Evaluating failure time probabilities for a Markovian wear process ☆

Jeffrey P. Kharoufeh[a],*, Jeffrey A. Sipe[b]

[a]Department of Operational Sciences, Air Force Institute of Technology, AFIT/ENS, Bldg 640, 2950 Hobson Way, Wright Patterson AFB OH 45433-7765, USA
[b]Headquarters Air Mobility Command, HQ AMC TACC/XONCY, 402 Scott Drive, Unit 2K1, Scott AFB IL 62225-5303, USA

Abstract

We present simplified analytical results for the numerical evaluation of failure time probabilities for a single-unit system whose cumulative wear over time depends on its external environment. The failure time distribution is derived as a one-dimensional Laplace–Stieltjes transform with respect to the temporal variable using a direct solution approach and by inverting an existing two-dimensional result with respect to the spatial failure threshold variable. Two numerical examples demonstrate that accurate cumulative probability values can be obtained in a straightforward manner using standard computing environments.

Scope and purpose

Reliability models that incorporate the effect of a stochastic and dynamic environment on a unit's lifetime have attracted a moderate amount of attention in the past decade. However, evaluating failure time probabilities using such models is nontrivial in all but a few cases. Kharoufeh [1] provided a closed-form lifetime distribution for a continuous Markovian wear process as a two-dimensional Laplace transform. The main purpose of this paper is to reduce the lifetime distribution to a one-dimensional Laplace transform in order to facilitate simpler numerical implementation.
Published by Elsevier Ltd.

☆ The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.
 * Corresponding author. Tel.: +1-937-255-6565; fax: +1-937-656-4943.
 E-mail address: jeffrey.kharoufeh@afit.edu (J.P. Kharoufeh).

1. Introduction

The need for an accurate assessment of a unit's reliability is apparent in many areas of operations research, and especially for the purpose of prescribing sound repair or replacement polices. Over the past four decades, parametric distributions (e.g., Weibull, exponential, normal) have been used to characterize the useful lifetime of single-unit systems. The assumption that the random lifetime follows a parametric distribution is usually employed as a result of goodness-of-fit tests using failure time observations obtained via accelerated life testing. However, these tests are usually performed in a controlled, static laboratory environment. Unfortunately, these testing conditions may not provide an accurate representation of the unit's true operating environment.

In light of these concerns, a relatively recent class of failure models has been proposed to capture the effects of the physical environment on the unit's cumulative degradation (wear) and lifetime. For instance, some recent models in the reliability literature, such as those due to Gillen and Celina [2], consider the degradation of materials. Meeker et al. [3] discuss general approaches to predicting lifetime distributions in accelerated life tests for highly variable environments. The models presented therein focus on specification of the degradation path as it depends explicitly on the environment. In that paper, the authors note the need for a formal stochastic-process model representing the ambient environment and for expedient numerical techniques for the estimation of relevant measures.

In the stochastic modelling community, numerous techniques focused on a stochastic-process approach to lifetime analysis have been reviewed in a cogent survey by Singpurwalla [4] who high-lights the most important contributions. The common thread running through each of these approaches is a mathematical characterization of the environment as a continuous-time stochastic process on a continuous or discrete state space. This approach is naturally appealing since it provides a means by which the stochastic evolution of the environment, and its ultimate effect on the operating unit, may be modelled. Singpurwalla [4] reviews failure models describing (i) the state (or wear) of the unit, (ii) the failure or hazard rate function, (iii) combinations of unit state and an external, co-variate driving process, (iv) stochastic shock models, and (v) a response variable (highly correlated with unit lifetime) modelled as a stationary, Gaussian process. As noted in [4], the computational requirements for implementation of these techniques remains a significant challenge.

In some cases, however, results exist in the recent literature as multidimensional Laplace transforms (see Kharoufeh [1]) that allow for the explicit calculation of the unit's lifetime distribution. The unit lifetime is closely related to the task completion time in the performability analysis of computer systems whose distribution also exists as a two-dimensional Laplace transform (cf. [5] and [6]). Numerical solutions for such models may be obtained in the time domain via multidimensional inversion of the Laplace transform. However, numerical procedures may be difficult to apply, can be computationally intensive and are, at times, unstable. For these reasons, it is desirable to obtain transform results in a single dimension, thereby allowing for the implementation of one-dimensional inversion algorithms, based on a Fourier-series representation, which have proven to be robust for a variety of problems. We shall further discuss such procedures in Section 4.

This paper provides simplified results for the numerical evaluation of the lifetime distribution of a single-unit system that accumulates wear over time due to the influence of its operating environment. The environment, which is often time-varying, is modelled as a continuous-time stochastic process denoted throughout by $\{Z(t): t \geqslant 0\}$. The wear rate of the unit depends explicitly on the state of its random environment. Assuming that repairs do not take place to restore the condition of

the system, the cumulative wear up to time t may be characterized by a nondecreasing stochastic process $\{X(t) : t \geqslant 0\}$. The system begins its lifetime in perfect working order but experiences wear under the influence of its random environment until the state of the system exceeds a fixed threshold value x, at which time it fails. We denote the time until failure by a random variable T_x. When the environment process is assumed to be a temporally homogeneous, continuous-time Markov chain (CTMC) on a finite state space, an explicit result exists for the unit lifetime distribution as a double Laplace transform (see Kharoufeh [1]). The two-dimensional result stems from the Laplace transform solution of a weakly-coupled matrix partial differential equation. The evaluation of numerical values for this model is nontrivial due to the computationally intensive (and sometimes unstable) process of multidimensional, numerical inversion. Hence, our aim is to simplify the procedure by analytically reducing the dimensionality of the problem and exploiting widely available single-dimension inversion algorithms.

The main result of this work is a closed-form, analytical expression for the system lifetime in the form of a one-dimensional Laplace transform as opposed to two dimensions. The distribution function can be computed through direct numerical inversion of the Laplace transform with respect to the temporal variable only. The simplified result is proven in two ways: (i) by directly solving an ordinary (matrix) differential equation via standard methods, and (ii) by analytically inverting the two-dimensional result of [1] with respect to the spatial dimension x. The moments of the system lifetime are also computed explicitly. The one- and two-dimensional distribution function results will be demonstrated via two numerical examples.

In the context of failure models in a dynamic environment, our model considers the state of the unit and an external covariate process (the environment process). Though the failure time distribution can be derived analytically, the issue of implementation remains a challenge. This work provides a viable approach that can be implemented in a relatively simplistic manner, particularly when compared with the previous two-dimensional results of [1]. We are able to compute failure time probabilities in a far more expedient manner with little to no degradation in performance. Moreover, our analytical result may be easily implemented with off-the-shelf software packages and known numerical inversion techniques (such as the algorithm in [7]).

The remainder of the paper is organized in the following manner. The next section reviews the formal mathematical model for the lifetime distribution as a two-dimensional Laplace transform. In Section 3, we present our main analytical result and prove the result in two ways. In Section 4, numerical examples are provided to compare the one- and two-dimensional results with simulated cumulative probability values. Finally, we provide some concluding remarks in Section 5.

2. Mathematical model

In this section, we briefly review the mathematical model and main results of Kharoufeh [1] from which the results of this paper will be derived. The single-unit system is subject to continuous and additive deterioration in time due to an explicit dependence on the state of an external random environment. Under normal operating conditions, the system accumulates wear until the magnitude of its cumulative wear exceeds a fixed threshold value x, at which time the system fails. Such failures are often referred to as "soft failures" (cf. Meeker and Escobar [8], p. 327). The rate of deterioration (or wear rate) of the system at time $t > 0$ is governed by a random environment that

is modelled as an ergodic, continuous-time Markov chain, $\{Z(t): t \geqslant 0\}$, on a finite state space $S := \{1, 2, \ldots, K\}$ where K is a positive integer. The evolution of the wear process can be described by a continuous-time stochastic process $\{X(t): t \geqslant 0\}$ that assumes values on the nonnegative real line denoted by $\mathbb{R}_+$.

The random variable $Z(t)$ denotes the state of the environment process at time $t \in \mathbb{R}_+$. Define $R(t)$ as the wear rate of the system at time $t \in \mathbb{R}_+$ and define a positive function $r : S \rightarrow \mathbb{R}_+ \setminus \{0\}$. The properties of the function $r(\cdot)$ are dictated by the type of system under consideration and its surrounding environment. Since the wear rate of the system is explicitly dependent on the environment process, the wear rate process $\{R(t): t \geqslant 0\}$ assumes values in the space $\mathscr{D} = \{r(1), \ldots, r(K)\}$. If $Z(t) = i \in S$, then $R(t) := r(Z(t)) = r(i) \in \mathscr{D}$. The environment transitions from state $i \in S$ to state $j \in S$, $j \neq i$, at time t according to a Markov transition function $\mathbf{P}(t) := [p_{i,j}(t)]$ where $p_{i,j}(t) := P\{Z(t) = j | Z(0) = i\}$. The infinitesimal generator matrix for the Z process shall be denoted by $\mathbf{Q}$, and its initial distribution vector shall be denoted by $\alpha := [\alpha_i]$ with $\alpha_i := P\{Z(0) = i\}$.

The cumulative wear process of the single-unit system, $\{X(t) : t \geqslant 0\}$ (abbreviated as X), is a continuous, additive functional of Z, and thus, (X, Z) constitutes a special case of a Markov additive process (cf. Çinlar [9]). The single-unit system will fail when the cumulative wear exceeds a fixed threshold level x. For this reason, the time until failure for the unit is a type of first passage time for the cumulative wear process X. Let $\Omega := \mathbb{R}_+ \times S$ denote the appropriate sample space for the Markov additive process $\{(X(t), Z(t)): t \geqslant 0\}$. For each sample path, $\omega \in \Omega$, the cumulative wear of the single-unit system up to time $t \in \mathbb{R}_+$ is defined by

$$X_\omega(t) = \int_0^t r(Z_\omega(u)) \, \mathrm{d}u. \tag{1}$$

For brevity, we shall suppress the dependence of (X, Z) on ω. The lifetime of the unit is defined by the random variable

$$T_x = \inf\{t : X(t) > x\}. \tag{2}$$

That is, the random variable T_x is that instant of time at which the degradation of the unit first exceeds the failure threshold value x. To obtain the distribution of this random variable, define the following joint probability distribution

$$V_{i,j}(x, t) = P\{X(t) \leqslant x, Z(t) = j | Z(0) = i\} \tag{3}$$

and the distribution matrix of $X(t)$ as $\mathbf{V}(x, t) = [V_{i,j}(x, t)]$. Due to the dual relationship of (2), it follows that $G(x, t) := P\{T_x \leqslant t\}$, the unconditional distribution of T_x, is given by

$$G(x, t) = 1 - \alpha \mathbf{V}(x, t) \mathbf{1} \tag{4}$$

where $\mathbf{1}$ denotes a K-dimensional column vector of ones. Let $\mathbf{V}_t(x, t)$ and $\mathbf{V}_x(x, t)$ denote the partial derivatives of $\mathbf{V}(x, t)$ with respect to t and x, respectively, and $\mathbf{R}_D := diag(r(1), r(2), \ldots, r(K))$ denote the diagonal matrix of wear rates. The following result from [1] is revisited here owing to its relevance to the main result of this paper.

Theorem 1 (Kharoufeh [1]). *The distribution matrix* $\mathbf{V}(x, t)$ *satisfies the matrix partial differential equation*

$$\mathbf{V}_t(x, t) + \mathbf{V}_x(x, t) \mathbf{R}_D = \mathbf{V}(x, t) \mathbf{Q}. \tag{5}$$

Proof. The result is proved via a standard conditioning argument. Let $\varepsilon > 0$. Then,

$$V_{i,j}(x, t + \varepsilon) = P\{X(t + \varepsilon) \leqslant x, Z(t + \varepsilon) = j | Z(0) = i\}$$

$$= \sum_k P\{Z(t + \varepsilon) = j | X(t + \varepsilon) \leqslant x, Z(t) = k, Z(0) = i\}$$

$$\times P\{X(t + \varepsilon) \leqslant x | Z(t) = k, Z(0) = i\} P\{Z(t) = k\}$$

$$= (1 + \varepsilon q_{jj}) V_{i,j}(x - \varepsilon r(j), t) + \sum_{k \in S \setminus \{j\}} \varepsilon q_{kj} V_{i,k}(x - \varepsilon r(k), t) + o(\varepsilon). \tag{6}$$

Simplifying Eq. (6), dividing by the time increment ε and letting $\varepsilon \downarrow 0$, $V_{i,j}(x, t)$ is seen to satisfy the partial differential equation

$$\frac{\partial V_{i,j}(x, t)}{\partial t} + \frac{\partial V_{i,j}(x, t)}{\partial x} r(j) = \sum_{k \in S} q_{kj} V_{i,k}(x, t), \quad j \in S \tag{7}$$

which may be written in matrix form as

$$\mathbf{V}_t(x, t) + \mathbf{V}_x(x, t) \mathbf{R}_D = \mathbf{V}(x, t) \mathbf{Q}. \qquad \square \tag{8}$$

Define $\mathbf{V}^*(x, s)$ as the Laplace transform (LT) of $\mathbf{V}(x, t)$ with respect to t and $\tilde{\mathbf{V}}^*(u, s)$, the Laplace–Stieltjes transform (LST) of $\mathbf{V}^*(x, s)$ with respect to x. Using these definitions, the double Laplace transform of the distribution function $G(x, t)$ is given by

$$\tilde{G}^*(u, s) = s^{-1} - \alpha \tilde{\mathbf{V}}^*(u, s) \mathbf{1}$$

$$= s^{-1} - \alpha (u \mathbf{R}_D + sI - \mathbf{Q})^{-1} \mathbf{1} \tag{9}$$

with $\text{Re}(s) > 0$ and $\text{Re}(u) > 0$.

The appealing aspect of this solution is that the double transform is available in closed form, lending itself to numerical inversion by the approaches of Choudhury et al. [10] and Moorthy [11]. However, the multidimensional inversion algorithms are difficult to apply and computationally expensive. Moreover, appropriate algorithm parameter selection is not immediately obvious. The main objective of this research is to circumvent the two-dimensional inversion process altogether by providing the transform of the lifetime distribution function in one dimension. Our main results for the distribution function and moments are provided in Section 3.

3. Main results

Owing to the fact that our result is in one transform variable, we adopt the following notation for the distribution function of the random lifetime:

$$G_x(t) := G(x, t) = P\{T_x \leqslant t\}. \tag{10}$$

Define the Laplace–Stieltjes transform of G_x by

$$\tilde{G}_x(s) = \int_{\mathbb{R}_+} e^{-st} G_x(dt).$$ (11)

Our main result is given by the following theorem.

Theorem 2. *Suppose the single-unit system is subject to a Markovian environment process Z with initial distribution, α, infinitesimal generator matrix, $\mathbf{Q}$, and wear rate matrix, $\mathbf{R}_D$. The Laplace–Stieltjes transform of the failure time distribution is*

$$\tilde{G}_x(s) = \alpha \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\mathbf{1}.$$ (12)

Theorem 2 will be proved in two ways. The first is a direct solution while the second involves analytical inversion of the two-dimensional result with respect to the spatial variable x.

Proof (Method 1). The result is first obtained by converting the partial differential equation:

$$\mathbf{V}_t(x,t) + \mathbf{V}_x(x,t)\mathbf{R}_D = \mathbf{V}(x,t)\mathbf{Q}$$

into an ordinary differential equation (ODE) and solving in the transform space. Taking the Laplace transform of both sides of the above equation with respect to t, noting that $\mathbf{V}(x,0) = I$ (the identity matrix), and rearranging terms yields the linear ODE with constant coefficients,

$$\frac{d\mathbf{V}^*(x,s)}{dx} + \mathbf{V}^*(x,s)(sI - \mathbf{Q})\mathbf{R}_D^{-1} = \mathbf{R}_D^{-1}.$$ (13)

We solve the ODE of (13) using an appropriate integrating factor which is given by

$$\mu(x) = \exp\left(\int (sI - \mathbf{Q})\mathbf{R}_D^{-1}\,dx\right)$$

$$= \exp\left((sI - \mathbf{Q})\mathbf{R}_D^{-1}x\right).$$ (14)

Multiplying both sides of (13) by $\mu(x)$ and integrating with respect to the spatial variable x shows that

$$\mathbf{V}^*(x,s)\exp((sI - \mathbf{Q})\mathbf{R}_D^{-1}x) = \mathbf{R}_D^{-1}(\mathbf{R}_D(sI - \mathbf{Q})^{-1})\exp((sI - \mathbf{Q})\mathbf{R}_D^{-1}x) + \psi,$$

where ψ is a (matrix) constant of integration. Due to the initial condition, $\mathbf{V}^*(0,s) = \mathbf{0}$, this matrix is given by

$$\psi = -(sI - \mathbf{Q})^{-1}.$$

By substituting the constant of integration and rearranging terms, we may write

$$\mathbf{V}^*(x,s) = (sI - \mathbf{Q})^{-1} - (sI - \mathbf{Q})^{-1}\exp((\mathbf{Q} - sI)\mathbf{R}_D^{-1}x).$$ (15)

By Eq. (4), the LST of G_x may be written as

$$\tilde{G}_x(s) = 1 - \alpha\mathbf{V}^*(x,s)s\mathbf{1}.$$ (16)

Substituting Eq. (15) into (16) shows that

$$\tilde{G}_x(s) = 1 - \alpha\{(sI - \mathbf{Q})^{-1} - \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x))(sI - \mathbf{Q})^{-1}\}s\mathbf{1}$$

$$= 1 - \alpha\{I - \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\}(I - \mathbf{Q}/s)^{-1}\mathbf{1}. \tag{17}$$

By applying the Neumann expansion (cf. [12]),

$$\left(I - \frac{\mathbf{Q}}{s}\right)^{-1} = \sum_{k=0}^{\infty} \frac{\mathbf{Q}^k}{s^k}, \tag{18}$$

and noting that $\mathbf{Q1} = \mathbf{0}$ and $\alpha\mathbf{1} = 1$, (17) reduces to

$$\tilde{G}_x(s) = \alpha \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\mathbf{1} \tag{19}$$

and the proof is complete. $\square$

Method 2: In the second method, the one-dimensional transform is obtained by analytical inversion of the two-dimensional result. Rewriting Eq. (9) as a LST with respect to both x and t, and then converting to a LT with respect to x, it is seen that

$$u\tilde{G}_u^*(s) = 1 - \alpha(u\mathbf{R}_D + sI - \mathbf{Q})^{-1}s\mathbf{1}. \tag{20}$$

Rearranging terms appropriately yields

$$u\tilde{G}_u^*(s) = 1 - \frac{\alpha}{u}\left(I - \frac{\mathbf{R}_D^{-1}(\mathbf{Q} - sI)}{u}\right)^{-1}\mathbf{R}_D^{-1}s\mathbf{1}.$$

Applying the Neumann expansion (18) to the bracketed inverse matrix above and dividing through by u gives

$$\tilde{G}_u^*(s) = \frac{1}{u} - \alpha\left(\frac{I}{u^2} + \sum_{k=1}^{\infty}(\mathbf{R}_D^{-1}(\mathbf{Q} - sI))^k\frac{1}{u^{k+2}}\right)\mathbf{R}_D^{-1}s\mathbf{1}. \tag{21}$$

Owing to the linearity of the inverse Laplace operator, the expression of (21) may be analytically inverted term by term. Performing this inversion with respect to the spatial variable x, we obtain

$$\tilde{G}_x(s) = 1 - \alpha\left(Ix + \sum_{k=1}^{\infty}(\mathbf{R}_D^{-1}(\mathbf{Q} - sI))^k\frac{x^{k+1}}{(k+1)!}\right)\mathbf{R}_D^{-1}s\mathbf{1} \tag{22}$$

which, after some manipulation, yields

$$\tilde{G}_x(s) = 1 - \alpha\left(x\mathbf{R}_D^{-1}(\mathbf{Q} - sI) + \sum_{k=1}^{\infty}(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)^{k+1}\frac{1}{(k+1)!}\right)(\mathbf{Q} - sI)^{-1}s\mathbf{1}$$

$$= 1 - \alpha\left(I - \sum_{k=0}^{\infty}(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)^k/k!\right)\left(I - \frac{\mathbf{Q}}{s}\right)^{-1}\mathbf{1}. \tag{23}$$

Applying Eq. (18) to the right-most inverse matrix of Eq. (23), using the definition of matrix exponentiation, and noting that $\mathbf{Q1} = \mathbf{0}$ and $\alpha\mathbf{1} = 1$, Eq. (23) reduces to

$$\tilde{G}_x(s) = 1 - \alpha \left(I - \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\right) \left(I - \frac{\mathbf{Q}}{s}\right)^{-1} \mathbf{1}$$

$$= 1 - \alpha \left(\mathbf{1} - \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\mathbf{1}\right)$$

$$= \alpha \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\mathbf{1}. \qquad \square$$

The main advantage of the one-dimensional result is that it requires numerical Laplace inversion with respect to the complex variable s only. It is worth mentioning that the resulting distribution function is of the matrix-exponential type as described by Bladt and Neuts [13].

For completeness, we derive the moments of the failure time of the unit using the one-dimensional result. Define the rth moment of the system lifetime as

$$m_r(x) = E[T_x^r]. \tag{24}$$

The Laplace–Stieltjes transform (LST) of this deterministic function is

$$\tilde{m}_r(u) = \int_{\mathbb{R}_+} e^{-ux} \, dm_r(x). \tag{25}$$

Our main result for the moments of the failure time is stated in Theorem 3. It should be noted that this result is also found in [1], but our derivation using Eq. (12) is simpler than the former.

Theorem 3. *The Laplace–Stieltjes transform of the rth moment of unit failure time is given by*

$$\tilde{m}_r(u) = r!\alpha(u\mathbf{R}_D - \mathbf{Q})^{-r}\mathbf{1}. \tag{26}$$

Proof. Define the matrix

$$\Phi_x(s) = \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x) \tag{27}$$

and its corresponding Laplace–Stieltjes transform with respect to x

$$\tilde{\Phi}_u(s) = \int_{\mathbb{R}_+} e^{-ux} \, d\Phi_x(s)$$

$$= (u\mathbf{R}_D - (\mathbf{Q} - sI))^{-1}(\mathbf{Q} - sI)$$

$$= (u\mathbf{R}_D + sI - \mathbf{Q})^{-1}u\mathbf{R}_D^{-1} - I. \tag{28}$$

The above result is obtained via standard matrix analysis operations as outlined in the appendix of Neuts [14]. In order to obtain the LST of the rth failure time moment, we evaluate the rth derivative of Eq. (28) with respect to s and then evaluate at $s = 0$. Doing so gives,

$$\left.\frac{\partial^r \Phi_u(s)}{\partial s^r}\right|_{s=0} = (-1)^r r!(u\mathbf{R}_D - \mathbf{Q})^{-r-1}u\mathbf{R}_D \tag{29}$$

so that the LST of the rth moment is finally obtained as

$$\tilde{m}_r(u) = (-1)^r r!\alpha \left.\frac{\partial^r \Phi_u(s)}{\partial s^r}\right|_{s=0} \mathbf{1} = r!\alpha(u\mathbf{R}_D - Q)^{-r}\mathbf{1}. \qquad \square \tag{30}$$

4. Numerical implementation and examples

In this section, we compare lifetime cumulative probability values obtained using the one- and two-dimensional transforms. First, we describe numerical implementation of the Laplace inversion algorithms.

To obtain the single dimension inverse Laplace transform values, we first pre-multiply the Laplace–Stieltjes transform result by s^{-1} in order to convert it to the Laplace transform (LT)

$$G_x^*(s) = s^{-1}\alpha \exp(\mathbf{R}_D^{-1}(\mathbf{Q} - sI)x)\mathbf{1}.$$

This conversion allows for direct implementation of the inversion algorithm of Abate and Whitt [7]. It remains to address the issue of matrix exponentiation in computing the Laplace transform. The preferred approach is the method of scaling and squaring with Padé approximations (cf. Moler and van Loan [15]). Vanden Bosch et al. [16] demonstrated that those algorithms using a Padé approximation are generally reliable. Let $\mathbf{C}$ denote any square matrix. The exponentiation of $\mathbf{C}$ may be written as

$$\exp(\mathbf{C}) = (\exp(\mathbf{C}/m))^m \tag{31}$$

where m is an integer power of two. The idea is to select m such that $\exp(\mathbf{C}/m)$ may be reliably computed. It is noted in [15] that m should be chosen as the smallest power of two satisfying $\|\mathbf{C}/m\| \leqslant 1$ where $\|\mathbf{C}\|$ denotes the norm of $\mathbf{C}$. In such case, the matrix $\exp(\mathbf{C}/m)$ may be reliably computed using Padé approximations. By squaring this matrix repeatedly, the desired result, $\exp((\mathbf{C}/m)m)$, is obtained. The scaling and squaring method is that which is implemented in the internal MATLAB® function **EXPM**.

In the case of two dimensions, the Laplace and inverse Laplace transforms for a function f are, respectively,

$$f^*(u,s) = \int_0^\infty \int_0^\infty e^{-ut_1-st_2} f(t_1,t_2)\, dt_1\, dt_2, \tag{32}$$

and

$$f(t_1,t_2) = \frac{1}{(2\pi i)^2} \int_{c_1-i\infty}^{c_1+i\infty} \int_{c_2-i\infty}^{c_2+i\infty} e^{ut_1+st_2} f^*(u,s)\, du\, ds, \tag{33}$$

where $f(t_1,t_2)$ is a real-valued function of t_1 and t_2, $f(t_1,t_2) = 0$ for t_1 or $t_2 < 0$, and $|f(t_1,t_2)| \leqslant Me^{\alpha_1 t_1 + \alpha_2 t_2}$, where α_1 and α_2 are real numbers and M is a positive constant. It is also assumed that $c_1 > \alpha_1$ and $c_2 > \alpha_2$. Moorthy [11] provides a technique to approximate $f(t_1,t_2)$ by extending the standard discrete Fourier cosine transform [17] to two dimensions as

$$f(t_1,t_2) \approx (2T^2)^{-1} \left\{ \frac{1}{2} f^*(c_1,c_2) + \sum_{m=1}^\infty \left[\text{Re}\left\{ f^*\left(c_1, c_2 + \frac{im\pi}{T}\right) \right\} \cos\left(\frac{m\pi t_2}{T}\right) \right.\right.$$

$$\left.\left. - \text{Im}\left\{ f^*\left(c_1, c_2 + \frac{im\pi}{T}\right) \right\} \sin\left(\frac{m\pi t_2}{T}\right) \right]\right.$$

$$+ \sum_{n=1}^{\infty} \left[\mathrm{Re} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 \right) \right\} \cos \left(\frac{n\pi t_1}{T} \right) \right.$$

$$\left. - \mathrm{Im} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 \right) \right\} \sin \left(\frac{n\pi t_1}{T} \right) \right]$$

$$+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left[\mathrm{Re} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 + \frac{im\pi}{T} \right) \right\} \cos \left(\frac{n\pi t_1}{T} + \frac{m\pi t_2}{T} \right) \right.$$

$$+ \mathrm{Re} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 - \frac{im\pi}{T} \right) \right\} \cos \left(\frac{n\pi t_1}{T} - \frac{m\pi t_2}{T} \right)$$

$$- \mathrm{Im} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 + \frac{im\pi}{T} \right) \right\} \sin \left(\frac{n\pi t_1}{T} + \frac{m\pi t_2}{T} \right)$$

$$\left. \left. - \mathrm{Im} \left\{ f^* \left(c_1 + \frac{in\pi}{T}, c_2 - \frac{im\pi}{T} \right) \right\} \sin \left(\frac{n\pi t_1}{T} - \frac{m\pi t_2}{T} \right) \right] \right\}. \tag{34}$$

The approximation requires appropriate selection of the parameters c_1, c_2, and T, as well as some appropriate criterion for terminating the infinite series. Suppose $\hat{f}_r$ is the r-term approximation. One approach is to compute the first N terms of the summation where N is the smallest integer for which $|\hat{f}_{N+1} - \hat{f}_{N+N/4}| < \delta$, and δ is some acceptable tolerance level. A different approach, based on Padé approximants, was used in this paper and will be described in what follows. The parameters c_1 and c_2 may be selected arbitrarily provided that $c_1 > \alpha_1$ and $c_2 > \alpha_2$. However, selection of the parameters α_1 and α_2 is not immediately obvious.

We compute the cumulative probability values $G_x(t)$ for various values of t by numerically inverting the two-dimensional result,

$$\tilde{G}_u^*(s) = s^{-1} - \alpha(u\mathbf{R}_D + sI - \mathbf{Q})^{-1}\mathbf{1}.$$

The two-dimensional Laplace–Stieltjes transform is converted to a two-dimensional Laplace transform by pre-multiplying both sides of the above equation by u^{-1}. A variant of Moorthy's algorithm [11] was implemented to evaluate distribution function values using Eq. (34). Let $T_{\max} := \max\{x, t\}$. The parameter T need only satisfy $T_{\max} < 2T$. However, Moorthy [11] notes that good results may be obtained when T is selected such that

$$0.5T_{\max} \leqslant T \leqslant 0.8T_{\max}.$$

One pragmatic choice for the parameter T is the mid-point of the interval, $0.65T_{\max}$. The technique also requires specification of the parameters c_1 and c_2. We note the equivalence of Moorthy's [11] parameters c_i, $i = 1, 2$, and the values a_i, $i = 1, 2$, of Choudhury et al. [10]. Consequently, we select $c_1 = A_1/2xl_1$ and $c_2 = A_2/2tl_2$ where the parameters, A_i and $l_i, i = 1, 2$, were chosen (as per the guidance of [10]) to be $A_1 = A_2 = 28.324$ and $l_1 = l_2 = 3$ to control the aliasing (discretization) and roundoff error. We implemented the ε-algorithm (which is described in detail in [18]) to accelerate convergence. The approximation for an infinite series is constructed by solving recursively

$$\varepsilon_{k+1}^n = \varepsilon_{k-1}^{n+1} + (\varepsilon_k^{n+1} - \varepsilon_k^n)^{-1}$$

upto the final term ε_{2m}^n. The initial conditions are given by $\varepsilon_{-1}^n = 0$ and ε_0^n is the nth partial sum of the infinite series. For these initial conditions, ε_{2m}^n is the $(m, m + n)$ Padé approximation to the infinite series. Our computational experience indicates that the values $m = 6$ and $n = 24$ are adequate for good results.

Distribution function values were compared to simulated values obtained via the Monte Carlo method. For each replication, we simulated 10,000 sample paths of the environment process $\{Z(t): t \geqslant 0\}$. In each case, we allowed the Markov process to evolve until the cumulative wear reached a critical threshold value ($x = 1.0$ in all experiments). The 10,000 observations of this first passage time were used to construct an empirical distribution function over a finite support. A total of 10 statistically independent replications were performed for each distribution, and 95% confidence intervals (CI) were constructed to bound each cumulative probability value. The Monte-Carlo simulation and one-dimensional inversion algorithms were coded in the MATLAB® computing environment while the two-dimensional algorithm was coded in the C programming language. All algorithms were executed on a single personal computer with a 2.0 GHz processor and 528 Mb of RAM.

4.1. Example 1: Fatigue crack dynamics

Assume $X(t)$ denotes the length of a crack in a metallic component (cf. [19]) at time t and assume that the (linear) rate at which the crack grows is subject to its random environment (applied stress, ambient conditions, and other factors). The process $\{Z(t): t \geqslant 0\}$ is a temporally homogeneous Markov chain that alternates between the two distinct states in its finite state space $S = \{1, 2\}$. Whenever the environment is in state $i \in S$, the crack grows at rate $r(i)$ units per unit time, $i \in S$.

The initial distribution of the environment process is arbitrarily chosen to be $\alpha = (1\ 0)$, i.e., the environment starts in state 1 with probability 1. The infinitesimal generator matrix is given by

$$\mathbf{Q} = \begin{bmatrix} -b & b \\ a & -a \end{bmatrix}$$

while the diagonal matrix of wear rates is

$$\mathbf{R}_D = \begin{bmatrix} r(1) & 0 \\ 0 & r(2) \end{bmatrix}.$$

The specific problem parameters chosen for this example are as follows. The threshold value is $x = 1.0$, and the values comprising the generator matrix are $a = b = 25/3$. The state-dependent crack growth rates are $r(1) = 1.0833$ and $r(2) = 0.250$. The cumulative probability values using one- and two-dimensional inversion were compared with Monte Carlo-simulated values at 94 distinct points. In Table 1, we report 30 of these values representing the entire range of the distribution.

The example demonstrates that both inversion approximations produce accurate results as compared to the simulated benchmark values (and corresponding confidence intervals). However, our computational experience indicates that the two-dimensional inversion algorithm exhibits instability when evaluating the inverse transform for smaller values of t (i.e., in the regime where $G_{1.0}(t)$ is close to zero). Such instability was not observed in the one-dimensional implementation. We also note a marked difference in the computation time of the two algorithms. In particular, the one-dimensional algorithm computed the 94 cumulative probability values in 0.73 s, while the two-dimensional

Table 1
$G_{1.0}(t)$ when Z is an on/off process

t	Simulated			Analytical	
	CI Lower	Mean	CI Upper	1-Dimensional	2-Dimensional
0.96	2.5208E-03	3.0000E-03	3.4792E-03	2.9661E-03	7.8100E-04
0.98	4.8639E-03	5.5600E-03	6.2561E-03	5.5638E-03	3.3400E-03
1.00	7.9532E-03	8.5000E-03	9.0468E-03	9.1149E-03	6.8870E-03
1.02	1.2645E-02	1.3160E-02	1.3675E-02	1.3812E-02	1.1781E-02
1.04	1.8665E-02	1.9580E-02	2.0495E-02	1.9890E-02	1.8308E-02
1.06	2.6660E-02	2.7340E-02	2.8020E-02	2.7533E-02	2.6333E-02
1.08	3.5420E-02	3.6600E-02	3.7780E-02	3.6839E-02	3.5643E-02
1.10	4.6350E-02	4.7650E-02	4.8950E-02	4.7857E-02	4.6700E-02
1.12	5.9091E-02	6.0960E-02	6.2829E-02	6.0632E-02	5.9064E-02
1.14	7.3114E-02	7.5390E-02	7.7666E-02	7.5224E-02	7.4658E-02
1.44	4.5978E-01	4.6273E-01	4.6568E-01	4.6630E-01	4.6620E-01
1.46	4.9076E-01	4.9366E-01	4.9656E-01	4.9699E-01	4.9681E-01
1.48	5.2087E-01	5.2448E-01	5.2809E-01	5.2726E-01	5.2715E-01
1.50	5.5167E-01	5.5528E-01	5.5889E-01	5.5696E-01	5.5684E-01
1.52	5.8169E-01	5.8490E-01	5.8811E-01	5.8598E-01	5.8575E-01
1.54	6.0991E-01	6.1317E-01	6.1643E-01	6.1419E-01	6.1417E-01
1.56	6.3615E-01	6.3895E-01	6.4175E-01	6.4149E-01	6.4148E-01
1.58	6.6198E-01	6.6514E-01	6.6830E-01	6.6779E-01	6.5239E-01
1.60	6.8673E-01	6.9019E-01	6.9365E-01	6.9303E-01	6.9301E-01
1.62	7.1130E-01	7.1502E-01	7.1874E-01	7.1713E-01	7.1696E-01
2.62	9.9970E-01	9.9981E-01	9.9992E-01	9.9981E-01	9.9980E-01
2.64	9.9976E-01	9.9986E-01	9.9996E-01	9.9985E-01	9.9986E-01
2.66	9.9980E-01	9.9988E-01	9.9996E-01	9.9988E-01	9.9990E-01
2.68	9.9980E-01	9.9989E-01	9.9998E-01	9.9990E-01	9.9992E-01
2.70	9.9988E-01	9.9993E-01	9.9998E-01	9.9992E-01	9.9993E-01
2.72	9.9991E-01	9.9995E-01	9.9999E-01	9.9994E-01	9.9994E-01
2.74	9.9994E-01	9.9997E-01	1.0000E+00	9.9995E-01	1.0000E+00
2.76	9.9995E-01	9.9998E-01	1.0000E+00	9.9996E-01	1.0000E+00
2.78	9.9997E-01	9.9999E-01	1.0000E+00	9.9997E-01	1.0000E+00
2.80	1.0000E+00	1.0000E+00	1.0000E+00	9.9997E-01	1.0000E+00

algorithm computed the 94 values in 9.93 s. Although the computation time for the two-dimensional algorithm is small (owing to the fact that $K=2$), this represents a 1260% increase in the computation time as compared to the one-dimensional algorithm.

4.2. Example 2: Stochastic tool wear model

Consider a machine cutting tool (e.g., an outside diameter grinding wheel) in which the work rate of the machine varies between five distinct settings. These settings may correspond to various cutting speeds, each of which corresponds to a particular workpiece composition. In order for the workpiece to have the appropriate surface finish, it is imperative that the cutting tool be in good form. We assume that a cutting tool is scrapped and replaced once its level of accumulated wear reaches a threshold value of x wear units.

Table 2
$G_{1.0}(t)$ when Z is a 5-state Markov process

t	Simulated			Analytical	
	CI Lower	Mean	CI Upper	1-Dimensional	2-Dimensional
1.00	3.4108E-03	3.7900E-03	4.1692E-03	3.6419E-03	1.5100E-03
1.04	6.2081E-03	6.8400E-03	7.4719E-03	7.0582E-03	5.3790E-03
1.08	1.1307E-02	1.2120E-02	1.2933E-02	1.2641E-02	1.1354E-02
1.12	1.9816E-02	2.0480E-02	2.1144E-02	2.1210E-02	2.0241E-02
1.16	3.1571E-02	3.2710E-02	3.3849E-02	3.3670E-02	3.2908E-02
1.20	4.7807E-02	4.9620E-02	5.1433E-02	5.0942E-02	5.0371E-02
1.24	6.9775E-02	7.2320E-02	7.4865E-02	7.3872E-02	7.3432E-02
1.28	9.8751E-02	1.0222E-01	1.0569E-01	1.0314E-01	1.0282E-01
1.32	1.3540E-01	1.3803E-01	1.4066E-01	1.3916E-01	1.3893E-01
1.36	1.7808E-01	1.8144E-01	1.8480E-01	1.8200E-01	1.8184E-01
1.52	4.0730E-01	4.1018E-01	4.1306E-01	4.0965E-01	4.0964E-01
1.56	4.7328E-01	4.7565E-01	4.7802E-01	4.7462E-01	4.7463E-01
1.60	5.3657E-01	5.3926E-01	5.4195E-01	5.3965E-01	5.3969E-01
1.64	5.9913E-01	6.0198E-01	6.0483E-01	6.0312E-01	6.0318E-01
1.68	6.6254E-01	6.6477E-01	6.6700E-01	6.6353E-01	6.6358E-01
1.72	7.1852E-01	7.2113E-01	7.2374E-01	7.1963E-01	7.1969E-01
1.76	7.6924E-01	7.7167E-01	7.7410E-01	7.7048E-01	7.7056E-01
1.80	8.1408E-01	8.1616E-01	8.1824E-01	8.1550E-01	8.1559E-01
1.84	8.5329E-01	8.5501E-01	8.5673E-01	8.5443E-01	8.5451E-01
1.88	8.8712E-01	8.8843E-01	8.8974E-01	8.8729E-01	8.8738E-01
2.48	9.9970E-01	9.9982E-01	9.9994E-01	9.9971E-01	9.9981E-01
2.52	9.9982E-01	9.9989E-01	9.9996E-01	9.9980E-01	9.9989E-01
2.56	9.9986E-01	9.9994E-01	1.0000E+00	9.9985E-01	9.9994E-01
2.60	9.9988E-01	9.9995E-01	1.0000E+00	9.9987E-01	9.9997E-01
2.64	9.9994E-01	9.9997E-01	1.0000E+00	9.9989E-01	9.9999E-01
2.68	9.9997E-01	9.9999E-01	1.0000E+00	9.9989E-01	9.9999E-01
2.72	9.9997E-01	9.9999E-01	1.0000E+00	9.9990E-01	1.0000E+00
2.76	9.9997E-01	9.9999E-01	1.0000E+00	9.9990E-01	1.0000E+00
2.80	9.9997E-01	9.9999E-01	1.0000E+00	9.9990E-01	1.0000E+00
2.84	1.0000E+00	1.0000E+00	1.0000E+00	9.9990E-01	1.0000E+00

Let $X(t)$ denote the amount of wear incurred by the cutting tool by time t and suppose cutting speeds are selected from the finite set $\{r(1), r(2), r(3), r(4), r(5)\}$. Let $Z(t) \in \{1, 2, 3, 4, 5\}$ denote the type of workpiece on the machine at time t such that if $Z(t) = k$, then the tool wears at a rate $r(k)$. Thus, we assume the environment process, $\{Z(t) : t \geqslant 0\}$, is a five-state CTMC. The specific wear rates in this problem were arbitrarily selected as

$$r(k) = 1.25/k, \quad k \in S,$$

so that the (i, j) entry of the matrix $\mathbf{R}_D$, denoted by $\mathbf{R}_D(i, j)$, is given by

$$\mathbf{R}_D(i, j) = \begin{cases} 1.25/i & j = i, \\ 0 & j \neq i. \end{cases}$$

With probability 1, the system begins in state 1 at time 0 so that the initial environment distribution is $\alpha = (1\ 0\ 0\ 0\ 0)$. The off-diagonal entries of the 5×5 infinitesimal generator matrix were drawn from a uniform distribution on the interval (3.333, 6.667). We set the failure threshold value to unity ($x = 1.0$). The cumulative probability value, $G_{1.0}(t)$ was computed for 54 distinct values of t, and Table 2 provides the results for 30 representative cases.

The second example also demonstrates that both the one- and two-dimensional techniques provide good results as compared to the simulated values. However, we again note a marked difference in the computation time between the two algorithms. The one-dimensional algorithm computed a total of 54 cumulative probability values in 0.55 s while the two-dimensional algorithm computed the same 54 values in 222.98 s (approximately 3.72 min). We surmise that, as the number of states in governing environment process increases, the difference in computation time will be even more marked.

5. Conclusions

Owing to the limitations of unit lifetime prediction via accelerated life testing, there exists a growing need for failure models that capture the effects of the physical environment on a unit's cumulative degradation (wear) and lifetime. Such models include a mathematical characterization of the environment as a continuous-time stochastic process on a continuous or discrete state space. This approach is naturally appealing since it provides a means by which the stochastic evolution of the environment, and its ultimate effect on the operating unit, may be modelled. However, the numerical implementation of such procedures lags far behind their analytical development. This paper has proposed a simpler means by which numerical results may be obtained for a specific class of failure models, namely continuous wear processes.

In particular, we have analytically proven simplified results for the evaluation of the lifetime cumulative distribution values for a single-unit system that continuously accumulates wear due to the influence of its time-varying operating environment. The need to incorporate the impact of the ambient environment has become prevalent in the reliability theory literature. This paper has considered the case in which the environment process is modelled as a finite state Markov process. We demonstrated the means by which to compute cumulative probability values by inverting a one-dimensional transform as opposed to the two-dimensional result of [1]. The one-dimensional result has the obvious advantage of requiring numerical inversion of a Laplace transform with respect to only a single complex variable, a process for which numerical algorithms abound. Though the simplified result requires matrix exponentiation, the method of scaling and squaring (using Padé approximations) appears to provide reliable results in a far more expedient manner. Moreover, the need to specify algorithm parameters is virtually eliminated. In the future, it will be instructive to consider numerical algorithms for models that additionally consider the effect of shocks occurring at random intervals such as the one reviewed by Singpurwalla [4]. To the authors' knowledge, no suitable numerical techniques exist in the current literature for such models.

Acknowledgements

References

[1] Kharoufeh JP. Explicit results for wear processes in a Markovian environment. Operations Research Letters 2003;31:237–44.

[2] Gillen KT, Celina M. The wear-out approach for predicting the remaining lifetime of materials. Polymer Degradation and Stability 2001;71:15–30.

[3] Meeker WQ, Escobar LA, Chan V. Using accelerated tests to predict service life in highly variable environments. In: Bauer DR, Martin JW, editors. Service life prediction: methodologies and metrologies. Washington, DC: American Chemical Society; 2002.

[4] Singpurwalla ND. Survival in dynamic environments. Statistical Science 1995;10:86–103.

[5] Bobbio A, Kulkarni VG, Telek M. Partial loss in reward models. In: Proceedings of the Second International Conference on Mathematical Methods in Reliability, Bordeaux, France: July 2000. p. 207–10.

[6] Bobbio A, Telek M. The task completion time in degradable sytems. In: Haverkort GRBR, Marie R, Trivedi KS, editors. Performability modelling: techniques and tools. New York: Wiley; 2001. p. 139–61.

[7] Abate J, Whitt W. Numerical inversion of Laplace transforms of probability distributions. ORSA Journal on Computing 1995;7:36–43.

[8] Meeker WQ, Escobar LA. Statistical methods for reliability data. New York: Wiley Inc.; 1998.

[9] Çinlar E. Shock and wear models and Markov additive processes. In: Shimi IN, Tsokos CP, editors. The theory and applications of reliability. New York: Academic Press; 1977. p. 193–214.

[10] Choudhury GL, Lucantoni D, Whitt W. Multidimensional transform inversion with applications to the transient M/G/1 queue. The Annals of Applied Probability 1994;4:719–40.

[11] Moorthy MV. Numerical inversion of two-dimensional Laplace transforms-Fourier series representation. Applied Numerical Mathematics 1995;17:119–27.

[12] Noble B, Daniel JW. Applied linear algebra, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall; 1988.

[13] Bladt M, Neuts MF. Matrix-exponential distributions: calculus and interpretations via flows. Stochastic Models 2003;19:113–24.

[14] Neuts MF. Algorithmic probability: a collection of problems. London: Chapman & Hall; 1995.

[15] Moler C, van Loan C. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years laters. SIAM Review 2003;45:3–49.

[16] Vanden Bosch PM, Dietz DC, Pohl EA. Choosing the best approach to matrix exponentiation. Computers and Operations Research 1999;29:871–82.

[17] Dubner H, Abate J. Numerical inversion of Laplace transforms by relating them to the finite Fourier cosine transform. Journal of the Association for Computing Machinery 1968;15:115–23.

[18] Macdonald JR. Accelerated convergence, divergence, iteration, extrapolation, and curve fitting. Journal of Applied Physics 1964;35:3034–41.

[19] Ray A, Tangirala S. A nonlinear stochastic model of fatigue crack dynamics. Probabilistic Engineering Mechanics 1997;12:33–40.

Jeffrey P. Kharoufeh is an assistant professor of Operations Research at the Air Force Institute of Technology. He holds a Ph.D. in Industrial Engineering and Operations Research from Pennsylvania State University. His research interests include the development and analysis of stochastic models in operations research. He is a professional member of INFORMS.

Jeffrey A. Sipe is Chief of the Tanker Airlift Control Center (TACC) Business Center Analysis Division at Scott Air Force Base, Illinois. He holds a B.S. in Mathematics from Angelo State University and a M.S. in Applied Mathematics from the Air Force Institute of Technology.