2014

Reliability growth by failure mode removal

Gaver, Donald P.; Jacobs, Patricia A.

Reliability Engineering and System Safety, Volume 130, pp. 27-32, 2014.
http://hdl.handle.net/10945/43156

Reliability Engineering and System Safety 130 (2014) 27–32

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Reliability growth by failure mode removal

Donald P. Gaver*, Patricia A. Jacobs

Operations Research Department, Naval Postgraduate School, Monterey, CA 93943, USA

ARTICLE INFO

Article history:
Received 1 April 2014
Accepted 23 April 2014
Available online 9 May 2014

keywords:
Reliability growth
Infinite server queue
Poisson thinning

ABSTRACT

Modern systems, civilian (e.g. automotive), and military (manned and unmanned aircraft, surface vehicles, submerged vessels), suffer initial design faults or failure modes (FMs), including software bugs, which detrimentally affect the system's reliability and availability. FMs must be removed or mitigated in impact during initial testing, including accelerated testing, in order for the system to meet its reliability requirements and operate satisfactorily in the field. This paper concerns models for reliability growth in which the behaviors of FMs are assumed independent, but of different types. Test effort is guided by prior information, expressed probabilistically, on the random number and tenacities of such FMs that are of various origins in the designs. Estimation of the numbers of FMs that will ultimately activate while in the field is considered here.

Published by Elsevier Ltd.

1. Introduction

Failure mode removal from any system, both hardware and software, is a dynamic uncertain process; see [1–19] for various discussions of problem approaches. In [5,6,8,9,16] an unknown number of FMs are supposed present in the system initially, and the subsequent random times until FM activations are independent and identically distributed; in [8,9] the unknown number of failure modes is assumed deterministic and asymptotic arguments are used for its estimation; in [5,6] the unknown number of FMs has a Poisson distribution, and estimation of the Poisson mean is discussed; [13] includes a Bayesian treatment of the general model; [16] considers a dynamic statistical model for mean number of FMs remaining. The models in [3,4] are widely used and FM activations occur according to a nonhomogeneous Poisson process; a Bayesian treatment of this model appears in Refs. [14,15]. Additional nonhomogenous Poisson process models have been suggested, including that of [17]. In [12] time series are used to summarize software failure data; parameter estimation uses a genetic algorithm; estimation is illustrated with small data sets. In [18,19] a neural network approach is discussed. In [7,9] reliability growth models are suggested for the management of system testing. In [10] a series system of subsystems with resulting FM masking is considered. The goal of failure avoidance, or *system reliability growth* remains a concern to military and civilian system designers, testers and operators; see Toyota automobile accelerator pedal occasional mishaps ([20] and also [21]).

This paper presents an approach to modeling and statistical analysis based on familiar applied stochastic process theory. The model notion is that of identifying failure mode creation and removal with an "infinite server queue", a generalized so-called $M/G/\infty$ system; here M refers to a general Markovian/memoryless "arrival process" of failure modes into a system; G represents the general distribution function of the "residence time", or "service time" in queueing language, of any FM in the system: either until discovered and rectified, or, if not discovered during test, activates in use, thus interrupting field operation usage and possibly causing fatality. The individual FM residence times are here assumed independent and identically distributed; however see [22] for plausible variation. Finally, "∞" refers to the practically infinite, or unlimited number of locations/sites in the system where FMs can reside; cf. [23]; these are "servers" in queueing context as in Ref. [24]. Note that here the items present are all eligible for service/removal when recognized. Later work will recognize congestive servers, and evaluate priority removal. The M-arrival process can include initial numbers of FMs of different types having independent Poisson distributions, with additional FMs that are inadvertently inserted during development according to nonhomogeneous Poisson processes (NHPPs). The assumptions that the unobservable initial number of FMs in a system have Poisson distributions and that the unobservable insertion of additional FMs in a system occur according to NHPPs is convenient and has been made before; see [24–26]. Since the presence of a FM in a system is a "rare event" the assumptions are reasonable, *prima facie*.

The $M/G/\infty$ queue can represent many features encountered in reliability growth data, as has been pointed out by [24–26]. In [24] the NHPP is exploited to describe single-type fault (FM) occurrence; our current results represent realistic recurrence of non-removed

* Corresponding author. Tel.: +1 8316562605; fax: +1 8316562595.
E-mail address: dgaver@nps.edu (D.P. Gaver).

http://dx.doi.org/10.1016/j.ress.2014.04.012
0951-8320/Published by Elsevier Ltd.

FMs. In [25] the discovery of FMs is modeled as occurring according to a NHPP and the times until removal of FMs as independently distributed random variables. The present paper extends the NHPP model of [26] to include fault (FM) type differences plus the realistic probability that actuated FMs are not removed until a removal success. A FM's residence time can include multiple occurrences of a FM due to unsuccessful attempts to remove the FM and may be summarized by a renewal process or even an epidemic process; see [27]. Our present work covers the likely random variation between activation-recurrence rates of different FMs; see (4)–(7) in this paper. These are seen to be natural and realistic extensions of [24–26]. The actual number of FMs is unknown and random, and is realistically controlled by the balance of the arrival and the service or fault-removal process, and so may actually be small, but can grow indefinitely, as in [28]. Note that the mathematical model can be time – or system age – dependent, so different, even new, FM types can be represented during a system design's lifetime.

The model proposed here, and the statistical methods based on it, does not explicitly represent the phenomenon of *mode masking*, meaning that early FM discovery, e.g. of a defect in a vehicle ignition system, or missile launch stage, does not here effect appearance of later FMs that may occur had the early FM not occurred, e.g. in vehicle steering, or missile guidance and detonation. We may view the present model as of one stage, s, of an S-stage series or sequentially operating system. The present model omits desirable mention of Prognostic Health Management (PHM), meaning anticipatory replacement of failure-imminent components or subsystems.

FMs remaining after testing detrimentally affect system field reliability and availability. The purpose of the model and its generalizations is to infer the properties of *FMs remaining* following system testing. The formal model is presented and discussed in Section 2 with examples of behavior that can be represented with the multi-type M/G/∞ queue. Of particular interest are *statistical models* that represent the inherent variation between FMs. A discussion of statistical inference is in Section 3. Section 4 illustrates issues of statistical inference using sample software testing data. The statistical analysis suggests that several different models summarize the data well and that more and extended software testing would be prudent; models with more parameters appear unneeded to summarize these data. The paper ends with conclusions in Section 5.

2. Model for failure mode increase and decrease

Let

- $\lambda_k(t)$ = Random arrival rate of failure modes (FMs) of type k into the system at time t, where $\lambda_k(t)$is the rate of a time-dependent Poisson process. These lie dormant until "flare up" or activation. In other, and subsequent work we can allow FMs to issue warnings or diagnostic symptoms that, if detected, can forestall serious failure. Such FMs can result from human intervention to repair others. $F_k(\tau)=P\{\boldsymbol{T}_k\le\tau\}$ = Probability distribution of $\boldsymbol{T}_k$, the random time until the activation of a single type k FM.

Initially it is assumed that such times can repeat themselves, to represent activations that occurred repeatedly but have not been successfully removed; the times between activations being independent and identically distributed. A special case of inter-activation time distribution is the exponential distribution, $\exp(\mu_k)$,

$$F_k(\tau,\mu_k)=1-\exp\{-\mu_k\tau\} \quad (1)$$

Note that the inter-activation time distribution function $F_k(\tau)$, and the arrival rate, $\lambda_k(t)$, can both be affected by environmental influences, including human maintenance or operator, by incorporation of suitable parameter sets and variables. Such important effects are not treated here; they are left for later work.

Next,

- ρ_k = Probability that a FM of type k is removed on any activation. This parameter is initially assumed constant no matter how many responses to activations have occurred; it is a candidate for modification to represent learning.
- $A_k(\tau)$ = Event that a failure mode of type k, is active, i.e. a latent failure, in the system at time τ after it "arrives" in a design or a copy thereof.

Then

$$P\{A_k(\tau)|F_k(\bullet),\rho_k\}=\sum_{n=0}^{\infty}\left[F_k^{*n}(\tau)-F_k^{*(n+1)}(\tau)\right](1-\rho_k)^n \quad (2)$$

where F^{*n} is the n-fold convolution of the distribution F with itself. This simply says that a fault that arrives in the system at $t=0$ has (independently) activated any number, n, times but has not been removed by time τ. In the special $\exp(\mu_k)$ case

$$P\left\{A_k(\tau)|\exp(\mu_k),\rho_k\right\}=\sum_{n=0}^{\infty}e^{-\mu_k\tau}\frac{(\mu_k\tau)^n}{n!}(1-\rho_k)^n=e^{-\mu_k\rho_k\tau} \quad (3)$$

Following [7], assume μ_k is a realization of independent identically distributed random variables with distribution function $H_k(\mu_k)=P\{\boldsymbol{\mu}_k\le\mu_k\}$; that is, while each FM has independent exponential times between activations, different FMs have different mean inter-activation times drawn from a mixing distribution, $H_k(\bullet)$. From (3), this then implies that

$$E\left[e^{-\boldsymbol{\mu}_k\rho_k\tau}\right]=\int_0^{\infty}e^{-\mu_k\rho_k\tau}dH_k(\mu_k)=\tilde{H}_k(\rho_k\tau), \quad (4)$$

where $\tilde{H}_k(s)$ is the Laplace-Stieltjes transform of the distribution function $H_k(\mu_k)$ evaluated at $s=\rho_k\tau$.

We propose two different forms for the mixing distribution, H_k

(A) Classical gamma.
(B) Positive stable law ([27]).

First, a simple explicit result for the transform of the *gamma distribution* assumed by [29] with scale β_k and shape α_k is

$$\tilde{G}_k(\rho_k\tau)=\left(1+\frac{\rho_k\tau}{\beta_k}\right)^{-\alpha_k} \quad (5)$$

Next, the positive *stable law* has relevant transform, for shape parameter $0<\alpha_k<1$,

$$\tilde{S}_k(\rho_k\tau)=\exp\left\{-\left(\frac{\rho_k\tau}{\beta_k}\right)^{\alpha_k}\right\} \quad (6)$$

notationally (only) matching the stable scale and shape parameters to those of the gamma for $0<\alpha_k<1$.

It is evident that (5) and (6) represent the distribution of residence time in the system of corresponding type k fault. Let H_k here represent either G_k or S_k: the probability the residence time is less than or equal to τ is

$$P(A_k(\tau)^c)=1-\tilde{H}_k(\rho_k\tau) \quad (7)$$

Put

- $\boldsymbol{N}_k(t)$=Random number of FMs, that *activate and are removed* from the system during exposure time t, i.e. within $(0,t]$. Note that this includes those initially within the system plus those that are introduced thereafter.
- $\boldsymbol{R}_k(t)$=Random number of *native* FMs that either have not yet activated or have activated but are not (yet) removed during exposure time t.

We assume that the FM removal times are independent. The theory of M/G/∞ service systems can now be invoked to show that $\boldsymbol{N}_k(t)$ and $\boldsymbol{R}_k(t)$ are statistically independent and each have independent components.

3. Preliminary results

$\boldsymbol{N}_k(t)$ is the sum of those $k-$ FMs that are initially in the design and removed before exposure time t, and those that migrate in thereafter, and are removed before t. Given $\boldsymbol{R}_k(0)$, the natives that are removed by t, $\boldsymbol{N}_k(t;o)$, has a Binomial distribution with $\boldsymbol{R}_k(0)$ trials and probability of removal $[1-\tilde{H}_k(\rho_k t)]$. The number of immigrant modes, $\boldsymbol{N}_k(t;i)$, that arrive and are removed during $(0,t]$ is Poisson distributed with mean

$$H_k^{\#}(t)=\int_0^t \left[1-\tilde{H}_k(\rho_k(t-x))\right]\lambda_k(x)dx \tag{8}$$

If $\boldsymbol{R}_k(0)$ is Poisson distributed with mean M_k, viewed either as a Bayes prior, or else as a mixing distribution, *then* the number removed has mean $M_k[1-\tilde{H}_k(\rho_k t)]+H_k^{\#}(t)$.

The number of FMs remaining in the system at t

• $\boldsymbol{R}_k(t)$, the sum of those that remain in the system which includes the survivors of those initially there, $\boldsymbol{R}_k(t;o)$, plus those immigrating and surviving until t, $\boldsymbol{R}_k(t;i)$. If $\boldsymbol{N}_k(0)$ is taken as Poisson (M_k) then the sum of $\boldsymbol{R}_k(t;o)+\boldsymbol{R}_k(t;i)$ is Poisson, and

$$E[\boldsymbol{R}_k(t;o)+\boldsymbol{R}_k(t;i)]=M_k\widetilde{H}_k(\rho_k t)+\overline{H}_k^{\#}(t) \tag{9}$$

where

$$\overline{H}_k^{\#}(t)=\int_0^t \tilde{H}_k(\rho_k x)\lambda_k(x)dx \tag{10}$$

Justification is by Poisson thinning; see [30].

4. Statistical inference

The purpose of the modeling described and its many imaginable generalizations is to infer properties of FMs remaining following Test (and Evaluation—in fact, this term refers to statistical/operational analysis). The practical issue is control of "field" FM incidence as it impacts system life-cycle cost and capability, including availability, cf. [11]. The jargon used here is military in origin, but has civilian meaning also; see [20].

Statistics are based upon data, which must come from realistically designed and executed Tests, both at basic engineering levels (contractor testing including accelerated testing and developmental testing, or DT), but importantly under field or operational environments (operational testing, or OT). Early test results should be used to predict the occurrence of FMs in future testing.

• Test observables

Envision the various observable test outcomes from a subsystem, or system test. We will encompass several of these in a formal process of statistical inference.

Test data on continuously-failure-prone subsystems often give test times of the ith occurrence of a FM of type $k: 0\le t_{1k}<t_{2k}<\cdots<t_{ik}<\cdots$; the times between occurrences will here be taken to be instances of iid absolutely continuous random variables.

At such an event we assume that repair or rectification occurs that is successful in removing the Type k FM with probability ρ_k, and leaves the FM "as bad as old" with probability $\overline{\rho}_k=1-\rho_k$.

4.1. *Example of estimation of the mixing distribution: Each FM activates at most once and all FMs are present initially*

Consider original FMs only, with exponentially distributed type k conditional activation times, and known removal probability $\rho_k=1$. Then for $r_k(0)$ FMs at time 0, the ith, $i=1,2..,r_k(0)$, observation time of activation of the ith type k FM, $t_i(k)$, suggests writing the conditional likelihood

$$L_i(\mu_k(i))=e^{-\mu_k(i)t_k(i)}\mu_k(i) \tag{11}$$

And the (conditional-on-i) estimate is clearly $\hat{\mu}_k(i)=t_k^{-1}(i)$. The latter will be treated as a random sample from a mixing distribution with finite mean and variance; (they can also be usefully studied by exploratory data-analytical methods). Either method of moments, maximum likelihood, invocation of Jeffries non-informative priors, or previous system adjusted priors may be used to fit the mixing distribution.

4.2. *Example of likelihood: Exponentially distributed inter-event (FM realization) times; all FMs are present initially*

If n_k activation times of a single type k FM are observed, then it is clear that there have been at least n_k-1 unsuccessful removal attempts by the end of test time, t. We can express the likelihood of the rate μ_k as, (using NR for *not removed* and R for *removed*)

$$L(\mu_k;t_{\bullet k},n_k;NR) \\ =(1-\rho_k)^{n_k}(\mu_k)^{n_k}\exp\{-\mu_k t\} \tag{12a}$$

$$L(\mu_k;t_{\bullet k},n_k;R) \\ =(1-\rho_k)^{n_k-1}\rho_k\,(\mu_k)^{n_k}\exp\{-\mu_k t\} \tag{12b}$$

Next introduce mixing distribution $H_k(\mu_k,\underline{\theta})$; this mixing/parameter randomization occurs initially before the present phase of testing, and represents inherent variation of activation among Type k FMs. Thus post-test likelihoods should remove the condition on μ_k and on ρ_k. Removing the condition on μ_k,

$$\tilde{L}(\theta_k;n_k)\propto\int_0^\infty \mu_k^{n_k}\exp\{-\mu_k t\}H_k(d\mu_k;\theta_k) \tag{13}$$

where a conjugate "mixer", $H_k(\bullet;\underline{\theta}_k)$, here is gamma (β_k,α_k), or positive stable.

Likewise one can mix on ρ_k; this step is omitted.

Note that the above represents a single (one) multi-time observation of activations of a single FM, of which there can have been several. Of course further uncertainty exists with respect to the removal probability, ρ_k, which can be estimated by Bayesian methods (depending on a satisfactory prior's availability). A deeper uncertainty is whether the final observation is actually the last, i.e. whether the FM has been removed or not. It will be convenient, but optimistic, to assume that a successful removal is *known to be successful* at the time. Otherwise, a non-activation run/sequence of tests can provide empirical evidence. *In what follows we assume the outcome of an attempted removal of a FM cause is known.*

4.3. *Sequential fault removal*

Adopt this specific model: decision-makers view a system S through a period of developmental or preliminary engineering test of operating time duration D, followed by a period of duration T, where the latter is the duration of more demanding operational test. We assume all FMs are present initially; the times between activations of a single FM are independent identically distributed having distribution (7). Focus on inferring status of an FM type k at the end of testing occurring at time $D+T$ using information from test results during D.

Refer now to (4)–(6), and treat $G_k(\bullet)$ and $S_k(\bullet)$ as alternative mixing distributions. Suppose $\pi_k(z;\underline{\theta})$ represents a mixing/prior density function, either $G_k(\bullet)$ or $S_k(\bullet)$. Note that during D, the rate of failure is influenced by controllable test and uncontrolled environmental conditions represented by vector explanatory variable $\underline{x}$, to make $\mu_k^D=\mu_k\theta_k(\underline{x})$; likewise for survivors of D into T, $\mu_k^T=\mu_k\theta_k(\underline{y})$, where vector $\underline{y}$ represents conditions during T. The controllable explanatory conditions, "factors" and levels, are essential to specify the tests; actual factor level values achieved during the tests should be measured. The controllable factors' levels chosen should be guided by modeling and simulation and prior tests to have a good predictive chance of success, but adjusted to challenge the system further after initial success is demonstrated.

A single FM of type k may experience these outcomes:

(a) Not removed (survives) T, given activates (at least once during D) and survives D;
(b) Removed during T, given activates and survives D;
(c) Survives T, given does not activate during D;
(d) Removed during T, given does not activate during D;
(e) Activates and removed during D.

Apply Bayes' formula to remove the condition on $\boldsymbol{\mu}_k$, given experience during D to obtain these explicit forms:

$$P(a)=\int_0^\infty e^{-z\rho_k\theta_k(\underline{y})T}\left[e^{-z\rho_k\theta_k(\underline{x})D}-e^{-z\theta_k(\underline{x})D}\right]\frac{\pi_k(z)}{\pi_k^{\#}(\rho_k\theta_k(\underline{x}),D)}dz \quad (14)$$

where

$$\pi_k^{\#}(\rho_k\theta_k(\underline{x}),D)=\int_0^\infty\left[e^{-z\rho_k\theta_k(\underline{x})D}-e^{-z\theta_k(\underline{x})D}\right]\pi_k(z)dz \quad (15)$$

$$P(b)=1-P(a) \quad (16)$$

$$P(c)=\int_0^\infty e^{-z\rho_k\theta_k(\underline{y})T}\frac{e^{-z\theta_k(\underline{x})D}\pi_k(z)}{\pi_k^{*}(\theta_k(\underline{x}),D)}dz \quad (17)$$

where

$$\pi_k^{*}(\theta_k(\underline{x}),D)=\int_0^\infty e^{-z\theta_k(\underline{x})D}\pi_k(z)dz \quad (18)$$

$$P(d)=1-P(c) \quad (19)$$

The expressions (14)–(19) can be evaluated for both G_k and/or S_k. Again, every event labeled Survive means that such is regarded as potentially active in the following testing stage, and those labeled removed may have failed, perhaps several times, but have been removed before the end of the current testing stage. If the outcome (a) is observed, this can be taken to mean that the FM in question survives to experience in-the-Field activation.

To estimate the mean number of FMs that have not yet activated, assume that the number of FMs of type k observed during D is the thinned result of a Poisson random variable having mean M_k. If the number of type k FMs observed to have activated at least once during D is $m_k(D)$, setting $m_k(D)$ equal to the expected number of FMs activated gives (method of moments)

$$m_k(D)=M_k[1-\pi_k^{*}(\theta_k(\underline{x}),D)] \quad (20)$$

so

$$\hat{M}_k(D)=\frac{m_k(D)}{1-\pi_k^{*}(\theta_k(\underline{x}),D)} \quad (21)$$

The estimated (mean) number of $k-$FMs that have not activated during D

$$\hat{M}_k(D)-m_k(D)=\hat{R}_k^0(0)=\frac{m_k(D)\pi_k^{*}(\theta_k(\underline{x}))}{1-\pi_k^{*}(\theta_k(\underline{x}),D)} \quad (22)$$

Of course it will be of greater interest to focus on separately estimating, or dealing with FMs in category (a) than for those not yet observed.

5. Estimation: One type of FM; all FMs present initially; each FM activates at most once

In this section we consider a model with one type of FM. The number of FMs at time 0 is modeled as Poisson distributed with mean M. No additional FMs are introduced. Each FM activates at most once. The activation times are independent identically distributed with distribution function F having survivor function $\overline{F}=1-F$, and density function f. Let $\boldsymbol{U}_i$ be the random time of activation of the ith FM; $\boldsymbol{N}(t)$ be the number of FMs that activate during $(0,t]$; and D be the total developmental test time. Standard calculations yield

$$P\{\boldsymbol{N}(D)=n,\boldsymbol{U}_1\in du_1,\ldots,\boldsymbol{U}_n\in du_n\}=e^{-MF(D)}\frac{[MF(D)]^n}{n!}\left[\prod_{j=1}^{n}\frac{f(u_j)}{F(D)}du_j\right] \quad (23)$$

The log-likelihood function for M and the parameters of F is

$$\ell\propto -MF(D)+n\ln(M)+\sum_{i=1}^{n}\ln(f(u_i)) \quad (24)$$

Note that

$$\frac{\partial\ell}{\partial M}=-F(D)+\frac{n}{M} \quad (25)$$

Setting the partial derivative equal to 0 results in

$$\hat{M}=\frac{n}{F(D)} \quad (26)$$

Evaluating ℓ at $\widehat{M}$ results in

$$\ell=-n+n\ln(n)-n\ln(F(D))+\sum_{i=1}^{n}\ln(f(t_i)) \quad (27)$$

Estimate the parameters of F using the conditional distribution of $\boldsymbol{U}$ given $\boldsymbol{U}\le D$.

To illustrate, consider the empirical data displayed in Table 1 of [31]; cf. [6]. The data are the times between failures (in days) of a piece of software (FM types not distinguished). There are 34 failures. Each FM (software bug) occurs at most once. The software was released after the 31st failure which occurred on day 540. The last 3 failures occurred after the software was fielded. The last observation is a FM activation on day 849. We estimate model parameters using FM activation times that occur at or before day 600. Then use these estimates to estimate the expected number of FMs that activate in the remaining time until day 849.

Three distributions F are considered: the often invoked *exponential*; the *gamma-mixed exponential* distribution, (5); and the inverse Weibull distribution, $F_{IW}(t)=\exp\{-[1/\beta t]^\alpha\}$. The exponential with parameter mixed by a stable law, (6), could not be used to summarize the data: the data do not support an estimated stable shape parameter less than 1.

Maximum likelihood is used to estimate the parameters of the distribution F using the 31 FM activation times less than or equal to 600 days. The mean initial number of FMs is estimated as

$$\hat{M}=\frac{31}{\hat{F}(600)} \quad (28)$$

The estimated expected number of FMs that activate between day 600 and the last recorded FM activation time of 849 days,

$$\hat{E}[\boldsymbol{N}(849)-\boldsymbol{N}(600)]=\hat{M}\left[\hat{F}(849)-\hat{F}(600)\right] \quad (29)$$

Table 1
Estimates using data consisting of failures occurring at or before day 600.

Distribution	Estimates [95% CI]	Estimated probability FM activates at or before 600 days [95% CI]	Estimated expected number of FMs remaining at time 600 [95% CI]	Estimated probability 0 FMs remain after day 600 [95% CI]	Estimated expected number of FMs that activate in the time interval (600, 849] [95% CI]
Exponential $F(t)=1-e^{-\beta t}$	$\hat{\beta}=0.0062$ [0.004, 0.009]	0.98 [0.91, 1.0]	0.8 [0.11, 2.9]	0.46 [0.06, 0.89]	0.61 [0.10, 1.7]
Exponential with parameter mixed with a gamma distribution $F(t)=1-[1+\xi t/\alpha]^{-\alpha}$	$\hat{\alpha}=11.3$, [1, 176] $\hat{\xi}=0.0064$ [0.004, 0.01]	0.99 [0.76, 0.99]	1.2 [0.14, 10.1]	0.31 [0, 0.87]	0.80 [0.12, 2.0]
Inverse Weibull $F(t)=\exp\{-(1/\beta t)^{\alpha}\}$	$\hat{\alpha}=0.78$, [0.6, 1.5] $\hat{\beta}=0.011$ [0.005, 0.17]	0.79 [0.61, 0.96]	8.0 [1.0, 18.3]	0.0003 [0, 0.35]	1.73 [0.39, 3.0]
Power law NHPP $E[\boldsymbol{N}(t)]=\gamma t^{\delta}$	$\hat{\gamma}=0.89$ [0.43, 1.58] $\hat{\delta}=0.56$ [0.47, 0.67]	NA	NA	NA	6.7 [5.4, 8.1]

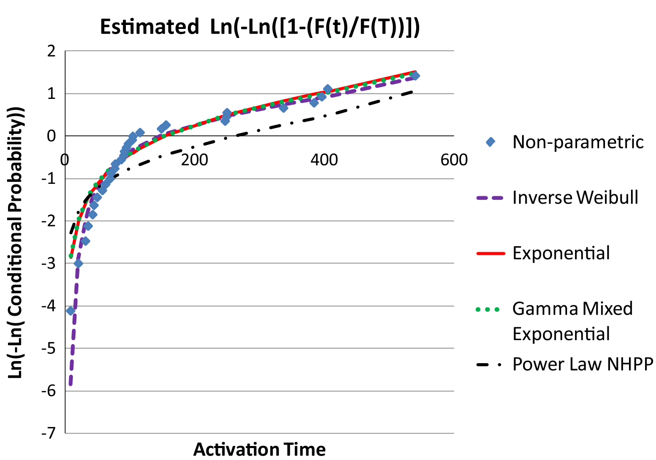

Fig. 1. Transformed estimated survivor function of FM activation time.

Another model often used to summarize data of this kind is the power law non-homogeneous Poisson process; that is, $\{\boldsymbol{N}(t); t \geq 0\}$ is a non-homogenous Poisson process (NHPP) with mean value function $E[\boldsymbol{N}(t)]=\gamma t^{\delta}$; cf. [32] which also discusses maximum likelihood estimation of γ and δ. The power law model is often referred to as the Duane-Crow-AMSAA model in the context of reliability growth; cf. [3,4,28]. The conditional distribution of the unordered activation times during the first 600 days given activation during the first 600 days is that of 31 independent and identically distributed random variables having distribution, (cf. [33]),

$$F_{NHPP}(t)=E[\boldsymbol{N}(t)]/E[\boldsymbol{N}(600)] \quad (30)$$

For the models that are not the power law NHPP, Table 1 displays the distributions' parameter estimates; the estimated expected mean number of FMs remaining after day 600; and the estimated probability a FM activates on or before Day 600. Also displayed are the estimated expected number of FMs that activate between day 600 and the last recorded FM activation time of 849 days, computed using (29). Table 1 also displays the estimated parameters for the power law NHPP and the resulting estimated expected number of FMs that activate between day 600 and day 849,

$$\hat{E}_{NHPP}[\boldsymbol{N}(849)-\boldsymbol{N}(600)]=\hat{\gamma}\left[849^{\hat{\delta}}-600^{\hat{\delta}}\right] \quad (31)$$

Bootstrap is used to assess the variability of the estimates; cf. [34]. A bootstrap replication consists of

1) The number of FMs that are observed during the first 600 days which is simulated from a Poisson distribution with mean equal to the number of observed FMs; in the example, the mean is 31.
2) A collection of activation times sampled with replacement from the observed activation times. The number in the collection is the number of observations, 31.

For the models that are not the power law NHPP, the parameters of the activation time distribution are estimated using the bootstrap sample of the activation times and the total test time of 600 days; the mean number of FMs at time 0 is estimated using (28). The bootstrap results for the power law NHPP use the bootstrap samples of 31 activation times.

100 bootstrap replications were conducted. An estimated 95% confidence interval is computed, with lower bound the 0.025 percentile and upper bound the 0.975 percentile of the bootstrap sample. The expected number of FMs that activate between day 600 and day 849 is computed using (29) and (31). The same bootstrap replications are used to estimate parameters for all the models. The results appear in Table 1.

Let u_i be the ith observed FM activation time. Fig. 1 displays $\ln(-\ln(1-[\hat{F}(u_i)/\hat{F}(600)]))$ versus u_i for each of the three models with an activation time distribution; u_i versus $\ln(-\ln(1-\hat{F}_{NHPP}(u_i)))$ is displayed; also displayed is u_i versus the non-parametric estimate $\ln[-\ln(1-[(i-0.5)/31])]$.

Fig. 1 suggests that the inverse Weibull distribution summarizes the 1st 31 activation times somewhat better than the other models; however the convenient simple exponential also summarizes well. Note that these data are used to estimate the distributions' parameters. From Table 1, the power law NHPP predicts a larger expected number of FMs occurring in the time period (600, 849] than the other 3 models. Among the other three models, the inverse Weibull distribution results in an estimated mean number of FMs remaining at day 600 about 6 times larger than those of the other models with activation time distributions. The Inverse Weibull distribution also estimates a mean number of FMs activating between day 600 and 849 over twice as large as the other two models with activation time distributions; the observed number of FMs activating during this time period is 3. However, there is no statistical difference between estimates of the expected number of FMs activating between day 600 and 849 for the three models with activation time distributions.

Since the inverse Weibull model appears to be the best summary of the data, we present its parameter estimates using all the data: The parameter point estimates for the inverse Weibull model using all the failure data are $\hat{\alpha} = 0.7$; $\hat{\beta} = 0.009$. The resulting estimated mean number of FMs remaining after the last recorded FM activation is 9.3. The estimated median of the distribution of the number of additional days until the next FM activation is 113 days; the 90th percentile is 487 days. Thus, even after 3 additional failures, it may be premature to release the software.

6. Conclusions

In this paper models are presented to assess system reliability growth. System reliability growth is achieved by testing the system to find failure modes (FMs) and mitigating their effects. Removing or mitigating the effect of FMs increases the system's field reliability and availability. The model notion is that of identifying failure mode creation and removal with an “infinite server queue”, a generalized so-called M/G/∞ system. The models represent the number of FMs; the time until a FM is activated/detected, and the subsequent experience of the FM until its removal. The number of FMs is assumed to have a Poisson distribution and the times to detect and subsequent experiences of the FMs are modeled as independent of each other and of the number of FMs. Models that represent the inherent variation between FMs are considered. The estimation of model parameters and assessment of model summarization of data are discussed and illustrated. The models proposed here appear to summarize the illustrative data better than the “traditional” non-homogeneous power law model.

References

[1] Cai K-Y, Wen C-Y, Zhang M-L. A critical review on software reliability modeling. Reliab Eng Syst Saf 1991;32(3):357–71.
[2] Fries A, Sen AA. Survey of discrete reliability growth models. IEEE Trans Reliab 1996;R-45(4):582–604.
[3] Crow LH. Reliability analysis for complex repairable systems. In: Proschan F, Serfling RJ, editors. Reliability and biometry. Philadelphia, PA: SIAM; 1974. p. 379–410.
[4] Crow LH. A Method for achieving an enhanced mission capability. IEEE Proc Reliab Maintainability Symp 2002:153–7.
[5] Raftery A. Inference for the binomial *N* parameter: a hierarchical Bayes approach. Biometrika 1988;75(2):223–8.
[6] Raftery A. Analysis of a simple debugging model. Appl Stat 1988;37(1):12–22.
[7] Ellner PM, Hall JB. An approach to reliability growth planning based on failure mode discovery and correction using AMSAA projection methodology. IEEE Proc Reliab Maintainability Symp 2006:266–72.
[8] Hall JB, Mosleh A. An analytical framework for reliability growth of one-shot systems. Reliab Eng Syst Saf 2008;93(11):1751–60.
[9] Hall JB, Ellner PM, Mosleh A. Reliability growth management metrics and statistical methods for discrete-use systems. Technometrics 2010;52(4): 379–89.
[10] Gaver DP, Jacobs PA, Glazebrook KD, Seglie EA. Probability models for sequential-stage system reliability growth via failure mode removal. Int J Reliab Qual Saf Eng 2003;10:15–40.
[11] Gaver DP, Jacobs PA. Comment on “reliability growth management and statistical methods for discrete use systems”. Technometrics 2010;52(4): 389–91.
[12] Yang B, Li X, Xie M, Tan F. A generic data-driven software reliability model with model mining technique. Reliab Eng Syst Saf 2010;95(6):671–8.
[13] Singpurwalla ND, Wilson SP. Statistical methods in software engineering. New York, NY: Springer; 1999.
[14] Yu JW, Tian GL, Tang ML. Statistical inference and prediction for the Weibull process with incomplete observations. Comput Stat Data Anal 2008;52: 1587–603.
[15] Tian GL, Tang ML, Yu JW. Bayesian estimation and prediction for the power law process with left truncated data. J Data Sci 2011;9:445–70.
[16] Martin OS. A dynamic risk model for filtering reliability and tracking survivability. Ph.D. dissertation, George Washington University, UMI number 3490816, ProQuest LLC, 2012.
[17] Chiu K-C, Huang Y-S, Lee T-Z. A study of software reliability growth from the perspective of learning effects. Reliab Eng Syst Saf 2008;93:1410–21.
[18] Littlewood B. Comments on ‘evolutionary neural network modeling for software cumulative failure time prediction'. Reliab Eng Syst Saf 2006;91:485–6.
[19] Hu QP, Xie M, Ng SH, Levitin G. Robust recurrent neural network modeling for software fault detection and correction prediction. Reliab Eng Syst Saf 2007;92:332–40.
[20] Bunkley N. Toyota issues a 2nd recall. New York Times; 2010.
[21] Davis TP. Science, engineering and statistics. Appl Stochastic Models Bus Ind 2006;22:401–30.
[22] Gaver DP. Random hazard in reliability problems. Technometrics 1963;5(2): 211–26.
[23] Tijms HC. Stochastic modeling and analysis: a computational approach. New York, NY: John Wiley & Sons; 1986.
[24] Dohi T, Osaki S, Trivedi KS. An infinite server queueing approach for describing software reliability growth. In: Proceedings of the 11th Asia-Pacific Software Engineering Conference 2004: 110-119.
[25] Huang C-Y, Huang W-C. Software reliability analysis and measurement using finite and infinite server queueing models. IEEE Trans Reliab 2008;57 (1):192–203.
[26] Kapur PK, Pham H, Anand S, Yadov K. A unified approach for developing software reliability growth models in the presence of imperfect debugging and error generation. IEEE Trans Reliab 2011;60(1):331–9.
[27] Feller W. An introduction to probability theory and its applications, 2. New York, NY: John Wiley & Sons; 1966.
[28] Duane JT. Learning curve approach to reliability monitoring. IEEE Trans Aerospace 1964;5:249–67.
[29] Ellner PM, Broemm WJ, Woodworth WJ. AMSAA reliability growth guide. Technical report no. TR-652, US Army Materiel Systems Analysis Activity: Aberdeen Proving Ground, Maryland 21005-5071, 2000.
[30] Feller W. 3rd ed. An introduction to probability theory and its applications, 1. New York, NY: John Wiley & Sons; 1968.
[31] Goel AL, Okumoto K. Time-dependent error-detection rate model for software reliability and other performance measures. IEEE Trans Reliab 1979;R-28(1): 206–11.
[32] Crowder MJ, Kimber AC, Smith RL, Sweeting TJ. Statistical analysis of reliability data. London: Chapman & Hall; 1991.
[33] Cox DR, Isham V. Point processes. London: Chapman & Hall; 1980.
[34] Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: Chapman & Hall/CRC; 1993.