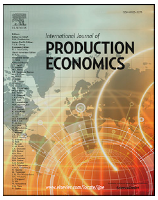
International Journal of Production Economics 196 (2018) 163–175

Contents lists available at ScienceDirect

International Journal of Production Economics

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



Reliability optimization for series systems under uncertain component failure rates in the design phase

Qianru Ge [a], Hao Peng [c], Geert-Jan van Houtum [a,*], Ivo Adan [a,b]

[a] Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands
[b] Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[c] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:
Capital goods
Reliability optimization
Performance-based contracting
Life cycle cost

ABSTRACT

We develop an optimization model to determine the reliability design of critical components in a serial system. The system is under a service contract, and a penalty cost has to be paid by the OEM when the total system down time exceeds a predetermined level, which complicates the evaluation of the expected cost under a given reliability design. Furthermore, in the design phase for each critical component, all possible designs are subject to uncertain component failure rates. Considering the computational intractability of evaluating the system performance, we develop approximate evaluation methods which take the system uncertainty into account. Numerical results show that the method which includes randomness in the number of failures, failure rates and repair times leads to efficient and accurate evaluations and to close-to-optimal design decisions when used in an enumeration procedure for the optimization problem. We also show that ignoring these three types of uncertainty may result in bad design decisions.

1. Introduction

Capital goods are machines or products that are used by manufacturers to produce their end-products or that are used by service organizations to deliver their services. Advanced technical systems such as medical systems, manufacturing systems, and defense systems, are examples of capital goods that are critical for the operational processes of their customers. System downtime of these capital goods can have serious consequences (e.g., millions of euros of reduced production output, extra waiting time of passengers, failure of military missions) and maintaining these high-tech systems is too challenging for customers to take care of by themselves. Original equipment manufactures (OEMs) can take care of the maintenance and guarantee high system availability levels. These OEMs can be seen as performance providers rather than only solution providers (Helander and Möller, 2007). The guaranteed system availability levels are generally specified in Service Level Agreements (SLAs) within service contracts. Customers pay a price for the service contracts. When an OEM fails to meet the predetermined level of availability, the OEM needs to pay a penalty cost to the customer. Different types of service contracts are mentioned in Cohen et al. (2006), among which performance-based contracts (PBCs). According to Guajardo et al. (2012), "Performance-based contracting compensates the supplier based on the same outcome that the customer cares about (i.e., product utilization), and hence the supplier is motivated to increase product performance, associated with metrics such as product reliability and availability". Therefore PBCs are a certain type of service contracts, in which system performance translates into financial bonuses and penalties (Selviaridis and Wynstra, 2015). In the main model studied in this paper, we include the penalty side of a PBC contract. We also show how the bonus side can be included (see Section 6.1).

Under a SLA on system availability, one of the OEM's major concerns is the life cycle cost (LCC), defined as the total cost incurred in the design/development, production, operation, maintenance, support, and final disposition of a system over its anticipated useful life span (Barringer and Weber, 1996). There is literature aiming at cost saving under such a SLA (Öner et al., 2010). Measurements reported in Öner et al. (2007) for an engineer-to-order system show that the sum of maintenance cost and downtime cost is larger than the acquisition cost and constitutes a significant portion of the LCC. The service cost is incurred by system failures which are highly determined by system designs. Therefore, it is important that the LCC is taken into account in the reliability design decision.

Customers of capital goods measure the availability of these complex systems at the end of service contract periods. In the literature (Al

---

[*] Corresponding author.
E-mail addresses: q.ge@tue.nl (Q. Ge), haopng@gmail.com (H. Peng), G.J.v.Houtum@tue.nl (G.-J. van Houtum), i.adan@tue.nl (I. Adan).

https://doi.org/10.1016/j.ijpe.2017.10.022
Received 9 November 2015; Received in revised form 21 October 2017; Accepted 23 October 2017

Hanbali and Van der Heijden, 2013; De Souza E Silva and Gail, 1986; Takacs, 1957), this availability during a service contract period is denoted as the *interval availability*. The realized interval availability of a capital good should meet the required performance levels. When the interval availability is below its target, or equivalently, when the interval downtime is above the target downtime, the OEMs will pay a penalty cost to the customer that is generally proportional to the exceedance. Therefore, it is important for the OEMs to have a method to determine the exceedance of the target downtime under a given design. This method can then be used when optimizing the reliability design.

In reality, engineers have to select a certain design from all possible alternatives for each critical component in a system during the design phase. The outcome of any development process for a certain design is uncertain with respect to its reliability level. For example, since the failure mechanisms of some emerging technologies (e.g., micro-electro-mechanical systems) are complex, it is often difficult to predict the actual reliability levels of the critical components before the development. Therefore newly-designed devices have been found to have component failure rates that deviate significantly from the expectations after the completion of a design. Brown and Burke (2000) and Brown et al. (2004) showed how individual system failure rates can deviate from the average failure rate of a population of similar equipment by collecting empirical data from the power delivery industry. This *uncertainty in component failure rates* also needs to be considered in the reliability design.

In this paper, we study the system reliability design problem by minimizing the LCC and considering both uncertain component reliability and interval availability under a SLA on system availability. So far, this problem has not been studied in the literature. For this new problem, our contribution is as follows. The key issue is the calculation of the exceedance of the target downtime under a given reliability design. In this calculation, one needs to take three uncertainties into account: the uncertainty in the component failure rates, the stochastic nature of the number of failures during a service contract period, and the stochastic repair times. This may be done by simulation, but that leads to a relatively long computation time per evaluation of a reliability design and to too long computation times for the optimization problem for instances of a reasonable size. Therefore, we propose a fast approximate evaluation method that takes all three uncertainties into account. This method is called the *full-uncertainty method*. We compare its performance to the performance of two other methods: a stochastic method ignoring the uncertain failure rates (partial method), and a deterministic method using expected total downtime as the actual total downtime (zero-uncertainty method). These approximate evaluation methods are also used in an optimization procedure for the reliability design, and hence lead to heuristic solutions for the reliability design problem. We show that the full-uncertainly method clearly outperforms the other two methods, both for the evaluation of a given design and when used in an enumeration procedure for the optimization problem. The use of the other two methods for the optimization problem may lead to bad reliability design decisions.

The remainder of the paper is organized as follows. In Section 2, we briefly review related literature. Section 3 gives the model description and model formulation. We describe the three approximate evaluation methods in Section 4. The numerical results are presented in Section 5. In Section 6, we first discuss how the bonus side of a PBC contract can be included. Next, we show the extension to a contract period with multiple subperiods and an interval availability target per subperiod. Conclusions and directions for future research are presented in Section 7.

2. Literature

We review the literature regarding the main characteristics of our problem: Interval availability, reliability optimization, and uncertain component failure rates.

As defined by Nakagawa and Goel (1973), *interval availability* is the fraction of time that a system is operational during a time period $[0,T]$. In the literature, several methods have been proposed to measure interval availability during a service period. Takacs (1957) considers a general stochastic on-off process and derives an exact expression for the distribution function of the total length of the off periods during a time period $[0,T)$. These off periods may represent the availability of a technical system. The exact expression has a high computational complexity, so that it can only be used for smaller problem instances. In Van der Heijden (1988), an approximate evaluation of the interval availability distribution is developed based on two-moment fits for the on and off periods. This method is accurate and can solve large problem instances in short computation times. De Souza E Silva and Gail (1986) consider a Markov process on a finite state space, where a subset of states represents 'good states' and they derive an exact expression for the distribution function of the total amount of time that the Markov process is in good states during a time period $[0,T)$. The good states may represent states in which a given technical system works. Also this exact expression of De Souza E Silva and Gail (1986) has a high computational complexity, so that it only can be used for smaller problem instances. Al Hanbali and Van der Heijden (2013) consider a two-echelon spare parts inventory system consisting of a single depot and multiple bases, and they study the logistics availability of technical systems installed at the bases during a time period $[0,T)$. They derive an approximate Markov process where a subset of states corresponds with logistics availability. They follow a numerical approach that can solve large problem instances within reasonable computation times. For more references and approaches on interval availability, we refer to the literature discussion in Al Hanbali and Van der Heijden (2013).

The evaluation of a given reliability design is related to the interval availability literature. Like Takacs (1957) and De Souza E Silva and Gail (1986), we derive an exact formula for exceedance of the target downtime, which is equivalent to deriving the distribution function of the interval availability. Our formula is different because we assume that the off periods are very short and they occur according to Poisson processes. In addition, we have to take uncertainty in the failure rates into account. Next we develop approximate, numerical approaches in order to be able to evaluate large problem instances. In these numerical approaches, we use two-moment fits, which is a general technique that was also used by Van der Heijden (1988). Nevertheless, our approach differs because it is based on a different exact formula.

Regarding *reliability optimization*, a lot of work has been done in this area since the 1990s (Kuo and Wan (2007)). For example, Mettas and Kallenberg (2000) determined the minimum required reliability for each component of a system in order to achieve a system reliability goal with minimum cost. The cost function for each component in this paper has been used in other papers as well; see e.g. Huang et al. (2007), Öner et al. (2010), and Jin and Wang (2012). Many papers maximize the system reliability by different techniques. For example, a random search process has been proposed by Beraha and Misra (1974) to determine the optimal reliability for each stage of a multi-stage system. Hwang (1975) used sequential unconstrained minimization, and Li and Haimes (1992) developed a 3-level decomposition approach to allocate the resources among subsystems optimally.

Some papers also built reliability allocation models to find optimal warranty policies for systems sold with traditional warranty contracts. For example, to minimize the system LCC, Monga and Zuo (1998) used genetic algorithms to solve the optimization problem and Öner et al. (2010) introduced a decision support model to jointly optimize the reliability level and spare parts inventory level of a single-component system. To maximize the profit, Huang et al. (2007) proposed a model to compute the optimal warranty policy under different market situations by using the maximum principle method. Park et al. (2015) formulated a model to determine the optimal warranty period for the manufacturer. They minimize the warranty cost consisting of the repair cost and downtime cost. In recent years a growing number of papers built reliability optimization models within the context of service contract. Jin