Optimising a general repair kit problem with a service constraint

Marco Bijvank[1], Ger Koole

Department of Mathematics, VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Iris F.A. Vis

Department of Information Systems and Logistics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

---

Field services are a particular type of after-sales service performed at the customer's location where technicians repair malfunctioning machines. The inventory decisions about which spare part types to take to the repair site and in what quantities is called the repair kit problem. This problem is characterized by an order-based performance measure since a customer is only satisfied when all required spare parts are available to fix the machine. As a result, the service level in the decision making process is defined as a job fill rate. In this paper we derive a closed-form expression for the expected service level and total costs for the repair kit problem in a general setting, where multiple units of each part type can be used in a multi-period problem. Such an all-or-nothing strategy is a new characteristic to investigate, but commonly used in practice. Namely, items are only taken from the inventory when all items to perform the repair are available in the right quantity. We develop a new algorithm to determine the contents of the repair kit both for a service and cost model while incorporating this new expression for the job fill rate. We show that the algorithm finds solutions which differ on average 0.2% from optimal costs. We perform a case study to test the performance of the algorithm in practice. Our approach results in service level improvements of more than 30% against similar holding costs.

Keywords: inventory; repair kit problem; multi-period; multi-item; closed-form expression

---

1 Introduction

When customers buy a machine or piece of equipment they expect to get more than just the physical product. They also expect to get after-sales service regarding any malfunctioning of the machine during its life cycle. Nowadays, offering a good after-sales service is becoming a more and more important differentiation strategy among competitors in a business market. One particular service that is offered is a local repair service at the customer's facility, for example, for copiers or coffee machines. This means that a service company is called when the machine breaks down. On a daily basis, the service company assigns repair jobs to technicians based on their skills. As a result, the technicians receive a list indicating the order of the call points that they have to visit and they travel around to repair the broken machines. Since it is not known in advance which parts of the machine have to be replaced, the repair person takes along a selection of the spare parts in the car. This set of parts is referred to as the *repair kit*. The technician can only complete a repair if all required spare parts are available in the right quantity in the repair kit. When one or more units of the required parts are unavailable, the technician cannot fix the machine and has to return when the car is restocked with all the items that are required to repair the machine. This extra visit is called a *return-to-fit* (RTF) visit. An important logistics decision problem for technicians is to determine which spare parts to put in the repair kit and in which quantities to avoid return-to-fit visits. In literature this problem is referred to as the *repair kit problem*. Notice that for an RTF visit it is known in advance which parts are required. Therefore, these repairs are not considered in the repair kit problem.

According to Bijvank [1] and as will appear from a case study in Section 6, companies usually base the contents of their repair kits upon experiences and practical limitations (e.g., the capacity of the car and the amount of money spent to purchase the contents of the repair kit). It would be more efficient to have a systematic procedure to determine the contents of a repair kit. Such a procedure could be based on costs but also on a service level granted to customers. We distinguish between two types of costs, namely holding costs and return-to-fit costs. A fixed amount of *holding costs* is incurred for each unit that is stored in the repair kit of the car. *RTF costs* are involved when a technician has to return because at least one of the required parts is not available in the repair kit. These RTF costs usually consist of the actual labour and driving costs, as well as costs due to loss of goodwill. As a result, a trade-off has to be made between holding costs and RTF costs. Two kinds of models are developed in the literature to make this trade-off, namely cost models and service models. In a *cost model* the sum of the expected holding and RTF costs is minimised, whereas in a *service model* the holding costs are minimised subject to a service level constraint. This latter model is preferred in practice due to the difficulty to quantify the extra cost for an RTF visit. Moreover, this type of model explicitly incorporates a customer service level criterion such that a minimum quality of service is guaranteed.

Only a few papers address the repair kit problem. These papers impose several unrealistic assumptions (see Section 2 for details). The objective of this paper is to solve the repair kit problem in a more general setting, where multiple units of each item can be required for a repair, multiple repairs are performed before the repair kit is restocked and no items are taken from the repair kit when a repair is not finished. We show how the problem can be solved for both a service and a cost model. We extend the work of Teunter [6] by firstly introducing an exact formulation for the service level, instead of an approximation. In this exact formulation we assume that items are only taken from the repair kit when a job can be completed. This representation is in correspondence with practice (see Section 6). To our knowledge, we are the first authors to investigate this characteristic instead of having items left behind at the customer as committed inventory in case of an RTF. Secondly, we propose a new solution algorithm that includes this new service level formulation.

In Section 2 we provide an overview of existing approaches and their assumptions that deal with the repair kit problem. We introduce our model and notations in Section 3. In this section we also derive a closed-form expression to calculate the service level. A new approach to solve both the service model and the cost model is required to include this exact expression. The algorithm to solve the repair kit problem in a more general setting is presented in Section 4. The performance of the algorithm is compared to previous results in Section 5. In Section 6 the results of a case study are presented in order to show the performance of the algorithm in real life situations. Finally, we end with our conclusions in Section 7.

2 Development to a General Model

In this section we show the context of our research by giving a short critical overview of papers addressing the repair kit problem, including their assumptions. As mentioned in Section 1, two different types of models have been considered in the literature to determine the contents of a repair kit. The cost model has been introduced by Smith *et al.* [5], while the service model has been introduced by Graves [2]. The main difficulty in formulating both types of models is to find an exact expression for the probability that a return-to-fit (RTF) visit will occur. We define this as the probability that a repairman is not able to finish the repair job due to an insufficient amount of required items in the repair kit. In the cost model, this probability is required to determine the expected RTF costs. In the service model, it expresses the service perceived by the customers. Clearly, a higher probability for an RTF to occur results in lower customer service and higher expected RTF costs. The reciprocal of this probability is referred to as the *job fill rate.* Thus, the job fill rate indicates the fraction of jobs (or orders) performed without a stock out or RTF visit.

The first models that have been developed assume that the technician returns to the warehouse after each job to restock the repair kit. Consequently, each job has the same probability of being completed during its first visit. These problems are called *single period* problems. Another assumption made in these models is that at most one unit of each spare part can be used for the repair. This is referred to as *single item* problems. When both assumptions hold (see, e.g., Smith *et al.* [5], Graves [2], Hausman [3]), at most one unit of each part type is added to the repair kit to obtain optimal contents. So far, all authors assume independence between the different part type failures causing the breakdown. Smith *et al.* [5] solve the problem for the cost model with an optimal marginal analysis procedure. Graves [2] transforms the service model into a knapsack problem.

The cost model of Mamer and Smith [5] considers a single period problem as well. However, they are the first authors to allow more than one unit of each part type to be required for a single job. They also relax the assumption of independence between the failure probabilities by defining a representative collection of job types where each job type corresponds to a set of demands for parts. They formulate the problem as a network problem and solve it with a max flow/min cut algorithm. Based on the network formulation, this technique is, however, only applicable to single period problems.

Heeremans and Gelders [4] are the first to relax the assumption of a single period model. They introduce the notion of a tour into their *multi-period* model. In a *tour* a sequence of jobs is performed before the repair kit is restocked. The number of jobs performed between two restock moments is called the *tour size.* The authors assume it to be fixed, but they do not impose an assumption on the maximum number of units that can be used of any part type to complete the repair (i.e. a *multi-item* model). However, instead of using the definition of job fill rate as service level, Heeremans and Gelders [4] use the fraction of tours performed without an RTF visit (i.e. *tour fill rate*, Teunter [6]). This does not correspond to the service perceived by customers. When tours can be of any fixed size, Teunter [6] gives an exact expression for the job fill rate under the assumption that at most one unit of each part type is used in a repair. Secondly, the author provides an approximation for the job fill rate in a multi-item, multi-period repair kit problem based on the tour fill rate definition. Teunter [6] is the first author to relax the independence assumption of part type failures in a multi-period problem. However, the assumption is made that all required items that are available on stock are always used (or left behind at the customer) despite the fact whether the repair can be finished. This is not in accordance with current practices, in which items are only used when all required items are available to complete the repair. Otherwise, the items that are available can be used in the subsequent repairs of the tour.

In comparison to previous research, we derive a closed-form expression for the job fill rate where items are only taken from the repair kit when a job can be completed (see Section 1). This requires dependency of the availability between the different part types. We derive the expression without making any assumptions on the tour size or on the maximum number of units that can be used during a repair. We model both aspects as stochastic random variables with a general distribution function. This more *general problem setting* is quite common in practice as noticed by Bijvank [1]. The only assumption we make in our definition of a general problem setting is independence between the failures of the different part types. This assumption is of less relevance in a lot of practical settings as demonstrated, for example, by the case study in Section 6. We did not find any significant correlation between the usage of part types within a repair. We also develop an improved solution algorithm in Section 4 to determine the contents of the repair kit based on the cost and service models which incorporate the new expression for the job fill rate in this more general setting.

3 Model Description

In this section, we use the same notations as in Teunter [6]. First, we introduce definitions for the cost model and the service model in Section 3.1. Both models use the same expression for the job fill rate. In Section 3.2, we derive this expression for the repair kit problem in a general setting where items are not kept aside when a job cannot be completed due to lack of required items and no assumptions are made on either the customer demand or the tour size. The notations are listed in Table 1.

3.1 Model Definition and Notation

As explained in Section 1, the repair kit problem deals with the selection of units of different part types that are put in the car of a technician. A repair kit is denoted by S, while n_i represents the number of units of part type i in repair kit S after it is restocked. The number of different part types that are considered to put in the repair kit is denoted by N, so $S = [n_1, \ldots, n_N]$. As explained in the introduction, the objective function in a cost model is to minimise the expected total costs consisting of holding and RTF costs. Calculating the total holding costs is trivial, since each part type i has its own fixed amount of holding cost H_i per tour (i.e., per replenishment cycle). The total holding costs of repair kit S is denoted by $C_H(S) = \sum_i n_i H_i$. As stated in Section 2, the expected total RTF costs is related to the expected job fill rate.

We define the tour size as a stochastic random variable denoted by M with a probability distribution function $P(M = m)$ and average $E[M]$. The maximum number of jobs that can be performed in a tour equals $\overline{M}$. For a given repair kit S and tour size m, the expected job fill rate is given by $\gamma^{job}(S, m)$. The expected job fill rate of a repair kit S equals

$$\gamma^{job}(S) = \sum_{m=1}^{\overline{M}} P(M = m) m \gamma^{job}(S, m) \Big/ E[M]. \tag{1}$$

The expected number of RTF visits equals the total expected number of jobs in a tour minus the expected number of completed jobs in a tour. Hence, the expected RTF costs equal

$$C_{RTF}(S) = P_{RTF} \sum_{m=1}^{\overline{M}} P(M = m) m \left(1 - \gamma^{job}(S, m)\right) = P_{RTF} E[M] \left(1 - \gamma^{job}(S)\right), \tag{2}$$

where P_{RTF} denotes the penalty cost for a return-to-fit visit. In (2), the penalty cost is multiplied with the expected number of RTF visits in a tour with size m and the probability for this to occur. The *cost model* is formulated as

$$\begin{aligned} \text{minimise} \quad & C_{RTF}(S) + C_H(S) \\ \text{subject to} \quad & n_i \geq 0, \end{aligned} \tag{3}$$

Input parameters	
N	number of different part types
H_i	holding cost per unit of part type i per tour
P_{RTF}	penalty cost per return-to-fit visit
β	minimum job fill rate in the service model
$P(M = m)$	probability that the number of jobs in a tour is m (with expectation $E[M]$ and maximum $\overline{M}$)
$p_i^{job}(j)$	probability that j units of part type i are required for a job (with maximum L_i^{max})
Stochastic variables	
M	number of jobs in a tour (tour size)
N_i^m	number of units of part type i available in the repair kit to perform the m-th job
T_i^r	number of units of part type i available in the repair kit to perform r jobs that are completed
U_i^r	number of units of part type i used to complete r jobs
V^m	number of completed jobs out of m jobs
Other notations	
$S = [n_1, \ldots, n_N]$	repair kit with n_i units of part type i
$\gamma(m)$	probability of finishing the m-th job
$\gamma^{job}(S, m)$	job fill rate for repair kit S and tour size m
$\gamma^{job}(S)$	job fill rate for repair kit S
$C_{RTF}(S)$	expected RTF costs for repair kit S
$C_H(S)$	expected holding costs for repair kit S
q_k^i	k-th quantity of part type i to consider in the repair kit
$\Delta_i^{job}(q_k^i, q_{k+1}^i)$	increase of the job fill rate when n_i increases from q_k^i to q_{k+1}^i
S'	repair kit used in the improvement procedure
S''	repair kit after the minimisation procedure
S^*	best found repair kit

Table 1: The notations used in this paper.

and the *service model* as

$$\begin{aligned} \text{minimise} \quad & C_H(S) \\ \text{subject to} \quad & \gamma^{job}(S) \geq \beta \\ & n_i \geq 0. \end{aligned} \tag{4}$$

A general demand process is considered, where $p_i^{job}(j)$ represents the probability of requiring j units of part type i to perform a job. At most L_i^{max} units of part type i are required in one job.

For ease of understanding, we first develop a closed-form expression for $\gamma^{job}(S, m)$ in a multi-period problem setting where $L_i^{max} = 1$ in Section 3.2. We relax this assumption at the second part of that section.

3.2 Job Fill Rate in General Models

Teunter [6] gives a closed-form expression for the expected job fill rate with a fixed tour size $\overline{M}$ (i.e., $P(M = \overline{M}) = 1$) and at most a single unit is used for each item (i.e., $L_i^{max} = 1$). Consequently, $p_i^{job}(0) + p_i^{job}(1) = 1$. For each of the $\overline{M}$ jobs the author calculates the expected probability to successfully repair the machine. The average job fill rate is then found by adding these probabilities and dividing the sum by $\overline{M}$. The expected probability to have enough units available in the m-th job for part type i depends upon the usage of that part type in the previous $m-1$ jobs. At least one unit should be available for each of the required part types after $m-1$ jobs to complete the m-th job. The probability to use l units of a particular part type in $m-1$ repairs equals the probability to replace that part type in l out of the $m-1$ jobs. This latter is true, because at most one unit is used in one job. This probability follows a binomial distribution function.

The assumption of a fixed tour size is relaxed by conditioning on the tour size q ($1 \leq q \leq \overline{M}$). Consequently,

$$\gamma^{job}(S, q) = \frac{1}{q} \sum_{m=1}^{q} \prod_{i=1}^{N} \left\{ (1 - p_i) + p_i \sum_{l=0}^{\min\{n_i - 1, m-1\}} \left[\binom{m-1}{l} p_i^l (1 - p_i)^{m-1-l} \right] \right\}, \tag{5}$$

where $p_i = p_i^{job}(1)$ and $1 - p_i = p_i^{job}(0)$. This equation can be substituted into (1) to find the average job fill rate. Notice that this expression assumes that a required part type is always removed from the repair kit, even if the job cannot be completed due to lack of other required parts. In the remainder of this section, we correct for this and relax the single item assumption.

The binomial distribution of (5) cannot be used anymore when more than one unit of a particular part type can be used in a single job. Therefore, a probability distribution function has to be formulated to express the probability that l units of part type i are available at the beginning of the m-th job. This expression should take the possibility into account that not enough units of a particular part type were available in the repair kit to complete a job before the m-th job, but enough units of the same part type are available to perform the m-th job. Take for instance a situation in which 3 units of part type A are required in the first job, but only 2 units are initially available in the repair kit. This will result in a return visit for this first job. During the second job only 2 units of this part type are required. Since no items are taken from the repair kit at the first job, the second job can be completed. Consequently, a stochastic variable N_i^m is defined as the number of units for part type i that are available in the repair kit to perform the m-th job. An expression for the probability distribution function of this random variable should be derived. This can be done by conditioning on the number of completed jobs V^m out of m jobs and the number of units used during these jobs. Let this latter variable be represented by U_i^r for part type i when r jobs are completed. When k units are used in r jobs that are completed, then $n_i - k$ units are left to perform the m-th job. When $m = 1$,

$$P(N_i^1 = l | V^0 = r) = \begin{cases} 1 & \text{, if } l = n_i \text{ and } r = 0 \\ 0 & \text{, otherwise.} \end{cases}$$

and when $m > 1$,

$$P(N_i^m = l | V^{m-1} = r) = P(U_i^r = n_i - l | T_i^r = n_i), \qquad \text{if } l \leq n_i, r < m,$$

where the probability distribution function of U_i^r depends on the number of items remaining in the repair kit to perform the r completed jobs (denoted by T_i^r). For example, if $L_i^{max} = 3$ and $n_i = 2$, then a job can only be completed if at most two units of item i are demand (or used). Consequently, U_i^r is only defined for 0, 1, and 2. When more than 2 units are demanded, the job cannot be completed and is therefore not included in U_i^r. Therefore, we only consider the conditional probabilities $P(U_i^r = u | T_i^r = j)$ for $u \leq j$ and $j \leq n_i$. Given the fact that a job can only be completed when all required items are available, we know for sure that the number of items demanded is also used in jobs that are completed and not more units are demanded than available (otherwise the job cannot be completed). Consequently,

$$P(U_i^r = u | T_i^r = j) = \begin{cases} \dfrac{p_i^{job}(u)}{\sum\limits_{k=0}^{\min\{j, L_i^{max}\}} p_i^{job}(k)}, & \text{if } r = 1, u \leq j, \\ \sum\limits_{l=0}^{\min\{L_i^{\max}, u\}} \dfrac{p_i^{job}(l)}{\sum\limits_{k=0}^{\min\{j, L_{max}^i\}} p_i^{job}(k)} P(U_i^{r-1} = u - l | T_i^{r-1} = j - l), & \text{if } r > 1, u \leq j, \\ 1, & \text{if } r = 0, u = 0, \\ 0, & \text{otherwise.} \end{cases}$$

We divide by $\sum_k p_i^{job}(k)$ to normalize the distribution function such that $\sum_u P(U_i^r = u | T_i^r = j) = 1$. Next, the probability distribution function for the number of completed jobs out of m jobs has to be specified, which is denoted by V^m. First, let us define $\gamma(m)$ as the probability of completing the m-th job and $\gamma(m|V^{m-1} = r)$ as the probability of completing the m-th job when r $(< m)$ jobs have already been completed. The latter probability depends on the number of units requested for each part type and the availability of these units,

$$\gamma(m|V^{m-1} = r) = \prod_{i=1}^{N} \left\{ \sum_{j=0}^{L_i^{max}} p_i^{job}(j) \left[\sum_{l=j}^{n_i} P(N_i^m = l | V^{m-1} = r) \right] \right\},$$

and $\gamma(m) = \sum\limits_{r<m} \gamma(m|V^{m-1} = r) P(V^{m-1} = r)$. Since the m-th job can either be completed or not,

$$P(V^m = r) = \begin{cases} 1, & \text{if } m = 0, r = 0, \\ 1 - \gamma(1|0), & \text{if } m = 1, r = 0, \\ \gamma(1|0), & \text{if } m = 1, r = 1, \\ P(V^{m-1} = r)[1 - \gamma(m|r)], & \text{if } m > 1, r = 0, \\ P(V^{m-1} = r)[1 - \gamma(m|r)] + P(V^{m-1} = r - 1)\gamma(m|r - 1), & \text{if } m > 1, 0 < r \leq m, \\ 0, & \text{if } r > m. \end{cases}$$

To calculate the job fill rate for a given repair kit S and tour size q, the probabilities to finish each of the q jobs are added and divided by q, similar to (5),

$$\gamma^{job}(S, q) = \frac{1}{q} \sum_{m=1}^{q} \gamma(m). \tag{6}$$

The mathematical formulation of the service model and the cost model is finished when (6) is substituted in (1) and put in (3) and (4), respectively.

4 Algorithm

In this section we develop an algorithm for the service model. The algorithm for the cost model is presented in Appendix A, due to the fact that it consists of the same steps as the algorithm for the service model with only a few minor adjustments. Another reason to focus on the service model is that not much research has been performed on this model in the general setting discussed in Section 2.

From (4), it can be noticed that the service model looks like a knapsack problem. Therefore, a greedy marginal analysis procedure is used in the literature to solve the service model (see, e.g., Graves [2], Teunter [6]). Such a procedure starts with an empty repair kit and adds one unit of a particular part type in each iteration until the predefined service level is satisfied. Determining which part type to add is based upon a ratio which measures the relative increase of the service level (i.e., the job fill rate) in relation to the increase of the holding costs. In previous papers only one unit was added in each iteration. However, when we consider multiple units of the same part type to be used in one job, it is unlikely that adding just one unit is most beneficial in all subsequent iterations. Namely, there are a lot of practical examples in which it is required to replace more than one unit of a part type to fix a job, while replacing only one unit is less likely. In such cases we would like to add more than one unit at a time to the repair kit. As a result, a new algorithm has to be developed to incorporate these possibilities.

The first step of the algorithm consists of the determination of the order of the number of units to add to the repair kit for each part type. In the second step, a greedy procedure similar to Teunter [6] is used to select the parts that are added to the repair kit based on the increase of the service level. The third, and final, step of the algorithm consists of improving the solution of step 2 with an improvement and minimisation procedure. Each step will be discussed in more detail below. The result of this algorithm is a near-optimal contents of the repair kit (see Section 5 for numerical results).

For the first step, we introduce q_k^i as the k-th quantity of part type i to consider in the repair kit, where $q_{k+1}^i > q_k^i$ for all k. The values of q_k^i are set such that the relative increase of the service level (or job fill rate) is decreasing for subsequent values of q_k^i. This is translated in the property formulated in (7).

$$\frac{\Delta_i^{job}(q_k^i, q_{k+1}^i)}{q_{k+1}^i - q_k^i} > \frac{\Delta_i^{job}(q_{k+1}^i, q_{k+2}^i)}{q_{k+2}^i - q_{k+1}^i}, \tag{7}$$

where $\Delta_i^{job}(q_k^i, q_{k+1}^i)$ represents the increase of the job fill rate when the number of units for part type i increases from q_k^i to q_{k+1}^i. So,

$$\Delta_i^{job}(q_k^i, q_{k+1}^i) = \gamma^{job}([n_1, \ldots, n_i = q_{k+1}^i, \ldots, n_N]) - \gamma^{job}([n_1, \ldots, n_i = q_k^i, \ldots, n_N]).$$

The values of q_k^i for a particular part type i can be found with the following pseudo-code:

$n_i = 0$, $q_0^i = 0$, $q_1^i = 1$, $j = 2$, $k = 1$
while $j \leq L_i^{max}\overline{M}$
    while $\frac{\Delta_i^{job}(q_{k-1}^i, q_k^i)}{q_k^i - q_{k-1}^i} \leq \frac{\Delta_i^{job}(q_k^i, j)}{j - q_k^i}$ and $k > 0$
        $k = k - 1$
    end while
    $q_{k+1}^i = j$, $k = k + 1$, $j = j + 1$
end while

This completes the first step of our algorithm. For the second step, we adjust the greedy procedure described at the beginning of this section such that the quantities q_k^i are considered and multiple units of the same part type can be added to the repair kit in one iteration. The pseudo-code for this step is found below.

1 $n_i = 0$ for all $i \in \{1, \ldots, N\}$, $S = [n_1, \ldots, n_N]$ (empty kit)
2 while $\gamma^{job}(S) < \beta$
3     $i^* = \text{argmax}_{\{i | n_i < L_i^{max}\overline{M}\}} \left\{ \frac{\Delta_i^{job}(n_i, q_1^i)}{(q_1^i - n_i) H_i} \right\}$
4     $n_{i^*} = q_1^{i^*}$, $S = [n_1, \ldots, n_{i^*}, \ldots, n_N]$
5     $k = 1$
6   while $q_1^{i^*} < L_i^{max}\overline{M}$ and $q_{k+1}^{i^*} < L_i^{max}\overline{M}$
7         $q_k^{i^*} = q_{k+1}^{i^*}$, $k = k + 1$
8   end while
9     $q_k^{i^*} = q_{k+1}^{i^*} = L_i^{max}\overline{M}$
10 end while

Line 1 represents the initialisation. In line 2 until line 10 items are added to the repair kit until the required job fill rate is met. In line 3 the part type i^* is selected which adds relatively the most to the repair kit (i.e., it has the highest increase of the job fill rate with respect to the increase of the holding costs). Line 4 adds the units $q_1^{i^*}$ of the selected part type i^* to the repair kit S. In line 5 until line 9 the ordering of $q_k^{i^*}$ is shifted one position, such that $q_1^{i^*}$ represents the next quantity to consider for part type i^*.

This greedy procedure immediately stops when the job fill rate is met. Even though the contents of the repair kit satisfies the service level constraint after performing step 2 of the algorithm, the total holding costs could be reduced when the last iteration is performed in a smarter way. This is the objective of step 3 in our algorithm. The improvement procedure starts with removing the units which were added to the repair kit S in the last iteration, resulting in repair kit S'. In order to satisfy the service level constraint, items have to be added to S' with the extra constraint $C_H(S') < C_H(S)$ to guarantee a better solution. The same greedy procedure of step 2 can be used to investigate whether a solution S' exists which satisfies the job fill rate criterion with lower holding costs. In the previous pseude-code, S has to be replaced by S' and line 3 of the pseudo-code should be replaced by

$$3 \qquad i^* = \operatorname{argmax}_{\left\{ i \,\middle|\, \substack{n_i < L_i^{\max}\overline{M}, \\ C_H(S') + (q_1^i - n_i)H_i < C_H(S)} \right\}} \left\{ \frac{\Delta_i^{job}(n_i, q_1^i)}{(q_1^i - n_i)H_i} \right\}.$$

If such a solution S' exists, it can be investigated for further improvements by setting S to S' and repeating the improvement procedure until no new and better solution is found.

Besides improving S we can also check whether units of the current solution S can be removed without replacing them with other parts to reduce the holding costs and still satisfy the job fill rate criterion. This procedure is referred to as the minimisation of S. A backtracking procedure is used to check whether the service level is still sufficient when one unit of the last added part type is removed and the one before, and so forth. The repair kit resulting from this minimisation procedure is denoted by S''. The overall best solution S^* is found by performing the different procedures in the order shown in Figure 1.

5 Results

In this section the performance of the algorithm described in Section 4 is tested by means of three test cases. Teunter [6] considers two kinds of test cases: small instances and large instances. In the definition of small instances at most 8 different part types are used and the maximum tour size is set to 4. For large instances at most 100 different part types are

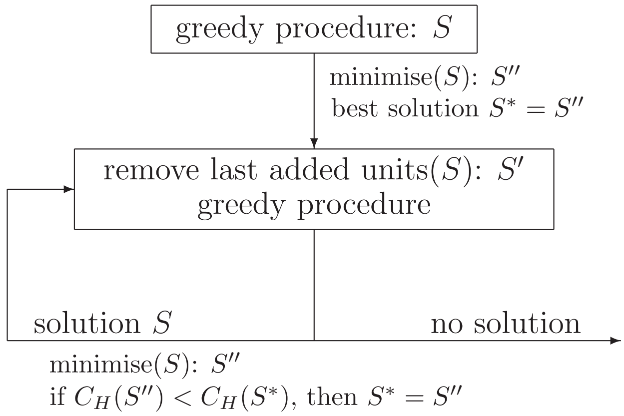

Figure 1: The structure of the solution procedure.

considered and the maximum tour size equals 12. As a third case, an additional setting is added which is more representative for reality. In this third setting the number of different part types ranges between 500 and 1,000. The test instances are drawn from (discrete) uniform distributions. In Appendix B we describe in more detail the specific distributions that are used for the different parameters to randomly generate 1,000 examples for each test case. The results of all three cases are discussed in this section.

In our analysis we compare the repair kits obtained by our algorithm with the repair kits resulting from the algorithm of Teunter [6]. We have tested two aspects of our algorithm: (1) the improvement and minimisation procedure (step 3 of our algorithm) and (2) the greedy procedure for the exact, closed-form expression to calculate the job fill rate. To test the first aspect, the contents of a repair kit is determined according to Teunter [6] and then our improvement and minimisation procedure of Section 4 modifies this solution. The relative reduction of the total holding costs for the different instances is shown in the first row of Table 2. In the second row the relative reduction of the holding costs is presented when the entire algorithm of Section 4 is used (including the exact formula for the job fill rate) and compared to the outcome of Teunter's [6] algorithm. The third row presents the relative deviation of the solution found with our algorithm compared to the optimal solution, which is found by enumeration. Optimal solutions can only be found for the small instances due to the complexity of the problem.

Based on the results shown in Table 2, we conclude that the improvement and minimisation procedure decreases the holding costs on average by almost 5% for the small instances. However, with our closed-form expression for the service level we even find a decrease of

	small		large		representative	
	average	standard deviation	average	standard deviation	average	standard deviation
approx. JFR + improvements	4.68%	8.43%	0.45%	0.55%	1.05%	1.41%
exact JFR + improvements	5.83%	9.16%	1.30%	0.76%	14.18%	4.30%
deviation from optimality	0.25%	1.27%	-	-	-	-

Table 2: Reduction of the total holding costs over 1,000 instances for each test setting, when the solution procedure of Teunter [6] is complemented with our improvement and minimisation step and when it is compared to our procedure which incorporates the exact job fill rate (JFR). The final row shows the deviation of the results found with our solution procedure from optimality.

the holding costs by 5.8% when the service constraint is satisfied for the small instances. This corresponds to an average deviation of 0.25% from the optimal solution. For the large instances the improvements are less significant. The results for the representative instances show the most significant cost reductions. The reason that the representative scenario benefits the most from the exact job fill rate expression is because of the different principles behind the two algorithms. The algorithm of Teunter [6] adds units to the repair kit based on the potential of each part type to increase the service level, contrary to our algorithm which adds units that immediately contribute (relatively) the most to the repair kit. In the representative scenario, the repair kit only contains at most one unit for most of the part types. The potential for each part type is, however, determined based on the contribution of adding more than one unit of that part type to the repair kit. Consequently, this potential is not always realised and other part types are selected in the next iterations of the algorithm. The average number of units per part type in the repair kit is much larger for the small and large instances. Therefore, the potential is a better representation of the actual contribution of the part types in these two scenarios. This is also the reason why the improvement and minimisation procedure of step 3 in our algorithm does not show big improvements in the representative instances.

Table 3 shows a number of statistics for the different scenarios. The first two rows show the size of the repair kit and the average number of units per part type in the repair kit.

		small instances	large instances	representative instances
average size of repair kit		9.27	290.26	397.47
average value of n_i		1.81	4.86	0.52
frequencies	approx. JFR + improvements is best	4.0%	0%	0.0%
	exact JFR + improvements is best	19.4%	97.4%	93.6%
	same solution	76.7%	2.6%	6.4%
	exact JFR + improvements is optimal	89.3%	-	-
p-values		<1E-06	<1E-06	<1E-06

Table 3: Several statistics about the solutions for the service model in the different test settings.

The *size* of a repair kit is defined as the number of units in the repair kit (i.e., $\sum_i n_i$). The results for the representative setting show repair kits with the largest size, but these repair kits also contain the most different part types. Consequently, the repair kits of the representative instances contain on average 0.52 units of each part type. Figure 2 also shows this relationship where improvements are more significant when the average number of units per part type (i.e., $\sum_i n_i/N$) is small.

Table 3 also presents the frequencies that the exact formula for the job fill rate results in a better solution in comparison to the approximation procedure of Teunter [6]. The best solution is found by the approximate job fill rate procedure in 4.0% of the small instances, while the exact formulation for the job fill rate finds the best solution in 19.4% of the instances. In the remaining 76.7% of the small instances, both methods result in the same solution. Notice that we included the minimisation and improvement procedure in Teunter's [6] algorithm to obtain these results. Otherwise, the approximation procedure of Teunter would never have resulted in a better solution. Table 3 also shows that our algorithm with the exact job fill rate finds the optimal solution in 89.3% of the small instances. For the large and more representative cases the best found solution is almost always found with the exact job fill rate. Therefore, we conclude that the algorithm with the exact expression for the job fill rate, as formulated in Section 3, significantly outperforms the algorithm with the approximation of Teunter [6]. This can also be concluded when a Wilcoxon test is performed in which the null

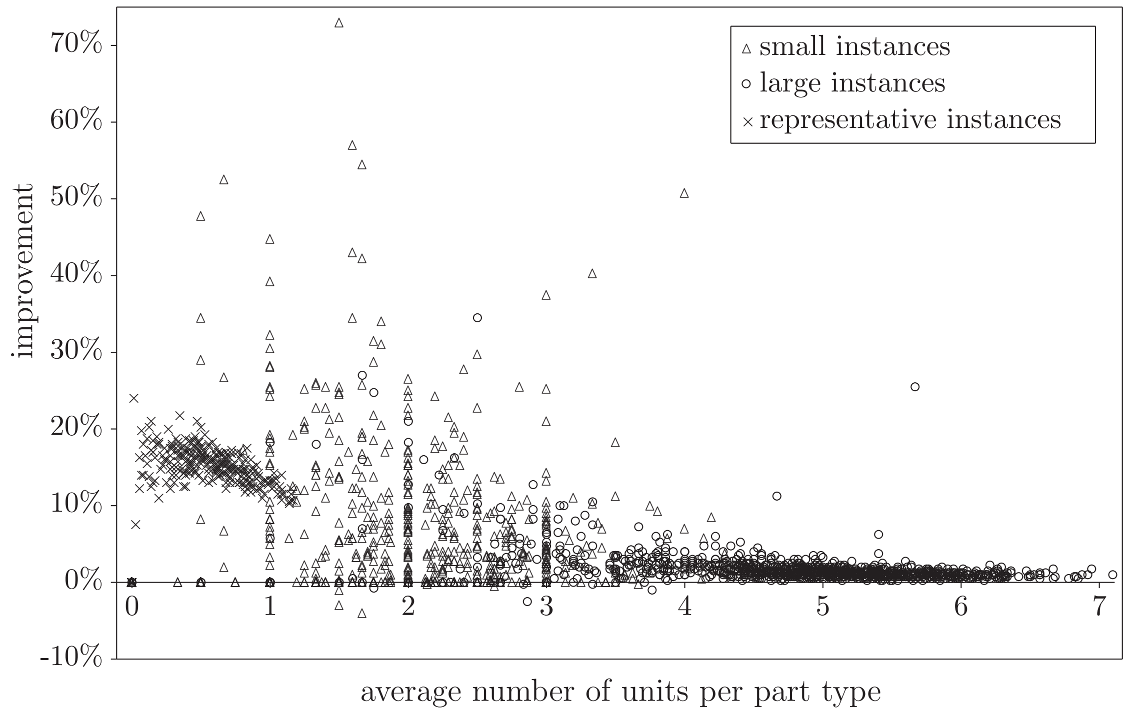

Figure 2: The relative improvement plotted against the average number of units in the repair kit per part type for the different test settings.

hypothesis specifies that Teunter's [6] algorithm performs better. Based on the p-values[2] shown in Table 3 we reject the null hypothesis and conclude that our algorithm performs significantly better.

The results for the cost model are discussed in Appendix A by performing the same set of experiments as described here. In the next section the performance of our algorithm is tested in a practical setting.

6 Case study

Besides the test instances of Section 5, we performed a case study to get a better feeling for the performance of the algorithm in practice. In this case study we solve the service model and the cost model, but we also perform a sensitivity analysis on the service level. This is important since the management of repair service companies wants to know the impact of a

[2]The p-value of a test refers to the probability of wrongly rejecting the null hypothesis if it is in fact true. Small p-values suggest that the null hypothesis is unlikely to be true.

particular service level criterion on the holding costs and the size of the repair kit.

This case study is based on real data from Ricoh Europe. Ricoh is a leading global manufacturer of office automation equipment. They offer products for businesses and for personal use. Ricoh performs the after-sales service to the customers as well. In this case study, we looked at multi-functional systems (combined copier/printer/fax/etc.). Ricoh Europe, located in Amstelveen, is the regional headquarter of Europe, Africa, and the Middle-East. Currently Ricoh Europe has subsidiaries and branches in fourteen countries and factories in France and the United Kingdom. Ricoh Netherlands is one of the subsidiaries. Ricoh Europe has about fifteen thousand distinct types of service parts for about three hundred different multi-functional systems. Ricoh Netherlands has more than 35 technicians driving around with a stock value of almost 6,000 Euros each. Ricoh charges RTF cost of 45 Euros if a repair cannot be performed in the first visit.

When we analyse the contents of the current repair kits used by the technicians we see rather low service levels of 53%. Therefore, the expected total return-to-fit costs are quite high. An overview of the current situation is shown in Table 4 as well as the results for applying the cost model and the service model and the associated algorithms.

The solution of the cost model shows an increase of the holding costs by 250%. Despite the fact that the total costs reduce significantly, this solution is undesirable for Ricoh because of high risk of theft. However, a solution with less holding costs and an improved service level can be found with the service model. Table 4 gives an overview on the costs for different values of the service level. Based on these results it is possible to increase the service level by 31% against current holding costs.

In Figure 3 we consider the relationship between the total holding costs and the service level. It shows a rapid increase of the service level when the size of the repair kit is small. This concave relationship is what is to be expected based on the property expressed in (7).

Figure 3 also shows the different costs when units are added to the repair kit. It shows a clear trade-off between the holding costs and RTF costs. Based on the results of this case study we conclude that our closed-form expression for the order fill rate and our algorithm work well in practice. It can help a company decide which parts to put in the repair kit, but it can also help them to analyse their current stock levels.

	job fill rate	holding costs	RTF costs	total costs
current contents	53.14%	1.26	22.23	23.48
cost model	95.99%	3.18	1.91	5.09
service model	84%	1.23	7.59	8.82
	85%	1.30	7.11	8.42
	86%	1.38	6.63	8.01
	87%	1.48	6.09	7.57
	88%	1.56	5.69	7.25
	89%	1.67	5.21	6.88
	90%	1.80	4.72	6.52
	91%	1.94	4.26	6.20
	92%	2.10	3.78	5.88
	93%	2.29	3.31	5.61
	94%	2.53	2.82	5.35
	95%	2.80	2.36	5.16
	96%	3.19	1.90	5.09
	97%	3.75	1.42	5.18
	98%	4.54	0.94	5.49
	99%	5.88	0.47	6.35

Table 4: Results for the current contents of the repair kit used in this case study, as well as the results for the cost model and the service model.

7 Conclusions

Customer-oriented markets become more and more important and, therefore, after-sales services as well. One particular service is a repair service on location, in which a customer is only satisfied when a repair is completed. This means that a technician should have enough spare parts taken along to the customer. If one of the required parts is missing, the

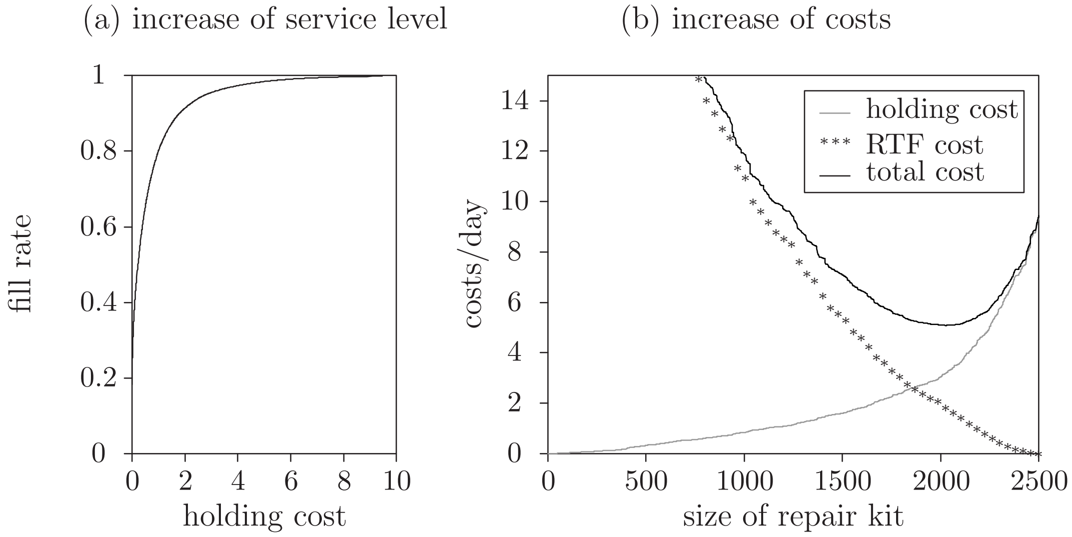

Figure 3: The results for the case study: (a) the increase of the service level when the holding costs increase, (b) the different costs when the number of items in the repair kit increases.

technician has to return later and none of the required parts that are available are taken out of the repair kit. This latter characteristic of the repair kit problem is not dealt with in previous literature. In this paper, we derived an exact, closed-form expression for the service level in a general setting where multiple units of each part type can be used in a job and multiple jobs are performed before the car is restocked. We also developed two algorithms to solve the service and cost model which incorporate this exact, closed-form expression for the job fill rate. The algorithm for the service model consists of three different steps: (1) finding an order for the quantities of a particular part type to be added to the repair kit, (2) getting a reasonable good solution for the service model with the use of a greedy, marginal analysis heuristic, and (3) improving the initial solution according to an improvement and minimisation procedure. The algorithm for the cost model is almost similar except for some minor modifications.

Based on test instances we concluded that our algorithm performs significantly better compared to other existing algorithms in the literature. Especially when only a few units are required in the repair kit for each part type. It outperforms the other algorithms in almost all examples and it finds solutions within a range of 0.2% from the optimal solution. We have also tested the applicability of the algorithm in practice. Based on a case study we have shown an increase of the service level by more than 30% without an increase in the holding costs. This shows that our algorithm and the closed-form expression can be used to find the near-optimal contents of repair kits in practice.

References

[1] M. Bijvank. Car stock management - the need to close the gap between theory and practice. In D. Kisperska-Moon, editor, *Tenth ELA Doctorate Workshop*, 2005.

[2] S.C. Graves. A multiple-item inventory model with a job completion criterion. *Management Science*, 28(11):1334–1337, 1982.

[3] W.H. Hausman. On optimal repair kits under a job completion criterion. *Management Science*, 28(11):1350–1351, 1982.

[4] D. Heeremans and L.F. Gelders. Multiple period repair kit problem with a job completion criterion: A case study. *European Journal of Operational Research*, 81(2):239–248, 1995.

[5] S.A. Smith, J.C. Chambers, and E. Shlifer. Optimal inventories based on job completion rate for repairs requiring multiple items. *Management Science*, 26(8):849–854, 1980.

[6] R.H. Teunter. The multiple-job repair kit problem. *European Journal of Operational Research*, 175(2):1103–1116, 2006.

A Cost Model

The cost model is defined by (3). We use the algorithm for the service model (see Section 4) to solve the cost model in a general setting. Step 1 of the algorithm is similar for both models. However, since the cost model does not have any restrictions upon the service level, the algorithm for the cost model needs a different stopping criterion. It also needs to keep track of the solution with the lowest expected total costs. The cost model does not need any improvement steps, since the algorithm does not stop immediately. Therefore, step 3 of the algorithm for the service model is removed for the cost model.

Only step 2 of the algorithm for the service model has to be adapted for the cost model. When we denote the solution with the lowest total costs by S^*, the pseudo-code for step 2 of the algorithm for the cost model is given below.

```
1   n_i = 0 for all i ∈ {1, ..., N}, S = [n_1, ..., n_N] (empty kit), S* = S
2   while C_H(S) < C_T(S*)
3      i* = argmax_{i: n_i < L_i^max} M̄  Δ_i^job(n_i, q_1^i) / ((q_1^i − n_i) H_i)
4      n_{i*} = q_1^{i*}, S = [n_1, ..., n_{i*}, ..., n_N]
5      if C_T(S) < C_T(S*) then
6         S* = S
7      end if
8      k = 1
9      while q_1^{i*} < L_i^max M̄ and q_{k+1}^{i*} < L_i^max M̄
10        q_k^{i*} = q_{k+1}^{i*}, k = k + 1
11     end while
12     q_k^{i*} = q_{k+1}^{i*} = L_i^max M̄
13  end while
```

The new stopping criterion in line 2 is to stop adding items when the holding costs are higher than (or equal to) the expected total costs of the best found solution so far, where $C_T(S^*) = C_H(S^*) + C_{RTF}(S^*)$.

The performance for the cost model is tested with the same set of experiments as described for the service model: small, large, and representative instances (see Section 5 and Appendix B). Since there is no improvement and minimisation procedure in the algorithm, we only compared the outcome of our algorithm to the outcome of the algorithm developed by Teunter [6]. Table 5 shows the relative savings on the expected total costs for all three test cases. This table also shows that there is hardly any deviation from the optimal solution. Table 5 also shows the frequencies how many times our algorithm results in a better solution compared to the algorithm of Teunter [6].

B Settings Sample Instances

The test instances are drawn from (discrete) uniform distributions. Table 6 shows the specific distributions that are used for the different parameters in the different test settings.

We remark that $P(M = m) > 0$ for $\overline{M} - 2 \leq m \leq \overline{M}$, $\overline{M} - 9 \leq m \leq \overline{M}$ and $\overline{M} - 1 \leq m \leq$

		small instances	large instances	representative instances
improvement	average	0.41%	0.50%	2.37%
	standard deviation	1.04%	0.52%	0.85%
deviation from optimality	average	0.00%	-	-
	standard deviation	0.00%	-	-
frequencies	approx. JFR is best	0%	0%	0%
	exact JFR is best	29.7%	87.9%	100%
	same solution	70.3%	12.1%	0%
	exact JFR is optimal	97.8%	-	-
average size of repair kit		10.06	287.20	280.83
average value of n_i		2.30	5.71	0.38

Table 5: The results for the cost model.

	small instances	large instances	representative instances
N	discrete uniform[1,8]	discrete uniform[1,100]	discrete uniform[500,1000]
L_i^{max}	uniform[1,4]	uniform[1,4]	uniform[1,3]
$p_i^{job}(j)$	uniform[0,0.2/L_i^{max}]	uniform[0,0.2/L_i^{max}]	uniform[0,0.0005/L_i^{max}]
H_i	uniform[0,0.35]	uniform[0,0.35]	uniform[0,0.05]
$\overline{M}$	discrete uniform[3,6]	discrete uniform[10,12]	discrete uniform[2,3]
$P(M=m)$	uniform[0,1/3]	uniform[0,1/10]	uniform[0,1/2]
β	uniform[85%,95%]	uniform[85%,95%]	uniform[85%,95%]
P	uniform[0,10]	uniform[0,100]	uniform[40,80]

Table 6: The distributions for the parameters used to generate the instances for the different test cases.

$\overline{M}$, respectively, for each test case. To ensure that $\sum_m P(M=m)=1$ we put the remaining probability mass on the middle tour size. Also notice that $p_i^{job}(0)=1-\sum_j p_i^{job}(j)$.