Constrained maintenance optimization under non-constant probabilities of imperfect inspections

J.P.C. Driessen[a,*], H. Peng[b], G.J. van Houtum[a],

[a]Eindhoven University of Technology, School of Industrial Engineering,the Netherlands
[b]Chinese Academy of Science, China

Abstract

Keywords: Maintenance Optimization, Delay-Time Model, False Positives, False Negatives, Human Factors

1. Introduction

Unexpected failures cause downtime for many advanced technical systems, such as airplanes, trains, baggage handling systems, and medical systems. To avoid such unexpected failures, maintenance is done during system operation, since maintenance increases conditional system reliability (Lewis, 1987). The costs of these maintenance activities comprise 15-60% of the total production costs in a manufacturer's facility (Mobley, 2002). Reducing the operating costs of such advanced technical systems can, therefore, be achieved by lowering the maintenance costs. To facilitate this, mathematical maintenance models and techniques are used to derive optimal maintenance policies.

The literature on maintenance optimization is rich and covers various areas such as system replacement, inspections, repair, and maintenance scheduling (van Oosterom et al., 2014). These areas of maintenance optimization underlie modeling techniques that describe system degradation. A commonly used technique for modeling system degradation is the Delay Time Model (DTM). This model distinguishes three system states: normal, defective, and failed. The system operates properly in the normal state; operates in the defective state as well, but needs maintenance; or has failed. The DTM is typically studied under inspection based maintenance policies, i.e., inspections are done to reveal the system's degradation. Elaborate literature overviews of the DTM up to 2012 are provided by Baker and Christer (1994), Christer (1999) and Wang (2012). The most recent advancements, since 2012, include postponements of maintenance actions when the defects are detected (van Oosterom et al., 2014), and the application of the DTM to systems that have redundant components (Wang, 2013). Furthermore, multiple different forms of preventive maintenance activities, such as routine service, preventive system replacement, and manual inspection, aided by condition-monitoring, are combined in two models, based on the DTM (Wang, 2013).

[*]Corresponding author, Tel.: +31 40 247 2601, E-mail: j.p.c.driessen@tue.nl

In this research, we focus on a single-component DTM, i.e., the system studied consists of a single component. These single-component DTMs are relevant for the DTM literature, because these provide important building blocks for further developments of multi-component DTMs (Wang et al., 2010). Our considered DTM is combined with an inspection based maintenance policy. Scarf et al. (2009) studied a general class of inspection policies, consisting of an inspection interval length T and a number of inspections before preventive system replacement M. This implies that the system is preventively replaced at time MT, whatever its state may be. The variable M denotes a general policy, capable of representing an age-based maintenance policy under $M = 1$, a pure inspection policy under $M = \infty$, and hybrid policies under a finite $M > 1$. We will employ this general inspection policy, and denote it by (M, T).

The DTM literature considering inspection policies, commonly assume that the inspections are perfect. However, due to (human) errors, inspections are usually not perfect in practice (Wickens, 1992). Therefore, imperfect inspections have been included in single-component DTMs by Okumura et al. (1996), and Berrade et al. (2013). Both works consider two types of imperfect inspection behavior, i.e., false positives and false negatives. A false positive corresponds to the judgment that a system is in its defective state, when it is actually in a normal state. False negatives correspond to the judgment that a system is in a normal state when, in fact, that system is in a defective state. We denote the probabilities of false positives and false negatives by α and β, respectively. For an overview, Table 1 is included.

Table 1: Probabilities of inspection behavior

		System State	
		Normal	Defective
Inspection outcome	Normal	$1-\alpha$	β
	Defective	α	$1-\beta$

The probabilities of false positives and false negatives, α and β, are assumed to be constant by Okumura et al. (1996) and Berrade et al. (2013). However, such an approach might be inaccurate, and therefore Wang (2010) explored the effects of non-constant probabilities of false negatives in a multi-component setting. Yet, a non-constant probability of false positives has not been explored in the DTM literature, nor in any of the other maintenance models, such as Markovian based maintenance models and stochastic degradation models. Therefore, our work extends the literature by proposing a non-constant probability of false positives under a single-component DTM. Furthermore, we extend the literature by relating the probability of false negatives to the system's duration in the defective state, relative to its delay time, as opposed to Wang (2010), who only considers the duration the defective state. Note that we consider a single-component system, whereas Wang (2010) considers a multi-component setting. For a schematic literature overview of single-component DTMs under imperfect inspections, see Table 2.

We borrow findings from different literature streams to base our non-constant probabilities

Table 2: Imperfect inspections in single-component DTM literature

		False Positives	
		Constant	Non-constant
False Negatives	Constant	Okumura et al. (1996); Berrade et al. (2013)	
	Non-constant		This research

on. In the literature on nondestructive testing (NDT), the Probability of Detection (POD) is studied. This POD is defined as the probability of finding a degradation when it is present (Berens, 1989). It reflects the probability of false negatives by β=1-POD (Nichols et al., 2008). As a system degrades, the POD increases; i.e., the probability of false negatives decreases with system deterioration (Forsyth and Fahr, 1998). Such a decrease in β implies a non-constant probability of false negatives with respect to system degradation. We will employ this type of relationship between system degradation and the probability of false negatives. The idea presented by Wang (2010) also captures such a relationship between system degradation and the probability of false negatives. In his work, system degradation is conceptualized by the duration that a system has been defective. In this research, we will conceptualize system degradation as the duration that a system has been defective relative to its delay time, which is a richer concept. For instance, weaker materials will have a lower delay time. Hence, under the same duration in the defective state, the weaker material (shorter delay time) will have a higher level of degradation, than the strong one.

With respect to the non-constant behavior of the false positive probability, we consider the field of psychology. As the industry, nowadays, pays more attention to system reliability aspects, the pressure on maintenance staffs increases. Such an increased pressure tends to motivate employees to violate inspection guidelines (Latorella and Prabhu, 2000), and rely on self-developed decision making instead of following inspection guidelines. Psychological literature labels this type of decision making, heuristic decision making. To avoid any confusion with the terminology of 'heuristic' in operations research, we will refer to it as self-developed decision making. This self-developed decision making is a way of decision making in which not all information is taken into account. Yet, it may provide a convenient way to tackle a problem, compared to that of complete information (Kahneman and Frederick, 2002). Multiple types of self-developed decision making are defined throughout the literature. In this paper, we consider attribute substitution (Kahneman, 2003). For this type of self-developed decision making, an alternative attribute, e.g., a system's duration of normal operation, may be used for the decision making, rather than the actual condition obtained through inspections. To the authors' best knowledge, such an attribute substitution, with respect to false positives, in inspections has not been studied in maintenance modeling up to date.

Besides the imperfectness of inspections in maintenance, reliability aspects are becoming increasingly important according to the industry's and European Union's agenda (European Union, 2009). Hence, companies explicitly include reliability measures in their maintenance analyses. As

Aven and Castro (2009) noted, reliability aspects can be included in the cost expression, but such an approach may be rather controversial, e.g., transforming human injuries into monetary terms. Therefore, they propose to include the reliability aspect in the form of a constraint.

In this research, we take into account the imperfectness of inspections, as well as an explicit reliability constraint. Hence, we propose a constrained single-component model, considering the non-constant behaviors of errors in inspections, thereby differing from the constant models of imperfect inspections in Okumura et al. (1996) and Berrade et al. (2013). Our analysis is directly transferable to a setting without a reliability constraint. Based on the POD literature, we model the changing probability of false negatives as being dependent on system degradation, i.e., the relative duration of defective state with respect to the delay time. In line with the theory of self-developed decision making in psychology, we model the changing probability of false positives as being dependent on the system has been operating in the normal state. Our objective is to minimize the average cost rate over an infinite time horizon, by optimizing the maintenance policy (M, T). The research's contributions are twofold: (1) we present an exact evaluation of a maintenance policy (M, T) for a single-component DTM including non-constant probabilities of false positives and false negatives; and (2) we propose a method for comparing a constrained inspection based model with non-constant probabilities of imperfect inspections, to a constrained inspection based model that has constant probabilities of imperfect inspections.

The remainder of this paper is organized as follows. In Section 2, we present the model. In Section 3, we give an exact evaluation of our policy, and we discuss the optimization procedure. In Section 4, we present a method for comparing a model with non-constant probabilities of false positives and false negatives, to a model with constant probabilities. In Section 5, we present the computational results, that compare both models (non-constant and constant). Section 6 concludes with some remarks and potential directions for future research.

2. Model Description

Let us consider a single-component system operating over an infinite time horizon. The system has three states: normal, defective and failed. In the normal operating state, the system is working properly, without any detectable defects. In the defective state, the system requires maintenance, but is still able to operate. The failed state of the system is self-announcing, and the system stops delivering its function immediately. To prevent the system from reaching its failed state, it is inspected periodically each $T > 0$ time units, is preventively maintained upon detection of the defective state, or is preventively replaced after $M \in \mathbb{N}$ inspections (at time MT). In other words, a (M, T) maintenance policy is followed. We assume that inspections are the only means to detect the normal and defective state. We do not need to reveal the failed state by inspections, as this state is self-announcing.

We denote the duration of the normal state, referred to as the time to defect, by the continuous random variable $X > 0$. This time to defect corresponds to the time between maintenance or replacement, and the arrival time of the defect. The random time the system takes from defect arrival to failure, without taking any maintenance or preventive replacement actions, is referred to as the delay time and denoted by the continuous random variable $H > 0$. The sum of both random variables is the system's time to failure. The cumulative distribution function (cdf) and the probability density function (pdf), corresponding to both state durations, are defined by $F_X(.)$ and $f_X(.)$ for the time to defect, and by $F_H(.)$ and $f_H(.)$ for the delay time, respectively. Both cdfs $F_X(.)$ and $F_H(.)$ are continuous functions.

The cost for doing an inspection is denoted by c_0. We assume perfect maintenance, implying that upon maintenance the system is renewed; i.e., the system is restored to the 'as-good-as-new' condition. Therefore, preventive maintenance is equivalent to preventive replacement. Consequently cost c_p are incurred for both preventive maintenance and preventive replacement. When the system unexpectedly fails in between inspections, corrective maintenance is immediately done with cost c_c, and the inspection schedule restarts; i.e., the first inspection is performed T time units after system failure. We assume that the failure cost are included in the corrective renewal cost c_c. Furthermore, we assume that $0 < c_0 < c_p < c_0 + c_p < c_c$, and the time for inspections and maintenance actions is negligible.

In our model, the inspections can be imperfect; i.e., an inspection error can occur, and the inspection outcome differs from the system's true state. We take two classes of imperfect inspections into account: (1) false positives; and (2) false negatives. For an overview, see Table 1. These two probabilities are assumed to be non-constant. Hence, we use two functions $\alpha(.)$ and $\beta(.)$.

For the non-constant probability of false positives, we assume that (some) engineers do not use the measurement outcome in the judgment, but replace this with a substitute attribute (Kahneman, 2003). We consider the time t that the system has been operating in its normal state as the substitute attribute. This time is counted from the last moment of system renewal, which may originate from a perfect preventive maintenance action, or from preventive system replacement. We assume that, if the time t of the inspection, approaches a threshold value a, the engineers (using the time t as a substitute attribute) will become more tempted to engage in a false positive. The variable a might relate to a temporal parameter at which the maintenance engineers believe the system typically becomes defective, e.g., the mean time to defect. If the time t of the inspection exceeds the threshold value a, we assume that the engineers using the attribute substitution, all send the system for preventive maintenance. The probability of $\alpha(t)$ is then increasing for $t < a$, and remains constant for $t \geq a$. For an illustration of $\alpha(t)$, see Figure 1.

When the defect appears at the realization x of the time to defect, we assume the system to start its degradation until system failure. If the realized delay time is h; i.e., the failure occurs at h time units after the defect arrival, the failure progress can be denoted by the relative duration $(t-x)/h$ in the defective state, for an inspection occurring at any time t, where $x \leq t \leq MT$. Based on the POD literature, we assume that the inspection engineers have difficulty in determining whether degradation exists for low values of the failure progress. The more the system degrades, the easier the detection of the degradation becomes; i.e., the probability of false negatives $\beta((t-x)/h)$ is nonincreasing in $(t-x)/h$. For a more detailed discussion, see Palmberg et al. (1987) and Berens (1989). An illustration for $\beta((t-x)/h)$ is presented in Figure 1.

We explicitly include a reliability constraint in our model, following the approach proposed by Aven and Castro (2009). The reliability constraint is commonly defined in industry by the maximum average number of failures per time unit over an infinite time horizon, which we denote by R_{max}. $R(M,T)$ corresponds to the average number of failures per time unit under policy (M,T), over an infinite time horizon.

Our aim is to minimize the average cost rate under a reliability constraint by optimizing the maintenance policy decision variables M and T. This yields the following formulation of our optimization model:

$$
\begin{aligned}
&\min_{M,T} \quad g(M,T) \\
&\text{s.t.} \\
&\qquad R(M,T) \le R_{max} \\
&\qquad T > 0,\ M \in \mathbb{N}.
\end{aligned}
\tag{1}
$$

2.1 Notation

X	:	Continuous non-negative random variable representing the system's time to defect
H	:	Continuous non-negative random variable representing the system's delay time
$F_X(.)$	:	Cumulative distribution function for the random variable X
$F_H(.)$	:	Cumulative distribution function for the random variable H
$f_X(.)$	:	Probability density function for the random variable X
$f_H(.)$	:	Probability density function for the random variable H
M	:	Maximum number of inspections before preventive system replacement
T	:	Fixed time of the inspection interval length
c_0	:	Cost per inspection
c_p	:	Cost for preventive maintenance
c_c	:	Cost for corrective maintenance
$\alpha(.)$	:	Non-constant probability of a false positive inspection
$\beta(.)$	:	Non-constant probability of a false negative inspection

3. Model Analysis

This section presents the analysis of the optimization model from Equation 1. We will first derive expressions for the cost rate $g(M,T)$ and the reliability $R(M,T)$ in Section 3.1. The solution procedure for solving the optimization model is discussed in Section 3.2.

3.1 Evaluation of the cost and reliability function

From renewal theory (Ross, 1983) we know that the average cost rate over an infinite time horizon, over which the system operates, equals the average cost rate $g(M,T)$ over one renewal cycle. By the same reasoning, we know that the average number of failures per time unit over an infinite time horizon (under policy (M,T)), equals to the average number of failures per time unit $R(M,T)$ in a renewal cycle. Hence, we will derive expressions for $g(M,T)$ and $R(M,T)$.

Since $g(M,T)$ and $R(M,T)$ are defined over renewal cycles, we will elaborate on the definition of a renewal cycle. For our problem setting, we define a renewal cycle as the time between two successive renewal points. Renewal points are constituted by the time points at which corrective maintenance, preventive maintenance, or preventive system replacement (at MT) is done. Then, we know from renewal theory that:

$$g(M,T) = \frac{C(M,T)}{L(M,T)}, \tag{2}$$

where $C(M,T)$ represents the expected renewal cycle cost (under the (M,T) policy), and $L(M,T)$ represents the expected renewal cycle length under (M,T). The reliability $R(M,T)$ is derived similarly, and according to renewal theory:

$$R(M,T) = \frac{F(M,T)}{L(M,T)}, \tag{3}$$

where $F(M,T)$ denotes the expected number of failures in a renewal cycle (under the policy (M,T)). The derivation procedure for the terms $C(M,T)$, $F(M,T)$ and $L(M,T)$ is based on various event paths that may occur, and end the cycle. Instead of considering all event paths separately, we consider types of event paths, which may define multiple event paths. We will address each of these types of event paths, and derive the occurrence probability expressions, which finally yields the expressions for $g(M,T)$ and $R(M,T)$.

Event Path Type 1 (E1). The system survives without any defect occurrence until time MT, at which the cycle ends. This implies that its time to defect has to exceed MT, and that no false positives (thus only true positives) occur on any of the inspections before M. Because the inspections are done every T time units, true positives occur on inspections $1,\dots,M-1$, corresponding to the times $T,\dots,(M-1)T$. Hence, the probability of true positives is evaluated at these time epochs. The probability of true positives can be written directly in terms of the probability of false positives. Since the probability of false positives is non-constant, a product series is included from inspections 1 to $M-1$. The probability expression for the event path of Type 1 equals

$$\pi_1 = \int_{MT}^{\infty} \prod_{n=1}^{M-1} \left(1-\alpha\left(nT\right)\right) f_X(x)dx.$$

Event Path Type 2 (E2). In this type of event paths, a false positive occurs on inspection $j \in \{1,\dots,M-1\}$, thereby ending the cycle. This implies that the time to defect exceeds time jT, and before inspection j no false positives have occurred. This yields the occurrence probability of a type 2 event path, characterized by inspection j:

$$\pi_{2,j} = \int_{jT}^{\infty} \prod_{n=1}^{j-1} \left(1-\alpha\left(nT\right)\right) \alpha\left(jT\right) f_X(x)dx.$$

Event Path Type 3 (E3). The system becomes defective in an interval $[(i-1)T, iT)$, characterized by inspection $i \in \{1, \dots, M\}$, and fails in this very interval, i.e., a false negative cannot occur. This implies that its delay time lies in the interval $[0, iT - x)$. Before defect arrival, no false positives occur. Then, the occurrence probability of an event paths of type 3 are characterized by inspection i is given by:

$$\pi_{3,i} = \int_{(i-1)T}^{iT} \prod_{n=1}^{i-1} (1 - \alpha(nT)) \int_{0}^{iT-x} f_H(h) dh f_X(x) dx.$$

Event Path Type 4 (E4). The system becomes defective in an interval $[(i-1)T, iT)$ characterized by inspection $i \in \{1, \dots, M-1\}$. In contrast to the third type of event paths, the system does not fail in this interval, but fails in an interval $[jT, (j+1)T), j \in \{i, \dots, M-1\}$. This leaves false negatives to occur on inspections i up to inspection j, and due to the non-constant false negative probability, we include a product series. Since the false negatives can only occur upon time instances of inspections, we consider kT in the product series, where $k = i, \dots, j$. No false positives occur before inspection i. Note that the event paths of type 4 are characterized by the inspections i and j, and therefore we obtain the probability expression:

$$\pi_{4,i,j} = \int_{(i-1)T}^{iT} \prod_{n=1}^{i-1} (1 - \alpha(nT)) \int_{jT-x}^{(j+1)T-x} \prod_{k=i}^{j} \beta\left(\frac{kT - x}{h}\right) f_H(h) dh f_X(x) dx.$$

Event Path Type 5 (E5). The system becomes defective in an interval $[(i-1)T, iT)$, where $i \in \{1, \dots, M-1\}$. No false positives occur before defect arrival. The system's defect is revealed upon inspection $j \in \{i, \dots, M-1\}$, denoting that the system's delay time has to exceed $jT - x$, and ending the cycle. Note that this detection of the defect (a true negative) at inspection j occurs at time jT. This means that for inspections i up to $j-1$ false negatives occur with a non-constant probability. The probability expression for an event path of type 5 corresponds to:

$$\pi_{5,i,j} = \int_{(i-1)T}^{iT} \prod_{n=1}^{i-1} (1 - \alpha(nT)) \int_{jT-x}^{\infty} \prod_{k=i}^{j-1} \beta\left(\frac{kT - x}{h}\right) \left(1 - \beta\left(\frac{jT - x}{h}\right)\right) f_H(h) dh f_X(x) dx.$$

Event Type 6 (E6). The system becomes defective in an interval $[(i-1)T, iT)$, where $i \in \{1, \dots, M\}$, and remains defective until the system is renewed at time MT. Before the defect arrives, no false positives occur. From the defect arrival at time x to MT, the system remains defective. This implies that inspections i up to $M-1$ undergo false negatives. This yields the probability expression for the final type of event paths, characterized by inspection i:

$$\pi_{6,i} = \int_{(i-1)T}^{iT} \prod_{n=1}^{i-1} (1 - \alpha(nT)) \int_{MT-x}^{\infty} \prod_{k=i}^{M-1} \beta\left(\frac{kT - x}{h}\right) f_H(h) dh f_X(x) dx.$$

From the six different types of event paths, we derive the expression for the expected renewal cycle costs $C(M, T)$. For types E3 and E4 corrective maintenance costs c_c are incurred, since these types end in system failure. Furthermore, by the definitions of E3 and E4, in total i and j inspections are done, respectively. Consequently the costs of E3 correspond to $ic_0 + c_c$, and for E4 are defined by $jc_0 + c_c$. Since all other types of event paths, E1, E2, E5, E6, end in preventive maintenance, costs c_p are incurred. Similar to types E3 and E4, the number of inspections and therefore the inspection costs can be determined. By considering the inspection variables i and j, all possible event paths can be captured by the six types E1-E6. To derive the expression for $C(M,T)$, we have to include all possible event paths. This implies that we sum the types of event paths over i and j appropriately, such that we cover all possible event paths that end a cycle. Then, we obtain Equation 4.

$$\begin{aligned} C(M,T) = (Mc_0 + c_p)\pi_1 + \sum_{j=1}^{M-1}(jc_0 + c_p)\pi_{2,j} + \sum_{i=1}^{M}(ic_0 + c_c)\pi_{3,i} + \sum_{i=1}^{M-1}\sum_{j=i}^{M-1}(jc_0 + c_c)\pi_{4,i,j} \\ + \sum_{i=1}^{M-1}\sum_{j=i}^{M-1}(jc_0 + c_p)\pi_{5,i,j} + \sum_{i=1}^{M}(Mc_0 + c_p)\pi_{6,i}. \end{aligned} \tag{4}$$

We derive the expression for the expected renewal cycle length $L(M,T)$ in a similar way to $C(M,T)$. For E3 and E4, the cycle ends in system failure and therefore takes $x + h$ time units. However, these are variables of integrals. Hence, we write the term $x + h$ inside both integrals. The other types of event paths are multiplied by the time epoch of the inspection ending the cycle. These inspection instances are directly transferable from Equation 4. By the same reasoning we sum appropriately over i and j, and obtain Equation 5.

$$\begin{aligned} L(M,T) = MT\pi_1 + \sum_{j=1}^{M-1} jT\pi_{2,j} + \sum_{i=1}^{M}\int_{(i-1)T}^{iT}\prod_{n=1}^{i-1}(1-\alpha(nT))\int_{0}^{iT-x}(x+h)f_H(h)dhf_X(x)dx \\ + \sum_{i=1}^{M-1}\sum_{j=i}^{M-1}\int_{(i-1)T}^{iT}\prod_{n=1}^{i-1}(1-\alpha(nT))\int_{jT-x}^{(j+1)T-x}\prod_{k=i}^{j}\beta\left(\frac{kT-x}{h}\right)(x+h)f_H(h)dhf_X(x)dx \\ + \sum_{i=1}^{M-1}\sum_{j=i}^{M-1} jT\pi_{5,i,j} + \sum_{i=1}^{M} MT\pi_{6,i} \end{aligned} \tag{5}$$

The expected number of failures in a renewal cycle $F(M,T)$ corresponds to the probability of corrective maintenance, by definition. Since corrective maintenance occurs upon system failure, we only consider E3 and E4. Under the similar reasoning as the ones from the derivations of $C(M,T)$ and $L(M,T)$, we sum appropriately over i and j. This way we obtain the expected number of failures in a renewal cycle $F(M,T)$:

$$F(M,T) = \sum_{i=1}^{M}\pi_{3,i} + \sum_{i=1}^{M-1}\sum_{j=i}^{M-1}\pi_{4,i,j}$$

3.2 Optimization procedure

The optimization problem from Equation 1 has a non-linear objective function and constraint. To reduce the problem complexity of this non linear constrained optimization problem with two decision variables M and T, we enumerate M and solve for the optimal inspection interval length T^* by using Sequential Quadratic Programming (SQP). The optimal inspection instance of preventive system replacement M^* can be found by selecting the value of M with the lowest average cost rate.

4. Comparing non-constant with constant probabilities

In this section we will present a method for comparing a model with non-constant probabilities, to a model with constant probabilities. In Section 4.1, we define expressions used to calculate the mean probabilities of false positives and false negatives. These will be used to make the comparison between both type of models (non-constant and constant) in Section 4.2.

4.1 Average false positive and false negative probabilities

To compare the model that has non-constant probabilities of false positives and false negatives with the model that has constant probabilities, we derive expressions for the average probabilities of false positives and false negatives, under a policy (M, T) in this section. We will not directly consider the type of event paths from Section 3, as these focus on the event paths' occurrence probability, instead of the probability of false positives or false negatives occurring. Hence, in this section we will introduce specific types of event paths for the probability of false positives and false negatives.

To determine the average probability of false positives μ_α, we condition to the occurrence of false positives; i.e., the system's time to defect has to exceed time T of the first inspection. Under this condition we consider two types of event paths for the false positives.

False Positive Event Path, Type 1 (FP1). The first type false positive event paths corresponds to the average probability of false positives, when the system becomes defective in the interval $[(i-1)T, iT)$, where $i \in \{2, \ldots, M-1\}$. Note that $i = 1$ is excluded from analysis by the condition that the time to defect has to exceed T. For this type of false positive event paths, $i-1$ inspections are done upon which a false positive may occur. We take the average over all these probabilities.

$$\theta_{1,i} = \frac{1}{i-1} \int_{(i-1)T}^{iT} \sum_{n=1}^{i-1} \alpha(nT) f_X(x) dx = \frac{F_X(iT) - F_X((i-1)T)}{i-1} \sum_{n=1}^{i-1} \alpha(nT).$$

False Positive Event Path, Type 2 (FP2). For the second type of false positive event paths, we consider the system's time to defect exceeding $(M-1)T$. Hence, we need to take $M-1$ false positive probabilities into account for the average, as a false positive upon inspection M, the last inspection, cannot occur (the system is renewed preventively, independent on the inspection outcome). The average false positive occurrence probability, then equals

$$\theta_2 = (1 - F_X((M-1)T)) \frac{1}{M-1} \sum_{n=1}^{M-1} \alpha(nT).$$

Summing both types of false positive event paths, and conditioning to the time to defect exceeding T, yields

$$\mu_\alpha = (1 - F_X(T))^{-1} \left(\sum_{i=2}^{M-1} \theta_{1,i} + \theta_2 \right). \tag{6}$$

The average probability of false negatives μ_β is derived similarly to the expression for μ_α. The average probability of false negatives is conditioned on the premise that false negatives can occur. This implies that when a defect occurs in an interval $[(i-1)T, iT)$ where $i \in \{1, \ldots, M-1\}$, its delay time has to exceed $iT - x$. Furthermore, since a defect can occur in any interval $[(i-1)T, iT)$ where $i \in \{1, \ldots, M-1\}$, we sum over all i for the condition. Similar to the false positives, we consider two events types for false negatives (under the condition).

False Negative Event Type 1 (FN1). This false negative event type denotes the average probability of false negatives when the system becomes defective in $[(i-1)T, iT)$, where $i \in \{1, \ldots, M-1\}$, and where the system fails in interval $[(jT, (j+1)T)$, with $j \in \{i, \ldots, M-1\}$. We take the average false negative probability over all inspections that can occur, yielding the probability expression for FN1:

$$\phi_{1,i,j} = \int_{(i-1)T}^{iT} \int_{jT-x}^{(j+1)T-x} \frac{1}{j-(i-1)} \sum_{k=i}^{j} \beta \left(\frac{kT-x}{h} \right) f_H(h) dh f_X(x) dx.$$

False Negative Event Type 2 (FN2). The second false negative event type corresponds to the average probability when the system becomes defective in $[(i-1)T, iT)$ where $i \in \{1, \ldots, M-1\}$, but it does not fail; i.e., the system's delay time exceeds $MT - x$. We only consider $i \in \{1, \ldots, M-1\}$ because a defect occurring in $[(M-1)T, MT)$ leads to a potential false negative upon inspection M. However, such a false negative cannot occur, as the outcome of the M^{th} inspection is irrelevant. Therefore, we exclude this case. Then the average probability of false negatives for FN2 becomes

$$\phi_{2,i} = \int_{(i-1)T}^{iT} \int_{MT-x}^{\infty} \frac{1}{(M-1)-(i-1)} \sum_{k=i}^{M-1} \beta \left(\frac{kT-x}{h} \right) f_H(h) dh f_X(x) dx.$$

Analogous to the expression for μ_α, here we also sum over all appropriate intervals characterized by inspections i and j, and we include the condition to obtain

$$\mu_\beta = \left(\sum_{i=1}^{M-1} \int_{(i-1)T}^{iT} (1 - F_H(iT-x)) f_X(x) dx \right)^{-1} \left(\sum_{i=1}^{M-1} \sum_{j=i}^{M-1} \phi_{1,i,j} + \sum_{i=1}^{M-1} \phi_{2,i} \right). \tag{7}$$

4.2 Comparison method

In this section, we present a method to compare a DTM with non-constant probabilities of false positives and false negatives, to a DTM with constant probabilities. We consider the DTM with constant probabilities for false positives and false negatives, to be an approximation of the model with non-constant probabilities. Note that our non-constant probabilities model can be easily altered to yield a constant probabilities model. For more details on a constant probabilities model, see Berrade et al. (2013).

For a proper comparison, the constant probabilities have to be set accordingly. For this purpose, we use the average probability of false positives μ_α, and the average probability of false negatives μ_β. Here, we assume that α_0, c_α, a, β_0, c_β, γ, and η are known. From Equations 6 and 7, we observe that μ_α and μ_β are both dependent on an inspection policy (M, T). For optimization, the optimal inspection policy is again dependent the probabilities of false positives and false negatives, μ_α and μ_β, respectively (see Section 3). Hence, for the model with constant probabilities, we propose to consider convergence in the policy (M, T), and consequently in the value of μ_α and μ_β. To this end, we will denote the optimal convergent policy by $(\widehat{M}, \widehat{T})$. Here, we assume that convergence to a single policy $(\widehat{M}, \widehat{T})$ exists. Furthermore, note that this convergent policy has to satisfy the reliability constraint $R(\widehat{M}, \widehat{T}) \leq R_{max}$. We obtain such a convergent policy iteratively. Therefore, we introduce the notation for the optimal policy $(\widehat{M}_i, \widehat{T}_i)$ of the i^{th} iteration, which has corresponding average probabilities of false positives μ_α^i, and false negatives μ_β^i. Remark that each $(\widehat{M}_i, \widehat{T}_i)$ also has to satisfy the reliability constraint, except for $i = 0$, since $i = 0$ is the initial policy.

We start by setting $i = 0$, and setting an initial inspection policy (M_0, T_0). From here we calculate μ_α^0 and μ_β^0, and start our iterative procedure to obtain the optimal convergent policy $(\widehat{M}, \widehat{T})$ (if it exists). The iterative procedure sets $i := i + 1$ and calculates the optimal inspection policy $(\widehat{M}_i, \widehat{T}_i)$ under the averages μ_α^{i-1} and μ_β^{i-1}. Then, based on $(\widehat{M}_i, \widehat{T}_i)$, the average probabilities μ_α^i and μ_β^i are calculated. The iterative procedure terminates when the absolute difference between the optimal inspection interval lengths of iterations i and $i-1$ is smaller than the threshold $\varepsilon = 0.01$. Furthermore, we require for termination that $\widehat{M}_i = \widehat{M}_{i-1}$. In case the policy does not converge to a single policy, we propose to consider the inspection policy which yields lowest costs, and compare this policy to the non-constant probability results.

Since the non-constant probabilities reflect the actual inspection behavior, we evaluate the actual cost rate $C(\widehat{M}, \widehat{T})$ and actual reliability $R(\widehat{M}, \widehat{T})$ under the convergent $(\widehat{M}, \widehat{T})$ policy. Note that for this evaluation, we take the non-constant probabilities, under the $(\widehat{M}, \widehat{T})$ policy, into account.

As we are considering constrained models, we will set the reliability $R(\widehat{M}, \widehat{T})$ as the reliability objective for the non-constant probabilities model, i.e., $R_{max} := R(\widehat{M}, \widehat{T})$. This ensures a one dimensional comparison (only costs). We propose to solve the optimization model with the modified constraint, yielding the optimal values for the decision variables M^* and T^* under the non-constant probabilities. The corresponding cost rate is denoted by $C(M^*, T^*)$ and can be compared to the cost rate $C(\widehat{M}, \widehat{T})$ by

$$\Delta C = \frac{C(\widehat{M}, \widehat{T}) - C(M^*, T^*)}{C(M^*, T^*)}(100\%).$$

To summarize, we propose the following method for comparing a constrained non-constant probabilities model to a constrained constant probabilities model. If we observe, in step 3, that the policy does not converge to a single policy $(\widehat{M}, \widehat{T})$, we select the policy that yields lowest costs.

1. Set $i := 0$, and set (M_0, T_0).

2. Calculate μ_α^0 and μ_β^0

3. Determine the optimal convergent $(\widehat{M}, \widehat{T})$ policy by:

   (a) Set $i := i + 1$

   (b) Calculate the optimal $(\widehat{M}_i, \widehat{T}_i)$ under μ_α^{i-1} and μ_β^{i-1}

   (c) Calculate μ_α^i and μ_β^i under $(\widehat{M}_i, \widehat{T}_i)$

   (d) If $\widehat{M}_i = \widehat{M}_{i-1}$, and $|\widehat{T}_i - \widehat{T}_{i-1}| \leq \varepsilon$, set $\mu_\alpha := \mu_\alpha^i$ and $\mu_\beta := \mu_\beta^i$, and go to 3. Else, go to (a)

4. Calculate $C(\widehat{M}, \widehat{T})$ and $R(\widehat{M}, \widehat{T})$

5. Set $R_{max} := R(\widehat{M}, \widehat{T})$ for the non-constant probabilities model

6. Calculate M^*, T^* and $C(M^*, T^*)$

7. Compare $C(\widehat{M}, \widehat{T})$ with $C(M^*, T^*)$ by calculating ΔC.

5. Computational study

In this section we present our testbed, based on data from the Dutch rail industry. The system's properties are set such to reflect actual systems in use. We will focus on four main factors, and we generate a testbed consisting of 81 instances, by considering three choices for each factor.

5.1 Testbed

Since we cannot capture all non-constant behavior of the false positive and false negative probabilities by a single variable for each, we include constant probabilities α_0 and β_0, respectively. These constant probabilities enable us to include the effects of other variables, yet in a constant way.

First, let us consider the probability of false positives. We relate this probability in a piecewise manner to the perceived defect progress t/a, where $t \leq x$. Recall that we consider some of maintenance engineers using the substitute attribute t/a instead of the measurement outcome, to determine whether to send the system to maintenance. For inspections occurring prior to a threshold value a, we assume that the probability of false positives increases linearly in time t. This behavior is based on the assumption that the engineers become more tempted in engaging in a false positive when the time approaches a. For the inspections occurring after the threshold time a, we assume that the system is always send to maintenance by the engineers using this attribute substitution. Since there are also engineers not using the perceived defect progress, the probability of false positives is linearly increasing until threshold a, and from this threshold remains constant at $\alpha_0 + c_\alpha$. The variable c_α directly relates to the relative size of engineers using the attribute substitution. Then,

$$\alpha(t) = \alpha_0 + c_\alpha \begin{cases} \frac{t}{a} & \text{if } t \leq a \\ 1 & \text{else} \end{cases} .$$

Regarding the probability of false negatives, we follow Probability of Detection (POD) literature, as discussed in Section 2. It is widely accepted that the relationship between the POD and system degradation is best modeled by a log odds distribution (Berens, 1989; Georgiou, 2006). Since β=1-POD (Nichols et al., 2008), there exists an explicit relationship (in log odds terms) between system degradation and the probability of false negatives. As we consider the failure progress $(t-x)/h$ to be an indicator of system degradation, we propose to model the explicit relationship (in terms of log odds) between the probability of false negatives and the failure progress $(t-x)/h$. Since we cannot capture the complete non-constant probability of false positives by only considering the failure progress, we include a constant probability β_0. Literature relating the false negative probability to system degradation, all assume that, under no degradation, the probability of a false negative equals 1, cf. Forsyth and Fahr (1998); Nichols et al. (2008). Based on the explicit relationship between the probability of false negatives and $(t-x)/h$), the inclusion of β_0, and a false negative probability of 1 under no degradation, we propose the following function for the probability of false negatives $\beta((t-x)/h)$:

$$\beta\left(\frac{t-x}{h}\right) = \beta_0 + \frac{1-\beta_0}{1+e^{\gamma+\eta\ln\left(\frac{t-x}{h}\right)}}, \qquad \text{if } x \leq t \leq x+h$$

Figure 1 illustrates the proposed functions for the probability of false positives related to t and for the probability of false negatives related to the failure progress $(t-x)/h$.

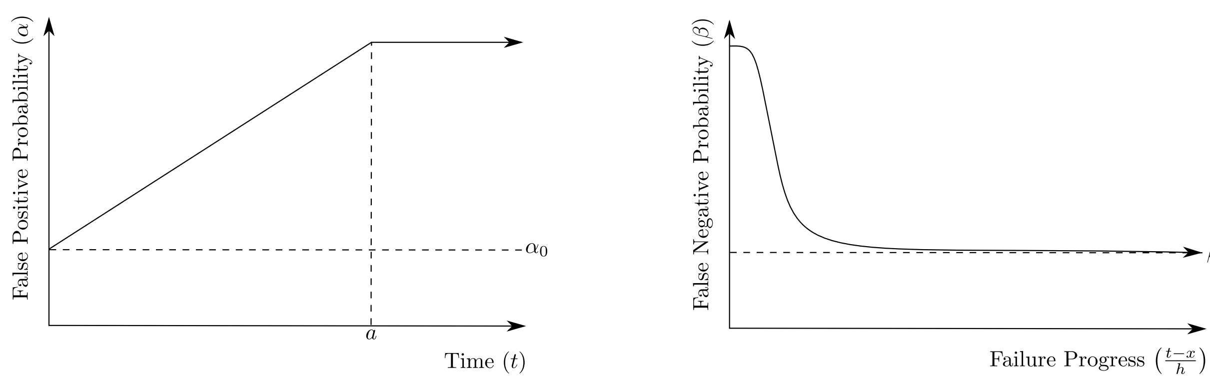

Figure 1: Illustration of non-constant probabilities α and β

We explore the effects of four factors on the cost differences between optimal policies of constant and non-constant probabilities of false positives and false negatives. These factors include: the reliability objective R_{max}, the inspection cost parameter c_0, the bandwidth of non-constant probability of false positives (and defining the slope) c_α, and the shape parameter η for false negatives. Note that decreases in γ yield a similar result as increases in η, as the ln() term is negative. All factor choices are presented in Table 4.

R_{max}	c_0	c_α	η
10^{-4}; 10^{-6}; 10^{-8}	50; 100; 200	0.25; 0.50; 0.75	1.5; 2.0; 2.5

Table 4: Factor choices

All other model parameters are fixed for our numerical study, and are listed in Table 5. The distributions used for the time to defect and delay time are Weibull distributions, with shape parameter $\delta_{\cdot}$ and scale parameter $\theta_{\cdot}$.

δ_X	θ_X	δ_H	θ_H	c_p	c_c	M_0	T_0	α_0	$a = E(X)$	β_0	γ
2.5	1234	2.5	203	1,000	2,000	12	60	0.05	1094.88	0.05	5

Table 5: Fixed parameter values

6. Conclusion

HERE IT COMES

Acknowledgements

The authors would like to express their gratitude to NedTrain for their input to this project.

Appendix

$$\mu_\alpha = \alpha_0 + \frac{c_\alpha}{1 - F_X(T)} \begin{cases} \sum_{i=1}^{M-1} \frac{F_X(iT) - F_X((i-1)T)}{i-1} \sum_{q=1}^{i-1} \frac{qT}{a} + \frac{1 - F_X((M-1)T)}{M-1} \sum_{q=1}^{M-1} \frac{qT}{a} & \text{if } MT \leq a \\ \sum_{i=2}^{\lfloor a/T \rfloor} \frac{F_X(iT) - F_X((i-1)T)}{i-1} \sum_{q=1}^{\lfloor a/T \rfloor} \frac{qT}{a} + & \text{if } MT > a \\ \sum_{\lceil a/T \rceil}^{M-1} \frac{F_X(iT) - F_X((i-1)T}{i-1} \left(\sum_{q=1}^{\lfloor a/T \rfloor} \frac{qT}{a} + i - \left\lceil \frac{a}{T} \right\rceil \right) + & \\ \frac{1 - F_X((M-1)T)}{M-1} \left(\sum_{q=1}^{\lfloor a/T \rfloor} \frac{qT}{a} + M - \left\lceil \frac{a}{T} \right\rceil \right) & \end{cases}$$

$$\mu_\beta = \beta_0 + \frac{1-\beta_0}{\sum_{i=1}^{M-1}\int_{(i-1)T}^{iT}\left(1-F_H\left(iT-x\right)\right)f_X(x)dx}$$

$$\left(\sum_{i=1}^{M-1}\sum_{j=i}^{M-1}\int_{(i-1)T}^{iT}\int_{jT-x}^{(j+1)T-x}\frac{1}{j-(i-1)}\sum_{q=i}^{j}\frac{1}{1+e^{\gamma+\eta\ln\left(\frac{qT-x}{h}\right)}}f_H(h)dhf_X(x)dx\right.$$

$$\left.+\sum_{i=1}^{M-1}\int_{(i-1)T}^{iT}\int_{MT-x}^{\infty}\frac{1}{M-i}\sum_{q=i}^{M-1}\frac{1}{1+e^{\gamma+\eta\ln\left(\frac{qT-x}{h}\right)}}f_H(h)dhf_X(x)dx\right).$$

References

Aven, T., Castro, I. T., 2009. A delay-time model with safety constraint. Reliability Engineering & System Safety 94 (2), 261–267.

Baker, R. D., Christer, A. H., 1994. Review of delay-time or modelling of engineering aspects of maintenance. European Journal of Operational Research 73 (3), 407–422.

Berens, A. P., 1989. NDE reliability data analysis. ASM Handbook. 17, 689–701.

Berrade, M. D., Scarf, P. A., Cavalcante, C. A. V., Dwight, R. A., 2013. Imperfect inspection and replacement of a system with a defective state: A cost and reliability analysis. Reliability Engineering & System Safety 120, 80–87.

Christer, A. H., 1999. Developments in delay time analysis for modelling plant maintenance. Journal of the Operational Research Society, 1120–1137.

European Union, 2009. Commission Regulation (EC) No 352/2009 of 24 April 2009 on the adoption of a common and assessment as referred to in Article 6(3)(a) of Directive 2004/49/EC of the European Parliament and of the Council. Official Journal of the European Union 52 (L108), 4–19.

Forsyth, D. S., Fahr, A., 1998. An evaluation of probability of detection statistics. In: Conference proceedings: Airframe Inspection Reliability under Field/Depot Conditions, RTO-MP-10, AC/323 [AVT] TP/2, Bruxelles.

Georgiou, G. A., 2006. Probability of detection (pod) curves: derivation, applications and limitations. Jacobi Consulting Limited Health and Safety Executive Research Report 454.

Kahneman, D., 2003. Maps of bounded rationality: Psychology for behavioral economics. American economic review, 1449–1475.

Kahneman, D., Frederick, S., 2002. Representativeness revisited: Attribute substitution in intuitive judgment. In: Heuristics and biases: The psychology of intuitive judgment. New York: Cambridge Univ. Press, pp. 49–81.

Latorella, K. A., Prabhu, P. V., 2000. A review of human error in aviation maintenance and inspection. International Journal of Industrial Ergonomics 26 (2), 133–161.

Lewis, E. E., 1987. Introduction to Reliability Engineering. John Wiley & Sons.

Mobley, R. K., 2002. An introduction to predictive maintenance. Butterworth-Heinemann.

Nichols, J. M., Trickey, S. T., Seaver, M., Motley, S. R., 2008. Using ROC curves to assess the efficacy of several detectors of damage-induced nonlinearities in a bolted composite structure. Mechanical Systems and Signal Processing 22 (7), 1610–1622.

Okumura, S., Jardine, A. K. S., Yamashina, H., 1996. An inspection policy for a deteriorating single-unit system characterized by a delay-time model. International Journal of Production Research 34 (9), 2441–2460.

Palmberg, B., Blom, A. F., Eggwertz, S., 1987. Probabilistic damage tolerance analysis of aircraft structures. In: Probabilistic fracture mechanics and reliability. Springer, pp. 47–130.

Ross, S. M., 1983. Stochastic Processes. Wiley, New York.

Scarf, P. A., Cavalcante, C. A. V., Dwight, R. A., Gordon, P., 2009. An age-based inspection and replacement policy for heterogeneous components. IEEE Transactions on Reliability 58 (4), 641–648.

van Oosterom, C. D., Elwany, A. H., Çelebi, D., van Houtum, G. J., 2014. Optimal policies for a delay time model with postponed replacement. European Journal of Operational Research 232 (1), 186–197.

Wang, W., 2010. Modelling planned maintenance with non-homogeneous defect arrivals and variable probability of defect identification. Maintenance and Reliability 46 (2), 73–77.

Wang, W., 2012. An overview of the recent advances in delay-time-based maintenance modelling. Reliability Engineering & System Safety 106, 165–178.

Wang, W., 2013. Models of inspection, routine service, and replacement for a serviceable one-component system. Reliability Engineering & System Safety 116, 57–63.

Wang, W., Banjevic, D., Pecht, M., 2010. A multi-component and multi-failure mode inspection model based on the delay time concept. Reliability Engineering & System Safety 95 (8), 912–920.

Wickens, C. D., 1992. Engineering psychology and human performance . HarperCollins Publishers.