Received: 1 December 2019 | Revised: 14 June 2020 | Accepted: 15 June 2020

DOI: 10.1002/qre.2707

SPECIAL ISSUE ARTICLE

WILEY

Optimal inspection for randomly triggered hidden deterioration processes

Ayman Hijazy[1,2] | **András Zempléni**[1,2]

[1]Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest, Hungary

[2]Faculty of Informatics, University of Debrecen, Debrecen, Hungary

Correspondence
Ayman Hijazy, Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest, Hungary.
Email: aymanh@cs.elte.hu

Present Address
András Zempléni, Pázmány Péter sétány 1/C, Budapest 1117, Hungary.

Funding information
European Social Fund, Grant/Award Number: EFOP-3.6.2-16-2017-00015

Abstract

This paper deals with degradation processes whose onset is triggered at a random time and which stay hidden until they are discovered through inspection or when they begin to show symptoms. This is applicable in many healthcare and industrial scenarios, for example, in the modeling of breast cancer or termite infestation. In our model, we assume that symptoms appear after hitting a random critical threshold and that inspections may have a sensitivity less than one as well as a nonzero false positive rate. The expected cost of repair is derived, and the inspection rate is optimized for a cycle (which lasts from degradation-free to repaired state). This gives results for three cases: the first is for a finite observation period with no degradation recurrence, the second for infinite time horizon allowing recurrence. In the third case, we derive an upper bound for the expected cost in a given constant time period. Finally, the model is applied to determine the optimal strategy for breast cancer screening with regard to the effects of different parametrizations.

KEYWORDS

maintenance, process monitoring and control, survival analysis

1 | INTRODUCTION

In order to prevent catastrophic failure of a system and minimize the cost of repairing it, maintenance programs are arranged for systems that are subject to deterioration. Using statistical modeling and inference techniques on degradation measures is useful for obtaining reliability information and understanding the wear process. After the statistical structure of the degradation has been identified, the appropriate maintenance policy can be selected. The two main maintenance techniques discussed in the literature[1] are time-based maintenance (TBM) and condition-based maintenance (CBM). TBM is calendar based where a maintenance plan is arranged for an item and is performed regularly. On the other hand, CBM is based on monitoring the actual condition of the asset to decide what maintenance needs to be done. This paper investigates TBM.

The gamma process[2] has been extensively used in the literature to describe the stochastic and monotone degradation accumulating over time.[3] Degradation models based on gamma process have been identified as the main way to model degradation processes where the increments are independent, nonnegative, and gamma distributed with an identical scale parameter. They are well suited for modeling the temporal variability of deterioration and have proven to be useful in determining optimal inspection and maintenance decisions.[4,5]

[*Corrections added on 15 August 2020, after first online publication: The word Appendix B1 have been corrected to Appendix A throughout the paper.]

Inspection has been optimized in different contexts, and for different assumptions, Kong and Park[6] minimized the long-run expected cost assuming that an item fails randomly but the failure rate depends on the accumulated damage. In their model, they assumed that an item is preventively replaced if it survives a certain damage limit at periodic inspections and is replaced immediately after failure. Many different approaches have been suggested to optimize inspection in different scenarios; Jia and Christer[7] jointly optimized the threshold of potential failure and inspection intervals to minimize the expected operating cost per unit time while differentiating between first and subsequent repairs. Another extension by Kallen and Noortwijk[8] includes imperfect measurements; they derived optimal inspection plans under uncertain deterioration, assuming a random failure condition that is dependent on uncertain operation conditions and material properties. A rich survey of the application of gamma processes in maintenance is presented by van Noortwijk.[3]

Markov chain-based models proved to be a good way to approach degradation modeling when it can be described by different states. Recently, Yang and Sørensen[9] presented a discrete Markov chain model as a simplified probabilistic model for damages in wind turbine blades by making use of a six-level damage categorization scheme applied by the wind industry. Degradation modeling also opens the door on cost minimization for preventive maintenance in healthcare applications. For instance, Dobi and Zempléni[10] used Markov chain-based control charts in order to optimize the cost for patients with cardiovascular event risk.

In this paper, deteriorations are assumed to start at a random time and stay hidden until detection. This is applicable in scenarios where the degradation is initiated by a triggering event, such as termite infestation or progression of breast cancer,[11] where the exact onset of the degradation is unknown. It is also not always known if an item is degrading or not; for example, a breast cancer patient would only find out that she is ill either by screening or after showing symptoms.

The process we consider can then be described by three states: starting from a degradation free state S_f, then moving to a hidden deterioration state S_h, until the deterioration is finally exposed in the state S_e where it is repaired and moves back to S_f, Figure 1 shows the structure of this three-state model. Moreover, assume that a gamma degradation process starts when the item moves to S_h where it stays hidden until moving to S_e. In this scenario, it is unknown if an item is already deteriorating or not until the degradation is either detected by inspection or shows some symptoms.

Now let us define the dynamics of propagation in this setup : suppose that the transition from S_f to S_h occurs at a random time point (onset); that is, the item stays in S_f till the deterioration triggering event occurs. For the movement between S_h and S_e, define the Sojourn time as the amount of time spent in the hidden state (S_h), where the deterioration is assumed to show symptoms and move to S_e after hitting a random critical threshold. In other words, the sojourn time is the amount of time needed for a degradation to show symptoms.

Assume that an inspection program is organized aiming for early detection of deterioration in the hope of minimizing the cost of repair. In this case, effective inspection takes place when a case that is already in S_h is detected. However, the inspection may not be perfect; that is, items that are deterioration free or are deteriorating may not be correctly identified. For that purpose, the sensitivity of an inspection is defined as the probability of detecting the deterioration given that a case is already in S_h. Let us suppose that the inspection sensitivity may depend on the level of degradation. Furthermore, we also assume that inspection may also have a false positive rate, which is defined as the probability of an inspection falsely detecting deterioration given that the case is in S_f.

After the degradation of an item is detected, whether by means of inspection or by showing symptoms, we suppose that it is repaired instantly and that the repair is perfect. This may not be realistic in some applications, but it can be true in some cases, for example, when repair means the actual replacement of a unit. Suppose that repaired items move back to S_f where they may propagate through the chain again. Under this setup, the aim is to establish a cost-optimal inspection policy. The cost is divided into four parts: the cost of repair of items detected at inspection, the cost of repair of items reaching S_e, the cost of inspection itself, and the cost of false positives.

The structure of the paper is as follows: in Section 2, the setup is formally established and the expected total costs for different scenarios are derived. In Section 3, we apply our approach to a breast cancer screening program. In Section 4, we

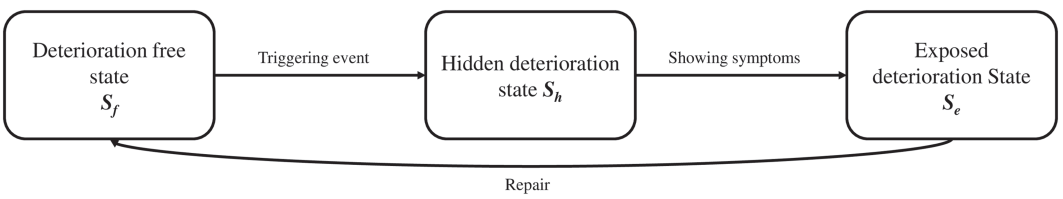

FIGURE 1 Propagation in the three-state model

state some concluding remarks and discuss the possible extensions for our approach. The derivation of the distributions of the degradation on inspections and between inspections is shown in Appendix A.

2 | SETUP

Before we lay out the setup of the model, notations are introduced and listed in Table 1. Starting with the initiation mechanism of the degradation, we assume that the degradation process is triggered by a random event occurring at a random time T_d. Let the lifetime risk r be the probability that the event causing the initiation of the degradation phenomenon under consideration occurs. T_d can then be viewed as the waiting time in S_f with density f_{T_d}, which is a mixture of a probability density function g and the Dirac delta function centered at $+\infty$:

$$f_{T_d}(t) = r \cdot g(t) + (1 - r) \cdot \delta_{+\infty}(t); \tag{1}$$

TABLE 1 List of notations

	Notation	Meaning
States	S_f	Deterioration-free state
	S_h	Hidden deterioration state
	S_e	Exposed deterioration state
Onset	T_d	Onset time
	r	Lifetime risk
	$f_{T_d}(t)$	Density of the onset time
Degradation process	$\eta(t)$	Shape parameter of the process after time t of the onset
	$a(t)$	Derivative of $\eta(t)$ w.r.t. time
	β	Scale parameter of the process
	$(X_t)_{t\geq 0}$	Gamma process with shape $\eta(t)$ and scale β
Inspection	τ_i	Time of the ith inspection
	Δ	Inter-inspection time
	$\Lambda(x)$	P(inspection detecting a deterioration | item is deteriorating with damage x)
	R_{FP}	P(inspection detecting a deterioration | item is not deteriorating)
	T_L	Observation period
	K	Number of inspections
	T_y	Given constant time period
Critical threshold	C	Gamma-distributed critical threshold
	Λ	Shape parameter of C
	ξ	Scale parameter of C
	f_C	Density of the critical threshold
Defined costs	$CR(x)$	Cost of repair (function of damage size x)
	c_I	Cost of inspection of a single item
	c_{FP}	Cost of identifying a single false positive item
Sojourn time	T_C	First hitting time of the random threshold C
	F_{T_C}, F_{T_C}	cdf and pdf of the first hitting time T_C
	$X_{T_C^-}$	left side limit of $(X_t)_{t\geq 0}$ at time T_C
	X_{T_C}	Deterioration after crossing the threshold C
Probabilities and densities related to the degradation	$f_{(i,j)}(x)$	Contribution of cases moving to S_h between (τ_{j-1}, τ_j) to the probability of the detected level of degradation at inspection τ_i to be between $(x, x+dx)$ (Equation A2)
	$f_{\tau_i}(x)$	Probability of the detected level of degradation at inspection and τ_i to be between $(x, x+dx)$ (Equation A3)
	$g_{(i,j)}(s,x,z)$	Defective joint density of $(T_c, X_{T_C^-}, X_{T_C})$ for cases moving into S_h between (τ_{j-1}, τ_j) and to S_e between (τ_{i-1}, τ_i) for a given $C = c$ (Equation A8)
	$g_{\tau_i}(s,x,z)$	Joint defective density of $(T_c, X_{T_C^-}, X_{T_C})$ for T_C between τ_{i-1} and τ_i for a given $C = c$ (Equation A7)
	$D(\tau_i)$	Probability of detecting the deterioration at inspection τ_i (Equation A4)
	$I(\tau_i)$	Probability of the deterioration to become exposed between τ_{i-1} and τ_i (Equation A9)
	$P_S(\tau_i)$	Probability of an item to participate in inspection τ_i
Costs	$C_D(\tau_i)$	Cost of repair of damaged items detected at inspection τ_i (Equation A5)
	$C_E(\tau_i)$	Cost of repair of exposed deterioration between τ_{i-1} and τ_i (Equation A10)
	$C_{ins}(\tau_i)$	Cost of inspection τ_i
	$C_{FP}(\tau_i)$	Cost of identifying false positives at inspection τ_i
	U	Number of repairs
	$V^{(i)}(\tau_1, \Delta)$	Total cost of the ith cycle
	$V^{(\infty)}(\tau_1, \Delta)$	Cost for an infinite time horizon

hence, $P(T_d < \infty) = r$. It is assumed that the evolution of the degradation process is independent of the onset age.

Regarding the degradation structure, suppose that the level of degradation can be described by a nonhomogeneous gamma process[2] denoted by $(X_t)_{t\geq 0}$. The considered process is assumed to have a constant scale parameter β and a time-dependent shape $\eta(t)$, where $\eta(t)$ is a continuous, monotone, nonnegative, and increasing function for $t \geq 0$ with $\eta(0) = 0$. Under this setup, the marginal distribution of X_t is the gamma distribution with shape $\eta(t)$ and scale β. Note that the gamma process has the following properties:

1. The increments of the gamma process in the interval (t_1, t_2) are independent random variables over disjoint time intervals. Let us denote them by $X_{(t_1,t_2)} = X_{t_2} - X_{t_1}; t_1 \geq 0; t_2 > t_1$.
2. Each increment $X_{(t_1,t_2)}$ follows a gamma distribution with constant scale parameter and time-varying shape parameter $\Delta\eta(t_1, t_2) = \eta(t_2) - \eta(t_1)$ for $t_2 > t_1$. The density of the increments is thus given by

$$f_{X(t_1,t_2)}(x) = \frac{1}{\Gamma(\Delta\eta(t_1,t_2))\beta^{\Delta\eta(t_1,t_2)}} x^{\Delta\eta(t_1,t_2)-1} \exp(-x/\beta)\mathbb{1}_{\mathbb{R}_+}(x). \tag{2}$$

The gamma process is a jump process; that is, it is a right-continuous stochastic process with left-side limits (càdlàg). The Markovianity is also a key property of the process. The Markov property implies that one can infer about the future of the process based solely on the present state just as well as one could by knowing the full history of the process; that is, $P\left(X_{t+s} = j | X_s = i, X_{u_1} = k_1, X_{u_2} = k_2 \ \ldots\ ; s > u_1 > u_2 > \ \ldots\ > 0\right) = P(X_{t+s} = j | X_s = i)\ , \forall t > s > 0$.

The degradation monitored in this setup is a gamma process triggered at a random time T_d. For a given $T_d = t_d$, let $z = t - t_d, t \geq t_d$ be the elapsed time after the onset. Hence, the density of the degradation after time z is simply given by $f_{X_z}(x)$ and the expected degradation is $E(X_z) = \beta \cdot \eta(z)$.

Moving to the means of detection of the deterioration, we assume that it is generally not known whether an item is deteriorating or not and suppose that the deterioration can be detected in one of two ways, either by hitting a random critical level or by preventive inspections aiming for early detection. Thus, our assumptions are the following:

- The deterioration shows symptoms and is detected when it reaches a critical threshold C. Depending on the application, one can choose between a probabilistic or a deterministic threshold. We opted for a random threshold as it is more general and suitable for our application, as there is some randomness in the time when the deterioration causes symptoms. Besides, there may be a propensity to delay medical attention even when symptoms are present. We suppose that C is a gamma-distributed random variable with shape Λ and scale ξ and denote its density by $f_C(c)$.
- Preventive periodic inspections are arranged at times $(\tau_1, \tau_2, \ \ldots\ )$ with an inter-inspection time Δ; that is, $\tau_i = \tau_1 + (i - 1)\Delta$ such that:
    1. Inspections may have a smaller than one level of sensitivity, which is defined as the probability of detecting deterioration given that the inspected item is deteriorating. In most applications, the probability of detection may depend on the level of degradation; that is, we may assume that the sensitivity of inspection $\Lambda(x)$ is a nondecreasing function of the deterioration level $x > 0$.
    2. On the other hand, the false positive rate R_{FP}, which is the probability of the inspection falsely detecting a deterioration given that there is no actual deterioration, is usually related to the diagnostic method, not the deterioration itself. For that reason, it is defined as a constant.

Next comes repair: we assume that items are repaired immediately and perfectly after detection. In other words, once an item is repaired, it goes back to S_f, from where it may propagate through the chain again. Figure 2 gives an overview of the model with some realizations: an item starts deteriorating at a random time triggering the gamma process; it may then be detected by inspection, keep progressing until it shows some symptoms, or the observation period may end without the degradation being detected.

Finally, we introduce the cost-related parameters: assume that the cost of repair $CR(x)$ is a function of the degradation level x. Denote the inspection cost of a single item by c_I and the further costs needed to identify a false positive case by c_{FP}. Define a cycle in the chain as the propagation from being deterioration free till repair and denote by T_L the observation period of an item. In this case, the maximal number of inspections in the observation period is given by

$$K = \text{int}((T_L - \tau_1)/\Delta) + 1, \ \text{ if } \ \tau_1 \leq T_L \ \text{ and } \ 0 \ \text{ otherwise.}$$

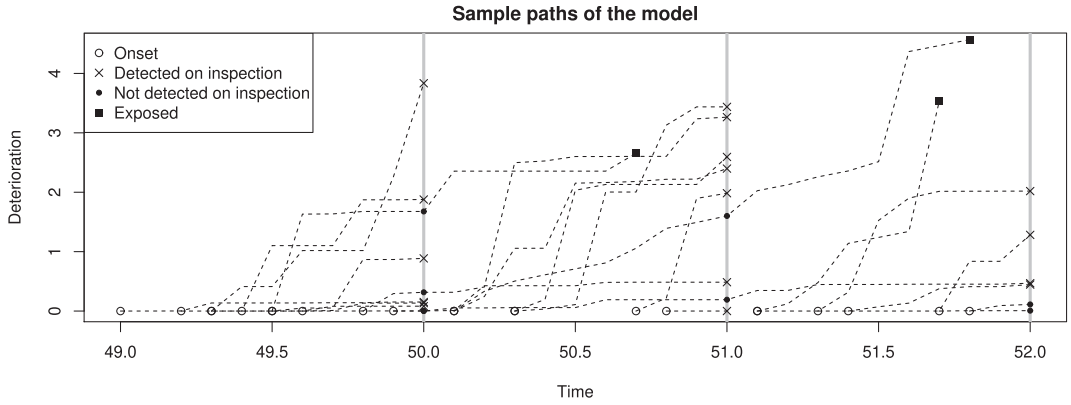

FIGURE 2 Sample paths of the model under the current setup

If more than one cycle is considered, then we assume that the repair of an item starts a new cycle. In other words, items whose degradation was detected will start a new inspection round of length T_L, where the first inspection after repair occurs τ_1 time units after detection. Hence, the item keeps propagating in the chain until another repair or the end of the observation period.

In this paper, three scenarios are considered: in Section 2.1.1, it is assumed that there is no recurrence; that is, the cost will be computed for one cycle through the chain. In this case, a repaired case goes out of the system. This serves as a building block for the other cases. In Section 2.1.2, recurrence is allowed and the expected cost for an infinite time horizon is derived. Finally, in Section 2.1.3, an upper bound of the expected cost within a given constant time period is derived.

Note that we do not consider failure in this setup; however, the model is quite flexible and can be modified to incorporate a failure mechanism. For example, one can change the exposed deterioration state to a failure state; in this case, inspection aims to prevent cases from reaching S_e and failing.

2.1 | Expected costs

Because we assumed the cost of repair to be a function of the deterioration level at detection, establishing the distribution of the size of detected damage on and between inspections will directly give the expected costs of repair of damaged items. Let us start by the cost of a single cycle; that is, here, we will only consider items moving to S_h once. This will serve as a building block for other cases. The derivation of the distributions of the deterioration on and between inspections in the first cycle is shown in Appendix A.

2.1.1 | Expected cost of a single cycle

For a single cycle, the expected costs of repair of deteriorations are derived in Appendix A. For those detected at inspection $E(C_D(\tau_i))$, see Equation (A5), and for exposed degradation $E(C_E(\tau_i))$, see Equation (A10). To determine the cost of inspection for a single cycle, note that repaired items will not be inspected again because repair ends the cycle. Then, the cost of inspection in a single cycle follows from the number of items participating in each inspection. For that purpose, define $P_{\tau_i}(S)$ as the probability of participating in an inspection at τ_i. Participating in an inspection means that the item's deterioration has not been previously detected. Hence, $P_{\tau_i}(S)$ is given by

$$P_{\tau_i}(S) = P_{\tau_1}(S) - \sum_{j=1}^{i-1} D(\tau_j) - \sum_{j=2}^{i} I(\tau_j) \quad , i = 2, \ldots, K,$$

where $P_{\tau_1}(S) = 1 - I(\tau_1)$ is the probability of participating in the first inspection.

The expected cost of inspection at τ_i denoted by $\mathrm{E}(C_{ins}(\tau_i))$, which is equal to $c_I \cdot P_{\tau_i}(S)$. On the other hand, the expected cost to identify a false positive at inspection τ_i, denoted by $E(C_{FP}(\tau_i))$, follows from the probability of being deterioration free at τ_i. Namely:

$$\mathrm{E}\left(C_{FP}(\tau_i)\right) = R_{FP} \cdot c_{FP} \cdot \left(1 - \int_0^{\tau_i} f_{T_d}(t_d) dt_d\right).$$

Putting everything back together, the total expected cost for a single cycle ($E(V^{(1)}(\tau_1, \Delta))$ for a given T_L is a function of the time of the first inspection and the interinspection time and is given by

$$\mathrm{E}\left(V^{(1)}(\tau_1, \Delta)\right) = \sum_{i=1}^{K+1} \mathrm{E}(C_E(\tau_i)) + \sum_{i=1}^{K} \left[\mathrm{E}(C_D(\tau_i)) + \mathrm{E}(C_{FP}(\tau_i)) + \mathrm{E}(C_{ins}(\tau_i))\right], \tag{3}$$

where $C_E(\tau_{K+1})$ is the cost of repair of items of showing symptoms between the last inspection τ_K and the end of the observation period T_L. The total expected cost is minimized in order to find the optimal first inspection time $\hat{\tau}_1$ and the optimal interinspection time $\hat{\Delta}$.

2.1.2 | Recurrence with an infinite time horizon

Assume now that we have an infinite time horizon, and let us compute the total cost based on the cycles for $r < 1$; that is, we assume that infinitely many inspection rounds with observation period T_L are organized. Let $T_C = \inf\{t \geq 0; \quad X_t \geq C\}$ be the first hitting time of the threshold C. We may assume that T_L is chosen to be large enough; for example, $P(T_d + T_C > T_L | T_C < \infty, T_d < \infty) \approx 0$. This gives enough time for the detection of the deterioration, so all long-term costs will be taken into account. Besides, we can assume that items that do not show deterioration by T_L will never move to S_h, therefore will not be monitored anymore. This prevents unnecessary inspections of deterioration-free items.

Proposition 1. *Let us denote the expected cost of the ith cycle by $E(V^{(i)}(\tau_1, \Delta))$. Then $E(V^{(i)}(\tau_1, \Delta)) \leq r \cdot E(V^{(i-1)}(\tau_1, \Delta))$.*

Proof. For a large enough T_L, the probability of an item to be repaired in the first cycle is bounded from above by the probability of starting to deteriorate; that is, $P(T_d < T_L) \leq r$. Then in expectation, at most r percent of items propagate through the chain in the second cycle. As we assumed "as good as new" repair, the propagation algorithm between states in the second cycle is identical to that of the first cycle. Therefore, the expected cost of the second cycle is at most $r \cdot \mathrm{E}\left(V^{(1)}(\tau_1, \Delta)\right)$. This is true for every cycle, so we get $E(V^{(i)}(\tau_1, \Delta)) \leq r \cdot E(V^{(i-1)}(\tau_1, \Delta))$. □

Let us denote by $E(V^{(\infty)}(\tau_1, \Delta))$ the total expected cost for an infinite time horizon. As a corollary of Proposition 1, we get

$$\mathrm{E}(V^{(\infty)}(\tau_1, \Delta)) \leq \sum_{i=0}^{\infty} r^i \cdot \mathrm{E}(V^{(1)}(\tau_1, \Delta)) = \frac{\mathrm{E}(V^{(1)}(\tau_1, \Delta))}{1 - r} \tag{4}$$

as a result; to get an optimal inspection program in this case, it is enough to optimize the inspections until the first repair and T_L. Note that equality in Equation (4) occurs when deteriorating items will be repaired in a cycle with probability 1.

2.1.3 | Cost within a given constant time period

Until now, we calculated costs on a cycle basis, not relating it to a fixed time point, and in practice, the latter is at least as important. For instance, screening programs or insurance policies may be valid for a predetermined constant time period, so their cost/price should be based on such a calculation. In order to get the expected cost for a given period $(0; T_y)$, the distributions of the deterioration on inspections and between inspections for the subsequent cycles need to be established, as the costs in $(0; T_y)$ may come from several cycles. Starting with the second cycle, the probability of detecting a deterioration at the first inspection is built from items that were repaired in the first cycle, moved to S_h again, and stayed there till they were detected by inspection. Hence, one needs to calculate separate contributions to the probability based on the time of repair as well as the time of the onset. Note that for the time period $(0; T_y)$, the maximal number of inspections K' in the first cycle is given by

$$K' = \mathrm{int}(\min(T_L, T_y) - \tau_1)/\Delta) + 1, \quad \text{if } \tau_1 \leq \min(T_L, T_y) \text{ and } 0 \text{ otherwise},$$

and the expected cost for the first cycle in the time period T_y, denoted by $V'^{(1)}$, is computed by substituting K by K' in Equation (3).

In order to get the expected costs, the contributions should be computed for inspections in several cycles. As a result, computing the exact costs would be infeasible, especially as it is usually not possible to get closed forms of the integrals in Equations (A5) and (A10). Instead of the computation of the distributions, we give an upper bound to the cost, based on U (the number of needed repairs until T_y) in Propositions 4 and 5. But first, we need some preparations. Let us denote the total expected cost for time period $(0, T_y)$ by $\mathrm{E}(V'^{(T_y)}(\tau_1, \Delta))$.

Proposition 2. *The expected cost for a longer observation period $T_{L'} > T_L$, denoted by $E(W^{(1)}(\tau_1, \Delta))$, is larger or equal to $E(V^{(1)}(\tau_1, \Delta))$, the cost associated with T_L.*

Proof. For the same parameters (τ_1, Δ), increasing the observation period to $T_{L'}$ results in additional costs for possible extra inspections, for identifying false positive items, and for the repair of detected deteriorations in $T_{L'} - T_L$. Hence, $E(W^{(1)}(\tau_1, \Delta)) \geq E(V^{(1)}(\tau_1, \Delta))$. □

Proposition 3. *Let us assume that the maximum number of repairs until T_y is u, then*

$$\mathrm{E}\left(V'^{(T_y)}(\tau_1, \Delta)\right) \leq \sum_{i=0}^{u-1} r^i \cdot \mathrm{E}(V'^{(1)}(\tau_1, \Delta)) \tag{5}$$

Proof. If the time horizon is extended into u cycles each of observation periods $\min(T_y, T_L)$, then by Section 2.1.2, an upper bound for the expected cost can be given by the terms of the expected cost for an infinite time horizon as $\mathrm{E}\left(V'^{(T_y)}(\tau_1, \Delta)\right) \leq \sum_{i=0}^{u-1} r^i \cdot \mathrm{E}(V'^{(1)}(\tau_1, \Delta))$. □

Proposition 4. $\mathrm{E}\left(V'^{(T_y)}(\tau_1, \Delta)\right) \leq (1 + M)\,\mathrm{E}\left(V'^{(1)}(\tau_1, \Delta)\right)$, *where* $M = \sum_{i=1}^{\infty} P\left(\sum_{j=1}^{i} t_d^{(j)} < T_y\right)$.

Proof. We get an upper bound of the expected costs using the maximum number of cycles within T_y along with the maximum expected cycle cost. First, observe that M is the sum of the probabilities that the deterioration is triggered i times until T_y, assuming all triggered deteriorations are detected immediately. This becomes an upper bound of the expected number of repairs. Namely, □

$$\mathrm{E}(U) \leq M = \sum_{i=1}^{\infty} P\left(\sum_{j=1}^{i} t_d^{(j)} < T_y\right), \tag{6}$$

where $t_d^{(j)}$ is the jth onset time and the probabilities $P\left(\sum_{j=1}^{i} t_d^{(j)} < T_y\right)$ can be derived using the jth order convolution of onset times.

For a deteriorating item repaired at time t, the next cycle will have an observation period of $\min(T_y - t, T_L)$, so by Proposition 2, the expected cost of the $(i + 1)$th cycle is less than or equal that of the ith one. Then, one can get an upper bound of the expected cost by multiplying the maximum expected number of cycles until T_y by the maximum expected cycle cost; thus,

$$\mathrm{E}\left(V'^{(T_y)}(\tau_1, \Delta)\right) \leq (1 + M)\,\mathrm{E}\left(V'^{(1)}(\tau_1, \Delta)\right). \tag{7}$$

Proposition 5. *For a given maximum number of needed repairs u until T_y, $(1 + M) \leq \sum_{i=0}^{u-1} r^i$.*

Proof. The probability of becoming onset i times within T_y: $P\left(\sum_{j=1}^{i} t_d^{(j)} < T_y\right) \leq r^i$. Hence, $(1 + M) \leq \sum_{i=0}^{u-1} r^i$. □

Note that because arbitrarily many repairs may be needed in $(0, T_y)$, higher order convolutions of the onset time are needed to compute M. However, one can stop at order u if the contribution to the cost beyond that is negligible.

Although perfect repair is assumed in this model, it is important to note that the assumption of "as good as new" repair (with age 0) is applicable in many industrial applications, for example, repairing a house after termite infestation. However, in healthcare applications, this assumption clearly does not hold. The model in its current form can be used only to optimize the screening program based on the total expected costs of the first treatment. Recurrence of the disease is usually influenced by age, treatment efficacy, and many other factors. Therefore, this implies a different triggering event distribution after repair. Hence, in order to compute the costs in this case, one must modify all cost elements after treatment, but we do not pursue this idea in this paper.

3 | APPLICATION

3.1 | Model

The model is applied to a hypothetical breast cancer screening program, as in that case a patient starts disease free and does not know about the illness until screening or symptoms (Zelen and Feinleib[11]). A good measure of the degradation at detection is the tumor size. The choice of the gamma process to model tumor growth is motivated by the results of Norton,[12] who suggested using a Gompertz function in which successive doublings occur at increasingly longer intervals, combined with the results of Speer et al,[13] who proposed modeling by a more generalized approach using the Gompertzian kinetics and included noisiness in the growth process.

A gamma process may fit these properties nicely. The shape $\eta(t) = m_1(1 - \exp(-m_2 t))$ with parameters m_1 and m_2 is adopted, which is the log of the shape used by Norton[12] assuming a starting cell size $N(0) = 1$. However, because the chosen $\eta(t)$ is bounded, there is a positive probability that the tumor will never reach the threshold. This leads to overdiagnosis, which is defined as a diagnosed condition that would otherwise not go on to cause symptoms. Reported estimates of breast cancer overdiagnosis range from 0% to 54%.[14] From a financial point of view, screening might lead to the diagnosis and treatment of cases that would have not needed any. On the other hand, this also implies that $E[T_C] = \infty$.[15] However, one is still able to get realistic estimates for the mean sojourn time by truncating the distribution of T_C at the maximum realistic value of the sojourn time.

The distribution of the onset age is chosen to be lognormal $LN(\mu, s^2)$. The choice of the lognormal distribution is based on the results of Lee and Zelen,[16] who found that the transition probability of breast cancer to the preclinical state is right skewed with a heavy tail. The lifetime risk of breast cancer is estimated by Feuer et al[17] to be around 12.5%; we used $r = 20\%$ to mimic high risk-based sampling. The density of the onset time T_d is

$$f_{T_d}(t_d) = \frac{0.2}{t_d s\sqrt{2\pi}} \exp\left(-\frac{(\ln(t_d) - \mu)^2}{2s^2}\right) + 0.8 \cdot \delta_{+\infty}(t_d), \quad t_d > 0.$$

Moreover, using the results of Michaelson et al,[18] let us assume that the sensitivity of the screen has a logistic form depending on the tumor size x:

$$\Lambda(x) = \frac{1}{1 + \exp(-b_0 - b_1 x)}, \quad x \geq 0;$$

because the logistic function takes values between 0 and 1 and is monotonically increasing in x, it is suitable for modeling sensitivity. The parameter b_0 determines the location of the curve, whereas b_1 is the growth rate or the steepness of the curve.

For the cost specifications, the results of Sun et al[19] show that the treatment costs of breast cancer are increasing with the disease stage at diagnosis. However, there are significant differences in treatment costs. Breast cancer treatment procedure is complex and adaptive, and there are different treatment protocols. Keeping things simple, we used a simple quadratic cost function (expressed in US$) based only on tumor diameter (cm); namely,

$$CR(x) = 10000(1 + x)^2, x \geq 0, \tag{8}$$

which is an increasing function that heavily penalizes larger tumors as well as giving a cost for small ones. The cost of a single inspection also varies with the type of exam (MRI, mammogram, or others).[20] We assume that a screen costs $C_I = \$300$. On the other hand, the false positive rate was chosen as $R_{FP} = 5\%$ with additional cost of identifying a false positive case as $c_{FP} = \$550$.[21]

The cost is minimized for one cycle $T_L = 100$ years using the function *nlm* in the statistical software ***R***. We also examine the costs for different fixed time intervals of the type $(0, T_y)$ with the aim of investigating the methods of Section 2.1.3.

Under the specified parametrizations, the integrals in Equations (A5) and (A10) do not have closed forms. Thus, the integration has to be evaluated numerically. The numerical integration can be carried out in ***R*** by the function *suave* from the package *cubature*.[22] However, high-dimensional numerical integration is computationally expensive, especially if one desires high accuracy. In order to lower the computation time, we suggest to limit the dimension of integrals based on a lag parameter α satisfying

$$\frac{\int_0^\infty f_{(i,i-\alpha)}(x)dx}{\int_0^\infty f_{\tau_i}(x)dx} < \epsilon,$$

where ϵ is a predefined value such that the chance that an item goes through α inspections and is not detected is very small. In this case, the contribution to subsequent inspections after α intervals can be omitted. We have found that $\epsilon = 0.01$ is suitable for the specific set of model parameters as it provides a balance between accuracy and speed of integration.

Another approach is to minimize the cost based on simulations, which might be more efficient than numerical integration. For that purpose, 100 000 cases were simulated in ***R***; for $r = 20\%$, this results in around 20 000 breast cancer patients. The onset age (T_d) is simulated by a lognormal random variable. After the deterioration is triggered, tumor growth is simulated by small increments over intervals of length 0.01 years by gamma random variables with scale β and shape:

$$\eta(t - (t_d + 0.01i)) - \eta(t - (t_d + 0.01(i-1))), \text{for } i = \frac{t_d}{0.01}, \ldots, \frac{T_L}{0.01}.$$

The critical threshold is simulated by a gamma random variable with shape Λ and scale ξ. After having the growth data, the inspections were simulated for items that did not cross their respective threshold using Bernoulli trials with $\Lambda(x)$ as the probability of success. Next, the cost of repair is calculated based on the deterioration at inspection or after crossing the threshold (including overshooting). The cost of inspection follows from the number of cases participating in the inspection, and the cost of false positives is calculated based on the number of erroneously identified healthy cases on each inspection. We carry out the minimization based on both numerical integration and simulation for the benchmark scenario as a comparison and minimize the cost of the rest of the scenarios based on simulations.

We have two aims: the first is to optimize screening for a benchmark scenario. The chosen parameters for this scenario are $\mu = 4$ and $s = 0.1$; these values result in an average age of transition into the hidden state to be 55 years old with a variance of 35. The parameters for the process are defined as $m_1 = 8$, $m_2 = 0.2$, and $\beta = 2.5$. The threshold parameters are $\Lambda = 4.5$ and $\xi = \frac{2}{3}$, leading to the mean of the critical threshold to be 3 cm with variance 2. This results in a mean sojourn time of 2.273 years (note that we truncated the distribution of sojourn time at 20 years to get a finite mean) close to the estimate of Duffy et al[23] (2.3 years), and the probability of never reaching the threshold is around 0.0077. The sensitivity parameters are $b_0 = -2.5$ and $b_1 = 3.5$, resulting in a steep sensitivity curve seen in Figure 3.

The second aim is to check the effects of changing the parameters governing the propagation in the chain on the optimal inspection program. For that reason, eight new scenarios are simulated, where one parameter is modified and others are kept at their baseline levels; the parameters for the scenarios S1–S8 are shown in Table 2. An overview of these scenarios can be seen in Figure 3.

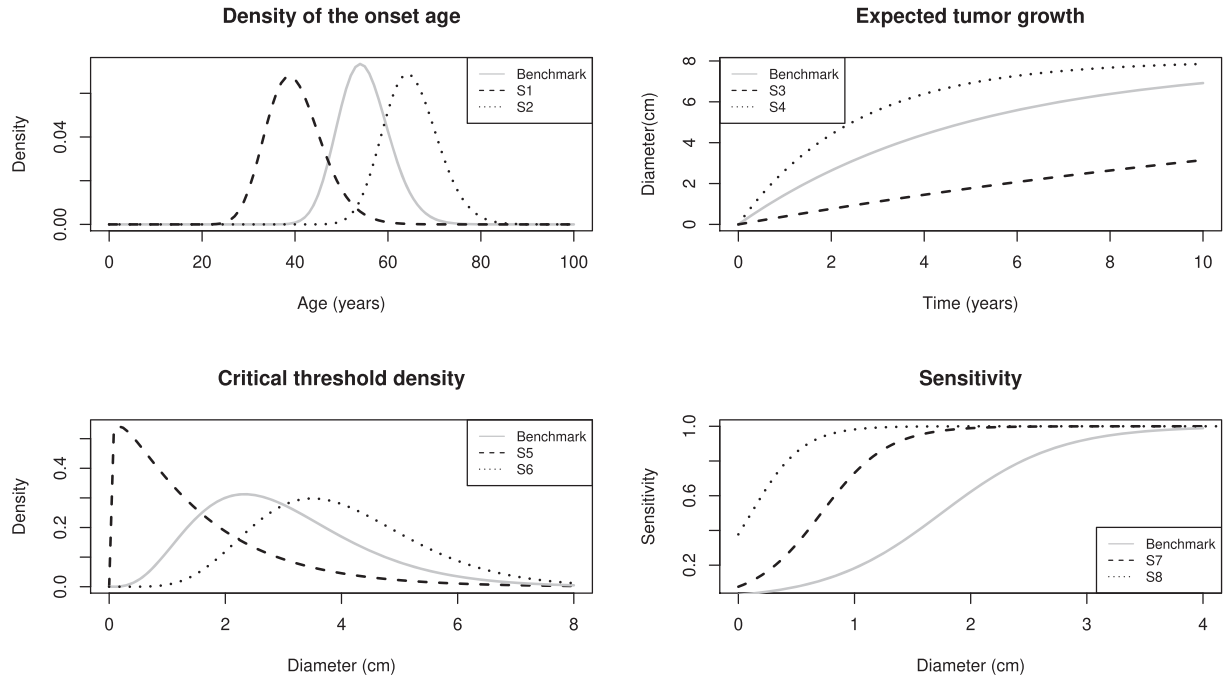

FIGURE 3 Overview of the change in parameters resulting in scenarios S1–S8

TABLE 2 Overview of the changed parameters in S1–S8

Benchmark	Changed parameters	Description
S1	$\mu = 3.68, s = 0.15$	Lowered average onset age to 40 years
S2	$\mu = 4.17, s = 0.09$	Raised average onset age to 65 years
S3	$m_2 = 0.05$	Decreased tumor growth rate
S4	$m_2 = 0.4$	Increased tumor growth rate
S5	$\Lambda = 1.125, \xi = 4/3$	Lowered average critical threshold to 1.5 cm
S6	$\Lambda = 8, \xi = 0.5$	Raised average critical threshold to 4 cm
S7	$b_0 = -3.5, b_1 = 2$	Decreased screening sensitivity
S8	$b_0 = -0.5, b_1 = 4.5$	Increased screening sensitivity

3.2 | Results

Let us start by discussing the results of the benchmark scenario. The first two lines of Table 3 show the costs under the optimal inspection plan that were obtained using simulation (line 1) and numerical integration (line 2). The results are quite close, the differences are likely caused by simulation and numerical integration errors. This shows that our model indeed performs well. The optimal inspection plan is based on starting screening when a patient is almost 46 years old and a periodic screen every 1.7 years. Screening seems to be cost efficient as it reduces the expected cost of treatment by around 24% compared with the cost without screening.

If we do not include overshooting in the model, that is, we base the cost of repair of symptomatic items on the threshold rather than the size after the simulations reached C, then, from the third line in Table 3, we notice that optimal screening starts almost a year later (46.9) with the same interinspection time. The cost composition is also noticeably changed, the biggest difference being in the cost of treatment of symptomatic cases, where the cost is almost doubled when including overshooting.

Figure 4 shows the contour plot of the expected cost; a dramatic increase in the cost is noticed when the interscreening time and the age at first inspection are too low (lower left part of the figure). Besides, it seems that there is a wide region around the optimum (black dot) in which the cost remains close to optimal. In that region, starting screening a bit earlier or later only slightly affects the cost, whereas the cost is pretty sensitive to the interscreening time. Note that we chose $\alpha = 3$ because we noticed that most cases are detected after three screens for $\Delta = 1$. In other words, the contribution to the probability of detection made by those cases that progressed for more than 3 years before being screened is almost 0.

The results of scenarios 1–8 are shown in Table 3, starting by S1 and S2. It seems that the average age of transition into S_h has a very strong positive correlation with the optimal time of the first screen τ_1; the difference in the optimal expected costs of the two scenarios is due to the inspection starting earlier in S1. This is also the reason behind the shifts in the optimal interscreening time Δ (1.8 years for S1 and 1.5 years for S2).

A lower rate of growth (S4) seems to affect τ_1, moving it to 47.5 years. Besides, an increase in the interscreening time is observable (2.5 years). The likely reason for this is that there is no gain in screening a slow growing tumor more often as it will be hard to detect; it is also noticed that screening is significantly more cost effective when tumors grow slower, saving around 50% compared with the cost without screening. On the other hand, a higher rate of growth causes a lower interscreening time, which is expected for a more aggressive tumor growth; the optimal interscreening time is every 1.7 years, but the expected savings by screening are reduced to around 10%.

The critical threshold parameters seem to be the most influential; a lower average critical threshold causes a huge increase in Δ. The optimal interscreening time for a low threshold is every 4.6 years. Besides, it is noticeable that here the costs are much lower and that expected savings by screening are reduced to around 3%. This is very interesting because it means that there is no significant difference in the cost compared with the cost without screening. This implies that spreading awareness about the symptoms could be more efficient than screening itself. On the other hand, a higher critical threshold causes a large increase in the costs, because symptomatic cases now have bigger tumors and cost more to treat. It is noticed that screening starts earlier at 45.5 years to prevent cases from reaching their threshold and that screening is more frequent (every 1.3 years). Screening in this case is very cost effective, saving around 41% compared with the cost without screening.

The last two lines of Table 3 show that screening sensitivity plays an important role on the optimal τ_1 and Δ: lower screening sensitivity means that small tumors are less likely to be detected by screening, and tumors then need to be larger to have a higher chance of detection resulting in an interscreening time of 2.4 years and a later age of starting inspection (47.4 years). The expected savings are also relatively low (around 10% compared with the cost without screening). On the

TABLE 3 Results for all scenarios

Scenario	Description	Optimal τ_1	Optimal Δ	Optimal total cost	Cost without screening	Cost of treatment of screened cases	Cost of treatment of symptomatic cases	Cost of screening	Cost of false positives
Benchmark	Benchmark scenario (simulation)	45.9	1.7	37 099.5	48 985.6	9673.5	17 248.1	9441.1	736.4
Benchmark	Benchmark scenario (numerical integration)	46.1	1.7	36 396.1	48 760.1	9375.7	16 954.2	9342.5	723.6
Benchmark	Benchmark scenario (no overshooting)	46.9	1.7	28 570.8	35 214.5	10 480.2	8541.6	8861.8	687.1
S1	Lower average age of transition to S_h	31.7	1.8	39 463.1	49 344.0	9649.7	17 880.3	11 068.8	864.3
S2	Higher average age of transition to S_h	56.9	1.5	35 523.1	49 874.6	9892.0	16 348.8	8611.5	670.9
S3	Lower rate of growth	47.6	2.5	24 778.3	49 022.9	9413.0	8641.3	6242.7	481.3
S4	Higher rate of growth	47.4	1.7	45 042.5	50 013.2	8019.8	27 236.0	9076.9	709.8
S5	Lower critical threshold	49.6	4.6	24 950.8	25 735.4	2632.3	18 847.8	3219.3	251.4
S6	Higher critical threshold	45.5	1.3	40 902.6	69 550.0	14 922.1	12 604.5	12 408.8	967.2
S7	Lower screening sensitivity	47.4	2.4	44 578.9	49 639.1	8557.8	29 036.2	6480.4	504.4
S8	Higher screening sensitivity	46.7	1.6	31 750.1	49 684.2	8517.8	12 461.7	9991.0	779.6

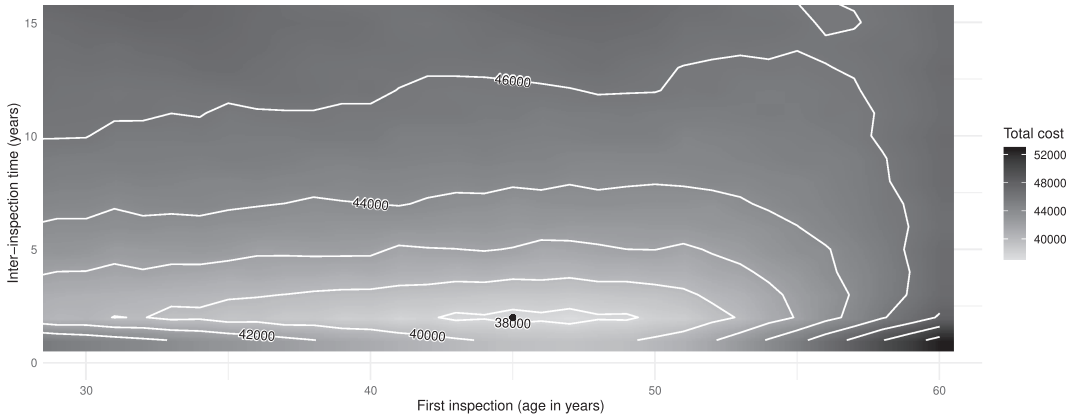

FIGURE 4 Contour plot of the expected total costs as a function of the first inspection time and interinspection time in the benchmark scenario; the optimum is achieved at the black dot

TABLE 4 Upper bounds of the costs assuming maximum of u recurrences for $T_y = 100$ in the benchmark scenario

Method	u	Multiplier	Total cost
$\sum_{i=0}^{u-1} r^i \cdot \mathrm{E}(V'^{(1)}(\tau_1, \Delta))$	2	1.24	\$46 003
	3	1.248	\$46 300
$(1+M)\mathrm{E}(V'^{(1)}(\tau_1, \Delta))$	2	1.2044	\$44 682
	3	1.2044	\$44 682

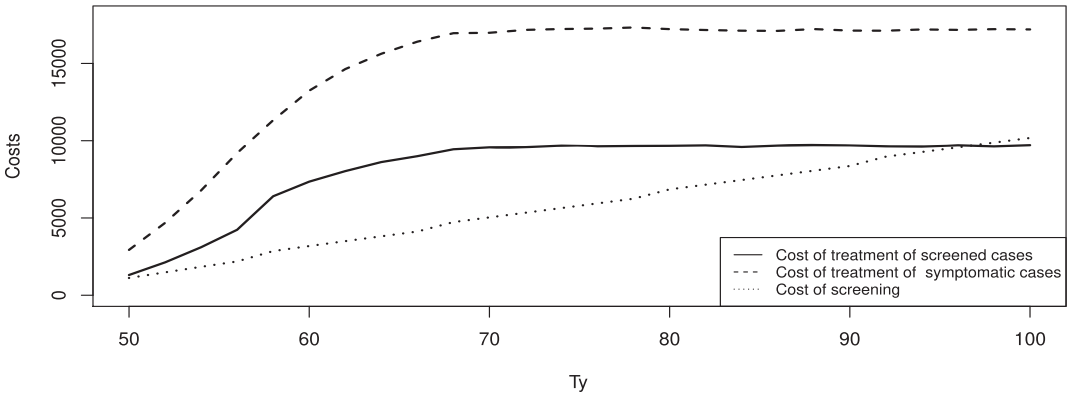

FIGURE 5 Costs for different time periods T_y in the benchmark scenario

other hand, higher sensitivity leads to a much more efficient program with expected savings of around 36% compared with the cost without screening. Screening also starts earlier as there is a greater chance to detect smaller tumors.

Regarding recurrence, it was found that the distribution of the time of the second onset is not the same of the first one and that there are complex factors governing the probability of recurrence, for example, tumor size at treatment and recurrence in other locations.[24] This means that our approach is inapplicable for more than one cycle. However, for the sake of demonstration, let us assume that onset after treatment follows the same distribution (lognormal) and suppose that the disease may be triggered a maximum of u times in the benchmark scenario for $T_y = 100$ years. The results are shown in Table 4; the first block of the table shows the bound based on Equation (5) assuming a maximum of two and three repairs, giving multipliers 1.24 and 1.248, respectively. Note that there is no significant increase in the cost beyond that. On the other hand, in the bottom block of the table, using the bound based on Equation (7) gives a smaller multiplier of 1.2044 for both cases, because the probability of three onsets within $(0, T_y)$ is negligible. Extending the model to different onset distributions and imperfect repair will be a future work.

Finally, Figure 5 shows the costs of treatment (symptomatic and screened cases) and the cost of inspections for different T_y. The figure shows that costs of treatment for inspection and symptomatic cases reach a peak around 70 years of follow-up and are almost constant after that. Hence, inspection is not needed beyond this point. In other words, it is enough to have the observation period to be 70 years. That being said, the choice of the follow-up period is a compli-

cated issue in real-life cases; death should also be taken into account as well as the difference in the cycles. This is just an illustration to show the properties of our model.

4 | CONCLUDING REMARKS

Our approach gave feasible models for cost-optimized inspection programs for the described degradation phenomena. We have demonstrated that both simulation- and numerical integration-based approaches were applicable in the investigated case. The model is also quite flexible; many specifications can be changed based on the desired inspection strategy.

Multiple extensions are possible; the case of nonperiodic inspection for a predefined number of inspections K is straightforward to incorporate. Namely, the expected cost can be minimized for $(\tau_1, \ldots, \tau_K)$. Besides, it is also possible to incorporate failure and base it on the level of degradation, which can be done by assuming an item fails when it reaches a fixed or a random critical threshold.

It is also possible to extend the form of η to include covariates as in Lawless and Crowder.[25] Besides, although we used a gamma deterioration process in the model, other types of degradation process can be used provided they are nondecreasing, have the Markov property and independent increments.

ACKNOWLEDGMENT

The work was supported by the project EFOP-3.6.2-16-2017-00015, which was supported by the European Union and cofinanced by the European Social Fund.

ORCID

Ayman Hijazy https://orcid.org/0000-0002-0812-3377
András Zempléni https://orcid.org/0000-0001-5773-5380

REFERENCES

1. Rosmaini A, Shahrul K. An overview of time-based and condition-based maintenance in industrial application. *Comput Ind Eng*. 2012;63(1):135-149.
2. Abdel-Hameed M. A gamma wear process. *IEEE Trans Reliab*. 1975;24(2):152-153.
3. Noortwijk JM. A survey of the application of gamma processes in maintenance. *Reliab Eng Syst Safe*. 2009;94(1):2-21.
4. Abdel-Hameed M. Inspection and maintenance policies of devices subject to deterioration. *Adv Appl Probab*. 1987;19(4):917-931.
5. Abdel-Hameed M. Correction to: Inspection and maintenance policies of devices subject to deterioration. *Adv Appl Probab*. 1995;27:584.
6. Kong MB, Park KS. Optimal replacement of an item subject to cumulative damage under periodic inspections. *Microelectron Reliab*. 1997;37(3):467-472.
7. Jia X, Christer A. A prototype cost model of functional check decisions in reliability-centred maintenance. *J Oper Res Soc*. 2002;53(12):1380-1384.
8. Kallen MJ. Optimal maintenance decisions under imperfect inspection. *Reliab Eng Syst Safe*. 2005;90(2–3):177-185.
9. Yang Y., Sørensen JD. Cost-optimal maintenance planning for defects on wind turbine blades. *Energies*. 2019;12(6):998.
10. Dobi B, Zempléni A. Markov chain-based cost-optimal control charts for health care data. *Qual Reliab Engng Int*. 2019;35:1379-1395.
11. Zelen M, Feinleib M. On the theory of screening for chronic diseases. *Biometrika*. 1969;56(3):601-614.
12. Norton L. A Gompertzian model of human breast cancer growth. *Cancer Res*. 1988;48:7067-7071.
13. Speer JF, Petrovsky VE, Retsky MW, Wardwell RH. A stochastic numerical model of breast cancer that simulates clinical data. *Cancer Res*. 1984;44:4124-4130.
14. Elmore JG, Fletcher SW. Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. *Annal Int Med*. 2012;156(7):536-537.
15. Paroissin C, Salami A. Failure time of non homogeneous gamma process. *Commun Stat Theo Method*. 2014;43(15):3148-3161.
16. Lee SJ, Zelen M. Scheduling periodic examinations for the early detection of disease: applications to breast cancer. *J Amer Stat Assoc*. 1998;93:1271-1281.
17. Feuer EJ, Wun LM, Boring CC, Flanders WD, Timmel MJ, Tong T. The lifetime risk of developing breast cancer. *J Natl Cancer Inst*. 1993;85(11):892-7.
18. Michaelson J, Satija S, Moore R, Weber G, Halpern E, Garland A, Kopans D, Hughes K. Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds. *J Women's Imag*. 2003;5(1):3-10.
19. Sun L, Legood R, Dos-Santos-Silva I, Gaiha SM, Sadique Z. Global treatment costs of breast cancer by stage: a systematic review. *PloS one*. 2018;13(11).

20. Moore SG, Shenoy PJ, Fanucchi L, Tumeh J, Flowers CR. Cost-effectiveness of MRI compared to mammography for breast cancer screening in a high risk population. *BMC Health Serv*. 2009;9(9).
21. Chubak J, Boudreau DM, Fishman PA, Elmore JG. Cost of breast-related care in the year following false positive screening mammograms. *Med Care*. 2010;8(9):815-20.
22. Hahn T. Cuba—a library for multidimensional numerical integration. *Comput Phys Commun*. 2005;168:78-95.
23. Duffy S, Chen H, Tabar L, Day N. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Stat Med*. 1995;14:1531-1543.
24. Lafourcade A, His M, Baglietto L, Boutron-Ruault MC, Dossus L, Rondeau V. Factors associated with breast cancer recurrences or mortality and dynamic prediction of death using history of cancer recurrences: the French E3N cohort. *BMC cancer*. 2018;18(1):171.
25. Lawless J, Crowder M. Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Anal*. 2004;10(3):213-227.
26. Kahle W, Mercier S, Paroissin C. Degradation Processes in Reliability. 1st ed., vol. 3, ISTE Ltd., 97–100; 2016.

AUTHOR BIOGRAPHIES

András Zempléni is an associate professor and the head of the Department of Probability Theory and Statistics at Eötvös Loránd University (ELTE), Budapest. He is also a senior research fellow at the University of Debrecen. He has a PhD in Mathematics by ELTE. His main research interests include statistical process control, medical applications, and extreme value models.

Ayman Hijazy is a PhD student at the Department of Probability Theory and Statistics at Eötvös Loránd University (ELTE), Budapest. He is also an assistant research fellow at the University of Debrecen. His main research areas is biostatistics and statistical modeling.

How to cite this article: Hijazy A, Zempléni A. Optimal inspection for randomly triggered hidden deterioration processes. *Qual Reliab Engng Int*. 2020;36:2660–2675. https://doi.org/10.1002/qre.2707

APPENDIX A: DISTRIBUTIONS OF THE DEGRADATION IN THE FIRST CYCLE

A.1 | Cost of repair of early detected deterioration

The cost of repair $CR(x)$ is a function of the level of degradation x; so to derive the expected cost of repair of inspected items, the probability distribution of the detected deterioration at inspections are needed. Starting with items detected on the first inspection (τ_1), their deterioration must have been triggered before τ_1, have not showed symptoms (did not hit their threshold until τ_1) and their inspection had to be successful.

Because the random threshold, the onset, and the process are assumed independent, for a given onset time $T_d = t_d$, the conditional probability of detecting level of deterioration between x and $x + dx$ at τ_1 is obtained by the density of the deterioration increments in (t_d, τ_1) weighted by the sensitivity $\Lambda(x)$ as it was detected by inspection together with the probability of not hitting the threshold before τ_1. Consequently, it is given by

$$f_{\tau_1|t_d}(x) = f_{X_{\tau_1-t_d}}(x) \cdot \Lambda(x) \cdot P(C > x) \cdot \mathbb{1}_{\mathbb{R}_+}(x). \tag{A1}$$

Hence, $f_{\tau_1}(x)$ is derived by applying the law of total probability to Equation (A1); therefore,

$$f_{\tau_1}(x) = \Lambda(x) \cdot P(C > x) \cdot \int_0^{\tau_1} f_{T_d}(t_d) \cdot f_{X_{\tau_1-t_d}}(x)\, dt_d \cdot \mathbb{1}_{\mathbb{R}_+}(x).$$

In general, for the detected level of degradation at the ith inspection τ_i, one will have to divide the timeline before τ_i into inspection intervals (τ_{j-1}, τ_j) for $j = 1, \ldots, i$ and apply the law of total probability to derive their contributions to f_{τ_i}. Let us

denote by $f_{(i,j)}$ the contribution of cases moving to S_h between (τ_{j-1}, τ_j) to the probability of the detected level of degradation at inspection τ_i to be between $(x, x+dx)$. Hence, $f_{(i,j)}$ is built up from items that had their triggering event between (τ_{j-1}, τ_j); their deterioration was not detected by inspection, and they did not show symptoms before τ_i when they were successfully inspected. Using the Markov property of the process and applying the law of total probability, $f_{(i,j)}$ can be computed as

$$\begin{aligned} f_{(i,j)}(x) = \int_{\tau_{j-1}}^{\tau_j} \int_0^x \int_{x_j}^x \cdots \int_{x_{i-2}}^x \Big[& f_{T_d}(t_d) f_{X_{(0,\tau_j - t_d)}}(x_j) \cdot (1 - \Lambda(x_j)) \cdot \\ & f_{X_{(\tau_j - t_d, \tau_{j+1} - \tau_j)}}(x_{j+1} - x_j) \cdot (1 - \Lambda(x_{j+1})) \ \cdots \ f_{X_{(\tau_{i-2} - t_d, \tau_{i-1} - \tau_{i-2})}}(x_{i-1} - x_{i-2}) \cdot (1 - \Lambda(x_{i-1})) \\ & f_{X_{(\tau_{i-1} - t_d, \tau_i - \tau_{i-1})}}(x - x_{i-1}) \cdot \ dx_{i-1} \ \cdots \ dx_{j+1} \ dx_j \ dt_d \Big] \cdot \Lambda(x) \cdot P(C > x) \mathbb{1}_{\mathbb{R}_+}(x) \, dx. \end{aligned} \tag{A2}$$

Thus, f_{τ_i} is given as

$$f_{\tau_i}(x) = \sum_{j=1}^{i} f_{(i,j)}(x), \tag{A3}$$

where $i = 1, \ \dots \ , K$. Consequently, the probability of being detected at inspection τ_i is given by

$$D(\tau_i) = \int_0^\infty f_{\tau_i}(x) \, dx. \tag{A4}$$

Finally, the expected cost of repair of cases detected at inspection τ_i, which we denoted by $E(C_D(\tau_i))$ is directly obtainable as

$$\mathrm{E}(C_D(\tau_i)) = \int_0^\infty CR(x) \cdot f_{\tau_i}(x) \, dx. \tag{A5}$$

A.2 | Cost of repair of exposed deterioration between inspections

A similar approach is adapted for symptomatic degradation. However, because the gamma process is a pure jump process, the threshold C is crossed by a jump; this exceedance $X_{T_C} - C$ is known as the overshoot[26] of the gamma process. Denote by $X_{T_C^-}$ the left-side limit of $(X_t)_{t \geq 0}$ at time T_C and by X_{T_C} the level of deterioration after crossing the threshold.

Assuming that $\eta(t)$ is almost surely differentiable and setting $a(t) = \eta'(t)$, Kahle et al[26] derived the joint probability density function of $(T_C, X_{T_C^-}, X_{T_C})$, for a given $C = c$. Translating their results into our setup, we get the joint defective density for a given onset t_d and threshold c as

$$f_{(T_c, X_{T_c^-}, X_{T_c} | T_d = t_d, C = c)}(s, x, z) = \mathbb{1}_{s > t_d}(s) \mathbb{1}_{0 < x < c \leq z} f_{X_{s - t_d}}(x) \frac{\exp\left(-\frac{z-x}{\beta}\right)}{z - x} a(s - t_d). \tag{A6}$$

Let us start by dealing with items showing symptoms before the first inspection; for a given threshold $C = c$, these have their deterioration triggered before τ_1 and their deterioration level crosses c before τ_1. Applying the law of total probability over t_d to Equation (A6), the defective joint density of $(T_C, X_{T_C^-}, X_{T_C})$ for a given threshold $C = c$ before the first inspection is given by

$$g_{\tau_1}(s, x, z) = \int_0^s f_{T_d}(t_d) f_{X_{s - t_d}}(x) \frac{\exp\left(-\frac{z-x}{\beta}\right)}{z - x} a(s - t_d) \, dt_d \ \mathbb{1}_{s < \tau_1}(s) \mathbb{1}_{0 < x < c \leq z}.$$

Hence, the cost of repair of exposed deterioration before the first inspection is given by

$$\mathrm{E}(C_E(\tau_1)) = \int_0^\infty \int_c^\infty \int_0^c \int_0^{\tau_1} CR(z) \cdot g_{\tau_1}(s, x, z) \cdot f_C(c) \, ds \, dx \, dz \, dc.$$

For the subsequent inspection intervals, the joint defective density for a given $C = c$ denoted by g_{τ_i} is determined by the sum of contributions of items that were not detected by previous inspections. Hence,

$$g_{\tau_i}(s,x,z) = \sum_{j \le i} g_{(i,j)}(s,x,z), \tag{A7}$$

where $g_{(i,j)}$ is the defective joint density of $(T_c, X_{T_c^-}, X_{T_c})$ for cases moving into S_h between (τ_{j-1}, τ_j) and to S_e between (τ_{i-1}, τ_i) for a given $C = c$ and is given by

$$\begin{aligned} g_{(i,j)}(s,x,z) = \int_{\tau_{j-1}}^{\tau_j} \int_0^x \cdots \int_{x_{i-2}}^x & f_{T_d}(t_d) \cdot f_{X(0,\tau-t_d)}(x_j) \cdot (1-\Lambda(x_j)) \cdot \\ & f_{X(\tau_j,\tau_j+1),1}(x_{j+1}-x_j) \cdot (1-\Lambda(x_{j+1})) \ \cdots \ f_{X(\tau_{i-2},\tau_{i-1})}(x_{i-1}-x_{i-2}) \cdot (1-\Lambda(x_{i-1})) \\ & f_{X_{(\tau_{i-1},s-\tau_{i-1})}}(x-x_{i-1}) \cdot \frac{\exp\left(-\frac{z-x}{\beta}\right)}{z-x} \cdot a(s-t_d) \quad dx_{i-1} \ \cdots \ dx_{j+1} \ dx_j \ dt_d \mathbb{1}_{\tau_{i-1}<s<\tau_i} \mathbb{1}_{0<x<c\le z}. \end{aligned} \tag{A8}$$

Hence, the probability of an item to show symptoms between τ_{i-1} and τ_i has the following form:

$$I(\tau_i) = \int_0^\infty \int_c^\infty \int_0^c \int_0^{\tau_1} g_{\tau_i}(s,x,z) \cdot f_C(c) \ ds \ dx \ dz \ dc. \tag{A9}$$

Finally, the expected cost of repair for cases moving into S_e in the interval (τ_{i-1}, τ_i) follows directly as

$$\mathrm{E}(C_E(\tau_i)) = \int_0^\infty \int_c^\infty \int_0^c \int_0^{\tau_1} CR(z) \cdot g_{\tau_i}(s,x,z) \cdot f_C(c) \ ds \ dx \ dz \ dc. \tag{A10}$$