

A Comparative Study of the Software NHPP Based on Weibull Extension Distribution and Flexible Weibull Extension Distribution

Tae-Jin Yang and Jea-Gun Park

Department of Electronic Engineering, Hanyang University, Seoul, Republic of Korea

Abstract: Software testing (debugging) to reduce testing costs is essential for the software reliability. In this study, the comparative problem for the reliability model using the Weibull extension distribution based on the generalized order statistics and the flexible Weibull extension distribution which made out efficiency about the software reliability was proposed. The maximum likelihood estimation and bisection iteration method for the estimating parameters were used. In addition, the model selection based on the mean square error and coefficient of the determination for the efficient model were offered. Analysis of the failure time based on the Weibull extension distribution and the flexible Weibull extension distribution was employed for the proposing reliability model. The Laplace trend test was employed for the assurance property about the failure time. The result of this study was obtained that the proposed Weibull extension distribution model than the flexible Weibull extension model is more efficient in terms of the reliability under a fixed shape parameter condition. Thus, the Weibull extension distribution model can be used as a reliability alternative model. From this study, the software developers must be filled to the growth model by a prior knowledge for the software to identify the failure modes which can be benefited.

Key words: Weibull extension distribution, flexible weibull extension distribution, NHPP, failure, Laplace

INTRODUCTION

Software failures were caused by failure of computer systems in our society can lead to huge losses. Thus, software reliability in the software development process is an important issue. These issues of the user requirements meet the cost of testing. Software testing (debugging) in order to reduce costs in terms of changes in the software reliability and testing costs, need to know in advance is more efficient. Thus, the reliability, cost and consideration of release time for software development process are essential. Eventually, the software to predict the contents of a defect in the product development model is needed. Until now, many software reliability models have been proposed. Non-Homogenous Poisson Process (NHPP) models rely on an excellent model (Gokhale and Trivedi, 1999; Goel and Okumoto, 1979) in terms of the error discovery process and if a fault occurs, immediately remove the debugging process and the assumption that no new fault has occurred.

In this field, enhanced non-homogenous Poisson process model was presented (Gokhale and Trivedi, 1999), proposed an exponential software reliability. In this model, the total numbers of defects have S-shaped or exponential-shaped with a mean value function was used

model (Goel and Okumoto, 1979). The generalized model relies on these models, delayed S-shaped reliability growth model and inflection S-shaped reliability growth model were proposed (Yamada *et al.*, 1983). Software reliability problems in change point were proposed software reliability problems in change point (Zhao, 1993) and the generalized reliability growth models proposed (Shyur, 2003). In testing measured coverage, the stability of model with software stability can be evaluated was presented (Pham and Zhang, 2003).

Relatively recently, generalized logistic testing-effort function and the change-point parameter by incorporating efficient techniques to predict software reliability were presented (Huang, 2005). The learning process that software managers to become familiar with the software and test tools for S-type model can be explained (Chiu *et al.*, 2008).

In this study, the comparative problem for the reliability model with the Weibull extension distribution and flexible Weibull extension distribution which made out efficiency for the software reliability was proposed using the nonhomogeneous Poisson process with an infinite number of faults.

The proposed process involves evaluation of the parameter of the mean value function and hence the mean

value function of infinite, finite failure model which is based on Non Homogeneous Poisson Process (NHPP) and curve regression models were considered.

Literature review

NHPP model: The mean value function and the intensity function (Gokhale and Trivedi, 1999) for Non-Homogeneous Poisson Process (NHPP) Model are given by:

$$m(t) = \int_0^t \lambda(s)ds, \frac{dm(t)}{dt} = \lambda(t) \quad (1)$$

Therefore, N(t) is known Poisson Probability Density Function (PDF) with the parameter m(t). In other words:

$$p[N(t) = n] = \frac{[m(t)]^n}{n!} e^{-m(t)} \quad n = 0, 1, \dots, \infty \quad (2)$$

These time domain models for the NHPP process can be described by the probability of failure are possible. This model is the failure intensity function $\lambda(t)$ expressed differently, also mean value function $m(t)$ will be expressed differently. These models are classified into categories, the finite failure NHPP models and infinite failure (Kuo and Yang, 1995). Finite failure NHPP models if they are given sufficient time to test, the expected value of faults has a finite expectation $\lim_{t \rightarrow \infty} m(t) = \theta < \infty$ value while infinite failure NHPP model assumes that the value is infinite. Let θ denote the expected number of faults that would be detected given finite failure NHPP Models. Then, the mean value function of the finite failure NHPP models can also be written as:

$$m(t) = \theta F(t) \quad (3)$$

Note that F(t) is a CDF (cumulative distribution function). From Eq. 3 the (instantaneous) failure intensity $\lambda(t)$ in case of the finite failure NHPP models is given by:

$$\lambda(t) = \theta F'(t) = \theta f(t) \quad (4)$$

In finite model, at the time of each repair, a new defect is assumed not to occur. However, the actual situation at the point of repair new failure may occur. Add to this situation, the RVS (Record Value Statistics) model could be used NHPP model and mean value function was as follows (Kuo and Yang, 1995).

$$m(t) = -\ln[1-F(t)] \quad (5)$$

Equation 5 is mean value function of infinite failure NHPP model. Therefore, from Eq. 5 using the related equations of NHPP in Eq. 1 intensity function can be the hazard function (h(t)). In other words:

$$\lambda(t) = m'(t) = f(t) / (1 - F(t)) = h(t). \quad (6)$$

Note that f(t) is a PDF (the probability density function). Let $\{t_n, n = 1, 2, \dots\}$ denote the sequence of times between successive software failures. Then t_n denote the time between (n-1)st and nth failure. Let x_n denotes nth failure time, so that:

$$x_n = \sum_{k=1}^n t_k \quad (k = 1, 2, \dots, n; \quad 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \quad (7)$$

The joint density or the likelihood function of can be written as (Gokhale and Trivedi, 1999; Kuo and Yang, 1995):

$$f_{x_1, x_2, \dots, x_n}(x_1, x_2, \dots, x_n) = e^{-m(x_n)} \prod_{i=1}^n \lambda(x_i) \quad (8)$$

For a given sequence of software failure times (x_1, x_2, \dots, x_n) that are realizations of the random variables (X_1, X_2, \dots, X_n) , the parameters of the software reliability NHPP models are estimated using the Maximum Likelihood Method (MLE) (Gokhale and Trivedi, 1999; Kuo and Yang, 1995).

Weibull extension distribution: This distribution was proposed from Weibull extension model (Xie *et al.*, 2002), this model has bathtub shaped failure rate function and asymptotically related to the traditional Weibull distribution. The Probability Density Function (PDF) and distribution function is given by:

$$f(t) = \lambda \beta (\alpha t)^{\beta-1} \exp\left[(\alpha t)^\beta + \lambda \frac{1}{\alpha} (1 - e^{(\alpha t)^\beta}) \right] A \quad (9)$$

$$F(t) = 1 - \exp\left[-\lambda \frac{1}{\alpha} (e^{(\alpha t)^\beta} - 1) \right] \quad (10)$$

Note that $\alpha, \lambda, \beta > 0, t \geq 0$. The hazard function is derived as follows:

$$h(t) = \frac{f(t)}{1-F(t)} = \lambda \beta (\alpha t)^{\beta-1} \exp\left[(\alpha t)^\beta \right] \quad (11)$$

Flexible Weibull extension distribution: This distribution was proposed a new two-parameter ageing distribution

called the flexible Weibull extension distribution (Bebington *et al.*, 2007). This new distribution is shown to be quite flexible, being able to model both IFR and IFRA ageing classes. The cumulative distribution function of the flexible Weibull extension distribution is given as follows:

$$F(t) = 1 - e^{-e^{\alpha t - \frac{\beta}{t}}}, t > 0, \alpha, \beta > 0 \quad (12)$$

So, the corresponding density function is given in the form:

$$f(t) = \left(\alpha + \frac{\beta}{t^2} \right) e^{\alpha t - \frac{\beta}{t}} e^{-e^{\alpha t - \frac{\beta}{t}}}, t > 0, \alpha, \beta > 0 \quad (13)$$

Note that α and β is shape parameter $t \in (0, \infty)$.

Comparison criteria for the efficient model: To investigate the usefulness of the proposed model, the comparison criteria can be used. It can be described as follows (Gokhale and Trivedi, 1999; Chen *et al.*, 2008): the Mean Square Error (MSE) measures the deviation about between the predicted values with the actual observations. It is well-defined as:

$$MSE = \frac{\sum_{i=1}^n [m(x_i) - \hat{m}(x_i)]^2}{n-k} \quad (14)$$

Note that m_i is the total cumulated number of errors observed within time is $(0, t_i]$ and \hat{m}_i cumulative number of errors at time x_i obtained from the fitting mean value is the estimated n is the number of observations and k is the number of parameters to can be estimated. R^2 can measure how successful the fit is in explaining the variation of the data. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^n [m(x_i) - \hat{m}(x_i)]^2}{\sum_{i=1}^n \left(m(x_i) - \frac{\sum_{j=1}^n m(x_j)}{n} \right)^2} \quad (15)$$

MATERIALS AND METHODS

Software reliability infinite NHPP model: NHPP Model based on Weibull extension distribution model. The mean value function $m(t)$ and failure intensity function of $\lambda(t)$ Weibull extension distribution using generalized order statistics software reliability model based on Eq. 8 and 9 are derived as follows:

$$\lambda(t) = h(t) = \frac{f(t)}{1-F(t)} = \lambda \beta (\alpha t)^{\beta-1} \exp[(\alpha t)^\beta] \quad (16)$$

$$m(t) = -\ln(1-F(t)) = \frac{\lambda}{\alpha} (e^{(\alpha t)^\beta} - 1) \quad (17)$$

Thus, using Eq. 8 the log likelihood of Weibull extension model can be written as:

$$\ln L_{NHPP}(\lambda, \alpha, \beta | \underline{x}) = n \ln \lambda + n \ln \beta + (\beta - 1) \sum_{i=1}^n \ln(\alpha x_i) + \sum_{i=1}^n (\alpha x_i)^\beta - \frac{\lambda}{\alpha} (e^{(\alpha x_n)^\beta} - 1) \quad (18)$$

Note $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$. Maximizing Eq. 18 for a fixed value of β with respect to λ, α , next conditions must be satisfied:

$$\frac{\partial \ln L_{NHPP}(\lambda, \alpha | \underline{x})}{\partial \lambda} = \frac{n}{\lambda} - \frac{1}{\alpha} (e^{(\alpha x_n)^\beta} - 1) = 0 \quad (19)$$

$$\frac{\partial \ln L_{NHPP}(\lambda, \alpha | \underline{x})}{\partial \alpha} = (\beta - 1) \left(\frac{n}{\alpha} \right) + \beta \alpha^\beta \sum_{i=1}^n x_i^\beta + \frac{\lambda}{\alpha^2} (e^{\alpha x_n} - 1 - \alpha^\beta \beta x_n^\beta e^{(\alpha x_n)^\beta}) = 0 \quad (20)$$

The reliability is derived as follows:

$$\hat{R}(\delta | x_n) = \exp[-\{m(\delta + x_n) - m(x_n)\}] \quad (21)$$

Note that δ is the mission time:

$$m(x_n + \delta) = \frac{\lambda}{\alpha} (e^{[\alpha(\delta + x_n)]^\beta} - 1) \quad m(x_n) = \frac{\lambda}{\alpha} (e^{(\alpha x_n)^\beta} - 1)$$

NHPP model based on flexible weibull extension distribution: The mean value function $m(t)$ and failure intensity function $\lambda(t)$ of Weibull extension distribution software reliability model based on Eq. 11 and. 12 are derived as follows:

$$\lambda(t) = h(t) = \frac{f(t)}{1-F(t)} = \left(\alpha + \frac{\beta}{t^2} \right) e^{\alpha t - \frac{\beta}{t}} \quad (22)$$

$$m(t) = -\ln(1-F(t)) = e^{\alpha t - \frac{\beta}{t}} \quad (23)$$

Thus, using Eq. 8 the log likelihood of flexible Weibull extension model can be written as:

$$\ln L_{NHPP}(\alpha, \beta | \underline{x}) = \sum_{i=1}^n \ln \left(\alpha + \frac{\beta}{x_i^2} \right) + \alpha \sum_{i=1}^n x_i - \beta \sum_{i=1}^n \frac{1}{x_i} - e^{\alpha x_n - \frac{\beta}{x_n}} \quad (24)$$

Note $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$.

Maximizing Eq. 24 for a fixed value of β with respect to, α next conditions must be satisfied:

$$\frac{\partial \ln L_{NHPP}(\alpha, \beta | \underline{x})}{\partial \alpha} = \sum_{i=1}^n \ln \left(\frac{x_i^2}{\alpha x_i^2 + \beta} \right) + \sum_{i=1}^n x_i - x_n e^{\alpha x_n - \frac{\beta}{x_n}} = 0 \quad (25)$$

The reliability is derived as follows:

$$\hat{R}(\delta | x_n) = \exp[-\{m(\delta + x_n) - m(x_n)\}] \quad (26)$$

Note that δ is the mission time:

$$m(x_n + \delta) = e^{\alpha(x_n + \delta) - \frac{\beta}{(x_n + \delta)}} m(x_n) = e^{\alpha - \frac{\beta}{x_n}}$$

RESULTS AND DISCUSSION

We construct the corresponding software reliability growth model by using the date in Table 1 (Hayakawa and Telfar, 2000). In general, the Laplace trend test analysis is used (Kanoun and Laprie, 1996) for reliability property. As a result of this test in this Fig. 1 as indicated in the Laplace factor is between 2 and -2, reliability growth shows the properties. Thus, using this data, it is possible to estimate the reliability (Kanoun and Laprie, 1996). The

Table 1: Failure time data (Hayakawa and Telfar, 2000)

Failure No.	Failure time (hours)	Failure No.	Failure time (hours)
1	0.479	16	10.771
2	0.745	17	10.906
3	1.022	18	11.183
4	1.576	19	11.779
5	2.610	20	12.536
6	3.559	21	12.973
7	4.252	22	15.203
8	4.849	23	15.640
9	4.966	24	15.980
10	5.136	25	16.385
11	5.253	26	16.960
12	6.527	27	17.237
13	6.996	28	17.600
14	8.170	29	18.122
15	8.863	30	18.735

estimation of parameters for the each model can be done using the maximum likelihood method. In this study, the numerical conversion data (Failure time (hour)×0.1) to facilitate the parameter estimation was used. The results of parameter estimation were obtained in Table 2. These calculations, solving numerically, the initial values given to 0.001 and 3.0 and the tolerance value for the width of interval given 10^{-5} using C language checking adequate convergent, were performed iteration of 100 times.

The result of parameter estimation, Mean Square Error (MSE) and coefficient of determination R^2 are exhibited in Table 2. For the software model judgment in Table 2, MSE which measures the difference about between the actual value and the predicted value shows that the case of the Weibull extension model using generalized order statistics than the flexible Weibull extension model has a small value. Therefore, the case of the Weibull extension model is appreciably better than the flexible Weibull extension model. In addition, R^2 which means that the predictive power of the difference for between about the predicted values shows that the case of the Weibull extension model than the flexible Weibull extension model has high value. As was expected, a case of the Weibull extension model than the flexible Weibull extension model is the utility model. Eventually, in terms of a deviation for the between about the predicted values with the actual observations, the Weibull extension model using generalized order statistics regard as the efficient model. The result of mean value functions are was summarized in Fig. 2.

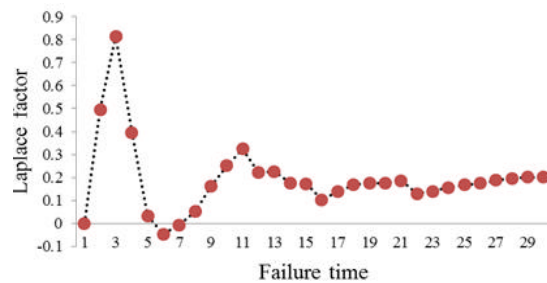


Fig. 1: The results of Laplace trend test

Table 2: MLE, MSE and R^2 for each model

Models	MLE	Model comparison	
		MSE	R^2
WEM	$\alpha = 0.5992$ $\beta = 16.9343$ $\lambda = 0.5$	253.8	0.98
FWEM	$\alpha = 1.7670$	111.6	0.85

WEM: Weibull Extension Model using generalized order statistics; FWEM: Flexible Weibull Extension Model; MLE: Maximum Likelihood Estimation; MSE: Mean Square Error; R^2 : coefficient of determination

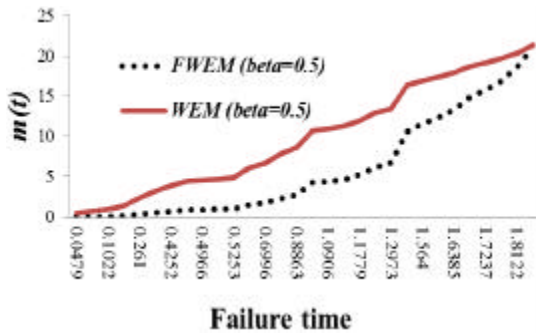


Fig. 2: Mean value function of each model

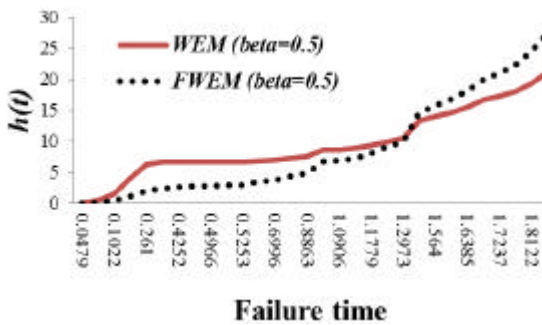


Fig. 3: Hazard function of the each model

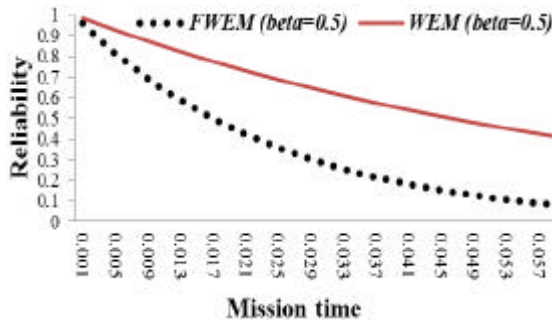


Fig. 4: Comparison reliability for each model

In Fig. 2, the patterns of mean value function have the tendency of the non-decreasing form. But, the result of intensity functions is summarized in Fig. 3. In Fig. 3, the case of the patterns of the intensity functions have the tendency non-decreasing form. In Fig. 4 in terms of the comparison of reliability, a case of the reliability for the assumed mission time shows that the case of the Weibull extension model than the flexible extension model is shown high reliability. Namely, the reliability has been sensitive to the mission time. As a result, the Weibull extension distribution than the flexible Weibull extension distribution model is judged more reliable model in this field.

CONCLUSION

In this study, the comparative problem for the reliability model with the Weibull extension distribution model and the flexible Weibull extension distribution model which made out efficiency for the software reliability was proposed.

In this study, the following conclusions were obtained. In terms of the deviation between the predicted values with the actual observations, the Weibull extension model regard as efficient model than the flexible Weibull extension distribution model. The result of a mean value functions has the tendency of the non-decreasing form. The result of intensity functions has the tendency increasing form. As a result, the Weibull extension distribution based on the generalized order statistics than the flexible Weibull extension distribution model is judged more reliable model in this field. An alternative study for this area will be valuable research.

ACKNOWLEDGEMENT

Funding for this study was provided by Namseoul University.

REFERENCES

Bebbington, M., C.D. Lai and R. Zitikis, 2007. A flexible Weibull extension. Reliab. Eng. Syst. Saf., 92: 719-726.

Chiu, K.C., Y.S. Huang and T.Z. Lee, 2008. A study of software reliability growth from the perspective of learning effects. Reliab. Eng. Syst. Safety, 93: 1410-1421.

Goel, A.L. and K. Okumoto, 1979. Time-dependent error-detection rate model for software reliability and other performance measures. Trans. Reliability, 28: 206-211.

Gokhale, S.S. and K.S. Trivedi, 1999. A time/structure based software reliability model. Ann. Software Eng., 8: 85-121.

Hayakawa, Y. and G. Telfar, 2000. Mixed poisson-type processes with application in software reliability. Math. Comput. Model., 31: 151-156.

Huang, C.Y., 2005. Performance analysis of software reliability growth models with testing-effort and change-point. J. Syst. Software, 76: 181-194.

Kanoun, K. and J.C. Laprie, 1996. Handbook of Software Reliability Engineering. McGraw-Hill, New York, USA., Pages: 437.

Kuo, L. and T. Yang, 1995. Bayesian computation of software reliability. J. Comput. Graphical Stat., 4: 65-82.

Pham, H. and X. Zhang, 2003. NHPP software reliability and cost models with testing coverage. Eur. J. Oper. Res., 145: 443-454.

- Shyur, H.J., 2003. A stochastic software reliability model with imperfect-debugging and change-point. *J. Syst. Software*, 66: 135-141.
- Xie, M., Y. Tang and T.N. Goh, 2002. A modified weibull extension with bathtub-shaped failure rate function. *Reliab. Eng. Syst. Saf.*, 76: 279-285.
- Yamada, S., M. Obha and S. Osaki, 1983. S-Shaped reliability growth modeling for software error detection. *IEEE Trans. Reliab.*, 12: 475-478.
- Zhao, M., 1993. Change-point problems in software and hardware reliability. *Commun. Stat. Theory Meth.*, 22: 757-768.