Bayesian Modelling for Product Testing and Release

John G. Wilson

▶ **To cite this version:**

John G. Wilson. Bayesian Modelling for Product Testing and Release. IFIP International Conference on Advances in Production Management Systems (APMS), Aug 2020, Novi Sad, Serbia. pp.63-70, 10.1007/978-3-030-57993-7_8 . hal-03630889

HAL Id: hal-03630889

https://inria.hal.science/hal-03630889

Submitted on 5 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Bayesian Modelling for Product Testing and Release

John G. Wilson

Ivey Business School
[2]University of Western Ontario, London, Ontario N6G ON1, Canada
jwilson@ivey.uwo.ca

Abstract. Deciding when to release a new product requires a tradeoff between costs, potential profits and the underlying reliability of a product. Many new products go through a "Test Analyze and Fix" process. When a failure occurs, an immediate design "fix" may occur or the product might undergo a minimal fix with design changes being delayed until later when many changes can be introduced at the same time. We introduce a Bayesian model that allows for the introduction of managerial knowledge and experience. Unlike most approaches, we do not build in an assumption that the product always improves throughout the process.

Keywords: Product Development, Prototype Testing, Product Release, Reliability Growth.

1 Introduction

1.1 Background

The mean time between failures of a new product (MTBF) is an important characteristic that has a direct impact on revenue streams and costs. For products designed for the US military, historically MTBF has to reach a specified level before a product will be accepted. In the commercial world, there is a balance between MTBF and the timing of a new product release: release the product too early and the reliability problems will damage a company's reputation and involve many repair and replacement costs; release the product after extensive and time intensive development to raise MTBF and market advantage and share might be lost. Before formally incorporating costs and revenues into a decision model of how to manage the development phase of a new product and the timing of product release, it is essential to have a flexible statistical process for modelling the phases of test development.

A complex newly designed system generally undergoes several stages of development testing before it is put into operation. After each stage of testing, changes are made to the design with the hope that the new design leads to a longer period of performance. This procedure is referred to as *reliability growth*. It is the result of iterative *Test Analyse and Fix* processes which are conducted to discover deficiencies and to verify that corrective action will prevent recurrence in the further test phases. A "fix" refers to a design change that improves the reliability of the system

After fixes have been implemented, the system reliability will jump to a higher value. Usually, the estimate of the reliability jump is not straightforward since the untested fixes require a specific prediction rule for them to "follow" in a growth model. In addition, the test data of a specific test prototype is only one set of data of all possible situations from the same design basis, e.g. design specifications and environmental specifications (Ireson et al. [1]).

1.2 Literature

The first reliability growth model appeared in the late 1950's. Duane [2] analysed the reliability data of five different complex systems and demonstrated that the cumulative failure rate versus the cumulative operation time fell close to a straight line when plotted on a log-log scale. Crow [3] at *Army Materiel Systems Activity* (AMSAA) suggested that Duane's postulate be stochastically represented as a Non-Homogeneous Poisson Process (NHPP) with Weibull intensity $\lambda(t) = \alpha\beta t^{\beta-1}$ where$\alpha, \beta > 0$. This is called the AMSAA model. For $0 < \beta$<1, $\lambda(t)$ is decreasing implying reliability growth. Bayesian extensions for the AMSAA model were proposed by Higgins and Tsokos [4] and Guida et al. [5]. Yu et al. [6] consider predictions using this model using non-informative priors. A pseudo-Bayes approach has been discussed by Singpurwalla [7] under the stochastic ordering of prior and posterior distributions. Robinson and Dietrich [8-9] proposed two nonparametric growth models in which no specific functional form is assumed for the change of process failure rate. Crow's *Engineering Judgement Model* [10] (EJM) considered techniques for *projection* of future expected reliability based on a delayed fixes testing plan. In the EJM model it is assumed that the *effectiveness factors* (EF) of fixes for all distinct failure modes in the system are given and all failures observed in RGDT are not fixed until the end of the testing program.

Wayne and Modarres [11] assume that the failure times between fixes are exponential and that a fix results in a known increase in mean time between failures. (MTBF's). Conjugate priors are used to model knowledge of the parameters. Wayne [12] discusses using a beta prior distribution for the factor by which MTBF increases after a fix. Our model generalises this to assuming that the increase in MTBF is itself a random variable and that MTBF's follow a pattern similar to that of AMSAA NHP approach. For software projects, predicting the expected number of software defects is important and Bayesian methods are making inroads. Rana et al. [13] discuss models including the Weibull model which is similar to the NHPP above and discuss finding prior distributions. Chen et al. [14] consider a multi-stage system with NHPP at each stage and models growth factors. Non-Bayesian approaches to modelling RDGT using power law growth models such as NHPP are still areas of active research--see, e.g., Xu et al. [15].

Bayesian methods for estimating the reliability of systems are examined in Li et al. [16] and Ruiz et al. [17]. A comparison of classical versus Bayesian methods can be found in Kamranfar et al. [18]. Pollo et al. [19] introduce an "objective Bayesian prior" that produces a posterior that is a product of gamma functions. Wayne et al. [20] model a posterior distribution for the failure intensity Optimal product release times for a product consisting of subsystems connected in series are modeled in Li et al [21].

In distinction to the above literature, we don't even require that the reliability grows…some fixes don't work! However, we do assume that on average fixes work and the *expected* reliability grows according to a power law. The parameters of this power law are unknown and we assume that we can obtain prior distributions for the parameters. However, we translate these parameters into quantities that are intuitive and will enable engineers to arrive at reasonable priors. This approach has the advantage of being transparent and intuitive. However, mathematically, arriving at conclusions will now require numerical methods.

2 A Bayesian Product Improvement Model

We will start with a description of the basic model and assumptions. For most of this paper, we consider the model where a fix is incorporated after each failure. Thus, as the machine fails and is repaired, its reliability grows. This is discussed in $2.1, where model assumptions are discussed in detail. The likelihood function for the data is formulated in $2.2. A prior density function $\pi(\alpha, \beta)$ over the parameters α and β of the Weibull intensity $\lambda(t) = \alpha\beta t^{\beta-1}$ is required. Ideally, this should incorporate expert opinion and knowledge. We will provide the likelihood function for the observed data. Then we will describe how to come up with reasonable prior distributions for the underlying parameters. This is an important feature of the model since it allows for the incorporation of expert managerial judgement. This is particularly important in product development where, initially, not much data is available. Most product developers will have a feel for the possible values of the reliability growth parameter β. This is discussed in $2.3. Information about the parameter α is not so intuitive. Instead, in $2.4, we propose the product developer provide information about the expected mean time between failures. This can then be translated into prior information for α. In $2.6, we combine the material from $2.3 and $2.4 to provide a joint prior for the underlying parameter.

2.1 Model Development

After each failure and fix, the time to the next failure is assumed to be exponential but with a new parameter since the machine has been "improved". This is the assumption made in Wayne and Modarres [11].

Suppose the system testing stops at some predetermined number of failures n. Let $t_1, t_2, \ldots, t_n$ be the cumulate test times of n successive failure modes observed up to t_n and let $x_1, x_2, \ldots, x_n$ be the times between failures i.e. $x_1 = t_1$and $x_i = t_i - t_{i-1}$, for i =2,…,n.

So the number of failures in each stage is limited to one. We note that the test plan of the most popular reliability growth model using NHP belongs to this class of problem.

During a particular stage, the distributions of the times to failure are assumed exponential. The exponential parameters are not assumed to be equal. In Wayne and Modarres [11], the assumption is that each fix results in a known fractional increase in meant time to failure. We generalise this by assuming that exponential parameter for stage i after a fix is assumed to be the value of a gamma random variable with mean

$\alpha\beta t_{i-1}{}^{\beta-1}$ and variance σ^2. We will assume that the value $\sigma^2$is provided—although extension to a prior over this parameter is straightforward. This allows for the incorporation of the major feature of the AMSAA model. In addition, in distinction to other models, the actual failure rate does not necessarily decrease from stage to stage. This is particularly important when considering the development of software products: a fix is new computer code which may itself be faulty and decrease the reliability of the product. However, overall the reliability tends to improve as fixes are incorporated. The parameters α, β and σ will depend on managerial specifications and knowledge and are key components. There are various ways to obtain information that can help set prior distributions. For instance, Wayne [11] discusses developing prior distributions for the factor by which MTBF increases after a fix and suggests the use of historical information such as that found in Trapnell [22] and in Brown [23]

On finding prior distributions, we need to obtain the posterior distributions for the parameters and the posterior expected value of various quantities of interest such as the expected time to failure once all the test data and fixes have been incorporated. Since we are making as few assumptions as possible and trying to capture managerial intuition, these are algebraically messy and ultimately need to be evaluated using numerical methods.

2.2 The Likelihood Function

At the *nth* failure the data observed will be the times $t_1, t_2, \ldots, t_n$. After simplification, the likelihood function –a constant times the probability of observing the data $t_1, t_2, \ldots, t_n$, can be written as:

$$\prod_{i=2}^{n} \alpha\beta t_{i-1}^{\beta-1} \left\{1 + \frac{(t_i - t_{i-1})\sigma^2}{\alpha\beta t_{i-1}^{\beta-1}}\right\}^{-\left(1+(\alpha\beta t_{i-1}^{\beta-1})^2\right)} .$$

Sufficient statistics of fixed dimensionality do not exist for the above likelihood function so that no natural conjugate prior can be found (Schlaifer and Raiffa [24]). Thus, inference based on a continuous joint prior generally requires numerical procedures.

From a practical viewpoint, Bayes methods are attractive because they allow for incorporating experience or technical considerations in the estimation procedure especially when the test data and the test procedure is expensive. This is greatly facilitated when a clear physical meaning can be attached to the model parameters. Generally, for selecting a prior distribution, it must adequately represent the state of prior knowledge about the parameter and not be a computational burden. Based on a technical judgement, if we believe the variability between prototypes will be relatively small it may be reasonable to take σ as a common constant parameter of the model and we do that in the sequel.

2.3 Prior for β

We note that a shape parameter β less than 1 means a system whose expected failure rate is decreasing (small β means rapid reliability growth) whereas a shape parameter β greater than 1 means a system whose expected failure rate is increasing. In many

practical situations, bounds for β can follow on the basis of prior knowledge of the underlying failure mechanism. For instance, guidance on practical ranges for β can be found in the following:

- The data collected by Duane [2] and Schafer et al. [25] show that β is "about 0.5 and is rarely less than 0.3 regardless of the initial and final MTBF's"
- For equipment exhibiting a good fit, the Crow's growth parameter β ranged from 0.579 to 0.794 (Gates et al. [26])
- The Duane model provides a reasonably good fit to most of the data sets and the estimated growth rates are within the generally accepted range or 0.3 to 0.9 (Gates et al [26])

When no further information exists, it is reasonable to represent prior knowledge by using a uniform prior pdf over the interval (0.3,0.9), the widest range for β as indicated by Gates et al. [26].

2.4 Prior for α

Since the scale parameter α in this model does not provide a clear physical meaning, we suggest an alternative approach to find a reasonable joint pdf for α and β as follows. For the failure process considered here, the expected failure time at time *t* is $\alpha\beta t^{\beta-1}$. We note that generally a development test program is never begun without some information regarding potential failure rates. So, the mean life time at some beginning time t_0 can be viewed as a process parameter μ_0. Without loss of generality, suppose that $t_0 = 1$. The physical meaning of μ_0 should allow an engineer to formulate prior knowledge. However, if no prior information is available from a similar system then a commonly used estimate of 10% to 20% of the requirement mean time between failures can be uses as an initial reliability (Morris and MacDiarmid [27]). Using this, the prior for μ_0.can be chosen as uniform over some interval (c_1, c_2):

$$\pi_{\mu_0}(\mu_0) \propto \frac{1}{c_2-c_1} \text{ for } c_1 < \mu_0 < c_2.$$

2.5 Joint prior for α and β

Assuming the random variables $\mu_0$and β are independent, the joint density function of $\mu_0$and β is given by

$$\pi_{\mu_0,\beta}(\mu_0,\beta) \propto \frac{1}{c_2-c_1}\frac{1}{\beta_2-\beta_1} \text{ for } c_1 < \mu_0 < c_2 \text{ and } \beta_1 < \beta < \beta_2.$$

By making a change of variable transformation, this induces an equivalent prior over (α, β) given by

$$\pi(\alpha,\beta) \propto \frac{1}{\beta\alpha^2}$$

over the region $0<\beta_1 < \beta < \beta_2<1$ and $0<c_1 \leq \frac{1}{\alpha\beta} \leq c_2$.

that μ_0.and β are independent, the joint density function is then equal to $\frac{1}{(c_2-c_1)(\beta_2-\beta)}$. Changing variables the joint distribution of α and β can be shown to be proportional to $\frac{1}{\beta\alpha^2}$ where $0 < \beta_1 \leq \beta \leq \beta_2 < 1$ and $0<c_1 \leq \frac{1}{\alpha\beta} \leq c_2$.

2.6 Putting it all together

The likelihood from 2.1 is now multiplied by the prior from 2.5 to form a quantity proportional to the posterior distribution. From this, various posterior quantities such as the posterior MTBF, posterior means of α and β may be obtained via integration. However, the resultant integrals have no closed form and the integrations, of necessity, must be numerical.

3 Applying the model

To illustrate the application of the Bayesian model, we use the 15 failure data points from Crow's [28] example. These data were generated by the computer simulation of a NHPP with α=0.42 and β = 0.5. These are (cumulative test times): 0.2, 11.2, 37.2, 39, 48.4, 53.4, 90.2, 91.6, 151.4, 159.4, 197.2, 240.2, 323.6, 361.2, 381.6.

It is important to note that we know a priori that the Bayes model does not fit this data since it was generated assuming the time between failures is non-exponential and there is no variability between prototypes. Thus, the simulated model is at best an approximation to the model which most authors suggest. As will be seen, even for this unfavourable data, the proposed Bayesian model performs well.

We assumed that β was uniform over (0.45, 0.55), μ_0 was uniform over (4.5 and 5) and σ was small.

The table below summarises the estimation results. Included are the results from applying the Bayesian model and the estimated using maximum likelihood estimation of α and β in Crow's model.

Table 1. Comparison of Model to Estimates from Crow's Model where underlying data is generated from a known distribution

Model	True σ	True β
True Model	0.42	0.5
Bayesian Estimates	0.40463	0.522
MLE for Crowe Model	0.4690	0.5829

4 Conclusions

Planning for new product release requires a reasonable model for the modelling of reliability growth. An approach to product reliability growth modelling that allows for managerial input, variability between prototypes and incorporates historically relevant observations has been presents. In the case of limited data, a Bayesian approach combined with expert judgement, often provides superior results. A limitation of the Bayesian approach is arriving at the prior distributions. This depends on the expert knowledge of the product developers. For cases where development time is short, misspecifications can lean to costly errors. Continuing work involves expanding the model to more general situations and performing extensive numerical comparisons. Now that a feasible model for tracking product reliability growth has been established, a next step is to explicitly incorporate costs and revenues to optimise the testing and fix process.

Acknowledgment
This work was accomplished with the assistance of T. Chu and forms part of his thesis

References

1. Ireson, W., Grant and Coombs, Clyde F., Jr.. *Handbook of Reliability Engineering and Management*, McGraw-Hill Book Company, 1988.
2. Duane, J. T. "Learning curve approach to reliability monitoring." *IEEE transactions on Aerospace* 2, no. 2 (1964): 563-566.
3. Crow, L.H "Reliability Analysis for Complex Repairable Systems" US Army Materiel Systems Analysis Activity, Technical Report 138, Dec. 1975, AD-A020296,1975.
4. Higgins, J. J., and C. P. Tsokos. "A quasi-Bayes estimate of the failure intensity of a reliability-growth model." *IEEE Transactions on Reliability* 30, no. 5 (1981): 471-475.
5. Guida, M., R. Calabria, and G. Pulcini. "Bayes inference for a non-homogeneous Poisson process with power intensity law (reliability)." *IEEE Transactions on Reliability* 38, no. 5 (1989): 603-609.
6. Yu, Jun-Wu, Guo-Liang Tian, and Man-Lai Tang. "Predictive analyses for nonhomogeneous Poisson processes with power law using Bayesian approach." *Computational Statistics & Data Analysis* 51, no. 9 (2007): 4254-4268.
7. Singpurwalla, N. "A Bayesian scheme for estimating reliability growth under exponential failure times." *TIMS Studies in the Management Science (19) pp* (1982): 281-296.
8. Robinson, David G., and Duane Dietrich. "A new nonparametric growth model." *IEEE Transactions on Reliability* 36, no. 4 (1987): 411-418.
9. Robinson, D., and Duane Dietrich. "A nonparametric-Bayes reliability-growth model." *IEEE Transactions on Reliability* 38, no. 5 (1989): 591-598.
10. Crowe, L.H. "Reliability Growth Assessments Utilizing Engineering Judgements." *Proc. Institute for Environmental Sciences*, pp 179-183, 1987.
11. Wayne, Martin, and Mohammad Modarres. "A Bayesian model for complex system reliability growth under arbitrary corrective actions." *IEEE Transactions on Reliability* 64, no. 1 (2014): 206-220.
12. Wayne, Martin. "Modeling Uncertainty in Reliability Growth Plans." In *2018 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-6. IEEE, 2018.

13. Rana, Rakesh, Miroslaw Staron, Christian Berger, Jörgen Hansson, Martin Nilsson, and Wilhelm Meding. "Analyzing defect inflow distribution and applying Bayesian inference method for software defect prediction in large software projects." *Journal of Systems and Software* 117 (2016): 229-244.
14. Chen, Guangyu, Liang Li, Jing Li, and Qiang Miao. "Multi-stage Bayesian evaluation method of reliability growth for large complex system." In *2015 Prognostics and System Health Management Conference (PHM)*, pp. 1-7. IEEE, 2015.
15. Xu, J. Y., D. Yu, M. Xie, and Q. P. Hu. "An approach for reliability demonstration test based on power-law growth model." *Quality and Reliability Engineering International* 33, no. 8 (2017): 1719-1730.
16. Li, Zhaojun Steven, Jian Guo, Ning-Cong Xiao, and Wei Huang. "Multiple priors integration for reliability estimation using the Bayesian melding method." In *2017 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-6. IEEE, 2017.
17. Ruiz, Cesar, Mohammadhossein Heydari, Kelly M. Sullivan, Haitao Liao, and Ed Pohl. "Data analysis and resource allocation in Bayesian selective accelerated reliability growth." *IISE Transactions* 52, no. 3 (2020): 301-320.
18. Kamranfar, H., J. Etminan, and M. Chahkandi. "Statistical inference for a repairable system subject to shocks: classical vs. Bayesian." *Journal of Statistical Computation and Simulation* 90, no. 1 (2020): 112-137.
19. Pollo, Marco, Vera Tomazella, Gustavo Gilardoni, Pedro L. Ramos, Marcio J. Nicola, and Francisco Louzada. "Objective Bayesian Inference for Repairable System Subject to Competing Risks." *arXiv preprint arXiv:1804.06466* (2018).
20. Wayne, Martin, and Mohammad Modarres. "A **B**ayesian Model for Complex System Reliability Growth under Arbitrary Corrective Actions." *IEEE Transactions on Reliability* 64, no. 1 (2015): 206-220.
21. Li, Meiyan, Dong Xu, and Zhaojun Steven Li. "A Joint Modeling Approach for Reliability Growth Planning Considering Product Life Cycle Cost Performance." *Computers & Industrial Engineering* (2020): 106541.
22. Trapnell, B. "Update on Reliability Fix Effectiveness Factors." *Interim Note R-86*, AMSAA, Aberdeen Proving Ground, MD 1984.
23. Brown, S.M., Development and validation of methodology for fix effectiveness during product development. Ph.D. Dissertation, University of Maryland, College Park, 2009.
24. Schlaifer, Robert, and Howard Raiffa. *Applied statistical decision theory*. 1961.
25. Schafer, R. E., R. B. Sallee, and J. D. Torrez. "Reliability growth study." *Hughes Aircraft Co., RADC-TR-75-253* (1975).
26. Gates, R. Gibson, G. and McLain. "Reliability Growth Projections" *Rome Air Development Study,* TR 86-148, Sept 1986.
27. Morris, Seymour, and Preston MacDiarmid. "Reliability Growth Testing Effectiveness." *The Journal of Environmental Sciences* 28, no. 4 (1985): 21-28.
28. Crow, Larry H. "A methodology for managing reliability growth during operational mission profile testing." In *2008 Annual Reliability and Maintainability Symposium*, pp. 48-53. IEEE, 2008.