A MULTI-ECHELON INVENTORY MODEL FOR A LOW DEMAND REPAIRABLE ITEM

Stephen C. Graves
A. P. Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA 02139

WP No. 1299-82 March 1982

A MULTI-ECHELON INVENTORY MODEL FOR A LOW DEMAND REPAIRABLE ITEM

ABSTRACT

The problem that we address is to determine the inventory stockage levels in a multi-echelon inventory system for a low demand repairable item. In its simplest form the multi-echelon system consists of a set of operating sites supported by a centrally-located repair depot. Each operating site requires a set of working items and maintains an inventory of spare items. The repair depot also holds an inventory of spare items. Item failures are infrequent and are replaced on a one-for-one basis. In this paper we present an exact model for finding the steady-state distribution of the net inventory level at each site. This model assumes that the failures are generated by a compound Poisson process and that the shipment time from the repair depot to each site is deterministic. No assumptions are made with regard to the repair cycle at the depot. We contrast this model with existing models for these systems. Based on the exact model we present an approximation for the steady-state distribution for the case with ample servers at the repair depot. We show that this approximation is very accurate on a set of test problems.

1.0 Introduction

In many industries and service organizations the reliance on multi-echelon logistic systems for repairing and supplying low-demand, recoverable items is becoming more and more prevalent. The military's use of and interest in these systems is well known (e.g. Demmy and Presutti [3], Clark [2]). Manufacturers in high technology fields are also resorting to multi-echelon inventory systems for supporting their field service operations for their products. For instance, most computer and office equipment manufacturers must maintain a service organization to service and repair their products in the field. Much of the service work done by these organizations requires the replacement of failed electronic modules that can ultimately be repaired. Consequently, these service organizations will have a multi-echelon inventory system consisting of field inventories for spare modules, and a centrally-located repair depot that repairs the modules and replenishes the field inventories. Communication networks also require multi-echelon inventory systems to ensure reliable service. For instance, a telephone system consists of a linked network of spatially-dispersed switching centers. Each switching center may have thousands of electronic modules that are each subject to infrequent failure. In most instances the failed modules are recoverable. To maintain system reliability, the telephone system must stock some spare modules at every switching center. Yet for reasons of economies of scale, the system will also have a centralized repair facility and possibly intermediate buffers of spare inventory.

A key component in the design of multi-echelon inventory systems for low-demand, recoverable items is the determination of the proper stockage levels of spare inventory at each echelon. Nahmias [7] provides an excellent review of the management science efforts at addressing this problem. A common approach to this problem (e.g. Sherbrooke's METRIC model [11]) has two components. The first component is to characterize the service performance (e.g. expected shortages) of the multi-echelon system for a given specification of the inventory stockage levels. This may be done with an exact or approximate steady-state model. The second component is to search systematically over the possible inventory stockage levels to find the best choice with regard to both inventory costs and service performance.

In this paper we focus on the first component, namely the characterization of system performance for a given specification of the inventory stockage levels. In the next section we present a general framework for determining the steady-state distribution of net inventory levels. This framework is a generalization of the model developed by Simon [12]. We also interpret the METRIC model [11] in the light of the general framework. In Section 3 we discuss the computational implications of the general framework, in particular with regard to finding the best stockage levels. We also pose an approximation to the steady-state distribution for a particular problem instance. We compare on a set of test problems this approximation with the approximation to the steady-state distribution used by METRIC, and with the exact distribution. Our approximation seems very close to the exact distribution, and seems to dominate the METRIC approximation on the set of test problems. In the final section we discuss extensions to the general framework as well as directions for future research.

2.0 A General Framework

Consider a two echelon inventory system for a repairable item where the system consists of a repair depot and N operating sites, as depicted in Figure 1. To be operable each site requires a number of identical working items. At each site items fail according to a specified failure process. All failed items are repairable, but only at the repair depot. Failures are handled by a one for one ordering system. Upon failure an item is replaced with a working item from the site's stock, if one is available; otherwise, there is a shortage at the site that will be filled when stock is available at the site. The failed item is then sent directly to the depot for repair. Upon repair the item is placed in stock at the depot or is used to fill a backorder on the depot. At the same time that the failed item is sent to the depot, the site requests a replacement item from the depot. If the depot has available stock, the depot immediately ships a replacement item; otherwise the request is backordered to be filled when stock is available at the depot, although not necessarily on a first-come, first-serve basis.

For this setting we require two critical assumptions. First we assume that at each site the failure process is a compound Poisson process that is independent of the site status. That is, the failure process for the site does not depend on the actual number of working items. This assumption is clearly violated whenever there are shortages at the site such that the number of working items drops below the normal requirements. Nevertheless, this assumption is common to the literature in this area (e.g. Sherbrooke [11], Simon [12], Muckstadt [6], Allen and D'Esopo [1], Simon and D'Esopo [13], Richards [9]), and seems reasonable provided that the expected number of shortages at a site is to be small relative to the required number of working items at that site.

The second assmuption is that the total shipment time from the repair

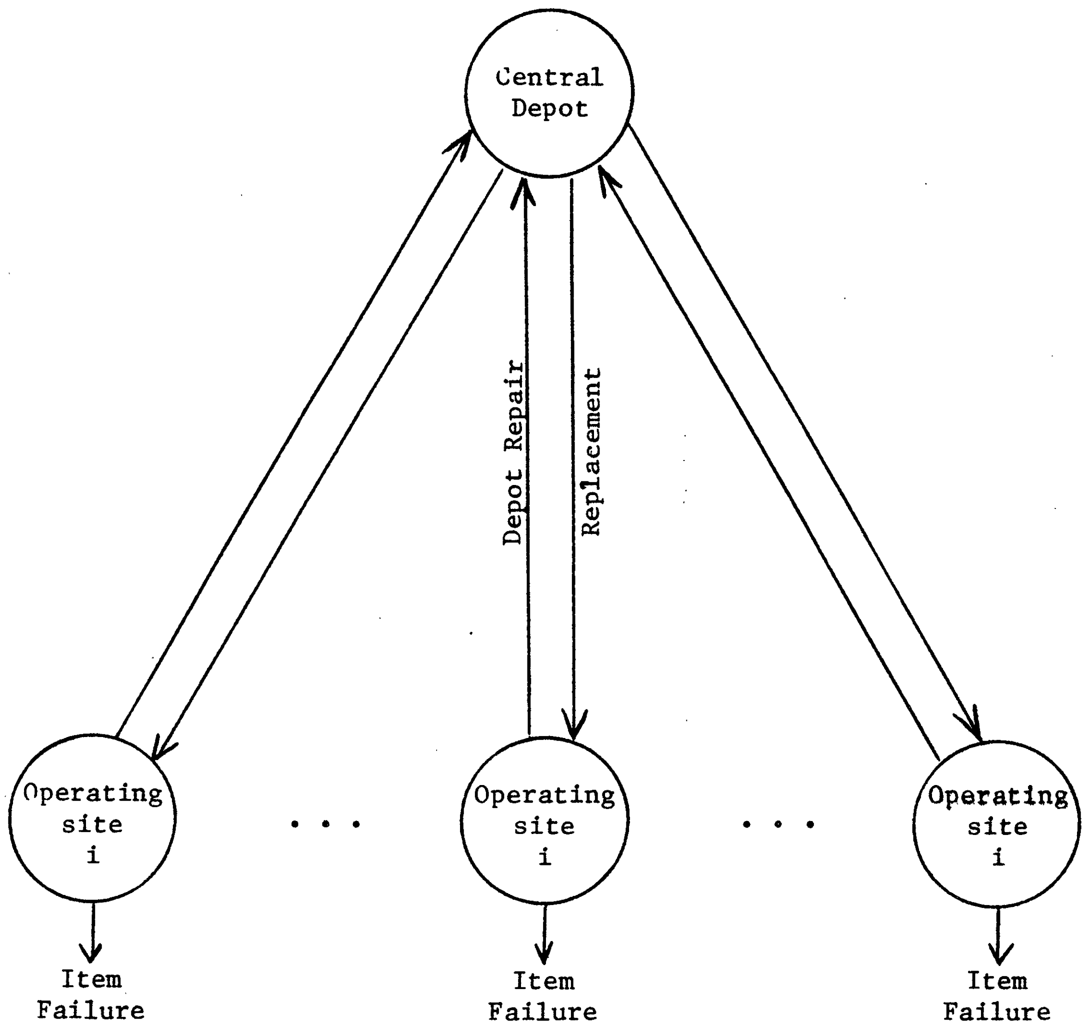

Figure 1: A Two Echelon System

depot to each site is deterministic. Simon [12] also makes this assumption, whereas the METRIC models do not ([6],[11]). Although we may permit this shipment time to vary across sites, for convenience we assume that the depot-to-site shipment time is the same for all sites and given by T_1.

We define $Q_i(t)$ to be a random variable denoting the outages at site i at time t. An outage corresponds to a replacement request that has yet to be filled. The unfilled request could be either intransit from the depot to the site or backordered at the depot. If s_i is the number of spare items stocked at the site, then $s_i - Q_i(t)$ is the net spare inventory on hand at time t, where negative inventory denotes a shortage level. We define Q(t) to be the aggregate outages at the sites, i.e.

$$Q(t) = \sum_{i=1}^{N} Q_i(t).$$

We note that $Q_i(t)$ does not depend on the stockage quantity at the site s_i, but does depend on the number of spares stocked at the depot, denoted by s_o. For notational convenience we will make this dependence explicit only when needed.

The primary result is

$$Q(t+T_1) = B(t|s_o) + D(t,t+T_1) \tag{1}$$

where $B(t|s_o)$ is the backorders at time t at the depot assuming the depot stocks s_o items, and where $D(t,t+T_1)$ is the total number of failures at all sites over the time interval $(t,t+T_1]$. Hence, if $D_i(t,t+T_1)$ is the number of failures at site i, then

$$D(t,t+T_1) = \sum_{i=1}^{N} D_i(t,t+I_1).$$

We note that $B(t|s_o)$ and $D(t,t+T_1)$ are independent random variables since the depot backorders at time t depend only on item failures that occur prior to time t and since we assume that the failure process at each site is a compound Poisson process that is independent of the status of the site.

We provide here an intuitive explanation of (1); a more formal proof is possible, but we believe to be unnecessary and cumbersome. At time t the aggregate outages at the sites consist of items that are intransit to the sites and of items that are backordered at the depot. By time $t+T_1$ all items that had been intransit at time t will have arrived at the sites and will no longer be outages. However, since the shipping time from depot to site is exactly T_1, those items that were backordered at time t cannot have arrived at the sites by time $t+T_1$. Hence, the depot backorders at time t, $B(t|s_o)$, remain as outages at time $t+T_1$. Any failure in the time interval $(t,t+T_1]$ generates a replacement request by the site that cannot feasibly be filled by time $t+T_1$ since the required shipping time is T_1. Hence all failures over this interval, $D(t,t+T_1)$, must be outages at time t. Furthermore, since $B(t|s_o)$ and $D(t,t+T_1)$ are independent, they do not double-count any outages, but do sum to give total outages at time t+T. We note that this explanation is very similar in spirit to that for finding the on-hand inventory in a continuous review inventory system with a constant lead time (Hadley and Whitin [15], pp. 181-188). Simon and D'Esopo [13] also use this type of argument to establish the steady-state distribution for on-hand inventory in a single-site model with both repairable and nonrepairable item failures.

This remarkably simple result requires only the assumptions of an independent compound Poisson failure process and deterministic shipment times from the depot to the sites. No assumptions were required with regard to the shipment times from the sites to the depot, or the repair process. Furthermore by using the same logic, we can extend the result to distribution networks with any number of echelons provided that the transit times from each echelon to its immediate successors are deterministic.

The utility of (1) for characterizing the site outages $Q_i(t)$ depends primarily on the resolution of two issues. First, we need the probability distribution for depot backorders $B(t|s_o)$. With this we obtain the distribution of the aggregated outages $Q(t)$ via a convolution with the distribution for $D(t,t+T_I)$ (assumed to be known). The second issue is how to disaggregate $Q(t)$ into $Q_i(t)$ for $i=1,2,...,N$. Knowledge of this distribution of $Q_i(t)$ permits us to find, for a given depot stockage level s_o, the best site stockage level s_i that minimizes inventory and shortage costs. Before exploring these issues we first take a look at the models of Simon [12] and Sherbrooke [11] in the light of (1).

Interpretation of Simon and METRIC Models

Simon [12] considers a two echelon system with a Poisson failure process at each site, deterministic transit times, and ample repair capacity (i.e. there is no queueing, and successive repair times are i.i.d. random variables). He derives the steady-state distribution for the net inventory level at each site. In terms of (1), we can interpret Simon's model as follows. The depot backorders would be written as

$$B(t|s_o) = [Q_o(t) - s_o]^+ \tag{2}$$

where $Q_o(t)$ is the number of outages at the depot at time t and $[x]^+$ denotes the nonnegative part of x. A depot outage is analogous to a site outage and represents a failed item that has not yet completed repair. These failed items can be either intransit from the sites to the depot or in the repair cycle at the depot. Whenever an item fails at a site, a request is made to the depot for a replacement. For this failed item the depot has an inventory outage until the failed item is repaired. The net inventory level at the depot is just the difference between the planned stockage level and the number of outages, i.e. $s_o - Q_o(t)$. Hence, backorders occur once depot outages exceed the stockage level.

By assuming ample repair capacity, $Q_o(t)$ is modelable as the occupancy level in a $M/G/\infty$ queue where the service time G includes both the intransit time to the depot and the repair time at the depot. Since the steady-state distribution of $Q_o(t)$ is well known to be Poisson (Palm's theorem [8]), we find the steady-state distribution of $B(t|s_o)$ to be just the tail of Poisson distribution. By convolving $B(t|s_o)$ with $D(t,t+T)$, we obtain the distribution of aggregate outages $Q(t)$.

Simon assumes that the depot backorders are filled on a first-come, first-serve basis. This implies that the "disaggregation" of the distribution of Q(t) is essentially a <u>random</u> disaggregation across the sites. That is, the likelihood that any aggregate outage is from a particular site i is directly proportional to that site's failure rate λ_i. Simon determines the distribution of $Q_i(t)$ by conditioning on Q(t) and by using the fact that the conditional distribution of $Q_i(t)$ is a binomial distribution. Specifically, Simon finds that

$$\Pr[Q_i(t)=j] = \sum_{k=j}^{\infty} \Pr[Q_o(t)=k] \binom{k}{j} \left[\frac{\lambda_i}{\lambda}\right]^j \left[\frac{\lambda-\lambda_i}{\lambda}\right]^{k-j} \qquad (3)$$

where $\lambda = \sum_{i=1}^{N} \lambda_i$ is the aggregate failure rate. Simon's original derivation also permits repair at the site and item condemnation [both features can be included in (1), but tend to obscure the simplicity of the result].
Kruse [4] provides a similar interpretation of Simon's model to that given here and shows that it is extendable to more than two echelons. Shanker [10] shows that Simon's model can permit compound Poisson demand, which is clear from this interpretation and the generalization of Palm's theorem by Feeney and Sherbrooke [14].

Sherbrooke's METRIC model [11] considers a two echelon system with a compound Poisson failure process at each site, ample repair capacity and general transit times. METRIC provides an approximate distribution for the net inventory level at each site. In that METRIC permits non-deterministic transit times, it does not fit exactly into the framework given by (1). However, for the case of a deterministic transit time to the sites, we can interpret the METRIC approximation in terms of (1). For ease of presentation assume that the demand processes at the sites are just Poisson. By modeling the depot backorders by (2), METRIC uses Palm's theorem to obtain the distribution of $Q_o(t)$, and the expected backorder level at the depot. METRIC then approximates the distribution of $B(t|s_o)$ as a Poisson distribution with the given mean. Since both $B(t|s_o)$ and $D(t,t+T_1)$ are assumed Poisson, their convolution $Q(t+T_1)$ is also Poisson and completely characterized by its mean. Now METRIC determines the distribution of $Q_i(t)$ by a random disaggregation (3) of $Q(t)$, identical to Simon. But for $Q(t)$ being Poisson, this random disaggregation results in each $Q_i(t)$ being Poisson so that only its mean need be computed. Hence, the one distributional approximation by METRIC yields an enormous computational simplification over the Simon exact model. The case with compound Poisson failures is completely analogous.

3.0 Computational Issues

In the previous section we have introduced a general probabilistic framework for characterizing the performance of a two-echelon inventory system for a repairable item. We also reexamined two noteworthy models for such inventory systems within the general framework. In this section we consider the use of this framework to address the standard design questions for these inventory systems: how many spare items are needed and where should they be stocked? In particular we explore both the use of exact and approximate models suggested by the framework, as well as the computational implications of these models.

Exact Models

To use our main result (1), we need a means for determining the backorder level at the depot $B(t|s_o)$, and a means for disaggregating aggregate outages $Q(t)$ into individual site outages $Q_i(t)$. We discuss these issues here. Throughout this discussion, we focus on steady-state distributions.

The equation (2) defines backorders at the depot in terms of the outages at the depot and the stockage level s_o. We note that the outages at the depot do not depend on the stockage level s_o. Rather the depot outage level depends only on the failure processes for the sites (which are independent of the site status), and the repair cycle where the repair cycle consists of the shipping time from the site to depot, queueing time at the depot and the repair time at the depot. This independence property implies that we need only determine the steady-state probability distribution of the depot outage level once to obtain the steady-state probability distributions of the depot backorder level for all values of s_o.

Hence the operationalization of (1) for any value of s_o requires just the determination of the probability distribution of $Q_o(t)$. In the previous section we saw that both Simon and Sherbrooke use Palm's theorem to characterize $Q_o(t)$ under the assumption of ample repair capacity and Poisson or compound Poisson failure processes. For several other common scenarios, we can also determine the probability distribution of the depot outage level. For instance, suppose the sites' failure processes are each Poisson, the shipment time from each site to the depot is a general random variable with finite mean, and there are k parallel repair lines, each with its repair time being exponentially distributed. A failed item arriving at the depot queues until a repair line is available. The distribution of the depot outage level is the convolution of the distribution of the number of items intransit to the depot with the distribution of the number of items either in repair at the depot or in the repair queue. The number of items intransit is equivalent to the occupancy level in an $M/G/\infty$ system and hence has a Poisson distribution. Since the output of an $M/G/\infty$ is a Poisson process that is independent of the occupancy level (Mirasol [5]), the number of failed items either in queue or in repair is modelable as an M/M/k system that is independent (in steady state) of the number of items intransit. The steady-state analysis of an M/M/k system is straightforward and well known. Although there is not a closed form result for the convolution of the two distributions, it is trivial to do this with a computer. Furthermore, this convolution need be done only once for the analysis of stockage levels for a particular scenario. Extensions to this problem setting that are also tractable would include Erlang repair times and more complex repair processes that require a series of repair steps at distinct facilities.

Once we have found for a depot stockage level s_o the distribution of the aggregate site outages via (2) and (1), we need to disaggregate this distribution into the outage distributions for individual sites. The general form for this disaggregation is

$$Pr(Q_i=j) = \sum_{k=j}^{\infty} Pr(Q_o=k)\ Pr(Q_i=j \mid Q_o=k)$$

where we need specify the conditional distribution $Pr(Q_i=j \mid Q_o=k)$ to reflect the depot priority scheme. If the depot fills item replacement requests on a first-come, first-serve basis, then we can perform the disaggregation by (3) where the conditional distribution is a binomial distribution. For other priority schemes, such as filling the most "urgent" request, the form of the conditional distribution is less clear, but presumably more complex. The computations required to obtain the site outage distribution are not trivial and must be done for every value of s_o that is investigated. To limit the amount of computational effort we present an appoximation to the disaggregation process in the next section.

Approximate Model

For the approximate model we assume that all failure processes are Poisson and that the depot services requests on a FCFS basis so that (3) gives the disaggregation of the aggregate outage process. From (3) and (2) we can easily show that the expected site outage level and its variance are given by

$$E\{Q_i\} = \frac{\lambda_i}{\lambda} E\{B(s_o)\} + \lambda_i T_1 \qquad (4)$$

$$Var\{Q_i\} = \left(\frac{\lambda_i}{\lambda}\right)^2 Var\{B(s_o)\} + \left(\frac{\lambda_i}{\lambda}\right)\left(\frac{\lambda-\lambda_i}{\lambda}\right) E\{B(s_o)\} + \lambda_i T_1 \qquad (5)$$

where λ_i is the failure rate for site i, $\lambda = \sum_{i=1}^{N} \lambda_i$ is the aggregate failure rate, and T_1 is the deterministic shipment time from depot to site. The notation $B(s_o)$ denotes the steady-state depot backorder level given s_o units stocked at the depot. Hence we can express the mean and variance of each site's outage level in terms of the mean and variance of the depot backorder level. Now from (2) we show in the appendix that

$$E\{B(s_o)\} = E\{B(s_o-1)\} - Pr(Q_o \geq s_o) \qquad (6)$$

$$\mathrm{Var}\{B(s_o)\} = \mathrm{Var}\{B(s_o-1)\} - \Big\{E\{B(s_o)\} + E\{B(s_o-1)\}\Big\} \cdot [1-\Pr(Q_o \geq s_o)] \qquad (7)$$

That is, we can compute the mean and variance of the depot backorder level recursively, provided we have knowledge of the distribution of the depot outage level. We propose to use (4) - (7) to compute the mean and variance of each site's outage level for all values of s_o of interest. We then intend to approximate the distribution of each site's outage level with a distribution that is fully specified by its first two moments. We illustrate this approximation strategy next.

We consider the case where all failure processes are Poisson and where there is ample repair capacity at the depot so that the depot outage level Q_o has a Poisson distribution. An empirical investigation of this case over a range of parameter sets suggests that the distribution of a site's outage level, Q_i, will be unimodal and will have its variance strictly greater than its mean. We propose to approximate the distribution of the site's outage level by a negative binomial distribution; that is

$$\Pr(Q_i=j) = \binom{r+j-1}{j} p^r (1-p)^j, \qquad \text{for } j=0,1,2,... \qquad (8)$$

where r and p are positive parameters $(0 < p < 1)$ such that

$$E\{Q_i\} = r(1-p)/p \quad , \qquad (9)$$

$$\mathrm{Var}\{Q_i\} = r(1-p)/p^2 \quad . \qquad (10)$$

The mean and variance of the site's outage level are found from (4) - (7).

To test the effectiveness of this approximation we compared this approximation with the METRIC approximation and with the exact distribution on a set of test problems. For all test problems, the shipment time T_1 from depot to site is exactly 3 days. The expected duration of the repair cycle (shipment time from site to depot plus repair time at depot) takes on one of the following four values: $E\{T_2\}$ = 1, 3, 6, or 9 days. We assume that there are four sites. The aggregate failure rate for the four sites is one of the four values: λ = .5, 1, 2, or 4 failures/day. The failure rates for the four sites are such that $\lambda_1/\lambda = .1$, $\lambda_2/\lambda = .2$, $\lambda_3/\lambda = .3$, and $\lambda_4/\lambda = .4$. To specify a test problem we need set the depot stockage level s_o. Since the depot outage level Q_o is a Poisson random variable with a mean and variance equal to $\lambda E\{T_2\}$, we need set the depot stockage level to be consistent with this expected depot outage level. To do this, we permit s_o to range from $\lambda E\{T_2\} - (\lambda E\{T_2\})^{1/2}$ up to $\lambda E\{T_2\} + 2(\lambda E\{T_2\})^{1/2}$ [i.e. from $\mu-\sigma$ to $\mu+2\sigma$]. In particular, we let s_o take on up to six integral values evenly-spaced over this range. Finally, for each test problem we specify the desired fill rate α for each site, where α = .84, .87, .90, .93, .96, .99. Hence, for a given fill rate α, the site stockage level s_i is the minimal quantity such that

$$\Pr\{Q_i \leq s_i\} \geq \alpha . \tag{11}$$

To recap the design of the test problems, we specify a test problem by setting the expected repair cycle time (4 candidates), by setting the aggregate demand rate (4 candidates), by setting the depot stockage level (up to 6 candidates), and by setting the site fill rate (6 candidates). This gives a maximum of 576 test problems. For each test problem, each analysis method (i.e. negative binomial approximation, METRIC approximation, and exact solution) generates the required stockage level for each of the four sites. Hence, a maximum of 4 * 576 = 2304 comparisons are possible.

The results from the test problems are summarized in Tables 1 - 4. In total there are 1968 problem instances. We report the number of instances that each approximation yields the wrong stockage quantity for a site. Overall both approximations are extremely effective. The METRIC approximation results in a "wrong decision" in 227 problem instances of 11.5% of the cases. In all of these instances the METRIC approximation recommends less stock than is actually required. The negative binomial approximation is even better. It errs in only 18 problem instances or 0.9% of the cases. Furthermore, the negative binomial approximation virtually dominates the METRIC approximation over this set of test problems. In all but 2 of the cases where the negative binomial approximation errs, the METRIC approximation also yields a wrong decision. In the two exceptional cases the negative binomial approximation recommends more stock than is actually needed. The computational requirements for the negative binomial approximation are comparable to that for the METRIC approximation. Whereas one requires both the mean and variance of each site's outage level [i.e. (4) - (5)], the other requires just the mean. The computation of the negative binomial distribution is the same complexity as the computation of the Poisson distribution required by METRIC.

Table 1: Error Incidents for Approximation Methods for Aggregate Demand Rate λ = .5 failures/day

$E\{T_2\}$	Number or problem instances per cell **	Site Demand Rate			
		$\lambda_1 = .1\lambda$	$\lambda_2 = .2\lambda$	$\lambda_3 = .3\lambda$	$\lambda_4 = .4\lambda$
1	6	0,0 *	0,0	0,0	0,0
3	18	0,0	0,0	0,0	0,0
6	36	0,0	2,0	0,0	2,0
9	36	1,1	4,0	4,0	4,0

* x,y : x = number of problem instances for which METRIC gave incorrect stockage quantity
y = number of problem instances for which negative binomial approximation gave incorrect stockage quantity.

** For each choice of $E\{T_2\}$ and λ_i, we considered six fill rates and up to six values for the depot stockage quantity.

Table 2: Error Incidents for Approximation Methods for Aggregate Demand Rate λ = 1.0 failures/day

$E\{T_2\}$	Number or problem instances per cell**	Site Demand Rate			
		$\lambda_1 = .1\lambda$	$\lambda_2 = .2\lambda$	$\lambda_3 = .3\lambda$	$\lambda_4 = .4\lambda$
1	12	0,0 *	0,0	0,0	1,0
3	36	3,1	1,0	4,0	4,1
6	36	1,0	5,0	4,0	7,0
9	36	3,0	5,0	6,1	12,2

---

* x,y : x = number of problem instances for which METRIC gave incorrect stockage quantity
y = number of problem instances for which negative binomial approximation gave incorrect stockage quantity.

** For each choice of $E\{T_2\}$ and λ_i, we considered six fill rates and up to six values for the depot stockage quantity.

Table 3: Error Incidents for Approximation Methods for Aggregate Demand Rate $\lambda = 2.0$ failures/day

$E\{T_2\}$	Number or problem instances per cell**	Site Demand Rate			
		$\lambda_1 = .1\lambda$	$\lambda_2 = .2\lambda$	$\lambda_3 = .3\lambda$	$\lambda_4 = .4\lambda$
1	24	0,0 *	0,0	0,0	3,1
3	36	0,0	4,0	2,0	5,0
6	36	2,0	1,1	10,0	10,1
9	36	2,0	3,0	9,1	10,0

---

* x,y : x = number of problem instances for which METRIC gave incorrect stockage quantity
y = number of problem instances for which negative binomial approximation gave incorrect stockage quantity.

** For each choice of $E\{T_2\}$ and λ_i, we considered six fill rates and up to six values for the depot stockage quantity.

Table 4: Error Incidents for Approximation Methods for Aggregate Demand Rate $\lambda = 4.0$ failures/day

$E\{T_2\}$	Number or problem instances per cell**	Site Demand Rate			
		$\lambda_1 = .1\lambda$	$\lambda_2 = .2\lambda$	$\lambda_3 = .3\lambda$	$\lambda_4 = .4\lambda$
1	36	2,1 *	0,0	0,0	2,0
3	36	2,0	3,1	5,0	8,1
6	36	4,1	5,0	8,0	13,1
9	36	4,1	6,0	13,0	18,2

---

* x,y : x = number of problem instances for which METRIC gave incorrect stockage quantity
y = number of problem instances for which negative binomial approximation gave incorrect stockage quantity.

** For each choice of $E\{T_2\}$ and λ_i, we considered six fill rates and up to six values for the depot stockage quantity.

4.0 Discussion

The contribution of this paper is twofold. First, we provide a reasonably general framework for determining the distribution of net inventory levels in a multi-echelon system for low-demand, recoverable items. Second, for a specific problem scenario, namely ample servers at the repair facility, we propose an approximate model that we show to be very accurate on a set of test problems. In this section we discuss extensions to the general framework and directions for future research.

The general model (1) requires a set of assumptions which may or may not be restrictive. The assumption of a compound Poisson failure process is likely to be general enough to capture most failure processes. The assumption of a deterministic shipment time from depot to site is slightly more restrictive. Shipment times do vary but are reasonably predictable. Indeed, many manufacturers may ship small, high-valued, electronic modules via an air cargo service that guarantees overnight service. Finally, the assumption that the failure process is independent of the site status is the most restrictive assumption. It clearly does not hold if there is a small population of working items that generate the item failures. However, even in this case, if there is a very small probability that an operating site has a shortage, then for modeling purposes it may be appropriate to assume that the failure process is independent of site status. Future work need examine the applicability of and extensions to the general model when its assumptions are not reasonable.

We presented the general model in its most simple context: only two echelons, all failed items are repairable, all repairs occur at the depot, and all depot-to-site shipment times are the same. The framework extends directly to distribution systems with more than two echelons provided we have deterministic shipment times from an echelon to its successors. As an example, suppose we have three echelons with a single repair depot, multiple intermediate sites, and multiple operating sites. Then using (1) we determine the aggregate outages at the intermediate sites, where s_o is the stockage level at the depot and T_1 is the shipment time from depot to the intermediate sites. The aggregate outages at the intermediate sites then need to be disaggregated, as in (3). Given a stockage level for each intermediate site, we can determine its backorder level. We now reapply (1) for each intermediate site and its successor operating sites to get the aggregate outages at the sites. Again the aggregate outages for the sites are disaggregated, as in (3). Admittedly, the details are quite involved, but the calculations are feasible.

The general model is also extendable to include site repair and to include item condemnation and procurement. In the first case, the model given by (1) just needs to be augmented to include the random variable for the aggregate number of items in site repair. In the second case, the determination of the backorders at the depot must reflect the effect of condemned units. Both Simon and D'Esopo [13] and Richards [9] illustrate how to determine the depot backorder level when item failures may or may not be recoverable.

If the depot-to-site shipment times are not the same, then we need modify model (1). The model should now be stated as

$$Q_i(t+T_i) = B_i(t|s_o) + D_i(t, t+T_i)$$

where $B_i(t|s_o)$ represents the number of depot backorders at time t that are outages at site i, and T_i is the shipment time from depot to site i. Hence, the outages at site i at time $t+T_i$ are the sum of the site's outages at time t that cannot have arrived by time $t+T_i$, plus the failures over the interval $(t, t+T_i]$. We determine $B_i(t|s_o)$ for each site by disaggregating $B(t|s_o)$, as in (3).

We have suggested and tested one approximation scheme for the two echelon system with ample servers at the repair facility. We have not examined this approximation for cases with limited service channels or with more than two echelons. We hope that future research will address these cases.

Acknowledgement

This research was conducted for and supported by the GTE Laboratories, Inc., Waltham, MA.

The author wishes to acknowledge and thank Dr. Anthony Kooharian and Frances Habib of the GTE Laboratories, Inc., and Prof. Julian Keilson, University of Rochester, for their helpful comments and advice throughout the duration of this research. The author also thanks Maqbool Dada for his insightful comments on an earlier draft of the paper.

APPENDIX

In this appendix, we establish the recursive formulas given by (6), (7) in the text. We are given that

$$B(s_o) = [Q_o - s_o]^+ \qquad \text{(A1)}$$

where $B(s_o)$ is the depot backorder level, Q_o is the depot outage level, and s_o is the depot stockage level. Define $\beta(u|s_o)$ to be the moment generating function for $B(s_o)$:

$$\beta(u|s_o) = \sum_{j=0}^{\infty} \Pr[B(s_o)=j]u^j \qquad \text{(A2)}$$

By induction we can show that

$$\beta(u|s_o) = \frac{1}{u}[\beta(u|s_o-1) - \Pr(Q_o \leq s_o-1)] + \Pr(Q_o \leq s_o-1) \qquad \text{(A3)}$$

for $s_o = 1,2,\ldots,$ where

$$\beta(u|s_o=0) = \sum_{j=0}^{\infty} \Pr(Q_o=j)u^j. \qquad \text{(A4)}$$

By differentiating (A3) and setting u=1, we obtain the result (6) in the text, namely

$$E\{B(s_o)\} = E\{B(s_o-1)\} - \Pr(Q_o \geq s_o). \qquad \text{(A5)}$$

By differentiating (A3) twice and setting u = 1, we obtain

$$E\{B^2(s_o) - B(s_o)\} = -2E\{B(s_o)\} + E\{B^2(s_o-1) - B(s_o-1)\}. \qquad \text{(A6)}$$

From (A5) and (A6) we obtain the desired result (7), namely

$$\mathrm{Var}\{B(s_o)\} = \mathrm{Var}\{B(s_o-1)\} - [E\{B(s_o)\} + E\{B(s_o-1)\}] \cdot [1 - \Pr(Q_o \geq s_o)],$$

(A7)

References

1. Allen, S. G. and D. A. D'Esopo, "An Ordering Policy for Repairable Stock Items", Operations Research, Vol. 16, No. 3 (1968), pp. 669-675.

2. Clark, A. J., "Experiences with a Multi-Indentured, Multi-Echelon Inventory Model", in Multi-Level Production/Inventory Control Systems: Theory and Practice, Studies in Management Sciences, Vol. 16, L. B. Schwarz (editor), North Holland, Amsterdam (1981), pp. 299-330.

3. Demmy, W. S. and V. J. Presutti, "Multi-Echelon Inventory Theory in the Air Force Logistics Command", in Multi-Level Production/Inventory Control Systems: Theory and Practice, Studies in Management Sciences, Vol. 16, L. B. Schwarz (editor), North Holland, Amsterdam (1981), pp. 279-297.

4. Kruse, W. K., "An Exact N Echelon Inventory Model: The Simple Simon Method", Technical Report, March 1979, U.S. Army Inventory Research Office, Philadelphia, PA.

5. Mirasol, N. M., "The Output of an $M/G/\infty$ Queueing System is Poisson", Operations Research, Vol. 11, No. 2 (1963), pp. 282-284.

6. Muckstadt, J. A., "A Model for a Multi-Item, Multi-Echelon, Multi-Identure Inventory System", Management Science, Vol. 20, No. 4 (1973), pp. 472-481.

7. Nahmias, S., "Managing Repairable Item Inventory Systems: A Review", in Multi-Level Production/Inventory Control Systems: Theory and Practice, Studies in Management Sciences, Vol. 16, L. B. Schwarz (editor), North Holland, Amsterdam (1981), pp. 253-277.

8. Palm, C., "Analysis of the Erlang Traffic Formula for Busy-Signal Arrangements", Ericsson Technics, No. 5 (1938), pp. 39-58.

9. Richards, F. R., "A Stochastic Model of a Repairable Item Inventory System with Attrition and Random Lead Times", Operations Research, Vol. 24, No. 1 (1976), pp. 118-130.

10. Shanker, K., "An Analysis of a Two Echelon Inventory System for Recoverable Items", Technical Report 341, School of Operations Research and Industrial Engineering, Cornell University, 1977.

11. Sherbrooke, C. C., "METRIC: A Multi-Echelon Technique for Recoverable Item Control", Operations Research, Vol. 16, No. 1 (1968), pp. 122-141.

12. Simon, R. M., "Stationary Properties of a Two Echelon Inventory Model for Low Demand Items", Operations Research, Vol. 19, No. 3 (1971), pp. 761-777.

13. Simon, R. M. and D. A. D'Esopo, "Comments on a Paper by S. G. Allen and D. A. D'Esopo: 'An Ordering Policy for Repairable Stock Items'", Operations Research, Vol. 19, No. 4 (1971), pp. 986-989.

14. Feeney, G. J. and C. C. Sherbrooke, "The (s-1, s) Inventory Policy Under Compound Poisson Demand", Management Science, Vol. 12, No. 5 (1966), pp. 391-411.

15. Hadley, G. and T. M. Whitin, Analysis of Inventory Systems, Prentice-Hall, Englewood Cliffs, NJ, 1963.