Inventory control in multi-echelon divergent systems with random lead times

Eindhoven University
of Technology

Department of Mathematics and Computing Science

Memorandum COSOR 96-36

Inventory control in multi-echelon divergent systems with random lead times

M.C. van der Heijden
E.B. Diks
A.G. de Kok

Eindhoven, December 1996
The Netherlands

Eindhoven University of Technology
Department of Mathematics and Computing Science
Probability theory, statistics, operations research and systems theory
P.O. Box 513
5600 MB Eindhoven - The Netherlands

Secretariat: Main Building 9.15 or 9.10
Telephone: 040-247 4272 or 040-247 3130
E-mail: wsbscaro@win.tue.nl or wscosor@win.tue.nl
Internet: http://www.win.tue.nl/win/math/bs/cosor.html

ISSN 0926 4493

Inventory control in multi-echelon divergent systems with random lead times

Matthieu van der Heijden

University of Twente, Department of Technology and Management

P.O. Box 217, 7500 AE Enschede, The Netherlands

e-mail: m.c.vanderheijden@sms.utwente.nl

Erik Diks

Eindhoven University of Technology, Department of Mathematics and Computer Science

P.O. Box 513, 5600 MB Eindhoven, The Netherlands

e-mail: diks@win.tue.nl

Ton de Kok

Eindhoven University of Technology, Department of Technology Management

P.O. Box 513, 5600 MB Eindhoven, The Netherlands

e-mail: a.g.d.kok@tm.tue.nl

Abstract: This paper deals with integral inventory control in multi-echelon divergent systems with stochastic lead times. The policy considered is an echelon stock, periodic review, order-up-to (R, S) policy. A computational method is derived to obtain the order-up-to level and the allocation fractions required to achieve given target fill rates. Extensive numerical experimentation shows that the accuracy of our approximate method is satisfactory. Further we show that variation in the lead times may have a significant effect on the stock levels required to achieve given target fill rates, hence this variation may not be ignored. Also, it appears that the correlation structure in the lead times may have a significant impact as well.

Key words: Inventory, multi-echelon, divergent, allocation, random lead times

1. Introduction

In the last decade considerable progress has been made with respect to the analysis of multi-echelon models. In a number of recent review papers this progress has been discussed. Nahmias and Smith [1993] give an overview of inventory models in one warehouse - N retailer systems. Axsäter [1993] reviews the literature on multi-echelon models with installation stock policies, i.e. each stockpoint is controlled based on local information about its inventory position and its demand. In Federgruen [1993] an overview is given of the papers focusing on the derivation of discounted cost-optimal echelon stock control policies. Zijm and Van Houtum [1996] build on this review to discuss average cost-optimal multi-echelon models. In Diks, de Kok and Lagodimos [1996] a review is given of multi-echelon models with service level constraints. They discuss both installation stock policies and echelon stock policies.

A problem that occurs when dealing with real-world supply chains is the randomness of lead times to stockpoints. Lead times to the most downstream stockpoints are in fact transportation times and thereby usually reliable, but lead times to more upstream stockpoints in a supply chain are random due to the fact that these lead times represent manufacturing throughput times. These throughput times are random variables due to e.g. capacity constraints, machine-breakdowns, rejects and batching decisions. We have to be aware that the multi-echelon models represent a supply chain for a set of products, but which share resources together with other sets of products. This sharing of resources is the major cause for randomness of lead times. So far no analysis is available of multi-echelon models with echelon stock policies and random lead times with the exception of the paper by Zipkin [1991], who deals with continuous review (S-1,S) models. It follows from the analysis in De Kok[1990], Verrijdt and De Kok[1995] and De Kok et al [1994] that random lead times to the most downstream stockpoints can easily be incorporated in the analysis of multi-echelon models with constant lead times. However, the more practical situation of random lead times to more upstream stockpoints cannot be straightforwardly incorporated.

In this paper we consider divergent N-echelon models, where each stockpoint is controlled according to a echelon order-up-to-policy. We define the echelon inventory position as the amount of the stock in the stockpoint itself plus all downstream stocks plus the stock in transit to the stockpoint minus all backorders its downstream stockpoints. Due to the fact that we deal with random lead times we have to be more specific about the replenishment mechanism of the order-up-to-policies. The most upstream stockpoint follows a periodic review (R,S)-policy, i.e. at each review moment the echelon inventory position is raised to the order-up-to-level S. However, a more downstream stockpoint raises its echelon inventory position to the order-up-to-level upon arrival of a replenishment at the (unique) stockpoint that precedes this stockpoint. Hence the ordering decisions are taken at random points in time. The rate at which ordering decisions are taken is obviously equal to 1/R. Through this way of timing the replenishment decisions we optimally synchronize these decisions.

The lead times to the stockpoints are identically distributed random variables. Subsequent replenishment orders do not overtake. This implies that the subsequent lead times are correlated. In this paper we model this (auto)correlation assuming that the lead times constitute an AR(1) process. Simulation experiments show that this modelling assumption yields good approximations for performance measures even if the lead time process is not AR(1). The demand process is assumed to be stationary and time-homogeneous, i.e. the demand during an interval (s,s+t] is independent of demand before s and after t+s and not dependent on s itself. Since we consider divergent N-echelon systems we have to cope with situations where stock at a stockpoint is insufficient to satisfy the demand of its successors. In that case we use a linear rationing rule as defined in Van der Heijden[1996], the Balanced Stock rationing rule.

In this paper we concentrate on the determination of the control parameters, i.e. the order-up-to-levels and the parameters of the BS rationing rule, such that target fill rates at downstream stockpoints are achieved. Here the fill rate is defined as the fraction of demand satisfied from stock on hand. We assume that so-called maximum stock levels at intermediate stockpoints are given. The maximum stock level is the difference between the order-up-to-level at the stockpoint and the sum of the order-up-to-levels at its successors. By varying the maximum stock levels we can varying the average physical stocks at all intermediate stockpoints. In principle one can use standard Newton methods, cf. Stoer and Bulirsch [1993], to find the cost-optimal maximum stock levels subject to the fill rate constraints. The motivation for using a service measure concept instead of a penalty cost concept is the fact that due to the use of business information systems data are available in practice about target fill rates, whereas hardly any data are available about penalty costs. Therefore a service measure concept is more appropriate.

We develop a heuristic algorithm to compute the control parameters. Therefore we have done extensive testing of the algorithm using discrete event simulation. These numerical experiments show that the algorithm performs quite well. Thereupon we use the algorithms to investigate whether the incorporation of autocorrelation of the lead times into the analysis is important for the accuracy of the algorithm. Furthermore we investigate the managerial issue of the impact of randomness of lead times on stock investments.

This paper is organized as follows. First we give a detailed model description (section 2). In the sections 3-5 we derive the analysis of this model. We start with the most simple case in section 3, the two-echelon model with stockless central depot. The analysis is extended to two-echelon systems with central stocks in section 4 and to general N-echelon systems in section 5. A summary of the algorithm is given in section 6. Next we validate our approximate method by comparison to simulation results, both for two-echelon and for three-echelon systems (section 7). Some sensitivity analysis is shown in section 8. Finally, we give our conclusions and suggest directions for further research in section 9.

2. The mathematical model

To define the mathematical model, we first describe the network structure and the material flow through the network (section 2.1). Next, we describe the inventory control policy (section 2.2). Finally, we give the model assumptions and an overview of notation (section 2.3).

2.1. Network structure and material flow

We consider a single item, N-echelon divergent system as shown in Figure 1. That is, each stockpoint receives products from exactly one supplier. These products are either distributed further to one or more successive stockpoints or used to satisfy (stochastic) external demand. This demand takes place at the *end* stockpoints (numbers 4-12 in Figure 1). The other nodes are called *intermediate* stockpoints. The material flows through the network from the single external supplier via one or more stockpoints to the final customers. For a given stockpoint i, we define the set of *upstream* stockpoints as all stockpoints between the external supplier and stockpoint i. Then the set of *downstream* stockpoints is defined as all stockpoints between stockpoint i and the final customers. To describe the network structure, we use the following notation:

i = stockpoint index, where i=0 denotes the most upstream stockpoint,
pre(i) = the single supplier of stockpoint i, e.g. pre(8)=2 in Figure 1,
succ(i) = the set of all stockpoints immediately supplied by stockpoint i, e.g. succ(2)={7, 8, 9} in Figure 1.

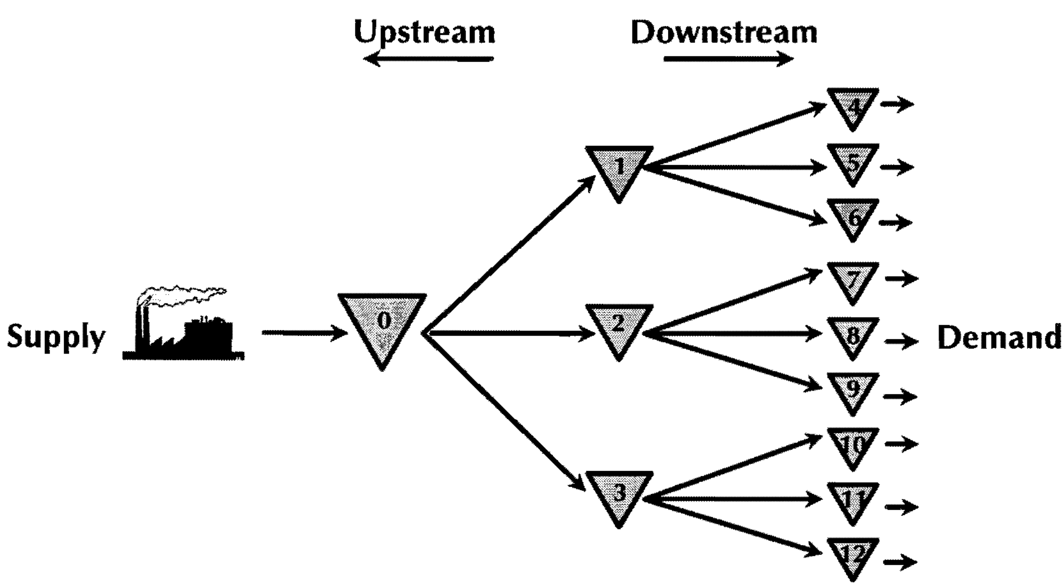

Figure 1. Divergent network structure

The lead times to move the material between two subsequent stages are random variables. Hence we have both demand- and supply uncertainty. The ultimate goal is to control the material flow such, that predetermined service levels to the final customers are attained. We will *not* optimize the stock levels at each node in the network, but we will derive a method to calculate the stock required at the end stockpoints given the maximum stock levels at intermediate stockpoints.

The service measure that we use is the *fill rate*, i.e. the fraction of demand that is satisfied immediately from stock on hand. We assume that demand that cannot be satisfied immediately is backordered. We allow that different target fill rates are used for different end stockpoints.

2.2. Inventory control mechanism

The material flow in the network is controlled by an echelon stock (R,S) policy, a periodic review, order-up-to policy. That is, the most upstream stockpoint releases a replenishment order each R time units, such that its echelon inventory position is raised to the order-up-to level S. A replenishment order arriving at an end stockpoint can be used to satisfy local customer demand. When a replenishment order arrives at an intermediate stockpoint, one should decide what to do with it. Of course, the replenishment order can be sent on immediately to downstream stockpoints to satisfy local customer demand. Then one should decide how the order should be allocated amongst the downstream stockpoints. Once this decision is taken, the downstream stocks cannot be reallocated anymore (in practice this may be possible, but at high costs). When the allocation quantities per stockpoint are calculated, negative quantities are theoretically possible. This may occur if the demand during the last period at one stockpoint has been considerably higher then expected, while the demand at an other end stockpoint has been considerably lower then expected. Such a situation is called *imbalance*. To reduce imbalance, one might keep back stocks at an intermediate stockpoint for allocation at a later point in time.

So the following decisions have to be taken after arrival of a replenishment order at an intermediate stockpoint:

- which part of the replenishment order should be kept back for allocation at a later point in time?
- how should the remaining part of the order be allocated amongst the downstream stockpoints?

To explain the control mechanism, we will first describe the situation in which the entire replenishment order is allocated. Next we will add the option of keeping back some stock for allocation later on.

Consider a replenishment order arriving at an intermediate stockpoint i. This order should be allocated to the succeeding stockpoints $j \in succ(i)$. The allocation decision is similar to the one described in Van der Heijden [1996]. This decision is based on:

- the size of the replenishment order Q,
- the echelon inventory positions of the successors j just before allocation, represented by z_j, $j \in succ(i)$,
- a set of order-up-to levels S_j^*, $j \in succ(i)$,
- a set of allocation fractions p_j, $j \in succ(i)$.

Now the allocation rule is as follows: Successor j gets an amount that raises its echelon inventory position from the level z_j before allocation to the level S_j after allocation, defined by

$$S_j = S_j^* - p_j \left\{ \sum_{k \in succ(i)} (S_k^* - z_k) - Q \right\} \tag{1}$$

Hence stockpoint j receives an amount $S_j - z_j$, which is a random variable because both the inventory positions before and after allocation, z_j and S_j, are random variables. Note that summation of both sides of (1) over j yields $\Sigma S_j = Q + \Sigma z_j$, which means that the replenishment order is allocated entirely indeed. We refer to Van der Heijden [1996] for a further explanation of this rule.

If no stock is kept back at stockpoint i, we have the relation $S_i^* = \sum_{j \in succ(i)} S_j^*$. We can extend allocation rule (1) to situations with intermediate stocks by introducing a parameter Δ_i, the maximum amount of stock to be hold in stockpoint i. As mentioned in section 2.1, we consider Δ_i as a given parameter. In fact, the relation between succeeding order-up-to levels changes to

$$S_i^* = \Delta_i + \sum_{j \in succ(i)} S_j^* \tag{2}$$

Now it may be possible that the size of the replenishment order Q exceeds the amount needed to raise all inventory positions to their maximum levels S_j^*. Negative rationing is not allowed, because then stock is not kept back at stockpoint i, so rule (1) has to be changed to

$$S_j = S_j^* - p_j * \max\left\{0, \sum_{k \in succ(i)} (S_k^* - z_k) - Q\right\} = S_j^* - p_j * \max\left\{0, S_i^* - (\Delta_i + Q + \sum_{k \in succ(i)} z_k)\right\} \tag{3}$$

This rule will be used in the sequel. The problem is to find the control parameters $\{p_j, S_j^*\}$ for each stockpoint j, given demand characteristics, lead time characteristics, target fill rates at the end stockpoints and given maximum intermediate stock levels Δ_i. For the most upstream stockpoint, say number 0, we have to find the order-up-to level S_0^* with allocation fraction p_0=1 of course.

One additional remark about the inventory control in the network has to be made. We assume that allocation decisions at intermediate stockpoints are taken immediately after an order arrives. This seems to be reasonable at first sight, because it is important to transfer products without delay to avoid shortages at the end stockpoints. It means also that the time between two subsequent shipments from intermediate stockpoints is a random variable as well (with mean R and positive variance). This may not be desirable in all cases, for example with respect to transport planning.

Further this policy is not optimal in the case of stochastic lead times. As an example, consider the arrival of two subsequent replenishment orders at an intermediate stockpoint. If the lead time of the first order is relatively short and the lead time of the second order is relatively long, the following may occur. When allocating the first order, the echelon inventory positions z_j of the successors are relatively high and a significant part of the replenishment order is kept back. If it takes a long time before the next replenishment order arrives, shortages may have occurred downstream whilst sufficient stock would have been available downstream. On the other hand, if we would have chosen for a model in which allocation decisions are made at fixed points in time as well, it may occur that the stock is allocated at the wrong moment, namely just *before* a replenishment order arrives instead of just after. This would give similar results. A solution to this problem might be a model with fixed allocation epochs, but at a higher frequency in the downstream part of the network. This is another model extension next to lead time variation, which should be subject for further research.

2.3. Assumptions and notation

The mathematical model uses the following assumptions

a) Customer demand only occurs at end stockpoints.
b) The demand per period is random and stationary in time.
c) The demand is both independent across end stockpoints and across periods in time.
d) All customer demand that can not be satisfied directly from stock on hand is backlogged.
e) Partial delivery of customer orders is allowed.
f) Replenishment orders do not cross in time. This implies that successive lead times can be correlated.
g) Lot sizing is not used, so any quantity can be ordered and delivered and any allocation rule for material rationing to the local depots is allowed (as far as the quantity allocated is nonnegative).
h) There are no capacity constraints on production, storage or transport.

Assumption a) can be made without loss of generality, since we can add an end stockpoint with lead time zero for each intermediate stockpoint facing customer demand. With respect to assumption c), remark that extension to demand that is correlated between end stockpoints is straightforward, see also Van der Heijden et al [1996]. This is however not true for demand that is correlated across periods in time. With respect to assumption f), we assume for a practical application that lead times can be measured, such that correlations between successive lead times can be estimated. In section 7 we will give numerical results on the effect of the lead time correlations.

Further we will use the following general balance assumption for the approximation of the order-up-to levels:

General balance assumption: Allocation using rule (3) yields nonnegative allocation quantities only.

This assumption is usually made and simplifies the analysis considerably (see e.g. De Kok [1990], De Kok et al. [1994] and Van der Heijden [1996]). Empirical evidence shows that even rather strong violation of this assumption has only a limited effect on the quality of the approximations, unless the variation of demand during lead time is very high (see Van Donselaar en Wijngaard [1987]).

For the mathematical analysis we use the notation as listed below.

stockpoint status

z_j = echelon inventory position of stockpoint j, just before allocation of a replenishment order that has arrived at its supplier i=pre(j).

S_j = echelon inventory position of stockpoint j, just after allocation of a replenishment order that has arrived at its supplier i=pre(j).

lead times and demand

L_j = lead time between the stockpoints pre(j) and j, a random variable with mean $E[L_j]$, variance $Var[L_j]$ and first-order autocorrelation ρ_j.

L_{jk} = k^{th} lead time between the stockpoints pre(j) and j, this is used to distinguish subsequent replenishment orders in the analysis.

R = review period.

$\overline{R}_j$ = the (stochastic) time period between the arrival of two replenishment orders at stockpoint j, the so-called *replenishment cycle.*

D_j = period demand at stockpoint j, a random variable with mean μ_j and standard deviation σ_j. If j is an intermediate stockpoint, D_j denotes all downstream demand, e.g. D_2 is the period demand at the stockpoints 7, 8 and 9.

$D_{j,Lj}$ = demand at stockpoint j during a lead time L_j.

$D_{j,Lj+\overline{Ri}}$ = demand at stockpoint j during a lead time L_j plus a replenishment cycle $\overline{R}_i$ of its predecessor.

X^+ = max{X,0}.

performance measures

β_j = target fill rate for end stockpoint j.

ψ_j = mean physical stock at stockpoint j.

control parameters

p_j = rationing fraction for stockpoint j.

S_j^* = maximum echelon inventory position of stockpoint j.

Δ_j = maximum physical stock allowed at stockpoint j, a prespecified parameter.

Note that the system order-up-to level equals S_0^*, because the external supplier is assumed to have sufficient capacity and rationing is not necessary, see assumption i).

3. Analysis of two-echelon systems with stockless central depot

In this section we will derive the mathematical expressions to calculate the system control parameters and the performance measures for the simplest model, namely a two-echelon system with stockless central depot. We consider one upstream stockpoint having index 0 and a number of end stockpoints $j \in succ(0)$. In section 4 we will add the option that the central depot is allowed to hold stock. Extension of the results to general N-echelon systems is discussed in section 5.

The mathematical model is analyzed using the approach of Van der Heijden [1996]. However, we have to account for the stochastic, correlated lead times. In section 3.1 we describe the calculation of the maximum inventory positions S_j^* assuming that the allocation fractions p_j are known. It will appear that some aggregate demand characteristics are required for the calculations, which are derived in section 3.2. In section 3.3 we derive an expression for the allocation fractions p_j, such that an approximate expression for the mean system imbalance is minimized. It will appear that we can calculate these allocation fractions based on demand and lead time characteristics only, so *independently* of fill rate requirements and order-up-to levels. This is attractive, both from an analytical and a computational point of view. Together, sections 3.1-3.3 give the control parameters $\{p_j, S_j^*\}$ of the system. Finally, we will show how to calculate the mean physical stock in section 3.4.

3.1. Maximum echelon inventory positions S_j^*

Our starting point is the simple singe location (R, S_j) system. It is well known that for this model the following equation should be solved for S_j (see Diks, de Kok and Lagodimos [1996]):

$$\frac{E[(D_{j,Lj+R} - S_j)^+] - E[(D_{j,Lj} - S_j)^+]}{R\mu_j} = 1 - \beta_j \tag{4}$$

The denominator equals the expected demand in a replenishment cycle. The numerator equals the difference between the expected shortage at the end and at the start of a replenishment cycle. In many (but not all) practical situations, the second term in the numerator can be ignored, see e.g. Hadley and Whitin [1963].

As shown in De Kok [1990] and Van der Heijden [1996], the expression to calculate S_j^* in a two-echelon system is a straightforward extension of (4), assuming that the allocation parameters p_j are known. Given the general imbalance assumption as stated in section 2.3, we find that we can obtain S_j^* by solving

$$\frac{E[\{D_{j,Lj+\overline{R0}} + p_j D_{0,L0} - S_j^*\}^+] - E[\{D_{j,Lj} + p_j D_{0,L0} - S_j^*\}^+]}{R\mu_j} = 1 - \beta_j \tag{5}$$

This expression is valid for general demand- and lead time distributions. To solve equation (5) numerically, it is convenient to use some simple two-moment approximation for the stochastic components in both terms in the numerator, $X_{1j} := D_{j,Lj+\overline{R0}} + p_j D_{0,L0}$ and $X_{2j} := D_{j,Lj} + p_j D_{0,L0}$, see Appendix for details. Then equation (5) can be solved for S_j^* using bisection. In this way the maximum echelon inventory positions S_j^* can be calculated one-by-one for each end stockpoint j, if the allocation fractions p_j are known. Next the order-up-to level S_0^* is obtained by summation of the S_j^*, since the central depot does not hold stock (see (2)).

Now the problem is the derivation of the mean and variance of X_{1j} and X_{2j}. We find that

$$E[X_{1j}] = E[D_{j,Lj+\overline{R0}}] + p_j E[D_{0,L0}] \tag{6a}$$

$$Var[X_{1j}] = Var[D_{j,Lj+\overline{R0}}] + p_j^2 Var[D_{0,L0}] + 2p_j Cov[D_{j,Lj+\overline{R0}}, D_{0,L0}] \tag{6b}$$

$$E[X_{2j}] = E[D_{j,Lj}] + p_j E[D_{0,L0}] \tag{7a}$$

$$Var[X_{2j}] = Var[D_{j,Lj}] + p_j^2 Var[D_{0,L0}] + 2p_j Cov[D_{j,Lj}, D_{0,L0}] \tag{7b}$$

We see that we need several demand characteristics to evaluate (6a)-(7b), among which two covariances. In the next section we show how these characteristics can be obtained.

3.2. Calculation of the demand characteristics

From the previous subsection we see that we need the following demand characteristics to calculate the control parameters S_j^*, given the allocation fractions p_j:

(i) the mean and variance of $D_{j,Lj}$ for all end stockpoints $j \in succ(0)$,

(ii) the mean and variance of $D_{0,L0}$,

(iii) the mean and variance of $D_{j,Lj+\overline{R0}}$ for all end stockpoints $j \in succ(0)$,

(iv) $Cov[D_{j,Lj+\overline{R0}}, D_{0,L0}]$ for all end stockpoints $j \in succ(0)$,

(v) $Cov[D_{j,Lj}, D_{0,L0}]$ for all end stockpoints $j \in succ(0)$.

We can easily show that $Cov[D_{j,Lj}, D_{0,L0}] = 0$. The demand $D_{j,Lj}$ relates to the time interval $[L_{01}, L_{01}+L_{j1}]$, so the length of this time interval is independent of L_{01}. Below we derive expressions for the other four demand characteristics.

First, we can easily derive the mean and variance of $D_{j,Lj}$ for each end stockpoint j by conditioning on L_j, cf. Silver and Peterson [1985].

$$E[D_{j,Lj}] = \mu_j E[L_j] \tag{8a}$$

$$Var[D_{j,Lj}] = \sigma_j^2 E[L_j] + \mu_j^2 Var[L_j] \tag{8b}$$

Second, the mean and variance of $D_{0,L0}$ can be obtained using the same equations, once we have the mean and variance of D_0. The latter can easily be obtained by summation, because the demand at different end stockpoints is independent, see assumption c):

$$\mu_0 = \sum_{j \in succ(0)} \mu_j \text{ and } \sigma_0^2 = \sum_{j \in succ(0)} \sigma_j^2 \tag{9}$$

Note that correlation between demand at different end stockpoints could be included easily here.

Third, the mean and variance of $D_{j,Lj+\overline{R0}}$ for all end stockpoints j should be calculated. To do this, we should carefully specify the time interval of the demand. Without loss of generality, suppose that at time 0 the central depot issues replenishment order 1 that raises the echelon inventory position to the level S_0^* . This order arrives after a random lead time L_{01} and is immediately allocated. The amount allocated to stockpoint j, together with the current physical stock and pipeline stock, should be sufficient to cover the demand until the arrival of the second replenishment order at stockpoint j, which occurs at time $R+L_{02}+L_{j2}$. Hence we are interested in the distribution of the demand at end stockpoint j in the time interval $[L_{01}, R+L_{02}+L_{j2}]$. Now we can derive by conditioning on the interval length $R+L_{j2}+L_{02}-L_{01}$ that

$$E[D_{j,Lj+\overline{R0}}] = \mu_j (R + E[L_j]) \tag{10a}$$

$$Var[D_{j,Lj+\overline{R0}}] = \sigma_j^2 (R + E[L_j]) + \mu_j^2 \{Var[L_j] + 2(1-\rho_0) Var[L_0]\} \tag{10b}$$

Fourth, we derive an expression for $Cov[D_{j,Lj+\overline{R0}}, D_{0,L0}]$, again by carefully specifying the time intervals of the demand. As stated above, $D_{j,Lj+\overline{R0}}$ is the demand at stockpoint j in the time interval $[L_{01}, R+L_{02}+L_{j2}]$. The other term is derived from the fact S_j^* should also cover a fraction p_j of the total demand during the first replenishment lead time L_{01}. Hence $D_{0,L0}$ is the total system demand in the time interval $[0, L_{01}]$. From this we see that $D_{j,Lj+\overline{R0}}$ and $D_{0,L0}$ are negatively correlated: If the first lead time L_{01} is relatively long, the total demand in $[0, L_{01}]$ will be relatively high, but also $[L_{01}, R+L_{02}+L_{j2}]$ is relatively short and hence the demand at stockpoint j during this interval will be relatively low. If we ignore this negative correlation, the variance of $X_{1j} = D_{j,Lj+\overline{R0}} + p_j D_{0,L0}$ will be overestimated, resulting in higher stock levels than actually necessary, see equation (5). We will return on this subject at our numerical analysis in section 8.

We obtain an expression for $Cov[D_{j,Lj+\overline{R0}}, D_{0,L0}]$ by conditioning on L_{01} and using that in general

$$Cov[X_1, X_2] = E[Cov(X_1, X_2 | Y)] + Cov(E[X_1 | Y], E[X_2 | Y]) \tag{11}$$

Expression (11) is derived analogously to the well-known equivalent for the variance as derived in Mood et al [1974]. We have that $Cov[D_{j,Lj+\overline{R0}}, D_{0,L0} | L_{01}] = 0$, since the demand in subsequent periods is independent (see assumption c). Therefore the first term of (11) is zero in this case. Assuming that the lead times L_{0k} follow an AR(1) process with first order autocorrelation ρ_0, we have that $E[L_{02} | L_{01}] = \rho_0 L_{01} + (1-\rho_0) E[L_0]$. Using these facts, we find a negative correlation indeed:

$$Cov[D_{j,Lj+\overline{R0}}, D_{0,L0}] = -(1-\rho_0) \mu_j \mu_0 Var[L_0] \tag{12}$$

3.3. Calculation of the allocation fractions p_j

From the results in section 3.1 we see that we can tune the S_j^* to the target service levels β_j for any arbitrary set of rationing parameters p_j, *provided that the general balance assumption is not violated.* It can be expected that the approximations are better if the general balance assumption is violated only slightly. This can be achieved choosing the allocation parameters p_j such, that the expected imbalance is minimized. We achieve this using the approach by Van der Heijden [1996], taking into account the stochastic, correlated lead times.

We define the amount of imbalance Ω_j at stockpoint j as the negative allocation quantity, so $\Omega_j = (z_j - S_j)^+$, see also De Kok [1990]. Assuming a balanced situation at the previous allocation epoch, we find that $z_j - S_j = p_j D_0[R, R + L_{02}] - p_j D_0[0, L_{01}] - D_j[L_{01}, R + L_{02}]$ where the numbers between squared brackets denote time intervals. Analogously to Van der Heijden [1996] and using the expressions for the various demand characteristics from section 3.2, we find that

$$E[z_j - S_j] = R\mu_j \tag{13a}$$

$$\mathrm{Var}[z_j - S_j] = 2p_j^2 E[M_0]\sigma_0^2 + \{R - 2p_j E[M_0]\}\sigma_j^2 + 2\mathrm{Var}[L_0](1-\rho_0)(\mu_j - p_j\mu_0)^2 \tag{13b}$$

where $M_0 = \min\{R, L_0\}$.

Now two approaches are suggested in Van der Heijden et al [1996]. Firstly, we can approximate z_j-S_j by a normal distribution and derive an explicit expression for $E[\Omega_j]$. Using a numerical method, the allocation fractions p_j can be obtained such that $E[\Omega_j]$ is minimized. Secondly, we can use $\mathrm{Var}[z_j-S_j]$ as a surrogate expression for the imbalance and minimize (13b) subject to $\Sigma p_j=1$. The latter approach was suggested by Van Donselaar [1996] for the model with deterministic lead times. Numerical tests in Van der Heijden et al [1996] revealed that the first approach is somewhat better, but the second approach is considerably simpler. Because our numerical experiments indicated that the difference between the two approaches decrease if stochastic lead times are introduced, we restrict ourselves here to the second approach. Taking into account the variance and correlation of L_0, we can derive by minimizing (13b) subject to $\Sigma p_j=1$ that

$$p_j = \frac{1}{2n} + \frac{E[M_0]\sigma_j^2 + \mu_0(1-\rho_0)\mathrm{Var}[L_0] * \left(2\mu_j - \frac{\mu_0}{n}\right)}{2E[M]\sigma_0^2 + 2\mu_0^2(1-\rho_0)\mathrm{Var}[L_0]} \tag{14}$$

where n denotes the number of end stockpoints. From (14) it can be shown that $0 \le p_j \le 1$ for all j and that $\Sigma p_j=1$. Note that $E[M_0]=E[\min\{R,L_0\}]$ should be calculated to evaluate (14). This can be done by fitting e.g. a mixture of Erlang distributions to the first two moments of L_0.

3.4. Physical stock

Just as for S_j^*, we can find expressions for the mean physical stock per local depot ψ_i that are simple extensions of the single depot (R,S) model. We have approximately that

$$\psi_j \approx \frac{\text{E[stock at start of replenishment cycle]} + \text{E[stock at end of replenishment cycle]}}{2}$$

which can be written as

$$\psi_j \approx S_j^* - p_j E[L_0]\mu_0 - (E[L_j] + \tfrac{1}{2}R)\mu_j + \frac{E[(D_{j,Lj} + p_j D_{0,L0} - S_j^*)^+] + E[(D_{j,Lj+\overline{R0}} + p_j D_{0,L0} - S_j^*)^+]}{2} \tag{15}$$

The two expectations represent the mean shortage at the start and at the end of a replenishment cycle, which are already calculated in (5).

4. Two-echelon system with central stock

We can extend the results from section 3 to the situation where the central depot is allowed to hold stock. The same subjects will be treated in the same order as in section 3.

4.1. Maximum echelon inventory positions S_j^*

Equation (5) can easily be modified such that the central depot holds central stock up to some prespecified level Δ_0, see e.g. De Kok et al [1994] and Van der Heijden [1996]:

$$\frac{E[\{D_{j,Lj+\overline{R}_0} + p_j(D_{0,L0} - \Delta_0)^+ - S_j^*\}^+] - E[\{D_{j,Lj} + p_j(D_{0,L0} - \Delta_0)^+ - S_j^*\}^+]}{R\mu_j} = 1 - \beta_j \tag{16}$$

Again, equation (16) can be solved numerically using two-moment approximations for both terms in the numerator, $X_{1j} := D_{j,Lj+\overline{R}_0} + p_j(D_{0,L0} - \Delta_0)^+$ and $X_{2j} := D_{j,Lj} + p_j(D_{0,L0} - \Delta_0)^+$. For the mean and variance of X_{1j} and X_{2j} we now find that

$$E[X_{1j}] = E[D_{j,Lj+\overline{R0}}] + p_j E[(D_{0,L0} - \Delta_0)^+] \tag{17a}$$

$$\text{Var}[X_{1j}] = \text{Var}[D_{j,Lj+\overline{R0}}] + p_j^2 \text{Var}[(D_{0,L0} - \Delta_0)^+] + 2p_j \text{Cov}[D_{j,Lj+\overline{R0}}, (D_{0,L0} - \Delta_0)^+] \tag{17b}$$

$$E[X_{2j}] = E[D_{j,Lj}] + p_j E[(D_{0,L0} - \Delta_0)^+] \tag{18a}$$

$$\text{Var}[X_{2j}] = \text{Var}[D_{j,Lj}] + p_j^2 \text{Var}[(D_{0,L0} - \Delta_0)^+] + 2p_j \text{Cov}[D_{j,Lj}, (D_{0,L0} - \Delta_0)^+] \tag{18b}$$

We see that we need several other demand characteristics to evaluate (17a)-(18b), among which two complex covariances. In the next section we show how these characteristics can be obtained.

4.2. Calculation of the demand characteristics

We need the following additional demand characteristics to calculate the control parameters S_j^*, given the allocation fractions p_j:

(i) the mean and variance of $(D_{0,L0} - \Delta_0)^+$

(ii) $\mathrm{Cov}[D_{j,Lj+\overline{R0}}, (D_{0,L0} - \Delta_0)^+]$ for all end stockpoints $j \in \mathrm{succ}(0)$

(iii) $\mathrm{Cov}[D_{j,Lj}, (D_{0,L0} - \Delta_0)^+]$ for all end stockpoints $j \in \mathrm{succ}(0)$

First, a common way to approximate the mean and variance of $(D_{0,L0} - \Delta_0)^+$ is to approximate $D_{0,L0}$ by a mixture of Erlang distributions, see Appendix.

Second, we need an expression for $\mathrm{Cov}[D_{j,Lj+\overline{R0}}, (D_{0,L0} - \Delta_0)^+]$. By considering the time intervals and using (11), we can derive that

$$\mathrm{Cov}[D_{j,Lj+\overline{R0}}, (D_{0,L0} - \Delta_0)^+] = -(1-\rho_0)\mu_j \mathrm{Cov}[L_0, g(L_0)] \tag{19}$$

where $g(L_0) = E[(D_{0,L0} - \Delta_0)^+ | L_0]$. This expression is difficult to evaluate in general. Therefore we approximate $\mathrm{Cov}[L_0,g(L_0)]$ by a first-order Taylor expansion around $E[L_0]$: $\mathrm{Cov}[L_0,g(L_0)] \approx g'(E[L_0])*\mathrm{Var}[L_0]$. In this way we obtain

$$\mathrm{Cov}[D_{j,Lj+\overline{R0}}, (D_{0,L0} - \Delta_0)^+] = -(1-\rho_0)\mu_j \mathrm{Var}[L_0] * g'(E[L_0]) \tag{20}$$

Although some approximation for g'(.) is possible, this function gets increasingly complicated if the number of echelons in the network increase, see section 5. Therefore we calculate g'(.) using a numerical derivative:

$$g'(E[L_0]) \approx \frac{g(E[L_0] + \varepsilon) - g(E[L_0])}{\varepsilon} \tag{21}$$

where ε is some small number (we took $\varepsilon = 0.001*E[L_0]$).

Third, we need an expression for $\mathrm{Cov}[D_{j,Lj}, (D_{0,L0} - \Delta_0)^+]$. Using the same reasoning as in section 3.2, we can show that this covariance is zero.

4.3. Calculation of the allocation fractions p_j

The derivation of near-optimal allocation fractions gets more complicated in the presence of central stock. Therefore we will use the allocation fractions (14) as obtained from the situation with stockless central depot. Van der Heijden et al [1996] showed that this approach yields satisfactory results.

4.4. Physical stock

For the mean physical stock in the end stockpoints, only a small modification of expression (15) is required. We find that

$$\psi_j \approx S_j^* - p_j E[(D_{0,L0} - \Delta_0)^+] - (E[L_j] + \tfrac{1}{2}R)\mu_j + \frac{E[(D_{j,Lj} + p_j(D_{0,L0} - \Delta_0)^+ - S_j^*)^+] + E[(D_{j,Lj+\overline{R0}} + p_j(D_{0,L0} - \Delta_0)^+ - S_j^*)^+]}{2} \qquad (22)$$

Again, all the terms required are already available.

However, now we have to calculate the mean physical stock in the central depot as well. Note that the central stock remains constant between the arrival of two successive replenishment orders, so at first sight we only need the mean physical stock at the central depot just after allocation, $E[(\Delta_0 - D_{0,L0})^+]$. However, the time period between two order arrivals $\overline{R}_0$ is a random variable now, which is correlated with the stock level just after allocation: If the first replenishment order lead time is relatively short, then:

- the time between two successive order arrivals is relatively long, and
- the total demand during the first lead time is relatively low, so that the central stock just after allocation is relatively high.

Hence we should account for the length of the time between order arrivals $\overline{R}_0 = R + L_{02} - L_{01}$ as well. The appropriate expression for the mean physical stock at the central depot is

$$\Psi_0 \approx \frac{E[(R + L_{02} - L_{01}) * (\Delta_0 - D_{0,L0})^+]}{R} \qquad (23)$$

Conditioning on L_{01}, we find

$$\Psi_0 \approx E[(\Delta_0 - D_{0,L0})^+] - \frac{1-\rho_0}{R} Cov[L_0, (\Delta_0 - D_{0,L0})^+] \qquad (24)$$

The latter covariance can be approximated similarly to (20)-(21).

5. General N-echelon systems

In this section we extend our results to general N-echelon models. For sake of convenience, the expressions are given for three-echelon systems. Modification to more than three echelons is straightforward.

5.1. Maximum echelon inventory positions S_j^*

We consider an end-stockpoint j with pre(j)=i and pre(i)=0. Van der Heijden et al [1996] show that then the following modification of (16) is required:

$$\frac{E\left[(X_{1j}-S_j^*)^+\right]-E\left[(X_{2j}-S_j^*)^+\right]}{R\mu_j}=1-\beta_j \tag{25}$$

where $X_{1j}=D_{j,Lj+\overline{Ri}}+p_j\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+$

and $X_{2j}=D_{j,Lj}+p_j\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+$

which can be solved numerically using two-moment approximations for X_{1j} and X_{2j}. Eliminating zero covariances, we find for the mean and variance of X_{1j} and X_{2j}:

$$E[X_{1j}] = E[D_{j,Lj+\overline{Ri}}]+p_jE\left[\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right] \tag{26a}$$

$$\mathrm{Var}[X_{1j}] = \mathrm{Var}[D_{j,Lj+\overline{Ri}}]+p_j^2\mathrm{Var}\left[\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right] + 2p_j\mathrm{Cov}\left[D_{j,Lj+\overline{Ri}},\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right] \tag{26b}$$

$$E[X_{2j}] = E[D_{j,Lj}]+p_jE\left[\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right] \tag{27a}$$

$$\mathrm{Var}[X_{2j}] = \mathrm{Var}[D_{j,Lj}]+p_j^2\mathrm{Var}\left[\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right] \tag{27b}$$

5.2. Calculation of the demand characteristics

We need the following additional demand characteristics to calculate the control parameters S_j^*, given the allocation fractions p_j:

(i) the mean and variance of $D_{j,Lj+\overline{R0}}$ for all end stockpoints $j\in \mathrm{succ}(0)$,

(ii) the mean and variance of $\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+$,

(iii) $\mathrm{Cov}\left[D_{j,Lj+\overline{Ri}},\left(p_i(D_{0,L0}-\Delta_0)^+ +D_{i,Li}-\Delta_i\right)^+\right]$ for all end stockpoints $j\in \mathrm{succ}(0)$.

First, the mean and variance of $D_{j,Lj+\overline{R0}}$ can be obtained analogously to the two-echelon system. Now we need the demand characteristics in a slightly different time interval, namely $[L_{01}+L_{i1}, R+L_{02}+L_{i2}+L_{j2}]$ instead of $[L_{01}, R+L_{02}+L_{j2}]$ as for the two-echelon system. Again, by conditioning on the interval length $R+L_{j2}+L_{i2}-L_{i1}+L_{02}-L_{01}$ that

$$E[D_{j,Lj+\overline{R0}}]=\mu_j(R+E[L_j]) \tag{28a}$$

$$\mathrm{Var}[D_{j,Lj+\overline{R0}}]=\sigma_j^2(R+E[L_j])+\mu_j^2\left\{\mathrm{Var}[L_j]+2(1-\rho_0)\mathrm{Var}[L_0]+2(1-\rho_i)\mathrm{Var}[L_i]\right\} \tag{28b}$$

Extension to general N-echelon systems is straightforward.

Second, the approximation of the mean and variance of $\left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+$ is a straightforward extension of the two-echelon system. First we approximate $D_{0,L0}$ by an Erlang mixture. Next we calculate the first two moments of $(D_{0,L0}-\Delta_0)^+$ as described in the Appendix. Using these two moments, we approximate $p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}$ by an Erlang mixture. Finally we obtain the desired mean and variance of $\left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+$ using the Appendix once more. It is clear that this approach can be extended to general N-echelon systems.

Third, we need an expression for $\mathrm{Cov}\left[D_{j,Lj+\overline{Ri}}, \left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+\right]$. By considering the time intervals and using (11), we can derive that

$$\mathrm{Cov}\left[D_{j,Lj+\overline{Ri}}, \left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+\right] = -(1-\rho_0)\mu_j \mathrm{Cov}[L_0, g_0(L_0)] - (1-\rho_i)\mu_j \mathrm{Cov}[L_i, g_i(L_i)] \tag{29}$$

where $g_0(L_0) = E\left[\left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+ \mid L_0\right]$

and $g_i(L_0) = E\left[\left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+ \mid L_i\right]$.

Similarly to (19), this expression can be approximated using first-order Taylor expansions. Also these calculations can straightforwardly be extended to general N-echelon systems.

5.3. Calculation of the allocation fractions p_j

Similar to Van der Heijden et al [1996], the allocation fractions are calculated according to (14), ignoring the effect of lead times in earlier stages. For the intermediate stockpoints, we also use (14) making the appropriate substitutions for μ_j, σ_j^2, μ_0 and σ_0^2 .

5.4. Physical stock

Similar to the preceding sections, we can make the following extension for the mean physical stock in the end stockpoints:

$$\psi_j \approx S_j^* - p_j E\left[\left(p_i(D_{0,L0}-\Delta_0)^+ + D_{i,Li}-\Delta_i\right)^+\right] - (E[L_j] + \tfrac{1}{2}R)\mu_j + \frac{E\left[(X_{1j}-S_j^*)^+\right] + E\left[(X_{1j}-S_j^*)^+\right]}{2} \tag{30}$$

where $E\left[(X_{1j}-S_j^*)^+\right]$ and $E\left[(X_{2j}-S_j^*)^+\right]$ denote the expected shortage at the end and at the start of a replenishment cycle respectively, see equation (25).

For the mean physical stock in the intermediate stockpoints we find analogously to section 4.4:

$$\Psi_i \approx E\left[\left(\Delta_i - D_{i,Li} - p_i(D_{0,L0} - \Delta_0)^+\right)^+\right] - \frac{1-\rho_0}{R}\mathrm{Cov}\left[L_0, \left(\Delta_i - D_{i,Li} - p_i(D_{0,L0} - \Delta_0)^+\right)^+\right]$$
$$- \frac{1-\rho_i}{R}\mathrm{Cov}\left[L_i, \left(\Delta_i - D_{i,Li} - p_i(D_{0,L0} - \Delta_0)^+\right)^+\right] \tag{31}$$

The latter covariances can be approximated similarly to (29).

Finally note that expression (24) for the stock in the most upstream stockpoint Ψ_0 remains valid.

6. Summary of the algorithm

The purpose of our algorithm is to calculate the allocation parameters p_j and maximum echelon inventory positions S_j^* for all stockpoints in the network, such that (different) target fill rates β_j for the end stockpoints are attained using minimal inventory imbalance. Then ordering decisions can take place using the order-up-to level S_0^* and allocation decisions can be made using both p_j and S_j^* according to (3).

We need the following information as a starting point for our algorithm:

- the length of the review period R
- the mean and standard deviation of the demand at each end stockpoint j: μ_j and σ_j^2
- the target fill rates β_j for all end stockpoints j.
- the mean, variance and first-order autocorrelation of all lead times $E[L_j]$, $Var[L_i]$ and ρ_j.
- the maximum inventory levels Δ_i for each intermediate stockpoint i.

Our algorithm consists of the following steps:

1. Calculate the following demand characteristics, working from the end stockpoints in the upstream direction:
   a) the mean and standard deviation of D_i for all intermediate stockpoints i using (9)
   b) the mean and variance of $D_{j,Lj}$ for all intermediate and end stockpoints j using (8a-b)
   c) the mean and variance of $D_{j,Lj+\overline{R0}}$ for all end stockpoints using (28a-b)
   d) the covariances required for the evaluation of the fill rate expressions (25), like $\mathrm{Cov}\left[D_{j,Lj+\overline{Ri}}, \left(p_i(D_{0,L0} - \Delta_0)^+ + D_{i,Li} - \Delta_i\right)^+\right]$ for a three-echelon system, where i=prec(j) and 0=prec(i). These can be obtained from (29) using a Taylor expansion and numerical differentiation. In the case of stockless depots, the simple expression (12) with extension to multiple echelons can be used.

e) similarly, the covariances required to evaluate intermediate stock levels, like $\text{Cov}\left[L_0,\left(\Delta_i - D_{i,Li} - p_i(D_{0,L0} - \Delta_0)^+\right)^+\right]$, $\text{Cov}\left[L_i,\left(\Delta_i - D_{i,Li} - p_i(D_{0,L0} - \Delta_0)^+\right)^+\right]$ and $\text{Cov}\left[L_0,(\Delta_0 - D_{0,L0})^+\right]$ for three-echelon systems, see (24) and (31)

2. Determine the allocation fractions p_j for all stockpoints i from equation (14), where stockpoint 0 is replaced by stockpoint i=prec(j).
3. Determine the maximum echelon inventory positions S_j^* for each end stockpoint j:

a) Approximate the distributions of X_{1j} and X_{2j} as given by (25) using (successive) approximation by Erlang mixtures as described in section 5.1. Again, i=prec(j) and 0=prec(i). The mean and variance of X_{1j} and X_{2j} are given by (26a)-(27b). The two-moment fitting is described in the Appendix.

b) Use bisection to find from (25) the maximum echelon inventory positions S_j^* that match the target fill rates β_j.

4. Determine maximum echelon inventory positions for the intermediate stockpoints using (2), working from end stockpoints in the upstream direction, until we have found the order-up-to level S_0^*.
5. Determine the mean physical stock per end stockpoint from equation (30) using the expected shortage at the start and at the end of a replenishment cycle as found in step 3.
6. Finally determine the mean physical stock per intermediate stockpoint from equations (24) and (31)

7. Numerical validation

The algorithm that we developed is approximate, therefore we establish the accuracy of our approximations using extensive numerical experimentation, both for two-echelon systems (experimental design in section 7.1, results in section 7.2) and for three echelon systems (experimental design in section 7.3, results in section 7.4). We use the difference between target fill rate and actual (simulated) fill rate as a measure of accuracy. Next to the mean absolute deviation from the target fill rate, we also consider the maximum deviation as a measure of robustness.

7.1. Experimental design for two-echelon models

In the first experiment we test two-echelon systems, in which one central depot supplies products to two so-called *service groups*. For sake of convenience we refer to these groups as service group A and B respectively. A service group consists of a number of local stockpoints with the same service, demand and lead time characteristics. Both service groups consist of three end stockpoints, so we have six end stockpoints in total. To normalise time and quantities, we made the following choices for all test runs:

- the review period equals R=1
- the mean demand per time unit for each local stockpoint in service group A equals $E[D_A]=10$

A special point of attention is the stochastic process by which successive lead times are generated. Assumption f) in section 2.3 stated that orders may not cross. This is reasonable in practice, but it complicates simulation. If we simply generate independent lead times, we can not prevent that orders cross in general. Therefore we choose the following mechanism. A replenishment order first enters a single server queue. If the order leaves the queue, it passes a pipeline with deterministic sojourn time. In this way, orders can not cross, while also a large range of lead time distributions can be modelled, see Diks and Van der Heijden [1996]. What remains is the choice how to divide the lead time in the stochastic queuing sojourn time L_S and the deterministic pipeline sojourn time L_D. We found that the division $L_S = c_L E[L]$ is a good choice, where c_L denotes the coefficient of variation of L (=ratio of standard deviation and mean). This division is only valid for $0 \le c_L \le 1$, since $L_D = E[L] - L_S < 0$ otherwise. However, lead times with $c_L > 1$ are exceptional.

Now we proceed as follows. First, we split the lead time in a deterministic and a stochastic part according to $L_S = c_L E[L]$. Next, we determine the parameters of the queuing model such, that the sojourn time of a customer in the single server queuing system plus L_D has mean E[L] and variance Var[L]. This is done using the method as described in Diks and Van der Heijden [1996]. We simulate this lead time process to estimate the first-order autocorrelation ρ. This value is used in our approximate method to determine the control parameters p_j and S_j^*. Finally we simulate the two-echelon system with the same stochastic lead time processes and the control parameters p_j and S_j^* to estimate the approximation accuracy. Note that this method is only used for validation purposes. Once we know that our method is accurate, we only have to measure the mean, variance and first-order autocorrelation of all lead times from actual data.

Let us return to the choice of the parameter values. Unfortunately, the number of parameters that can be varied in our experiment is still quite large. To keep the size of the experiment within reasonable limits, we take the following parameters fixed:

- The mean lead time to the central depot equals $E[L_0]=3$. Reason for this is that most upstream (production) lead times are usually larger than a review period. Besides, the experiments for deterministic lead times in Van der Heijden et al [1996] revealed that the accuracy of the approximation was better for $L_0=1$ than for $L_0=3$, so the latter should be tested.
- The downstream lead times are usually small, because these lead times represent usually order picking, handling and transport times. Therefore we take $E[L_j]=1$ in all test runs.

Eight other parameters are varied in our experiment. We choose two different values for each parameter (see Table 5.1), except for the variation in the upstream lead time L_0. Since the variation of L0 is largest in practice, we should carefully examine the effect of this variance. To choose the maximum amount of central stock Δ_0, we proceed as follows. Equation (24) shows that the amount of central stock heavily depends on

$(\Delta_0 - D_{0,L0})^+$. Therefore it is convenient to express Δ_0 in the mean system demand during the lead time L_0, say $\Delta_0 = a_0 * E[D_{0,L0}]$ for some constant c. We have significant central stock if $a_0>1$, so we choose $a_0=1.2$. Also, we consider the situation with stockless central depot, $a_0=0$. Using the value of the constant a_0, we determine the appropriate value of Δ_0 for each case.

We tested all possible parameter combinations, yielding $3*2^7=384$ cases. The performance of the algorithm is tested by an extensive simulation of 75,000 time periods for each case to ensure high simulation accuracy.

para-meter	**description**	**values in test runs**
$E[D_B]$	the mean demand per period at an end stockpoint in service group B	10, 30
$c[D_A]$	coefficient of variation of demand per period at an end stockpoint in service group A	0.4, 0.8
$c[D_B]$	coefficient of variation of demand per period at an end stockpoint in service group B	0.4, 0.8
β_A	target fill rate at an end stockpoint in service group A (%)	90, 99
β_B	target fill rate at an end stockpoint in service group B (%)	90, 99
$c[L_0]$	coefficient of variation of L_0	0, 0.25, 0.5
$c[L_j]$	coefficient of variation of L_j	0, 0.5
a_0	constant, describing the level of stock at the central warehouse $\Delta_0 = a_0 * E[D_{0,L0}]$	0, 1.2

Table 1. Parameter values in the experiment with two-echelon systems.

7.2. Results for two-echelon models

The accuracy of our algorithm is shown in the tables 2-4 below. This accuracy is expressed as mean and maximum absolute deviation from target in percent points. Because a deviation from the target service level has usually more serious consequences in the case of a high target service level, we separately give the rationing policy performance for each fill rate level in Table 2. Further we show the performance for stockless and stock holding central depot separately (Table 3). Reason for this is that we may expect a better performance in the case of a stockless central depot, since we do not need a Taylor expansion then, cf. equations (12) and (20). Finally we show the accuracy of the approximation for the system stock in Table 4, expressed as the mean and maximum relative deviation from simulation results. Here the system stock is defined as the sum of the mean physical stock in the central depot and the end stockpoints plus the pipeline stock between central depot and end stockpoints. The pipeline stock from the external supplier to the central depot is *not* included, because this is usual for external account.

Target fill rate (β)	**Mean absolute deviation**	**Maximum absolute deviation**
90	0.52	1.80
99	0.41	1.06
ALL	0.45	1.80

Table 2. Fill rate accuracy per target fill rate

Central stock level	**Mean absolute deviation**	**Maximum absolute deviation**
$a_0 = 0$	0.39	1.06
$a_0 = 1.2$	0.54	1.80
ALL	0.45	1.80

Table 3. Fill rate accuracy per central stock level

Central stock level	Mean absolute deviation	Maximum absolute deviation
$a_0 = 0$	0.4%	1.6%
$a_0 = 1.2$	2.2%	5.0%
ALL	1.3%	5.0%

Table 4. Stock accuracy per central stock level

The overall results show that the accuracy is sufficient for practical applications, although the accuracy for the stock levels is clearly better if $a_0 = 0$. Of course, zero central stock is easy to 'approximate'. For more detailed numerical analysis we refer to section 8.

7.3. Experimental design for three-echelon models

To design an experiment for three-echelon models, we proceed from the design that is developed in Van der Heijden et al [1996] for the situation with deterministic lead times. We consider a system consisting of one central depot, supplying 4 intermediate stockpoints. Each intermediate stockpoint supplies 6 end stockpoints, so the system consists of 24 end stockpoints totally. This is the largest system considered in Van der Heijden et al [1996], showing the highest approximation errors in the case of deterministic lead times. Because the accuracy is usually better for smaller systems, we chose this large three-echelon system with 1+4+24=29 stockpoints.

To keep the number of test runs within reasonable limits, we take the following parameters fixed:

- the review period equals R=1
- the lead time to the central depot 0 has mean $E[L_0]=3$ and coefficient of variation $c[L_0]=0.5$
- the lead time to each intermediate stockpoint i from its supplier pre(i)=0 has mean $E[L_i]=1$
- the lead time to each end stockpoint j from its supplier pre(j) has mean $E[L_j]=1$

For the other parameters, we selected the values as shown in Table 5.

parameter	description	values in test runs
E[D]	the mean demand per period at an end stockpoint	10, 30
c[D]	coefficient of variation of demand per period at an end stockpoint	0.4, 0.8
β	target fill rate at an end stockpoint (%)	90, 99
$c[L_i]$	coefficient of variation of the lead time L_i to each intermediate stockpoint i	0, 0.5
$c[L_j]$	coefficient of variation of the lead time L_j to each end stockpoint j	0, 0.5
a_0	constant, describing the level of stock at the central warehouse $\Delta_0 = a_0 * E[D_{0,L0}]$	0, 1.2
a_i	constant, describing the level of stock at each intermediate stockpoint $\Delta_i = a_i * E[D_{i,Li} + p_i(D_{0,L0} - \Delta_0)^+]$	0, 1.2

Table 5. Parameter values in the experiment with two-echelon systems.

For the demand and service characteristics, we used the experimental design as described in Van der Heijden et al [1996], where it is shown that 87 parameter sets are sufficient to cover this part of the design. We combined these 87 sets with all possible combinations of $c[L_i]$, $c[L_j]$, a_0 and a_i, resulting in 2^4*87=1392 test runs. The accuracy of our approximation is tested by a simulation of 25,000 time periods for each case.

7.4. Results for three-echelon models

Similar to section 7.2, we present results on the accuracy of our algorithm for the three-echelon systems in the tables 6-8 below. We give the fill rate accuracy both per target fill rate (Table 6) and per intermediate stock level (Table 7). The stock accuracy per intermediate stock level is shown in Table 8.

Target fill rate (β)	Mean absolute deviation	Maximum absolute deviation
90	1.14	2.48
99	0.75	1.60
ALL	0.95	2.48

Table 6. Stock accuracy per target fill rate

Central and intermediate stock level	Mean absolute deviation	Maximum absolute deviation
$(a_0, a_i) = (0, 0)$	0.78	1.69
$(a_0, a_i) = (0, 1.2)$	0.96	2.48
$(a_0, a_i) = (1.2, 0)$	1.01	2.01
$(a_0, a_i) = (1.2, 1.2)$	1.04	2.40
ALL	0.95	2.48

Table 7. Fill rate accuracy per intermediate stock level

Central and intermediate stock level	Mean absolute deviation	Maximum absolute deviation
$(a_0, a_i) = (0, 0)$	2.0%	2.9%
$(a_0, a_i) = (0, 1.2)$	4.6%	6.3%
$(a_0, a_i) = (1.2, 0)$	3.8%	5.5%
$(a_0, a_i) = (1.2, 1.2)$	4.3%	5.6%
ALL	3.7%	6.3%

Table 8. Stock accuracy per intermediate stock level

Although the results are worse, we think that our method is still accurate enough for practical applications. This is especially true for systems with stockless intermediate stockpoints, such as the hierarchical planning procedure as described by De Kok [1990].

7.5. Computational effort

It appears that the approximate method is fast. CPU time using a Pentium 100Mhz PC equals 0.13 seconds per case on average for the two-echelon systems and 0.30 seconds on average for the three-echelon systems. Little computational effort is important, because often control rules have to be established for hundreds or thousands of products.

8. Sensitivity analysis

Now that a tool is available, it can be used for some sensitivity analysis to get some insight in the effects of lead time variation and autocorrelation. We focus on the following two questions:

1. Which effect have lead time variation on the amount of stock required to obtain prespecified target service levels?
2. Is it really important to include correlations in the approximate method? Life becomes considerable easier if we ignore it, so can't we keep it simple?

We will answer these questions by more detailed analysis of the results for two-echelon models only to keep the results clear.

To start with the first question, we take a look at Table 9 in which the average amount of system stock is shown for various combinations of lead time variation. To make the various cases comparable, we expressed the average stock in weeks of total customer demand.

	c[L$_i$]=0	**c[L$_i$]=0.5**	**ALL**
c[L$_0$]=0	3.4	3.7	**3.5**
c[L$_0$]=0.25	4.1	4.4	**4.3**
c[L$_0$]=0.5	5.4	5.6	**5.5**
ALL	**4.3**	**4.6**	**4.4**

Table 9. Average number of weeks stock, depending on lead time variation

We see that especially the upstream lead time variation has a significant impact on the amount of stock required to reach the target fill rates. The effect of the lead time variation between central depot en intermediate stockpoints is less, which can partly be explained by the fact that $E[L_0]=3$ and $E[L_i]=1$, so $Var[L_0]$ is larger than $Var[L_i]$. Anyway, we see that we can not ignore the lead time variation, because in practice significant lead time variation is present in the upstream part of the network.

Now the question raises what has more impact, demand variation or lead time variation? Usually the attention is focused on demand variation, but is this justified? This question can be answered by looking at Table 10. Here we consider the most important source of lead time variation c[L$_0$] and the various combinations of demand variation in the two service groups.

c[D]→ c[L$_0$]↓	**0.4, 0.4**	**0.4, 0.8**	**0.8, 0.8**	**ALL**
0	2.8	3.5	4.3	**3.5**
0.25	3.6	4.3	4.9	**4.3**
0.5	4.9	5.5	6.1	**5.5**
ALL	**3.8**	**4.4**	**5.1**	**4.4**

Table 10. Effect of lead time variation and demand variation on the average stock in weeks.

Table 10 shows that in our experiment the effect of lead time variation even dominates the effect of demand variation on the stock levels required to achieve the target fill rates! Although this can be parameter dependent, it is clear that both supply and demand variation should be taken into account when determining safety stock levels in multi-echelon divergent systems. This conclusion is in line with the results obtained by Gross and Soriano [1969] for a single stockpoint.

The second question is important to judge whether our method can be simplified considerably by excluding the correlations, such as given by the equations (12), (20) and (29). We recalculated the control parameters $\{p_j, S_j^*\}$ for all 384 cases, ignoring all these correlations. Also, we simulated the 384 cases with the modified control parameters. In Table 11 below we compare the accuracy of the simplified method to the original results. As can be seen from the formulas, this difference is only relevant if $c[L_0]>0$.

performance measure	c[L0]	original	ignore correlations
mean absolute deviation	**0.25**	0.45	1.39
from target fill rate	**0.5**	0.67	1.54
maximum absolute deviation	**0.25**	1.41	4.03
from target fill rate	**0.5**	1.80	3.45

Table 11. The effect of neglecting correlations:

We see that the performance of our method is considerably better if we take into account the correlation effects. The price paid is that of considerable additional complexity.

9. Conclusions

In this paper we developed an algorithm to analyze multi-echelon divergent networks with integral (R, S) inventory control under both stochastic demand and lead times. Validation of our method by extensive comparison to simulation results, both for two-level and for three-level systems, shows that our algorithm is sufficiently accurate for practical applications. This is in particular true for situations where the intermediate depots do not carry stocks and are only used as allocation points. An example is the hierarchical planning procedure as described by De Kok [1990]. Using our method, it is easy to include stochastic production lead times in this procedure, yielding accurate results.

Our method can be simplified by ignoring the (complicating) correlations involved, but at considerable loss of accuracy, as is shown by numerical experimentation. Further we showed the importance of including lead time variation in the model. Although frequently the attention is focused on demand variation, we showed that the effect of lead time variation may be larger than the effect of demand variation on the stock levels required to obtain prespecified target fill rates.

One of the differences between our model and practical situations is the fact that the replenishment frequency is usually not the same throughout the logistic chain in practice. For example, production orders are released monthly, while transport from warehouses to local stockpoints may be carried out weekly or multiple times a week. Therefore, subsequent research will be focused on the inclusion of differentiated replenishment frequencies in our model.

References

1. Axsäter, S., Continuous review policies for multi-level inventory systems with stochastic demand, in: Graves, S.C, A.H.G. Rinnooy Kan and P.H. Zipkin (eds.), *Logistics of production and inventory*, Handbooks in Operations research and management science 4, Elsevier science publishers, Amsterdam, 1993

2. Diks, E.B., A.G. de Kok and A.G. Lagodimos, "Multi-echelon systems: a service measure perspective", *European Journal of Operational Research* 95 no. 2, 241-263, 1996.

3. Diks, E.B. and M.C. van der Heijden, "Modelling stochastic lead times in multi-echelon systems", *Memorandum COSOR* 96-33, Eindhoven University of Technology, 1996 (submitted for publication in *Probability in the Engineering and Informational Sciences*).

4. Donselaar, K. van, and J. Wijngaard, "Commonality and safety stocks", *Engineering Costs and Production Economics* 12, 197-204, 1987.

5. Donselaar, K. van, *Personal communication*, 1996.

6. Federgruen, A., "Centralized planning models for multi-echelon inventory systems under uncertainty", in: Graves, S.C, A.H.G. Rinnooy Kan and P.H. Zipkin (eds.), *Logistics of production and inventory*, Handbooks in Operations research and management science 4, Elsevier science publishers, Amsterdam, 1993.

7. Gross, D. and A. Soriano, "The effect of reducing leadtime on inventory levels - simulation analysis", *Management Science* 16, 61-76, 1969.

8. Hadley, G. and T.M. Whitin, *Analysis of inventory systems*, Prentice-Hall, Englewood Cliffs, NJ, 1963.

9. Heijden, M.C. van der, "Supply rationing in multi-echelon distribution systems", to appear in the *European Journal of Operational Research.*

10. Heijden, M.C. van der, E.B. Diks and A.G. de Kok, "Allocation policies in general multi-echelon distribution systems with (R, S) order-up-to policies", to appear in the *International Journal of Production Economics.*

11. Houtum, G.J. van, K. Inderfurth and W.H.M. Zijm, "Materials coordination in stochastic multi-echelon systems", *European Journal of Operational Research* 95 no. 1, 1-23, 1996.

12. Kok, A.G. de, "Hierarchical production planning for consumer goods", *European Journal of Operational Research* 45, 55-69, 1990.

13. Kok, A.G. de, A.G. Lagodimos and H.P. Seidel, "Stock allocation in a two-echelon distribution network under service constraints", Working paper Eindhoven University of Technology TUE/BDK/LBS/94-03, 1994.

14. Mood, A.M., F.A. Graybill and D.C. Boes, *Introduction to the theory of statistics*, 3rd edition, McGraw-Hill, Tokyo, 1974.

15. Nahmias, S. and A.A. Smith, "Mathematical models of retailer inventory systems", in: *Perspectives in operations management*, ed. R.K. Sarin, Kluwer, Dordrecht, 1993.

16. Silver, E.A. and R. Peterson, *Decision Systems for Inventory Management and Production Planning*, Wiley, New York, 1985.

17. Stoer, J. and R. Bulirsch, *Introduction to numerical analysis*, 2nd edition, Springer Verlag, Berlin, 1993.

18. Tijms, H.C., *Stochastic models: an algorithmic approach*, Wiley, Chichester, 1994.

19. Verrijdt, J.H.C.M., and A.G. de Kok, "Distribution planning for a divergent N-echelon network without intermediate stocks under service restrictions", *International Journal of Production Economics 38*, 225-243, 1995.

20. Zipkin, P., "Evaluation of base-stock policies in multiechelon inventory systems with compound-Poisson demands", *Naval Research Logistics* 38, 397-412, 1991.

Appendix. Approximation of $E[(X\text{-}S)^+]$ and $Var[(X\text{-}S)^+]$

In this appendix we give a method to calculate the first two moments of $(X\text{-}S)^+$, where X is a random variable with mean E[X] and variance Var[X] and S is a nonnegative constant. We approximate the density of X by a mixture of Erlang densities with the same scale parameter, see e.g. Tijms [1994]:

$$h_{r,q,\lambda}(x) = q\frac{\lambda^{r-1}x^{r-2}e^{-\lambda x}}{(r-2)!} + (1-q)\frac{\lambda^{r}x^{r-1}e^{-\lambda x}}{(r-1)!} \tag{A1}$$

with the corresponding probability distribution function $H_{r,q,\lambda}(x)$, defined by

$$1 - H_{r,q,\lambda}(x) = \sum_{i=0}^{r-2}\frac{(\lambda x)^{i}e^{-\lambda x}}{i!} + (1-q)\frac{(\lambda x)^{r-1}e^{-\lambda x}}{(r-1)!} \tag{A2}$$

Defining the squared coefficient of variation of a random variable X as $c^2[X]=Var[X]/E^2[X]$, we can determine the parameters of the Erlang mixture using a two-moment fit as follows (see e.g. Tijms [1994]):

$$q = \frac{r\,c^2[X] - \sqrt{r(1+c^2[X]) - r^2c^2[X]}}{1+c^2[X]} \tag{A3}$$

$$\lambda = \frac{r-q}{E[X]} \tag{A4}$$

where r is chosen such, that $r^{-1} \le c^2[X] < (r-1)^{-1}$. After some algebra, we obtain from (A3)-(A4) that

$$E[(X-S)^+] = \frac{r}{\lambda}\,[1 - H_{r+1,q,\lambda}(S)] - S[1 - H_{r,q,\lambda}(S)] \tag{A5}$$

Similarly, we can derive for the second moment

$$E[\{(X-S)^+\}^2] = \frac{r(r+1)}{\lambda^2}\,[1 - H_{r+2,q,\lambda}(S)] + \frac{2Sr}{\lambda}\,[1 - H_{r+1,q,\lambda}(S)] - S^2[1 - H_{r,q,\lambda}(S)] \tag{A6}$$

Of course, we have the variance of $(X-S)^+$ now that we have the first two moments.