Efficient heuristics for two-echelon spare parts inventory systems with an aggregate mean waiting time constraint per local warehouse

Efficient Heuristics for Two-Echelon Spare Parts Inventory Systems with An Aggregate Mean Waiting Time Constraint Per Local Warehouse

Hartanto Wong[a], Bram Kranenburg[b], Geert-Jan van Houtum[b*], Dirk Cattrysse[a]

[a] *Katholieke Universiteit Leuven, Centre for Industrial Management, Celestijnenlaan 300A, 3001 Heverlee, Belgium*

[b] *Technische Universiteit Eindhoven , Department of Technology Management, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Abstract: This paper presents solution procedures for determining close-to-optimal stocking policies in a multi-item two-echelon spare parts inventory system. The system we consider consists of a central warehouse and a number of local warehouses, and there is a target for the aggregate mean waiting time per local warehouse. We develop four different heuristics and derive a lower bound on the optimal total cost. The effectiveness of each heuristic is assessed by measuring the relative gap between the heuristic's total cost and the lower bound. The results of the computational experiments show that a greedy procedure performs most satisfactorily. It is accurate as indicated by relatively small gaps, easy to implement, and furthermore, the computational requirements are limited. The computational efficiency can be increased by using Graves' approximate evaluation method instead of an exact evaluation method, while the results remain accurate.

Keywords: inventory; spare parts; multi-item; multi-echelon; target aggregate mean waiting times, heuristics

1. Introduction

We consider a multi-item, two-echelon inventory system for spare parts that consists of a central warehouse and a number of local warehouses. Such a setting is commonly used by many manufacturers of "high-tech" products (airplanes, complex machinery, medical equipment, mainframes, etc.) to support their after-sales services. In this paper, we specifically consider environments in which the *aggregate mean waiting time* is used as the measure of service quality. This measure is defined as the average time required to fulfill an arbitrary request at a local warehouse, and is a weighted average over average delays for individual items. This measure is directly related to *availability*. The expected time that a technical system is not available per year is equal to the average number of failures per year multiplied by the aggregate mean waiting time. The problem to be dealt with by the manufacturer is determining optimal stocking levels of spare parts at the central and local

* *Correspondence to:* G. J. van Houtum (g.j.v.houtum@tm.tue.nl)

warehouses so that target aggregate mean waiting times are met against minimal system-wide inventory holding costs.

An appropriate model to study the spare parts problem sketched above is the METRIC (Multi-Echelon Technique for Recoverable Item Control) model developed by Sherbrooke [12]. METRIC is widely adopted as the basis for many other models dealing with multi-echelon spare parts inventory systems. This model considers multi-item inventory systems with which repair and distribution processes are controlled in a two-echelon system, consisting of a central warehouse and a number of local warehouses. It assumes (S-1,S) continuous review policies and ample repair capacity for each local warehouse, and provides an approximate distribution for inventory on-hand and backorders at each warehouse. We can in fact distinguish two types of problems when analyzing the METRIC model: *evaluation* and *optimization* problems. A lot of work has been done to develop evaluation methods. Muckstadt [9] developed MOD-METRIC, an extension to METRIC, which allows for multiple levels of indenture, i.e., spares requirements for an end item and its components. Another variant of METRIC is VARI-METRIC developed by Slay [14]. In METRIC it is assumed that the mean number of items in repair equals the variance of the number of items in repair. Slay developed an approximation to obtain an expression for the variance. Next he fitted a negative binomial distribution on the first two moments. Graves [5] independently explored the negative binomial distribution for the approximation method while he also developed an exact method. He reported the improvement of the negative binomial approximation over the METRIC approximation. Graves' exact and approximate method has been extended to multi-echelon, multi-indenture systems by Rustenburg et al. [11]. Muckstadt and Thomas [10] extended the METRIC model by allowing emergency replenishments in case of stock-out situations at the central or local warehouses. They also show that managing multi-echelon systems by using adaptations of single-location methods can be dramatically inferior to methods designed for taking advantage of a system’s structure. Hausman and Erkip [6] extended the work of Muckstadt and Thomas by presenting an improved single-echelon model to approximate multi-echelon model performance. Axsäter [1] developed an alternative, exact evaluation procedure, and in addition developed an exact optimization procedure for the single-item problem.

There are a few results available for multi-item optimization problems. Sherbrooke [12, 13], for example, looked at the minimization of the *total expected backorders across all local warehouses* subject to a budget constraint. He proposes the use of exchange curves of system availability versus inventory investment, rather than offering a single “optimal value”. Different from Sherbrooke, a few researchers explicitly consider service constraints in their optimization problems. Cohen et al. [4] developed a model called *Optimizer* to determine optimal inventory policies of IBM. They use fill rate as the service measure and solve the problem by decomposing the model into three stages. The decomposition starts with the lowest echelon where demand occurs and passes up to the next level. Hopp et al. [7] and Caglar et al. [3] consider a system similar to the one that we are analyzing. In both papers, heuristics are developed to minimize inventory holding costs subject to aggregate mean waiting time constraints. In Caglar et al. [3], (S-1,S) policies are assumed at both the central warehouse and the local warehouses, just like in our model. Hopp et al. [7] are somewhat more general as they assume the more general (r, Q) inventory policy for the central warehouse (together with a constraint for the total order frequency). Caglar et al. [3] show that their heuristic is more accurate than the one of Hopp et al. [7].

In this paper, we consider a multi-item, two-echelon model, controlled by (S-1,S) policies at both the central and local warehouses and we will develop heuristics for the minimization of system-wide inventory holding costs. As stated above, we explicitly consider an *aggregate mean waiting time constraint at each local warehouse*. The consideration of aggregate mean waiting time per local warehouse instead of the average over all local warehouses (such as in Sherbrooke [12, 13]) is motivated by practice. E.g., for commercial technical systems such as large-scale computers and medical equipment, spare parts are often only kept on stock in a network managed and centrally controlled by the Original Equipment Manufacturer who sets targets in terms of availability or related measures *per service region* (or per sales area). The targets are agreed with sales departments or directly with customers, and different regions may have different targets. The main difference between our work and the work of Hopp et al. [7] and Caglar et al. [3] is that we use exact evaluation (cf. Graves [5]) within our heuristics, while Hopp et al. [7] and Caglar et al. [3] use relatively inaccurate approximate evaluations. Hopp et al. develop a own approximate evaluation and the heuristic of Caglar et al. is based on the METRIC approximation of Sherbrooke [12]. The use of the METRIC approximation in an optimization method may result in a generated solution that is *not* feasible, and in many cases even far from feasible, as we shall show in Section 5.

In this paper, our focus is on developing efficient heuristics to determine close-to-optimal stocking policies. Motivated by the results presented in Wong et al. [15] that indicate a quite satisfactory performance of the greedy approach for solving similar problems allowing lateral transshipments, we are interested to see how this approach performs in solving the problem analyzed in this paper. This approach is quite easy to implement and hence, attractive from a practical point of view. In addition, we also present a local search method that can be used to improve the solution obtained by the greedy approach. To obtain lower bounds on the optimal total costs, we develop a procedure based on a decomposition and column generation approach similar to the Dantzig-Wolfe decomposition. By the same method, we also obtain an alternative heuristic, which also may me combined with local search. Our *main contribution* is that we show that the greedy procedure (without local search) performs very satisfactorily. In addition, we show that it is safe to use Graves' approximate evaluation within the greedy procedure. That results in a heuristic method that is accurate and sufficiently simple and computationally efficient for implementation in practice.

The rest of our paper is organized as follows. In Section 2, we present the model formulation. We introduce the basic assumptions and notation used in the model, and present the mathematical formulation of our problem. In Section 3, we describe all the basic techniques used in the development of all heuristic and lower bounding procedures. Section 4 presents our computational experiments for the evaluation of heuristics. Further study investigating the accuracy of the approximate evaluation methods is presented in Section 5. Finally, some concluding remarks are presented in Section 6.

2. The model

2.1 Model description

We have a non-empty set N_{loc} of local warehouses. These local warehouses are numbered $n = 1, 2, \ldots, |N_{loc}|$. Each local warehouse serves a number of technical systems of the same or at least similar type. A technical system consists of several critical components, each of which may fail incidentally, and a failure of a component implies that the full system (or at least a significant part) fails. The components are at such levels in the material breakdown structure of the machine that they can be replaced as a whole by spare parts. They are also called assemblies, and we also refer to them as stock-keeping units (SKU-s). Let I denote the non-empty set of all SKU-s that may occur in the configurations of the technical systems, and the SKU-s are numbered $i = 1, 2, \ldots, |I|$. We assume that the total stream of failures of SKU i as observed by local warehouse n constitutes a Poisson process with a constant rate l_{in} (≥ 0). This assumption is standard in METRIC type models (and a key factor for obtaining a tractable model). For many real-life systems, lifetimes of components are (close-to) exponential, or lifetimes are not precisely exponential but the total stream of failures is a composition of sub-processes coming from relatively many technical systems that are supported by a local warehouse. In those cases it is reasonable to assume a Poisson failure process. Further, in practice, one does not allow long down times of technical systems, and thus, it is reasonable to assume constant failure rates. Apart from the local warehouses, there exists a central warehouse, denoted by index 0. Let N denote the set of all warehouses; i.e., $N = \{0\} \cup N_{loc}$.

Suppose a component of SKU i of a technical system at some local warehouse n fails. Then the technical system goes down. The malfunctioning component is replaced by a spare part stocked at the local warehouse, if the local warehouse has stock on-hand. Otherwise, the component is backordered and the technical system remains down until a ready-for-use component becomes available at the local warehouse. The malfunctioning component is shipped to the central warehouse, where all failed components are repaired. At the same time, a request for a ready-for-use component is placed at the central warehouse. The order and ship time for a component i from the central warehouse to local warehouse n is denoted by T_{in} . This order and ship time is excluding a possible waiting time at the central warehouse in case a ready-for-use component is not immediately available there, and is assumed to be *deterministic*. For returned malfunctioning components at the central warehouse, it takes a *random* repair lead-time with mean T_{i0} until the component is returned to the spare parts stock at the central warehouse. Notice that implicitly a one-for-one replenishment policy has been assumed for all SKU-s at all local warehouses including the central warehouse. I.e., each SKU i at each warehouse n is controlled according to a base stock policy. The corresponding base stock level is denoted by S_{in} . As we deal with critical components at the assembly level and the items are expensive in general, the assumption of one-for-one replenishments is reasonable. Further, we assume that backordered demands for each SKU $i \in I$ at each warehouse $n \in N$ are treated in FCFS order. A holding cost h_i is counted for each unit of spare part of SKU i.

At local warehouse j, there is a maximum level W_n^{max} given for the aggregate mean waiting time per request for a ready-for-use component. Our goal is to determine a system's stocking policy $\underline{S}$ to minimize the total holding cost subject to the aggregate mean waiting time constraint per local warehouse, where $\underline{S}$ is represented as a $|I| \times |N|$ matrix.

2.2 Overview of assumptions

The main assumptions of the model are as follows:

(i) At each of the local warehouses, the failures for the different components occur according to independent Poisson processes with a constant rate.

(ii) All components are repairable and there is no condemnation.

(iii) For each SKU, the repair lead times of all items of that SKU are independent and identically distributed random variables.

(iv) For each SKU, the order and ship times are assumed to be deterministic.

(v) A one-for-one replenishment policy is applied for all SKU-s at all warehouses.

(vi) Replenishment orders at the central warehouse are fulfilled in FCFS order.

(vii) There is no lateral supply in the distribution network.

2.3 Evaluation

For a given base stock policy $\underline{S}$, evaluation of the steady-state behavior can be done exactly, as described in Graves [5]. In this subsection, we summarize this method for our system (we follow the formulae of Rustenburg et al. [11], in which Graves' exact recursion has been generalized to general multi-echelon, multi-indenture systems).

Define $l_{i0} = \sum_{n \in N_{loc}} l_{in}$, $i \in I$, as the total demand for SKU i at the central warehouse. Let Y_{in}, $i \in I$, $n \in N$, denote total demand during a time interval of length T_{in}. This Y_{in} is Poisson distributed with parameter $l_{in} T_{in}$, i.e., $P\{Y_{in} = y\} = (l_{in} T_{in})^y e^{-l_{in} T_{in}} / y!$, $y \geq 0$. For $n \in N_{loc}$, this follows directly from the Poisson distribution of l_{in}. For $n = 0$, this follows from Palm's theorem and the property that the total demand process at the central warehouse is Poisson. Define X_{i0} as the total amount on order at the central warehouse in steady state, i.e., the total amount in the pipeline from the supplier to the central warehouse. It holds that $X_{i0} = Y_{i0}$. From the distribution of X_{i0}, we can derive the distribution of $I_{i0}(S_{i0})$, the physical stock for SKU *i* at the central warehouse, and $B_{i0}(S_{i0})$, the backorder position at the central warehouse, as a function of the base stock level S_{i0}:

$$P\{I_{i0}(S_{i0}) = x\} = \begin{cases} \sum_{y=S_{i0}}^{\infty} P\{X_{i0} = y\}, & x = 0, \\ P\{X_{i0} = S_{i0} - x\}, & x \in \{1, \ldots, S_{i0}\}, \end{cases} \tag{1}$$

$$P\{B_{i0}(S_{i0}) = x\} = \begin{cases} \sum_{y=0}^{S_{i0}} P\{X_{i0} = y\}, & x = 0, \\ P\{X_{i0} = S_{i0} + x\}, & x > 0. \end{cases} \tag{2}$$

From this, we can easily determine $\overline{I}_{i0}(S_{i0})$, the expected inventory on hand of SKU i at the central warehouse at base stock level S_{i0}, and $\overline{B}_{i0}(S_{i0})$, the expected backorder level of SKU i at the central warehouse at base stock level S_{i0}.

Define $B_{i0}^{(n)}(S_{i0})$, $i \in I$, $n \in N_{loc}$, as the number of backorders of local warehouse n in the backorder queue at the central warehouse. As each backordered demand at the central warehouse stems from local warehouse n with probability l_{in}/l_{i0}, the probability distribution of $B_{i0}^{(n)}(S_{i0})$ is obtained by

$$P\{B_{i0}^{(n)}(S_{i0}) = x\} = \sum_{y=x}^{\infty} \binom{y}{x} \left(\frac{l_{in}}{l_{i0}}\right)^{x} \left(1 - \frac{l_{in}}{l_{i0}}\right)^{y-x} P\{B_{i0}(S_{i0}) = y\}. \tag{3}$$

Define $X_{in}(S_{i0})$ as the total amount on order at local warehouse n, i.e., the total amount in the pipeline from the central warehouse to the local warehouse n, under a given base stock level S_{i0}. It holds that $X_{in}(S_{i0}) = B_{i0}^{(n)}(S_{i0}) + Y_{in}$. From the distribution of $X_{in}(S_{i0})$, we can derive the distribution of $I_{in}(S_{i0}, S_{in})$, the physical stock for SKU i at local warehouse n, and $B_{in}(S_{i0}, S_{in})$, the backorder position at local warehouse n, as a function of the base stock levels S_{i0} and S_{in}, like in the central warehouse. Further, we can determine $\overline{I}_{in}(S_{i0}, S_{in})$, the expected inventory on hand of SKU i at local warehouse n, and $\overline{B}_{in}(S_{i0}, S_{in})$, the expected backorder level of SKU i at local warehouse n, as a function of the base stock levels S_{i0} and S_{in}.

The mean waiting time for getting a ready-for-use component of SKU $i \in I$ at local warehouse $n \in N_{loc}$ when the base stock level of SKU i is S_{i0} at the central warehouse and S_{in} at the local warehouse, $W_{in}(S_{i0}, S_{in})$, can be determined by Little's formula: $W_{in}(S_{i0}, S_{in}) = \overline{B}_{in}(S_{i0}, S_{in}) / l_{in}$. Taking all SKU-s together, the aggregate mean waiting time $W_n(\underline{S})$ at local warehouse $n \in N_{loc}$ is:

$$W_n(\underline{S}) = \sum_{i\in I} \text{Prob \{an arbitrary component demand at local warehouse } n \text{ is of SKU } i\} \times$$
$$\text{(mean waiting time for a component of SKU } i \text{ at local warehouse } n)$$

$$= \sum_{i\in I} \frac{l_{in}}{\sum_{k\in I} l_{kn}} \times \frac{\overline{B}_{in}(S_{i0}, S_{in})}{l_{in}} = \sum_{i\in I} \frac{\overline{B}_{in}(S_{i0}, S_{in})}{\sum_{k\in I} l_{kn}}. \tag{4}$$

2.4 Problem formulation

Our optimization problem can be formulated as:

$$\underset{\underline{S}}{Min} \quad Z(\underline{S}) = \sum_{i\in I} h_i \overline{I}_i(S_{i0}) + \sum_{i\in I} \sum_{n\in N_{loc}} h_i \overline{I}_{in}(S_{i0}, S_{in}) \tag{5}$$

$$\text{subject to} \quad W_n(\underline{S}) \le W_n^{\max}, \quad n \in N_{loc}, \tag{6}$$

$$S_{in} \text{ integer}, \qquad i \in I, \; n \in N. \tag{7}$$

An optimal policy is denoted by $\underline{S}^*$ and the optimal costs by $Z(\underline{S}^*)$. Note that in the above formulation, the pipeline stock is excluded in the calculation of the inventory holding costs. Note also that in comparison to the formulation of Caglar et al. [3], our formulation is slightly different as no upper bound is assumed for S_{in}.

3. Basic procedures

In this section, we describe the basic procedures that are used as building blocks in our heuristic methods and for the computation of the lower bound on the optimal total cost. In Section 4, we will discuss how these basic procedures are combined in the different methods. The basic procedures include the exact and approximate evaluation (subsection 3.1), the decomposition and column generation method (subsection 3.2), the greedy approach (subsection 3.3), and the local search improvement (subsection 3.4).

3.1 Exact and approximate evaluation

Exact evaluation can be done following the method developed by Graves [5], which has been summarized in Subsection 2.3. A computational issue occurs, since $P\{Y_{in} = y\}$, $i \in I$, $n \in N$, should be calculated for all values $y \geq 0$. In practice, however, for each $i \in I$, $n \in N$, we limit ourselves to $y \in \{0,..., y_{in}^{\max}\}$, with $y_{in}^{\max} = \min\{y \mid P\{Y_{in} \leq y\} \geq 1 - e\}$ and $\in = 10^{-6}$, and allocate the remaining probability mass $1 - P\{Y_{in} \leq y_{in}^{\max}\}$ to $P\{Y_{in} = y_{in}^{\max}\}$. Furthermore, this computation must be done for every value of S_{i0} that is investigated.

To reduce the computational burden, we can use an approximate evaluation method like METRIC or Graves' approximation. The METRIC approximation assumes that successive replenishments at the local warehouses are independent processes, which leads to a Poisson distribution. Graves [5] proposed a different approximation that uses the two-parameter negative binomial distribution to fit the distribution of the backorders at the local warehouses. This two-parameter approximation is, in general, more accurate than the METRIC approximation. In Section 5 we present the results of experiments evaluating the accuracy of both approximate methods when used for executing the greedy procedure.

3.2 A decomposition and column generation method

A lower bound on $Z(\underline{S})$ can be obtained by a *decomposition and column generation (DCG) method* which reveals close similarity to Dantzig-Wolfe decomposition for linear programming problems. In Kranenburg and Van Houtum [8], this method is described in detail for another multi-item spare parts problem. A decomposition and column generation method is appropriate for problems that have a complicated aggregation constraint (like constraints (6) in our problem), but by decomposition of the problem can be reduced to relatively simple sub-problems (in the problem under consideration: per SKU i). In this subsection, we limit ourselves to a general description of the method for our problem.

Like in the Dantzig-Wolfe decomposition, a linear programming problem, the *Master Problem*, is introduced in which the variables of our original problem (base stock levels) are expressed as convex combination of columns that contain possible values for the decision variables in the original problem. Besides the Master Problem, a *Restricted Master Problem* is defined that only considers a subset of all possible columns. The method starts with some initial columns which constitute a feasible solution for the Restricted Master Problem, and solves the Restricted Master Problem with the simplex method. Next, the method iteratively solves a *Sub-Problem* for each SKU $i \in I$, to determine if there exists a column for that SKU that would improve the solution, adds this column to the Restricted Master Problem, and solves the Restricted Master Problem.

The Sub-Problem for one SKU i is

$$\underset{\underline{S}}{Min} \quad h_i \overline{I}_i(S_{i0}) + \sum_{n \in N_{loc}} h_i \overline{I}_{in}(S_{i0}, S_{in}) - \sum_{n \in N_{loc}} u_n \frac{\overline{B}_{in}(S_{i0}, S_{in})}{\sum_{i \in I} l_{in}} - v_i \tag{8}$$

$$\text{subject to} \quad S_{in} \text{ integer}, \qquad i \in I,\ n \in N, \tag{9}$$

where u_n is the shadow price of the waiting time constraint for local warehouse $n \in N_{loc}$ in the Restricted Master Problem, and v_i is the shadow price for a constraint in the Restricted Master Problem that assures that for SKU i a convex combination of policies is chosen. Shadow price $u_n \leq 0$ by definition for all $n \in N_{loc}$. This Sub-Problem comes down to solving a single-item cost minimization problem with linear inventory and backordering costs (but without service level constraints). This optimization problem is *precisely the problem studied by Axsäter [1]* and we solve this problem by his method. If the resulting policy for a SKU i has a negative reduced cost (i.e. a negative value of the objective function (8)), this policy is added as a column to the Restricted Master Problem. The method ends if for none of the SKU-s a policy with a negative reduced cost is found.

The method results in a lower bound on $Z(\underline{S}^*)$. The corresponding policy, however, in general will not be a base stock policy, but a convex combination of base stock policies (columns in the Restricted Master Problem), which in fact constitutes a randomized policy. The costs of this randomized policy are a lower bound on $Z(\underline{S}^*)$.

In the heuristic methods described later, the DCG may be followed by a greedy approach to obtain a feasible solution for our problem formulated in (5)-(7). For that approach we need a policy $\underline{S}$ as starting policy. We use the following starting policy. For each $i \in I$ and $n \in N$, we select the smallest value for S_{in} that is found among the base stock levels of the convex combination.

3.3 A greedy approach

A feasible solution may be obtained in an efficient way via a greedy procedure similar to the procedure described in Wong et al. [14] for a multi-item, multi-location problem with lateral transshipments. The basic idea of this procedure is to add units of stock in an iterative way. At each iteration, we add one unit of stock for an SKU $i \in I$ at a warehouse $n \in N$ such that we gain the largest decrease in *distance to the set of feasible solutions* per extra unit of additional cost. The procedure is terminated when a feasible solution is obtained.

Let U_{ab} be a $|I| \times |N|$ matrix containing zero values at all cells except for the cell (a,b) that has a value of one. The procedure starts by setting zero stock for all SKU-s and warehouses, i.e. $\underline{S} = \mathbf{0}$. We define for each solution $\underline{S}$ the distance to the set of feasible solutions as $\sum_{n \in N_{loc}} (W_n(\underline{S}) - W_n^{max})^+$ where $(x)^+ = \max(0, x)$. In each iteration, for each combination of $i \in I$ and $n \in N$, we calculate the ratio $r_{in} = \frac{\Delta W_{in}}{\Delta Z_{in}}$ where:

$$\Delta W_{in} = \sum_{n \in N_{loc}} \left((W_n(\underline{S}) - W_n^{max})^+ - (W_n(\underline{S} + U_{in}) - W_n^{max})^+ \right), \tag{10}$$

and

$$\Delta Z_{in} = Z(\underline{S} + U_{in}) - Z(\underline{S}). \tag{11}$$

One unit of stock is then added for the combination with the largest ratio. Notice that the formulae for ΔW_{in} and ΔZ_{in} are easily simplified based on the structure in the functions $W_n(\underline{S})$ and $Z(\underline{S})$. This is exploited in the computations. A formal description of the greedy procedure is as follows:

A greedy procedure

Step 1: Set the initial solution $\underline{S} = \mathbf{0}$; Calculate $W_n(\mathbf{0})$ for all local warehouses $n \in N_{loc}$.

Step 2: For all combinations $i \in I$ and $n \in N$: Calculate ΔW_{in}, ΔZ_{in}, and r_{in}.

Step 3: Let i^* and n^* be the combination with the highest ratio r_{in}. Set $\underline{S} = \underline{S} + U_{i^*n^*}$.

If $W_n(\underline{S}) \le W_n^{max}$ for all $n \in N_{loc}$ go to *END*; Otherwise go to *Step* 2.

END

3.4 Local search

Once a feasible solution has been obtained, one may apply a local search method to obtain a further improved solution. We apply a greedy (steepest descent) local search method that allows us to explore the entire neighborhood at each iteration. At each iteration, all possible neighbors of the current solution are evaluated, and the one with the minimum total cost is selected. If the new total cost is less than the current total cost, the selected solution becomes the current solution. Otherwise, no local improvement is possible and we take the current solution as the heuristic's solution.

For each solution $\underline{S}$, we define the neighborhood of $\underline{S}$ as follows:

$NE(\underline{S}) = NE_1(\underline{S}) \bigcup NE_2(\underline{S}) \bigcup NE_3(\underline{S}) \bigcup NE_4(\underline{S})$ where

$$NE_1(\underline{S}) = \left\{ \text{all } \underline{S}^{'} \in \mathbb{S} \mid \underline{S}^{'} = \underline{S} - U_{in}, i \in I, n \in N \right\};$$

$$NE_2(\underline{S}) = \left\{ \text{all } \underline{S}^{'} \in \mathbb{S} \mid \underline{S}^{'} = \underline{S} + U_{in}, i \in I, n \in N \right\};$$

$$NE_3(\underline{S}) = \left\{ \text{all } \underline{S}^{'} \in \mathbb{S} \mid \underline{S}^{'} = \underline{S} + U_{in} - U_{i'n'},\ i \in I,\ i' \in I,\ i \neq i',\ n \in N,\ n' \in N \right\};$$

$$NE_4(\underline{S}) = \left\{ \text{all } \underline{S}^{'} \in \mathbb{S} \mid \underline{S}^{'} = \underline{S} + U_{in} - U_{in'},\ i \in I,\ n \in N,\ n' \in N,\ n \neq n' \right\};$$

where $\mathbb{S}$ is the set consisting of all feasible solutions.

The neighborhood of a solution can be seen as an integration of four sub-neighborhoods. The first sub-neighborhood is formed by reducing one unit of stock in the system. Obviously, since the total cost for our problem is a function of the expected inventory on hand, any possible move to the first sub-neighborhood would always give better solutions. In contrast, exploring the second sub-neighborhood, which is formed by adding one unit of stock, would always lead to more expensive policies. Hence, any move to the second sub-neighborhood would never be accepted. However, to give a general structure of the entire neighborhood of a solution, the second sub-neighborhood is included in the definition as described above. The third sub-neighborhood is formed by removing one unit of an SKU and putting another SKU as a replacement. A cost reduction may be obtained here as an expensive component is removed and replaced with a less expensive component. Lastly, exploring the fourth sub-neighborhood may be useful to obtain the best stock allocation across all warehouses.

It can be shown that the upper bound on the neighborhood size of a solution is equal to $|I||N|(1+|I||N|)$. At the first iteration, we need to evaluate at most $2|I||N|$ neighboring solutions for the first and second sub-neighborhoods, and $|I||N|(|N|-1)$ solutions for the fourth sub-neighborhood. To evaluate a solution lying in the third sub-neighborhood, we can use the results obtained from the first and second sub-neighborhood. At the subsequent iterations, we only need to evaluate one or a few new neighbors since any changes of the stock levels for a given SKU do not affect the results for the other SKU-s.

4. Computational Experiments

In this section, we describe different heuristic methods to solve our optimization problem (subsection 4.1) and present the set up and results of the computational experiments for the evaluation of these heuristic methods (subsections 4.2 and 4.3).

4.1 Description of heuristics

We now describe how the basic procedures described in the previous section are combined to form heuristic methods. There are four different heuristic methods that we would like to test, namely:

- *Heuristic 1*: Greedy approach
- *Heuristic 2*: Greedy approach + Local search
- *Heuristic 3*: DCG (+ Greedy approach)
- *Heuristic 4*: DCG (+ Greedy approach) + Local search

In Heuristic 1 we only apply the greedy approach. This means that the procedure is terminated when a feasible base stock policy is obtained. In Heuristic 2, we continue the procedure of Heuristic 1 by applying a local search method that may lead to base stock policies with lower total costs. Comparisons between these two heuristic methods would give us information on how far solutions obtained by the greedy approach are from local optima. The next two heuristic methods, Heuristic 3 and Heuristic 4, are based on the lower bounding procedure (decomposition and column generation). As previously explained, the resulting policy of this procedure in general will not be a base stock policy, but a convex combination of base stock policies. A starting base stock policy $\underline{S}$ is obtained by selecting the smallest value of S_{in} (for each $i \in I$ and $n \in N$) that is found among the base stock levels of the convex combination. Two possibilities exist with regard to the resulting base stock policy $\underline{S}$. First, $\underline{S}$ is a feasible policy. In that case $\underline{S}$ becomes the solution of Heuristic 3 or the initial solution for the local search procedure applied in Heuristic 4. Second, $\underline{S}$ is not a feasible policy, and in that case a greedy approach is applied to obtain a feasible policy.

In our experiments, we do not include the method presented in Caglar et al. [3] as a benchmark because they developed their algorithm based on the METRIC approximation. We show in the next section that the use of METRIC approximation results in infeasible solutions in many cases. As we would like to have an objective evaluation of all the heuristics in the sense that comparisons between heuristics are made based on the resulting *feasible* solutions, it is therefore not possible to compare the heuristic methods listed above to the heuristic method of Caglar et al [3].

4.2 Experimental test beds

Four test beds are considered in our experiments. In the first test bed, we consider symmetrical cases in which the demand rates across all the local warehouses are identical, but they are varied for different SKU-s. In the second test bed, we consider asymmetrical cases in which the demand rates for different local warehouses are different. For those two test beds, the target aggregate mean waiting times of all local warehouses are identical. To see how the heuristic methods perform when the local warehouses have different targets, we did experiments based on the third and fourth test beds.

Test bed 1

We consider inventory systems with two different numbers of local warehouses ($|N_{loc}|$ = 5 and 20) and two different numbers of SKU-s ($|I|$ = 20 and 100). With regard to the demand rates l_{in} for $i \in I$ and $n \in N_{loc}$, a uniform distribution U(0.002, 0.08) is used to generate the demand rates for all SKU $i \in I$. Values for the inventory holding costs were generated from two uniform distributions U(100, 1000) and U(100, 10000), representing two different variability levels of the inventory holding costs of all SKU-s. The order and ship time from the central warehouse to local warehouse is fixed at one day and assumed to be identical for all local warehouses and all SKU-s. For the repair lead time at the central warehouse, we tested two values ($T_{i0} = 1$ day and $T_{i0} = 10$ days) for all SKU-s. Further, two values were used for the target aggregate mean waiting time ($W_n^{\max} = 0.1$ day and $W_n^{\max} = 0.3$ day) and we consider symmetrical cases in which all local warehouses have the same targets. In our experiments, we generated five data samples of the demand rates for each combination of all other parameters (the same holding cost parameters are used for each set of these five data samples). These parameter settings result into 160 instances. Table 1 summarizes the parameter settings used in the first test bed.

Table 1. Parameter values for test bed 1

Name of the parameter	Unit	Number of values	Values
Number of local warehouses $\vert N_{loc} \vert$		2	5 and 20
Number of SKU-s $\vert I \vert$		2	20 and 100
Demand rate l_{in}	unit/day	1	U(0.002, 0.08)
Inventory holding cost h_i	\$/unit/day	2	U(100,1000) and U(100, 10000)
Order and ship time from the central warehouse T_{in}	days	1	1
Repair lead time T_{i0}	days	2	1 and 10
Maximum waiting time W_j^{max}	days	2	0.1 and 0.3

Test bed 2

In the second test bed, we consider cases with asymmetrical demand rates. The same uniform distribution U(0.002, 0.08) is used to generate demand rates for all SKU-s $i \in I$. Next, for each SKU, the demand rate at each local warehouse is determined by multiplying the generated demand rate of this SKU with a factor generated from the second uniform distribution U(0.2, 2). The other parameters are set in the same way as for the first test bed. There are 160 instances experimented for the second test bed.

Test bed 3

In the third test bed, we consider cases with five local warehouses in which the demand rates are identical across the local warehouses. The target aggregate mean waiting times for the five local warehouses are set at 0.1, 0.15, 0.2, 0.25 and 0.3, respectively. For the other parameters, we used the same data as in the instances of the first test bed with $|N_{loc}| = 5$. Hence, 40 instances are obtained.

Test bed 4

This test bed is similar to the third test bed except that now asymmetrical demand rates for the five local warehouses are taken. For this test bed, we used all the demand rates of the instances of the second test bed with $|N_{loc}| = 5$. Hence, again 40 instances are obtained.

For the evaluation of heuristic methods, we measured the relative difference between the total cost obtained by the heuristic and the corresponding lower bound *(%GAP)*. That is,

$$\%GAP = \frac{\text{heuristic's total cost - lower bound}}{\text{lower bound}} \times 100$$

4.3 Computational results

The results of our experiments for the four test beds are summarized in Tables 2-5. In each table, we present the performance of each heuristic method in terms of the average and maximum value of *%GAP*, where we first distinguish subsets of instances with the same value for a specific input parameter and in the bottom line the results for all instances together are presented. For example, in test bed 1, the average and maximum *%GAP* obtained by the greedy approach for all instances with $|N_{loc}|$ = 5 are equal to 7.46% and 15.35%, respectively (see the results in the first line of Table 2). For the same test bed, the average and maximum *%GAP* obtained by the greedy approach for all 160 instances together are equal to 7.11% and 23.29%, respectively (see the last line in Table 2). The performance of each heuristic in terms of computation time is presented in Table 6 (*programs for executing all heuristics are written in MATLAB and all the experiments were run on a PC with a Pentium4 2.8 GHz processor and 3.37 GB RAM*).

Table 2. Experiment results for test bed 1 (symmetric demand rates, symmetric target aggregate mean waiting times)

Parameter		Heuristic							
		1: Greedy		2: Greedy + LS		3: DCG		4: DCG + LS	
		Avg	Max	Avg	Max	Avg	Max	Avg	Max
$\lvert N_{loc} \rvert$	5	7.46	15.35	5.43	14.62	1.14	7.78	0.61	7.46
	20	6.75	23.29	4.68	14.94	1.00	8.82	0.76	4.68
$\lvert I \rvert$	20	8.28	23.29	5.93	14.94	2.01	8.82	1.29	7.46
	100	5.93	9.65	4.18	8.21	0.13	0.64	0.07	0.35
h_i	U(100,1000)	6.98	14.37	5.13	14.12	0.90	7.23	0.65	7.23
	U(100,10000)	7.24	23.29	4.98	14.94	1.23	8.82	0.72	7.46
t_0	1	7.31	15.35	5.21	14.62	1.30	8.82	0.91	7.46
	10	6.91	23.29	4.90	14.94	0.84	4.51	0.45	3.76
W^{max}	0.1	6.31	15.04	4.41	11.35	0.92	8.82	0.56	7.46
	0.3	7.91	23.29	5.70	14.94	1.22	7.23	0.80	7.23
All		7.11	23.29	5.06	14.94	1.07	8.82	0.69	7.46

Table 3. Experiment results for test bed 2 (asymmetric demand rates, symmetric target aggregate mean waiting times)

Parameter		Heuristic 1: Greedy Avg	1: Greedy Max	2: Greedy + LS Avg	2: Greedy + LS Max	3: DCG Avg	3: DCG Max	4: DCG + LS Avg	4: DCG + LS Max
$\lvert N_{loc} \rvert$	5	3.36	12.19	2.86	10.39	1.34	5.03	1.12	4.77
	20	2.55	6.61	2.34	6.35	1.68	5.35	1.50	5.05
$\lvert I \rvert$	20	4.42	12.20	3.86	10.39	2.76	5.35	2.39	5.05
	100	1.49	3.37	1.34	2.89	0.27	0.71	0.23	0.43
h_i	U(100,1000)	2.79	7.93	2.46	7.91	1.45	5.25	1.33	5.05
	U(100,10000)	3.12	12.20	2.74	10.39	1.58	5.35	1.29	4.61
t_0	1	3.70	12.20	3.26	10.39	1.88	5.35	1.62	5.05
	10	2.21	5.39	1.94	5.15	1.15	4.10	1.00	3.42
$W^{\max}$	0.1	2.60	7.60	2.24	6.57	1.28	4.29	1.07	3.74
	0.3	3.31	12.20	2.96	10.39	1.75	5.35	1.55	5.05
All		2.96	12.20	2.60	10.39	1.51	5.35	1.31	5.05

Table 4. Experiment results for test bed 3 (symmetric demand rates, asymmetric target aggregate mean waiting times)

Parameter		Heuristic 1: Greedy Avg	1: Greedy Max	2: Greedy + LS Avg	2: Greedy + LS Max	3: DCG Avg	3: DCG Max	4: DCG + LS Avg	4: DCG + LS Max
$\lvert I \rvert$	20	7.81	10.98	6.55	10.04	2.82	5.09	2.36	4.30
	100	4.56	5.97	3.21	4.80	0.32	0.48	0.27	0.43
h_i	U(100,1000)	6.06	9.94	5.09	9.94	1.41	4.78	1.24	4.30
	U(100,10000)	6.31	10.98	4.68	10.04	1.73	5.09	1.39	3.97
t_0	1	6.93	10.98	5.78	10.04	1.95	5.09	1.76	4.30
	10	5.44	10.57	3.98	7.71	1.19	3.24	0.87	3.97
All		6.19	10.98	4.88	10.04	1.57	5.09	1.32	4.30

Table 5. Experiment results for test bed 4 (asymmetric demand rates, asymmetric target aggregate mean waiting times)

Parameter		Heuristic 1: Greedy		2: Greedy + LS		3: DCG		4: DCG + LS	
		Avg	Max	Avg	Max	Avg	Max	Avg	Max
$\mid I \mid$	20	4.28	6.78	3.61	6.62	2.68	5.05	2.20	4.87
	100	1.94	2.92	1.61	2.89	0.28	0.40	0.24	0.40
h_i	U(100,1000)	2.99	5.57	2.55	5.02	1.09	3.65	0.99	3.58
	U(100,10000)	3.23	6.78	2.66	6.62	1.87	5.05	1.46	4.87
t_0	1	3.79	6.78	3.14	6.62	1.94	5.05	1.60	4.87
	10	2.43	4.20	2.07	3.61	1.01	3.27	0.85	2.44
All		3.11	6.78	2.61	6.62	1.48	5.05	1.22	4.87

Table 6. Average computation time for each heuristic (seconds)

Parameter		Heuristic 1: Greedy	2: Greedy + LS	3: DCG	4: DCG + LS
$\mid I \mid = 20$	$\mid N_{loc} \mid = 5$	4.80	7.24	90.35	92.79
	$\mid N_{loc} \mid = 20$	66.37	119.48	1132.50	1185.60
$\mid I \mid = 100$	$\mid N_{loc} \mid = 5$	23.92	38.40	437.05	490.16
	$\mid N_{loc} \mid = 20$	321.50	1015.12	6728.24	7421.95

The main observations drawn from these tables can be summarized as follows:

- The DCG heuristics (heuristics 3 and 4) perform very well. In all four test beds the average *%GAP* is below 2% and the maximum below 10%.
- The greedy heuristics (heuristics 1 and 2) perform also very well in the test beds 2 and 4 with asymmetrical demand rates. The average *%GAP* is around or below 3% in these test beds, and the maximum below 12%. When limiting ourselves to the instances with 100 items in these test beds, we see that the average *%GAP* is even below 2% and the maximum below 4%. The performance of the greedy heuristics in the test beds 1 and 3 with symmetrical demand rates is less good: the average *%GAP* is around or somewhat below 7%. We think that this phenomenon is due to how the greedy heuristic works. With symmetrical demand rates, we get the property that if in a given iteration an item of a specific SKU is stocked at one local warehouse, then also an item of the same SKU is

stocked at all other local warehouses in the succeeding iterations. This behavior strengthens the discrete character of our optimization problem. Once there is some asymmetry in the demand rates, as one will always have in practical applications, the phenomenon will disappear.

- The improvements obtained by local search are quite limited for both the greedy procedure (compare the results of the heuristics 1 and 2) and the DCG method (compare the results of the heuristics 3 and 4).
- The average values of *%GAP* tend to decrease as the problem size (in terms of number of items and local warehouses) or the required stock levels get larger. The latter occurs when the average repair lead time is higher (10 as opposed to 1) or when the target average waiting time is lower (0.1 as opposed to 0.3). This observation is in line with the findings in Wong et al. [15] for a single-echelon, multi-location system with lateral transshipments.
- It is shown in Table 6 that the greedy method is the most efficient heuristic in terms of computation time. Significant additional computation times are required when the local search method is applied to improve the solution obtained from the greedy heuristic (compare e.g. 321.50 and 1015.12 seconds for the biggest problem size in our experiments). As expected, the computation time of the DCG method is considerably high.

The results of our computational experiments indicate that the greedy procedure (heuristic 1) is a very appropriate approach for solving the optimization problem in a multi-item two-echelon spare parts system as analyzed in this paper. This approach has been proven to be quite effective particularly for solving large-sized problems, which are indeed the type of problems typically faced in practice, and furthermore, this approach is easy to implement. To reduce the computational burden of the procedure, it would be worthwhile to use approximate evaluations instead of the exact method. We analyze this issue in the following section.

5. Applying approximate evaluation methods

As the computational requirements of the exact evaluation method are rather extensive, using an approximate instead of exact evaluation method is one way to increase the speed of the greedy procedure. An accurate approximate method will lead us to walk through about the same solutions (as with the exact method) while executing the greedy heuristic, and the generated solution will approximately satisfy the aggregate mean waiting time constraints. We know already that Graves' approximation based on two-moments fits is quite accurate. This has been tested in Graves [5]. We therefore are particularly interested in evaluating the accuracy of the greedy heuristic when executed using this evaluation method. For this purpose, we conducted experiments using the 320 instances of test bed 1 and test bed 2 presented in the previous section. For each instance, the greedy heuristic was executed using Graves' approximate evaluation method. At the termination of the procedure, the solution obtained was then evaluated using the exact method. We recorded the result with regard to whether or not the generated solution is feasible. If the solution is not feasible, we are also interested in measuring the distance to the set of feasible solutions. Such a measure is calculated by a similar expression to the one used in the greedy procedure. For a solution $\underline{S}$, we calculate the relative distance as: $\sum_{n \in N_{loc}} (W_n(\underline{S}) - W_n^{max})^+ \Big/ \sum_{n \in N_{loc}} W_n^{max}$ where $(x)^+ = \max(0, x)$.

The METRIC approximation is another approximate evaluation method that is less accurate (see Graves [5]) but widely used in practice. In our experiments, we also applied the METRIC method as an alternative for Graves' method so that the effect of using that method is also evaluated.

The results of our computational experiments are reported in Table 7. Information on the average computation time under the use of the two approximate evaluation methods and under exact evaluation (as a function of the number of SKU-s and local warehouses) is presented in Table 8.

Table 7. Performance of approximate evaluation methods

Parameter		METRIC			GRAVES		
		number of feasible solutions	distance to feasible region (avg)	distance to feasible region (max)	number of feasible solutions	distance to feasible region (avg)	distance to feasible region (max)
$\lvert N_{loc} \rvert$	5	13 / 160	4.56 %	23.60 %	53 / 160	0.11 %	0.75 %
	20	13 / 160	1.82 %	10.03 %	62 / 160	0.02 %	0.16 %
$\lvert I \rvert$	20	26 / 160	2.71 %	23.60 %	82 / 160	0.04 %	0.49 %
	100	0 / 160	3.67 %	21.11 %	33 / 160	0.09 %	0.75 %
Demand	symmetric	23 / 160	3.65 %	23.60 %	71 / 160	0.09 %	0.75 %
	asymmetric	3 / 160	2.74 %	10.03 %	44 / 160	0.04 %	0.16 %
All		26 / 320	3.19 %	23.60 %	115 / 320	0.07 %	0.75 %

Table 8. Average computation times for executing the greedy procedure (seconds)

		METRIC	Graves	Exact
$\|I\| = 20$	$\|N_{loc}\| = 5$	0.93	2.39	4.80
$\|I\| = 100$	$\|N_{loc}\| = 5$	6.39	13.33	23.92
$\|I\| = 20$	$\|N_{loc}\| = 20$	9.49	13.07	66.37
$\|I\| = 100$	$\|N_{loc}\| = 20$	68.04	84.57	321.50

The results show us several important observations:

- As expected, METRIC is inferior to the Graves method with respect to the number of feasible solutions and the distance to the set of feasible solutions. METRIC is only able to provide feasible solutions in 26 out of 320 data sets. The worst case is observed for data sets with 100 items in which METRIC never gives a feasible solution. The average and maximum relative distances to the set of feasible solutions for METRIC are 3.19% and 23.6%. Graves' method performs much better by giving 115 feasible solutions with 0.07% and 0.75% for the average and maximum relative distances.

- We observe too that both methods are more accurate when more local warehouses are involved. For Graves' method, we can see that all the measures suggest that higher accuracy is obtained in problems with 20 local warehouses than in problems with 5 local warehouses. For METRIC, we observe a reduction in relative distances although the number of feasible solutions remains unchanged. This is in line with what was pointed out by Axsäter [2]: the METRIC approximation will be more accurate as long as the demand at each local warehouse is low relative to the total demand.

- The results for both methods deteriorate when dealing with a larger number of items. The distribution of demand across local warehouses also seems to be an influencing factor. Both methods are more likely to generate a feasible solution when used for problems with identical demands across local warehouses. In terms of the average relative distance, however, both methods are more accurate when used for problems with non identical demands.

The results of our experiments show that Graves' approximate evaluation method is very appropriate to be used within the greedy approach. This method gives highly accurate results and requires much less computational efforts than exact evaluations. We do not recommend to use the METRIC approximation as it may lead to solutions that are far from the feasible region.

6. Conclusions

In this paper, we have developed solution procedures for the optimization of base-stock levels of a multi-item, two-echelon spare parts system. The problem being dealt with is determining a close-to-optimal stocking policy that minimizes the system wide inventory holding cost while satisfying aggregate mean waiting time constraints per local warehouse. Four different heuristic methods have been developed and evaluated based on the relative distance between the heuristic's total cost and its corresponding lower bound. Exact evaluation has been used instead of approximate evaluation to compare the performance of these heuristic methods. To calculate the lower bounds, we have developed a decomposition and column generation method which reveals close similarity to the Dantzig-Wolfe decomposition for linear programming problems. In particular, our computational results show that the greedy procedure, which is quite simple to implement, is a very appropriate heuristic. It performs extremely well when used for solving large-sized problems with non-identical demand rates across local warehouses. An average distance to the lower bound of less than 2% was observed in our experiments for the problem instances with 100 SKU-s and non identical demand rates across the local warehouses. Further, additional experiments have been conducted to test the accuracy of approximate evaluation methods (METRIC and Graves' method) when used within the greedy heuristic. The results of these experiments suggest that Graves' method can be safely used instead of the exact method when there is a need to reduce the computational burden, e.g., when dealing with large-sized problems encountered in practice. Graves' method has proven to give accurate results while the computational efforts can be reduced significantly. The use of METRIC is not recommended as it may lead to solutions that are far from the feasible area.

References

[1] S. Axsäter, Simple solution procedures for a class of two-echelon inventory problems. *Oper Res* 38 (1990), 64-69.

[2] S. Axsäter, "Continuous review policies for multi-level inventory systems with stochastic demand,", Logistics of Production and Inventory, S.C. Graves, A.H.G. Rinnooy Kan, and P.H. Zipkin (editors), North-Holland, New York, 1993.

[3] D. Caglar, C.L. Li, and D. Simchi-Levi, Two-echelon spare parts inventory system subject to a service constraint, *IIE Trans* 36 (2004), 655-666.

[4] M. Cohen, P.V. Kamesam, P. Kleindorfer, H. Lee, and A. Tekerian, Optimizer: IBM's multi-echelon inventory system for managing service logistics, *Interfaces* 20 (1990), 65-82.

[5] S.C. Graves, A multi-echelon inventory model for a repairable item with one-for-one replenishment, *Management Sci* 31 (1985), 1247-1256.

[6] W.H. Hausman and N.K. Erkip, Multi-echelon vs. single-echelon inventory control policies for low demand items, *Management Sci* 40 (1994), 597-602.

[7] W.J. Hopp, R.Q. Zhang, and M.L. Spearman, An easily implementable hierarchical heuristic for a two-echelon spare parts distribution system. *IIE Trans* 31 (1999), 977-988.

[8] A.A. Kranenburg, and G.J. van Houtum, Effect of commonality on spare parts provisioning costs for capital goods. *Int J Prod Econ*, forthcoming.

[9] J.A. Muckstadt, A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Sci* 20 (1973), 472-481.

[10] J.A. Muckstadt and L.J. Thomas, Are multi-echelon inventory methods worth implementing in systems with low-demand-rate items? *Management Sci* 26 (1980), 483-494.

[11] J.W. Rustenburg, G.J. van Houtum, and W.H.M Zijm, "Exact and approximate analysis of multi-echelon, multi-indenture spare parts systems with commonality," *Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains*, J.G Shantikumar, D.D. Yao, and W.H.M. Zijm (Editors), Kluwer, Boston, 2003.

[12] C.C. Sherbrooke, METRIC: a multi-echelon technique for recoverable item control. *Oper Res* 16 (1968), 122-141.

[13] C.C. Sherbrooke, *Optimal inventory modeling of systems*, Kluwer, Boston, 2004.

[14] F.M. Slay, VARI-METRIC: An approach to modelling multi-echelon resupply when the demand process is Poisson with a gamma prior. *Technical Report AF301-3*, Logistics Management Institute, Washington D.C., 1984.

[15] H. Wong, G.J. Van Houtum, D. Cattrysse, and D. Van Oudheusden, Simple, efficient heuristics for multi-item multi-location spare parts systems with lateral transshipments and waiting time constraints. *J Oper Res Soc*, forthcoming (doi:10.1057/palgrave.jors.2601952)