Integrated service engineers and spare parts planning in the maintenance Logistics

A. Al Hanbali, and W.H.M. Zijm
Department of Industrial Engineering and Business Information Systems (IEBIS), Faculty of Behavioral, Management, and Social Sciences, University of Twente, The Netherlands
A. Sleptchenko
Department of Mechanical and Industrial Engineering, College of Engineering, Qatar University, Qatar
Corresponding author: a.alhanbali@utwente.nl

Abstract: We analyze the integrated tactical capacity planning of spare parts supply and workforce allocation in maintenance logistics of advanced equipment. The equipment time-to-failure, spare parts replenishment time, and equipment repair time are random and independent of each other.

Keywords: Spare parts; Service Engineers; Discrete Stochastic Optimization;

INTRODUCTION

As a result of technological progress, equipment used in a number of important economic sectors such as discrete manufacturing, chemical production, transport, defense, and health care, become more and more capital intensive. For that reason, but also for safety and security reasons, the uninterrupted processing of these equipment is essential. If an unplanned downtime due to failure occurs, it is of utmost importance to keep it as short as possible. Best thing of course is to avoid unplanned downtime using preventive maintenance, but unfortunately that is not always possible due to reasons that will be explained in the following. In that case, everything should be done to get the system up and running again as soon as possible. In order to do so, failing parts are often replaced by ready-for-use ones, since repair of the complete system on site requires too much time. Now, it is the availability of ready-for-use spare parts that counts, as well as the waiting time for service engineers and tooling needed. To minimize any delay due to absence of these resources, it is of utmost importance that the latter are readily made available. This leads to our key and complex integrated multi-resource planning problem. The added complexities are due to the definition of the right KPIs (key performance indicators) and the differentiation between various spare parts characteristics.

Despite the overwhelming literature on maintenance and service, the approaches are still fragmented. The researchers either concentrate on server availability as in the classical machine repairmen problem or on spare parts availability, and with limited focus on tools, whereas it is clear that any integrated, or complete, solution encompasses all three resources. However, to that end one needs models that have not been developed so far, although we may borrow from some partial approaches in other fields. Capacitated resources and uncertainty due to unplanned downtime are two inherent characteristics that are central to our key problem.

Due to the primary importance of advanced equipment to the system owners (customers in the sequel), performance-based service contracts were introduced. Such a contract is an agreement between the customer and the manufacturer that contains precise information on the target service levels that the system should satisfy in addition to the involved costs. The most commonly used Service Level Agreements (SLAs) are based on targets on average availability of equipment and maximum time to fix a failure. Different types of contracts can be classified into different classes, e.g., platinum, gold, and silver, see (Cohen et al. 2006). It is clear that the contract value increases with the tightness of the targets specified. For example, 99\% equipment availability or 2 hours maximum time to fix are familiar targets used in the semiconductor and the aircraft industries. Moreover, penalties are paid by the manufacturer if SLAs are violated. Therefore, it is essential that the manufacturer differentiates its services offered to the different classes of customers. This can be done by assigning a higher priority to tighter SLAs. For the manufacturer it is necessary to avoid penalties that damage its reputation and lead to a loss of future businesses. Therefore, the manufacturer should concentrate on improving its products quality and reliability from the early design phase. This can be done by monitoring the performance of previous equipment in order to feed back this information to subsequent designs. These actions must be complemented in the operational phase by efficient and effective maintenance planning and operations.

Maintenance tasks are divided into two categories: preventive and corrective, see (Ebeling 2004). It is well known that preventive maintenance fails in eliminating all possible equipment failures. This is due to two facts. First, for some components, e.g., electronic components, it is hard to predict a future failure. This is because their wear-out process is weakly correlated with the usage. Moreover, when a failure occurs it is not linked to a specific operation time. Second, when failures are predictable, e.g., failures of mechanical components, there are inherent errors in the statistical and physical models used. Advanced equipment are for instance mechatronic systems in which components are electronic, mechanical, or hybrid. In this case of unavoidable failures, corrective maintenance comes to the picture to keep up the equipment, usually by replacing the failed parts by ready-for-use ones.

In this paper, we shall assume that the service tools are not extremely expensive which makes it economical for a service engineer to own the service tools needed for the corrective maintenance tasks. We will focus on the following service strategy inspired by common practices in the after-sales maintenance logistics of advanced equipment. Upon arrival of a corrective maintenance request if any of the required resources (e.g., spare parts or engineers) is not readily available the request will be satisfied via an emergency channel with an ample supply of resources. The emergency channel has much shorter lead time with a much higher costs as compared to the regular replenishment channel. The objective of the service provider is to minimize the total costs of spare parts, service engineers pool, and of emergency. This will be done under Poisson random demand of maintenance requests, and both an exponentially distributed repair time and spare part's replenishment time.

Our key contributions in this paper are as follow:

- We formulate a Mixed-Integer Linear Program (MILP) for the exact optimisation of the integrated spare parts and service engineers capacity planning.
- We propose a computational efficient and accurate optimisation heuristics with a less than 1% relative errors compared to MILP results and with a guaranteed convergence.
- We show using simulation the insensitivity of the emergency cost to the repair time of equipment and the replenishment time of spare parts.

LITERATURE REVIEW

A quick and effective equipment repair means that upon a customer call for repair the manufacturer determines the required part(s) to replace, assigns a priority to the call according to the service contract, assigns one or more qualified engineers, and plans the necessary parts and tools. The integrated multi-resource planning problem of spare parts, engineers and tools, as we will show, has not been considered so far in the literature despite the overwhelming studies on the planning of these three resources individually.

In spare parts management, the main objective is to meet the target service levels specified in the contracts at minimum costs. Note that, the value of a single part may amount to several hundred thousand dollars in advanced equipment. An important feature that is relevant when dealing with this problem: the spare parts and the stocking locations are both hierarchically structured. Spare parts can be broken down into modules, sub-modules, and piece parts, each with a different cost and a specific time to replace. Repairing the system by replacing a failed piece part is much cheaper than replacing a complete module. However, this usually comes with a significantly longer time to repair(replace). Hence, there is a tradeoff between the value of a spare part and the time to repair an equipment. Similarly, there is a tradeoff between the costs involved in stocking parts very close to the customers' sites (often called bases) and centrally (often called depot). A central depot can support multiple customers at different locations. Moreover, due to the reduction of risk with the pooling of demands of different customers it is desirable to position a number of the stocks centrally. However, having a strict SLA may force a manufacturer to move some spare parts closer to the bases. Sherbrooke (1968) was among the first to tackle the spare parts management problem. He proposed the quantitative METRIC model that considers these tradeoffs, and came up with close-to-optimal decisions on what and how many spare parts (modules, submodules, or piece parts) to keep in each location. Several extensions and new model features of the basic METRIC model are explained in (Sherbrooke 2004) and (Muckstadt 2005).

The manpower planning problem focuses on the number of service engineers that should be hired for each service region in a network. In the literature, the manpower planning problem for field maintenance services is studied in (Agnihothri and Karmarkar 1992) and (Agnihothri et al. 2003). The author highlights that a service territory size in which the workload can be managed with one service engineer provides a major advantage of a good relationship between the customer and the service engineer.

Tools and service engineers as resources share few characteristics with the spare parts resource. For example, tools are usually demanded in sets and they are not consumed. This means, after a repair activity they become available for possible future usage. However, spare parts can be either consumable or repairable. This means, they have to be first replenished or repaired (usually this is done in a different location) before they become available for a future use. Therefore, the planning of service tools and engineers needs different approaches from those used for spare parts. Few studies in the literature are dedicated to the planning of tools, (Vliegen 2010).

To the best of our knowledge, the integration of the all key resources for equipment maintenance has not been found in the literature, although it is essential to tackle our problem. Similar problems arise in other fields, for example in flexible manufacturing, and multi-resource project scheduling. The main difference with the problem studied in this paper is that we deal with optimal decisions under uncertainty. Another field where the multi-resource capacity planning is considered under demand and supply uncertainty is the Assemble-To-Order (ATO) production systems, see, (Song et al. 1999) and the reference therein. Song et al. (1999) considered a multi-product multi-component assembly system with a Poisson demand of products and a stochastic replenishment time of components.

MODEL DESCRIPTION

In this paper, we consider a single-site multi-item problem. The single-site is responsible for a specific service region. The arrival of requests for repairs is according to a Poisson process with an arrival rate λ. A request for a repair consists of a simultaneous demand for two resources, namely, a service engineer and a single ready-for-use item. Upon a request arrival if any of these two resources is not available the costumer's request is satisfied via an emergency channel with a high cost. We assume that this emergency channel has an ample supply of both parts and engineers and will not affect further analysis. In Rahimi-Gharoodi et al., we consider the service strategy with emergency shipment for (only) spare parts.

The objective is to minimize the total costs of service engineers, service parts, and emergency costs (related to the loss probability of requests). We assume that there are in total N different (types of) service parts. The probability that a part of type-*i* is requested is equal to $r_i, i = 1, \dots, N$, with $\sum_{i=1}^{N} r_i = 1$. The inventory of type-*i* parts, $i = 1, \dots, N$, is managed according to the base-stock policy, referred to as $(S_i - 1, S_i)$. The stock replenishment time of a type-i part is an exponentially distributed random variable with a rate μ_i. Finally, a mission of an engineer to a customer takes in total an exponentially distributed time with a rate γ and is independent of other engineers. The mission duration includes the time to go from the site to the customer's location and backward in addition to the system's repair time. The team size of service engineers in the region under consideration is equal to E.

Let $n_e(t)$ denote by the number of engineers in the field on missions at time *t*. Let $n_i^s(t)$, $i = 1, \dots, N$, denote the number of type-*i* parts on-order to replenish the stock at *t*, i.e., the pipeline size of type-*i* part at *t*. Under the above assumption, the joint process $M_t := \{\left(n^e(t), n_1^s(t), \dots, n_N^s(t)\right);\ t > 0\}$ is a continuous-time finite-state Markov chain with a state space $\Omega = \{0, \dots, E\} \times \{0, \dots, S_1\} \times \cdots \times \{0, \dots, S_N\}$. In the following, we shall denote by $(l; \boldsymbol{s})$, where $\boldsymbol{s} = (s_1, \dots, s_N)$, an element of Ω. If the Markov chain is in state $(l; \boldsymbol{s})$ it means that there are l busy engineers and s_i parts of type-*i* are on-order (stock replenishment orders). In Figure 1, we show the transition rate diagram of M_t in the single item case. The Kolmogorov forward balance equations of the Markov chain $M_t$in steady state read, $\forall (l; \boldsymbol{s})$,

$$\left(\lambda \sum_{i=1}^{N} r_i + l\gamma + \sum_{i=1}^{N} s_i\mu_i \right)p_{(l;\boldsymbol{s})} = \lambda \sum_{\{i=1\}}^{N} r_i p_{(l-1;\,\boldsymbol{s}-\,e_i)} + (l+1)\ \gamma\ p_{(l+1;\,\boldsymbol{s})} + \sum_{\{i=1\}}^{N} (s_i+1)\mu_i\ p_{(l;\,\boldsymbol{s}+\boldsymbol{e_i})}\,, \quad (1)$$

where e_i is thw unit vector with i-th entry equal to one and the rest equal to zero, and $p_{(l;\boldsymbol{s})} = 0 = 0$ if $(l;\boldsymbol{s}) \notin \Omega$. The balance equations together with the normalization condition $\sum_{\forall(p;\boldsymbol{s})\in\Omega} p_{(l;\boldsymbol{s})} = 1$ gives the steady state probabilities of the Markov chain.

In the literature, it is common that a performance analysis of the system is first carried out to find a closed-form result of the key indicators, e.g., in this paper the loss probability is the main performance indicator. When it is hard to find a closed-form result the problem is solved numerically using one of the standard algorithms. By rearranging the states appropriately, the transition generator matrix of the Markov chain can be seen as a three-diagonal level-dependent finite-state quasi-birth process. The steady-state probabilities can be computed numerically using a standard algorithm. These numerical results help in exploring the behavior of the performance indicators as a function of the system parameters which facilitate the design of an exact/approximate procedures devised for the performance optimization. Instead of following the traditional approach, we shall try to directly build an exact optimization model with the objective of minimizing the total costs that include spare parts, engineers, and emergency costs under the constraints that the balance equations of the Markov chain should be satisfied.

OPTIMISATION

From the perspective of a service provider it is necessary to insure a high offered service to the customer at a minimal total cost. There is a trade-off between the cost of holding the resources and the emergency cost. The higher the former cost component (i.e., more resources are available upon request) the lower the latter and vice versa. Our objective is to find the optimal balance between all these cost components. Let us first introduce the cost parameters:

- c_e cost of hiring a service engineer per time unit, e.g., hourly wage.
- c_i^s ost of holding a part of type-i per time unit, this cost may include cost of capital, storage and risk, including the obsolescence risk.
- c_i^{loss} emergency of a request of type-i.

The emergency cost of a type-i request per time unit is the product of c_i^{loss} and the type-i request (emergency) loss rate. The latter is equal to λr_i, the arrival rate of type-i requests, multiplied by its (emergency) loss probability. To find the optimal number of engineers (E) and items on stock $(S_1, \dots, S_N)$ that minimize the total costs, we will introduce the following decision variables:

- I_l^e ($I_{i,l}^s$) binary variable that is equal to one if the system has l active engineers (l parts of type-i) and zero otherwise.

- $y^i_{l;\boldsymbol{s}}$ the probability that the Markov chain is in state $(l;\boldsymbol{s})$ and a new request of type-i can be still accepted. This happens when $l < E$ and $s_i < S_i$. That is $y^i_{l;\boldsymbol{s}}$ is either equal to $p_{l;\boldsymbol{s}}$ when the system has both an available-for-mission engineer and an on-shelf part of type-*i* or it is zero otherwise.

As a function of the decision variables and the costs parameters, the objective function reads:

$$min \sum_{l=1}^{\bar{E}} c^e I_e^l + \sum_{i=1}^{N} \sum_{s_i=1}^{\bar{S}_i} c_i^s I_{i,s_i}^s + \sum_{i=1}^{N} c_i^{loss} \lambda r_i \sum_{\forall(l;\boldsymbol{s})} \left(p_{l;\boldsymbol{s}} - y^i_{l;\boldsymbol{s}}\right),$$

where $\bar{E}$ and $\bar{S}_i$ are the maximum allowed numbers of engineers and service parts in the system (truncation parameters). Note, as indicated above, the $y^i_{l;\boldsymbol{s}}$ variable is equal to the $p_{l;\boldsymbol{s}}$ only when we have an available engineer and on-shelf part of type-i. Therefore, the last term of the objective function ($\sum_{\forall(l;\boldsymbol{s})}\left(p_{l;\boldsymbol{s}} - y^i_{l;\boldsymbol{s}}\right)$) gives us the loss probability of type-i requests.

The costs minimization is subject to the following constraints. The balance equation of the Markov chain and the normalization condition should be satisfied. Constraint to ensure that if the number of engineers is not equal to *l* then the probability $p_{l;\boldsymbol{s}} = 0$. Constraint to ensure that if $I_l^e = 0$ then all $I_{l+1}^e, I_{l+2}^e, \dots, I_{\bar{E}}^e = 0$. Similarly, there are a number of constraints on (I_{i,s_i}^s). The last constraints ensure that the variables $y^i_{l,\boldsymbol{s}}$ is either equal to $p_{l;\boldsymbol{s}}$ when the system has an available engineer and a part of type-i or it is 0 otherwise. The obtained Mixed-Integer Linear Problem (MILP) has $(N+1) \times (\bar{E}+1) \times (\bar{S}_1+1) \times \cdots \times (\bar{S}_N+1)$ continuous variables (*p*'s and *y*'s) and $\bar{E} + \bar{S}_1 + \cdots + \bar{S}_N + N + 1$ binary variables (in Al Hanbali et al. (2016) we give the full problem formulation). It is easy to see that the size of the problem grows exponentially with the number of parts in the system N and with the truncation limits $\bar{E}$ and $\bar{S}_i$. Application of the proposed MILP to real-life problems is impossible. Therefore, in the following we shall propose a fast approximate method to evaluate the loss probability and a fast optimization heuristic based on local search algorithm to find the near-optimal number of engineers and spare parts in the system.

OPTIMISATION HEURISTICS

In this section, we introduce two heuristics to reduce the computation time of the MILP for large problem size. In the first heuristic, we propose a local search procedure while computing the loss probability in an exact way using a numerical algorithm to solve the balance equations described. In second heuristics, we propose to use the same local search procedure while computing the loss probability in an approximate way.

Local Search with Exact Loss Probability

Let us refer to the loss probability of type-i requests as $p^{loss}_{i;E;\boldsymbol{S}}$ when there are E engineers and $\boldsymbol{S} := (S_1, \dots, S_N)$ parts in the system. We propose a local search based heuristic, which at every point in time we decide to add or remove a unit of the resources that leads to the highest total cost reduction. Note that, the loss probability can be evaluated by solving the balance equations (1) and summing up the state probabilities for e=E or $s_i = S_i$. The exact evaluation of the loss probability is time consuming since the state space grows really fast with N and E. For that reason, in the following we propose an efficient approximate method to evaluate the loss probability. In order to prevent the deadlock of the search heuristic one can make a dedicated (taboo) list of all solutions explored so far by the search heuristic and only new solutions are evaluated.

Approximation of the Loss Probabilities

By plugging a product form solution in the balance equations in (1), we find that: (i) the number of busy engineers in steady state, $n^e(\infty)$, is distributed as number of customers in an $M/M/E/E$ with an arrival rate $\lambda(1-\sum_{i=1}^{N} r_i\, P(n^{s_i}(\infty) = S_i))$ and a service rate γ, (ii) the number of parts on order (being replenished), $n^{s_i}(\infty)$ is distributed as number of customers in an $M/M/S_i/S_i$ queue with an arrival rate $\lambda\, r_i\big(1-P(n^e(\infty) = E)\big)$ and a service rate μ_i. To compute $\big(P(n^e(\infty) = E), P(n^{s_1}(\infty) = S_1), \dots, P(n^{s_N}(\infty) = S_N)\big)$, we propose the following stepwise iterative procedure: (i) Initialize, p_{eng}^{loss}, the blocking probability of a customer's request due to an unavailable engineer in the case there is an ample stock of spare parts. The latter probability is equal to the blocking probability in the $M/M/E/E$ queue (Erlang-B formula). (ii) For each part's type, compute the arrival rate of type-i requests admitted in the system as $\lambda\, r_i(1-p_{eng}^{loss})$. Find $p_{S_i}^{loss}$, the loss probability of type-*i* requests, as the blocking probability in the $M/M/S_i/S_i$ queue with an arrival rate $\lambda\, r_i(1-p_{eng}^{loss})$ and a service rate μ_i. (iii) Compute the arrival rate of requests admitted in the system as $\lambda\big(1-\sum_{i=1}^{N} r_i\, p_{S_i}^{loss}\big)$, using results in (ii). Find a new estimate of p_{eng}^{loss} as the blocking probability in the $M/M/E/E$ queue with an arrival rate $\lambda\big(1-\sum_{i=1}^{N} r_i\, p_{S_i}^{loss}\big)$ and a service rate γ. (iv) Repeat steps (ii) and (iii) until convergence of the loss probabilities.

The iterative approximation of the probabilities converges to a unique solution if

$$\max\left(\frac{\lambda}{\gamma\, E}, \frac{\lambda\, r_1}{\mu_1\, S_1}, \cdots, \frac{\lambda\, r_N}{\mu_N\, S_N}\right) < 1.$$

For a detailed proof see (Al Hanbali et al. 2016). Note, we believe that for most practical cases the previous condition is satisfied in order to guarantee a small loss probability of a request. This is especially for companies with a strategy focused on offering a high service level to the customers. Note that when $S_i = 0$, the probability $p_{S_i}^{loss} = 1$ and should not be included in the iterative calculations.

Quality and Efficiency of the Heuristics

To evaluate the quality of the approximation we performed a number of experiments with different numbers of part types (N). Due to the exponential growth of the number of balance equations in the exact case we could test our exact evaluations only for a very limited numbers of N. All the experiments are performed using Python based implementations on a computer with Intel Xeon E5-2697v2 2.70GHz CPU and 64GB RAM. The MILP optimization model is solved using Gurobi 6.0 optimizer.

The input data for these experiments are based on data from an OEM of advanced equipment. However, we had to modify that input data in order to satisfy the requirement of the experiments presented in this section. Namely, we scaled the failure rates such optimal solutions require about 2 SKU's for each type of failure. In Table 1we compare the total system cost for different numbers of part types *N*. We conclude that the our heuristics are accurate and time efficient.

Table 1 - Accuracy and efficiency of the heuristics

	MILP		E+H		A+H		Relative error	
N	**Costs**	**time (sec)**	**Costs**	**time (sec)**	**Costs**	**time (sec)**	**E+H**	**A+H**
1	61626.11	0.05	61626.11	0.02	61634.41	0.37	0.000%	0.013%
2	61877.33	0.22	61877.33	0.03	61885.66	0.47	0.000%	0.013%
3	73410.04	2.17	73410.02	0.12	73439.66	0.72	0.000%	0.040%
4	86213.11	110.10	86213.14	0.43	86276.56	0.97	0.000%	0.074%
5	99919.88	6117.56	99919.88	1.71	100018.69	1.30	0.000%	0.099%
6	-	-	105525.58	4.35	105628.00	1.35	-	0.097%
8	-	-	114363.87	42.08	114478.41	1.62	-	0.100%
10	-	-	121535.43	2341.09	121653.77	1.89	-	0.097%
15	-	-	-	-	163923.43	2.67	-	-
20	-	-	-	-	388558.17	3.29	-	-
30	-	-	-	-	425964.67	3.80	-	-
50	-	-	-	-	504623.00	6.83	-	-
70	-	-	-	-	1047400.18	10.62	-	-
90	-	-	-	-	1324498.79	24.45	-	-

Based on the results of this experiments we can conclude the following:

- Results of MILP and E+H are almost the same. There is a slight difference that occurs due to numerical errors in MILP solver. This suggests that the optimization heuristic is very efficient and gives the optimal solution in most cases.
- Results of E+H and A+H are comparable. There is a slight difference in the optimum cost due to the approximation of the loss probability in A+H. The optimum costs of A+H are always larger than those of E+H.
- If we start with no item on stock, we rarely remove a resource unit in E+H and A+H.

CASE STUDY AND MANAGERIAL INSIGHTS

We consider a scenario with 60 parts from the company case. The holding cost per part per year is equal to 15% of the new part price, the engineer wage per year is equal to 50,000, and the penalty cost per request is equal to 200,000 (unless it is specified in the experiment). We performed a number of experiment in order to answer the following questions: (i) How the penalty costs will affect the total costs and the system fill rate? (ii) How the system KPI's depend on the service rate of the repair process (μ) and of the replenishment process (γ)? (iii) How the system KPI's depend on the variability of the items failure, and repair and replenishment?

According to simulations, we find the loss probability is insensitive to the coefficient of variations of the repair time and the replenishment time. For example, by increasing the latter coefficient of variations from 0.7 to 2.2 the loss probability stays almost the same. This is true for different values of the stock level and the number engineers. Note, the loss probability is sensitive to coefficient of variations the time to failures. In the following, we answer the first two questions.

Impact of Penalty Cost

It is known for many inventory systems with penalty costs for backorders that the optimal inventory level can be found using "news-boy" type relations between cost parameters and backorder probabilities. Similar results are also possible for the lost-sales type models. This allows finding an appropriate penalty costs when the desired loss probability is known. For the presented model such relation is not really possible due to the relation between numbers of engineers and stock levels, as will be shown in the examples below. This fact strengthens the role of the presented analysis and optimization heuristics. In Figure 1, we show the total costs and the achieved fill rate (percentage of satisfied orders from stock on-shelf) changes with the increase of the penalty costs.

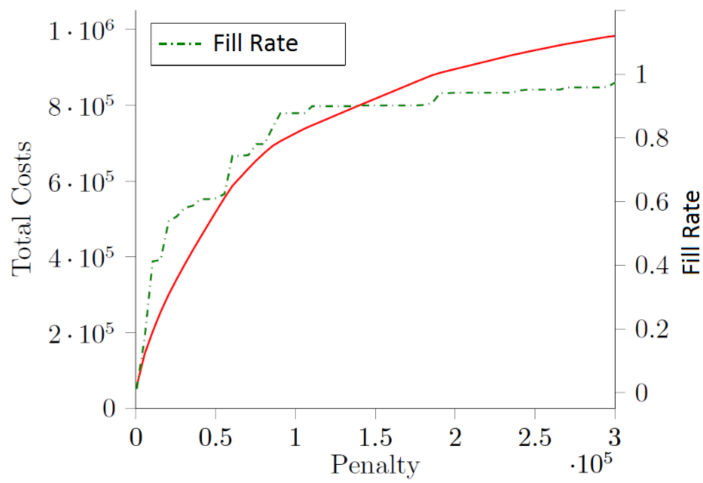

Figure 1: Impact of penalty cost on the total cost and fill rate

Faster Repair Process vs Closer Supplier

In this section, we vary the replenishment (e.g., closer vs farther supplier) rates by multiplying $\mu_i, i = 1, \ldots, 60$ by a factor $K_\mu \in [0.01; 2]$. In a similar way, we multiply γ the repair rate (e.g., slower vs faster repair process) by a factor $K_\gamma \in [0.05; 2]$. Figure 2 shows the total optimal costs obtained using the A+H heuristic as function of K_μ and K_γ. In this specific case study, we conclude it is more beneficial for the OEM to reduce the repair rate than the replenishment rate.

ACKNOWLEDGMENT

This publication was made possible by the NPRP award [NPRP 7-308-2-128] from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the author[s].

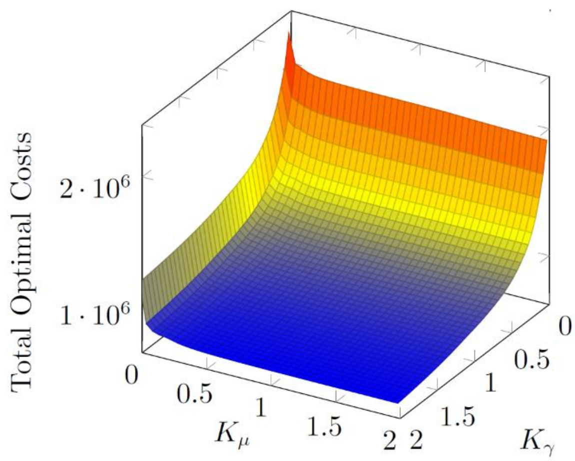

Figure 2 **-** *Sensitivity of of the total optimal cost to repair time (K_γ) and replnishment time (K_μ)*

BIBLIOGRAPHY

Cohen, M. A., N. Agrawal, V. Agrawal. 2006. Winning in the aftermarket. *Harvard business review* **84**(5): 129.

Ebeling, C. E. 2004. *An introduction to reliability and maintainability engineering*. Tata McGraw-Hill Education.

Sherbrooke, C. C. 1968. Metric: A multi-echelon technique for recoverable item control. *Operations Research* **16**(1): 122-141.

Muckstadt, J. A. 2005. *Analysis and algorithms for service parts supply chains*. Springer.

Sherbrooke, C. C. 2004. *Optimal inventory modeling of systems: multi-echelon techniques*. Springer.

Agnihothri, S. R., U. S. Karmarkar. 1992. Performance evaluation of service territories. *Operations Research* **40**(2): 355-366.

Agnihothri, S. R. , A. K. Mishra, D. E. Simmons. 2003. Workforce cross-training decisions in field service systems with two job types. *Journal of the Operational Research Society* **54**(4): 410-418.

Vliegen, I. M. H., P. A. M. Kleingeld, G.-J. Van Houtum. 2010. Separate tools or tool kits: An exploratory study of engineers preferences. *International Journal of Production Economics* **125**(1): 173-184.

Song, J.-S., S. H. Xu, B. Liu. 1999. Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes. *Operations Research* **47**(1): 131-149.

Rahimi-Ghahroodi, S., A. Al Hanbali, A., W.H.M Zijm, J.-K. W. van Ommeren, A. Sleptchenko. Integrated Planning of Service Engineers and Spare Parts Inventory in Maintenance Logistics with spare parts emergency shipment. Under preparation.

Al Hanbali, A, Sleptchenko, A., Zijm W.H.M. Integrated service engineers and spare parts planning in the maintenance Logistics with engineers and parts emergency shipment. Under preparation.