Integrating optimisation, priority setting, planning and combining of maintenance activities

Rommert Dekker

Department of Econometrics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, Netherlands

Abstract

In this paper we present an integration of optimisation, priority setting, planning and combining of maintenance activities. We use a framework which covers several optimisation models, like the block replacement, a minimal repair and an efficiency model, and develop a uniform analysis for all these models. From this analysis penalty functions are derived which can act as priority criterion functions. These penalty functions also serve as basic elements in a method to determine optimal combinations of activities and in maintenance planning. The framework further assists in setting up an elicitation procedure, in case deterioration modelling is based on expert judgement rather than on statistical analysis of data.

Keywords: Maintenance; Replacement; Optimization; Probability

1. Introduction

Every few years new surveys appear on maintenance optimisation, showing that it is a lively field and that many interesting mathematical problems can be found in the maintenance area (e.g. McCall, 1965; Pierskalla and Voelker, 1979; Sherif and Smith, 1981; Valdez Flores and Feldman, 1989; Cho and Parlar, 1991). Applications follow, but at a slow rate. These are stimulated by the advent of decision support systems (d.s.s.) for maintenance optimisation. One of the problems encountered in building a d.s.s. is which of the many optimisation models to select for incorporation and how to assist a user in choosing the right model.

Another problem encountered in practice is that many relationships exist between components to be maintained. Modelling these relations directly yields large models, which are difficult to analyse as they suffer from the curse of dimensionality. A decomposition approach is then to be preferred. In such an approach one applies simple models for individual components and uses the outcomes as input in a comprehensive model. This requires a certain structure of the underlying models. Little work is present in this respect.

Other problems are encountered in the implementation of maintenance policies for individual components. Urgent corrective maintenance work usually sets preventive maintenance aside, and

priorities have to be set. Furthermore, it can be profitable to combine maintenance activities, thereby saving common preparation work. Finally, maintenance plans have to be made in agreement with production plans, which may result in a restriction to certain time windows where only a limited number of activities can be executed.

This paper tackles these problems in an integrated way. To facilitate in selecting a model we present an overall framework for time-based preventive maintenance and analyse it with respect to the questions when and where there exists an average-cost minimum. The framework is based on the marginal costs of deferring preventive maintenance. These costs may be directly estimated by a user or specified through a number of models, including the block replacement model. The framework further allows an extension to priority setting, combination and planning of maintenance. It is based on experience with developing two decision support systems for maintenance optimisation (see Dekker and Smeitink, 1994) in which a number of these problems had to be tackled.

Few papers attempt to unify maintenance optimisation models. Aven and Bergman (1986) do, and in fact our framework is a simpler version of theirs. Yet they only consider optimisation and do not make links with combination, priority setting or finite horizon planning. The central notion in this paper, the marginal cost rate of deferring maintenance, was first introduced by Berg (1980). It was fruitfully applied in Berg and Cléroux (1982) and Berg (1995) for repair-limit models and Dekker and Smeitink (1991) and Dekker and Dijkstra (1992) for opportunity maintenance.

The structure of this paper is as follows. After introducing the framework in Section 2 we provide a basic analysis and state which models can be incorporated. Next we treat elicitation of parameters by means of expert judgement in Section 3. Penalty functions are introduced in Section 4. Combining execution of maintenance activities is considered in Section 5, priority setting in Section 6, and maintenance planning in Section 7.

2. The framework

2.1. Introduction

As there is quite a variety in maintenance activities there are many optimisation models. A method for planning, priority setting and combining of preventive maintenance activities should therefore embrace as many types of activities as possible. Priority setting refers to determining the order of execution of activities which have to be carried out. Planning encompasses the timing of activities in coordance with production and manpower requirements. Finally, combining refers to shifting originally planned execution times to allow joint execution at possibly different moments. Notice that all these processes share timing as dominant aspect and that they are not done once but repeatedly. Being able to plan and shift execution times is in fact one of the most important advantages of preventive maintenance over failure-based maintenance. As a result of the planning the work can be prepared beforehand and necessary spare parts can be ordered. Shifting work in time also allows a more evenly spread workload and thus a higher efficiency.

In a method for integrating optimisation with planning, priority setting and combining of maintenance activities it seems necessary to restrict oneselves to activities whose execution can be planned in advance. We therefore restrict ourselves in first instance to maintenance activities whose next execution moment is determined from its last one. To that end we formulate a framework for optimisation models and derive the results necessary for the existence and position of an average-cost minimum. It is a simpler version of the framework from Aven and Bergman (1986) and of the marginal cost analysis from Berg (1980, 1995). Yet our results are somewhat stronger. An extension to other types of maintenance will be discussed in Section 2.5.

2.2. The framework in the continuous time

Here we present the general structure of the framework for a continuous time setting; extensions follow later. Consider a component (for ease of terminology we use this term, it may also be a part of a system) which deteriorates in time and which can be returned to the as-good-as-new condition by a preventive maintenance activity. The main question the framework focuses at is *when to execute the activity* and the answer will be based on cost considerations. We primarily consider long-term average costs as objective criterion, as that best reflects what one should do on a long term. The central notion in the framework is the so-called *marginal expected cost* of deferring the execution of the activity for an infinitesimally small interval. We first consider the case in which the activity can be carried out at any moment against the same cost c^p. In this case it is natural to speak of the marginal deterioration cost rate, denoted by $m(\cdot)$, which is assumed to be a continuous and piecewise differentiable function of the time t since the previous execution of the activity. We will now show that these assumptions are sufficient to determine an average optimal maintenance interval.

Let $M(t) := \int_0^t m(x)\,\mathrm{d}x$, i.e. the total expected costs due to deterioration over an interval of t time units, when the component was new at the start. It easily follows from renewal theory, that the average costs $g(t)$ per time unit when executing the activity every t time units amount to

$$g(t) = (c^p + M(t))/t. \tag{2.1}$$

Our first objective is to find the t value which minimises $g(t)$. To this end we take the derivative and notice that $g'(t) = [m(t) - g(t)]/t$. Hence

$$g'(t) = 0 \iff m(t) - g(t) = 0 \iff tm(t) - \int_0^t m(x)\,\mathrm{d}x = c^p. \tag{2.2}$$

Eq. (2.2) is the key for the analysis of $g(t)$. Let $\Psi(t) \equiv tm(t) - M(t) = \int_0^t m(t) - m(x)\,\mathrm{d}x$. Notice that $\Psi'(t) = tm'(t)$. The following theorem summarises the relations between the behaviour of $m(t)$ and the existence of an average cost minimum. Part (iv) is a generalization of results for existing models (see e.g. Barlow and Proschan, 1965, for the block replacement model), the other parts are simple new results which are added for completeness.

Theorem 1. (i) *If $m(t)$ is nonincreasing on $[t_0, t_1]$ and $m(t_0) < g(t_0)$, then $g(t)$ has no minimum on $[t_0, t_1]$.*

(ii) *If $m_2(t) = m_1(t) + c$, for some c and all $t > 0$, then $g_2(t)$ and $g_1(t)$ have the same extremes.*

(iii) *If $m_1(t)$ is nonincreasing on $(0, t_l)$ and increases thereafter, then $g_1(t)$ has the same minima as $g_2(t)$, where $m_2(t) = m_1(t_l)$ for $t \le t_l$ and $m_2(t) = m_1(t)$ else and $c_2^p = c_1^p + \int_0^{t_l}(m_1(t) - m_1(t_l))\,\mathrm{d}t$.*

(iv) *If $m(t)$ increases strictly for $t > t_0$, where $m(t_0) < g(t_0)$, and either*
(a) $\lim_{t \to \infty} m(t) = \infty$, *or*
(b) $\lim_{t \to \infty} m(t) = c$ *and* $\lim_{t \to \infty}[ct - M(t)] > c^p$, *for some $c > 0$, then $g(t)$ has a minimum, say g^*, in t^*, which is unique on $[t_0, \infty]$; moreover,*

$$m(t) - g(t) \begin{cases} < 0 & \text{for } t_0 < t < t^*, \\ = 0 & \text{for } t = t^*, \\ > 0 & \text{for } t > t^*, \end{cases} \tag{2.3}$$

and

$$m(t) - g^* \begin{cases} < 0 & \text{for } t_0 < t < t^*, \\ = 0 & \text{for } t = t^*, \\ > 0 & \text{for } t > t^*. \end{cases} \tag{2.4}$$

(v) *If $m(t)$ is convex on $[t_0, T]$, where $m(t_0) < g(t_0)$ and $\frac{1}{2}(T + t_0)[m(T) - m(t_0)] > c^p + \int_0^{t_0} (m(t) - m(t_0))\, dt$, then $g(t)$ has a minimum, say g^*, in t^*, which is unique on (t_0, T) and (2.3) and (2.4) hold on $[t_0, T]$. If $t_0 = 0$, it is sufficient that $\frac{1}{2}T[m(T) - m(0)] > c^p$.*

Proof. (i): Notice that $m(t_0) < g(t_0)$ implies that $\Psi(t_0) < c^p$. If $m(t)$ is nonincreasing, then $\Psi'(t) \leq 0$ and the result is immediate.

(ii): If $m_2(t) = m_1(t) + c$, then $\Psi_2(t) = \Psi_1(t)$ and the result is immediate.

(iii): According to (i) neither $g_1(\cdot)$ nor $g_2(\cdot)$ have a minimum before t_l. Notice next that for $t > t_l$ we have $\Psi_1(t) = c_1^p \Leftrightarrow \Psi_2(t) = c_2^p$, from which the assertion follows.

(iv): Notice that

$$\Psi(t) - \Psi(t_0) = \int_{t_0}^{t} (m(t) - m(x))\, dx + t_0[m(t) - m(t_0)] > (t_1 - t_0)[m(t) - m(t_1)], \quad t > t_0,$$

for some $t_1 \in (t_0, t)$. Hence, $\Psi(t)$ increases strictly to infinity if $m(t)$ does so. Since $\Psi(t_0) < c^p$ by (i), $\Psi(t)$ passes the level c^p only once for $t > t_0$, which guarantees the uniqueness.

If $\lim_{t \to \infty}[ct - M(t)] = d$ for some $d > c^p$, then it easily follows that $\lim_{t \to \infty} g(t) = c$. Moreover, for t large enough, say $t > t_\varepsilon$ we have $M(t) < ct - d + \varepsilon$ and $g(t) < c + [c^p - d + \varepsilon]/t$ for any $\varepsilon > 0$. Hence if $c^p - d < 0$, then $g(t)$ approaches c from below, implying that it must have a finite minimum. The uniqueness of the minimum follows from the fact that (2.2) implies that $m(t)$ intersects $g(t)$ in minima from below and in maxima from above. As $m(t)$ is strictly increasing on $[t_0, \infty)$, there can be no maxima in that region.

(v): If $m(t)$ is convex on $[t_0, T]$, then $M(T) - M(t_0) < \frac{1}{2}(T - t_0)[m(T) + m(t_0)]$. Inserting this in $\Psi(T)$ and using assertion (iii) shows after some algebra that $\Psi(T) > c^p$, from which the results follow in the same way as part (iv).

Remark. A decreasing $m(t)$ may be due to burn-in failures. Part (iii) of this theorem shows that we only need to estimate their contribution to the total costs and that we can leave the burn-in failures out of the modelling of $m(t)$ provided that a compensation is made for them in c^p. In this way we can take care of the bathtub curve.

Relation (2.4) can be interpreted in the following way (Berg, 1980, was the first one to introduce it). Consider at time t the two options: (a) maintain now or (b) defer maintenance for a time dt. For option (b) the expected costs over $[t, t + dt]$ amount to $m(t)\, dt + c^p$. For option (a) there are direct costs c^p, and the renewal is dt time units before the option (b). To compensate for this time difference we associate costs $g^*\, dt$ to the interval, which gives a total expected costs of $c^p + g^*\, dt$ for option (a). Subtraction then yields that maintaining is cost-effective if $m(t) - g^* > 0$. The myopic stopping rule: Maintain if $m(t) - g^* \geq 0$ is therefore average optimal.

Although a simple enumeration to locate the average-cost minimum usually satisfies in practice, one can speed up calculations by using relations (2.3) and (2.4) and applying a bisection procedure.

Special cases: (i) If $m(t) = \alpha t^{\beta - 1}$, $\alpha > 0$, then $\Psi(t) = \alpha(1 - 1/\beta)t^\beta$, which increases if $\beta > 1$. In that case $t^* = \sqrt[\beta]{\beta c^p / \alpha(\beta - 1)}$.

(ii) If $m(t) = at + b$, $a, b > 0$, then $\Psi(t) = \frac{1}{2}at^2$, and $t^* = \sqrt{2c^p/a}$.

Eq. (2.2) also allows us to do some sensitivity analysis. We have

Theorem 2. (i) *If $m_2(t) = \lambda m_1(t)$ with $\lambda > 1$ and $c_1^p = c_2^p$, then $t_2^* < t_1^*$.*
(ii) *If $m_2(t) - m_1(t)$ increases in t and $c_1^p = c_2^p$, then $t_2^* < t_1^*$.*
(iii) *If $c_2^p > c_1^p$ and $m_2(t) = m_1(t)$, then $t_2^* > t_1^*$.*

Proof. Notice that $\Psi_2'(t) - \Psi_1'(t) = t[m_2'(t) - m_1'(t)]$. For case (ii) we now have that $\Psi_2(t) - \Psi_1(t)$ increases in t and that $\Psi_2(t)$ reaches the level c_1^p earlier than $\Psi_1(t)$ from which the assertion follows. In case (i), $\Psi_2(t) - \Psi_1(t)$ is increasing if $\Psi_1(t)$ increases and the same argument holds. Assertion (iii) is also a direct consequence of (2.2).

2.3. Determination of the deterioration costs

In this section we present a number of models that can be captured by the framework.

The block replacement model
In this model a component is replaced upon failure and preventively after a fixed interval of length t (see, e.g. Barlow and Proschan, 1965). The total deterioration costs over that interval, $M(t)$, are made up of the failure renewals, each against cost c^f. Let $H(t)$ denote the expected number of failures in $[0, t]$. Then we have $M(t) = c^f H(t)$. It is well-known that for $H(t)$ the following asymptotic expansion holds: $\lim_{t \to \infty}[H(t) - t/\mu] = \frac{1}{2}(\sigma^2/\mu^2 - 1)$. Hence Theorem 1(iv) implies the existence of a minimum provided that $c^p/c^f < \frac{1}{2}(1 - \sigma^2/\mu^2)$, which is exactly the condition derived in Barlow and Proschan (1965).

Minimal repair model with block replacements
In this model failures of a system occur according to a nonhomogeneous Poisson process with rate $\lambda(t)$. Upon failure the system undergoes a minimal repair, which brings it back to an as-good-as-before condition. Next to that the system may be replaced, which has to be planned in advance and cannot be combined with a failure repair. Let c^r, c^p denote the costs of a minimal repair and a preventive replacement respectively, hence $m(t) = c^r \lambda(t)$. An average-cost minimum exists if either $\lim_{t \to \infty} \lambda(t) = \infty$ or in case $\lim_{t \to \infty} \lambda(t) = c$ for some $c > 0$, if $\lim_{t \to \infty}[ct - \Lambda(t)] > c^p$, where $\Lambda(t) = \int_0^t \lambda(s) \, ds$. Notice that if $\lambda(t)$ follows a bathtub pattern, we may add by Theorem 1(iii) the costs associated to initial failures to the preventive maintenance costs, and consider for optimisation only the increasing part of $m(t)$. There is also a more general version of the minimal repair model in which replacements may be combined with failure repairs and in which the repair costs may vary. It is not yet clear whether that can be incorporated in the framework.

A standard inspection model
In this inspection model a component is inspected every t time units against costs c^p, which reveals whether the component is functioning. We assume that inspection is always accompanied by corrective actions which brings the component back to an as-good-as-new condition (e.g. calibration of instruments). A failure of the component can only be detected by inspection. Let $F(\cdot)$ denote the c.d.f. of the time to failure X and let c^u be the costs associated to a non-functioning component per time unit. In this case the deterioration costs $M(t)$ consist of the expected costs due to unavailability of the component over $[0, t]$. Hence $M(t) = c^u \int_0^t F(x) \, dx$, and $m(t) = c^u F(t)$. It is easy to show that we have $\lim_{t \to \infty}[M(t) - c^u t] = -c^u EX$, where EX denotes the expected lifetime. Hence there exists a unique minimum by Theorem 1 provided that $c^p < c^u EX$ (the unavailability costs during a lifetime of a component are more than the inspection and repair costs).

An efficiency model
Assume that the efficiency of a system drops in course of time and that preventive maintenance restores the efficiency to that of the as-good-as-new condition. The efficiency can be measured in terms of output vs input and either the output can go down, or the input may have to be increased to sustain the same output. The efficiency, $E(\cdot)$, as function of time, is scaled on the interval $[0, 1]$, where $E(0) \equiv 1$. Let c^e denote the costs per time unit associated to a zero efficiency. The deterioration costs $M(t)$ consist

of the total efficiency losses over $[0, t]$, i.e. $M(t) = \int_0^t c^e (1 - E(x)) \, dx$. Let $E(\infty)$ denote the limiting efficiency in case no maintenance is ever carried out. Similarly to the previous model we can establish that $\lim_{t \to \infty} [M(t) - c^e (1 - E(\infty)) t] = -c^e \int_0^\infty (E(\infty) - E(t)) \, dt$. Hence an optimal preventive maintenance interval exists provided that $c^P < c^e \int_0^\infty (E(\infty) - E(t)) \, dt$. This model is mathematically equivalent to that of running costs (see also Berg and Epstein, 1979), yet they do not derive conditions for optimality.

A combined model

All aforementioned models may be combined as deterioration costs may consist of failure costs, repair costs, efficiency losses and unavailability penalties together.

2.4. Extensions

In this section we will give a number of extensions of the framework.

Discrete time case

In the discrete time case actions may only be taken at discrete points in time. The only change for the framework is that all functions have to be discretised: i.e. $m(t)$ indicates the expected deterioration costs until the next time moment.

Scrapping value

Suppose a system is replaced by a new one every t time units against costs c^P and that it has a scrapping value $S(t)$ at age t. We assume that $S(t)$ is decreasing in course of time. Let $M(t)$ again denote the deterioration costs of the system and let $\tilde{M}(t) = M(t) - S(t)$. Notice that the total costs over a replacement cycle $[0, t]$ amount to $c^P - S(t) + M(t)$, which equals $c^P + \tilde{M}(t)$. Hence a scrapping value can be taken care of by adjusting the deterioration cost function. Finally, we would like to remark that this model is mathematically equivalent to the block replacement model with time-dependent replacement costs.

Discounted cost case

The analysis for the average costs case is easily extended to the discounted costs case, as is also shown in Berg's (1980) marginal cost analysis. Assuming a discount rate λ we remark that the expected discounted deterioration cost rate at time t is given by $m(t) e^{-\lambda t}$. Hence total expected discounted costs over an interval of length t, starting with a new system/component amount to $\int_0^t m(y) e^{-\lambda y} \, dy$. Total expected discounted costs $v(t)$ over an infinite horizon when replacements are made very t time units and starting with preventive maintenance, amount to

$$v(t) = \frac{c^P + \int_0^t m(y) \, e^{-\lambda y} \, dy}{1 - e^{-\lambda t}}, \quad t > 0. \tag{2.5}$$

In this case we have

$$\frac{dv(t)}{dt} = 0 \iff m(t) - \lambda v(t) = 0,$$

which leads to a similar analysis as for the average costs (see also Section 2.5).

Opportunity maintenance

Suppose that preventive maintenance can only be carried out at opportunities which are generated according to a renewal process. Let the generic r.v. Y denote the time between successive opportunities.

We assume that the opportunity process is independent of the deterioration process. In this case maintenance can no longer be planned, instead we consider policies of the control-limit type, i.e. maintain at the first opportunity after a threshold time t since the last execution of the maintenance. Let the r.v. Z_t denote the forward recurrence time, i.e. the time from t to the first opportunity. It easily follows from renewal theory that for the average costs $g_Y(t)$ we now have

$$g_Y(t) = \frac{c^p + \int_0^\infty M(t+x)\, dP(Z_t \le x)}{t + EZ_t}, \quad t > 0$$

(see Dekker and Smeitink, 1991). In a similar, but far more complicated way, they derive inequalities (2.3) and (2.4) with $m(t)$ replaced by $\eta(t) \equiv \int_0^\infty m(t+z)\, dP(Y \le z)$, the expected deterioration costs until the next opportunity.

2.5. An extension of the framework for age replacement

In the well-known age replacement model the marginal deterioration cost rate $m(\cdot)$ is a function of the *age* of a *component* rather than of the time since the last execution of the preventive maintenance activity. The age is set back to zero upon any renewal of the component, including failure renewals. This implies that the renewal cycle has a variable length. The framework can be extended in the following way (see also Berg, 1995). Suppose the time to a system renewal, possibly caused by a breakdown, is stochastic with c.d.f. $F(t)$ and p.d.f. $f(t)$. The long-run average costs are now given by

$$g(t) = \left(c^p + \int_0^t m(x)(1 - F(x))\, dx\right) / (L(t)) \tag{2.6}$$

where $L(t) = \int_0^t(1 - F(x))\, dx$ indicates the expected cycle length. It is easily shown that $g'(t) = (m(t) - g(t))/\{(1 - F(t))L(t)\}$. Let $\Phi(t)$ be the analogon of $\Psi(t)$, i.e. $\Phi(t) = m(t)L(t) - \int_0^t m(x)(1 - F(x))\, dx$. Hence $g'(t) = 0 \Leftrightarrow m(t) - g(t) = 0 \Leftrightarrow \Phi(t) = c^p$. Notice further that $\Phi'(t) = m'(t)[1 - F(t)]$. We are now in a position to formulate a theorem similar to Theorem 1 and of which the proof is analogous.

Theorem 3. (i) *If $m(t)$ is nonincreasing on $[t_0, t_1]$ and $m(t_0) < g(t_0)$, then $g(t)$ has no minimum on $[t_0, t_1]$.*
   (ii) *If $m_2(t) = m_1(t) + c$, for some c and all $t > 0$, then $g_2(t)$ and $g_1(t)$ have the same extremes.*
   (iii) *If $m(t)$ is nonincreasing on $(0, t_1)$ and increases thereafter, then $g(t)$ has the same minima as $g_2(t)$, where $m_2(t) = m(t_l)$ for $t \le t_l$ and $m_2(t) = m(t)$ else and $c_2^p = c^p + \int_0^{t_l}(m(t) - m(t_l))(1 - F(t))\, dt$.*
   (iv) *If $m(t)$ increases strictly for $t > t_0$, where $m(t_0) < g(t_0)$, and either*
(a) $\lim_{t \to \infty} m(t) = \infty$, or
(b) $\lim_{t \to \infty} m(t) = c$ where $c > \lim_{t \to \infty} g(t)$, for some $c > 0$,
then $g(t)$ has a minimum, say g^*, in t^*, which is unique on $[t_0, \infty)$; moreover,

$$m(t) - g(t) \begin{cases} < 0 & \text{for } t_0 < t < t^*, \\ = 0 & \text{for } t = t^*, \\ > 0 & \text{for } t > t^*, \end{cases}$$

and

$$m(t) - g^* \begin{cases} < 0 & \text{for } t_0 < t < t^*, \\ = 0 & \text{for } t = t^*, \\ > 0 & \text{for } t > t^*. \end{cases}$$

In case of the age replacement model the marginal deterioration cost rate $m(x)$ amounts to $(c^f - c^p)r(x)$, where $r(x)$ denotes the hazard rate, $r(x) = f(x)/(1 - F(x))$. In that case the numerator of (2.6) equals $c^p + (c^f - c^p)F(t)$.

Notice that the discounted cost case (see Section 2.4) can be regarded as a special case of the extended framework by considering discounting as a truncation of the system lifetime. Hence the cdf of the time to system renewal, $F(x)$, should be defined as $F(x) = 1 - e^{-\lambda x}$, where λ is the continuous discount rate. In this case the expected lifetime equals $1/\lambda$ and hence the total discounted costs per unit of time equal $\lambda v(t)$, where $v(t)$ is given by (2.5).

The main problem to use the age replacement extension for planning and combining is that we no longer can predict in advance whether we will replace at some time t, as that depends upon the possible occurrence of failures in between. Doing a correct analysis implies that we have to condition on all possible events between the moment of planning and the expected moment of execution. This directly leads to intractable models in case of multiple components. An heuristic way out is to do a conditional planning, assuming that no failures occur in the planning horizon and taking the actual ages into account. This is a reasonable approach since numerical experiments show that in cases where preventive maintenance is really cost-effective, $F(t^*)$ is quite small (up to 20%). Implementing this approach on a rolling horizon basis (i.e. adapting the planning in course of time with the occurrence of events) takes care of failures. This idea was pursued in Dekker, Wildeman and Van Egmond (1993) in a discrete time case.

2.6. The delay time model

In the delay time or two-phase model, an item passes through an observable intermediate state (often called fault) before failing (see, e.g. Baker and Christer, 1994). Inspections are undertaken to see whether deterioration has progressed that far, that the intermediate state is visible. If so, a repair is carried out immediately, which is also the case upon a failure. After the repair the state is as-good-as-new. Suppose that faults occur according to a Poisson process with a rate λ per time unit and that a c.d.f. $F(\cdot)$ of the so-called delay time is available, i.e. the time left between the occurrence of the intermediate state and a failure. Let c^i, c^r and c^f denote the costs of inspection, repair in the intermediate state (to the as-good-as-new condition) and failure repair respectively. Notice that both after an inspection and repair the item is as good-as-new. Let $G(\cdot)$ be the cdf of the time between the as-good-as-new state and failure if no inspections are carried out. Accordingly, $G(t) = \int_0^t \lambda e^{-\lambda u} F(t - u) \, du$. Assume further that inspections are carried at intervals of length t since the previous inspection or repair. The deterioration cost rate $m(t)$ now amounts to $c^r \lambda + (c^f - c^r)G'(t)/(1 - G(t))$. The delay-time model is especially suited for cases where c^f is high compared to c^p and c^r, hence an asymptotic criterion w.r.t. existence of an optimum makes little sense.

2.7. Relation with the framework from Aven and Bergman

Aven and Bergman (1986) argue that the objective function in many maintenance optimisation models is of the form

$$\left(E\left[c(0) + \int_0^T a(s)h(s) \, ds\right]\right) \Big/ \left(E\left[p(0) + \int_0^T h(s) \, ds\right]\right)$$

where T is a stopping time based on the information about the condition of the system, $a(t)$ is a nondecreasing stochastic process, $h(t)$ a nonnegative stochastic process and both $c(0)$ and $p(0)$ are nonnegative r.v.'s. The expectation is taken w.r.t. all r.v.'s and stochastic processes. In our case $h(s) = 1$, T is set at a prefixed value t, $c(0)$ has the constant value c^p, $p(0) = 0$ and $a(s)$ represents the

deterioration costs $m(s)$. They show that (2.3) and (2.4) hold, but give no further results. We also consider the case that $m(s)$ is first decreasing and next increasing (the bathtub pattern). Although their framework is more general, one can not predict the replacement in advance and it is not yet clear how their framework can be used for planning and combining.

3. Elicitation of model parameters

The crucial problem to apply the framework (and in fact any maintenance optimisation model) is to obtain a functional form for either $m(t)$ or $M(t)$. If one has many data on failures and maintenance of systems at an aggregate level, one may directly fit a curve to the deterioration costs versus the system age. This approach has been followed by Kamath et al. (1984) in a study for ambulance maintenance. They fit a Duane model: $M(t) = \alpha t^{\beta}$, for some parameters α and β, to the maintenance costs.

If many data on failures and maintenance are available on a single deterioration mechanism, e.g. through tests, one may fit a statistical lifetime distribution or a stochastic process. Most models for replacement and inspection assume the knowledge of such a lifetime distribution. One has to check whether the failures were independent and what the effect is of maintenance on the deterioration mechanism(s). A problem is that the effectiveness of replacements depends on the reduction in failure rate, and not on the average rate; this requires fitting of at least two parameters.

In many cases however, there are either too few data, or data are too unreliable to be used (e.g. when failure registration does not specify the failure mode or the precise equipment part involved). In these cases one has to rely on expert judgement. Decision support systems assist in this respect, as they provide insight into the importance of data, by enabling sensitivity analyses. A major problem in the development of their userinterfaces is the modelling of deterioration in such a way that it is both suited for the problem and understandable for the expert. Asking for parameters of lifetime distributions is not advisable since a second parameter is often related to the variance and one hardly sees a complete lifetime in case of maintained items. Van Noortwijk et al. (1992) therefore propose the use of interval techniques to assess a truncated discrete lifetime distribution. The main drawback of their approach is that they only focus on severe failures causing a replacement. Quite often there are also minor failures after which a minimal repair is done, and one will have severe interpretation problems (e.g. did the failure repair renew the equipment part or did it bring it back to an as-good-as before condition?). By generalising their approach to our framework, these problems can be overcome, because we cover multiple models at the same time. We therefore suggest the following procedure.

First an appropriate interval should be selected to which experience is related, e.g. an historic maintenance interval. Next the deterioration costs are elicited for this interval and multiples or fractions of it. This can best be done in an interactive way in which the user specifies the intervals one by one and in which feedback is provided on whether the optimum lies within one of the intervals or beyond all of them. Within an interval, $m(\cdot)$ can be assumed to be linear or constant, and an extrapolation is made for the area right of the last interval, assuming e.g. linear, quadratic, polynomial or exponential shape. We will demonstrate the procedure here for two (the minimum required number) of intervals. Suppose that the historic maintenance interval was of length T. The following questions should be asked:
(a) Estimate the deterioration costs in the interval $[0, T]$.
(b) Estimate the deterioration costs in the interval $[T, T + \Delta T]$ for some ΔT (which may also be negative, depending on which side one has most experience; for negative ΔT the analysis is similar, but the expressions are slightly different).
Both questions can either be answered directly, or a split up can be made according to:
 – probability of system breakdown requiring replacement and the associated costs;
 – expected number of failures requiring a minimal repair and the associated costs;

- efficiency loss and financial consequences over the interval;
- increased probability of a hidden system failure and possible financial consequences.

This way of questioning is less model specific and leads to less interpretation problems between failures and breakdowns. Let c_a, c_b denote the answers to questions (a) and (b). Notice that $c_a = M(T)$ and that $c_b = M(T + \Delta T) - M(T)$. The next step is to assume a functional form for $m(t)$, fit it to the three values (assuming also that $M(0) = 0$) and have the user agree with the results. For example,

(i) $m(\cdot)$ piecewise constant and linear, i.e. $m(t) = c_1$, $t < T$, and $m(t) = c_1 + c_2(t - T)/\Delta T$ for $t > T$.
   We obtain: $c_1 = c_a/T$ and $c_2 = 2c_b/\Delta T - 2c_a/T$. This leads to $t^* = T\sqrt{(c_b - c_a + c^p)/(c_b - c_a)}$.

(ii) $m(\cdot)$ piecewise linear, i.e. $m(t) = c_1 t$, $t < T$, and $m(t) = c_1 T + c_2(t - T)$ for $t > T$. We obtain:
   $c_1 = 2c_a/T^2$ and $c_2 = 2\{c_b - (2c_a\Delta T)/T\}/\Delta T^2$. If $c_a > c^p$ then the optimum interval t^* equals
   $T\sqrt{c^p/c_a}$, else we have $t^{*2} = T^2 + (\Delta T)^2(c^p - c_a)/(c_b - 2c_a\Delta T/T)$.

(iii) $m(\cdot)$ linear, i.e. $m(t) = c_1 + c_2 t$, $t > 0$, and $\Delta T = T$ (to simplify formulas). We obtain $c_1 = (3c_a - c_b)/2T$ and $c_2 = (c_b - c_a)/T^2$. This yields $t^* = T\sqrt{2c^p/(c_b - c_a)}$.

It is evident that this procedure can be extended when more values for $M(t)$ are elicited. Once one agrees on a functional form, it is easy to optimise.

Example. Consider a pump which undergoes a short maintenance check (lubrication, cleaning, etc.) every year. We will elicit costs associated to this interval and its double ($T = \Delta T = 1$ yr). Suppose that there is a 4% probability that the pump fails within the first year after a check and that the failure probability increases to 6% for the second year after the check. Failure induces costs of about $1000 (mainly due to downtime) and a check costs about $40. Hence $c_a = \$40$ and $c_b = \$60$. Assuming a linear deterioration rate (see case (iii)) yields: $m(t) = \$20/\text{yr}^2 \cdot t + \$30/\text{yr}$. The optimum maintenance interval follows from $\Psi(t) = c^p$, hence $10\$/\text{yr}^2 \cdot t^2 = 40\$$, implying that $t^* = 2$ yr and $g^* = \$70/\text{yr}$. Notice also that $g(1 \text{ yr}) = \$80/\text{yr}$ and that the expected costs of deferring maintenance from year 1 to the 2nd year amount to $60 minus $80, implying that deferring is cost-effective. From the data one can directly infer that maintenance should be deferred.

Remark. This example shows how simple maintenance optimisation in fact can be. With simple means (estimation of probabilities through interval classes and separating costs into manhours, downtime and materials) the comparison can be carried out manually by the maintenance engineers themselves. It also shows that we do not need to fit a full statistical distribution to the time to failures, estimations of only two probabilities suffice. In fact a similar procedure has been used in a company to rationalise (i.e. double or halve the interval) thousands (!) of routine maintenance activities.

4. Penalty cost function for shifting from the optimum

One important aspect of the framework is that it allows the derivation of penalty costs for deviating from the individual optimal execution interval. These penalty costs are input in comprehensive models for combination of maintenance and for maintenance planning. Three different types of deviation are possible: a short-term shift, a long-term shift and finally, a permanent shift. Here we assume that a preventive maintenance activity is carried out at regular intervals of length t. The short term shift changes one interval to $t + x$, where x may be positive or negative, the next one to $t - x$, so that only one execution moment is changed. The long-term shift changes one interval to $t + x$ and all following intervals remain constant. Finally, the permanent shift changes all intervals to $t + x$. Which shift is most appropriate, depends on how the preventive maintenance program is incorporated in the maintenance

management system and whether the shifts are permanent or not. If the maintenance management system calculates all future execution dates from the initially planned dates, then the short-term shift is appropriate. If it does so from the actual execution dates, then the long-term shift should be used. From the deterioration costs $M(\cdot)$ penalty cost functions for each of the shifts can be derived. Let $h_S(x)$, $h_L(x)$ and $h_P(x)$ denote the penalty functions for deviating x time units from the optimum t^* for a short-term shift, long-term shift and permanent shift respectively. It is easy to see that

$$h_S(x) = M(t^* + x) + M(t^* - x) - 2M(t^*), \tag{4.1}$$

$$h_L(x) = M(t^* + x) - M(t^*) - xg^* = \int_{t^*}^{t^* + z}(m(y) - g^*)\, dy, \tag{4.2}$$

$$h_P(x) = g(t^* + x) - g(t^*) = h_L(x)/(t^* + x), \tag{4.3}$$

where g^* denotes the minimum long term average costs. These penalty functions can not only be used to assess the cost-effectiveness of any special sales offer, but also for priority setting and to assist in combining activities or integrating maintenance planning with production planning. Notice that the penalty functions have the following properties: they are always nonnegative and they are zero for $x = 0$. Furthermore, $h_S(\cdot)$ is symmetric round zero.

These penalty functions indicate the expected cost for deviating from the optimum interval. It may happen, however, that the present state already deviates from the optimum and that one does not need to take the costs into account for arriving in the present state, but that one is interested in the extra costs for deviating even further. More specifically, suppose one is at t time units, $t > t^*$ since the last execution of the activity. The expected costs for deferring (in this case there is no other option) the activity for another x time units amount to (we only consider the long-term shift)

$$h_L^t(x) = M(t + x) - M(t) - xg^* = \int_t^{t+z}(m(y) - g^*)\, dy, \quad x > 0. \tag{4.4}$$

In case of the extended framework we have to condition on the present age and only consider the case where the component survives. Hence the penalty costs for deferring preventive maintenance at age t to age $t + x$, where it is normally executed at age t^* (only the long-term shift is relevant) amount to

$$h_L^t(x) = \int_t^{t+x}(m(y) - g^*)\frac{(1 - F(y))}{(1 - F(t))}\, dy, \quad x > 0. \tag{4.5}$$

5. Combining execution of activities

One way of reducing maintenance costs is to combine the execution of maintenance activities. In many cases preparatory work, such as shutting down a unit, scaffolding, travelling of the maintenance crew, has to take place before maintenance can be done. Combining activities allows savings on this work. On the other hand, combining mostly implies that one deviates from the originally planned execution moments, which is not free. Combining activities can both be done on a long-term (e.g. creating maintenance packages) or on a short-term, taking all once-off combinations into account. Here we will consider short-term combining and show that the penalty functions derived in the previous section allow a cost-effectiveness evaluation of combinations and assist in the timing of the execution. The main idea is to apply a decomposition approach, that is, we first determine for each activity its preferred execution moment and derive its penalty function. Next we consider groups of activities, for which the preferred moment of execution follows from a minimisation of the sum of the penalty function involved. If this sum

Table 1
Example data for combining 8 activities

Activity	λ_i (days)	β_i	c_i^{r}	c_i^{p}	t_i (days)	x_i^* (days)
1	2380	1.70	46	60	0	229
2	2380	1.70	91	120	15	230
3	1900	2.00	14	180	32	681
4	2850	2.00	15	90	60	698
5	1620	1.70	86	300	100	278
6	2850	2.00	15	180	160	987
7	1950	1.75	45	60	180	195
8	1350	1.75	25	180	212	354

is less than the set-up savings because of a joint execution, combining is cost-effective. Corrective maintenance work can also be involved in the combination, provided that it is known at the outset of planning, deferrable and a penalty function for deferring can be assessed. Determining the optimal groups can be formulated as a set-partitioning problem (see Dekker, Smit and Loosekoot, 1992). Wildeman, Dekker and Smit (1992) show that under certain conditions the optimal grouping consists of groups with consecutive initial planning moments, which allows the formulation of an $O(n^2)$ dynamic programming algorithm (n being the number of activities).

Example. Table 1 provides data on 8 maintenance activities, which each replace a unit. Deterioration costs of unit i are primarily due to small failures upon which a minimal repair is done. These occur independently of the state of other units and the cost rate amounts to: $m_i(x_i) = c_i^{\mathrm{r}}100(\beta_i/\lambda_i)(x_i/\lambda_i)^{\beta_i-1}$, where x_i denotes units i's age. Special case (i) (see Section 2.2) gives a formula for the individually optimal replacement age, which we denote by x_i^*. Finally let t_i be the resulting initial planning moment (counted from the start of the planning horizon).

The resulting penalty functions are shown in Fig. 1 (the numbers refer to the activities).

We consider combining under short-term shifts, in which case the penalty costs are given by Eq. (4.1). The planning horizon is $[0, 220]$. We assume that combining execution of k activities saves $k - 1$ times
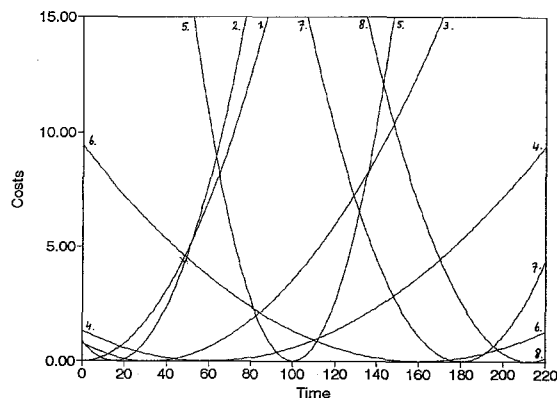

Fig. 1. Penalty cost functions.

the set-up work (for any k), which is estimated at 15 cost units (about 10% of the preventive maintenance costs of an activity). Using the algorithm of Wildeman, Dekker and Smit yields as optimal groupings: {1, 2, 3} executed at day 12.6, {4, 5} at day 97.9 and {6, 7, 8} at day 192.9. The savings (set-up cost reduction minus penalty costs) for the combinations amount to 29.4, 14.4 and 28.2 respectively. Total savings amount to 72.0, which constitutes 6% of total preventive maintenance costs.

Dekker, Wildeman and Van Egmond (1993) give an analysis of the performance of this combination method for a more complex case where components are replaced using a discrete time age replacement. They apply a conditional planning (assuming no failures in the planning horizon) on a rolling horizon basis (implement the decision for the current epoch, observe the new state at the next epoch and make a new planning). They use the discrete version of the penalty functions (4.5). They consider combining both for a finite and infinite horizon and compare their planning method with an optimal solution obtained by solving a large scale Markov decision chain numerically (which was tractable up to four identical components only). It appears that for high set-up costs and many components the component decomposition has to be changed because components are almost always replaced together. When that has been done the loss of their strategy compared to the optimal one is less than 1%.

6. Priority setting

Maintenance is usually classified into corrective and preventive work. The first originates from a directly foreseeable, or already observed malfunctioning of systems, and the latter from a preconceived plan to keep systems in a good condition in the long run. Often the first type of work is the most urgent one. The maintenance capacity needed to take care of that may fluctuate severely in time, due to the random character of failures. Hence preventive work is often delayed in favour of the corrective work. Accordingly, there is usually a large backlog of preventive work, with the implication that an individual preventive maintenance activity is either delayed for an unknown time or even never carried out. Most maintenance organisations have problems in managing the backlog. It will be clear that the results of maintenance optimisation decrease in value if the maintenance organisation is not able to do the work on time, which is especially a problem for the many small maintenance activities. Priority criterion functions, embedded in management information systems, can be of help.

Here we propose the use of the penalty functions $h_L(x)$ (or $h_S(x)$ if appropriate, see Section 4) as priority functions, where the long-term objective is the average costs. Although they are formulated for a continuous time setting, where at each moment a decision can be taken, they can easily be extended to discrete intervals between decision moments, depending on how often one wants to reset priorities. The same holds for a discounted costs objective function. Before we give some pro's and cons of these functions, we first introduce some other priority criteria which have been used in practice (see Pintelon, 1990, for a review):

 (i) a fixed priority according to the importance of the machine to be maintained;

 (ii) a machine importance factor multiplied by the waiting time for execution.

It will be clear that these criterion functions are heuristic and not related to an optimisation model. The penalty function $h_L(x)$ on the other hand, has the following properties: it is negative before the optimal execution time, zero exactly at that time and increases thereafter. It is expressed in money terms, has an easy interpretation and is *additive*. The latter means that the priority criterion for a group is the sum of the individual priority criteria. Hence, splitting up activities into smaller activities does not effect the priority for the group. This is not the case for the criteria of type (i) and (ii): these are expressed on a cardinal scale only! The penalty based criteria can therefore also be used for the groups which are the result of the combination of activities (just use the sum of the penalty functions minus their minimum value). Furthermore, they can used in more sophisticated planning (see the next section). We like to

remark that the penalty based priority criteria only express how important it is to execute a certain activity. It does not express how much of a scarce resource is needed for execution.

Setting priorities between corrective maintenance and preventive is in principle possible using the above ideas, since the priority functions are expressed in money. Corrective maintenance has to be separated into deferrable and nondeferrable and for the first category a cost rate for deferring has to be estimated, which can be compared with the priority function for preventive maintenance.

Dekker and Smeitink (1994) provide an analysis of similar priority criteria, though for a case where preventive maintenance could only be executed at randomly occurring opportunities of restricted duration. Hence at each opportunity priorities for execution had to be set. They computed long-term average costs for a twenty-four component system under four different priority criteria, including one based on the penalty functions. This one performed best in all cases considered.

Finally, in case of the extended framework for the age replacement model (see Section 2.5) we may use a similar priority function. Given a present age of x, the expected cost rate of delaying the maintenance activity amounts to $m(x)$ whereas we save by deferring on the average g^* per time unit. Hence the priority function reads $m(t) - g^*$ for the continuous time case.

The same reasoning has been followed to derive a replacement criterion for a maintenance package consisting of multiple preventive maintenance activities which could only be executed together, whereas upon failure only the respective activity was carried out. As in that case the minimal average costs g^* are very difficult to compute, they were approximated by those of a block replacement policy, g_b^*. Such a marginal cost reasoning yielded a simple but very effective replacement criterion which performed only slightly worse than the (complex) optimal policy (see Dekker and Roelvink, 1995).

7. Maintenance planning

In this section we will show that the penalty functions can be used as basic elements in sophisticated versions of maintenance planning, using a rolling horizon. The difference with the previous section is that we now provide an explicit modelling of the planning, including a single capacity constraint, but this can easily be extended to multiple constraints. The approach we propose, is to apply a decomposition, in which one first optimises at the individual maintenance activity level and then matches the results with the higher level constraints. Although the decomposition approach is not exact (in the first stage one neglects the higher level constraints), one obtains a more stable solution and insight into the answer. In case the higher level constraints are very tight or if considerable cost savings can be obtained in combining, one may have to take these effects into account in the optimisation for each individual activity (see also Dekker, Wildeman and Van Egmond, 1993, for a discussion of this effect). The proposed way of planning is also especially suited for visual interactive modelling, as the effect of changing the planning of each activity follows directly from the penalty functions.

We consider constraints on the execution moments and on the available maintenance capacity. Let the planning interval be $[0, T]$, in which activities $1, \ldots, n$ are to be planned. Suppose time windows $W_1, \ldots, W_m$ are available for maintenance and that no planning within a window is necessary at this stage. Hence if an activity is carried out in window W_j we assume it is carried out at its midpoint, say t_j. Let K_j be the available maintenance capacity in window j. Window m, at time T, is unrestricted in capacity; it therefore represents the backlog of the activities to be carried at the end of the planning horizon. In the presence of constraints the optimal solution no longer bears the consecutiveness property. Hence we will formulate an integer programming problem. Let the decision variable x_{ij} be 1 if activity i is carried out during window j and 0 else. Suppose that $t_i^{(0)}$ is the preferred execution moment of activity i when restricted to the time windows t_j, $j = 1, \ldots, m$ (in the uncapacitated case) and that $h_i(\cdot)$ indicates the associated penalty cost function for activity i (w.r.t. $t_i^{(0)}$). Define c_{ij} as the value of this penalty function

at t_j, i.e. $c_{ij} \equiv h_i(t_j - t_i^{(0)})$. Finally, let k_i denote the capacity requirements for activity i. We now have the following integer programming problem for the planning of maintenance under constraints:

$$\min \quad \sum_{i,j} c_{ij} x_{ij}$$

$$\text{st.} \quad \sum_j x_{ij} = 1, \quad i = 1, \ldots, n \qquad \text{(every activity is executed once)}$$

$$\sum_i x_{ij} k_i \leq K_j, \quad j = 1, \ldots, m \qquad \text{(capacity constraint in window } j\text{)}$$

$$x_{ij} \in \{0, 1\}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m.$$

The solution for the uncapacitated case is straightforward: $x_{ij} = 1$ if $t_j = t_i^{(0)}$ and zero else; apart from $x_{.T}$, it does not depend on the planning horizon. Other formulations which apply deterioration costs instead of penalty costs, such as in Stinson and Khumawala (1987), have a solution which does depend on the planning horizon. This is only correct if a system is indeed terminated at the end of the planning horizon. Justifying preventive replacements only over the planning horizon is not correct, as their effect may last much longer. Else, if in fact a rolling horizon is applied, one either has to compensate for the state at the end of the horizon, or to apply our procedure. The first may still give a very unstable policy because of horizon effects, whereas our procedure more closely follows a long-term policy (see, e.g. Dekker et al., 1993). A further advantage of our formulation is that the uncapacitated solution can serve as starting point for heuristics:

(i) A single stage heuristic: consider the windows successively. If in window j the required capacity exceeds the available, delay those activities which have the lowest $(c_{ij+1} - c_{ij})/k_i$ ratio (penalty increase over capacity use).

(ii) As (i) but also consider forwarding activities one window, in case of spare capacity.

As it is our main intention to show that a maintenance planning problem can be formulated with the penalty functions, we will not go into details about solving the problem. A final remark is that in practical applications it is mostly not the optimal solution which counts, but the insight which is obtained in a structured planning process. Multiple (not modelled) objective criteria and uncertainty about e.g. execution durations may make optimal solutions less desirable. The 'optimal solution' is then only used for comparison reasons.

8. Conclusions

In this paper we presented a framework for optimisation models which allows integration with priority setting, planning and combination of activities. It is also useful in setting up an appropriate elicitation procedure in case model parameters are based on expert judgement. The latter is important for developing decision support systems for maintenance optimisation. Further research is required to investigate whether more models can be incorporated into the framework, and whether other models can be converted to allow combining and planning as done in this paper.

References

Aven, T., and Bergman, B. (1986), "Optimal replacement times – A general set-up", *Journal of Applied Probability* 23, 432–442.

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

Barlow, R.E., and Proschan, F. (1965), *Mathematical Theory of Reliability*, Wiley, New York.

Berg, M. (1980), "A marginal cost analysis for preventive replacement policies", *European Journal of Operational Research* 4, 136–142.

Berg, M. (1995), "The marginal cost analysis and its application to repair and replacement policies", *European Journal of Operational Research* 82, 214–224.

Berg, M., and Cléroux, R. (1982), "A marginal cost analysis for an age replacement policy for units with minimal repair", *Infor* 20, 258–263.

Berg, M., and Epstein, B. (1979), "A note on a modified block replacement strategy with increasing running costs", *Naval Research Logistics Quarterly* 26, 157–159.

Cho, D.I., and Parlar, M. (1991), "A survey of maintenance models for multi-unit systems", *European Journal of Operational Research* 51, 1–23.

Dekker, R. (1992), "Applications of maintenance optimisation models: A review and analysis", Report Econometric Institute 9228/A, Erasmus University Rotterdam, Netherlands.

Dekker, R., and Dijkstra, M.C. (1992), "Opportunity-based age replacement: exponentially distributed times between opportunities", *Naval Research Logistics* 39, 175–190.

Dekker, R., and Roelvink, I.F.K. (1995), "Marginal cost criteria for preventive replacement of a group of components", *European Journal of Operational Research* 84/2, to appear.

Dekker, R., and Smeitink, E. (1991), "Opportunity-based block replacement: The single component case", *European Journal of Operational Research* 53, 46–63.

Dekker, R., and Smeitink, E. (1994), "Preventive maintenance at opportunities of restricted duration", *Naval Research Logistics* 41, 335–353.

Dekker, R., Smit, A.C.J.M., and Loosekoot, J.E. (1992), "Combining maintenance activities in an operational planning phase", *IMA Journal of Mathematics Applied in Business and Industry* 3, 315–332.

Dekker, R., Wildeman, R.E., and Van Egmond, R., (1993), "Joint replacement in an operational planning phase", Report Econometric Institute 9438/A (revised version), Erasmus University Rotterdam, Netherlands.

Kamath, A.R.R., Al-Zuhairi, A.M., Keller, A.Z., and Selman, A.C. (1984), "A study of ambulance reliability in a metropolitan borough", *Reliability Engineering* 9, 133–152.

McCall, J.J. (1965), "Maintenance policies for stochastically failing equipment: A survey", *Management Science* 11, 493–524.

Noortwijk, J.M. van, Dekker, R., Cooke, R.M., and Mazucchi, T.A. (1992), "Expert judgment in maintenance optimization", *IEEE Transactions on Reliability* 41, 427–432.

Pierskalla, W.P., and Voelker, J.A. (1979), "A survey of maintenance models: The control and surveillance of deteriorating systems", *Naval Research Logistics Quarterly* 23, 353–388.

Pintelon, L. (1990), "Performance reporting and decision tools for maintenance management", Ph.D. Thesis, University of Leuven.

Sherif, Y.S., and Smith, M.L. (1981), "Optimal maintenance models for systems subject to failure – A review", *Naval Research Logistics Quarterly* 28, 47–74.

Stinson, J.P., and Khumawala, B.M. (1987), "The replacement of machines in a serially dependent multi-machine production system", *International Journal of Production Research* 25, 677–688.

Valdez Flores, C., and Feldman, R.M. (1989), "A survey of preventive maintenance models for stochastically deteriorating single unit systems", *Naval Research Logistics Quarterly* 36, 419–446.

Wildeman, R.E., Dekker, R., and Smit, A.C.J.M. (1992), "Combining activities in an operational planning phase: A dynamic programming approach", Report Econometric Institute 9424/A (revised version), Erasmus University Rotterdam, Netherlands.