Int. J. Production Economics 118 (2009) 260–268

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Joint optimisation of spare part inventory, maintenance frequency and repair capacity for *k*-out-of-*N* systems

Karin S. de Smidt-Destombes [a,*], Matthieu C. van der Heijden [b], Aart van Harten [b]

[a] *TNO Defence, Security and Safety, The Hague, The Netherlands*
[b] *University of Twente, Enschede, The Netherlands*

ARTICLE INFO

Available online 10 October 2008

Keywords:
Maintenance
Spare parts
Repair capacity
k-Out-of-*N* systems
Availability
Installed base

ABSTRACT

To achieve a high system availability at minimal costs, relevant decisions include the choice of preventive maintenance frequency, spare part inventory levels and spare part repair capacity. We develop heuristics for the joint optimisation of these variables for (a) a single *k*-out-of-*N* system under condition-based maintenance and (b) an installed base of multiple identical *k*-out-of-*N* systems under block replacement. We show that a straightforward extension of the METRIC method for spare part inventory optimisation yields inferior results, because both the availability and costs are not necessarily monotonous functions of the decision variables. We develop an adjusted marginal analysis and show that it performs considerably better in numerical experiments.

1. Introduction

Many of today's capital assets require a high availability, because the consequences of downtime can be serious. For example, the failure of a wafer stepper in the semiconductor industry usually leads to a production stop. This yields reduced output and hence reduced revenues, so the consequence of a failure is serious. The system availability is influenced by many tactical and operational decisions, such as the maintenance frequency, the amount of maintenance resources like service engineers and test equipment, and spare part inventories. A common approach is to decompose the overall trade-off in a set of subproblems. However, we can argue that there are clear relations between these decision variables.

For example, consider the interaction between spare part inventories and maintenance frequency. Demand for spare parts arises from both preventive and corrective maintenance. A higher preventive maintenance frequency leads to higher maintenance cost, but at the same time to a better predictable demand for spare parts and hence to a lower spare parts safety stock. Also, the interaction between repairable spare part inventories and the capacity needed to repair these spares cannot be neglected, see e.g. Sleptchenko et al. (2002, 2003): low repair shop capacity means a high utilisation, so long spare part repair leadtimes. As safety stocks should cover the demand during the leadtime, this means that savings on repair capacity lead to a need for more spare parts and vice versa.

In this paper, we discuss heuristics for the joint optimisation of maintenance frequency, spare part inventories, and spare part repair capacity. We focus on *k*-out-of-*N* systems with hot stand-by redundancy. That is, a system consists of N identical components of which only $k<N$ are required for system operation. The $N-k$ stand-by components have the same failure behaviour as the k operational components. We construct our optimisation heuristics based on approximations that we have developed before to calculate the system availability as function of the maintenance frequency, spare part inventories and repair capacity (De Smidt-Destombes et al., 2004, 2006, 2007). A complication is that the availability might not be a monotonous function of the maintenance

* Corresponding author.
E-mail address: ksmidt@feweb.vu.nl (K.S. de Smidt-Destombes).

0925-5273/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2008.08.058

frequency. When the frequency decreases, the probability that the system fails before maintenance starts increases and this pushes the availability down. On the other hand, the cycle length increases and the expected uptime in a cycle increases as well, which pushes the availability up. The aggregate effect may both be a decrease or an increase in the system availability. Therefore, the development of a joint optimisation method for spare part inventories, repair capacity and maintenance frequency is not straightforward.

The remainder of this paper is structured as follows. In the next section, we discuss the related literature. In Section 3, we develop an optimisation heuristic for a single k-out-of-N system with condition-based maintenance. We evaluate the quality of our heuristic in a numerical experiment in Section 4. We give our conclusions and directions for further research in Section 5.

2. Related literature

Although there is a lot of literature on spare part management (e.g. Sherbrooke, 2004; Muckstadt, 2005) and maintenance optimisation (e.g. Dinesh Kumar et al., 2000), relatively little has been published on the interaction between maintenance, spares and repairs. Most of the current literature deals with the interaction of two out of these three components in a specific setting.

The combination of maintenance and spare parts has been analysed by several authors. For example, Kabir and Al-Olayan (1996), Kabir and Farrash (1996) and Park and Park (1986) deal with an age-based maintenance strategy and non-repairable components. Brezavšček and Hudoklin (2003) present a model with a joint optimisation of a block replacement interval and the maximum inventory level. In Chelbi and Aït-Kadi (2001) the block replacement interval, the optimal stock level as well as the replenishment cycle are optimised simultaneously. Again the components are not repairable, which is encountered in most models that are concerned with joint optimisation of a maintenance policy and a spares provisioning policy.

For the interaction between spare parts and repair capacity, some models have been developed as well. Finite repair capacity is usually modelled by (multi-class) multi-server queues. Gross et al. (1985) were among the first to realise that the combination of inventory and queueing models might lead to insights in the trade-off with respect to maintenance flexibility achieved either through stocks or through sufficient capacity. They attempt to find a cost-optimal combination of the number of spare parts and the number of repair channels, under the constraint that a target service level is met. Kim et al. (2000) have presented an iterative algorithm to determine a cost optimal combination of repair capacities and spare part levels in a single item, multi-echelon model. Avsar and Zijm (2003) consider more general multi-echelon resource structures in which each repair facility may be a queueing network, and show how under Poisson failure rates stock levels at all echelons can be optimised. A similar approach can be used for multi-indenture structures and for combinations of multi-echelon and multi-indenture structures, see Zijm and Avsar (2003). Sleptchenko (2002) deals with the optimisation of the number of spare parts and repair capacity in a multi-item system. Sleptchenko et al. (2005) show that repair priorities may seriously reduce the spare parts investment needed to obtain a target supply availability.

Although the importance of integrating the maintenance strategy, spare parts and repair capacity is recognised, only a few papers describe quantitative models. Natarajan (1968) considers a single unit with spares and a number of repair facilities. By calculating the time to failure the availability is determined. Furthermore, Wang (1993, 1995) consider a single system consisting of operational and stand-by components. They optimise simultaneously the number of stand-by components, number of spares and the number of repairmen. These models are the ones that come the closest to our problem definition. The strongest resemblance is found in Wang (1993) in which there is a number of operating units, a number of warm stand-by units and a number of cold stand-by units (i.e. spare units). Choosing the failure rate of the operating and warm stand-by units to be equal, we have a redundant system in which replacements are done after each component failure (one warm stand-by component turns into an operating unit and a cold stand-by unit becomes warm stand-by). However, they do not cover the interactions we consider in this paper. They do consider a parameter affecting the time until a system failure, namely, the number of warm stand-by units; but they do not have a parameter for the maintenance frequency. So, the number of maintenance set-ups is fixed (maintenance is done after every unit failure). As a consequence the cost involved with the maintenance set-ups is fixed. In this paper we do consider the maintenance frequency as a parameter and we can influence the total maintenance costs by choosing the maintenance frequency.

3. Single system

We use the following model for a single k-out-of-N system. Maintenance is initiated when the system has $m \leqslant N-k+1$ failed components. After a deterministic leadtime L (which can also be equal to zero), maintenance is performed. The decision variables are the maintenance initiation level m, the spare parts stock level S and the repair capacity c. The expected costs per time unit $C_{m,S,c}$ include (i) the holding and depreciation costs of a spare part per time unit C_{spare}, (ii) the cost of repair capacity per time unit C_{capacity}, (iii) the maintenance set-up cost C_{init}. The goal is to minimise these costs $C_{m,S,c}$ given a lower bound Av^* for the expected operational availability $Av_{m,S,c}$. So, we formulate our problem as

$$\begin{aligned} \min \quad & C_{m,S,c} = \frac{C_{\text{init}}}{E[T_m] + L + E[D_{m,S,c}]} + SC_{\text{spare}} + cC_{\text{capacity}} \\ \text{s.t.} \quad & Av_{m,S,c} \geqslant Av^* \end{aligned} \tag{1}$$

Here $E[T_m]$ denotes the expected time until maintenance initiation (at m failed components) and $E[D_{m,S,c}]$ the expected maintenance duration. For the approximation