Optimal FCFS allocation rules for periodic-review assemble-to-order systems

Citation for published version (APA):
Huang, K., & Kok, de, A. G. (2015). Optimal FCFS allocation rules for periodic-review assemble-to-order systems. *Naval Research Logistics*, *62*(2), 158-169. https://doi.org/10.1002/nav.21620

DOI:
10.1002/nav.21620

Document status and date:
Published: 01/01/2015

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Optimal FCFS Allocation Rules for Periodic-Review Assemble-To-Order Systems

Kai Huang,[1] Ton de Kok[2]

[1] *Operations Management, DeGroote School of Business, McMaster University, Ontario, Canada*

[2] *Operations, Planning, Accounting, and Control (OPAC), School of Industrial Engineering, Eindhoven University of Technology, The Netherlands*

Received 25 November 2012; revised 12 February 2015; accepted 13 February 2015
DOI 10.1002/nav.21620
Published online in Wiley Online Library (wileyonlinelibrary.com).

Abstract: In Assemble-To-Order (ATO) systems, situations may arise in which customer demand must be backlogged due to a shortage of some components, leaving available stock of other components unused. Such unused component stock is called remnant stock. Remnant stock is a consequence of both component ordering decisions and decisions regarding allocation of components to end-product demand. In this article, we examine periodic-review ATO systems under linear holding and backlogging costs with a component installation stock policy and a First-Come-First-Served (FCFS) allocation policy. We show that the FCFS allocation policy decouples the problem of optimal component allocation over time into deterministic period-by-period optimal component allocation problems. We denote the optimal allocation of components to end-product demand as multimatching. We solve the multi-matching problem by an iterative algorithm. In addition, an approximation scheme for the joint replenishment and allocation optimization problem with both upper and lower bounds is proposed. Numerical experiments for base-stock component replenishment policies show that under optimal base-stock policies and optimal allocation, remnant stock holding costs must be taken into account. Finally, joint optimization incorporating optimal FCFS component allocation is valuable because it provides a benchmark against which heuristic methods can be compared. © 2015 Wiley Periodicals, Inc. Naval Research Logistics 62: 158–169, 2015

Keywords: assemble-to-order; periodic-review; cost minimization; mathematical programming; stochastic programming

1. INTRODUCTION

Assemble-To-Order (ATO) systems are increasingly important in contemporary manufacturing as they facilitate mass customization without a need for high inventory capital investments, which can create a competitive advantage with regard to customer service and profitability [3, 10]. For example, delayed differentiation allows for assembly of final products within the customer lead times thereby eliminating finished goods stock, that is, inventory consists only of components. When orders arrive, components are withdrawn from inventory and assembled into final products. ATO systems exploit component commonality, which leads to risk pooling among the component inventory [14]. However, in ATO systems there are also the potential of component mismatches and component allocation issues when different products share common components. To optimize system performance, the decision problems of inventory replenishment and component allocation must both be addressed. In [2], it is shown that both types of decisions have similar impacts on the aggregate fill rate in a periodic-review ATO system. Therefore, in the context of ATO systems, we face a joint optimization problem. Unfortunately, the joint optimization problem is hard and the optimal policy for jointly determining inventory replenishment and component allocation is unknown, except for specific small scale systems (cf. [11, 16]).

The literature concerning the optimization of ATO systems can be classified according to three aspects. First, with respect to inventory review moments, we have single period (cf. [8]), periodic-review (cf. [22, 1, 5, 2]) and continuous-review (cf. [20, 12, 23]) models. Second, with respect to the objectives of the optimization problem, the literature may be partitioned in terms of service level (cf. [22, 1, 2, 20, 13]), or cost minimization (cf. [8, 5, 12]) goals. Thirdly, the decisions in the optimization problem involve either inventory replenishment optimization (cf. [22, 1, 5, 20, 12, 23]), or joint optimization of inventory replenishment and component allocation (cf. [8, 2]).

With regard to the continuous and periodic-review literature, there are very few cost minimization models that incorporate both inventory holding and order-based backlogging costs. For example, in [5], inventory costs are minimized

Correspondence to: Kai Huang (khuang@mcmaster.ca)
Additional Supporting Information may be found in the online version of this article.

given customer service level constraints, while in [12], an order-based backorder cost minimization model is proposed and compared to a classical newsvendor model with item-based backorder costs. In [17], inventory holding costs are incorporated into the objective, but there is no backlogging cost, and the focus of the study is on the asymptotic behavior of the system under high order volume. Dogru et al. [7] used a two-stage stochastic program to provide an inventory replenishment policy whose approximation results in a lower bound on the total inventory cost. Importantly, they found that the optimal component allocation policy obtained for a special W system violated the FCFS principle. van Jaarsveld and Scheller-Wolf [21] developed an algorithm for the joint optimization of inventory replenishment and component allocation, and studied the performance of this algorithm for industrial-scale ATO systems. Nevertheless, all these studies assume continuous review ATO systems. As such, the analysis cannot be directly applied to periodic-review ATO systems.

The difficulty of accounting for the operational cost of inventory holding is associated with the challenge of evaluating *remnant stock*. As first discussed in [1], remnant stock refers to components that are allocated to certain products according to the First-Come-First-Served (FCFS) rule, but can not immediately be assembled into a final product, because complementary components are missing. In [6] the concept of an ideal ATO system is introduced, and remnant stock is explicitly analyzed and accurate approximations for the long-run average remnant inventories are obtained. In [2], it is shown that base-stock level optimization and component allocation optimization have a similar impact on increasing the order fill rate. In [17], an infinite horizon model is proposed where a periodic-review policy is shown to be the asymptotically optimal component rationing policy, even though it ignores within-period order sequencing and myopically minimizes the single period cost.

In this article, we consider an ATO system consisting of multiple end-products that are assembled to order from multiple components. We assume that end-product demand is stochastic and periodic, that is, at the end of each period end-product demand for that period reveals itself (cf. [1]). We assume that components are replenished according to an installation stock policy. More precisely, the component replenishment policy uses its inventory position as state variable that must be controlled. The inventory position of a component is defined as the sum of outstanding replenishment orders and net stock. The net stock equals physical inventory minus backorders. In the context of ATO systems, we need to be careful concerning the definition of component backorders. Each component is controlled in isolation based on its inventory position, as if in a single-item single-echelon system. Component backorders are defined from this single-item single-echelon perspective. In the remainder of this article, we use the term inventory-position-based policy instead of installation stock policy, as the usage of the inventory position as state for the component inventory replenishment policy is the key to the results derived. Note that inventory-position-based replenishment policies include (R, S), (R, s, S), and (R, s, nQ)-policies (cf. [19]). We assume that backlogged demand is filled on a FCFS basis. We assume linear holding costs for components and linear penalty costs for end-products. The penalty costs for an end-product are linear in the amount backordered of this end-product. We allow for partial backorders and equivalently partial fulfillment of end-product demand, that is, we allow that end-product demand in a particular period is satisfied over multiple periods in time. We assume that (part of) end-product demand is fulfilled as soon as the necessary components have been allocated to it. Under this convention the lead time of the final assembly process is irrelevant for our analysis. This also implies that we can assume without loss of generality that the assembly lead time is zero. This convention is helpful when formulating the model in Section 2.

It is important to note here that we do not make any specific assumption regarding the demand process when exploring the problem of optimally allocating components to end-products, other than that demand reveals itself at the end of the period. Assumptions on the demand process do become important when determining the optimal inventory-position-based replenishment policies. In our formulation of the joint optimization of replenishment and allocation policies in Section 3, and in our experimental study in Section 4, we assume i.i.d. demands from different periods and base-stock component replenishment policies.

Our contributions to the existing literature are as follows. First, we extend the concept of matching in [15], which refers to the coupling of a single supply unit and a single customer. We consider the coupling of multiple component units and end-product (demand) units, called multimatching. The concept of multimatching allows us to characterize the remnant holding cost exactly. Second, we explore the inherent structure of the mathematical model underlying multimatching. In particular, we show that under the FCFS rule, the component allocation problem can be solved period by period in the form of a Linear Program (LP) for continuous demand, and in the form of an Integer Linear Program (ILP) for discrete demand. In the context of ATO systems it is appropriate to assume that end-product demand is discrete. We derive an iterative algorithm to efficiently solve the component allocation problem. Third, we investigate the joint optimization of the base-stock levels and the component allocation policy, modeled as a two-stage stochastic integer nonlinear program. We propose an approximation scheme that ignores the remnant stock holding cost. This formulation is much easier to solve, and provides lower and upper bounds on the optimal cost value. Fourth, in numerical experiments we show that the optimal FCFS allocation policy yields a significant improvement over existing FCFS heuristics.

The remainder of the paper is organized as follows. In Section 2, we formulate the component allocation problem as a deterministic optimization problem. In Section 3, the joint optimization problem is formulated as a two-stage stochastic program, where simple upper and lower bounds on the optimal objective are provided. Section 4 presents findings from numerical experiments. Our conclusions and future research ideas are presented in Section 5.

2. THE ATO MODEL

We consider a general periodic-review ATO system consisting of m components and n products. We define $\mathcal{M}$ as the set of all components $1, 2, \cdots, m$, and $\mathcal{N}$ as the set of all products $1, 2, \cdots, n$. We have summarized further relevant notation in Table 1. The Bill Of Materials (BOM) is defined by the matrix $(a_{i,j})$, where $a_{i,j}$ equals the number of components i needed to produce one unit of end-product j. Components are ordered at the start of each period according to an inventory-position-based policy. As stated above, the inventory position is defined as the sum of net stock and all outstanding replenishment orders, where the net stock is defined as the physical stock minus backorders. The replenishment lead time of component i equals L_i periods, but the production (final assembly) lead time is zero. We assume linear holding costs for components. Demand $P_{j,t}$ for end-product j in period t reveals itself at the end of the period; this implies demand $D_{i,t} = \sum_{j \in \mathcal{N}} a_{i,j} P_{j,t}$ for component i. Unless stated otherwise we do not make any assumption regarding the demand process $\{P_{j,t}\}$. End-product demand is produced as soon as all required and allocated components are available. In case one or more components do not have sufficient availability, customer orders are backlogged and linear backlog costs are incurred. We allow partial fulfillment of demand $P_{j,t}$ over subsequent periods, yet component demand must be satisfied and allocated to product demand on a FCFS basis. This implies that, even though demand for component i_0 from demand $P_{j_0,t}$ for some end-product j_0 could be filled with physically available component inventory, it will only be served after component availability over time resulting from replenishments triggered by $D_{i_0,s}, \forall s < t$, has been allocated to satisfy product demand before $t, \forall s < t$. The FCFS assumption together with the general BOM matrix $(a_{i,j})$ leads to the possibility of remnant stocks. Remnant stock is stock allocated to demand that cannot yet be produced due to the lack of availability of one or more components. The type and amount of remnant stock and its associated holding costs, as well as the type and amount of demand backlogged, depend on how component availability is allocated to demands of various products over time. Whereas in single item single location systems such an allocation is trivial, this is not the case for a general ATO system. This allocation of component availability is denoted as *multimatching* as we try to match demand for multiple products with availability of multiple components over time.

Table 1. ATO system notation

t	index of periods. Period t is defined as the duration [t,t+1); $t \geq 0$.
i	index of components, where $i \in \mathcal{M} = \{1, \cdots, m\}$.
j	index of products, where $j \in \mathcal{N} = \{1, \cdots, n\}$.
L_i:	lead time of component i.
S_i:	base-stock level of component i.
L	$\text{Max}_{i \in \mathcal{M}} L_i$, maximum lead time.
$a_{i,j}$	the number of component i used in each unit of end-product j.
h_i:	unit inventory holding cost for component i.
b_j:	unit backlogging cost for end-product j.
$P_{j,t}$	demand for end-product j in period t.
$D_{i,t}$	demand for component i in period t; $D_{i,t} = \sum_{j \in \mathcal{N}} a_{i,j} P_{j,t}$.
$X_{i,t}$	net stock of component i at the end of period t.
$O_{i,t}$	replenishment order of component i placed at the start of period t.
$A_{i,t}$	replenishment order of component i arriving at the start of period t.
$Y_{i,t}$	inventory position of component i at the start of period t (after event 1).

2.1. Sequence of Events

As we consider a periodic-review system, we must carefully define the sequence of events during a period. At the start of period t we have the following sequence of events:

1. Order $O_{i,t}$ is placed for all i.
2. Order $A_{i,t} = O_{i,t-L_i}$ is received for all i.

At the end of period t we have the following subsequent events:

3. Demand $D_{i,t}$ reveals itself.
4. Determine order $O_{i,t+1}$ to be placed at the start of period $t+1$.
5. Allocate component availability to $P_{j,s}, s \leq t$.
6. Satisfy (partially) end-product demand for which all required components are available.
7. Incur penalty and holding costs.

As stated above, we assume that the production lead time is zero. Thus we state that end-product demand, to which components are allocated at the end of period t in step 5, is satisfied at time t. The costs charged in step 7 only consider the customer demands received until the end of period t that have not been satisfied in step 6. Furthermore, end-product demand that uses components received in step 2 at the start of period $t + 1$, is assumed to be satisfied in step 6 at the end of period t. This is because of the fact that the replenishment quantities received at the start of period $t + 1$ are known at the end of period t. To enable the multimatching technique, we need to specify a condition on the inventory-position-based policies that ensures that the maximum waiting time until demand for component i is satisfied equals L_i.

CONDITION 1: The inventory-position-based component replenishment policies satisfy

$$Y_{i,t+1} \geq 0, \quad \forall i \in \mathcal{M}, \forall t \geq 0.$$

The above condition states that after ordering a component its inventory position is non-negative. This holds for base-stock policies with non-negative base-stock levels and for reorder point policies with non-negative reorder points.

2.2. Multimatching

In this section, we define the multimatching problem associated with an arbitrary period t. Due to the FCFS assumption and the assumption that component orders for period $t + 1$ are immediately determined after demands $P_{j,t}$ reveal themselves (event 3 and 4), we have the following component net stock $X_{i,t}$ immediately after demand revelation,

$$X_{i,t} = Y_{i,t-L_i} - D_i[t - L_i, t], \quad \forall i \in \mathcal{M},$$

where $Y_{i,t-L_i}$ is the inventory position at the start of period $t - L_i$ (after event 1); and $D_i[s,t] = \sum_{u=s}^{t} D_{i,u}$ for $s \leq t$.

Now we distinguish between two cases, viz., demand $P_{j,t}, \forall j \in \mathcal{N}$, can be satisfied immediately, or not. The first situation occurs when

$$X_{i,t} \geq 0, \quad \forall i \in \left\{k : a_{k,j} > 0\right\}. \tag{1}$$

In this case, we do not incur backlog costs for end-product j. However, in case Eq. (1) does not hold, we face an allocation problem. In what follows we show that under the FCFS assumption and inventory-position-based component replenishment policies satisfying Condition 1, this allocation problem can be formulated as an LP. The main idea behind the derivation of this result is that the two assumptions mentioned above (Condition 1 and FCFS) separate the allocation problems in different periods. Firstly, the allocation problem to be solved in period t is not affected by demands in periods $s > t$ due to the FCFS assumption. Secondly, by Condition 1, the demand in period t is satisfied from component replenishment orders known after event 4 in period t, whereby the allocation problem becomes a deterministic problem.

The FCFS assumption implies that we can assume without loss of generality that $P_{j,s} = 0, \forall j \in \mathcal{N}, \forall s > t$, when solving the allocation problem at the end of period t. Let $X^0_{i,t+s}$ be defined as

$$\begin{aligned} X^0_{i,t+s} := {} & \text{net stock of component } i \text{ at the end of} \\ & \text{period } t+s, \text{ assuming no demand} \\ & \text{occurs after time } t. \end{aligned}$$

Using this assumption it follows that the inventory balance equation for periods $t+s$, $s > 0$ can be formulated as follows,

$$X^0_{i,t+s} = X_{i,t} + \sum_{u=1}^{s} A_{i,t+u}, s \geq 1.$$

Now note that at the start of period $t + 1$ the inventory-position-based replenishment policy generates a, possibly zero, replenishment order $O_{i,t+1}$ that arrives at the start of period $t + L_i + 1$. Thus we find that

$$\begin{aligned} Y_{i,t+1} &= Y_{i,t} + O_{i,t+1} - D_{i,t} \\ &= Y_{i,t} + A_{i,t+L_i+1} - D_{i,t} \\ &= X_{i,t} + \sum_{s=1}^{L_i} A_{i,t+s} + A_{i,t+L_i+1} \\ &= X_{i,t} + \sum_{s=1}^{L_i+1} A_{i,t+s} \\ &= X^0_{i,t+L_i+1}. \end{aligned}$$

We emphasize here that $X^0_{i,t+L_i+1}$ equals the cumulative availability of component i immediately after arrival of order $O_{i,t+1}$. This availability can be used to allocate component i to satisfy demand at time t. From the above equality and Condition 1 it follows that $X^0_{i,t+L_i+1} \geq 0$. Now note that $Y_{i,t+1}$ represents the inventory position of component i after ordering at time t, which is already known after event 4 of period t, that is, after revelation of demand $D_{i,t}$. This and the assumption that no demand occurs after time t implies that $D_{i,t}$ has been satisfied at or before time $t + L_i$. This reasoning results into the following Theorem:

THEOREM 1: Under Condition 1 the following statements hold:

1. $D_{i,t}$ can be satisfied before or in period $t + L_i$ for all $i \in \mathcal{M}$ and $t \geq 0$.
2. All component orders from which demand $P_{j,t}$ is satisfied are known at time t before component allocation, $\forall t \geq 0, \forall j \in \mathcal{N}$.

3. $P_{j,t}$ can be satisfied before or at time t+L, $\forall t \geq 0, \forall j \in \mathcal{N}$.

With respect to property 1 in Theorem 1 we note that under Condition 1, the worst case is that component demand is satisfied by the order it triggers. Indeed, since this order arrives L_i periods later, the worst case is that demand in period t for component i is satisfied in period $t + L_i$.

From Theorem 1, we derive the following corollary concerning base-stock policies:

COROLLARY 1: If component inventory is controlled by base-stock policies with $S_i \geq 0, \forall i \in \mathcal{M}$, then

1. All component orders from which demand $P_{j,t}$ is satisfied are known at time t before component allocation, $\forall t \geq 0, \forall j \in \mathcal{N}$.
2. $P_{j,t}$ is satisfied before or at time $t+L$, $\forall t \geq 0, \forall j \in \mathcal{N}$.

With the above result, we can formulate the component allocation problem at the end of period t, as a deterministic problem, since all relevant information is known.

The decisions to be taken at time t to satisfy demand $P_{j,t}$ concern allocation of component availability at the end of periods $t, t+1, \cdots, t+L$. Define

$$x_{j,k}^{t} := \text{amount of end} - \text{product demand } P_{j,t} \text{ satisfied at time } t+k, \quad j \in \mathcal{N}, 0 \leq k \leq L, t \geq 0.$$

The decision variables $x_{j,k}^{t}$ ($t \geq 0$) must satisfy the following constraints,

$$\sum_{k=0}^{L} x_{j,k}^{t} = P_{j,t}, j \in \mathcal{N} \tag{2}$$

$$\sum_{u=0}^{k} \sum_{j \in \mathcal{N}} a_{i,j} x_{j,u}^{t} \leq \text{Max}\left\{0, X_{i,t} + D_{i,t} + \sum_{u=1}^{k+1} A_{i,t+u}\right\}, \quad i \in \mathcal{M}, 0 \leq k \leq L_i \tag{3}$$

$$\sum_{u=0}^{k} \sum_{j \in \mathcal{N}} a_{i,j} x_{j,u}^{t} \leq \text{Max}\left\{0, X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u}\right\}, \quad i \in \mathcal{M}, L_i < k \leq L \tag{4}$$

Equation (2) ensures that demand for end-product j is satisfied at or before time $t+L$. Inequality (3) states that the total amount of demand for component i in period t, satisfied until period $t+k$ ($0 \leq k \leq L_i$), cannot exceed the cumulative availability of component i at time $t+k$ for satisfying demand from period t. This cumulative availability may be negative in case of backorders at time $t+k$ due to end-product demand before time t using component i. The imposed nonnegativity of $x_{j,k}^{t}$ requires that we take the maximum of 0 and component i availability at time $t+k$. Inequality (4) is similar to inequality (3): it expresses the fact that at time $t + L_i$ all component i replenishment orders needed to satisfy component i demand from end-product demand $P_{j,t}$ have been received in stock. Note that the consumption of component i for satisfying end-product demand in period t may take place later than time $t+L_i$ due to lack of availability of components with lead times longer than L_i.

Given the allocation decision $x_{j,k}^{t}$ we can compute the holding and backorder costs associated with end-product demand $P_{j,t}$ in period t. Let us first consider the holding costs incurred after allocating components over time to demand. For all components $i \in \mathcal{M}$ the free available physical stock left for satisfying future demands (demands after period t) after allocation at time t equals $X_{i,t}^{+}$,

$$X_{i,t}^{+} = \text{Max}\left\{0, Y_{i,t-L_i} - \sum_{u=t-L_i}^{t} D_{i,u}\right\}.$$

We incur an immediate holding cost associated with period t equal to $\sum_{i \in \mathcal{M}} h_i X_{i,t}^{+}$ (called *classical* inventory holding cost). Furthermore, some of the *allocated* stock of component i may be carried additional periods, due to lack of availability of other components. Thus additional holding costs will be incurred in future periods that we should associate with the decisions $x_{j,k}^{t}$ at time t. We denote physically available stock already allocated to demand in period t as *remnant stock*. Let us define

$$\tilde{X}_{i,t}(t+s) := \text{amount of remnant stock of component } i \text{ carried at time } t+s \text{ to satisfy demand in period } t, 0 \leq s \leq L.$$

Let us derive an expression for $\tilde{X}_{i,t}(t+s)$. Note that, according to equality (2), the total amount of component i allocated at time t equals to $D_{i,t}$. This implies that the remnant stock associated with the allocation at the end of period $t+s$ can never exceed $D_{i,t}, 0 \leq s \leq L$. Furthermore, we take into account the availability of component i at time $t, t+1, \ldots, t+L_i$ under the assumption that no demand occurs after time t. It follows from (3) and (4) that component i stock availability at time $t+s$ equals $X_{i,t}+D_{i,t}+\sum_{u=1}^{s+1} A_{i,t+u}$ for all $0 \leq s \leq L_i$, and equals $X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u}$ for all $L_i + 1 < s \leq L$. As long as there is a backlog of component i at time $t+s, 0 \leq s \leq L_i$, there cannot be remnant stock. As soon as there is positive availability, then the remnant stock is the difference between this availability and the cumulative amount of component i consumed by satisfying end-product demand from period t. These observations yield the following expressions for $\tilde{X}_{i,t}(t+s)$.

$$\tilde{X}_{i,t}(t+s) = \left(\text{Min}\left\{D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{s+1} A_{i,t+u}\right\}\right)^{+} - \sum_{j\in\mathcal{N}}\sum_{u=0}^{s} a_{i,j}x_{j,u}^{t}, \quad 0 \le s \le L_i \tag{5}$$

$$\tilde{X}_{i,t}(t+s) = \left(\text{Min}\left\{D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u}\right\}\right)^{+} - \sum_{j\in\mathcal{N}}\sum_{u=0}^{s} a_{i,j}x_{j,u}^{t}, \quad L_i + 1 < s \le L \tag{6}$$

Clearly, constraints (2)–(4) will guarantee the nonnegativity of $\tilde{X}_{i,t}(t+s)$ for all $0 \le s \le L$.

Next, we determine the amount of demand $P_{j,t}$ still backlogged at time $t+s$ with similar reasoning. Define

$\tilde{B}_{j,t}(t+s) :=$ amount of demand for end − product j in period t that has not been satisfied at time $t+s$.

Then we have

$$\tilde{B}_{j,t}(t+s) = P_{j,t} - \sum_{u=0}^{s} x_{j,u}^{t}, \quad 0 \le s \le L. \tag{7}$$

Now that we have expressions for physically available stock for future demands at time t, and remnant stocks and backlogs associated with the allocation decisions at time t, we can derive an expression for the total cost $C_1^{\mathcal{P}}(t)$ incurred by the *ordering decisions* associated with the inventory-position-based policy $\mathcal{P}$ at the start of period $t - L_i, i \in \mathcal{M}$, yielding $Y_{i,t-L_i}$, and the *allocation decisions* at the end of period t.

$$\begin{aligned}
C_1^{\mathcal{P}}(t) &= \sum_{i\in\mathcal{M}} h_i \left(X_{i,t}^{+} + \sum_{s=0}^{L} \tilde{X}_{i,t}(t+s)\right) + \sum_{j\in\mathcal{N}} b_j \sum_{s=0}^{L} \tilde{B}_{j,t}(t+s) \\
&= \sum_{i\in\mathcal{M}} \left\{ h_i \text{Max}\left\{0, Y_{i,t-L_i} - \sum_{u=t-L_i}^{t} D_{i,u}\right\} \right. \\
&\quad + \sum_{s=0}^{L_i} h_i \left[\left(\text{Min}\left\{D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{s+1} A_{i,t+u}\right\}\right)^{+} - \sum_{j\in\mathcal{N}}\sum_{u=0}^{s} a_{i,j}x_{j,u}^{t}\right] \\
&\quad \left. + \sum_{s=L_i+1}^{L} h_i \left[\left(\text{Min}\left\{D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u}\right\}\right)^{+} - \sum_{j\in\mathcal{N}}\sum_{u=0}^{s} a_{i,j}x_{j,u}^{t}\right]\right\} \\
&\quad + \sum_{j\in\mathcal{N}} b_j \sum_{s=0}^{L} \left(P_{j,t} - \sum_{u=0}^{s} x_{j,u}^{t}\right).
\end{aligned} \tag{8}$$

Thus the component allocation problem at time t consists of an objective function and constraints that are linear in our decision variables $x_{j,u}^{t}$, while all problem-relevant exogenous information has revealed itself at or before time t. This implies that the component allocation problem is a Linear Program. Under continuous demand the component allocation problem can be solved with standard methods. In the context of Assemble-To-Order systems end-product demand is typically low volume, which makes it more appropriate to assume that end-product demand is discrete. In that case the component allocation problem becomes an Integer Linear Program (ILP) for which we develop an efficient solution method below.

2.3. Mathematical Programs for Component Allocation

Let $\mathcal{L} = \{0, \cdots, L\}$. According to the analysis in Subsection 2.2, the component allocation problem at period t for the demand in period t under inventory-position-based policy P can be formulated as:

$$\begin{aligned}
\text{Min} \quad & C_1^{\mathcal{P}}(t) \\
\text{s.t.} \quad & (2), (3), (4) \\
& x_{j,k}^{t} \in \mathbb{Z}_{+} \qquad \forall j \in \mathcal{N}, k \in \mathcal{L},
\end{aligned} \tag{9}$$

We would like to emphasize the fact that in the objective function of (9), there are costs determined by policy $\mathcal{P}$, yielding the component inventory positions $Y_{i,t-L_i}$, as well as costs determined by the allocation decisions.

We can rearrange the coefficients before the decision variables and discard the terms not influenced by the decision variables in (8) such that (9) is transformed to:

$$\begin{aligned}
\text{Max} \quad & C_2^{\mathcal{P}}(t) = \sum_{j\in\mathcal{N}}\sum_{k=0}^{L}\left(\sum_{i\in\mathcal{M}} h_i a_{i,j} + b_j\right)(L+1-k)x_{j,k}^{t} \\
\text{s.t.} \quad & (2), (3), (4) \\
& x_{j,k}^{t} \in \mathbb{Z}_{+} \qquad \forall j \in \mathcal{N}, k \in \mathcal{L},
\end{aligned} \tag{10}$$

The explanation of $C_2^{\mathcal{P}}(t)$ is that we collect reward $(\sum_{i\in\mathcal{M}} h_i a_{i,j} + b_j)(L+1-k)$ for each unit of end-product j demand satisfied at time $t+k$. This reward depends on the moment the end-product demand is satisfied, which is in contrast with the reward for each unit satisfied at time t used in [2].

An interesting observation of $C_2^{\mathcal{P}}(t)$ is that its value does not depend on the specific definitions of h_i's and b_j's. In (10), if we let $h'_i = 0$ and $b'_j = \sum_{i\in\mathcal{M}} h_i a_{i,j} + b_j$ for all i, j, then the same optimal FCFS component allocation will be obtained. The equivalence of (9) and (10) implies that if the new cost scheme of h'_i and b'_j is implied, the optimal FCFS component allocation in (9) will be obtained as well. Using $\sum_{k=0}^{L} x_{j,k}^t = P_{j,t}$, we can easily show that (9) is equivalent to:

$$\begin{aligned} \text{Min} \quad & C_3^{\mathcal{P}}(t) = \sum_{j\in\mathcal{N}} \sum_{k=0}^{L} k b'_j x_{j,k}^t \\ \text{s.t.} \quad & (2), (3), (4) \\ & x_{j,k}^t \in \mathbb{Z}_+ \qquad \forall j \in \mathcal{N}, k \in \mathcal{L}, \end{aligned} \tag{11}$$

We emphasize that the equivalence among (9), (10), and (11) is for *component allocation* optimization only, that is, under the same inventory-position-based component replenishment policies. Thus we have proven the following theorem.

THEOREM 2: Mathematical programs (9), (10), and (11) are equivalent in terms of optimal solutions.

2.4. An Iterative Algorithm for Component Allocation

According to Theorem 2, we can transform (9) to an equivalent mathematical program with the same feasible region and objective $C_3^{\mathcal{P}}(t) = \sum_{j\in\mathcal{N}} \sum_{k=0}^{L} k b'_j x_{j,k}^t$. In the following we show that an iterative algorithm can be used to solve this mathematical program.

To do so we reconsider the expressions for the remnant stocks (5) and (6). From the non-negativity of the remnant stocks we find

$$\sum_{u=0}^{k} \sum_{j\in\mathcal{N}} a_{i,j} x_{j,u}^t \le \left(\text{Min} \left\{ D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{k+1} A_{i,t+u} \right\} \right)^+, \quad i \in \mathcal{M}, 0 \le k \le L_i \tag{12}$$

$$\sum_{u=0}^{k} \sum_{j\in\mathcal{N}} a_{i,j} x_{j,u}^t \le \left(\text{Min} \left\{ D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u} \right\} \right)^+, \quad i \in \mathcal{M}, L_i + 1 < k \le L \tag{13}$$

In the presence of constraint (2), which ensures that the cumulative amount of component i allocated to demand in period t can never exceed $D_{i,t}$, we find that the feasible region determined by constraints (2), (3), and (4) is identical to the feasible region determined by constraints (2), (12), and (13).

Let us define the variables $O_{i,k}^t$ as

$$\begin{aligned} O_{i,k}^t &:= \left(\text{Min} \left\{ D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{k+1} A_{i,t+u} \right\} \right)^+, \quad 0 \le k \le L_i \\ O_{i,k}^t &:= \left(\text{Min} \left\{ D_{i,t}, X_{i,t} + D_{i,t} + \sum_{u=1}^{L_i+1} A_{i,t+u} \right\} \right)^+, \quad L_i + 1 \le k \le L. \end{aligned} \tag{14}$$

Then we can reformulate the component allocation problem as

$$\begin{aligned} & \text{Min } C_3^{\mathcal{P}}(t) \\ & \text{s.t.} \ \sum_{k=0}^{L} x_{j,k}^t = P_{j,t} \qquad j \in \mathcal{N} \\ & \sum_{u=0}^{k} \sum_{j\in\mathcal{N}} a_{i,j} x_{j,u}^t \le O_{i,k}^t \ i \in \mathcal{M}, k \in \mathcal{L} \\ & x_{j,k}^t \in \mathbb{Z}_+ \qquad \forall j \in \mathcal{N}, k \in \mathcal{L}. \end{aligned} \tag{15}$$

We can design a Benders decomposition algorithm to solve (15). In each iteration, this algorithm solves a variant of (15) where only a subset of the second set of constraints in the feasible region is used. We add such a constraint only when it is violated. As presented in Figure 1, it is clear that the algorithm will stop in a finite number of steps with the optimal solution. In Figure 1, we define

$\hat{O}_{i,k}^t :=$ amount of component i allocated at the end of period $t + k$ to satisfy demand in period t.

As we add only binding constraints, we have that $\hat{O}_{i,k}^t = O_{i,k}^t - O_{i,k-1}^t$ (let $O_{i,-1}^t = 0$). Therefore, we always have $\sum_{u=1}^{k} \hat{O}_{i,u}^t = O_{i,k}^t$ and $O_{i,L}^t = \sum_{k=0}^{L} \hat{O}_{i,k}^t = \sum_{j\in\mathcal{N}} a_{i,j} P_{j,t} = D_i$.

3. BASE-STOCK LEVEL OPTIMIZATION

In the derivations above, we assumed that the inventory-position-based component replenishment policies are given. In this section we assume that a base-stock component replenishment policy $\mathcal{B}$ is used. To optimize the entire system, base-stock levels and component allocation must be jointly determined.

Let S_i denote the base-stock level for component $i, i \in \mathcal{M}$. Under a base-stock policy $\mathcal{B}$, (9) can be simplified. According to definition (14), we have $O_{i,k}^t =$

Algorithm 1

1: Initialization: Define φ as the set of all active second set constraints $\sum_{u=0}^{k}\sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t \le O_{i,k}^t$ in (15). Let $\varphi = \emptyset$.
2: Solve the variant of (15), where the second set constraints are replaced by constraints in φ, and $b'_j = \sum_{i\in\mathcal{M}} h_i a_{i,j} + b_j$.
3: For all $i \in \mathcal{M}$, do Step 4 to 6.
4: Let $k = 0$.
5: Compare $\sum_{u=0}^{k}\hat{O}_{i,u}^t$ and $\sum_{u=0}^{k}\hat{Q}_{i,u}^t$ where $\hat{Q}_{i,u}^t = \sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t$. If $\sum_{u=0}^{k}\hat{O}_{i,u}^t \ge \sum_{u=0}^{k}\hat{Q}_{i,u}^t$, let $k = k+1$. If $k > L$, go to Step 3.
6: If $\sum_{u=0}^{k}\hat{O}_{i,u}^t < \sum_{u=0}^{k}\hat{Q}_{i,u}^t$, add the constraint $\sum_{u=0}^{k}\sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t \le O_{i,k}^t$ corresponding to i and k to φ. Go to Step 2.
7: If for all $i \in \mathcal{M}$, no constraint $\sum_{u=0}^{k}\sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t \le O_{i,k}^t$ is added to φ in Step 3 to 6, then the solution in Step 2 is optimal.

Figure 1. An iterative algorithm for ATO system component allocation

(Min $\left\{X_{i,t} + D_{i,t} + \sum_{s=1}^{k+1} A_{i,t+s}\right\}, D_{i,t})^+$ for $0 \le k \le L_i$, and $O_{i,k}^t = ($Min $\left\{X_{i,t} + D_{i,t} + \sum_{s=1}^{L_i+1} A_{i,t+s}\right\}, D_{i,t})^+$ for $L_i < k \le L$. Consider $0 \le k \le L_i$. Note that a base-stock policy requires that $Y_{i,t+1} = S_i$. Therefore $S_i = Y_{i,t+1} = X_{i,t} + \sum_{s=1}^{L_i+1} A_{i,t+s}$ (cf. the derivation before Theorem 1), which implies that $X_{i,t} + \sum_{s=1}^{k+1} A_{i,t+s} = S_i - \sum_{s=k+2}^{L_i+1} A_{i,t+s} = S_i - \sum_{s=t+k+1-L_i}^{t} D_{i,s}$ (note that $A_{i,t+s} = D_{i,t+s-L_i-1}$ under the base-stock policy). Thus $O_{i,k}^t = (\text{Min}\left\{(S_i - \sum_{s=t+k+1-L_i}^{t} D_{i,s}) + D_{i,t}, D_{i,t}\right\})^+$ for $0 \le k \le L_i$. This implies that $O_{i,L_i}^t = D_{i,t}$. Similarly, when $k > L_i$, we have $O_{i,k}^t = D_{i,t}$.

With these notations, we can express the component allocation problem for policy $\mathcal{B}$ as:

$$\begin{aligned} &\text{Min } C_1^{\mathcal{B}}(\mathbf{S}, \mathbf{x}^t) \\ &\text{s.t. } \sum_{k=0}^{L} x_{j,k}^t = P_{j,t} \qquad \forall j \in \mathcal{N} \\ &\quad \sum_{u=0}^{k}\sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t \le O_{i,k}^t \; \forall i \in \mathcal{M}, k \in \mathcal{L} \\ &\quad x_{j,k}^t \in \mathbb{Z}_+ \qquad \forall j \in \mathcal{N}, k \in \mathcal{L}, \end{aligned} \tag{16}$$

where $\mathbf{S} = (S_i)_{i\in\mathcal{M}}, \mathbf{x}^t = (x_{j,k}^t)_{j\in\mathcal{N},k\in\mathcal{L}}$ and

$$\begin{aligned} C_1^{\mathcal{B}}(\mathbf{S}, \mathbf{x}^t) := &\sum_{i\in\mathcal{M}} h_i(S_i - D_i[t-L_i, t])^+ \\ &+ \sum_{i\in M}\sum_{k=0}^{L} h_i\left(O_{i,k}^t - \sum_{u=0}^{k}\sum_{j\in\mathcal{N}} a_{i,j}x_{j,u}^t\right) \\ &+ \sum_{j\in\mathcal{N}}\sum_{k=0}^{L} b_j\left(P_{j,t} - \sum_{u=0}^{k} x_{j,u}^t\right). \end{aligned} \tag{17}$$

Note that by period t when (16) needs to be solved, the random demands $P_{j,s}$ for $t - L_i \le s \le t$ have already been realized. Thus, (16) is a *deterministic* optimization problem. Moreover, the base-stock levels are determined before the random demands are realized. To emphasize that the optimal objective value of (16) would depend on the random demands, we can write it as $Q(\mathbf{S}, \xi(\omega))$, where $\xi = (P_{j,s})_{t-L_i\le s\le t, j\in\mathcal{N}}$ is the vector of all random variables (ω represents the random outcome). Then the joint optimization problem can be modeled as a two-stage stochastic program (cf. [4, 18]):

$$\begin{aligned} &\text{Min} \quad \mathbb{E}_\xi[Q(\mathbf{S}, \xi(\omega))] \\ &\text{s.t.} \quad S_i \in \mathbb{Z}_+ \qquad \forall i \in \mathcal{M}. \end{aligned} \tag{18}$$

Note that the $O_{i,k}^t$'s are piecewise linear *nonconvex* functions of S_i's. In (16), these nonlinear terms appear in the right hand side and in the objective function. Thus, the joint optimization is a stochastic integer nonlinear program, which is notoriously difficult to solve [4]. Here, from (17), it can be seen that the $O_{i,k}^t$ terms appear only in the expression of the remnant stock holding cost. Therefore, as in [1], the technical difficulty in the joint optimization problem arises mainly from the cost of remnant stock. To avoid this difficulty, we can drop the remnant stock holding cost, that is, replace $C_1^{\mathcal{B}}(\mathbf{S}, \mathbf{x}^t)$ by $C_4^{\mathcal{B}}(\mathbf{S}, \mathbf{x}^t)$,

$$\begin{aligned} C_4^{\mathcal{B}}(\mathbf{S}, \mathbf{x}^t) := &\sum_{i\in\mathcal{M}} h_i(S_i - D_i[t-L_i, t])^+ \\ &+ \sum_{j\in\mathcal{N}}\sum_{k=0}^{L} b_j\left(P_{j,t} - \sum_{u=0}^{k} x_{j,u}^t\right) \end{aligned}$$

In the objective function, we can easily linearize the nonlinear term $(S_i - D_i[t-L_i, t])^+$ by introducing a continuous variable. Then, as in [2], the Sample Average Approximation algorithm (cf. [9]) can be used to solve the problem without the remnant stock costs. The approximation scheme is not only easier to solve, but can provide lower and upper bounds on the optimal objective as well.

THEOREM 3: Let the objective function of the original joint optimization problem (16) be $G(\mathbf{S})$, where $\mathbf{S} = (S_i)_{i\in\mathcal{M}}$ is the vector of base-stock levels; and let the objective function of the problem without remnants stocks be $\tilde{G}(\mathbf{S})$. Let $\mathbf{S}^*$ be the optimal solution of the original problem, and $\tilde{\mathbf{S}}$ be the optimal solution of the problem without remnant stocks. Assume the percentage of the remnant stock holding cost associated with $\tilde{\mathbf{S}}$ is ϵ, that is, $(1-\epsilon)G(\tilde{\mathbf{S}}) = \tilde{G}(\tilde{\mathbf{S}})$, then we have

$$(1-\epsilon)G(\tilde{\mathbf{S}}) \le G(\mathbf{S}^*) \le \frac{1}{1-\epsilon}\tilde{G}(\tilde{\mathbf{S}}). \tag{19}$$

PROOF: For the first inequality, we have $\frac{G(\tilde{\mathbf{S}})-G(\mathbf{S}^*)}{G(\mathbf{S}^*)} = \frac{G(\tilde{\mathbf{S}})}{G(\mathbf{S}^*)} - 1 \le \frac{G(\tilde{\mathbf{S}})}{\tilde{G}(\mathbf{S}^*)} - 1 \le \frac{G(\tilde{\mathbf{S}})}{\tilde{G}(\tilde{\mathbf{S}})} - 1 = \frac{1}{1-\epsilon} - 1 = \frac{\epsilon}{1-\epsilon}$. For the second inequality, we have $\frac{G(\mathbf{S}^*)-\tilde{G}(\tilde{\mathbf{S}})}{G(\mathbf{S}^*)} = 1 - \frac{\tilde{G}(\tilde{\mathbf{S}})}{G(\mathbf{S}^*)} = 1 - \frac{(1-\epsilon)G(\tilde{\mathbf{S}})}{G(\mathbf{S}^*)} \le 1 - \frac{(1-\epsilon)G(\mathbf{S}^*)}{G(\mathbf{S}^*)} = \epsilon$. Therefore (19) follows. □

With similar reasoning as in Theorem 3, the same lower and upper bounds can be derived for general inventory-position-based policies. This implies that if we find an optimal solution for the problem without remnant stocks having a small percentage remnant stock costs, then the bounds are tight. Note that the proposed approximation scheme is still a nonconvex optimization problem. However, computationally, this approximation will remove a part of the nonlinearity of the original program, and can be solved much faster than the original program.

4. NUMERICAL RESULTS

In this section, we report the results of the numerical experiments concerning the proposed models and solutions. For all the experiments, we assume base-stock component replenishment policies. Importantly, we do not need any assumption on the demand process for deriving the formulations in Section 2 for the multimatching problem. For our numerical study we need further assumptions, basically for convenience regarding simulating the demand process. Given the base stock policies and the allocation policy we can evaluate the costs using discrete event simulation, sampling from the demand distributions assumptions. In principle we can approximate the costs arbitrarily close by increasing the simulation run-length. Unfortunately we have to deal with a nonconvex cost function, and for the optimal allocation we must solve an ILP in each period. The latter implies that evaluating the costs for a given base stock policy can be quite time consuming. The former implies that the search for the optimal base-stock policies can be quite time consuming here as well. Thus we need to compromise on run-length for each cost evaluation for given base stock policies and allocation policy.

First, we study the behavior and impact of remnant stock. Second, the effectiveness and efficiency of the solution strategies are tested. Finally, we examine the impact of component allocation optimization in relation to base-stock level optimization, with emphasis on the role of component allocation heuristics in ATO system optimization. The algorithms are implemented in ANSI C, by incorporating the ILOG CPLEX12.4 Callable Library optimization package. A WINDOWS workstation with 2.4GHz Intel Xeon processor and 24GB RAM was used for all experiments. In the following, we report the major results of our numerical experiments. For more details, please refer to Supporting Information.

Table 2. Cost distribution with respect to instance size

Instance	Cost Percentage			Correlation Coefficient	
Size	I	II	III	I & II	II & III
$n4m4$	78.38	6.06	15.55	−0.36	0.65
$n16m8$	34.40	19.86	45.73	−0.51	0.82
$n32m16$	2.56	35.68	61.76	−0.39	0.89
$n8m16$	69.25	12.72	18.03	−0.58	0.98
$n16m32$	26.36	31.81	41.83	−0.65	0.98

4.1. The Behavior and Impact of Remnant Stock

In the first experiment, the component allocation schemes are tested, with base-stock levels that are derived from component safety stocks that have the same safety factor ν,

$$S_i = (L_i + 1)E[D_i] + \nu\sqrt{L_i + 1}\sigma(D_i).$$

For each ATO system instance, 100 realizations of random demands are generated and the corresponding component allocation problems are optimally solved. The three types of costs (classical inventory holding, remnant stock holding, and backlogging costs; Cost I, II and III, respectively) are collected from the optimal solution. The percentages of each cost are then calculated, as well as the correlation coefficients among them. See Table 2, where instance size refers to the ATO system size. For example, $n4m4$ is indicative of 4 products and 4 components. Note that in our Supporting Information, the findings concerning the influences of base-stock levels and lead time dissimilarity in components, are presented as well.

The percentage of the remnant stock holding cost ranges from 6.06% to 35.68%. Table 2 and the experiments in Supporting Information show that remnant stock holding cost becomes a meaningful contributor to operational cost as: (a) the size of the ATO system increases; (b) the base-stock levels are low, or (c) the dissimilarity in the lead times is high. Also, in all instances, the correlation coefficients between classical inventory holding cost and remnant stock holding cost are uniformly negative, while the correlation coefficients between the remnant stock holding cost and the backlogging cost are uniformly positive. Thus, the behavior of remnant stock is similar to that of backorders but deviates from classical inventory. This is due to the multimatching nature of the inventory system.

Table 3. Comparison of optimal FCFS allocation and simple FCFS heuristic allocation rules.

In.	De.	OA obj.	OA gap (%)	Local opt. gap (%) PBP	Local opt. gap (%) FS	Local opt. gap (%) OBCA	Global opt. gap (%) PBP	Global opt. gap (%) FS	Global opt. gap (%) OBCA
M1	−0.1	245.21	9.18	9.9	14.85	18.13	19.08	24.03	27.31
	−0.05	230.71	2.73	8.41	18.24	19.83	11.14	20.96	22.56
	0	224.59	0	6.85	17.13	20.4	6.85	17.13	20.4
	+0.05	233.37	3.91	5.03	15	21.04	8.94	18.91	24.95
	+0.1	247.11	10.03	3.46	15.65	18.38	13.49	25.67	28.41
M2	−0.1	616.87	34.22	24.26	46.81	8.82	58.47	81.02	43.04
	−0.05	508.28	10.59	20.26	38.44	7.74	30.85	49.03	18.33
	0	459.6	0	13.35	32.14	5.94	13.35	32.14	5.94
	+0.05	497.43	8.23	10	22.88	5.47	18.23	31.11	13.7
	+0.1	563.28	22.56	5.71	16.8	2.29	28.27	39.36	24.85

4.2. Efficiency of the Iterative Algorithm for Component Allocation

In Option I, the standard CPLEX Mixed-Integer-Programming (MIP) solver is used. In Option II, Algorithm 1 in Section 2.4 is applied. Eight ATO instances with sizes $n16m32, n32m64, n64m128, n128m256, n32m16$, $n64m32$, $n128m64, n256m128$ are tested. For each instance, five demand realizations are generated and the corresponding component allocation problems are solved. Option II proves faster in 38 of the 40 component allocation problems. For all six of the larger instances ($n \geq 32$ or $m \geq 32$), Option II consistently takes 20% less time than Option I. Thus, Algorithm 1 is effective and improves solution efficiency.

4.3. Optimal FCFS Allocation Versus Simple FCFS Heuristic Allocation

In our third experiment, the effects of simple-to-implement FCFS allocation heuristics are compared to optimal FCFS allocation. Specifically, three FCFS heuristic policies, Product-Based-Priority (PBP, see [22]); Fair-Share (FS, see [1]); and the Order-Based-Component-Allocation (OBCA, see [2]) are evaluated relative to the optimal FCFS component allocation mathematical program we propose.

Two small instances denoted as M1 ($n2m2$) and M2 ($n3m2$) are used. Under PBP, backlogging costs are used to rank the products, that is the priority list reflects decreasing backlogging costs. With the FS rule, we round the allocated components down to the closest integer number. A greediness coefficient of 1 was used under the OBCA rule. One hundred realizations are generated for each ATO instance. We find the globally optimal base-stock levels first by enumeration. We then deviate from the optimal base-stock levels by a uniform percentage and compare the SAA objectives of the three FCFS heuristics to optimal FCFS allocation.

For a given set of base-stock levels, we not only compute the optimality gaps between the heuristic and optimal allocation rules, but also obtain the optimality gaps between these solutions and the globally optimal solution. We refer to these as the local and global optimality gaps, respectively. Table 3 summarizes the results for the M1 and M2 instances.

Regarding Table 3, the first column (In.) shows the instance index, followed by (De.), the percentage deviation of the base-stock levels from the globally optimal solution. For example, for M1, when deviation equals 0, the base-stock levels 66 and 40 are the global optimal, whereas when the deviation is 0.1, the base-stock level of component 1 is $\lfloor 66 \times (1 + 0.1) \rfloor =$ 72. Column (OA obj.) refers to the objective value (for the SAA problem of 100 realizations) of the optimal FCFS allocation under a given base-stock level, while OA gap (%) reflects the percentage difference between the objective and the globally optimal solution. Local opt. gap (%) refers to the percentage difference between the objective of a given FCFS heuristic allocation rule and that of the optimal FCFS allocation, while Global opt. gap (%) is the difference relative to the global optimal objective. To standardize the comparisons, when the gap percentages are calculated, the global optimal objective is always used as the denominator. Thus, the global optimality gap percentage (Global opt. gap (%)) is always the sum of the optimal allocation gap [OA gap (%)] and the local optimality gap percentages [Local opt. gap (%)].

The experiment findings clearly demonstrate the value of optimal allocation, especially the necessity of joint optimization. Table 3 shows that as the base-stock levels increase, the importance of optimal FCFS component allocation decreases, and the simple FCFS heuristic allocation rules perform well. However, in terms of joint optimality gap, no matter whether the base-stock levels deviate from the optimal solution from above or from below, larger deviations always make the global optimality gap increase. The global optimality gap of a FCFS heuristic allocation rule can be viewed as the optimality gap of the optimal FCFS allocation (OA gap) plus the local optimality gap. When the base-stock levels deviate from above the optimal solution, although the local optimality gap tends to decrease, the optimality gap of the optimal FCFS allocation always increases. Thus, the global optimality gap will still increase. This is crucial because it shows that joint optimization with optimal FCFS component allocation is always important. Finally, note that local optimization becomes more important as the base-stock levels decrease.

5. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

We evaluated cost minimization schemes for a periodic-review ATO system with independent component inventory-position-based replenishment policies and FCFS allocation. Several contributions to the literature concerning these systems were made. First, we showed that the component allocation problem under FCFS is a period-by-period deterministic problem. Second, the concept of multimatching was introduced wherein an exact cost accounting scheme for the ATO system under consideration was proposed. In our system, classical inventory holding, remnant stock holding, and backlogging costs are all determined simultaneously. Third, a new mathematical program for component allocation was proposed in which an iterative algorithm was used to improve the solution efficiency. An approximation scheme was also proposed for the joint optimization of inventory replenishment and component allocation, which is a two-stage stochastic integer nonlinear program. Both lower and upper bounds are developed for the joint optimization problem. Fourth, the importance of component allocation was evaluated relative to base-stock level optimization. This was viable due to our exact model for component allocation. Our experiments showed that although component allocation is of less concern than base-stock level optimization in a cost minimization setting, it is nonetheless of general importance. Finally, we showed that joint optimization incorporating the optimal FCFS component allocation will always be of value.

There are several avenues for future study. First, the concept of multi-matching could be extended to non-FCFS component allocation for certain cases. Second, as Theorem 1 holds for all inventory-position-based component replenishment policies under Condition 1, we should explore the impact of component order lot-sizing on the performance of ATO systems. Finally, our computational studies clearly demonstrate that component allocation problems can be optimally solved for small to medium size ATO systems. Since the bottleneck is the joint optimization problem, its solution strategy is worth further study.

ACKNOWLEDGMENTS

The authors acknowledge two anonymous referees, the Associate Editor and the Editor-In-Chief for their critical comments and valuable suggestions that helped us greatly improve the presentation of this paper. We also thank Joachim Arts for his valuable comments and suggestions. This research has been supported by the National Science and Engineering Research Council of Canada under the Discovery Grant (RGPIN/386124-2010).

REFERENCES

[1] N. Agrawal and M.A. Cohen, Optimal material control in an assembly system with component commonality, Naval Res Logist 48 (2001), 409–429.
[2] Y. Akcay and S.H. Xu, Joint inventory replenishment and component allocation optimization in an assemble-to-order system, Manage Sci 50 (2004), 99–116.
[3] R.P. Bagozzi, Principles of Marketing Management, Science Research Associates, Illinois, 1986.
[4] J.R. Birge and F. Louveaux, Introduction to Stochastic Programming, Springer, New York, 1997.
[5] F. Cheng, M. Ettl, G. Lin, and D.D. Yao, Inventory-service optimizaiton in configure-to-order systems, Manufact Serv Oper Manage 4 (2002), 114–132.
[6] A.G. de Kok, "Evaluation and optimization of strongly ideal ato systems," In G.J. Shanthikumar, D.D. Yao, and W.H.M. Zijm, (Editors), Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains, Chapter 9, Kluwer, Boston, 2003, pp. 203–240.
[7] M.K. Dogru, M.I. Reiman, and Q. Wang, A stochastic programming based inventory policy for assemble-to-order systems with application to the w model, Oper Res 58 (2010):849–864.
[8] V.N. Hsu, C.Y. Lee, and K.C. So, Optimal component stocking policy for assemble-to-order systems with lead-time-dependent component and product pricing, Management Science, 52 (2006):337–351.
[9] A.J. Kleywegt, A. Shapiro, and T. Homem-de Mello, The sample average approximation method for stochastic discrete optimization, SIAM J Opt 12 (2001), 479–502.
[10] H.L. Lee and C.S. Tang, Variability reduction through operations reversal, Manage Sci 44 (1998), 162–172.
[11] L. Lu, J.S. Song, and H. Zhang, Optimal policies for an assemble-to-order N-system, Working Paper.
[12] Y.D. Lu and J.S. Song, Order-based cost minimization in assemble-to-order systems, Oper Res 53 (2005), 151–169.
[13] Y.D. Lu, J.S. Song, and D.D. Yao, Backorder minimization in multiproduct assemble-to-order systems, IIE Trans 37 (2005), 763–774.
[14] P. Mirchandani and A.K. Mishra, Component commonality: Models with product-specific service constraints, Prod Oper Manage 11 (2002), 199–215.
[15] A. Muharremoglu and J.N. Tsitsiklis, A single-unit decomposition approach to multiechelon inventory systems, Oper Res 56 (2008), 1089–1103.
[16] E. Nadar, M. Akan, and A. Scheller-Wolf, Optimal structural results for assemble-to-order generalized m-systems, Oper Res 62 (2014), 571–579.
[17] E.L. Plambeck and A.R. Ward, Optimal control of a high-volume assemble-to-order system, Math Oper Res 31 (2006), 453–477.

[18] A. Ruszczynski and A. Shapiro (Editors), Stochastic Programming, Elsevier, Amsterdam, 2003.

[19] E.A. Silver, D.F. Pyke, and R. Peterson, Inventory Management and Production Planning and Scheduling, Wiley, New York, 1998.

[20] J.S. Song and D.D. Yao, Peformance analsyis and optimization of assemble-to-order systems with random lead times, Oper Res 50 (2002), 889–903.

[21] W. van Jaarsveld and A. Scheller-Wolf, Optimization of industrial-scale assemble-to-order systems, Working Paper.

[22] A.X. Zhang, Demand fulfillment rates in an assemble-to-order system with multiple products and dependent demands, Prod Oper Manage 6 (1997), 309–324.

[23] Y. Zhao and D. Simchi-Levi, Peformance analsyis and evaluaation of assemble-to-order systems with stochastic sequential lead times, Oper Res 54 (2006), 706–724.