International Journal of Soft Computing 12 (5-6): 308-312, 2017
ISSN: 1816-9503

A Capability Comparative of Failure Predicting Time Using Software NHPP Finite Failure Reliability Model

Hong-Kyu Kwon and Hee-Cheul Kim
Department of Industrial and Management Engineering,
Namseoul University, Cheonan, Republic of Korea

Abstract: This study aims to analyze the predict capability of some of the popular software NHPP finite failure reliability models (Goel-Okumo Model, delayed S-formed reliability model and Rayleigh distribution model). The predict capability analysis was performed on two key factors, one pertaining to the degree of goodness of fit on applied failure time data and the other is comparison of predict capability. The estimation of parameters for the each model was used maximum likelihood estimation using first 80% of the applied failure time data. The comparison of predict capability of models was selected by validating against the remaining 20% of the applied failure time data. Through this study, the findings can be used as priori information for the administrator to analyze the failure time of the software.

Key words: NHPP, delayed S-formed reliability model, prediction of failure time, laplace trend test, capability, parametres

INTRODUCTION

In software development field, the software reliability can be defined as probability that can operate without the failure during the period. Therefore, software reliability is a key challenge in the software development course. These main points must be satisfied the user's necessities and must be reduced the testing costs. If know the variation in the reliability of the software in advance, it can be reduced the cost in terms of software testing. Therefore, the software development process that satisfies the reliability, cost and the release time can be required.

Until present, many software reliability models have been suggested. These models depend on Non-Homogenous Poisson Process (NHPP) (Goel and Okumoto, 1979; Gokhale and Trivedi, 1999) in terms of the error discovery process and if a fault occurs, immediately remove during the debugging process and has the assumption that no new fault has encountered. The generalized logistic testing-determination functions and the change-point parameter by incorporating efficient techniques to the 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 was explained (Chiu *et al*., 2008). This study aims to analyze the predictive ability of the software reliability NHPP model (Goel-Okumo Model, delayed S-formed type reliability model and Rayleigh distribution model).

Literature review

Goel-Okumoto software reliability model: The most elementary model in software reliability ground is Goel-Okumoto Model (Goel and Okumoto, 1979; Yamada *et al*., 1983). The Goel-Okumoto Model is the modest Non-Homogenous Poisson Process (NHPP) model with the mean value function (Goel and Okumoto, 1979). Note that the parameter $m(t|\theta, \beta_1) = \theta(1 - e^{-\beta_1 t})$ is the number of early faults in the software and the parameter $\lambda(t|\theta, \beta_1) = \theta\beta_1 e^{-\beta_1 t}$ is the fault detection degree. The matching failure intensity function is $f(t|\beta_1) = \beta_1 e^{-\beta_1 t}$. The probability density function of Goel-Okumoto Model requires the form $f(t|\beta_1) = \beta_1 e^{-\beta_1 t}$. The matching cumulative distribution function is $F(t|\beta_1) = 1 - e^{-\beta_1 t}$. Note that $\beta_1(>0)$ is the shaping parameter $t\in(0, \infty)$. Using $f(t|\beta_1) = \beta_1 e^{-\beta_1 t}$ and $F(t|\beta_1) = 1 - e^{-\beta_1 t}$, the hazard function (Kim and Kim, 2014) can be resulted next equation:

$$h(t) = \frac{f(t)}{1 - F(t)} = \beta_1 \quad (1)$$

The likelihood function (Gokhale and Trivedi, 1999) was well-known to next form:

$$L_{NHPP}(\Theta|\underline{x}) = \prod_{i=1}^{n} \lambda(x_i)\, e^{-m(x_n)} = \left[\prod_{i=1}^{n} \theta\beta e^{-\beta x_i}\right] \exp\left[-\theta(1 - e^{-\beta x_n})\right] \quad (2)$$

Using the expressions (Eq. 2) $\hat{\theta}_{MLE}$ and $\hat{\beta}_{1MLE}$ must be contented the following equations for the maximum likelihood approximation of the individually parameter:

Corresponding Author: Hong-Kyu Kwon, Department of Industrial and Management Engineering, Namseoul University, Cheonan, Republic of Korea

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\theta} = \frac{n}{\theta} - 1 + e^{-\beta_1 x_n} = 0 \quad (3)$$

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\beta_1} = \frac{n}{\beta_1} - \sum_{i=1}^{n} x_i - \theta x_n e^{-\beta_1 x_n} = 0 \quad (4)$$

Where:

$\underline{x}$ = $x_1, x_2, x_3, \ldots, x_n$

Θ = The parameter space

The expression (Eq. 3 and 4) can be summarized as follows:

$$\hat{\theta} = \frac{n}{1-e^{-\hat{\beta}_1 x_n}}, \ \frac{n}{\hat{\beta}_1} = \sum_{i=1}^{n} x_i + \hat{\theta} x_n e^{-\hat{\beta}_1 x_n} \quad (5)$$

MATERIALS AND METHODS

Delayed S-formed software reliability model: In finite NHPP Model of the delayed S-formed reliability growth model (Yamada *et al*., 1983), the mean value function and intensity function were well-known to next shapes:

$$\lambda(t|\theta, \beta_2) = \theta F'(t) = \theta f(t) = \theta\beta_2^2 t\, e^{-\beta_2 t} \quad (6)$$

$$m(t|\theta, \beta_2) = \theta F(t) = \theta\left[1-(1+\beta_2 t)e^{-\beta_2 t}\right] \quad (7)$$

Where:

θ = The number of early faults in the software

β_2 = Shaping parameter

Similarly, by means of Eq. 6 and 7, the hazard function expression can be resulting as follows:

$$h(t) = \frac{f(t)}{1-F(t)} = \frac{F'(t)}{1-F(t)} = \frac{\beta_2^2}{1+\beta_2 t} \quad (8)$$

Also, using the Eq. 6 and 7, the likelihood function can be derived next form (Kim, 2013, 2015):

$$L_{NHPP}(\Theta|\underline{x}) = e^{-m(x_n)} \prod_{i=1}^{n} \lambda(x_i) = \left[\prod_{i=1}^{n} \theta\beta_2^2 x_i e^{-\beta_2 x_i}\right] \times \exp\left(-\theta\left[(1-(1+\beta_2 x_n)e^{-\beta_2 x_n}\right]\right) \quad (9)$$

The estimator $\hat{\theta}_{MLE}$ and $\hat{\beta}_{2MLE}$ must be satisfied the following equations for the maximum likelihood estimation of the individually parameter by means of Eq. 9:

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\theta} = \frac{n}{\theta} - 1 + e^{-\beta_2 x_n} + \beta_2 x_n e^{-\beta_2 x_n} = 0 \quad (10)$$

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\beta_2} = \frac{2n}{\beta_2} - \theta\beta_2 x_n^2 e^{-\beta_2 x_n} - \sum_{i=1}^{n} x_i = 0 \quad (11)$$

Equation 10 and 11 can be summarized as follows:

$$\hat{\theta} = \frac{n}{1-e^{-\hat{\beta}_2 x_n}(1+\hat{\beta}_2 x_n)} \quad (12)$$

$$\frac{2n}{\hat{\beta}_2} = \theta\hat{\beta}_2 x_n^2 e^{-\hat{\beta}_2 x_n} + \sum_{i=1}^{n} x_i \quad (13)$$

Rayleigh distribution software reliability model: In probability and statistics, the Rayleigh distribution (Shin and Kim, 2014; Sheldon, 2000) is a continuous probability distribution. This distribution is a singular form (shaping parameter is 2) of Weibull distribution.

The Rayleigh distribution was known as a model for the distribution of life test and reliability concept in the ground of software reliability. The probability density function and the distribution function for Rayleigh distribution known to next forms:

$$f(t|\beta_3) = 2\beta_3 t \exp(-\beta_3 t^2) \quad (14)$$

$$F(t|\beta_3) = 1 - \exp(-\beta_3 t^2) \quad (15)$$

Note that $\beta_3(>0)$ is scaling parameter $t\epsilon(0, \infty)$. Using the Eq. 14 and 15, the hazard function can be derived next expression:

$$h(t) = \frac{f(t)}{1-F(t)} = 2\beta_3 t \quad (16)$$

Thus, the mean value function and intensity function of the finite failure NHPP model can be uttered as follows using the Eq. 14 and 15:

$$m(t) = \theta F(t|\beta_3 = \theta\left[1-\exp(-\beta_3 t^2)\right] \quad (17)$$

$$\lambda(t) = \theta f(t|\beta_3) = 2\theta\beta_3 t \exp(-\beta_3 t^2) \quad (18)$$

In this case, the likelihood functions can be expressed as follows (Kim and Kim, 2014):

$$L_{NHPP}(\Theta|\underline{x}) = \left[\prod_{i=1}^{n} 2\theta\beta_3 x_i e^{-\beta_3 x_i^2}\right] \times \exp\left[-\theta(1-e^{-\beta_3 x_n^2})\right] \quad (19)$$

Using the expressions (Eq. 19) $\hat{\theta}_{MLE}$ and $\hat{\beta}_{3MLE}$ must be satisfied the following equations for the maximum likelihood estimation of the each parameter:

$$\hat{\theta} = \frac{n}{1-e^{-\hat{\beta}_3 x_n^2}} \quad (20)$$

Table 1: Failure time data (Prasad *et al.*, 2011)

Failure number	Failure time (h)	Failure number	Failure time (h)
1	30.02	16	151.78
2	31.46	17	177.50
3	53.93	18	180.29
4	55.29	19	182.21
5	58.72	20	186.34
6	71.92	21	256.81
7	77.07	22	273.88
8	80.90	23	277.87
9	101.90	24	453.93
10	114.87	25	535.00
11	115.34	26	537.27
12	121.57	27	552.90
13	124.97	28	673.68
14	134.07	29	704.49
15	136.25	30	738.68

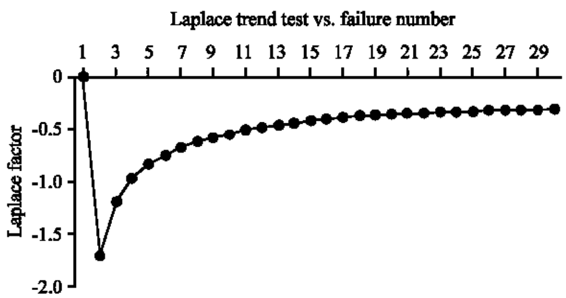

Fig. 1: Results of Laplace trend test

$$\frac{n}{\hat{\beta}_3} = \sum\nolimits_{i=1}^{n} x_i^2 + \hat{\theta} x_n^2 e^{-\hat{\beta}_3 x_n^2} \qquad (21)$$

RESULTS AND DISCUSSION

Data analysis and forecasting for software failure time: In this study, the characteristics of NHPP finite failure reliability models (Goel-Okumo Model, delayed S-formed reliability model and Rayleigh distribution model) were analyzed using software failures time data. This data (Prasad *et al.*, 2011) was listed in Table 1. The trend test distribution (Kim, 2015) must be offered about the data for the faith models. Thus, the Laplace trend test analysis is used in this study. As the Laplace factor is specified in between 2.0 and -2.0 in Fig. 1, the reliability evolution shows the properties. Thus, it is likely to estimate the reliability using this data.

Predicting approach: Predictive validity process was composed of the following process (Williams, 2006).

Step 1: Using first 80% of the application failure data, it shall be estimated the parameter for the each model. Thus, the analysis first 80% (30×0.8 = 24) of the application failure data for the Goel-Okumo Model, delayed S-formed reliability model and Rayleigh distribution model was performed.

Table 2: Parameter estimation (MLE), MSE and R^2 for the each model

		Model comparison	
Models	MLE	MSE	R^2
Goel Okumoto	$\hat{\beta}_{1\,MLE} = 7.086\times10^{-1}$, $\hat{\theta}_{MLE} = 25.0026$	7.2647	0.9784
Delayed S-shaped reliability	$\hat{\beta}_{2\,MLE} = 16.074\times10$, $\hat{\theta}_{MLE} = 24.1357$	2.0932	0.9845
Rayleigh distribution	$\hat{\beta}_{3\,MLE} = 3.707\times10^{-1}$, $\hat{\theta}_{MLE} = 24.0116$	2.1878	0.9851

MLE: Maximum Likelihood Estimation; MSE: Mean Square Error; R^2: Coefficient of determination

Step 2: From the remaining 20% (30×0.2 = 6) of the application failure data, the failure times were predicted for the each model.

Step 3: The comparative analysis to predict the performance difference was carried out using the Mean Square Error (MSE) and coefficient of determination by means of the predicted time using result of Step 2. The Mean Square Error (MSE) and R square (R^2) are defined as Kim (2013):

$$MSE = \frac{\sum_{i=1}^{n}[m(x_i) - \hat{m}(x_i)]^2}{n-k} \qquad (22)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}[m(x_i) - \hat{m}(x_i)]^2}{\sum_{i=1}^{n}\left(m(x_i) - \sum_{j=1}^{n} m(x_j)/n\right)^2} \qquad (23)$$

Where:

- $m(x_i)$ = The over-all cumulated number of the errors can be observed
- $\hat{m}(x_i)$ = The estimated over-all cumulated number of the errors within time is $(0, x_i)$
- n = The number of observations
- k = The number of parameters to can be estimated

Using the method of maximum likelihood estimation and bisection method for nonlinear equations, the parameter estimation values for first 80% of the application failure data summarized in Table 2.

In this study, the numerical translation data (Failure time (h)$\times10^{-2}$) to simplify the parameter estimation was used. A consequence of the parameter approximation was summarized in Table 2. These calculations, solving numerically, the initial values given to 0.001 and 5.0 and tolerance value for the width of an interval was given to 10^{-5} using C-language checking tolerable convergent were accomplished iteration of 100 times.

Also, the estimation values for the Mean Square Error (MSE) and coefficient of determination (R^2) are

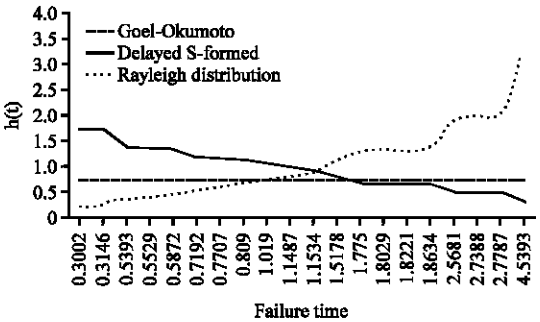

Fig. 2: Type of the hazard function

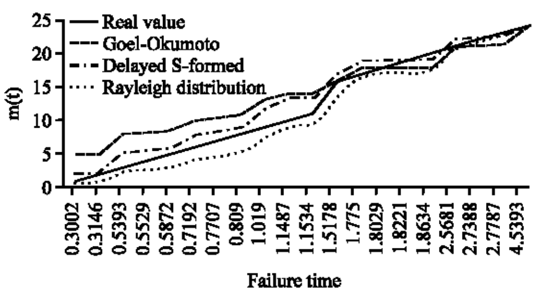

Fig. 3: Types of the mean value function

summarized in Table 2. These estimation values may be utilized as a basis for investigating the efficiency of the model. From the first 80% of the application failure data, the delayed S-formed model may appear smaller than the Rayleigh Model and the Goel-Okumoto Model. Thus, delayed S-formed than others can be considered as efficient model.

In terms of R^2, Rayleigh Model may appear significantly. That is Rayleigh Model than other models appears (for the difference between the predicted values) the highly prognostic power. Thus, Rayleigh Model than other models can be regarded as well-organized model.

The hazard functions for each model are summarized in Fig. 2. In this Fig. 2, the case of Goel-Okumoto Model is independent of the time t as having constant pattern. The case of delayed S-formed model have non-increasing pattern. On the other hand, Rayleigh Model shows non-decreasing pattern.

Figure 3 shows the mean value function from the first 80% of the application failure data. A similar pattern was seen in all models in this figure. In terms of the compare for the actual value, delayed S-formed model and Goel-Okumoto Model show the overestimation. But, Rayleigh Model appears underestimation.

The estimation values for the Mean Square Error (MSE) and coefficient of determination (R^2) using the selected data (the remaining 20% of the application failure

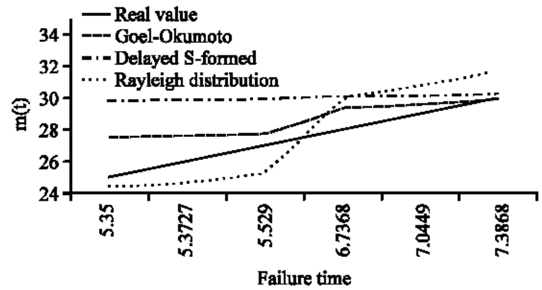

Fig. 4: Types of the mean value function of predicted value

Table 3: Estimation of MSE, R^2 for the each model

	Model comparison	
	MSE	R^2
Goel-Okumoto	3.0449	0.8955
Delayed S-shaped reliability	13.2221	0.8847
Rayleigh distribution	3.9414	0.8938

MSE: Mean Square Error; R^2: Coefficient of determination

data, 30×0.2 = 6) for each model was provided in Table 3. In Table 3, for MSE value, the Goel-Okumoto Model appears to be smaller than other models and value of R^2 appears to be higher than other models using the selected data (the remaining 20%). Thus, Goel-Okumoto Model can be described as relatively efficient model than other models.

The pattern of the predicted mean value function of the predicted value was summarized in Fig. 4 using the selected data (the remaining 20%). In terms of the compare for the actual value, delayed S-formed model and Goel-Okumoto Model show the overestimation. But, Rayleigh Model shows non-decreasing pattern (S-type) over failure time.

CONCLUSION

Software which includes a lot of big data that cannot avoid the occurrence of defects in the process that changes and modifications in the course of execution occurs frequently environment.

After all, using the failure information in the test process and the actual use phase of software development, environmental failure, the model may be evaluated for efficiency to its function.

By testing or run-time, efficient management of the relationship with the number of malfunctioning and failure may be enhanced for the software reliability. This process was known as software growth property.

This study aims to analyze the predictive ability of the software reliability model NHPP (Goel-Okumo Model, delayed S-formed type reliability model and Rayleigh distribution model.

From the first 80% of the application data in terms of mean square error, the delayed S-formed model may appear smaller than the Rayleigh Model and the Goel-Okumoto Model. Thus, delayed S-formed than others can be considered as efficient model.

In terms of the coefficient of determination, Rayleigh Model may appear significantly. That is a highly prognostic power to appear for the difference between the predicted values, Rayleigh Model than other models can be regarded as well-organized model.

Using the remaining 20% of the application data, the mean square error of the Goel-Okumoto Model appears to be smaller than other models and coefficient of determination appears to be higher than other models. Thus, Goel-Okumoto Model can be described as relatively well-organized model than other models.

The study by software designers or managers are assuming diets can be used as guidance to determine the basis of prior information to understand the existing knowledge of the software fault type by seeing the case of multiple software environments downtime.

REFERENCES

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 measure. IEEE Trans. Reliab., R-28: 206-211.

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

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

Kim, H.C. and J.B. Kim, 2014. The comparative study of generalized goel-NHPP model using software reliability cost model. Intl. J. Appl. Eng. Res., 9: 25471-25477.

Kim, H.C., 2013. The comparative study of NHPP delayed S-shaped and extreme value distribution software reliability model using the perspective of learning effects. Int. J. Adv. Comput. Technol., 5: 1210-1218.

Kim, H.C., 2015. The property of learning effect based on delayed software S-Shaped reliability model using Finite NHPP Software cost model. Indian J. Sci. Technol., 8: 1-7.

Prasad, R.S., K.R.H. Rao and R.R.L. Kantha, 2011. Software reliability measuring using modified maximum likelihood estimation and SPC. Int. J. Comput. Applic., 21: 1-5.

Sheldon, M.R., 2000. Introduction to Probability and Statistics for Engineers and Scientists. Academic Press, Cambridge, Massachusetts, USA., ISBN:9780125984737, Pages: 578.

Shin, H. D. and H.C. Kim, 2014. Assessing rayleigh software reliability model based on NHPP using SPC. J. Next Gen. Inform. Technol., 5: 77-84.

Williams, D.R.P., 2006. Prediction capability analysis of two and three parameters software reliability growth models. Inform. Technol. J., 5: 1048-1052.

Yamada, S., M. Obha and S. Osaki, 1983. S-Shaped reliability growth modeling for software error detection. IEEE Trans. Reliab., 12: 475-478.