Optimal maintenance policy for permanently monitored infrastructure subjected to extreme events

Mauricio Junca, Mauricio Sanchez-Silva

1. Introduction

The design criteria for new projects should lead to optimal investment decisions that balance the benefits derived from the existence of the project, the economic investment (inspection and maintenance) and the consequences in case of failure (e.g., reconstruction costs). Thus, designing, constructing and maintaining infrastructure may be viewed as a decision problem where the maximum economic benefit derived from the lifecycle of the project is achieved, while the reliability requirements are fulfilled simultaneously at the decision point [1]. In the literature, this type of analysis is referred to as life-cycle cost analysis. In the design and operation of public and essential infrastructure systems, the structural performance over time plays a key role. It defines the actual cost of the project, beyond the construction investments and the operation program (inspection and maintenance). However, in practice, a significant effort has been traditionally given to time-invariant design without considering the time-dependent structural performance; i.e., problems related to deterioration.

Infrastructures deteriorate as a result of the normal use and due to external demands imposed by adverse environmental conditions (e.g., earthquakes, hurricanes). One of the main challenges in life-cycle cost analysis is modeling the damage accumulation mechanisms and the associated uncertainties. Deterioration mechanisms can be divided into progressive (e.g., corrosion, fatigue) and shock-based (e.g., earthquakes, blasts) [2]. In the particular case of large civil infrastructure, progressive deterioration can be caused by, for instance, corrosion of steel structures or of the reinforcement in RC structures due to chloride ingress, reduction of structural stiffness due to concrete cracking, fatigue, creep and so forth [3–5]. On the other hand, deterioration caused by extreme events is usually associated with earthquakes, hurricanes or blasts (including both accidents and terrorists attacks). Extensive research has been carried out on mathematical models for shock degradation in infrastructure and in other types of engineered artifacts; for more details see [6–10,2].

The management of the physical aspects of infrastructure is frequently linked to inspection and intervention programs; the selection of a particular strategy is also called a maintenance policy. A maintenance policy consists of a set of actions directed to keep the system (e.g., building, bridge or pavement) operating above a pre-specified level of service; thus, maintenance is carried out to improve the availability or to extend the life of the system [11,12]. The long-term benefits of an optimum maintenance policy include: minimizing the management costs, increasing the system availability (un-interrupted operation), i.e., reducing the system's downtime and improving the time-dependent reliability [12]. Frequently, a comprehensive maintenance program includes preventive and/or corrective or reactive actions [13,14]. Preventive maintenance involves all actions directed to avoid...
failure or to avoid higher cost at a later stage by keeping the component in a safe or operational condition. While preventive maintenance is commonly carried out at fixed time intervals, corrective maintenance is performed at unpredictable intervals because failure times cannot be known a priori [15,16].

For infrastructure systems with long expected lifetimes, e.g., bridges that last on average 50–75 years; standard maintenance policies are not realistic. For example, over time, infrastructure’s deterioration cannot be predicted accurately, the structure may be exposed to some unplanned demands (e.g., increment of traffic loading) and the technology involved in inspection and intervention may change. It is then very likely that these unpredictable changes end up forcing modifications to the original maintenance plans. Furthermore, the cost-efficiency of any long-term maintenance program is difficult to verify. Therefore, the best maintenance policy should be based on a permanent monitoring strategy that leads to optimum interventions.

This paper presents a maintenance strategy based on impulse control models in which the time at which maintenance is carried out and the extent of interventions are optimized simultaneously to maximize the cost–benefit relationship. In the model, the optimal time and size of interventions are executed according to the system state, which is obtained from permanent monitoring. The model assumes that an infrastructure maintenance policy is mainly dominated by its mechanical performance.

The paper is organized as follows. The basic life-cycle cost problem is described in Section 2. In Section 3, we present an overview of infrastructure deterioration modeling strategies. Here, a novel deterioration model that takes into account the effect of damage accumulation is presented. The basic formulation and the conceptual aspects of the impulse control approach is presented in Section 4.1, where a numerical routine to calculate the optimal policy is also included. Finally, two illustrative examples are presented in Sections 5 and 6.

2. Life-cycle cost formulation

The cost-based analysis of the investments in the design and operation of a structure throughout its lifetime is based on the following basic cost–benefit formulation:

$$ Z(p, T) = B(p, T) - C_0 - \sum_{i=1}^{N_i(t)} C_i(p, \zeta_i), $$

where $T$ is the structure’s lifetime; $p$ is a vector parameter that takes into consideration the structural properties (e.g., geometry and material properties); $B(p, T)$ is the total benefit derived from the existence of the structure; $C_0$ is the initial investment cost (e.g., design and construction); $C_i(p, \zeta_i)$ is the cost associated to the $i$-th intervention (e.g., maintenance or failures) with $\zeta_i$ being the extent of the intervention, which does not necessarily takes the structure to an ‘as good as new’ condition; $N_i(t)$ is the number of interventions in the time window $T$. Clearly, both benefits and costs (losses in particular) are random variables (usually time-dependent) and therefore, $Z$ is also random. Therefore, according to statistical decision theory, Eq. (1) should be evaluated using the discounted expected net present value.

The continuous discounting function can be expressed as:

$$ \gamma(t) = \exp(-\delta t), $$

where $\delta$ is the discount rate. The discount rate is in general difficult to estimate since it depends on many factors but typical values are within the range $0 < \delta < 7\%$. For instance, for bridge investments in the United Kingdom a common discount rate varies between 4% and 6% [17]. An interesting discussion on the selection of the discount factor can be found in [18,19].

An investment in the construction and operation of the facility makes sense only if the discounted (to the decision point, e.g., $t=0$) expected value of the objective function, $E[Z(p, T)] > 0$; and it is financially optimal for the value of $p$ that maximizes Eq. (1).

3. Structural deterioration

Infrastructure degradation is referred as the process of decay, or loss of value, of one or more structural properties (e.g., stiffness, resistance). Degradation is measured mainly in physical units (e.g., inter-storey drift, loss of stiffness), although analytical assessments such as the reliability index [5], may be used also to describe the overall performance of the system. There are two distinct types of degradation models: (1) progressive (graceful) and (2) shock-based, which may or may not occur simultaneously (Fig. 1) [2].

3.1. Progressive degradation

Progressive (graceful) degradation is the result of a continuous reduction in the structure’s capacity/resistance. Most progressive degradation models available in the literature assume that the form of the degradation process is known, but the parameters are uncertain. The solution to this problem conveys to a parameter estimation problem. Thus, if $R_p(t)$ is the state of the system at a given time $t$, which in practice, may be expressed in terms of, for example, remaining capacity, reliability, safety, durability, etc., then these type of models have the following general form:

$$ R_p(t) = \begin{cases} r_0 & 0 \leq t \leq t_e \\ r_0 - D_p(t - t_e) & t > t_e \end{cases}, $$

where $r_0$ is the remaining life of the system at time $t=0$ and $t_e$ is the time of degradation initiation (e.g., time of corrosion initiation). The function $D_p$ may take a linear, non-linear, exponential

![Fig. 1. Description of degradation mechanisms.](image-url)
or any other form based on the appropriate selection of the vector parameter \( \mathbf{p} \), which depends upon the problem at hand.

Progressive degradation can be handled also in terms of a degradation rate, i.e., \( \delta_i(t) \), \( t \geq 0 \) that may or may not change (randomly) over time. Thus, the system state (e.g., remaining capacity measured in physical units per unit time) at time \( t \) can be expressed as

\[
R_D(t) = r_0 - \int_0^t \delta_i(t) \, dt,
\]

where the rate may be dependent or independent of time.

Finally, several models have been proposed to approximate continuous degradation as a discrete process with random changes at fixed/random time intervals; for example, by using gamma processes [20].

3.2. Degradation caused by shocks

Shocks can be defined as events that cause a significant change in a system’s physical property in a small time interval. Then, shock-based degradation occurs when a fixed amount of capacity/resistance is removed from the system at discrete points in time [2].

If, \( Y_i \) is a random variable describing the degradation caused by shock \( i \), the total degradation by time \( t \) is

\[
D(t) = \sum_{i=1}^{N(t)} Y_i,
\]

where \( N(t) \) represents the number of shocks by time \( t \). Note that in many practical applications the time between shocks is also random; therefore, \( N(t) \) is also a random variable. If deterioration is caused by shocks, which are assumed \( iid \), then, the remaining structural capacity/resistance by time \( t \) can be computed as

\[
R_s(t) = r_0 - D(t) = r_0 - \sum_{i=1}^{N(t)} Y_i,
\]

Extensive research has been carried out on developing mathematical models for shock degradation. Additional details about this methods can be found in [6,7,5,2].

Frequently, the assumption that shock sizes are \( iid \) is too strong or not realistic. For instance, consider a structure located in a seismic region subject to a series of earthquakes that cause damage to accumulate with time. Then, the damage caused by a given earthquake is conditioned on the damage state of the structure just before its occurrence. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state. This means that the probability distribution of next shock size (i.e., damage) is somehow dependent on the current state.

3.3. General degradation model

Frequently, progressive and shock-based deterioration occur simultaneously. Thus, for a structural component with initial capacity \( r_0 \), subject to both continuous and sudden damaging events acting independently, the remaining capacity/resistance by time \( t \) can be computed as

\[
R(t) = r_0 - D(t) = r_0 - \sum_{i=1}^{N(t)} Y_i(R(T_i-)).
\]

It is important to stress that in some cases both events are not independent and, therefore, the coupled effect should be taken into consideration.

4. Basic formulation of the model

4.1. Impulse control model

Impulse control models have been used in areas such as finance to optimize a portfolio of risky assets with transaction costs, or to find the best strategy to execute a position in a risky asset [22,23], inventory control to find the optimal order placement [24] and insurance to find the optimal dividend payment for an insurance company [25]. Recently, in [26], this model has been used in the context of optimal maintenance policies. The following model is partially based on that work.

Let us assume that a system (e.g., structure, bridge) is subjected to degradation caused by shocks that occur according to a Poisson process. Every shock causes a random amount of damage according to the function \( g \) as described in Section 3.2. Thus, given \( (\Omega,\mathcal{F},\mathbb{P}) \) the probability space in which we define all the stochastic quantities, we define the process \( R(t) \):

\[
R(t) = r_0 - \sum_{i=1}^{N(t)} g(Y_i(R(T_i-)),
\]

where \( N(t) \) is a homogeneous Poisson process with intensity \( \lambda > 0 \). We assume that \( Y_i, i \in \mathbb{N} \) are independent and identically distributed random variables with probability distribution \( F \) on \((0,\infty)\) and independent of the Poisson process \( N(t) \). The initial reliability level is \( r_0 \) (Fig. 2a).

Let us now define a maintenance policy for the system as a double sequence \( r = (\tau_i,\zeta_i), i \in \mathbb{N} \) of intervention times \( \tau_i \) at which the performance is improved an amount \( \zeta_i \). The policy is said to be an impulse control if it satisfies the following conditions:

1. \( 0 \leq \tau_i < \tau_{i+1} \) for all \( i \in \mathbb{N} \),
2. \( \tau_i \) is a stopping time with respect to the filtration \( \mathcal{F}_i = \sigma(R(s-): s \leq t) \) for \( t \geq 0 \),
3. \( \zeta_i \) is a \( \mathcal{F}_\tau \) measurable random variable.
Then, given an impulse control $v$, the controlled process $R'(t)$ is defined by (Fig. 2b)

$$R'(t) = r_0 - \sum_{i=1}^{\infty} g(Y_i, R(T_i-)) + \sum_{\Gamma \Delta} \xi_i.$$  

(9)

Since we are interested in keeping the system performance above a pre-defined threshold $k^*$, with $0 \leq k^*$, we assume total failure when the system performance falls below this level. At this time the process is stopped. Thus, the time of total failure of the controlled process is denoted by $\tau^* = \inf\{t > 0 | R'(t) \leq k^*\}$.  

(10)

We denote by $\tau$ the time of total failure of the uncontrolled process $R$. For simplicity, it is assumed that $k^* = 0$. While $k^*$ denote a lower limit for the process, we also assume that there is a maximum (i.e., optimum) performance level $\hat{r}$ that cannot be improved. Therefore, any intervention (i.e., maintenance) at time $\tau_i$ must satisfy that $\xi_i \in \{0, \hat{r} - R(T_i-), 0\}$, where $R(T_i-)$ is the state of the system just before the intervention. In this case we say that the policy is admissible.

4.2. Value function

Let us now consider the gains (benefits) and costs associated with the process. Thus, if we denote $g(x) := E_x[R'(0-)-r_0]$, for a given admissible $v$ and initial component state $r_0 \in [0, \hat{r}]$, then the expected profit (Benefits–Costs) derived from the operation of the project can be computed as

$$J(r_0, v) = E_x\left[\int_0^\tau e^{-\delta s} G(R(s)) ds - \sum_{\tau_i < \tau} e^{-\delta \tau_i} C(R(T_i-), \xi_i)\right],$$

(11)

where $g$ is a non-negative continuous, increasing and concave function on $[0, \hat{r}]$ with $G(0) = 0$, $C$ is a continuous function, increasing in both variables and $\delta$ is the discount factor (see Section 2). Note that the first term in Eq. (11) corresponds to the discounted benefits; where the function $G$ can be interpreted as an utility function. On the other hand, the second term describes the discounted costs of interventions with $C(r, \xi)$ the cost of bringing the system from level $r$ to level $r+\xi$.

The objective of the model is to find the policy that maximizes the profit among all admissible impulse controls, that is,

$$V(r_0) = \sup_v J(r_0, v),$$

(12)

for a given level $r_0$ in the state space $[0, \hat{r}]$. It is very difficult to calculate $V(r_0)$ directly from (12). The main difficulty is that, first, given a policy $v$, we need to use simulations to estimate the expected profit $J(r_0, v)$. Then, we have to repeat this process for all possible policies $v$, which clearly cannot be obtained at a reasonable computational cost.

Instead, we will solve the problem for all $r \in [0, \hat{r}]$ at the same time, that is, we want to find the value function

$$V(r) = \sup_v J(r, v).$$

(13)

Although apparently this is a harder problem, we will characterize $V$ as the unique solution of certain equation (that does not involve expectation) and solve this equation numerically. Then, from the definition of the value function, $V$, we can easy see that $V \geq 0$ since we can choose to do nothing. Also, $V(0) = 0$ and $V$ is bounded. Below, we will use these properties to characterize the function $V$.

4.3. Intervention and infinitesimal operators

In this section we will present some definitions and fundamental concepts that are required to find $V$ in Eq. (13). We start by letting $T$ be a stopping time with respect to the filtration $\mathcal{F}_t$, then, for all $r \in [0, \hat{r}]$ (see proof in [26])

$$V(r) \geq E_r\left[\int_0^{T \wedge \tau} e^{-\delta s} G(R(s)) ds + e^{-\delta \tau} V(R(T)) I_{[T < \tau]}\right].$$

(14)

Note that we may have equality in (14) if it is not optimal to intervene the system before $T$. Next, in order to characterize the value function $V$ in Eq. (13) we need to define two important operators. The first one is the intervention operator $\mathcal{M}$ defined as

$$\mathcal{M}f(r) = \sup_{0 \leq \xi \leq \hat{r}-r} f(r + \xi) - C(r, \xi),$$

(15)

for a given function $f$ defined on $[0, \hat{r}]$ and $r$ in the same interval. Note that we take the supremum over the interval $[0, \hat{r}-r]$ in order to consider only admissible policies. We are interested in applying $\mathcal{M}$ to the function $V$. Hence, if we consider any policy $v$ such that $\tau_1 = 0$ and write $v = (0, \xi) \cup \{(\tau_2, \xi_2)\}_{2 \geq 2} = (0, \xi) \cup \tilde{v}$, then by Eqs. (13) and (11)

$$V(r) \geq f(r, v) = -C(r, \xi) + f(r + \xi, \tilde{v}).$$

Since $\tilde{v}$ is arbitrary we can take the supremum over all controls $v$ and obtain

$$V(r) \geq V(r + \xi) - C(r, \xi).$$

(16)

Now, taking the supremum over all admissible $\xi$, we obtain

$$V(r) \geq \sup_{0 \leq \xi \leq \hat{r}-r} V(r + \xi) - C(r, \xi) = \mathcal{M}V(r).$$

We will use this inequality in the characterization of the function $V$. The second operator is the infinitesimal generator of the uncontrolled process $R(t)$, that is,

$$\mathcal{A}f(r) = \partial_t f(r) + (\partial_y g(y, r)) dF(y) - f(r).$$

(17)
It is known (see [27,25]) that given $T_1 \leq T_2$ almost sure (a.s.) finite stopping times, then

$$
E[e^{-\gamma T_2}f(R(T_2)) - e^{-\gamma T_1}f(R(T_1))] = E\left[e^{\int_{T_1}^{T_2} e^{-\gamma s}[\mathcal{A}f(R(s)) - \delta f(R(s))] \, ds}\right].
$$

(18)

Below, we will use this so-called Dynkin’s Formula with $f$ replaced by $V$ to completely describe the value function.

### 4.4. Optimal maintenance policy

Since the process $R(t)$ is Markovian, to obtain an optimal policy it is necessary to only look at the present state of the system and not how the system arrived to the present state. So, given an state $r$ we want to know if an intervention is required or not. We now use the intervention operator $\mathcal{M}$ to answer this question. From Eq. (16) $V \geq MV$, we can divide the state space $[0,Opt]$ into the subsets

$$
A = \{ r \in [0,Opt] : V(r) = MV(r) \}
$$

and

$$
B = \{ r \in [0,Opt] : V(r) > MV(r) \}.
$$

For $r \in A$ we must intervene the system immediately and improve the performance process by $z^*$, where

$$
\mathcal{M}V(r) = V(r + z^*) - C(r, z^*) = \sup_{0 \leq z \leq Opt - r} V(r + z) - C(r, z).
$$

(19)

Therefore, we call the set $A$ as the intervention region. For the other states, i.e., $B$, we do nothing and let the system evolve. We call the set $B$ as the no intervention region (Fig. 3). It is important to stress that because of the Markovian property, this classification will always be the same and will not change in time.

Now, for $r \in B$ it is optimal not to intervene the system, therefore, we obtain equality in (14), and using Dynkin’s Formula we have that $\delta V(r) - AV(r) = 0$. Then, it can be proved [26] that the value function $V$ solves the equation,

$$
\min_r \{ \delta V(r) - AV(r) - G(r) \} = 0,
$$

(20)

for all $r \in [0,Opt]$. Furthermore, if $f$ is a non-negative bounded function on $[0,Opt]$ that solves (20) such that $f(0) = 0$, then $f = V$.

These are existence and uniqueness results for Eq. (20). Now, with respect to existence of optimal controls, it is possible that $z^*$ is not attainable in Eq. (19). In this case there is no attainable optimum policy, but we can find controls with expected profit as close as possible to the value function $V$.

#### 4.5. Numerical solution

To obtain the optimal policy it is necessary to find the value function $V$ by solving Eq. (20). Once we have $V$, we can compute the intervention and the no intervention regions. We start by analyzing the set $B$. We know that for $r \in B$, we have that $\delta V(r) - AV(r) - G(r) = 0$.

Using the definition of the infinitesimal operator $\mathcal{A}$ (Eq. (17)) with $f = V$ and solving the above equation for $V(r)$ we obtain

$$
V(r) = \frac{1}{\lambda + \delta} \int_0^\infty (r - g(y,r)) \, df(y).
$$

(21)

Note that $V$ is defined only for $r \in [0,Opt]$, so for $y > y^*$ with $g(y^*,r) = r$, $r - g(y,r) < 0$. Therefore, we can replace the upper limit in the integral by $y^*$. On the other hand, for the region where interventions are required, i.e., $r \in A$, we have that

$$
V(r) - AV(r) = 0.
$$

Then, using the definition of the intervention operator (Eq. (15)), $V(r)$ satisfies

$$
V(r) = \sup_{0 \leq z \leq Opt - r} V(r + z) - C(r, z).
$$

(22)

To approximate $V$ we follow the Jacobi iteration method described in [28]. We initialize $V_0(r)$ for $r \in [0,Opt]$ (an option is to initialize with 0) and using Eqs. (21) and (22), we define recursively

$$
V_{n+1}(r) = \max \left\{ \frac{1}{\lambda + \delta} \int_0^\infty (r - g(y,r)) \, df(y), \sup_{0 \leq z \leq Opt - r} V_n(r + z) - C(r, z) \right\}.
$$

(23)

### 5. Example 1: Comparison of progressive deterioration intensity

#### 5.1. Problem description

Consider an infrastructure component (e.g., bridge) whose performance, $R$, which should be in practice measured in physical units (e.g., resistance, reliability), is normalized and evaluated within the interval $[0,1]$, where $r_0 = 1$ means that the system is in “as good as new” condition and $r_{\text{min}} = k^0$ indicates that it is not in operating condition. Furthermore, let us assume that the structure is located in a seismic region where earthquake times follow a Poisson process with intensity $\lambda = 0.5$.

As a result of every earthquake, the structure may be damaged; earthquake damage is represented by shocks. The size of shocks is the result of the earthquake motion characteristics and the structural response. For the purpose of this example, shock sizes $Y$ are iid random variables log-normally distributed. Although shock size distributions are iid random variables, damage accumulation is a monotonically increasing process for which we need to define the function $g(y,r)$ (Eq. (8)). For the purpose of this example,

$$
g(y,r) = \beta \frac{Y}{T},
$$

(24)

where $\beta$ is an arbitrary constant; in this example $\beta = 1$. In other words, the condition of the structure does not depend only on the shock sizes but also on the structural condition at the time of the earthquake. An illustrative sample path of this process is shown in Fig. 4. Note also that it is not easy to find an analytical expression for the probability distribution of $Y_t/R(T_1-)$ and the corresponding convolution after $N(t)$ shocks.
The objective of the study is to define an optimal maintenance policy of a structure that is monitored permanently. This means that it is possible to know the state of the system whenever it is required (e.g., at periodic inspection times). Then, the optimal solution requires the assessment of benefits and costs. Therefore, the benefit (utility) function derived from the existence of the project is given by (Eq. (11))

\[ G(r) = (\phi C_0) \frac{1}{2} (1 - e^{-r}), \]

where \( C_0 = 100, \phi = 0.275 \) and \( z = 0.5 \). Note that this curve has the form of an exponential risk aversion utility function. On the other hand, it is assumed that the costs associated to an intervention are given by the following function (Eq.(11)):

\[ C(r; \zeta) = C_0 \frac{1}{2} + k C_0 (1 - r), \]

where the constant \( k = 0.1 \) reflects the fixed costs of any intervention. Note that the intervention costs are proportional to the size of the intervention. For both utility and cost, these values are discounted to the time of the decision by using a discount factor \( \delta = 0.05 \).

5.2. Results and discussion

The analysis consists of two parts. First, it is necessary to define, for every structural state \( r \), the intervention intensity \( \zeta \) that maximizes the expected profit (Eq. (11)). This requires dividing the state space in two regions: a region where no intervention is required and a set of values for which the intervention is necessary. Thus, at a given time \( t_{\text{fail}} \), the state (condition) of the structure is obtained from an inspection; the intervention level that maximizes the expected profit at that particular time can be obtained from Fig. 5b. The second result is the value function \( V \) that provides the maximum expected profit if the intervention program, defined previously (Fig. 5b), is implemented. Then, the x-axis in Fig. 5a corresponds to the initial state of the system, i.e., \( r_0 \) and \( V \) is the maximum profit for the intervention program shown in Fig. 5b.

For the data of this problem, the division between the intervention and the not intervention states can be observed in Fig. 5b. Clearly, if the system is operating at a level \( r > 0.42 \) there is no need for an intervention. However, if an inspection indicates that the system state is \( r < 0.42 \) and intervention is required. For instance, if after an inspection the system state is \( r = 0.7 \) no intervention is required, but if \( r = 0.3 \) and intervention of magnitude \( \zeta = 0.7 \) will be necessary to maximize the profit. This means that immediately after the intervention the system will be at state 1 (i.e., \( 0.3 + 0.7 = 1 \)). It is important to note that the results shown from Fig. 5 are time-independent. If maintenance is carried out under this policy, the maximum expected profit can be obtained in Fig. 5a. Therefore, if, for example, the initial system state is \( r_0 = 0.4 \), the maximum expected profit that we can obtain, by following the above policy, is \( V(0.3) = 870 \).

The sensitivity of the maintenance policy with respect to the discount rate is shown in Fig. 6. For comparison purposes, two different deterioration functions \( g(r; \gamma) \) were considered. In Fig. 6a, the function \( g \) was selected as defined in Eq. (24), while in Fig. 6b the analysis was carried out for \( g(r; \gamma) = \gamma \), which means that shock sizes are iid and the damage accumulation does not depend on the previous state of the system. It should be first noted that, for both functions, as the discount rate becomes larger, the range of structural states for which an intervention is required becomes smaller. This is justified by the fact that interventions are only required if the system state is closer to failure then, although interventions are more expensive, they are discounted with a higher rate. In addition, it can be observed also that if the effect of damage accumulation is taken into account, the region of system states where an intervention is required is larger than the region for the case of no damage accumulation.

Finally, the effect of the shock sizes on the maintenance policy for the case in which damage accumulation is taken into consideration is presented in Fig. 7. For given mean value it is clear that larger coefficients of variation (COV) imply larger failure probabilities and, therefore, the region where interventions are required becomes also larger. In addition, the effect of the mean, for a fixed COV, is similar than in the previous case. However, the intervention space is larger than in the first case.

6. Example 2: structure subject to cyclic loading

6.1. Problem description

Consider now the case of a structure whose performance is described by a bilinear constitutive model as shown in Fig. 8, where \( K = 2, K_c = 0.2 \) and \( \gamma_c = 0.25 \).
The structure is subjected to successive extreme events (e.g., earthquakes). If the demand is not large enough to take the structure out of the elastic range, no damage will be reported. The excursions into the inelastic range will define the degradation process by redefining the initial displacement state and the extension of the elastic range for next iteration. Damage in this case will be measured in terms of the residual displacement; then, after a shock of size \( y \), the change in the residual displacement \( r \) can be computed as

\[
g(y,r) = \begin{cases} 0 & \text{if } y \leq Kc - Kc(1 - \alpha) \frac{K}{K + Kc} \\ \frac{y}{K} (1 - (1 - \alpha) \frac{K}{K + Kc}) & \text{if } y > Kc - Kc(1 - \alpha) \frac{K}{K + Kc} \end{cases}
\]

(27)

where \( \epsilon_r \) is as indicated in Fig. 8. Note that if an intervention \( \zeta \) is carried out (i.e., retrofitting), the initial displacement for the subsequent loading cycle will be reduced. The purpose of this example is to identify the optimum maintenance policy.

Let us assume that both the utility and the intervention cost functions have the same form as in Example 1, i.e., Eqs. (25) and (26) with the following parameters: \( C = 100, k = 0.1, \delta = 0.05 \). Shock sizes were assumed to be lognormal with parameters \( \mu = 0.4 \) and \( \text{COV} = 0.35 \).

For comparison purposes, the analysis was carried out for three different event occurrence rates \( \lambda = 0.1, \lambda = 1 \) and \( \lambda = 10 \); the results are shown in Fig. 9. The maximum expected benefit is shown in Fig. 9a, while the optimum maintenance policy for all the three cases considered is presented Fig. 9b. The results show that the effect of the occurrence rate \( \lambda \) on the total profit is as expected; lower occurrence rates lead to larger profits and to a smaller intervention region. Note that when the rate becomes very small, the value of the objective function reaches a maximum value of $1100. On the other hand, the intervention policies also change depending upon the occurrence rate. In this case, the state space for which maintenance actions are required is larger for higher rates (see Fig. 9b). In this case, it is interesting to observe that for \( \lambda = 0.1 \) interventions do not require to take the structure to its original condition (i.e., “as good as new”) but to a lower level. For instance, for \( \lambda = 0.1 \), if the condition of the system is \( r = 0.1 \) the size of the intervention would be \( \zeta = 0.3 \) and the final state of the system would be \( r = 0.1 + 0.3 = 0.4 \). The main reason for this is that since events are highly spaced in time, the structure can operate for long time without failure.
7. Summary and conclusions

The paper presents an approach to define the optimum maintenance policy of a system that deteriorates with time as a result of shocks. The deterioration process is modeled as a compound Poisson process. The proposed maintenance strategy follows an impulse control model that requires the permanent (or at least frequent) monitoring of the system state. Then, at every inspection time the model can be used to make a decision as to whether the system should be intervened or not. In case of requiring an intervention, the extent of the optimum repair can be obtained from the model. The decisions based on the model guarantee that the net present value of the utility at the time of the intervention is maximum. It is suggested in the paper that this maintenance approach is of particular importance for engineering systems that operate under adverse environments for long time periods (e.g., > 25 years), for instance, physical infrastructure (e.g., bridges, highways). Traditional maintenance strategies define long term inspection and maintenance plans at a given point in time (usually t = 0) without considering possible variations in the system’s use or condition and the technology available at the time of the decisions. Thus, the best maintenance policy in these cases can be obtained by combining both permanent monitoring and optimum interventions.

References