Q), (s, S)\}$, and one safety stock option from the set $S = \{None, Proportional, Piecewise, Explicit, Guaranteed\}$. For example, a model can be generated using the base model (Eqs. (1)–(11)) with an (r,Q) policy, and explicit risk-pooling safety stock.

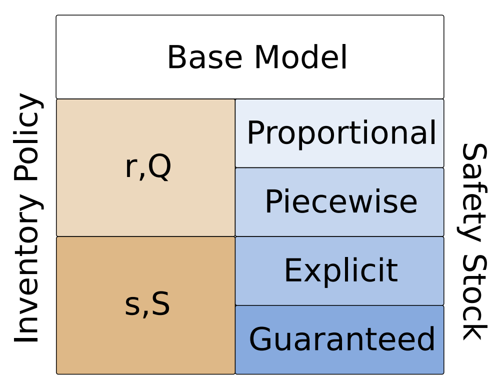

Figure 3. Schematics of the different options to formulate supply chain planning models with inventory policies and safety stock

3 Modeling of Inventory Policies

Traditionally, the decision parameters for an inventory policy are obtained from historical data or calculations based on average demand and lead times. Here, we present formulations for planning models that explicitly optimize the parameters of inventory policies. The development from this section is general enough to accommodate other policies tailored to a specific the application.

3.1 Continuous-review (r,Q) Policy

Under the (r, Q) policy the inventory level is continuously reviewed. When the on-hand inventory reaches the level r an order is placed for Q units of product. An example of an inventory curve using this policy is presented in Fig. 4

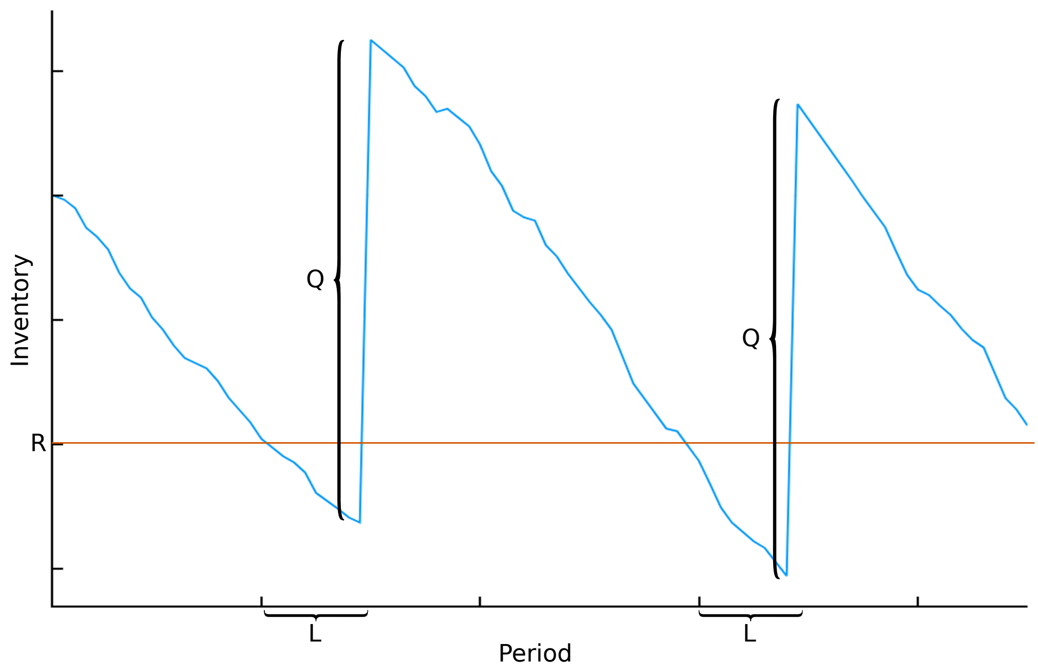

Figure 4. (r,Q) policy example

Since the (r,Q) policy requires continuous review of the inventory level, it is better suited for continuous time models. For discrete-time planning models considered in this paper the condition of having the inventory level at r to place an order must be relaxed to place an order when the inventory is lower than or equal to r.

From the modeling standpoint, the policy can be formulated using the following logic: if the inventory is less than or equal to level r at a given period a replenishment order is placed, unless an order had been already placed, which is represented by Eqs. (12)–(13)

$$inv_{jt-L} \leq r_j \Longleftrightarrow \bigvee_{\tau=t-L+1}^{t} y_{j\tau} \qquad \forall j, t > 1+L \tag{12}$$

$$y_{jt} \Rightarrow \sum_i x_{ijt} = Q_j \qquad \forall j, t \tag{13}$$

which can be reformulated with the mixed-integer constraints:

$$inv_{jt-L} - r_j \leq MaxInv_j(1 - z_{jt}) \qquad \forall j, t > 1+L \tag{14}$$

$$\sum_{\tau=t-L+1}^{t} y_{j\tau} \leq z_{jt} \qquad \forall j, t > 1+L \tag{15}$$

$$-MaxInv_j \left(\sum_{\tau=t-L+1}^{t} y_{j\tau} \right) + \epsilon \leq inv_{jt-L} - r_j \qquad \forall j, t > 1+L \tag{16}$$

$$\sum_i x_{ijt} - Q_j \leq MaxInv_j(1 - y_{jt}) \qquad \forall j, t \tag{17}$$

$$-\sum_i x_{ijt} + Q_j \leq MaxInv_j(1 - y_{jt}) \qquad \forall j, t \tag{18}$$

$$\sum_i x_{ijt} \leq MaxInv_j y_{jt} \qquad \forall j, t \tag{19}$$

$$z_{jt} \in \{0, 1\} \qquad \forall j, t \tag{20}$$

3.2 Periodic-review (s,S) Policy

In many applications, continuously checking the inventory level might not be practical or even possible. In others, stocks are replenished on certain days of the week. In this cases the inventory is reviewed in defined intervals. The decision of placing a replenishment order or not, and the amount of the order depend on the inventory level at the moment the inventory is reviewed. When the inventory is below a threshold level, an order to replenish the inventory position to a specified base stock level is placed. For example, the inventory level can be reviewed every p days. Let us call the inventory level at this period inv_t. If the on-hand inventory is below a defined level s, an order is placed to restore the inventory position to level S, i.e. $S - inv_t$ units are ordered. An example inventory curve using this policy is presented in Fig. 5

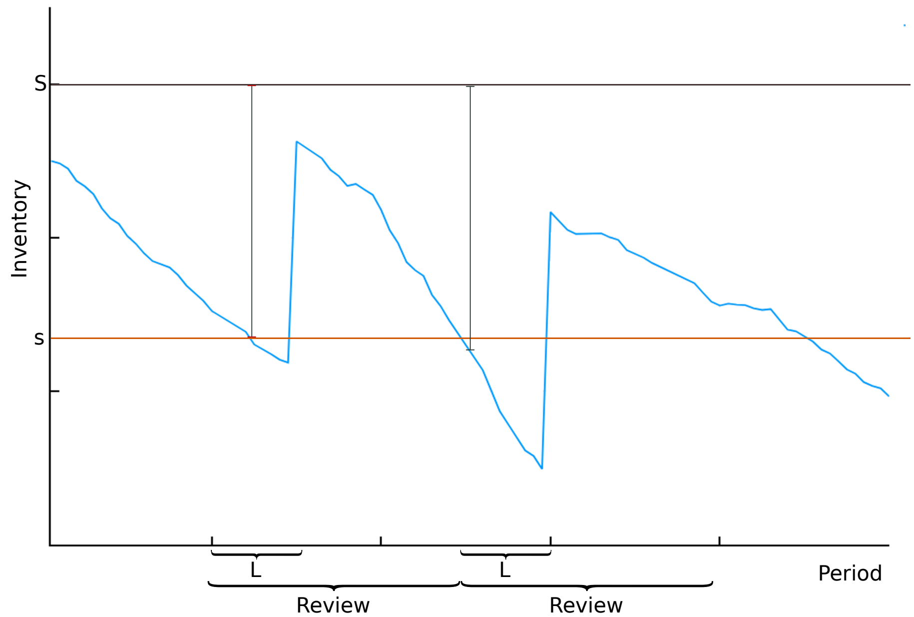

Figure 5. s,S inventory policy example

We now derive a formulation to obtain the review frequency, starting day, and levels s and S as outputs of a supply chain planning model. The derivation is presented for a distribution center j. However, the steps are analogous for a retailer k.

First, we define a replenishment patterns matrix that indicates the days on which a replenishment is allowed (Table 1). The review frequency and the starting day identify each pattern. The reader may notice that any custom review pattern can be accommodated.

Next, a binary variable, rp_{jn}, associated with the use of a pattern is defined. The variable takes a value of 1 if review pattern n is chosen at distribution center j. The selection of a given pattern dictates if a replenishment is allowed in a given period. A new variable, ra_{jt}, is used to indicate a replenishment is allowed at distribution center j in period t. The two variables are linked by the constraint from Eq. 21. Since a higher frequency of review indirectly increases the operational costs, a penalty term associated to the pattern chosen is included. For example, in the case of

Table 1. Replenishment policy matrix

Pattern	Mo	Tu	We	Th	Fr	Mo	Tu	We	Th	Fr	Mo	Tu	We	Th	Fr
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
3	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
4	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
5	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
6	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
7	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0
8	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
9	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
10	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
11	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
12	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0
13	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
14	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0
15	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1

supply contracts, the ability of receiving frequent replenishments is a premium.

$$\sum_{n \in N_t} rp_{jn} = ra_{jt} \qquad \forall j, t \tag{21}$$

where N_t is the set of patterns that allow a replenishment on period t. Flow into the distribution center j is only allowed if ra_{jt} is 1 (Eq. 22).

$$\sum_{i} x_{ijt} \leq M ra_{jt} \qquad \forall j, t \tag{22}$$

Clearly, only one replenishment pattern must be chosen.

$$\sum_{n} rp_{jn} = 1 \qquad \forall j \tag{23}$$

Having defined the constraints to ensure the replenishment in the days allowed by the selected replenishment pattern, one has to enforce that the replenishment is done when the on-hand inventory is below s. For this purpose, a new variable, rr_{jt}, is defined to indicate when a replenishment is required at distribution center j in period t. With the new variable the logical condition that represents the inventory requirement is given as Eq. 24, and formulated as the constraints in Eqs. 25–26.

$$inv_{jt-L} \leq s_j \iff rr_{jt} \qquad \forall j, t > 1 + L \qquad (24)$$

$$-M\, rr_{jt} + \epsilon \leq inv_{jt-L} - s_j \qquad \forall j, t > 1 + L \qquad (25)$$

$$inv_{jt-L} - s_j \leq M(1 - rr_{jt}) \qquad \forall j, t > 1 + L \qquad (26)$$

Where, ϵ represents a small number. A replenishment will happen when it is required and allowed. The variable yr_{jt} indicates if a replenishment is actually performed at period t. The corresponding logical condition is given by Eq. 27, and represented as constraints by Eqs. 28–30.

$$ra_{jt} \wedge rr_{jt} \iff yr_{jt} \qquad \forall j, t \qquad (27)$$

$$ra_{jt} + rr_{jt} \leq yr_{jt} + 1 \qquad \forall j, t \qquad (28)$$

$$yr_{jt} \leq ra_{jt} \qquad \forall j, t \qquad (29)$$

$$yr_{jt} \leq rr_{jt} \qquad \forall j, t \qquad (30)$$

Finally, once a replenishment has been decided, the replenishment amount must bring back the inventory position to S. This is represented by the disjunction in Eq. 31, and reformulated using Big-M in Eqs. 32–34.

$$\begin{bmatrix} Y_{jt} \\ \sum_j x_{ijt} = S_j - inv_{jt-L} \end{bmatrix} \vee \begin{bmatrix} \neg Y_{st} \\ \sum_j x_{ijt} \leq 0 \end{bmatrix} \qquad \forall j, t > 1 + L \qquad (31)$$

$$\sum_i x_{ijt} - S_j + inv_{jt-L} \leq M(1 - yr_{jt}) \qquad \forall j, t > 1 + L \qquad (32)$$

$$-\sum_i x_{ijt} + S_j - inv_{jt-L} \leq M(1 - yr_{jt}) \qquad \forall j, t > 1 + L \qquad (33)$$

$$\sum_i x_{ijt} \leq M yr_{jt} \qquad \forall j, t \qquad (34)$$

where Y_{jt} is the boolean variable associated with yr_{jt}. Note that introduction of Eq. 34 makes

Eq. 22 redundant, our experiments indicate that removing the constraint does not have an effect in the LP relaxation of the problem. The inventory curve for the example using this formulation is presented in Fig. 6.

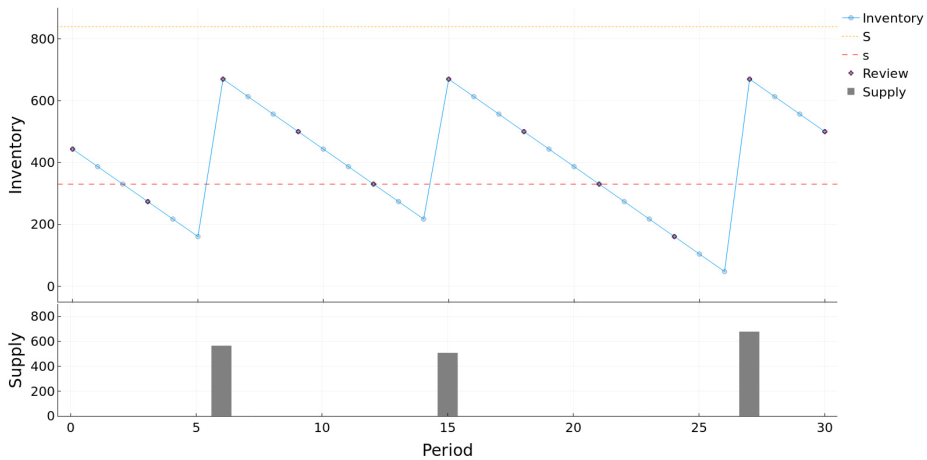

Figure 6. Inventory curve for retailer 2 using the (s,S) policy

4 Safety Stock Modeling

To handle the uncertainty on the demand and prevent stockouts a safety stock must be included. In this section we analyze four alternatives to incorporate safety stock in supply chain models.

4.1 Safety Stock under Normally Distributed Demand

When the demand is normally distributed with mean μ and variance σ^2, a safety stock must be kept to hedge against the variation of the demand during the replenishment lead time (L). When combined with an (s,S) policy, the safety stock must hedge against the variation during L + P periods, where P is the review frequency defined by the replenishment policy selected from Table 1. When demands are assumed independent and identically distributed (iid), the safety stock is then proportional to the number of standard deviations that must be kept in inventory (Eppen and Martin, 1988). For the rest of the paper, all demands are assumed to be iid.

$$ss = z\sqrt{L}\sigma \tag{35}$$

where ss is the safety stock, and z is the inverse cumulative normal distribution coefficient for a given service level required. For example, to obtain a service level of 95%, the value of z is 1.65.

4.2 Proportional to Throughput

Analyzing Eq. 35, it can be observed that the safety stock is proportional to the service level, the standard deviation of the demand, and the square root of the lead time. Considering that the service level is given, the lead time is constant and that the standard deviation is proportional to the throughput, we establish the following relationship for the safety stock

$$ss_k = \frac{\beta\sqrt{L}}{|T|}\sum_c \sum_t x_{kct} \qquad \forall k \tag{36}$$

The proportionality of the standard deviation with the throughput is explained by two factors:

1. The magnitude of the variability in the demand of a given customer is characterized by its coefficient of variation $\frac{\sigma}{\mu}$. When serving a single customer an increase in demand is also reflected with an increase in the mean demand, and the standard deviation.

2. An increase in the throughput is also an indirect indicator of an increase of the number of customers that are being served by a given warehouse, which also increases the standard deviation at the warehouse (see Section 4.3).

This relationship indicates that a retailer serving a larger volume is more at risk of being affected by steep demand variabilities. Therefore, it should hold a higher safety stock. The parameter β indicates the level of stockout risk defined, for example we can consider β to be 20%. In this case the safety stock would be 20% of the daily throughput at the retailer multiplied by the square root of the replenishment lead time. Eq. 36 provides a method to optimize the safety stock based on the throughput through the retailer. In previous works the safety stock was not a variable but a given parameter (Jackson and Grossmann, 2003; Lim and Karimi, 2003; Varma et al., 2007). In the proposed formulation, the safety stock is used as a lower bound for the inventory (Eq. 37).

$$inv_{kt} \geq ss_k \qquad \forall k, t \tag{37}$$

4.3 Piecewise Linear with Risk-pooling

Eppen (1979), showed that when more customers are served by the same warehouse, unexpected increases in the demand of one customer are offset by the decrease in the demand of another customer. Thus the joint standard deviation decreases. This is known as the ”risk-pooling effect”.

Brunaud et al. (2017) proposed a piecewise linear formulation to capture this effect. It is based on recognizing that the higher the throughput, the likelihood to be serving more customers increases, and thus the standard deviation is reduced. Mathematically, the idea is to decrease the β parameter from Eq 36 with the increase in the throughput. This leads to a piecewise linear representation of the safety stock, using Special Ordered Sets of Type II (SOS2) (Beale and Tomlin, 1970). Besides being ordered, the variables in these kind of sets meet an additional adjacency condition: no more than two variables can be non-zero at a time. The non-zero variables must be adjacent. Variables of these kind are set to each of the breakpoints of the piecewise-linear function, and are used to interpolate and obtain the safety stock. The piecewise-linear approximation, which is computationally more efficient than the logarithmic approximation proposed by Cafaro and Grossmann (2014), is given by the following constraints,

$$\sum_c \sum_t x_{kct} = \sum_n \lambda_{kn} SSx_n \qquad \forall k \tag{38}$$

$$ss_k = \sum_n \lambda_{kn} SSy_n \qquad \forall k \tag{39}$$

$$\sum_n \lambda_{kn} = 1 \qquad \forall k \tag{40}$$

$$\lambda_{kn} \in [0, 1], SOS2 \qquad \forall k, n \tag{41}$$

where (SSx, SSy) are the breaking points of the piecewise linear function (Fig. 7).

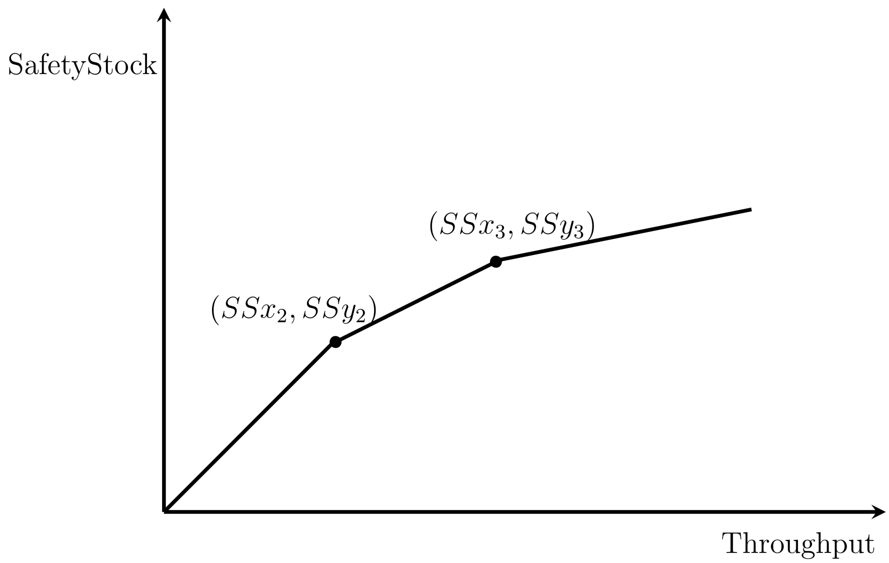

Figure 7. Safety stock piecewise linear function

The key assumption made is that the throughput level relates to the number of customers served, and thus to the level of risk-pooling employed. To illustrate the validity of this assumption consider an example with a single warehouse serving up to six customers. Each customer has a given demand with variance σ_c^2, and let y_c be a binary variable indicating that the warehouse is serving customer c. y is a vector containing each y_c as its components. For example, a value $y = (1, 0, 1, 0, 0, 0)$ indicates that the warehouse serves the first and third customers. For each of the 64 possible values of the vector y it is possible to compute the safety stock with the following expression (You and Grossmann, 2008).

$$ss = z\sqrt{L}\sqrt{\sum_c \sigma_c^2 y_c} \tag{42}$$

The value of the safety stock for each value of the vector y is presented in Fig. 8. The shape of the point represents the number of customers being served by the warehouse

The values from Fig. 8 can be approximated by a piecewise-linear function relating the through-

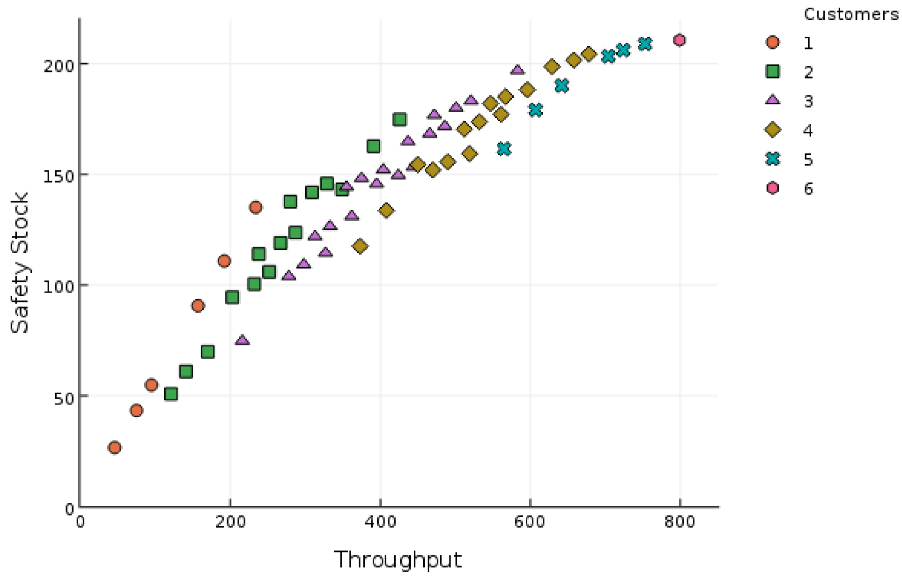

Figure 8. Safety stock value for each possible value of the vector y

put to the safety stock (see Fig. 9).

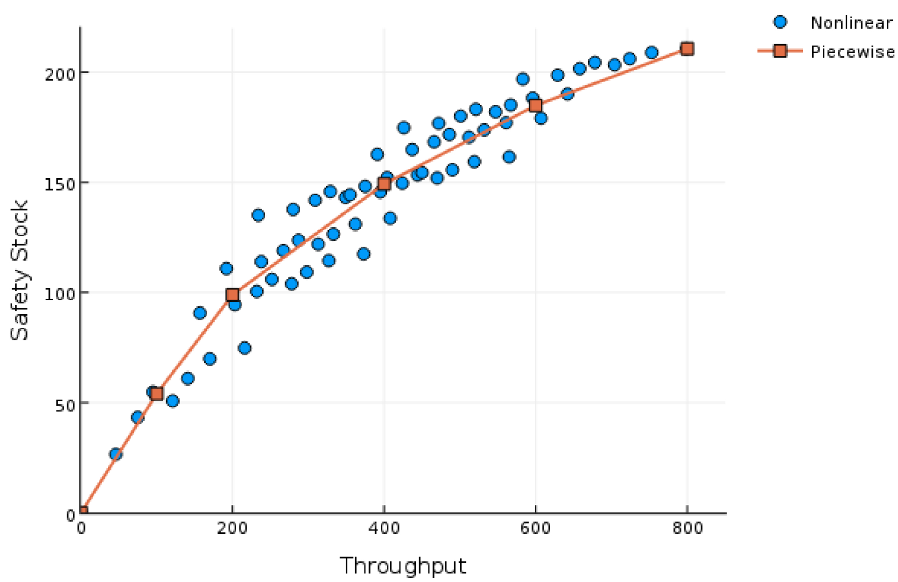

Figure 9. Piecewise-linear approximation of the safety stock as a function of throughput

4.4 Explicit Risk Pooling

Daskin et al. (2002) incorporate safety stock with risk-pooling effect in an MINLP formulation, and solved it using Lagrangean relaxation(Fisher (1985)). You and Grossmann (2008), and Miranda and Garrido (2009), also used a similar formulation.

In this formulation, the standard deviation at a retailer is the geometric sum of the standard deviations of the customers served,

$$\sigma_k = \sqrt{\sum_{c \in C_k} \sigma_c^2} \tag{43}$$

where k represents the retailer and C_k is the set of all customers served by retailer k. In order to correctly allocate the portion of the standard deviation affecting each retailer, the fraction of demand served must be defined (w_{kc}), a continuous variable between 0 and 1. This is an extension of the work of You and Grossmann (2008), which considered the case of w_{kc} being binary variables.

$$\sum_t x_{kct} = w_{kc} \sum_t D_{ckt} \tag{44}$$

Then, the safety stock can be explicitly expressed by Eq. 45,

$$ss_k = z\sqrt{L}\sqrt{\sum_c w_{kc}\sigma_c^2} \tag{45}$$

The inclusion of Eq. 45 leads to a nonconvex MINLP formulation, which require for its solution global optimization methods that are limited in problem size.

4.5 Guaranteed Service Time

In the previous formulation the replenishment lead time was considered constant. The guaranteed service time approach (Graves and Willems, 2000) allows to explicitly include the quoted service time in the Net Lead Time (NLT). For a retailer k quoting a service time ST_{kc} to its customers, it is the committed time to have an order ready to be shipped. The NLT is given by Eq. 46.

$$NLT_{kc} = ST_{jk} + L_{jk} - ST_{kc} \qquad \forall j, k \tag{46}$$

The NLT for a given distribution center depends on the offered service time, the upstream service time, and the lead time. In the extreme case, when $NLT = 0$, there is no need to hold inventory. The system operates in a just-in-time mode. This term can be included in the safety stock computation (Eq. 47).

$$ss_k = z\sqrt{\sum_c w_{kc} NLT_{kc} \sigma_c^2} \qquad \forall k \tag{47}$$

In the case direct shipments from distribution centers to customers are allowed, the safety stock expression at the distribution center must also include a term to account for this demand (Eq. 48). The demand fraction of a distribution center servicing a retailer is defined by Eq. 49.

$$ss_j = z\sqrt{\sum_c w_{jc} NLT_{jc} \sigma_c^2 + \sum_k w_{jk} NLT_{jk} \sigma_k^2} \qquad \forall j \tag{48}$$

$$w_{jk} \sum_c \sum_t x_{kct} = \sum_t x_{jkt} \qquad \forall j, k \tag{49}$$

where σ_k is calculated with a weighted average of the variance of the customers being served by the retailer k.

$$\sigma_k^2 = \sum_x w_{kc} \sigma_c^2 \tag{50}$$

The service time for a customer ST_{kc}, ST_{jc} and the service time offered by the supplier ST_{ij} are normally exogenous parameters. However, the internal service time ST_{jk} can be considered as a parameter or as an optimization variable.

5 Case Study

The proposed models are designed to achieve better coordination between tactical planning and inventory management. To analyze the applicability of these formulations, a case study is optimized with each of the policies and proposed safety stock models. The case study includes a single supplier, two distribution centers, two retailers, and four customers, as seen in Fig. 10. Two of them are served exclusively by retailers while the other two can be either served by a retailer or a distribution center. The planning horizon is 30 days. Data for the case study is included in Appendix A. A single inventory policy must be established for the entire period. The problems are modeled using Pyomo (Hart et al., 2012) on an Intel i7 quad-core computer with 16 Gb RAM. Gurobi 7.5 is used to solve the mixed-integer programming models (MILP), and Baron 17.10 (Tawarmalani and Sahinidis, 2005) is used to solve the MINLP models.

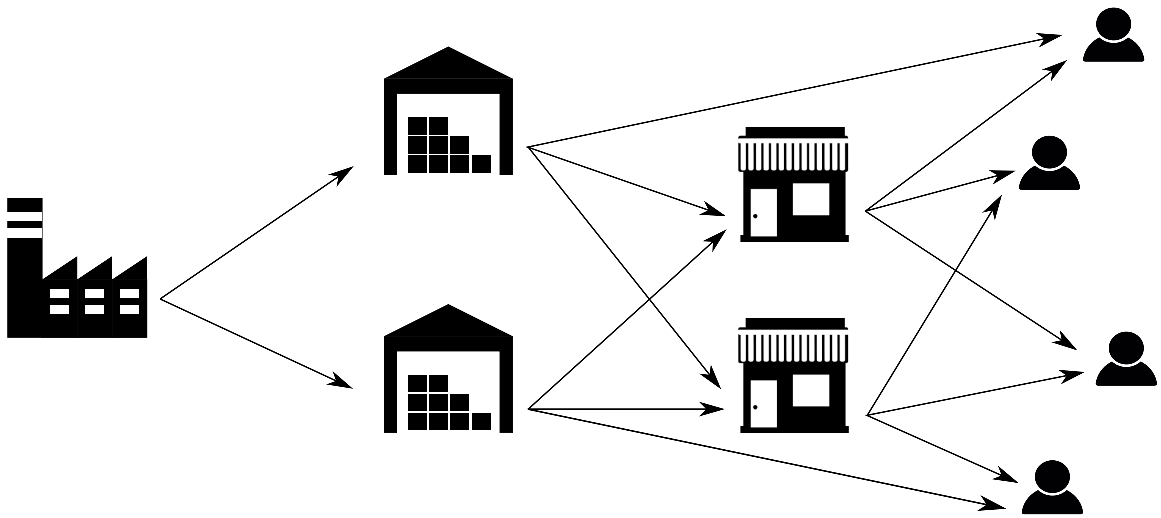

Figure 10. Supply chain planning topology for the case study

5.1 Inventory Policies

When no inventory policy is employed, the inventory curve from Fig. 11 is obtained for retailer 1, and Fig. 12 for retailer 2.

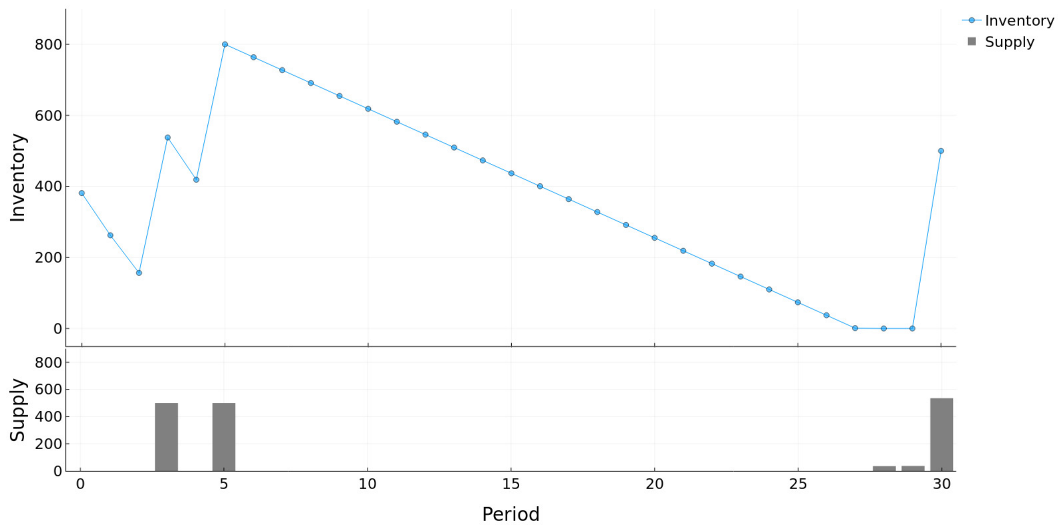

Figure 11. Inventory curve for retailer 1 without inventory policy

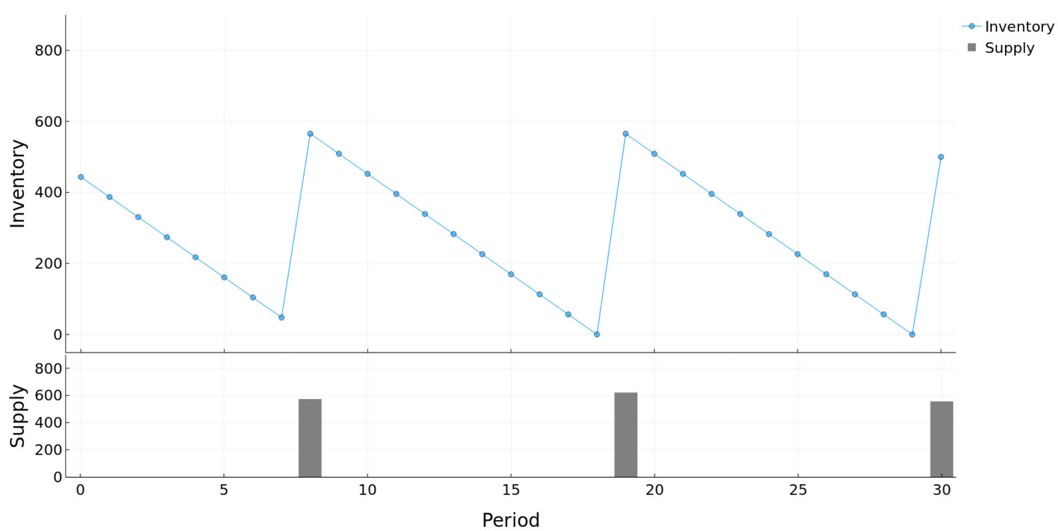

Figure 12. Inventory curve for retailer 2 without inventory policy

The supply operation for retailer 1 is uneven in amount and frequency. In periods 3 and 5, two replenishments are made, then there is no replenishment for a long part of the month, to replenish again with three orders in a row in periods 28, 29, and 30. Furthermore, the orders for periods 28 and 29 are very small. It is unlikely that an inventory manager can execute that plan in practice.

Coincidently, the plan obtained for retailer 2 is reasonable (see Fig. 12), but the result does not provide any directive in how to proceed if the demand changes.

To improve the inventory management operations a policy can be employed. The inventory curve obtained by applying the (r,Q) policy to retailer 1 is shown in Fig. 13. The replenishment orders are now the same size and correctly spaced. A replenishment order for 590 units is placed every time the inventory reaches or goes under 109 units.

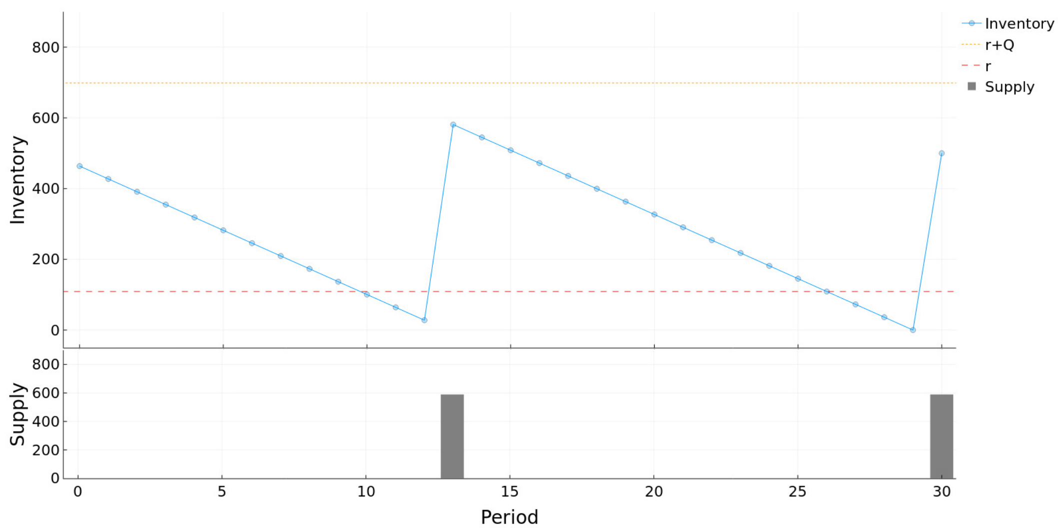

Figure 13. Inventory curve for retailer 1 with (r,Q) policy

If the replenishment frequency is constrained, an (s,S) policy is more appropriate. The inventory curve obtained for retailer 1 using this policy is shown in Fig. 14.

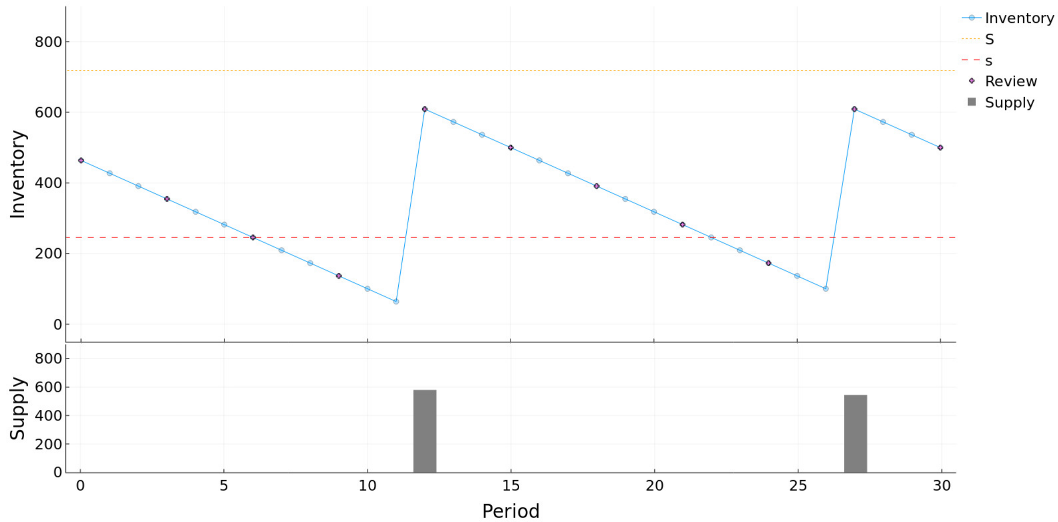

Figure 14. Inventory curve for retailer 1 with (s,S) policy

The policy obtained indicates that the inventory must be reviewed every 3 days. If the inventory is found to be under 245 units, a replenishment order to bring the inventory back to 718 units is placed. In the example, the first replenishment order is for 580 units, and the second for 545 units, because the inventory level at period 9 was lower than the inventory level at period 24.

The policy parameters obtained for all facilities are presented in Table 2. For the DC 2, the value of r obtained using the (r,Q) policy is close to zero because only one replenishment is done during the entire horizon.

Table 2. Policy parameters for all the facilities

	(r,Q) policy		(s,S) policy		
	r	Q	s	S	Frequency
DC 1	925	1,243	852	1,134	2
DC 2	0	3,089	345	1,304	2
Retailer 1	109	590	246	718	3
Retailer 2	170	500	330	839	3

The model statistics and results are shown in Table 3. All models are MILP, solved with Gurobi 7.5 with a 0.5% gap.

Table 3. Computational statistics of the models for inventory policies

	none	(r,Q)	(s,S)
Constraints	767	1,347	1,935
Variables	1,106	1,238	1,546
Binaries	155	279	583
Objective (k\$)	4,559	4,542	4,620
Solution time (s)	0.1	71.4	65.4

5.2 Safety Stock

In several of the examples shown in the previous section the planned inventory reached a zero or close to zero level. Because the demand is uncertain allowing such low inventories can lead to stockouts. To prevent stockouts from happening a safety stock is considered. The inventory curves obtained with proportional and piecewise safety stock without inventory policy are shown in Fig. 15. Figure 16 features the inventory curves using explicit and guaranteed safety stock formulations, also without inventory policy.

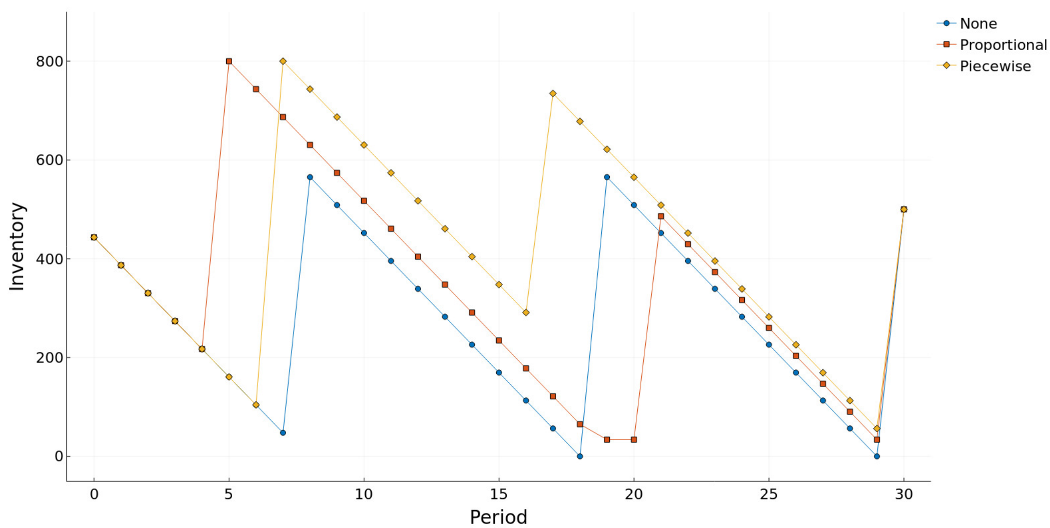

Figure 15. Inventory curves using MILP formulations for safety stock, ("None" indicates the base case)

As it can be seen in the curves, the MINLP formulations in Fig. 16 give similar results in the safety stock level, and are more conservative than the result obtained with MILP formulations. The piecewise-linear formulation in Fig. 15 gives a higher safety stock for the retailers than the proportional formulation. The opposite is true for the distribution centers. Both models have

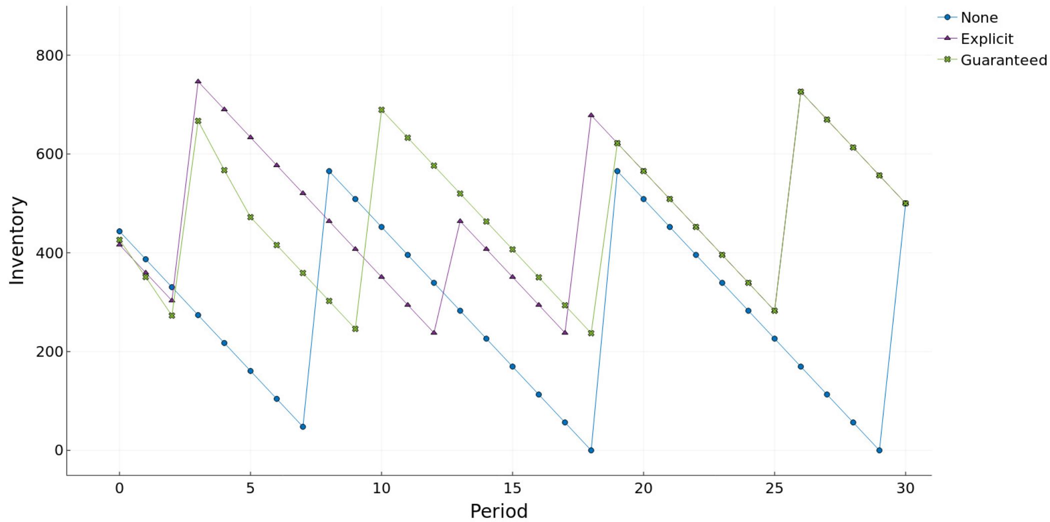

Figure 16. Inventory curves using MINLP formulations for safety stock

adjustable parameters to change the risk level. The safety stocks obtained for all facilities are presented in Table 4, and the computational statistics in Table 5.

Table 4. Safety stock values obtained for all facilities with each formulation

	Proportional	Piecewise	Explicit	Guaranteed
DC 1	462.3	147.1	337.3	345.7
DC 2	240.7	121.8	427.2	436.0
Retailer 1	21.8	36.3	165.4	165.4
Retailer 2	33.9	56.3	237.9	237.2

As seen in Table 4 the MILP models yield low safety stock values for the retailers and larger for the distribution centers. Since these models relate safety stock to throughput, DCs require a larger safety stock. The MILP formulations (proportional and piecewise), yield lower values of safety stock than their MINLP counterparts (explicit and guaranteed). However, as mentioned before, the risk level at the MILP policies can be easily adjusted. Furthermore, the proportional policy gives similar result to the piecewise safety stock formulation for the retailers and very different for the DCs. The consideration of the risk-pooling effect makes the facilities with higher throughput hold a lower safety stock than the proportional case. This indicates that the proportional formulation can be good enough for single-echelon systems, but for multi-echelon systems a piecewise formulation might be more appropriate. The explicit and guaranteed formulations gave very similar results for the example. The effect of considering the net lead time is more important in systems with larger lead times, which is not the case of the case study.

Table 5. Computational statistics of the models for safety stock to reach 0.5% gap

	None	Proportional	Piecewise	Explicit	Guaranteed
Constraints	767	771	1,351	785	785
Variables	1,106	1,110	1,242	1,120	1,126
Binaries	155	155	155	155	155
Objective (k\$)	4,517	4,511	4,499	4,526	4,531
Solution time (s)	0.1 s	0.1 s	0.1 s	1.8 s	2.5 s
Model type	MILP	MILP	MILP	MINLP	MINLP
Solver	Gurobi	Gurobi	Gurobi	Baron	Baron

Since the safety stocks obtained using the MINLP formulations were significantly larger than the MILP formulations, the objective value (cost) is also larger. They also have an important impact in the solution time. Even for a small example, the MINLP formulations take longer to solve, although the times are actually quite reasonable. On the other hand, the proportional and piecewise formulations take less time to solve and might be better suited for larger problems.

5.3 Models with Safety Stock and Inventory Policies

Inventory policies can be combined with safety stock formulations to yield optimal supply chain plans with clear inventory management policies that hedge against stockouts. The inventory curves obtained from using (s,S) policy together with a piecewise safety stock formulation are shown in Fig. 17 for retailer 1, and in Fig. 18 for retailer 2.

The consideration of safety stock increases the computational burden of the model. However, the inclusion of an inventory policy does not have the same impact. The combination of a safety stock formulation with an inventory policy verifies the same observation. The models combining a MINLP safety stock formulation with an inventory policy could not be solved in under 1 hour, which limits their scope of application. Table 6 summarizes the solution times for the combination of (s,S) policy with piecewise safety stock.

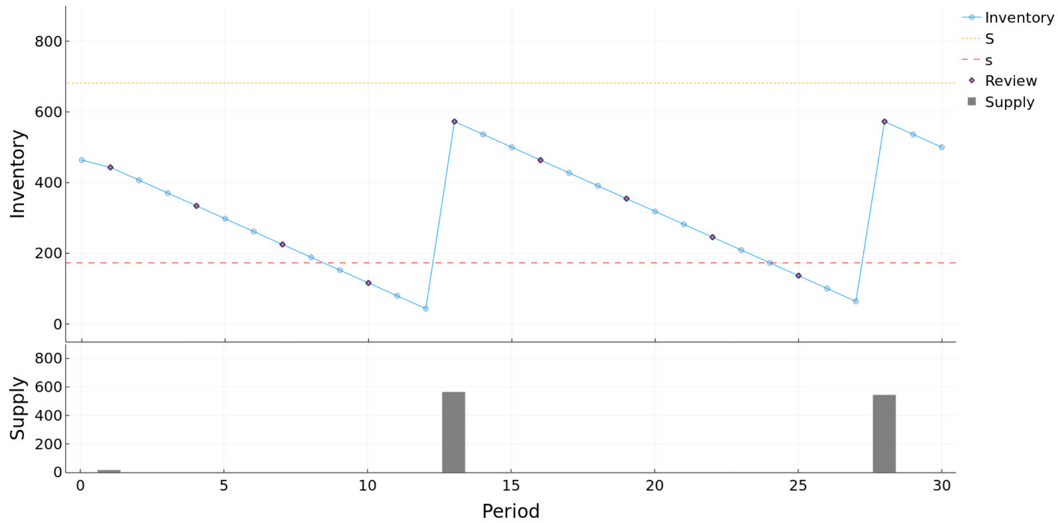

Figure 17. Inventory curve for retailer 1 with (s,S) policy with piecewise safety stock

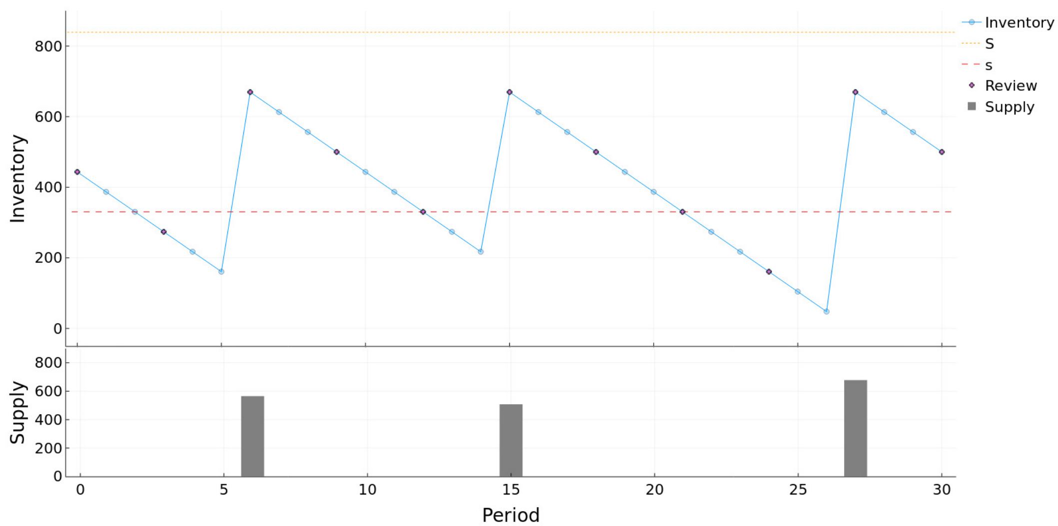

Figure 18. Inventory curve for retailer 2 with (s,S) policy with piecewise safety stock

6 Simulation of the Supply Chain

The inventory curves of the model are meant to be used as a guideline for the monthly plan. At the same time, the policies provide specific instructions on how to actually manage the warehouse. To assess the quality of the models proposed, the process was simulated in Python using each of the policies and safety stock formulations. For each simulation, 1,000 runs were executed with demands randomly generated from the probability distribution of demand of the example. Inventory policy parameters and a set of prioritized suppliers is obtained from the planning model. When inventory is not available at the primary location for replenishment the secondary supplier is used. The

Table 6. Solution times for combinations of piecewise safety stock and (s,S) inventory policy

Policy	Safety Stock	Solution time (s)
none	none	0.1
none	piecewise	0.1
(s,S)	none	65
(s,S)	piecewise	73

service level is measured as the percentage of customer orders that can be fulfilled on time. The results of the simulation are presented in Table 7. The target service level at the time of optimizing the system was 95%.

Table 7. Results of the supply chain simulation with and without safety stocks. Average service level (SL) and standard deviation (σ)

	(r,Q)		(s,S)	
	SL	σ	SL	σ
None	94.8 %	0.027	93.8 %	0.035
Proportional	95.2 %	0.027	95.1 %	0.032
Piecewise	95.5 %	0.027	96.9 %	0.025
Explicit	98.1 %	0.018	98.8 %	0.016
Guaranteed	98.4 %	0.017	99.0 %	0.015

When no safety stock is considered, the plan does not reach the target service level. With proportional and piecewise safety stock formulations the target service level is reached, and surpassed in the case of the piecewise formulation with (s,S) policy. However, the standard deviation for that case is also lower than the proportional formulation, making it more consistent in achieving the required service level. Using the MINLP formulations the target service level was exceeded, even though the target was explicitly considered through the z value in the formulation. The result is a worse objective function and an overly conservative plan.

7 Conclusions

Decisions related to amounts and timings of inventories are critical to have a responsive and efficient supply chain. To minimize cost, optimization models are employed to prescribe optimal stock plans. The supply chain planning models have been improved with the inclusion of inventory policies to make them usable in real-world applications. Models for two of the main inventory policies were proposed. The (r,Q), a continuous review policy, was adapted to a discrete time forumulation relaxing the requirement of having the inventory level exactly at r to trigger a replenishment. The derivation of the (s,S) policy is general enough to accommodate other custom policies. The computational efficiency of models including the (s,S) or the (r,Q) policy are similar. As expected, the simulation results favor the (r,Q) policy because continuously reviewing the inventory allows to react faster against increases in the demand. Both inventory policies bring a significant computational burden to the base model. However, they also combine two decision levels: tactical and operational. The combination of inventory planning with inventory policies achieves the integration of tactical with operational decisions for distribution centers, warehouses, and retailers. The models presented yield optimal planning solutions with clear management policies to manage the warehouse inventory. They were put in practice with the simulation case study with results meeting the target service levels.

Additionally, the different options to address uncertainty through safety stock have been considered and extended. From the simulation results it is possible to conclude that the proportional formulation yield good results for single-echelon systems, and that the piecewise-linear formulation yields good results for both single and multiple echelon systems. The MINLP formulations, explicit risk-pooling and guaranteed service time, give conservative solutions and take a long time to solve. Because of the significant computational burden added by MINLP formulations, either a proportional or a piecewise-linear-formulation are recommended. The piecewise-linear is preferred when the risk-pooling effect is specified.

This research takes an important step into bringing inventory management theory closer to mathematical programming. Several additional concepts from inventory management could be considered in the future to develop richer supply chain planning models.

References

Arrow, K. J., Karlin, S., and Scarf, H. (1958). The nature and structure of inventory problems. *Studies in the mathematical theory of inventory and production. Stanford University Press, Stanford, Ca.*

Bashyam, S. and Fu, M. C. (1998). Optimization of (s, s) inventory systems with random lead times and a service level constraint. *Management Science*, 44(12-part-2):S243–S256.

Beale, E. M. L. and Tomlin, J. A. (1970). Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. *Operations Research*, 69(99):447–454.

Bradley, J. R. and Arntzen, B. C. (1999). The simultaneous planning of production, capacity, and inventory in seasonal demand environments. *Operations Research*, 47(6):795–806.

Brunaud, B., Bassett, M. H., Agarwal, A., Wassick, J. M., and Grossmann, I. E. (2017). Efficient formulations for dynamic warehouse location under discrete transportation costs. *Computers & Chemical Engineering*.

Cafaro, D. C. and Grossmann, I. E. (2014). Alternate approximation of concave cost functions for process design and supply chain optimization problems. *Computers & Chemical Engineering*, 60:376–380.

Daskin, M. S., Coullard, C. R., and Shen, Z.-J. M. (2002). An inventory-location model: Formulation, solution algorithm and computational results. *Annals of Operations Research*, 110(1):83–106.

Diabat, A. and Theodorou, E. (2015). A location–inventory supply chain problem: Reformulation and piecewise linearization. *Computers & Industrial Engineering*, 90:381–389.

Enke, S. (1958). On the economic management of large organizations: a laboratory study. *The Journal of Business*, 31(4):280–292.

Eppen, G. D. (1979). Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science*, 25(5):498–501.

Eppen, G. D. and Martin, R. K. (1988). Determining safety stock in the presence of stochastic lead time and demand. *Management Science*, 34(11):1380–1390.

Federgruen, A. and Zheng, Y.-S. (1992). An efficient algorithm for computing an optimal (r, q) policy in continuous review stochastic inventory systems. *Operations Research*, 40(4):808–813.

Fisher, M. L. (1985). An applications oriented guide to Lagrangian relaxation. *Interfaces*, 15(2):10–21.

Galliher, H., Morse, P. M., and Simond, M. (1959). Dynamics of two classes of continuous-review inventory systems. *Operations Research*, 7(3):362–384.

Garcia-Herreros, P., Agarwal, A., Wassick, J. M., and Grossmann, I. E. (2016). Optimizing inventory policies in process networks under uncertainty. *Computers & Chemical Engineering*, 92:256–272.

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

Hart, W. E., Laird, C., Watson, J.-P., and Woodruff, D. L. (2012). *Pyomo-optimization modeling in python*, volume 67. Springer.

Jackson, J. R. and Grossmann, I. E. (2003). Temporal decomposition scheme for nonlinear multisite production planning and distribution models. *Industrial & Engineering Chemistry Research*, 42(13):3045–3055.

Kapuscinski, R. and Tayur, S. (1999). Optimal policies and simulation-based optimization for capacitated production inventory systems. In *Quantitative Models for Supply Chain Management*, pages 7–40. Springer.

Lim, M.-F. and Karimi, I. (2003). Resource-constrained scheduling of parallel production lines using asynchronous slots. *Industrial & Engineering Chemistry Research*, 42(26):6832–6842.

Linderoth, J. (2017). Mixed-integer (linear) programming: Recent advances, and future research directions. *In Proceedings of FOCAPO 2017*.

Miranda, P. A. and Garrido, R. A. (2009). Inventory service-level optimization within distribution network design problem. *International Journal of Production Economics*, 122(1):276–285.

Platt, D. E., Robinson, L. W., and Freund, R. B. (1997). Tractable (q, r) heuristic models for constrained service levels. *Management Science*, 43(7):951–965.

Raman, R. and Grossmann, I. E. (1994). Modelling and computational techniques for logic based integer programming. *Computers & Chemical Engineering*, 18(7):563–578.

Relvas, S., Matos, H. A., Barbosa-Povoa, A. P. F., Fialho, J., and Pinheiro, A. S. (2006). Pipeline scheduling and inventory management of a multiproduct distribution oil system. *Industrial & Engineering Chemistry Research*, 45(23):7841–7855.

Shen, Z.-J. M., Coullard, C., and Daskin, M. S. (2003). A joint location-inventory model. *Transportation Science*, 37(1):40–55.

Tawarmalani, M. and Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103(2):225–249.

Varma, V. A., Reklaitis, G. V., Blau, G. E., and Pekny, J. F. (2007). Enterprise-wide modeling & optimizationan overview of emerging research challenges and opportunities. *Computers & Chemical Engineering*, 31(5):692–711.

You, F. and Grossmann, I. E. (2008). Mixed-integer nonlinear programming models and algorithms for large-scale supply chain design with stochastic inventory management. *Industrial & Engineering Chemistry Research*, 47(20):7802–7817.

You, F. and Grossmann, I. E. (2010). Integrated multi-echelon supply chain design with inventories under uncertainty: MINLP models, computational strategies. *AIChE Journal*, 56(2):419–440.

Zheng, Y.-S. and Federgruen, A. (1991). Finding optimal (s, s) policies is about as simple as evaluating a single policy. *Operations Research*, 39(4):654–665.

Appendix

A Data for Case Study

Table 8. Transportation costs

	DC1	DC2	Ret1	Ret2	Cus1	Cus2	Cus3	Cus4
DC1	-	0.003	0.002	0.003	0.287	0.288	-	-
DC2	0.065	-	0.043	0.052	0.287	0.302	-	-
Ret1	-	-	-	0.101	0.284	0.299	0.0001	-
Ret2	-	-	0.025	-	0.287	0.287	-	0.0001

Table 9. Warehouse and customer data

	HC	*InitInv*	*MaxInv*	*L*		Daily Demand	σ^2	ST
DC1	0.0165	2,000	3,000	7	Cus1	245.06	3,455	1
DC2	0.1	2,000	3,000	7	Cus2	82.5	2,627	1
Ret1	0.1	500	800	3	Cus3	56.53	6,814	0
Ret2	0.1	500	800	3	Cus4	36.32	3,349	0

For proportional safety stock $\beta = 0.2$

Table 10. Piecewise safety stock breaking points

SSx	0	1,564	3,128	4,692	6,516	13,033
$SSy\ DC1$	0	71	120	138	147	147
$SSy\ DC2$	0	71	120	138	147	147
$SSy\ Ret1$	0	30	51	59	63	63
$SSy\ Ret2$	0	30	51	59	63	63