RC 20942 (07/30/97)
Engineering and Technology

IBM Research Report

Improving Performance Evaluation in DRP Systems

Gerald E. Feigin, Kaan Katircioglu
IBM Research Division
T.J. Watson Research Center
Yorktown Heights, New York 10598

David D. Yao
IEOR Department
Columbia University
New York, NY 10027-6699

Research Division
Almaden · T.J. Watson · Tokyo · Zurich

Improving Performance Evaluation in DRP Systems

Gerald E. Feigin and Kaan Katircioglu
IBM Research Division, T.J. Watson Research Center
Yorktown Heights, NY 10598

David D. Yao*
IEOR Department, 302 Mudd Building
Columbia University, New York, NY 10027

July 1997

Abstract

Distribution Resource Planning (DRP) is a general framework for planning and managing inventory in distribution networks. The DRP framework can be applied to complex distribution networks with thousands of unique stock-keeping units (SKUs) and hundreds of stocking locations. It allows for non-stationary (e.g. seasonal) demand patterns and a wide variety of user-specified inventory control rules including all standard inventory policies such as (S, s) and fixed order quantity rules. A number of software implementations of DRP are commercially available and are widely used in industry. In this paper, we describe the logic underlying DRP and point out some its limitations. In particular, we show how the performance evaluation capability of DRP can be substantially enhanced by some simple analytical formulas, derived as approximations from base-stock and (S, s) control schemes.

*Research undertaken while an academic visitor at IBM Research Division, T.J. Watson Research Center; also supported by NSF Grant DDM-95-23029.

1 Introduction

In the retail and distribution sectors of many industries, inventory management activities are made difficult by the sheer enormity of the number of distinct SKUs that are sold. For example, it is not unusual for a large department store to stock more than 100,000 items. In such an environment, inventory management can be reasonably carried out only with the aid of decision support software which automates many of the planning activities, and enables decision makers to view at a glance what their inventory requirements are over time at different levels of product aggregation and at different points in the network. In addition, the software gives estimates of future on-hand inventory, backorders, and service levels, and provides the decision maker with the ability to evaluate quickly how changes in planning parameters such as safety stock levels and order quantities will affect inventory performance, measured primarily by inventory turnover and order fill rate. Because this software can typically be linked directly with order entry and purchasing systems, the entire replenishment decision process, if desired, can be automated. The role of the inventory analyst then becomes one of tuning parameters governing the replenishment decisions to achieve desired business performance, as well as responding to exceptions that the software brings to the attention of the analyst.

A class of inventory decision support software widely used in industry is based on Distribution Resource Planning (DRP), which refers to a general framework for planning and managing inventory in distribution networks. In particular, it allows for the setting of inventory control parameters and the calculation of time-phased inventory requirements, based on user-specified goals, such as service level targets or total inventory investment. It can take into consideration a host of different constraints, including supply capacity constraints, minimum and maximum order quantity constraints, etc. Since the early 1980's, DRP systems have been implemented in a variety of commercial software packages, offering a wide range of decision support capabilities to inventory managers and procurement personnel, and have been widely used in industry. For descriptions of DRP as well as examples of how it is used, refer to Martin [2], Mlot et al. [3], Silver and Peterson [6], Stenger [7], and Smith [8]. Examples of popular software packages that implement DRP include those sold by Manugistics, American Software, i2, Numetrix, and LPA.

One of the advertised strengths of DRP is its ability to project future on-hand inventory and future replenishment requirements at all echelons of a distribution network. Indeed, Stenger [7] states that "the projection plan represented by any one [DRP] table is of great value in its own right because it shows the inventory levels we need to maintain to meet demand and provide customers with the service needed (through the safety stock). It also indicates when replenshments – and their quantities – need to be scheduled to make the plan work." The projection of future on-hand inventory and replenishment requirements is accomplished using standard DRP logic, described below. However, as we shall explain, standard DRP logic provides these estimates without taking into consideration the uncertainty of future demands. Specifically, DRP logic, in its calculation of projected on-hand inventory and replenishment requirements, assumes that the demand that will be realized in the future is exactly equal to the forecast and this can lead to misleading performance projections.

In this paper we present a detailed analysis on the inner workings of DRP so as to pinpoint its shortfalls and limitations, in particular, in estimating future performance measures as mentioned above. Second, we want to illustrate how the performance evaluation capability of DRP can be substantially enhanced by incorporating into its logic some simple analytical formulas. We shall derive these formulas through analyzing certain familiar inventory systems, such as the base-stock control and the (S, s) control schemes. Third, in the context of a multi-echelon distribution network, we want to illustrate how safety stock levels can be set at upstream locations so as to meet a target service level at the retail level using an echelon inventory mechanism and how this mechanism can be implemented within the DRP framework. Throughout the paper, wherever exact analysis becomes intractable or requires heavy computation, we will not hesitate to resort to approximation, seeking to derive closed-form formulas that are easy to evaluate.

The use of simple approximate formulas are justified by the enormous volume of SKUs in any inventory system for which DRP-based software is meant to generate fast decision alternatives. An alternative would be to use discrete-event simulation but experiments indicate that the required computation time for simulation, even on a powerful computer, makes analysis of large numbers of SKUs infeasible. Approximate formulas are also justified simply because in many industrial applications, demand streams fluctuate significanly over time, and exact analytical performance estimates in the case of non-stationary demands are mostly intractable and unavailable in the inventory literature.

The remainder of the paper is organized as follows. In §2 we review DRP logic and point out its weaknesses. In §3, we derive analytical approximations that enhance DRP's projections on inventory performance, under the lot-for-lot ordering rule. In §4, we review the stationary (S, s) model, focusing on bringing out approximate formulas that can be adapted to the non-stationary setting and are easy to evaluate. In §5, we show that in general, DRP ordering rules can be unified as a dynamic (S, s) rule — with time-dependent

S and s values. We then develop improved estimates on inventory performance based on the approximations in §4. In §6, we extend our discussion to multi-echelon networks, describe how our approximation formulas apply in this more general setting, and specify how to calculate safety stock levels at the different echelons so as to achieve a target fill rate at the retail level. Finally, §7 presents some numerical results comparing our analytical formulas with standard DRP logic and with simulation.

2 How Does DRP Work?

2.1 The DRP Logic

By DRP logic, we mean the mathematical algorithm used in standard DRP procedures to calculate future on-hand inventory and future replenishment requirements. We begin with a mathematical description of DRP logic. Our description is a formal abstraction from standard DRP procedures, descriptions of which are widely available in professional references such as Stenger [7], as well as in manuals of commercial software, such as those mentioned earlier. We explain the logic in terms of a single product at a single location; later, we provide an example illustrating how the logic carries over to distribution networks with multiple locations and multiple products.

DRP logic is based on a set of recursive equations that characterize the inventory dynamics over time. To be more precise, the principal logic of DRP involves the recursive calculation of the replenishment requirements, the on-hand inventory, and the backordered demand in the future for a specific product at a specific location. The first step in describing the recursion is to divide the future time horizon into a set of discrete periods of equal length, indexed by $t = 1, 2, \ldots, n$, where n is the time horizon of interest (As a practical matter, DRP logic is typically implemented using weekly or monthly periods with a time horizon of 6 months to one year.)So, assuming that we are currently at the beginning of period 1, the objective of DRP logic is to project, for the product and location of interest, the following quantities for all future periods, $t,\ t = 1, 2, \ldots, n$:

I_t := the on-hand inventory at the *end* of time period t.

B_t := the backordered demand at the *end* of time period t.

A_t := the required quantity of product needed at the beginning of period t.

Q_t := the constrained (or feasible) quantity of product needed at the beginning of period t.

R_t := the recommended order quantity at the beginning of period t.

The distinction between the requirements, A_t, and the constrained order quantity, Q_t, is important. Whereas A_t represents the quantity that is needed at the beginning of period t, Q_t reflects what is feasible, taking into account lead time constraints and order quantity restrictions. For example, if $A_t = 40$ and the maximum order quantity is 30, then Q_t would be set to 30.

To carry out the recursive calculation of these quantities, DRP assumes that the following information is known (again, for the specific product and location) at the start of time period 1:

D_t := the demand in future time period t, $t = 1, 2, \ldots, n$. Although the actual future demand is unknown, DRP assumes that the future demand in period t is known through its mean, $\mathsf{E}[D_t]$, and standard deviation, $\mathsf{sd}[D_t]$. (The standard deviation presumably comes from some forecasting procedure but how this information is generated is not relevant to DRP logic.)

T_t := the scheduled receipts at the *beginning* of time period t. This is the quantity of the product that is currently (i.e. at the beginning of period 1) scheduled to arrive at the location at the start of period t.

K_t := the safety stock requirement for period t. This is the quantity of the product that should be kept at this location in this period to protect against uncertainty in demand. (Most DRP sofware implementations provide a variety of methods to automatically calculate safety stock requirements for each period. These methods vary in degree of sophistication but, just as with forecast generation, the particular approach used is not relevant to DRP logic.)

I_0 := the on-hand inventory at the *beginning* of time period 1. This is the *current* on-hand inventory.

B_0 := the backordered demand at the *beginning* of time period 1.

The scheduled receipts, T_t, should not be confused with the constrained order quantity, Q_t. The former is an input to DRP and reflects what is actually going to be delivered at the start of period t; the latter is an output of DRP and reflects an assessment of what is additionally needed at the start of period t subject to order constraints. The scheduled receipts are sometimes referred to as "in-transit" to suggest that the goods are physically en route to the location in question.

At the beginning of each period t, the following sequence of events takes place. First, replenishments (if any) arrive, including both Q_t (the constrained order quantity, which relates to A_t and will be specified below) and T_t (scheduled receipt). These are used to first satisfy the backorders, if any, left from the previous period. Next, demand of the period, D_t, is realized and filled, which brings us to the *end* of the period, when I_t (on-hand inventory) and B_t (backorders) are updated.

To project future requirements, A_t, $t = 1, \ldots, n$, DRP makes use of the following recursive relation:

$$A_t = [B_{t-1} - I_{t-1} - T_t + D_t + K_t]^+, \tag{1}$$

where $[x]^+$ denotes $\max\{x, 0\}$. The above equation says that the replenishment requirement in each period along with any on-hand inventory from the last period and scheduled receipts due at the beginning of this period, should be able to supply the demand of the current period and any backlog from the last period, and still result in a surplus that is equal to the required safety stock for this period.

The next step in DRP logic is to modify A_t to reflect what can feasibly be obtained by the start of period t. To this end, the constrained order quantities, Q_t, are derived from A_t, by applying a set of prespecified order-size restrictions, which we refer to as *order policies* or *order rules*. Several commonly used rules are listed below. Note that all of the rules apply if and only if $A_t > 0$; otherwise, we set $Q_t = A_t = 0$.

(i) *Lot-for-Lot*: $Q_t = A_t$, i.e., there is no restriction on the order quantity.

(ii) *Min-Max*: with $Q_{\min}$ and $Q_{\max}$ being the (given) lower and upper limit on order quantities,

$$Q_t = \min\{\max(Q_{\min}, A_t), Q_{\max}\},$$

becomes the order quantity planned for period t.

(iii) *DOS* ("days of supply"): Q_t is equal to the (projected) demand over a given number of periods.

(iv) *EOQ* (economic order quantity): $Q_t = \sqrt{2\mu C/h}$, the classical EOQ formula (e.g., [6]), where μ is the average demand per period, C is a fixed cost for placing a replenishment order, and h is the inventory holding cost per period.

Note that all the rules above with the exception of (i) impose some restrictions on the order quantity. In general, we can assume there is some prespecified Q_t, which is a function of A_t (and other parameters).

The recursive equations governing the on-hand inventory, I_t, and the backorders, B_t, are simple to describe as well. First, denote the *net* inventory level at the beginning of period t as

$$Y_t = I_{t-1} - B_{t-1} + T_t, \quad t = 1, \ldots, n. \tag{2}$$

That is, Y_t accounts for the (net) balance left over from the previous period and the scheduled receipts for the current period. This amount, if positive, can be used to supply demand in the current period, along with the constrained order quantity, Q_t, which is derived from the requirement A_t following a prespecified order policy. If Y_t is negative, then it must be first deducted from A_t, before any remaining units can be used to supply demand. Hence, clearly,

$$I_t = [Q_t + Y_t - D_t]^+, \quad \text{and} \quad B_t = [D_t - Q_t - Y_t]^+. \tag{3}$$

The final calculation performed by DRP is the recommended orders. Suppose that the order lead time for a specific product at a specific location is L periods. In DRP, the recommended order quantity in period t corresponds to the calculated order quantity $L + t$ periods later:

$$R_t = Q_{t+L}, \quad t = 1, \ldots, n - L. \tag{4}$$

So, the recommended order quantity at the start of period t is whatever the constrained order quantity is for period $t + L$.

Since the future demand D_t is known only through its forecasted mean and standard deviation, to implement the recursions defined in (1, 2, 3, 4), DRP logic uses the forecasted mean demand, $\mathsf{E}[D_t]$, in place of D_t. Therefore, these quantities should be viewed as DRP's estimates of unknown future quantities. For example, R_t is DRP's estimate of the quantity that will be ordered at the start of period t standing at the beginning of period 1. In particular, R_1 is the current recommended order quantity; whereas R_5, for instance, is DRP's best guess at the quantity that is most likely to be ordered at the start of period 5.

2.2 An Example

To help illustrate DRP logic, we consider a simple distribution network, illustrated schematically in Figure 1. The network consists of three retailers supplied by one warehouse location. The retailers obtain replenishments from the warehouse, who in turn obtains replenishments from a supplier. For the purpose of this example, we focus on one product that is sold by all three retailers. We assume that the replenishment lead time for this product at each of the retailers is 5 weeks and that the replenishment lead time at the warehouse is 15 weeks. This means, for example, that if a retailer places a replenishment order at the start of week 1, then it will arrive at the start of week 6. Also, we assume that the retailers place orders in quantites of 30 and that the warehouse places orders in quantities of 200.

Let us suppose that we are currently at the beginnning of week 1 (i.e. $t = 1$) and we are interested in using DRP to project forward 26 weeks into the future. DRP logic assumes that the following information is known at this time:

- The mean and standard deviation of each retailer's future weekly demand over the horizon in question.

- The current on-hand inventory at each location, namely at all three retailers and at the warehouse.

- The scheduled receipts – i.e. a list of all shipments scheduled to arrive at each location by week over the horizon in question.

- The safety stock to be held at each location.

For this example, we assume that the weekly demand at each retailer for this product follows a discrete uniform distribution on $[0, 20]$ – that is, the weekly demand equals n with probability $p = 1/21$, $n = 0, 1, \ldots 20$. Thus, the weekly demand at each retail location has a mean of 10 and a standard deviation of $\sqrt{110/3} \approx 6.06$. We also assume that at the start of week 1, we have 60 units of on-hand inventory available at each retailer and 100 units available at the warehouse. Next, let us suppose that there are currently no scheduled receipts. Finally, we assume that the retailers wish to hold 10 units of safety stock and the warehouse wishes to hold 15 units of safety stock.

To project on-hand inventory and replenishment requirements for the next 26 weeks at each location in this network, DRP logic first starts with the retail locations and constructs a table for each retailer that reflects the weekly operation of the retailer over the next 26 weeks. The table consists of 26 columns (one for each week) and the following rows:

1. Expected (i.e. mean) demand ($\mathsf{E}[D_t]$)

2. Scheduled receipts (T_t)

3. Safety stock requirements (K_t)

4. Replenishment requirements (A_t)

5. Constrained order quantity (Q_t)

6. On-hand inventory (I_t)

7. Backordered demand (B_t)

8. Recommended orders (R_t)

Table 1 shows a completed DRP table for one retailer in our example. (Since all the retailers are identical in this example, the tables generated for the retailers are identical).

As the reader can verify, all of the projected quantities, A_t,Q_t, I_t, B_t and R_t, have been calculated using the DRP recursions specified above while respecting the order quantity restriction.

To continue with our example, let us assume that the DRP tables have been constructed for all three retailers following the DRP logic described above. The next step is to construct an analogous DRP table for the warehouse. To do this, DRP sums up the recommended order quantities by week from all three retailers and this becomes the weekly demand forecast used to drive the DRP logic at the warehouse level. That is, row 1 of the DRP table at the warehouse corresponding to the expected weekly demand at the warehouse is constructed by summing the corresponding recommended order quantities from the three retail DRP tables. Once the weekly demand is specified, the DRP recursive calculations are carried out exactly as for the retailers, resulting in a completed DRP table for the warehouse. Table 2 shows a completed DRP table for the warehouse in our example.

To summarize, the end result of the DRP calculations is a set of tables, one for each product and location in the distribution network. The key quantities that these tables contain are *projections* (equivalently, *estimates*) about the future on-hand inventory and replenishment requirements. While we have illustrated the construction of these tables using a simple example, the logic we have described carries over in a straightforward manner to much more complex networks.

2.3 Critique of DRP Logic

The primary criticism of DRP logic is that it ignores the randomness of future demand in its calculations by substituting the mean forecasted demand in the recursive equations. In our example, the mean demand at the retail level is 10 units per week and this is the number that is used in carrying out the DRP calculations. However, the standard deviation of demand is 6.06 and nowhere in the DRP logic is this fact used. (The calculation of safety stock may make use of the standard deviation of demand, but this calculation falls outside the scope of DRP logic. As stated earlier, the safety stock is assumed to be given as an input to the DRP calculations. See the Appendix for a description of how safety stock levels are calculated to achieve a target fill rate within DRP.) In fact, the implicit assumption in DRP logic is that the standard deviation of demand is 0.

Another way to view the situation is that the formal recursive equations governing DRP logic as defined in (1), (2), (3), and (4), are statements about *random* quantities. Viewed in this light, it is reasonable to ask if the DRP projections specified in the recursive equations correspond to the correct *expected values* of these random quantities. In fact, since $f(x) = (x)^+$ is a convex function of x, it is easy to see (following Jensen's inequality, or the so-called "variability ordering" — [4]) that in replacing D_t by its mean, the DRP logic will *underestimate* the expected values of A_t, I_t and B_t.

Furthermore, assuming a lot-for-lot order policy, i.e., $Q_t = A_t$, we can rewrite (3) as follows:

$$I_t = \max\{A_t + Y_t - D_t, 0\} = \max\{K_t, Y_t - D_t, 0\},$$

and

$$B_t = -\min\{A_t + Y_t - D_t, 0\}, = -\min\{\max\{K_t, Y_t - D_t\}, 0\}.$$

Note that the safety stock, K_t, need *not* always be positive.

Therefore, if the safety factor is positive, then DRP always estimates positive on-hand inventory (i.e., with zero backlog), which is at or above the level of safety stock, since $I_t \geq K_t$ from the above. On the other hand, when the safety factor is negative, DRP prescribes a backlog strategy of not keeping any on-inventory, but instead keeping a level of backlog that is at least as much as the absolute value of the (negative) safety stock. That is,

$$I_t = K_t \qquad \text{or} \qquad B_t = -K_t$$

according to $K_t > 0$ or $K_t < 0$. These predictions, of course, can be quite remote from the reality of any typical inventory system.

3 Improved DRP Projections: The Lot-for-Lot Rule

In this section, we discuss how to improve the DRP performance estimates under the lot-for-lot ordering rule. Assume that the demand in period t can be expressed as

$$D_t = \mu_t + \sigma_t Z_t$$

where μ_t is the mean demand in period t, σ_t is the standard deviation of demand in period t and $Z_t, t = 1, \ldots, n$ are i.i.d. random variables following any distribution with zero mean and unit variance. For ease of discussion, we focus on the case in which Z_t follows a standard normal distribution. The results extend readily to more general demand distributions.

To start with, assume zero lead time: $L = 0$. Suppose we are at the beginning of period t, with Y_t available as a realized (deterministic) value. We want to know what should be the requirement A_t, which once determined and ordered will arrive immediately (due to zero lead time). We want to make this decision so that B_t, the backorder at the end of the period, will satisfy the desired service level. In addition, we also want to have a good estimate on I_t, the on-hand inventory at the end of the period.

What we do *not* have access to, in determining A_t, is D_t, the demand to be realized over the period: it's actual value will only become available by the end of the period. At our decision epoch (the beginning of the period), we only know the distribution of D_t, which, as specified above, follows a normal distribution with mean μ_t and standard deviation σ_t.

Substituting (2) into (1), we can rewrite the latter as follows:

$$A_t = [D_t + K_t - Y_t]^+. \tag{5}$$

Since D_t is treated as deterministic, and replaced by μ_t in the DRP logic, we have $A_t = [\mu_t + K_t - Y_t]^+$. Therefore, with a lot-for-lot ordering policy, DRP logic basically follows a base-stock policy, with $\mu_t + K_t$ as the base-stock level: if the net inventory at the beginning of t, Y_t, falls below the base-stock level, $\mu_t + K_t$, then $A_t > 0$, and an order quantity of $Q_t = A_t$ will bring the inventory position back to the base-stock level. On the other hand, since Y_t can never exceeds the base-stock level (this can be shown inductively), $A_t = 0$ implies that Y_t is equal to the base-stock level (which could happen when there is no demand in period $t-1$). In either case, after the review (at the beginning of period t), since $L = 0$, with the instantaneous arrival of A_t (possibly zero), the on-hand inventory is brought back to the base-stock level. The inventory will then be depleted to supply the demand D_t. Write

$$K_t = k_t \sigma_t,$$

where k_t is the "safety factor." At the end of the period, we have

$$\mathsf{E}[B_t] = \mathsf{E}[D_t - \mu_t - k_t\sigma_t]^+ = \sigma_t \mathsf{E}[Z_t - k_t]^+ = \sigma_t G(k_t);$$

and

$$\mathsf{E}[I_t] = \mathsf{E}[\mu_t + k_t\sigma_t - D_t]^+ = \sigma_t \mathsf{E}[k_t - Z_t]^+ = \sigma H(k_t).$$

where

$$G(k) := \mathsf{E}[Z - k]^+ = \int_k^\infty (x - k) d\Phi(x),$$

where $\Phi(x)$ denotes the distribution function of Z, and $H(k) := k + G(k)$.

Observe that $H(x)$ is an increasing function, $G(x)$ is decreasing, and both are convex. In addition, when Z is the standard normal variate, for x reasonably large (say, $x \geq 3$), we have, approximately, $H(x) = x$, $H(-x) = 0$; and $G(x) = 0$, $G(-x) = x$. These properties will become useful in the analysis below.

Next, consider the more general case of $L \geq 0$. Suppose again, we want to determine the requirement A_t for period t. We need to make this decision at the beginning of period $t - L$ so as to place the order in time for it to arrive in period t.

Position ourselves at the beginning of period $t - L$. What we know are the following: (a) the net inventory left from the previous period, Y_{t-L}; and (b) the scheduled arrivals, $A_{t-L}, A_{t-L+1}, ..., A_{t-1}$, which had already been decided and ordered. What we do *not* know is the sequence of demand over the lead time, $D_{t-L}, D_{t-L+1}, ..., D_t$.

Denote:

$$D(s,t) := D_s + \cdots + D_t, \qquad A(s,t) := A_s + \cdots + A_t;$$

$$\mu(s,t) := \mathsf{E}[D(s,t)], \qquad \sigma(s,t) := \mathsf{sd}[D(s,t)].$$

We claim that the right A_t value should be the following:

$$A_t = [\mu(t-L,t) + k_t\sigma(t-L,t) - Y_{t-L} - A(t-L,t-1)]^+. \tag{6}$$

First, suppose the quantity inside $[\cdot]^+$ on the right hand side above is positive, and hence so is A_t. Then, at the beginning of period $t - L$, the net inventory is Y_{t-L}, which, along with the scheduled arrivals over the lead time — a total of $A(t - L, t)$, is going to supply all the demands over the lead time — a total of $D(t - L, t)$. Hence, at the end of period t, the expected backorder is:

$$\begin{aligned}
\mathsf{E}[B_t] &= \mathsf{E}[D(t-L,t) - Y_{t-L} - A(t-L,t)]^+ \\
&= \mathsf{E}[\mu(t-L,t) + Z_t\sigma(t-L,t) - \mu(t-L,t) - k_t\sigma(t-L,t)]^+ \\
&= \sigma(t-L,t)\ \mathsf{E}[Z_t - k_t]^+ \\
&= \sigma(t-L,t)\ G(k_t),
\end{aligned} \tag{7}$$

where the second equality follows from (6), assuming $A_t > 0$:

$$Y_{t-L} + A(t-L,t) = \mu(t-L,t) + k_t\sigma(t-L,t).$$

Similarly,

$$\begin{aligned}
\mathsf{E}[I_t] &= \mathsf{E}[Y_{t-L} + A(t-L,t) - D(t-L,t)]^+ \\
&= \sigma(t-L,t)\ \mathsf{E}[k_t - Z_t]^+ \\
&= \sigma(t-L,t)\ H(k_t).
\end{aligned} \tag{8}$$

Next, suppose $A_t = 0$ in (6). This implies

$$Y_{t-L} + A(t-L,t-1) \geq \mu(t-L,t) + k_t\sigma(t-L,t).$$

Hence,

$$\begin{aligned}
\mathsf{E}[B_t] &= \mathsf{E}[\mu(t-L,t) + Z \cdot \sigma(t-L,t) - Y_{t-L} - A(t-L,t-1)]^+ \\
&= \sigma(t-L,t)\ G(k_t'),
\end{aligned} \tag{9}$$

with

$$k'_t := \frac{Y_{t-L} + A(t-L,t-1) - \mu(t-L,t)}{\sigma(t-L,t)} \geq k_t. \tag{10}$$

Hence, $\mathsf{E}[B_t]$ in (9) is no greater than $\mathsf{E}[B_t]$ in (7), since $G(\cdot)$ is a decreasing function. (Intuitively, in this case, since the available (net) inventory Y_{t-L} is higher, the effective saftety factor k'_t is also higher; hence, the (projected) backorder is lower.)

Similarly, when $A_t = 0$, we have

$$\mathsf{E}[I_t] = \sigma(t-L,t)\; H(k'_t), \tag{11}$$

with k'_t following (10); and $\mathsf{E}[I_t]$ above is larger than $\mathsf{E}[I_t]$ of (8).

To summarize, we have

Proposition 1 When the order rule is lot-for-lot, i.e., $Q_t = A_t$ for all t, the required quantity needed at the beginning of period t (A_t), and hence the recommended order quantity at the beginning of period $t-L$ (R_{t-L}), should be

$$R_{t-L} = Q_t = A_t = [s_{t-L} - Y_{t-L} - A(t-L,t-1)]^+,$$

where

$$s_{t-L} = \mu(t-L,t) + k_t\sigma(t-L,t),$$

In this case, DRP follows a base-stock control mechanism, with s_{t-L} being the base-stock level (or, reorder point) for period $t-L$. The estimates for expected on-hand inventory and backorders (for period t) follow (8,11) and (7,9), respectively.

It is important to note that the reorder point, s_{t-L}, in the above proposition has a time index $(t-L)$ that is consistent with the time when the order is placed. On the other hand, we choose to index the safety factor k_t by t, since it relates more closely to the on-hand inventory and backorders in period t.

We should point out that the expression in (6) that specifies the requirement A_t can also be derived directly from the DRP recursion. From (2,3), with $Q_t = A_t$, we have

$$\begin{aligned} Y_t &= I_{t-1} - B_{t-1} + T_t \\ &= Y_{t-1} + A_{t-1} - D_{t-1} \\ &= Y_{t-2} + A_{t-2} - D_{t-2} + A_{t-1} - D_{t-1} \\ &= \cdots \\ &= Y_{t-L} + A(t-L,t-1) - D(t-L,t-1). \end{aligned} \tag{12}$$

Substituting the above into (5), we recover (6), noticing that since DRP treats demand as deterministic, $D_t + D(t-L, t-1) = \mu(t-L, t)$. In other words, if DRP uses (6) directly to generate the requirements A_t, instead of going through the recursions that involves I_t and B_t, the estimates on these requirements will be protected from the inaccuracies involved in estimating I_t and B_t.

Following the expressions derived above for A_t, B_t and I_t, the DRP recursion can be carried over to future periods through $Y_{t+1} = I_t - B_t$. Specifically, starting from period 1, we first derive A_{L+1}, followed by B_{L+1} and I_{L+1}, and then Y_{L+1}. Continue this for $t = L+2, ..., 2L$, assuming that $Y_1, Y_2, ..., Y_L$ have all been prespecified, i.e., given at beginning of period 1. (Note that nothing in periods 1 through L can be affected by any replenishment decision, at the beginning of period 1, due to the lead time.) In return, we will derive $Y_{L+2}, ..., Y_{2L}$. Next, we will derive A_{2L+1}, at the beginning of period $L+1$. At that point, we will need the value of Y_{L+1}, which has already been derived; and in return, we will derive Y_{2L+1}. The recursion then continues.

4 The (S, s) Inventory System

Having analyzed the lot-for-lot rule, we want to study next other rules, which impose restrictions on the order quantities. It will become evident below (§5) that all these rules can be unified into a dynamic (S, s) control scheme — dynamic in the sense that the control parameters, S and s, in general change over time. In order to analyze the dynamic (S, s) control rule, here we first review the standard (S, s) inventory model under stationary demand, and develop some simple approximations for its performance.

The (S, s) policy is known to be optimal in a quite general setting, in terms of miminizing the total costs of placing orders, keeping inventory, and paying penalty of backorders. Refer to Scarf [5]. Our concern here is quite different, however; our focus is on performance evaluation, rather than cost minimization. Our objective is to develop certain simple approximate formulas that are not only easy to evaluate but also useful in suggesting approximations in the non-stationary (i.e., dynamic) setting. More refined approximations for the stationry (S, s) systems are available in the literature, e.g., Tijms [9], although it is not clear how they can be adapted to the non-stationary setting.

4.1 Stationary Analysis

Consider an (S, s) inventory model with periodic review. A review takes place at the beginning of each period t, at which time the inventory position is updated and a replenishment decision is made following the (S,s) rule — order up to S if and only if the inventory position (on-hand plus on-order minus back-order) falls below s.

Observe that, defined as above, the inventory position, X_t, always takes values between s and S. In particular, if an arriving (batch) order brings the on-hand inventory level to *below* s, the inventory position is immediately brought up to S.

Suppose, without loss of generality, that we start with $X_0 = S$. Then, we have a regenerative structure, with the cycle being the time between two consecutive replenishment orders. Given x ($x = s, s+1, ..., S$), let

$$T_x = \min\{t:\ D_1 + \cdots + D_t > S - x\}$$

be the time, in a cycle, for the inventory position to drop from S to below x. In particular, T_s is the cycle length.

Since demands are i.i.d., we can view $T_x - 1$ as the number of renewals by time $S - x$ of a renewal process with "interarrival time" D_t. (Refer to, e.g., Ross [4], §3.4, in particular pp. 69-70.) Hence,

$$\mathsf{E}[T_x] = M(S-x) + 1, \qquad \mathsf{E}[T_s] = M(S-s) + 1,$$

where

$$M(x) = \sum_{n=1}^{\infty} F_{*n}(x)$$

is the renewal function, with $F_{*n}(\cdot)$ denoting the n-fold convolution of the (per-period) demand distribution, $F(\cdot)$.

Furthermore, from standard renewal theory ([4], Chapter 3), we know

$$\mathsf{P}[X \geq x] = \lim_{t\to\infty} \mathsf{P}[X_t \geq x] = \frac{\mathsf{E}[T_x]}{\mathsf{E}[T_s]} = \frac{1 + M(S-x)}{1 + M(S-s)},$$

for $x \in [s, S]$. Alternatively,

$$\begin{aligned}
\mathsf{P}[X = x] &= \frac{M(S-x) - M(S-x-1)}{1 + M(S-s)}, \quad x = s, ..., S-1; \\
\mathsf{P}[X = S] &= \frac{1}{1 + M(S-s)}.
\end{aligned}$$

The mean then follows:

$$\mathsf{E}[X] = \frac{S + sM(S-s) + \sum_{x=1}^{S-s-1} M(x)}{1 + M(S-s)}.$$

To characterize the on-hand inventory and the backorders, the key is to observe that by time t all orders that are placed before or at $t - L$ will have arrived. In other words, the on-order quantities included in the inventory position at $t-L$ will all have arrived by t. As before, let $D(t-L,t)$ denote the total demand over the $L+1$ periods: $t-L, t-L+1, ..., t$. We have

$$I_t = [X_{t-L} - D(t-L,t)]^+, \tag{13}$$

where $[x]^+ = \max\{x, 0\}$. Note that in (13), $D(t-L,t)$ is independent of X_{t-L}. Similarly, we have

$$B_t = [D(t-L,t) - X_{t-L}]^+. \tag{14}$$

Letting $t \to \infty$, and denoting the limits by omitting the time index, we have

$$I = [X - D(L+1)]^+, \quad \text{and} \quad B = [D(L+1) - X]^+, \tag{15}$$

where $D(L+1)$ denotes a random variable that is equal in distribution to $D(t-L,t)$ and is independent of X.

4.2 Approximations

The renewal function is known to have a linear asymptote:

$$M(x) \sim \frac{x}{\mu} + \frac{m_2}{2\mu^2} - 1, \tag{16}$$

for large x, where $\mu = \mathsf{E}(D)$ and $m_2 = \mathsf{E}(D^2)$ are the first two moments of the (per period) demand distribution. (More precisely, the difference between $M(x)$ and the linear function on the right hand side above goes to zero when $x \to \infty$.)

Since in most applications the (S,s) values are large or moderately large, we can use the linear asymptote in (16) as an approximation for the renewal function. This way, the probability distribution of the inventory position X in §4 becomes:

$$\begin{aligned} \mathsf{P}[X = x] &= \frac{1}{\frac{m_2}{2\mu} + S - s}, \qquad x = s, ..., S-1; \\ \mathsf{P}[X = S] &= \frac{\frac{m_2}{2\mu}}{\frac{m_2}{2\mu} + S - s}; \end{aligned}$$

where note that $\frac{m_2}{2\mu}$ is the average "undershoot" — the gap between the inventory position when an order is placed (which is necessarily smaller than s) and the reorder point s. Hence,

$$\mathsf{E}[X] = \frac{\frac{1}{2}(S+s-1)(S-s) + \frac{m_2}{2\mu}S}{\frac{m_2}{2\mu} + S - s} = \frac{\mu(S-s)(S+s-1) + m_2 S}{m_2 + 2\mu(S-s)}. \tag{17}$$

In some special cases, the above reduces to a uniform distribution over the integers $\{s, ..., S\}$, and $\mathsf{E}[X] = (S+s)/2$. For instance, this is the case when the (per period) demand follows a truncated normal distribution with unit mean and unit coefficient of variation (hence, $\mu = 1$ and $m_2 = 2$). Note that for truncated normal distributions, just like normal distributions, assuming a unit coefficient of variation loses no generality; hence, unit mean is the only significant assumption here.

Next, making use of (15), along with the normal approximation of $D(L+1)$, we get the following approximation:

$$\mathsf{E}[B] = \sum_{x=s}^{S} \sigma\sqrt{L+1}G(\frac{x-(L+1)\mu}{\sigma\sqrt{L+1}})\mathsf{P}[X=x].$$

Substituting the uniform distribution of X into the above, we have

$$\begin{aligned}\mathsf{E}[B] &= \frac{1}{1+S-s}\sum_{x=s}^{S}\sigma\sqrt{L+1}G(\frac{x-(L+1)\mu}{\sigma\sqrt{L+1}}) \\ &\leq \sigma\sqrt{L+1}G(\frac{s-(L+1)\mu}{\sigma\sqrt{L+1}}), \qquad (18)\end{aligned}$$

where the inequality follows from the fact that $G(x)$ is decreasing in x.

Once $\mathsf{E}[B]$ is derived, $\mathsf{E}[I]$ follows from (15), we have

$$\mathsf{E}[I] = \mathsf{E}[B] + \mathsf{E}[X] - \mu(L+1).$$

For instance, from (18), taking into account $G(k) = H(k) - k$, and approximating X with a uniform distribution, we have

$$\mathsf{E}[I] = \sigma\sqrt{L+1}H(\frac{s-(L+1)\mu}{\sigma\sqrt{L+1}}) + \frac{S-s}{2}. \qquad (19)$$

Alternatively, $\mathsf{E}[X]$ can follow the approximation in (17).

5 Improved DRP Projections: The Dynamic (S, s) Rule

Observe that when following any policy other than lot-for-lot, the equation in (6) governing the requirements should be modified as follows:

$$A_t = [\mu(t-L,t) + k_t\sigma(t-L,t) - Y_{t-L} - Q(t-L,t-1)]^+ \qquad (20)$$

where

$$Q(t-L,t-1) = Q_{t-L} + \cdots + Q_{t-1}$$

replaces $A(t-L,t-1)$. Hence, from $A_t > 0$, we have

$$\begin{aligned}A_t &= \mu(t-L,t) + k_t\sigma(t-L,t) - Y_{t-L} - Q(t-L,t-1) \\ &= s_{t-L} - Y_{t-L} - Q(t-L,t-1).\end{aligned}$$

That is, s_{t-L} is the reorder point, since following the DRP logic, an order (R_{t-L}) is placed (at $t-L$) if and only if $A_t > 0$. From the above expression, A_t is the required quantity to bring the inventory position (again, at $t-L$) back to s_{t-L}. (This is consistent with the base-stock mechanism when there is no order-size restrictions.)

Now, suppose the ordering policy is a dynamic (S,s) rule. Specifically, when $A_t > 0$, we want to bring the inventory position to S_{t-L} ($> s_{t-L}$), and an additional amount, $S_{t-L} - s_{t-L}$, is needed. Hence, the order quantity is:

$$Q_t = A_t + S_{t-L} - s_{t-L} = S_{t-L} - Y_{t-L} - Q(t-L, t-1), \tag{21}$$

when $A_t > 0$. ($Q_t = 0$ when $A_t = 0$).

On the other hand, given any order rule with a prespecified Q_t (as a function of $A_t > 0$, as discussed in §2.1), we can implement this rule as a dynamic (S,s) rule by letting [cf. (21)]

$$S_{t-L} = Q_t + s_{t-L} - A_t = Q_t + Q(t-L, t-1) + Y_{t-L},$$

when $A_t > 0$. When $A_t = 0$, we have $Q_t = 0$, and from the first equation above, we have $S_{t-L} = s_{t-L}$.

Implementation of a dynamic (S,s) policy within DRP is straightforward: once the parameters S_t and s_t for all t are specified, following (21) requires knowledge only of the quantities A_t, which are directly available within DRP. Therefore, here we focus on the dynamic (S,s) rule as a unified ordering rule for DRP when there are order-quantity restrictions.

We now turn to performance projections for DRP under the dynamic (S,s) rule. Note that the inventory position at $t-L$ is:

$$X_{t-L} = Y_{t-L} + Q(t-L, t), \tag{22}$$

i.e., the net inventory plus the on-order quantities (orders that have been placed but have yet to arrive). In particular, when $Q_t > 0$, from (21) and (22), we have,

$$S_{t-L} = Y_{t-L} + Q(t-L, t) = X_{t-L},$$

i.e., after the order is placed, the inventory position is brought up to S_{t-L}.

Just as in (12), we can iterate on (2,3) to obtain

$$Y_t = Y_{t-L} + Q(t-L, t-1) - D(t-L, t-1),$$

which, upon substitution back into (3), yields:

$$I_t = [Y_{t-L} + Q(t-L, t) - D(t-L, t)]^+ = [X_{t-L} - D(t-L, t)]^+,$$

and

$$B_t = [D(t-L,t) - Y_{t-L} - Q(t-L,t)]^+ = [D(t-L,t) - X_{t-L}]^+.$$

These are exactly the same formulas as in (13,14).

Therefore, we can adapt the approximations in the stationary (S,s) model. For instance, based on the backorder approximation in (18), and taking into account the formulas in (7,9) for the lot-for-lot case, we develop the following approximation:

$$\mathsf{E}[B_t] = \sigma(t-L,t)\cdot G(k_t), \quad \text{or} \quad \mathsf{E}[B_t] = \sigma(t-L,t)\cdot G(k_t'), \tag{23}$$

according to whether $A_t > 0$ or $A_t = 0$, where

$$k_t' := \frac{Y_{t-L} + Q(t-L,t-1) - \mu(t-L,t)}{\sigma(t-L,t)},$$

which, again, can be verified to dominate k_t.

To approximate $\mathsf{E}[I_t]$, just as in the stationary case, we make use of the identity [from (13]

$$I_t - B_t = X_{t-L} - D(t-L,t).$$

Hence,

$$\mathsf{E}[I_t] = \mathsf{E}[B_t] + X_{t-L} - \mu(t-L,t),$$

where $\mathsf{E}[B_t]$ follows the approximations in (23), and X_{t-L} follows the expression in (22). According to the two cases in (23), and making use of the expression in (22) and the relation: $H(k) = k + G(k)$, we have:

$$\mathsf{E}[I_t] = \sigma(t-L,t)\cdot H(k_t) + S_{t-L} - s_{t-L}, \tag{24}$$

when $A_t > 0$ (note in this case $X_{t-L} = S_{t-L}$); and when $A_t = 0$:

$$\begin{aligned} \mathsf{E}[I_t] &= \sigma(t-L,t)\cdot H(k_t') - k_t'\sigma(t-L,t) + X_{t-L} - \mu(t-L,t) \\ &= \sigma(t-L,t)\cdot H(k_t'). \end{aligned} \tag{25}$$

Approximating $H(k_t')$ by k_t' in the above equation, we have

$$\mathsf{E}[I_t] = k_t'\sigma(t-L,t) = X_{t-L} - \mu(t-L,t).$$

The above has the intuitive interpretation that if no order is placed at $t-L$ (i.e., $R_{t_L} = Q_t = A_t = 0$), then the expected on-hand inventory at t is simply the inventory position — of which all the on-order quantities will have arrived by t — minus the demand over the lead time, period $t-L$ through period t.

To implement the above approximations requires the derivation of Y_t, which is involved in both X_t and A_t. Recursively, Y_t can be approximated by its mean:

$$\mathsf{E}[Y_t] = \mathsf{E}[I_{t-1}] - \mathsf{E}[B_{t-1}].$$

Alternatively, a cruder approximation is to forego the distinction between the two cases $A_t > 0$ and $A_t = 0$. (Observe that this distinction is not present in the stationary case: it is averaged out in each regenerative cycle.) Specifically, ignore the k_t' case in (23), and approximate X_{t-L} by $(S_{t-L} + s_{t-L})/2$. This way, we have

$$\mathsf{E}[B_t] = \sigma(t-L,t) \cdot G(k_t), \tag{26}$$

and

$$\begin{aligned} \mathsf{E}[I_t] &= \sigma(t-L,t) \cdot H(k_t) - k_t\sigma(t-L,t) + (S_{t-L} + s_{t-L})/2 - \mu(t-L,t) \\ &= \sigma(t-L,t) \cdot H(k_t) + (S_{t-L} - s_{t-L})/2. \end{aligned} \tag{27}$$

Both are consistent with the stationary approximations in (18) and (19).

Proposition 2 With order-quantity restrictions, the implementation of DRP can be unified into a dynamic (S,s) control rule. Under this rule, the calculated quantity of product needed at the beginning of period t (Q_t), and hence the recommended order quantity at the beginning of period $t-L$ (R_{t-L}), should be

$$R_{t-L} = Q_t = [A_t + S_{t-L} - s_{t-L}] \cdot 1[A_t > 0],$$

where $1[\cdot]$ denotes the indicator function, and

$$A_t = [s_{t-L} - Y_{t-L} - Q(t-L,t-1)]^+,$$

following (20), with

$$s_{t-L} = \mu(t-L,t) + k_t\sigma(t-L,t).$$

The relationship between Q_t and S_{t-L} follows (21); in particular, when $A_t > 0$, $Q_t = S_{t-L} - Y_{t-L} - Q(t-L,t-1)$; and when $A_t = 0$, $Q_t = 0$ and $S_{t-L} = s_{t-L}$. The estimates on the expected on-hand inventory and backorders (for period t) follow (24,25) and (23), respectively; or alternatively, follow (27) and (26), respectively;

6 Multi-Stage Models

Here we extend the models developed above to a distribution network. For ease of exposition, we consider a model that consists of a central stocking facility (e.g., a warehouse, or depot) supplying a set of local stocking facilities (e.g., retailers, or outlets), each of which, in turn, supplies it own customer demand. The analysis applies as well to distribution networks with more levels.

Same as in the single-stage models, our focus here is on performance evaluation, via simple approximations. We do not address the issue of cost optimization, for which many of the models can be found in the survey article by Federgruen [1].

Suppose the central warehouse is numbered as stage 0, and the set of retailers, numbered as stages $1, ..., c$. Quantities that relate to these stages will be subscripted or superscripted (if there is already a subscript for time) by their stage indices. For instance, the demand stream to retailer i, $i = 1, ..., c$, is $\{D_t^i,\ t = 1, 2, ...\}$. Assume the demands are independent among the retailers, and independent over the time periods.

Suppose the lead time at each retailer i is L_i periods, a deterministic constant, for $i = 1, ..., c$. For instance, this can be the transportation time for a replenishment order to travel from the warehouse to the retailer; or, in the case that stage i is a plant, the cycle time to build a customer order. Suppose the lead time at the warehouse is L_0.

The analysis of each retailer i follows the same discussion as before. In particular, write

$$s_t^i = \mu_i(t, t + L_i) + k_{t+L_i}^i \sigma_i(t, t + L_i);$$

and if a set of target fill rates at each of the retailers, $\beta_i, i = 1, \ldots, c$, has been specified, then the safety factor k_t^i is obtained from solving the following equations [cf. (28, 29) in the Appendix]:

$$\sigma_i(t, t + L_i)G(k_{t+L_i}^i) = (1 - \beta_i)[\Delta_t^i + \frac{\mathsf{E}(D_t^i)^2)}{2\mathsf{E}(D_t^i)}],$$

(where $\Delta_t^i := S_t^i - s_t^i$ is assumed given); or, from

$$\sigma_i(t, t + L_i)G(k_{t+L_i}^i) = (1 - \beta_i)\mathsf{E}(D_t^i),$$

in the special case of i following a lot-for-lot rule.

The expected number of backorders is:

$$\mathsf{E}[B_t^i] = \sigma_i(t - L_i, t)G(\frac{s_{t-L}^i - \mu_i(t - L_i, t)}{\sigma_i(t - L_i, t)}),$$

corresponding to the k_t case in (23), i.e., when $A_t^i > 0$. For the other case ($A_t^i = 0$), the expression is similar, corresponding to the k_t' case in (23). The expected on-hand inventory is

$$\mathsf{E}[I_t^i] = \mathsf{E}[B_t^i] + X_{t-L_i}^i - \mu_i(t - L_i, t).$$

where $X_{t-L_i}^i$ denotes the inventory position.

To analyze the warehouse, we follow the standard notion of "echelon stock." That is, we aggregate stage 0, along with stages $i = 1, ..., c$, into a single system, indexed by E (for "echelon"). This aggregated system supplies the superposition of all c demand streams. Note that each demand process D_t^i still has its own lead time L_i. Let

$$L_{\max} := \max\{L_i,\ i = 1, ..., c\},$$

Then, (13) and (14) should be modified, with L replaced by $L_E := L_0 + L_{\max}$.

Therefore, the (lower) s value for this aggregated system should be set, in period t, as follows:

$$s_t^E = \mu_E(t, t + L_E) + k_{t+L_E}^E \sigma_E(t, t + L_E),$$

where,

$$\mu_E(t, t + L_E) := \sum_{i=1}^{c} \mu_i(t, t + L_E),$$

$$\sigma_E(t, t + L_E) := [\sum_{i=1}^{c} \sigma_i^2(t, t + L_E)]^{1/2}.$$

As for the echelon safety factor, $k_{t+L_E}^E$, it can either be prespecified by the user, or, if the safety stock levels are to be set automatically based on target fill rates specified at the retail level, it can be obtained from the solution to [cf. (28) of the Appendix]:

$$\sigma_E(t, t + L_E) G(k_{t+L_E}^E) = (1 - \beta_{\max})(\Delta_t^E + \gamma_t^E),$$

with $\Delta_t^E := S_t^E - s_t^E$ assumed given, $\beta_{\max} := \max\{\beta_i,\ i = 1, ..., c\}$, where β_i is the target fill rate at retailer i, and

$$\gamma_t^E := \frac{\mathsf{E}[(\sum_{i=1}^c D_t^i)^2]}{2\mathsf{E}(\sum_{i=1}^c D_t^i)} = \frac{\sum_{i=1}^c (\sigma_t^i)^2 + (\sum_{i=1}^c \mu_t^i)^2}{2\sum_{i=1}^c \mu_t^i}.$$

That is, stage 0 should place an order at the beginning of period t if and only if the *echelon* inventory position falls below s_t^E. This echelon inventory position is the inventory position with the aggregated system considered as a single stage. It includes all the on-hand inventory at both the warehouse and all the retailers, plus any replenishment orders placed by the warehouse that have not yet arrived, minus any customer demand that is backorded at all the retailers. Hence, this echelon inventory position will only change whenever there is a customer order arriving at a retailer, and whenever the warehouse places a replenishment order. It will not change when a retailer takes supply from or registers a backorder with the warehouse. When stage 0 issues a replenishment order, say, at the beginning of period t, the order quantity is to bring the echelon stock position to $S_t^E = \Delta_t^E + s_t^E$. Hence, $\Delta_t^E + \gamma_t^E$ is the estimate of this order quantity.

Estimates of $\mathsf{E}[B_t^E]$ and $\mathsf{E}[I_t^E]$, like those at the retailers, follow in the same vein as those in Proposition 2.

7 Numerical Examples

We present results from some numerical experiments designed to compare standard DRP estimates to those obtained using the approximations developed in §3 and §5. In all of the experiments, we use discrete event simulation as a benchmark for our comparison. The purpose of the experiments is to compare the accuracy of the DRP estimates as well as the computation time required to obtain the estimates. We conducted two sets of experiments. The first set is based on the example described in §2.2 in which demand is stationary. The second set explores the accuracy of the approximations in the case of non-stationary demand and is based on actual demand history for four different SKUs obtained from a large retailer.

For the first set of experiments, we modify the example of §2.2 so that the order quantity rule is lot-for-lot. We then compare estimates of on-hand inventory at the start of week 10 at one of the retailers obtained using three approaches: simulation, standard DRP, and improved DRP estimates based on approximations developed in §3 and §5. We carry out the comparison at several different safety stock levels. Figure 2 plots the estimates of on-hand inventory for all three methods. Note that the simulation and approximation values are essentially indistinguishable. For high safety stock values, the standard DRP estimates are close to those obtained from simulation. However, at lower values, the estimates deteriorate significantly. The reason for this behavior can be understood by referring back to the recursive equations governing DRP logic. When the safety stock levels are high, stockouts occur rarely and thus little distortion results when interchanging the expectation operation with the "max" operation in (5). At lower safety stock levels, this interchange leads to greater distortion, and, thus, poorer estimates of on-hand inventory. The situation for backorders, although not displayed here, is exactly analogous.

In the second set of experiments, we are largely interested in seeing how well our approximations for the dynamic (S, s) order policy work in the case of non-stationary demands. To this end, we compare estimates of on-hand inventory and backorders by week for four different SKUs at one retail location, each with a distinct non-stationary demand pattern.

For all of these experiments, the coefficient of variation of demand is set to be 0.3 and demand per week is assumed to follow a truncated normal distribution. (We varied the coefficient of variation and found that the quality of the estimates is not significantly affected.) We use the method outlined in the Appendix to set s_t in order to meet a target fill rate of 95% each week. S_t is set to be s_t plus one week of demand. The review period is one week. Product 1 has a lead time of 5 weeks, Product 2 a leadtime of 1 week, Product 3 a lead time of 4 weeks, and Product 4 a lead time of 4 weeks. We apply the formula (27) in §5 to obtain estimates of on-hand inventory at the end of each week for 47 weeks and compare these to estimates obtained via simulation. Figures 3–6 plot week-ending on-hand inventory estimates for each week obtained via analytical approximation and simulation.

In general, the approximation appears "smoother" than the simulation results: When sharp shifts in demand occur from one week to the next the results from simulation indicate a more dramatic shift in on-hand inventory than the approximation estimates. One explanation for this phenomenon is that the approximation formula in (27) is based on the approximation in the stationary case, which contribute to the smoothing effect by averaging out the two cases $A_t > 0$ and $A_t = 0$. As for computational effort, the time required to run 30 replications for all four SKUs is approximately 0.5 seconds on an IBM RISCSystem6000 Model 43P-140. To apply the approximation formulas for these SKUs required 0.009 seconds, approximately 55 times faster than simulation. Hence, for instance, in evaluating the performance of a system of 100,000 SKUs, the simulation would take about 3.5 hours, while the approximation only needs less than 4 minutes.

A Appendix: Setting Safety Stock Levels in DRP

In the preceding sections, we have assumed that the safety stock levels, K_t, are prespecified – that is, they are given as input to the DRP logic. Most DRP systems allow a number of methods for automatically setting the safety stock levels. One such method, which is of particular importance because of its wide use, sets safety stock levels based on achieving a target fill rate. For example, the DRP system may set safety stock levels based on a user-input target fill rate of 95% at the retail level. Here, for completeness, we present the formulas used to set this type of safety stock. We focus on the case of a single location; the analysis extends readily to distribution networks such as those in §6.

For ease of discussion, here we assume stationary demand, and omit the time index t wherever possible. Suppose the service requirement is that the fraction of demand backordered should be limited to $1-\beta$, where β is the required fill rate.

To characterize the fraction of backordered demand, we need to pick a "typical" time frame. In the standard inventory literature, this is taken to be the time between two consecutive orders — known as a regenerative cycle (when demands are independent and identically distributed). The required fraction is then the ratio of the average number of backorders to the average number of demand units, both over the regenerative cycle.

In the stationary (S, s) model, the average number of backorders, following (18), is approximated by

$$\sigma\sqrt{L+1}G(\frac{s-(L+1)\mu}{\sigma\sqrt{L+1}});$$

whereas the average demand per cycle is

$$(\frac{S-s}{\mu}+\frac{m_2}{2\mu^2})\mathsf{E}(D) = S - s + \frac{m_2}{2\mu},$$

where the first factor on the left hand side is the expected cycle length $\mathsf{E}[T_s] = 1 + M(S-s)$ (refer to §4.1) with $M(S-s)$ approximated by the linear asymptote in (16). Note that the right hand side above is nothing but the expected order quantity; in particular, $\frac{m_2}{2\mu}$ is the expected undershoot.

Therefore, based on the above, under the dynamic (S, s) rule, k_t should be the solution to:

$$\sigma(t-L,t)G(k_t) = (1-\beta)[\Delta_{t-L} + \frac{\mathsf{E}(D^2_{t-L})}{2\mathsf{E}(D_{t-L})}], \tag{28}$$

where $\Delta_{t-L} := S_{t-L} - s_{t-L}$ is assumed given.

The lot-for-lot rule is equivalent to $S_{t-L} = s_{t-L}$, or $\Delta_{t-L} = 0$. Also, since in every period an order is placed, the order quantity is simply equal to the demand. Hence, the equation in (28) is reduced to:

$$\sigma(t-L,t)G(k_t) = (1-\beta)\mathsf{E}(D_{t-L}). \tag{29}$$

Note that the equations in (28) and (29) are easily solved through the Newton's method. Write the equations in the form of $G(k) = c$. The Newton's iteration, indexed by the superscript (n), is as follows:

$$k^{(n+1)} = k^{(n)} - \frac{G(k^{(n)}) - c}{G'(k^{(n)})}.$$

When demand follows a normal distribution, $G'(k) = -1 + \Phi(k)$, where Φ denotes the distribution function of the standard normal variate.

References

[1] Federgruen, A., Centralized Planning Models for Multi-Echelon Inventory Systems under Uncertainty, in: S.C. Graves, A.H.G. Rinnooy Kan and P.H. Zipkin (eds.) *Logistics of Production and Inventory,* North-Holland, 133-173, 1993.

[2] Martin, A.J., *DRP Distribution Resource Planning: Distribution Management's Most Powerful Tool*, Prentice-Hall, Englewood Cliffs, NJ, and Oliver Wight Ltd Publications, INC., Essex Junction, VT, 2nd ed., 1990.

[3] Mlot, B., DiFrancesco, L., Perry, D., Landvater, D. and Martin, A., Distribution Resource Planning: The Critical Link from Final Point of Manufacture to Final Point of Sale, *The 39th APICS International Conference Proceedings,* 1996, pp. 294-297.

[4] Ross, S.M. *Stochastic Processes,* Wiley, New York, 2nd edition, 1996.

[5] Scarf, H., The Optimality of (s, S) Policies in the Dynamic Inventory Problem, in: K. Arrow, S. Karlin and P. Suppes (eds.), *Mathematical Methods in the Social Sciences,* Stanford University Press, Stanford, CA, 1960.

[6] Silver, E.A. and Peterson, R., *Decision Systems for Inventory Management and Production Planning,* Wiley, New York, 2nd ed., 1985.

[7] Stenger, A.J., Distribution Resource Planning, in: *The Logistics Handbook,* J.F. Robeson and W.C. Copacino (eds.), The Free Press, New York, 1994.

[8] Smith, B.T., Using Distribution Resource Planning to Manage Inventories in Multiple Locations, *The 39th APICS International Conference Proceedings,* 1996, pp. 27-29.

[9] Tijms, H., *Stochastic Modeling and Analysis: A Computational Approach*, Wiley, New York, 1994.

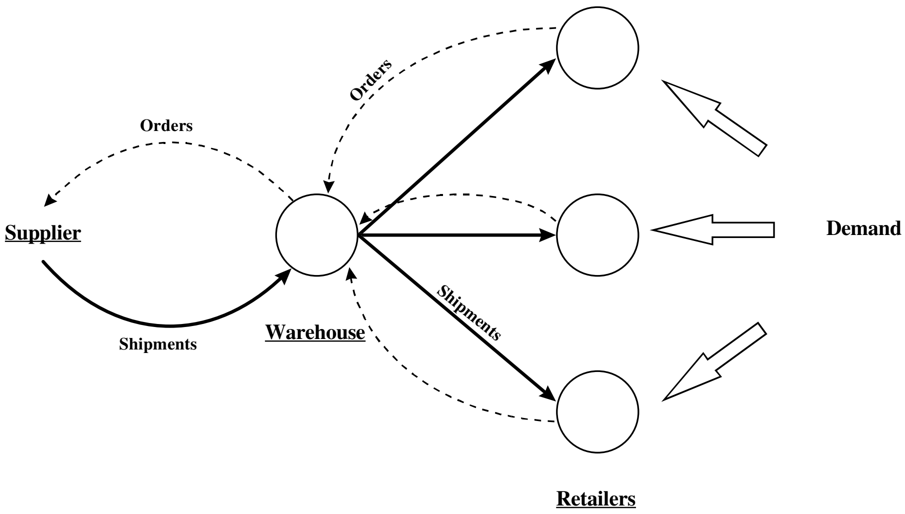

Figure 1: Layout of distribution network example

Planning parameters

Order Quantity	30
Lead Time	5
Safety Stock	10

Retailer #1

Period	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Demand		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Scheduled Receipts		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Safety Stock		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
On-Hand (pd end)	60	50	40	30	20	10	30	20	10	30	20	10	30	20	10	30	20
Requirements		0	0	0	0	0	10	0	0	10	0	0	10	0	0	10	0
Planned Receipts		0	0	0	0	0	30	0	0	30	0	0	30	0	0	30	0
Backorders		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rec. Orders		30	0	0	30	0	0	30	0	0	30	0	0	30	0	0	30

Table 1: Sample DRP Table for one retail location

Planning parameters

Order Quantity	200
Lead Time	15
Safety Stock	15

Warehouse

Period	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Demand		90	0	0	90	0	0	90	0	0	90	0	0	90	0	0	90
Scheduled Receipts		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Safety Stock		15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
On-Hand (pd end)	200	110	110	110	20	20	20	0	0	0	0	0	0	0	0	0	0
Requirements		0	0	0	0	0	0	85	85	85	175	175	175	265	265	265	355
Planned Receipts		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200
Backorders		0	0	0	0	0	0	70	70	70	160	160	160	250	250	250	140
Rec. Orders		200	200	0	200	0	0	0	0	0	200	0	0	0	0	0	200

Table 2: Sample DRP Table for warehouse

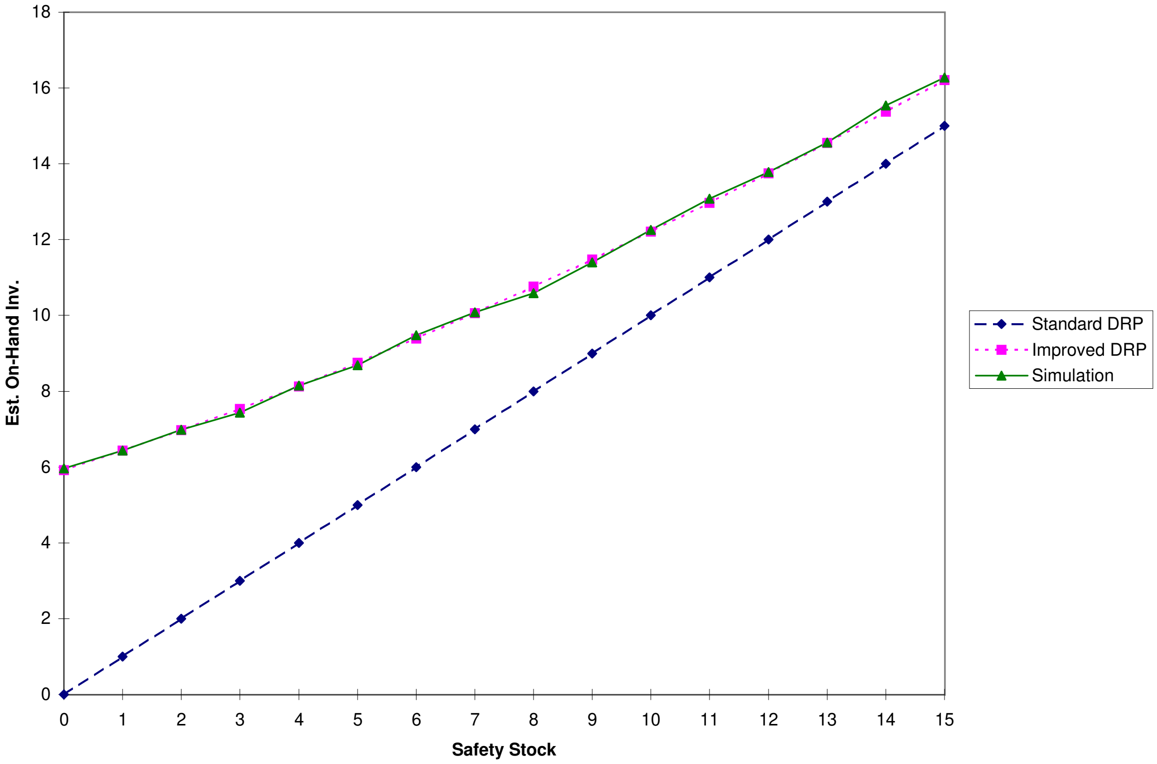

Figure 2: Estimates of Week 10 On-Hand Inventory as Function of Safety Stock for Stationary Demand Example using Lot-For-Lot Order Policy

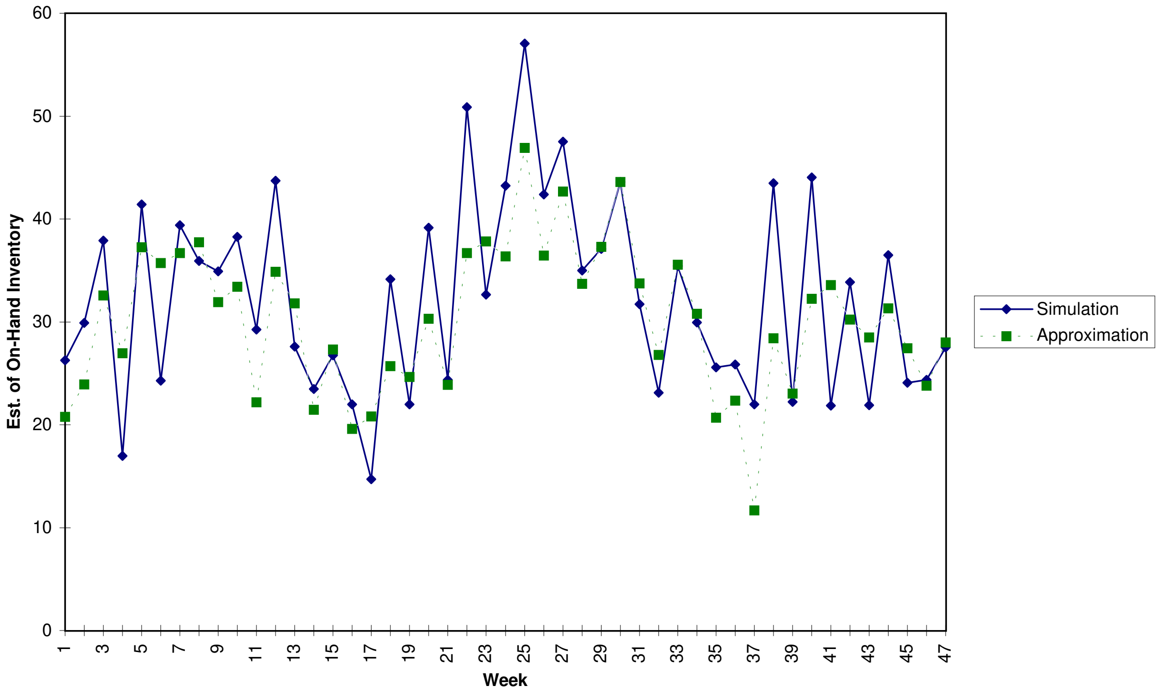

Figure 3: Estimates of Week-Ending On-Hand Inventory Levels for Product 1 with Non-Stationary Demand and Dynamic (S, s) Order Policy

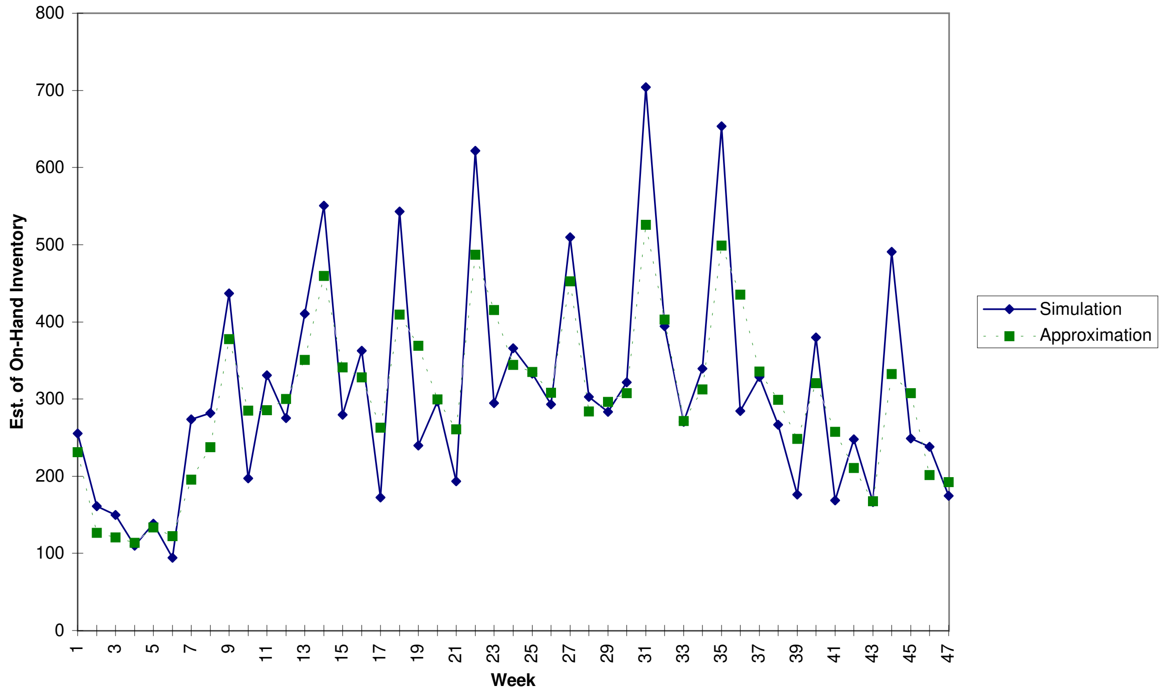

Figure 4: Estimates of Week-Ending On-Hand Inventory Levels for Product 2 with Non-Stationary Demand and Dynamic (S, s) Order Policy

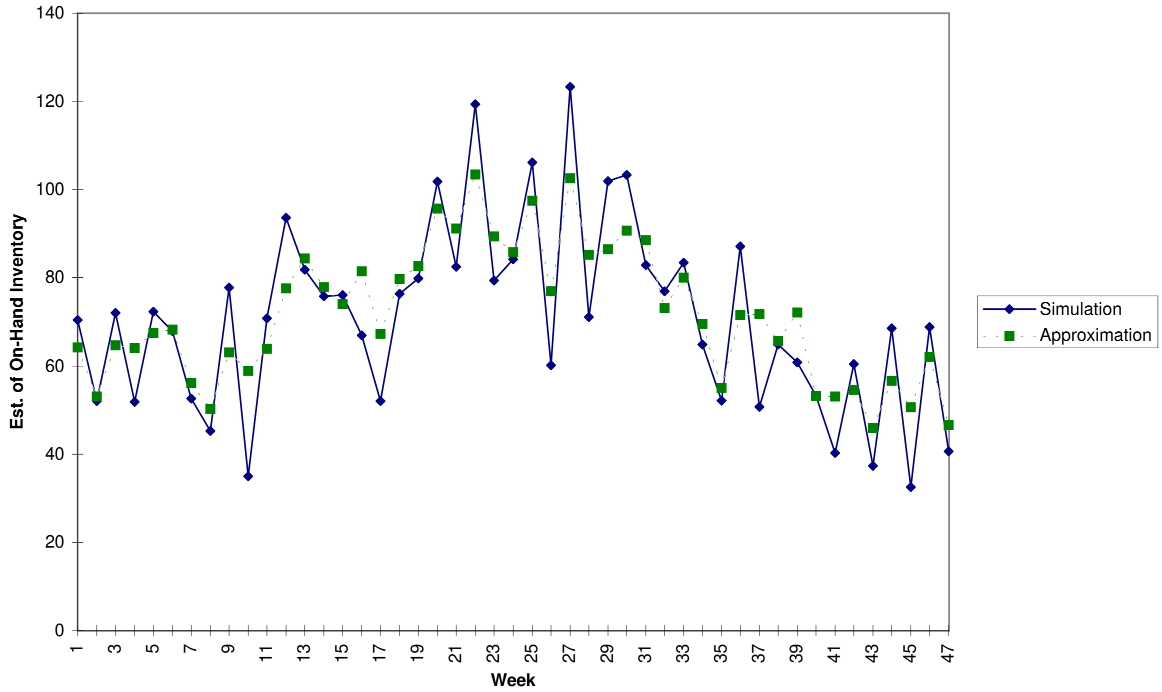

Figure 5: Estimates of Week-Ending On-Hand Inventory Levels for Product 3 with Non-Stationary Demand and Dynamic (S, s) Order Policy

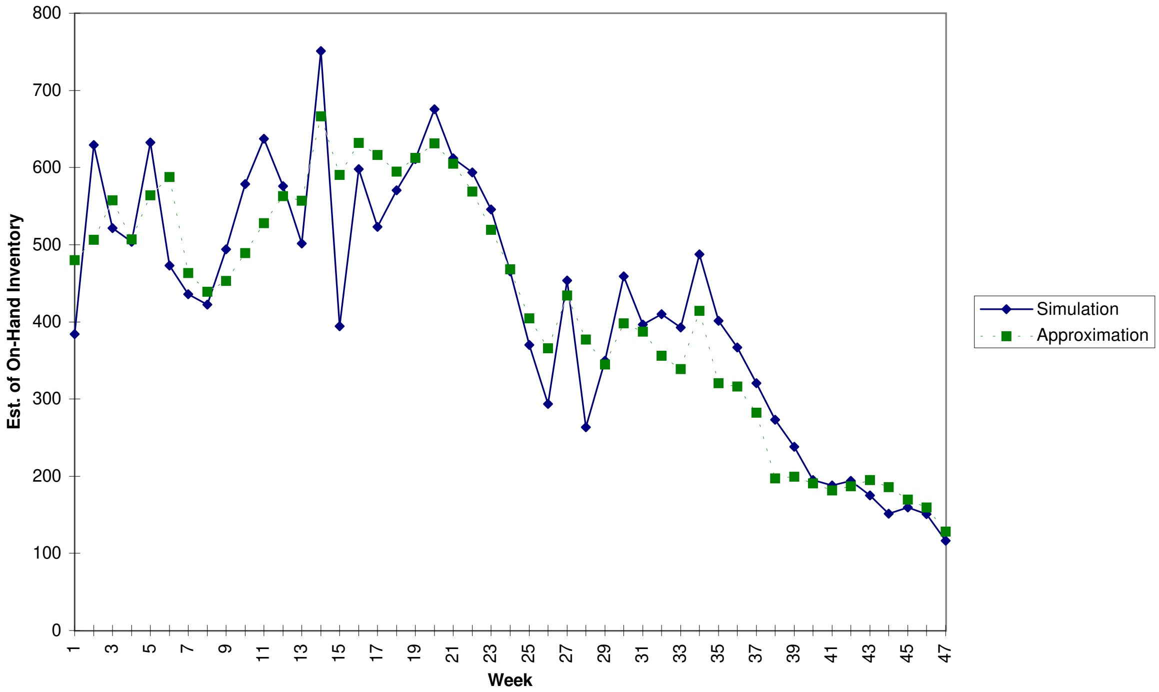

Figure 6: Estimates of Week-Ending On-Hand Inventory Levels for Product 4 with Non-Stationary Demand and Dynamic (S, s) Order Policy