Optimal Replacement Decisions for Structures under Stochastic Deterioration[1]

Jan M. van Noortwijk

HKV Consultants, P.O. Box 2120, 8203 AC Lelystad, The Netherlands, and
Delft University of Technology, Faculty of Mathematics and Computer Science,
P.O. Box 5031, 2600 GA Delft, The Netherlands.

1 Introduction

The subject of this paper is determining optimal replacement decisions for components under stochastic deterioration. Using the discrete renewal theorem, three cost-based criteria for comparing maintenance decisions over unbounded horizons are determined in Sec. 2: (i) the average costs per unit time, (ii) the discounted costs over an unbounded horizon, and (iii) the equivalent average costs per unit time. By using these criteria, the cost of preventive maintenance can be balanced against the cost of corrective maintenance. In structural engineering, a distinction can often be made between a structure's resistance and its design stress. A failure may then be defined as the event in which - due to deterioration - the resistance drops below the design stress or the failure level. Since deterioration is uncertain, it can best be regarded as a stochastic process. On the basis of a stochastic gamma process, the probabilities of failure per year, i.e. the probabilities that the resistance drops below the failure level per year, are calculated in Sec. 3. A case study on the maintenance of cylinders shows the usefulness of the replacement model in Sec. 4.

2 Cost-based Criteria for Comparing Maintenance Decisions

Usually, *maintenance* is defined as a combination of actions carried out to restore the structure's component to, or to "renew" it to, the original condition. Inspections, repairs, replacements, and lifetime-extending measures are possible maintenance actions. Through lifetime-extending measures, the deterioration can be delayed as such that failure is postponed and the component's lifetime is extended. Roughly, there are two types of maintenance: *corrective maintenance* (mainly after failure) and *preventive maintenance* (mainly

[1]This work was partly supported by the Civil Engineering Division of the Dutch Ministry of Transport, Public Works, and Water Management. This paper was published in Andrzej S. Nowak, editor, *Proceedings of the Eighth IFIP WG 7.5 Working Conference on Reliability and Optimization of Structural Systems, Kraków, Poland, 1998*, pages 273-280. Ann Arbor: University of Michigan, 1998.

before failure). Corrective maintenance can best be chosen if the cost arising from failure is low (like for instance replacing a burnt-out light bulb); preventive maintenance if this cost is high (like for instance heightening a dyke). In structural engineering, the consequences of failure are generally so large that mainly expensive preventive maintenance is applied. The use of maintenance optimisation models is therefore of considerable interest.

During the lifetime of a structure, we can roughly identify four phases: the design, the building, the use, and the demolition. There are mainly two phases in which it is worth applying maintenance optimisation techniques: (i) the *design* phase and (ii) the *use* phase. In the design phase, the initial cost of investment has to be balanced against the future cost of maintenance. In the use phase, the cost of inspection and preventive replacement have to be balanced against the cost of corrective replacement and failure.

Since the planned lifetime of most structures is very long, maintenance decisions may be compared over an unbounded time-horizon. According to Wagner [11, Ch. 11], there are basically three cost-based criteria that can be used to compare maintenance decisions:

1. the *expected average costs per unit time*, which are determined by averaging the costs over an unbounded horizon;
2. the *expected discounted costs over an unbounded horizon*, which are determined by summing the (present) discounted values of the costs over an unbounded horizon, under the assumption that the value of money decreases in time;
3. the *expected equivalent average costs per unit time*, which are determined by calculating the discounted costs per unit time.

The notion of equivalent average costs relates to the notions of average costs and discounted costs. The cost-based criteria of discounted costs and equivalent average costs are most suitable for balancing the initial building cost optimally against the future maintenance cost. The criterion of average costs can be used in situations in which no large investments are made (like inspections) and in which the time value of money is of no consequence to us. Often, it is preferable to spread the costs of maintenance over time and to use discounting.

Examples of optimising maintenance in the design phase are: determining optimal dyke heightenings and optimal sand nourishments whose expected discounted costs are minimal (see Speijker et al. [5] and Van Noortwijk & Peerbolte [8], respectively). Examples of optimising maintenance in the use phase are: determining cost-optimal rates of inspection for dykes, berm breakwaters, and the sea-bed protection of the Eastern-Scheldt barrier (see Van Noortwijk et al. [6, 10, 7, 9]), and determining cost-optimal preventive maintenance intervals (see Sec. 4 of this paper).

The maintenance of structures can often be modelled as a discrete renewal process, whereby the renewals are the maintenance actions that bring a component back into its original condition or "as good as new state". After each renewal we start, in statistical sense, all over again. A discrete renewal process $\{N(n) : n \in \mathbb{N}\}$ is a non-negative integer-valued stochastic process that registers the successive renewals in the time-interval $(0, n]$. Let the renewal times

$T_1, T_2, \ldots,$ be non-negative, independent, identically distributed, random quantities having the discrete probability function $\Pr\{T_k = i\} = p_i(d)$, $i \in \mathbb{N}$, with $\sum_{i=1}^{\infty} p_i(d) = 1$, where $p_i(d)$ represents the probability of a renewal in unit time i when the decision-maker chooses maintenance decision d. We denote the costs associated with a renewal in unit time i by $c_i(d)$, $i \in \mathbb{N}$. The above-mentioned three cost-based criteria will be discussed in more detail in the following subsections.

The expected average costs per unit time.
The expected average costs per unit time are determined by simply averaging the costs over an unbounded horizon. They follow from the expected costs over the bounded horizon $(0, n]$, denoted by $C(n, d)$, which solve the recursive equation

$$C(n,d) = \sum_{i=1}^{n} p_i(d)[c_i(d) + C(n-i,d)] \tag{1}$$

for $n \in \mathbb{N}$ and $C(0, d) \equiv 0$, when the decision-maker chooses maintenance decision d. To obtain this equation, we condition on the values of the first renewal time T_1 and apply the law of total probability. The costs associated with occurrence of the event $T_1 = i$ are $c_i(d)$ plus the additional expected costs during the interval $(i, n]$, $i = 1, \ldots, n$. Using the discrete renewal theorem (see Feller [1, Ch. 12 & 13] and Karlin & Taylor [3, Ch. 3]), the expected *average costs per unit time* are

$$\lim_{n\to\infty} \frac{C(n,d)}{n} = \frac{\sum_{i=1}^{\infty} c_i(d) p_i(d)}{\sum_{i=1}^{\infty} i p_i(d)} = C(d). \tag{2}$$

Let a renewal cycle be the time-period between two renewals, and we recognise the numerator as the expected cycle costs and the denominator as the expected cycle length (mean lifetime). Eq. (2) is a well-known result from renewal reward theory (see e.g. Ross [4, Ch. 3]). If $c_i(d) \equiv 1$ for all $i \in \mathbb{N}$ in Eq. (2), then the expected *average number of renewals per unit time* is the reciprocal of the mean lifetime.

The expected discounted costs over an unbounded horizon.
Discounting expected costs over an unbounded horizon is based on the assumption that the utility of money decreases in time from the standpoint of the present. Since the future cost can be discounted to its present value on the basis of a discount rate, we can compare the value of money at different dates. In mathematical terms, the *(present) discounted value* of the costs c_n in unit time n is defined to be $\alpha^n c_n$ with $\alpha = [1 + (r/100)]^{-1}$ the *discount factor* per unit time and $r\%$ the *discount rate* per unit time, where $r > 0$. The decision-maker is indifferent to the costs c_n at time n and the costs $\alpha^n c_n$ at time 0. Therefore, the higher the discount rate, the better it is to postpone expensive maintenance actions.

The expected discounted costs over a bounded time-horizon can be obtained with a recursive formula similar to that of the expected costs in Eq. (1). Again, we condition on the values of the first renewal time T_1 and apply the law of total probability. In this case, however, we want to account for the discounted value of the renewal costs $c_i(d)$ plus the additional expected discounted costs in time-interval $(i,n]$, $i=1,\ldots,n$. Hence, the expected discounted costs over the bounded horizon $(0,n]$ can be written as

$$C_\alpha(n,d)=\sum_{i=1}^{n}\alpha^i p_i(d)\,[c_i(d)+C_\alpha(n-i,d)] \tag{3}$$

for $n\in\mathbb{N}$, and $C_\alpha(0,d)\equiv 0$, when the decision-maker chooses maintenance decision d. By using Feller [1, Ch. 13], the expected *discounted costs over an unbounded horizon* $C_\alpha(d)$ can be written as

$$C_\alpha(d)=\lim_{n\to\infty}C_\alpha(n,d)=\frac{\sum_{i=1}^{\infty}\alpha^i c_i(d)p_i(d)}{1-\sum_{i=1}^{\infty}\alpha^i p_i(d)}. \tag{4}$$

We recognise the numerator of $C_\alpha(d)$ as the expected discounted cycle costs, while the denominator can be interpreted as the probability that the renewal process terminates due to discounting. Such a renewal process is called a *terminating renewal process* since infinite inter-occurrence times can cause the renewals to cease. The inter-occurrence times $Z_1,Z_2,\ldots$, of our imaginary terminating renewal process have the distribution $\Pr\{Z_k=i\}=\alpha^i p_i(d)$, $i\in\mathbb{N}$, and $\Pr\{Z_k=\infty\}=1-\sum_{i=1}^{\infty}\alpha^i p_i(d)$.

The expected equivalent average costs per unit time.
The expected equivalent average costs per unit time relate to the two notions of average costs and discounted costs. To determine this relation, we construct a new infinite stream of identical costs with the same present discounted value as the expected discounted costs over an unbounded time-horizon $C_\alpha(d)$. This can be easily achieved by defining an infinite stream of costs appearing at times $i=0,1,2,\ldots$, which are all equal to $(1-\alpha)C_\alpha(d)$. Using the geometric series, we can write

$$\textstyle\sum_{i=0}^{\infty}\alpha^i(1-\alpha)C_\alpha(d)=C_\alpha(d) \tag{5}$$

when the decision-maker chooses maintenance decision d, $0<\alpha<1$. We call $(1-\alpha)C_\alpha(d)$ the *equivalent average costs per unit time.* As α tends to 1, from below, the equivalent average costs approach the average costs per unit time:

$$\lim_{\alpha\uparrow 1}(1-\alpha)C_\alpha(d)=C(d), \tag{6}$$

using L'Hôpital's rule.

The initial cost of investment.
For cost-optimal investment decisions, we are interested in finding an optimum balance between the initial cost of investment and the future cost of maintenance, being the area of life cycle costing. In this situation, the monetary losses over an unbounded horizon are the sum of the initial cost of investment $c_0(d)$ and the expected discounted future cost $C_\alpha(d)$:

$$L_\alpha(d)=c_0(d)+C_\alpha(d), \tag{7}$$

when the decision-maker chooses decision d, and the discount factor is α, $0<\alpha<1$. The corresponding expected equivalent average costs per unit time are $(1-\alpha)L_\alpha(d)$. For investment decisions, we cannot use the criterion of the expected average costs per unit time,

$$L(d) = \lim_{n \to \infty} \frac{c_0(d)}{n} + C(d) = C(d), \tag{8}$$

because the contribution of the initial cost to the average costs is completely ignored.

3 Stochastic Deterioration and Time of Failure

A difficulty in modelling maintenance is that the process of deterioration and the time of failure are uncertain. In structural engineering, a distinction can often be made between a structure's resistance (e.g. the crest-level of a dyke) and its design stress (e.g. the maximal water level to be withstood). A failure may then be defined as the event in which - due to deterioration - the resistance drops below the design stress or the failure level.

Stochastic deterioration.
Since deterioration is uncertain, it can best be regarded as a stochastic process. At first glance, it seems possible to represent the uncertainty in a deterioration process by the normal distribution. This probability distribution has been used for modelling the exchange-value of shares and the movement of small particles in fluids and air. A characteristic feature of this model, also denoted by the Brownian motion with drift, is that a structure's resistance alternately increases and decreases, like the exchange-value of a share. For this reason, the Brownian motion is inadequate in modelling deterioration which proceeds in one direction. For example, a dyke of which the height is subject to a Brownian deterioration can, according to the model, spontaneously rise up, which of course cannot occur in practice!

In order for the stochastic deterioration process to proceed in one direction, we can best consider it as a so-called 'gamma process'. In mathematical terms, a gamma process is a stochastic process with independent non-negative increments (e.g. the increments of crest-level decline of a dyke) having a gamma distribution with a known average rate of deterioration. In the case of a gamma deterioration, dykes can only sink. An advantage of modelling deterioration processes through gamma processes is that the required mathematical calculations are relatively straightforward. Recall that a random quantity X has a gamma distribution with shape parameter $a > 0$ and scale parameter $b > 0$ if its probability density function is given by:

$$\text{Ga}(x|\,a,b) = [b^a/\Gamma(a)]\,x^{a-1}\exp\{-bx\}\,I_{(0,\infty)}(x),$$

where $I_A(x) = 1$ for $x \in A$ and $I_A(x) = 0$ for $x \notin A$, and $\Gamma(a) = \int_{t=0}^{\infty} t^{a-1}e^{-t}\,dt$ is the gamma function for $a > 0$. A gamma process with stationary increments is defined as follows. The gamma process with shape function $at > 0$, $t \geq 0$, and scale parameter $b > 0$, is a continuous-time stochastic process $\{Y(t) : t \geq 0\}$ with the following properties:

1. $Y(0) = 0$ with probability one;
2. $Y(\tau) - Y(t) \sim \text{Ga}(a(\tau - t), b)$ for all $\tau > t \geq 0$;
3. $Y(t)$ has independent increments.

The characteristic function of the gamma distribution Ga(a, b), which is given by

$$\phi(u) = [b/(b - iu)]^a = \exp\left\{\int_0^\infty (e^{iux} - 1)\; dM(x)\right\}$$

where $M(x) = -\,a\int_x^\infty (e^{-by}/y)\,dy$ for $x > 0$, shows us that the gamma process is an integral of compound Poisson processes with jump intensity $M(x)$ (see Gnedenko & Kolmogorov [2, pp. 86-87]). Hence, the gamma process is a pure jump process. Let $X(t)$ denote the amount of deterioration at time t, $t \geq 0$, and let the probability density function of $X(t)$ be given by

$$p_{X(t)}(x) = \text{Ga}\left(x \,\middle|\, [\mu^2 t]/\sigma^2, \mu/\sigma^2\right) \tag{9}$$

for $\mu, \sigma > 0$ with

$$E(X(t)) = \mu t, \qquad \text{Var}(X(t)) = \sigma^2 t.$$

Due to the stationarity of the above stochastic deterioration process, both the mean value and the variance of the deterioration are linear in time. For expected deterioration being non-linear rather than linear in time, we refer to Van Noortwijk & Klatter [10].

Stochastic time of failure.
A component is said to fail when its deterioration exceeds a certain failure level, say y, where y is defined as the initial resistance r_0 minus the failure level s. Let the time at which the failure level is crossed be denoted by the lifetime T (in years). Due to the gamma distributed deterioration, Eq. (9), the lifetime distribution can then be written as:

$$F(t) = \Pr\{T \leq t\} = \Pr\{X(t) \geq y\} = \int_{x=y}^{\infty} p_{X(t)}(x)\,dx = \frac{\Gamma([\mu^2 t]/\sigma^2, [y\mu]/\sigma^2)}{\Gamma([\mu^2 t]/\sigma^2)}, \tag{10}$$

where $\Gamma(a, x) = \int_{t=x}^\infty t^{a-1}e^{-t}\,dt$ is the incomplete gamma function for $x \geq 0$ and $a > 0$. The probability of failure per year, denoted by p_i, $i = 1, 2, 3, \ldots$, follows immediately from Eq. (10):

$$p_i = F(i) - F(i-1), \qquad i \;=\; 1, 2, 3, \ldots \tag{11}$$

A useful property of the gamma process is that the gamma density in Eq. (9) transforms into an exponential density if $t = (\sigma/\mu)^2$. When the unit-time length is chosen to be $(\sigma/\mu)^2$, the increments of deterioration are exponentially distributed with mean σ^2/μ and the probability of failure in unit time i reduces to a Poisson distribution (see e.g. Van Noortwijk et al. [6]):

$$p_i = \frac{1}{(i-1)!}\left[\frac{y\mu}{\sigma^2}\right]^{i-1}\exp\left\{-\frac{y\mu}{\sigma^2}\right\}, \qquad i \;=\; 1, 2, 3, \ldots \tag{12}$$

This unit time facilitates the algebraic manipulations considerably and, moreover, often results in a very good approximation of the optimal decision (see the example of Sec. 4).

4 Example: Maintenance of a Cylinder

A well-known preventive maintenance strategy is the age replacement strategy. Under an age replacement policy, a replacement is carried out at age k (preventive replacement) *or* at failure (corrective replacement), whichever occurs first, where $k = 1, 2, 3, \ldots$ A preventive replacement entails a cost c_P, whereas a corrective replacement entails a cost c_F, where $0 < c_P \leq c_F$.

As a simplified example, we study the maintenance of a cylinder on an existing swing bridge. Preventive maintenance of a cylinder mainly consists of replacing the guide bushes and plunger, and replacing the packing of the piston rod. In the event of corrective maintenance, the cylinder has to be replaced completely because too much damage has occurred. The cost of preventive maintenance is $c_P = 30,000$ Dutch guilders, whereas the cost of corrective maintenance is $c_F = 100,000$ Dutch guilders. Both maintenance actions bring the cylinder back into its "as good as new state". The expected deterioration is assumed to degrade linearly in time from the initial condition of 100% down to the failure level of 0%. The rate of deterioration is based on periodic lifetime-extending maintenance, in terms of cleaning and sealing the cylinder, at a frequency of once in $w = 5$ years. The cost of lifetime-extending maintenance is approximately $c_L = 20,000$ Dutch guilders. The time at which the expected condition equals the failure level is 15 years with uncertainty parameters $\mu = 6.67$ and $\sigma = 1.81$. Suppose either a preventive replacement or a corrective replacement is carried out at unit time i, then the cylinder's lifetime is extended at the units of time $w, 2w, 3w, \ldots, \lfloor (i-1)/w \rfloor w$, where $\lfloor x \rfloor$ is the largest integer less than or equal to x. Using the geometric series, the discounted costs of a lifetime extension at unit time i can be written as

$$\alpha^{1w} c_L + \alpha^{2w} c_L + \ldots + \alpha^{\lfloor (i-1)/w \rfloor w} c_L = \frac{\alpha^w - \alpha^{(\lfloor (i-1)/w \rfloor + 1)w}}{1 - \alpha^w} c_L = \chi_{iw} c_L,$$

$w = 1, 2, 3 \ldots$. Then, it follows from Eq. (4) that the expected discounted costs of lifetime

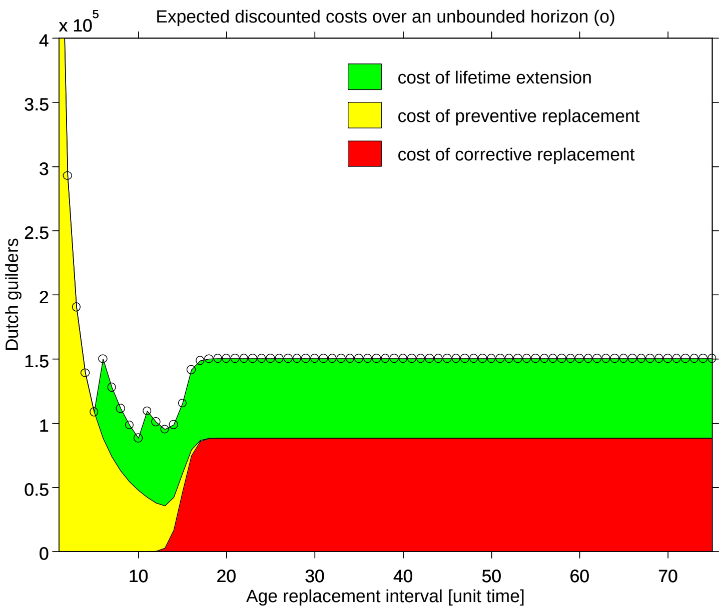

Figure 1: Maintaining a cylinder on an existing swing bridge. The expected discounted costs over an unbounded horizon as a function of the age replacement interval k, $k = 1, \ldots, 75$.

extension and age replacement over an unbounded horizon are

$$C_\alpha(k) = \lim_{n\to\infty} C_\alpha(n,k) = \frac{\sum_{i=1}^k \left[\chi_{iw}c_L + \alpha^i c_F\right] p_i + \left[\chi_{kw}c_L + \alpha^k c_P\right]\left(1 - \sum_{i=1}^k p_i\right)}{1 - \left[\left(\sum_{i=1}^k \alpha^i p_i\right) + \alpha^k \left(1 - \sum_{i=1}^k p_i\right)\right]},$$

where $C_\alpha(n,k)$ are the expected discounted costs in the bounded time-interval $(0,n]$. The optimal age replacement interval k^* is an interval for which the expected discounted costs over an unbounded horizon are minimal, i.e. for which $C_\alpha(k^*) = \min_{k=1,2,3,\ldots} C_\alpha(k)$. Note that the replacement model can also be applied for determining the optimal initial resistance of a structure, which balances the initial cost of investment c_P optimally against the future cost of maintenance $C_\alpha(k^*)$.

On the basis of an annual discount rate of 5%, the expected discounted costs of life-time extending, preventive and corrective maintenance are displayed in Fig. 1. The expected discounted costs over an unbounded horizon are minimal for a preventive replacement interval of $k^* = 10$ years. If the cost of lifetime-extending maintenance is *not* taken into account, then the optimal preventive replacement interval is 13 years. The latter result also follows when the expected discounted costs are calculated with respect to units of time of length $(\sigma/\mu)^2 = 0.074$ and a discount factor of $\alpha^{0.074} = 0.9964$ per unit time (see Sec. 3).

References

[1] William Feller. *An Introduction to Probability Theory and its Applications; Volume 1.* New York: John Wiley & Sons, 1950.

[2] B.V. Gnedenko and A.N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables.* Cambridge: Addison-Wesley Publising Company, 1954.

[3] Samuel Karlin and Howard M. Taylor. *A First Course in Stochastic Processes;* Second Edition. San Diego: Academic Press, 1975.

[4] Sheldon M. Ross. *Applied probability models with optimization applications.* New York: Dover Publications, 1970.

[5] Lennaert J.P. Speijker, Jan M. van Noortwijk, Matthijs Kok, and Roger M. Cooke. Optimal maintenance decisions for dikes. *Probability in the Engineering and Informational Sciences*, 14:101–121, 2000.

[6] Jan M. van Noortwijk, Roger M. Cooke, and Matthijs Kok. A Bayesian failure model based on isotropic deterioration. *European Journal of Operational Research*, 82(2):270–282, 1995.

[7] Jan M. van Noortwijk, Matthijs Kok, and Roger M. Cooke. Optimal maintenance decisions for the sea-bed protection of the Eastern-Scheldt barrier. In Roger Cooke, Max Mendel, and Han Vrijling, editors, *Engineering Probabilistic Design and Maintenance for Flood Protection*, pages 25–56. Dordrecht: Kluwer Academic Publishers, 1997.

[8] Jan M. van Noortwijk and E. Bart Peerbolte. Optimal sand nourishment decisions. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 126(1):30–38, 2000.

[9] Jan M. van Noortwijk and Pieter H.A.J.M. van Gelder. Optimal maintenance decisions for berm breakwaters. *Structural Safety*, 18(4):293–309, 1996.

[10] J.M. van Noortwijk and H.E. Klatter. Optimal inspection decisions for the block mats of the Eastern-Scheldt barrier. *Reliability Engineering and System Safety*, 65:203–211, 1999.

[11] Harvey M. Wagner. *Principles of Operations Research;* Second Edition. Englewood Cliffs, NJ: Prentice-Hall, 1975.