SRGM using Testing-Effort Function with Uncertainty in Operating Environment

Ramgopal Dhaka¹, Bhoopendra Pachauri¹ and Anamika Jain¹

¹Department of Mathematics & Statistics, Manipal University Jaipur, Jaipur, Rajasthan, India

E-mail: ramgopaldhaka605@gmail.com

Abstract: With increasing pace of technological advancement and new tech introduction in today’s word, reliability of the software has become vital. For software reliability assessment, many software reliability growth models (SRGMs) have been discussed. In literature, many of the existing SRGMs have considered uncertainty of the operating environment but very little work has included testing effort function (TEF) with uncertainty in operating environment. In this research, an SRGM incorporating Gompertz TEF has been investigated with the uncertainty of operating environment. Software testing environment is usually a controlled one with variables known to the developer, but operating environment may introduce uncontrolled and unknown variables. This model has considered a constant fault detection rate in perfect debugging environment. Further, sensitivity analysis has been done. The numerical results have been compared with existing SRGMs.

1. Introduction

Software has become integral to the modern way of life in form of mobile phones, computers, smart wearables, smart house-hold appliances, introduction of Internet of Things (IoT) to every aspect of industry. With this growing demand and rapid changes to software operating environment, a reliable software that is easier to use, is of lower cost and can work in different environments has become very important. The reliability of a software means, the probability that the system will perform without failure under certain conditions for a specific period. Software reliability also needs improvement with the operating environment and technology. However, it is a very complicated process and the existing SRGMs are not sufficient for handling today’s complex situations.

In published literature, there are two types of NHPP based SRGMs, perfect debugging (PD) models and imperfect debugging (ID) models. In perfect debugging SRGMs, removing the detected defect will not cause new defects in the further testing. Whereas, in imperfect debugging SRGMs, a new defect can be introduced, when a detected defect is removed. Further various models have been introduced based on perfect and imperfect debugging environment. Such as SRGM with fault detection rate, SRGM with testing effort function, SRGM considering testing coverage factor, SRGM with considering uncertainty in operating environments, one-dimensional and two-dimensional SRGMs, SRGM using time lag in fault removal process, etc.

The first NHPP based SRGM in PD environment was developed by Goel and Okumoto [1]. This model is also called exponential NHPP model. A large number of SRGMs were also studied in perfect debugging environment. Yamada et al. [2] introduced an SRGM in which fault detection rate function is S-shaped. A few researchers, Pham, Lee et al., Li and Pham, and Song et al. [3][4][5][6][7][8][9][10]
discussed SRGMs with different fault-detection rate factors. Chang et al. [11] have given an SRGM with testing-coverage function. Pachauri et al. [12] discussed an inflection S-shaped curve and considered fault reduction factor in a PD and ID environment. A two-dimensional multi-release SRGM developed by Kapur et al. [13]. Some models were also studied using testing effort function (TEF).

The SRGM with exponential TEF was developed by Yamada et al. [14]. Pachauri et al. [15] studied a TEF based SRGM with optimal release time and cost-reliability in imperfect debugging environment was calculated. Jin and Jin [16] introduced a S-shaped TEF based SRGM and improved swarm intelligent optimization was used for parameter optimization. Pachauri et al. [17] gave an SRGM with GMW TEF, the software release time with optimum cost computed using multi-attribute utility theory as well as genetic algorithm. Rafi and Akthar [18] discussed an SRGM using Gompertz TEF and software release time determined by enhancing testing efficiency. Jain et al. [19] developed an SRGM with GMW TEF and fault reduction factor.

Recently, many SRGMs have been discussed considering the uncertain environment factor [3][20]. Song et al. [10] introduced a SRGM with testing coverage factor in perfect debugging. Li and Pham [20] discussed an SRGM considering the testing coverage in imperfect debugging environment, they also analyzed optimal release time and sensitivity analysis. Song et al. [6] introduced a SRGM using three-parameter fault-detection. These are some other software reliability models [21][22][23][24][25][26] that do not involve uncertainty in environment factor.

In this paper, a SRGM is introduced with Gompertz TEF. The uncertainty in operating environment also has been considered. The calculation of mean value function (MVF) for an SRGM with mathematical derivation is shown in Section 2. Results through numerical examples are discussed in Section 3 and concluding remark is in section 4.

2. Model Formulation

Here, we considered the Gompertz TEF with perfect debugging and uncertainty in operating environment. The assumptions are being made for this model [3][18] as,

a. The software failure phenomena follow NHPP.

b. The effect of uncertainty factor is represented by product of fault detection rate \( b(t) \) and a random variable \( \eta \).

c. The \( b(t) \) is proportional to the remaining faults in the system at any time.

d. A software failure can occur during operation, caused by unresolved faults in the system.

In the model of Pham [3], the rate of change in cumulative faults at time \( t \) with uncertainty factor is:

\[
\frac{d}{dt} m(t) = \eta b(t) [N - m(t)], \tag{1}
\]

where \( m(t) \) represents collective number of faults at time \( t \), \( \eta \) is assume to have the PDF of gamma distribution with parameter \( \alpha \geq 0 \) and \( \beta \geq 0 \), and \( N \) is the total number of faults in the system [3]. The mean value function (MVF) based on above differential equation is:

\[
m(t) = \int_{\eta} N \left( 1 - e^{-\eta \int_{0}^{t} b(x) \, dx} \right) \, dg(\eta), \tag{2}
\]

then the MVF, \( m(t) \) after applying the random variable \( \eta \) in the differential equation (1) is,

\[
m(t) = N \left( 1 - \frac{\beta}{\beta + \int_{0}^{t} b(s) \, ds} \right)^{\alpha}. \tag{3}
\]

A SRGM with TEF and uncertainty of operating environment in a perfect debugging,
\[
\frac{d m(t)}{d(t)} = \eta [w(t) \times b(t)][N - m(t)],
\]
(4)

where \(w(t)\) is a TEF. Then the solution of differential equation is,
\[
m(t) = N \left( 1 - \frac{\beta}{\beta + \int_0^t w(s) \times b(s) ds} \right)^\alpha.
\]
(5)

Where the distribution function of Gompertz TEF [18,19] is defined as,
\[
W(t) = a(e^{-\gamma e^{-ct}}),
\]
(6)

where \(c\) is growth pattern indicator, \(a\) is the total effort expenses and \(\gamma\) is a scale parameter. The current testing effort in time \((0,1]\) is,
\[
\frac{d W(t)}{dt} = w(t) = ay(c - \gamma e^{-ct}),
\]
(7)

when \(b(t) = b\) then, the MVF is,
\[
m(t) = N \left( 1 - \frac{\beta}{\beta + ba(e^{-\gamma e^{-ct}})} \right)^\alpha.
\]
(8)

In the next section, parameter estimation with numerical results is discussed.

3. Numerical Description
A few of the existing models have used the maximum likelihood estimation technique parameter estimations. Here, in this paper, a nonlinear least square estimation (LSE) technique is used. We have taken two historical data sets to justify the performance and compared the introduced model with existing models. The summary of data sets (DS) is shown in Table 1. To justify the results, we used the mean square error (MSE). A relative error (RE) curve is used as a measure of precision. The estimated parameter values for this TEF are as shown in Table 2.

| Data set | Time (t) (weeks) | Testing CPU hours | No. of faults | Description | References |
|----------|------------------|-------------------|---------------|-------------|------------|
| DS1      | 19               | 10,272            | 120           | Tandem computer software data project | [25]       |
| DS2      | 12               | 5053              | 61            | Tandem computer software data project | [25]       |
Table 2. Estimated Parameters of TEF

| Datasets | Parameters |
|----------|------------|
| DS1      | $\hat{a} = 11620, \hat{\gamma} = 3.78, \hat{c} = 0.1886$ |
| DS2      | $\hat{a} = 6314, \hat{\gamma} = 4.874, \hat{c} = 0.2695$ |

For DS1, the results with comparison are shown in Table 3. From the Table 3, this model performs better in term of MSE that is 4.5893. The curves of mean value function of all models are shown in Figure 1 and results closer to the actual values. From the Figure 2, the relative error of the proposed model is closer to zero.

Table 3. Comparative Study with Existing Models for DS1

| No. | Model                  | Estimated value                        | MSE       |
|-----|------------------------|----------------------------------------|-----------|
| 1   | GO [1]                 | $\hat{a} = 183, \hat{b} = 0.0615$     | 26.002    |
| 2   | Y-DS [2]               | $\hat{a} = 127.4, \hat{b} = 0.2417$  | 14.6880   |
| 3   | O-IS [26]              | $\hat{a} = 124.4, \hat{b} = 0.2535, \hat{\beta} = 3.779$ | 7.1268   |
| 4   | K-SRGM 3[21]           | $\hat{A} = 1.858, \hat{\beta} = 3.791, \hat{b} = 2.413, \hat{a} = 0.9873$ | 18.550   |
| 5   | R-M-D [22]             | $\hat{a} = 98.13, \hat{\beta} = 1.353, \hat{b} = 0.215, \hat{\beta} = 0.1835$ | 13.4322  |
| 6   | C-TC [11]              | $\hat{a} = 0.0787, \hat{\alpha} = 46.24, \hat{\beta} = 26.63, \hat{\beta} = 1.474, \hat{N} = 125.7$ | 13.2656  |
| 7   | P-Vtub [3]             | $\hat{N} = 118.7, \hat{\alpha} = 60.55, \hat{\beta} = 0.0246, \hat{\beta} = 8.282, \hat{b} = 1.157$ | 65.594   |
| 8   | S-3PFD [6]             | $\hat{a} = 0.673, \hat{c} = 39.4, \hat{\beta} = 0.3264, \hat{b} = 0.2555, \hat{N} = 130.6$ | 8.2599   |
| 9   | Proposed model         | $\hat{a} = 0.9151, \hat{\beta} = 11.11, \hat{b} = 0.0004469, \hat{N} = 368.2$ | 4.5893   |

Figure 1. MVF Curve for Various Models for DS1
Figure 2. RE Curve for All Models for DS1

For DS2, estimated parameter values and MSE values of models are given in Table 4 and the value of MSE = 5.7938 of proposed model is less than the others. The curves of mean value function of all models are shown in Figure 3, which shows the better performance of the proposed model near to actual data. From Figure 4, relative error of the model is again near to zero.

Based on given results, it can be said that the model performs better for both data sets.

| No. | Model               | Estimated value                      | MSE   |
|-----|--------------------|--------------------------------------|-------|
| 1   | GO [1]             | $\hat{a} = 244.3, \hat{b} = 0.02651$ | 20.894|
| 2   | Y-DS [2]           | $\hat{a} = 76.25, \hat{b} = 0.2741$ | 9.897 |
| 3   | O-IS [26]          | $\hat{a} = 64.4, \hat{b} = 0.4832, \hat{\beta} = 11.36$ | 6.345 |
| 4   | K-SRGM 3[21]       | $\hat{A} = 3.057, \hat{\rho} = 1.942, \hat{b} = 2.097, \hat{a} = 0.966$ | 15.296|
| 5   | R-M-D [22]         | $\hat{a} = 59.05, \hat{\alpha} = 1.353, \hat{\beta} = 0.2526, \hat{b} = 0.1996$ | 20.512|
| 6   | C-TC [11]          | $\hat{a} = 0.04391, \hat{a} = 1070, \hat{\beta} = 94.38, \hat{b} = 1.921, \hat{N} = 64.69$ | 10.738|
| 7   | P-Vtub [3]         | $\hat{a} = 1.923, \hat{\alpha} = 59.87, \hat{\beta} = 852.4, \hat{b} = 0.782, \hat{N} = 62.07$ | 7.4792|
| 8   | S-3PFD [6]         | $\hat{a} = 29.31, \hat{c} = 228.1 \hat{\beta} = 2.857, \hat{b} = 0.4653, \hat{N} = 64.48$ | 9.4741|
| 9   | Proposed model     | $\hat{a} = 0.9475, \hat{\beta} = 2.665, \hat{b} = 0.0003012, \hat{N} = 162$ | 5.7938|
3.1. Sensitivity Analysis
Here, we perform a sensitivity analysis (SA) to evaluate the effect of individual parameter on robustness of the mean value function. Sensitivity analysis is performed keeping one parameter variable, whereas the remaining parameters are fixed [10]. We investigate, how the MVF alters for the parameter values estimated from DS1. Sensitivity is defined as follows,

\[ S_{\theta,P} = \frac{m(P + \theta P) - m(P)}{m(P)} \]

where \( m \) represents the mean value function, \( P \) refers to the parameter of the model and \( \theta \) is the comparative difference in parameter. When \( P \) is changed by 100\( \theta \)% then \( S_{\theta,P} \) is the relative change in the MVF. Similarly, \( S_{\theta,a} \), \( S_{\theta,b} \), \( S_{\theta,y} \), \( S_{\theta,n} \) and \( S_{\theta,N} \) can be found in same manner. Table 5, Figure 5 - 11, and Figure 12 show that the changes in MVF with alteration in each parameter.
Table 5. SA of Parameter of the Proposed Model

| Parameter | -20% | -10% | 0     | 10%  | 20%  |
|-----------|------|------|-------|------|------|
| α         | 0.3346 | 0.1546 | 0     | -0.1319 | -0.2452 |
| β         | 0.1661 | 0.0763 | 0     | -0.0657 | -0.1230 |
| γ         | 0.0895 | 0.0428 | 0     | -0.0393 | -0.0754 |
| a         | -0.1489 | -0.0725 | 0     | 0.0689  | 0.1345  |
| b         | -0.1489 | -0.0725 | 0     | 0.0689  | 0.1345  |
| c         | -0.1396 | -0.0649 | 0     | 0.0565  | 0.1056  |
| N         | -0.2000 | -0.1000 | 0     | 0.1000  | 0.2000  |

Figure 5. SA of the Model for Parameter $\alpha$

Figure 6. SA of the Model for Parameter $\beta$
**Figure 7.** SA of the Model for Parameter $\gamma$

**Figure 8.** SA of the Model for Parameter $a$

**Figure 9.** SA of the model for Parameter $b$
Figure 10. SA of the Model for Parameter $c$

Figure 11. SA of the Model for Parameter $N$
4. Conclusion
In this article, a SRGM with Gompertz TEF by considering uncertainty in operating environment has been investigated in perfect debugging environment. Sensitivity analysis has also been done to show the changes in MVF with respect to change in every other parameter of the proposed model. Comparison of proposed model with various SRGMs has been done using real data sets and demonstrated that the proposed model has improved performs. MSE criteria has been used to compare the results and non-linear least square estimation (LSE) was used for parameter estimation. The results of the proposed model have improved in terms of mean square error and relative error. In future, the proposed models can be extended to a two-dimensional model and results may be improved using soft-computing techniques. Some new factors such as fault reduction factor, change point, etc. may be introduced with uncertainty.

5. References
[1] Goel A L and Okumoto K 1979 Time-dependent error-detection rate model for software reliability and other performance measures. IEEE Trans. Reliab. 28 206–211.
[2] Yamada S, Ohba M and Osaki S 1983 S-shaped reliability growth modeling for software fault detection. IEEE Trans. Reliab. 32 475–484.
[3] Pham H 2014 A new software reliability model with Vtub-Shaped fault detection rate and the uncertainty of operating environments. Optimization 63 1481–1490.
[4] Lee D H, Chang I H, Pham H and Song K Y 2018 A software reliability model considering the syntax error in uncertainty environments, optimal release time and sensitivity analysis, Appl. Sci. 8 1483, doi: 10.3390/app8091483.
[5] Li Q and Pham H 2019 A generalized software reliability growth model with consideration of the uncertainty of operating environments, IEEE Access 7 84253–84267.
[6] Song K Y, Chang I H and Pham H 2017 A Three-parameter fault-detection software reliability model with the uncertainty of operating environments. J. Syst. Sci. Syst. Eng. 26 121–132.
[7] Song K Y, Chang I H and Pham H 2019 NHPP software reliability model with inflection factor of the fault detection rate considering the uncertainty of software operating environments and predictive analysis. Symmetry 11 521, doi:10.3390/sym11040521.
[8] Song K Y, Chang I H and Pham H 2018 Optimal release time and sensitivity analysis using a new NHPP software reliability model with probability of fault removal subject to operating environments. Appl. Sci. 8 714, doi: 10.3390/app8050714.
[9] Song K Y, Chang I H and Pham H 2017 A software reliability model with a Weibull fault detection rate function subject to operating environments. *Appl. Sci.* 7 983, doi: 10.3390/app7100983.

[10] Song K Y, Chang I H and Pham H 2019 A testing coverage model based on NHPP software reliability considering the software operating environment and sensitivity analysis, *Mathematics* 7 450, doi:10.3390/math7050450.

[11] Chang I H, Pham H, Lee S W and Song K Y 2014 A testing-coverge software reliability model with the uncertainty of operation environments. *Int. J. Syst. Sci. Oper. Logist.* 1 220–227.

[12] Pachauri B, Kumar A and Dhar J 2015 Incorporating inflection S-shaped fault reduction factor to enhance software reliability growth. *Appl. Math. Model* 39(5) 1463–1469.

[13] Kapur P K, Aggarwal A G and Kaur G 2012 Two dimensional multi-release software reliability modeling and optimal release planning. *IEEE Trans. Reliab.* 61 1–11.

[14] Yamada S, Ohtera H and Narihisa H 1986 Software reliability growth models with testing-effort. *IEEE Trans. Reliab.* 35(1) 19-23. doi: 10.1109/TR.1986.4335332.

[15] Pachauri B, Kumar A and Dhar J 2013 Modeling optimal release policy under fuzzy paradigm in imperfect debugging environment. *Inf. Soft. Tech.* 55(11) 1974–1980.

[16] Jin C, Jin S 2016 Parameter optimization of software reliability growth model with S-shaped testing-effort function using improved swarm intelligent optimization. *App. Soft. Comp.* 40 283-291.

[17] Pachauri B, Kumar A and Dhar J 2014 Software reliability growth modeling with dynamic faults and release time optimization using ga and maut. *Appl. Math. Comput.* 242 500–509.

[18] Rafi S M and Akhtar S 2010 Software reliability growth model with Gompertz TEF and optimal release time determination by improving the test efficiency. *Int. J. of Comp. App.* 7 34–43.

[19] Jain M, Agarwal P and Solanki R 2020 NHPP-based SRGM using time-dependent fault reduction factors (FRF) and Gompertz TEF. In: Kapur P., Singh G., Klochkov Y., Kumar U. (eds) Decision Analytics Applications in Industry. Asset Analytics (Performance and Safety Management). Springer, Singapore. https://doi.org/10.1007/978-981-15-3643-4_6.

[20] LI Q and Pham H 2017 NHPP software reliability model considering the uncertainty of operating environments with imperfect debugging and testing coverage, *Appl. math. Model.* 51 68–85.

[21] Kapur P K, Pham H, Anand S and Yadav K 2011 A unified approach for developing software reliability growth models in the presence of imperfect debugging and error generation. *IEEE Trans. Reliab.* 60 331–340.

[22] Roy P, Mahapatra G S and Dey K N 2014 An NHPP software reliability growth model with imperfect debugging and error generation. *Int. J. Reliab. Qual. Saf. Eng.* 21 1–3.

[23] Pham H. 2006 System Software Reliability. *Springer*: London, UK.

[24] Tohma Y, Jacoby R, Murata Y and Yamamoto M 1989 Hyper-Geometric distribution model to estimate the number of residual software fault, *COMPSAC* pp. 610–617.

[25] Wood A 1996 Predicting software reliability. *IEEE Computers* 29(11) 69–77.

[26] Ohba M 1984 Inflexion S-shaped software reliability growth models. In: Osaki S and Hatoyma Y (eds) Stochastic Models in Reliability Theory. *Springer*: Berlin, Germany. pp. 144–162.