Study on Two Dimensional Three stage Software Reliability Growth Model

Rashmi Upadhyay, Prashant Johri

Abstract—Software reliability is defined as the probability of failure-free software operation for a specified period of time in a specified environment. Software release time is a crucial decision to make and has been extensively studied in the literature. A number of models have been proposed in the last two decades to discuss the software release time problem. From the earlier studies, it has been observed that more and more faults are detected as testing grows and as additional (testing) effort are used. In this paper, we reviewed the existing NHPP models for fault removal and proposed three stage software reliability growth model for minimizing the faults using Cobb Douglas production function for accumulating the effect of testing time and effort. The total fault removal phenomenon is the superposition of the three processes. The model has been validated, evaluated and compared with other existing NHPP models by applying them on actual failure data sets cited from real software development projects. The results show that the proposed model provides improved goodness of fit for software failure data and estimate results in SPSS and forecasting method is also proposed for the models.

Index Terms—Forecasting, Non homogeneous Poisson Process (NHPP), Software Reliability, Software Reliability Growth Model (SRGM), Statistical Package for Social Sciences (SPSS), Testing Effort.

1 INTRODUCTION

Software plays a critical role not only in scientific and business related enterprises, but also in daily life where it helps to run devices such as phones, cars, and television sets etc. Although advances have been made towards the production of defect free software, but this can be a costly, time consuming process. To effectively manage their budgets, managers require accurate information about the software development process and how the software reliability grows during this process. It is anticipated that software is made more reliable with the effect of this process, and can be modeled through the use of Software Reliability Growth Models (SRGMs).

SRGMs are applicable to the later stages of testing process. They can provide very useful information about how to improve reliability. Some important metrics, such as the number of initial faults, failure intensity, reliability within a specific time period, number of remaining faults, mean time between failures (MTBF), and mean time to failure (MTTF), can be easily determined through SRGMs. However, choosing a good model that can be used to explain the current and past failure behavior most adequately is very important. From our study, we find that many authors considered a Non-homogeneous Poisson Process (NHPP) as a stochastic process to describe the fault process. Ideally, these models provide a means of characterizing the development process, and enable software reliability practitioners to make predictions about the expected future reliability of software under development. Software development is a time consuming, costly process with high quality targets.

In this paper, We develop a model for two dimensional software reliability modeling and its failure process, which is a relatively unexplored area in the field of software reliability. Development of the software process for a firm is dependent on three major assets: time, effort, and tangible resources. All assets are limited, and precious. The software development process consists of requirements and design, coding, testing, and maintenance phases. Though in building quality software each phase plays its role, it is the testing phase which is always marked as the most crucial phase of the development process. In the last phase of software development process, the testing is carried out to detect and fix software faults introduced by human work, prior to its release for the operational use. The software faults that cannot be detected and fixed remain in the released software system after the testing phase. Thus, if a software failure occurs during the operational phase, then a computer system stops working and it may cause serious damage to our daily life. If the length of software testing is long, we can remove many software faults in the system and its reliability increases. However, it leads to increase the testing cost and to delay software delivery. In addition to modeling the software fault detection process, the total fault removal process is the superposition of the two NHPP. The main assumption of this model is that the simple faults removed immediately on their detection (the time between the fault detection and removal is negligible), while the hard faults need more cause analysis and consequently require more time to be removed. Kapur et al [2, 3] proposed an SRGM with three types of fault. For each type, the FRR per remaining faults is assumed to be time independent. The first type is modeled by an Exponentional model of Goel and Okumoto [1]. The second type is modeled by Delayed S-shaped model of Yamada et al. [5]. The third type is modeled by three stage Erlang model proposed by Khoshoghoftaar [4]. Later they extended their model to cater for more types of faults [3]. In their model authors ignored the role of the learning process during

• Rashmi Upadhyay is currently pursuing M.Tech Computer Science and Engineering, School of computing science and engineering, Galgotias University, Uttar Pradesh, India. E-mail: rurashmiupadhyay1@gmail.com

• Prashant Johri is currently working with School of computing science and engineering, Galgotias University, Uttar Pradesh, India E-mail: jojri.prashant@gmail.com
the testing phase by not accounting for the experience gained with the progress of software testing. The proposed model in this paper considers learning of the fault removal team by incorporating time dependent logistic learning function during the removal phase as it is expected the learning will grow with time. To cater the need of high exactitude of software reliability we require a software reliability growth model which supply not only the testing time but also the testing effort. For this we develop a two dimensional software reliability growth model integrating the joint effect of testing time and testing effort on the number of faults removed in the software.

The Proposed three stage model is validated with using two datasets DS-II and MUSA-413 and compares with other two models i.e. Goel Okumoto model and yamada model and with comparison table models has been proposed. The model considers that software system consists of a finite number of components and takes into account different types of errors, the time lag between the failure and fault removal processes. We have also presented numerical examples based on real data set obtained from real software development project. The results of the proposed model are very encouraging when compared with other model developed under similar environment. The paper is organized as follows. In Section 2, Literature review is given. In Section 3, Software reliability modeling with basic and proposed model framework is presented. The parameter estimation results and data validation of the formulated model is shown in Section 4. In Section 5 model evaluation is presented. Finally, conclusions and scope for future research are penned in Section 6.

2 LITERATURE REVIEW

According to the ANSI definition [6], “Software reliability is the probability of failure-free software operation for a specified period of time in a specified environment”. An SRGM provides a mathematical relationship between time span of testing or using the software and the cumulative number of faults detected. It is used to assess the reliability of the software during testing and operational phases.

The reliability of any software relies on testing phase. The testing phase focuses on detecting and correcting the faults in the software. It has been observed that more and more faults are detected as testing grows and as additional testing efforts are used. Testing phase of software development life cycle increases the reliability of software using the tools of Software Reliability Engineering. The software reliability growth model (SRGM) describes the relationship between the cumulative number of detected software errors and the time span of software testing.

The software faults are categorized into three types according to the number of stages needed to remove them. The fault is classified as simple if the time delay between failure observation, fault isolation and fault removal is negligible. For the simple faults, the fault removal phenomenon is modeled by the exponential model of Goel and Okumoto [1] with the logistic learning function. If there is a time delay the fault is classified as a hard fault. For the hard faults, the fault removal phenomenon is modeled similar to Delayed S-shaped model of Yamada et al. [7] with the logistic learning function during the removal phase as it is expected the learning will grow with time. If the removal of a fault after it is isolation involves even a greater time delay, it is classified as a complex fault. For the complex faults, the fault removal phenomenon is modeled by three-stage Erlang model proposed by Khoshogoftaar [6] with the logistic learning function. Models due to Bittanti et al. [13], and Kapur & Garg [11] are examples of flexible models. Later, Kapur et al. [12] proposed an SRGM with three types of faults. The first type was modeled by an Exponential model. The second type was modeled by a delayed S-shaped model. The third type is modeled by a three-stage Erlang model. The total removal phenomenon is again modeled by the superposition of the three SRGMs. Pham et al. [19] proposed a general imperfect software debugging model with an S shaped fault-detection rate. Recently, Kapur et al. [17] proposed a unified modeling framework for developing software reliability growth models under imperfect debugging. A number of models have been derived using this framework. Recently, Inoue et al. [10] also proposed a two-dimensional software reliability growth model with a change-point that described an actual phenomenon being related to the software reliability growth process.

Kapur et al. [9] also proposed a two dimensional modeling framework which was applied in determining optimal allocation of testing time and resources simultaneously to a modular software system. In 1991, Kapur and Garg [16] formulated release policies incorporating the effect of testing resource expenditure for an exponential SRGM under the added assumption that testing resource curves are described by either exponential, Rayleigh, or Weibull curves. Huang and Lyu [2] proposed an SRGM with a generalized testing effort function, and studied optimal release policies based on cost and reliability considering testing effort and efficiency. In 2007, Kapur et al. [14] proposed an SRGM with two types of imperfect debugging, and determined the optimal release time of the software. In this paper, we develop a three stage model which incorporates the combined effect of testing time and effort to remove the faults lying dormant in the software. The model developed is based on the Cobb-Douglas production function [18]. Recently, two dimensional software reliability models have been developed to appraise the software quantitatively. For solving software reliability problem the software engineer innovating a two dimensional model at the place of one dimensional since in one dimensional analysis the object variable is dependent on one basic variable although the object takes on many different roles based upon its dependence on various other factors. Two dimensional models are used to capture the joint effect of testing time and testing effort on the number of faults removed in the software. In economics, the Cobb-Douglas functional form of production functions is extensively used to personify the relationship of an output to inputs.

3 SOFTWARE RELIABILITY MODELLING

3.1 Model Development

The proposed model of this paper is based on categorizing the
faults in the software according to their removal complexity. The removal complexity is assumed to be proportional to the amount of time required to remove the fault. This complexity / time lag is represented by the number of stages required to remove the fault after the failure observation / fault isolation (with delay between the stages). This approach allows the proposed model to be flexible and simple. The model assumes that the testing phase consists of three processes namely, failure observation, fault isolation and fault removal.

Basic Model Assumptions
1. Failure observation / fault removal phenomenon is modeled by NHPP [2].
2. Software is subject to failures during execution caused by faults remaining in the software
3. The faults existing in the software are of three types: simple, hard, and complex faults.
4. Each time a failure is observed, an immediate (delayed) effort takes place to decide the cause of the failure and to remove it.
5. During the fault isolation / removal, no new fault is introduced into the software.
6. Failure rate of the software is equally affected by faults remaining in the software.

The time dependent behavior of fault removal process is explained by a Software Reliability Growth Model (SRGM). In literature, several SRGMs have been proposed to measure the reliability during the testing phase. Most of these can be categorized under Non Homogeneous Poisson Process (NHPP) models. The assumption that governs these models is, ‘the software failure occurs at random times during testing caused by faults lying dormant in software.’ And, for modeling the software fault detection phenomenon, counting process \( \{N(t); t \geq 0\} \) is defined which represents the cumulative number of software faults detected by testing time \( t \). The SRGM based on NHPP is formulated as:

\[
Pr\{N(t) = n\} = \frac{m(t)^n \cdot \exp(-m(t))}{n!}, \quad n = 0, 1, 2, ... \]

Where \( m(t) \) is the mean value function of the counting process \( N(t) \).

The model proposed in this paper incorporates the effect of these two vital factors simultaneously using a Cobb Douglas production function. The Cobb-Douglas function presents a simplified outlook of the economy in which production output is obtained by the amount of labor occupied, and the amount of capital invested. While there are many factors influencing economic performance, their model demonstrated remarkable accuracy. The mathematical form of the production function is specified as

\[
Y = AL^aK^{1-v} \]

Where, \( Y \) is the total production (the monetary value of all goods produced in a year), \( L \) is the labor input, \( K \) is the capital input, \( A \) is total factor productivity, and \( v \) is the elasticity of labor. The value of \( v \) is constant, and is determined by available technology.

To cater the combined effect of testing time and effort, we use the Cobb-Douglas production function [18] of the form

\[
\tau \equiv s' \tau^{1-r} \quad 0 \leq r \leq 1
\]

Let \( \{N(s, u), s \geq 0, u \geq 0\} \) be a two-dimensional stochastic process representing the cumulative number of software failures by time \( s \) and testing coverage \( u \). A two-dimensional NHPP with a mean value function \( m(s, u) \) is formulated as:

\[
Pr(N(s, u) = n) = \frac{(m(s, u))^n}{n!} \exp((-m(s, u))), \quad n=0,1,2...
\]

Notations:

- \( a \) Initial number of faults.
- \( b \) Fault detection rate per remaining fault.
- \( l \) Logistic learning factor.
- \( t \) Testing time.
- \( s \) Effort.
- \( r \) Resource elasticity to testing time.

\( m(s, t) \) Cumulative number of faults removed by time \( t \) and with using effort \( s \).

3.2 Modelling framework

3.2.1 GO Model Modelling Framework

GO model were used the combined effect of time and effort by representing the rate of change of the cumulative number of faults is given as

\[
m_r(\tau) = \frac{b}{1+l\exp(-br)}(a - m_r(\tau)) \tag{1}
\]

Integrating the above equation with initial condition \( m(\tau = 0) = 0 \), and using (1), we get

\[
m(s, t) = \frac{a(1 - \exp(-br))}{1 + l \exp(-br)}
\]

Replace \( \tau \equiv s' \tau^{1-r} \) and we get this equation.

\[
m(s, t) = \frac{a(1 - \exp(-bs')^{1-r})}{1 + l \exp(-bs')^{1-r}} \tag{2}
\]

3.2.2 Two Stage model Modelling framework

In Two stage model, we combined effect of time and effort by representing the rate of change of the cumulative number of faults is given as

\[
m_r(\tau) = b(a - m_r(\tau)) \tag{3}
\]

\[
m_r(\tau) = \frac{b}{1+l\exp(-br)}(m_r(\tau) - m_r(\tau)) \tag{4}
\]
Integrating the above equation with initial condition \(m(\tau = 0) = 0\), and using (3), we get

\[
m(s, t) = \frac{a(1 - \exp(-br(1 + br)'))}{1 + l * \exp(-br)}
\]

Replace \(\tau \equiv s't^{1-r}\) and we get this equation.

\[
m(s, t) = \frac{a(1 - \exp(-br(1 + br(1 + br)))}{1 + l * \exp(-br)}
\]

(5)

### 3.3 Proposed model Modelling framework

The combined effect of testing time and effort, we use the Cobb-Douglas production function.

Under the above assumptions, the differential equation representing the rate of change of the cumulative number of faults detected w.r.t. to the combined effect of time and effort is given as

\[
m_r'(\tau) = b(m(\tau) - m_r(\tau))
\]

(6)

\[
m_r'(\tau) = b(m(\tau) - m_r(\tau))
\]

(7)

\[
m_r'(\tau) = \frac{b}{1 + l * \exp(-br)}(m(\tau) - m_r(\tau))
\]

(8)

Integrating the above equation with initial condition \(m(\tau = 0) = 0\), and using (6), we get

\[
m(s, t) = \frac{a(1 - \exp(-br(1 + br(1 + br)))}{1 + l * \exp(-br)}
\]

Replace \(\tau \equiv s't^{1-r}\) and we get this equation.

\[
m(s, t) = \frac{a(1 - \exp(-br(1 + br(1 + br)))}{1 + l * \exp(-br)}
\]

(9)

In the above model, if \(l = 0\) the above model reduces to a 2-dimensional exponential model. For other values, it is S-shaped. If \(r = 0\) then it becomes a flexible time dependent SRGM proposed by Kapur et al [15].

### 4 Estimation Parameters

The SRGMs only become useful if it is possible to estimate their parameters. To validate the model described in Section 3.1, parameter estimation on actual software failure data is done.

In our study, we have used the Statistical Package for Social Sciences (SPSS). For the estimation of the parameters of the proposed model, nonlinear regression (NLR) modules of SPSS have been used to estimate the unknown parameters of the proposed model given in equation (5). Since the data set used is given in the form of pairs \((t_i, x_i, s_i)(i = 1, 2, ..., k)\), where \(x_i\) is the cumulative number of faults removed by time \(t_i(0 < t_1 < t_2 < ... < t_k)\) and \(s_i\) is the accumulated time spent to remove \(x_i\) faults and \(s_i\) is the cumulative execution time. The SPSS modules use the iterative estimation algorithms namely sequential quadratic programming and Levenberg-Marquardt method to find the least square estimates of the parameters.

#### 4.1 Model validation

To check the validity of the proposed model to describe the software reliability growth, it has been tested on two Data Sets DS-II and Musa-413.

The data set DS-II have 1301 cumulative faults and and working for 35 weeks and use 1846.92 cumulative CPU hours for remove these faults and the data set Musa-413 have 136 cumulative faults and and working for 21 weeks and use 25.26327 cumulative CPU hours for remove these faults.

In this paper we assume that the effort data set were better goodness of fit in three stage model as compared with GO model and 2 stage model. Parameters of SRGM were estimated using SPSS.

<table>
<thead>
<tr>
<th>Parameter Estimation DS-II</th>
<th>a</th>
<th>b</th>
<th>r</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td></td>
<td>.010</td>
<td>.0623</td>
<td>1.9</td>
</tr>
<tr>
<td>2stage</td>
<td>1350</td>
<td>.066</td>
<td>.209</td>
<td>1.9</td>
</tr>
<tr>
<td>Proposed</td>
<td>3 stage</td>
<td>1403.395</td>
<td>.166</td>
<td>.020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter Estimation Musa-413</th>
<th>a</th>
<th>b</th>
<th>r</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>135.00</td>
<td>.224</td>
<td>.399</td>
<td>4.872</td>
</tr>
<tr>
<td>2stage</td>
<td>140.00</td>
<td>.258</td>
<td>.178</td>
<td>9.874</td>
</tr>
<tr>
<td>Proposed</td>
<td>3 stage</td>
<td>175.46</td>
<td>.203</td>
<td>.128</td>
</tr>
</tbody>
</table>
The performance of an SRGM is judged by its ability to fit the past software failure / fault removal data (goodness of fit) and to predict satisfactorily the forecasted method for the validated data for analysis of forecasting.

In DS – II and MUSA – 413 the failure curves are exponential but again the proposed SRGM gives very good fits proving its flexibility. The plot of actual values and the estimated values of fault removal and faults remaining for the two data sets are depicted in figures 1 and 2 respectively.

5.1 The Goodness of Fit Criterion

1. The Mean Square Fitting Error (MSE): The model under comparison is used to simulate failure data, the difference between the expected values, \( \hat{m}(t_i) \) and the observed data \( y_i \) is measured by MSE as follows

\[
MSE = \frac{\sum_{i=1}^{k} (\hat{m}(t_i) - y_i)^2}{k}
\]

Where \( k \) is the number of observations. Lower value of MSE indicates less fitting error, thus better goodness of fit.

2. Coefficient of Multiple Determination (R\(^2\)) : This measure can be used to investigate whether a significant trend exists in the observed failure intensity. This coefficient is defined as the ratio of the Sum of Squares (SS) resulting from the trend model to that from a constant model subtracted from 1, that is:

\[
R^2 = 1 - \frac{\text{residual SS}}{\text{corrected SS}}
\]

\( R^2 \) measures the percentage of the total variation about the mean accounted for by the fitted curve. It ranges in value from 0 to 1. Small values indicate that the model does not fit the data well.

3. Coefficient of ADJ \( R^2 \): This coefficient is defined as the ratio of the Sum of Squares (SS) resulting from the trend model to that from a constant model subtracted from 1, that is:

\[
ADJ \, R^2 = 1 - \frac{(\text{residual SS})(n-k)}{(\text{corrected SS})(n-d)}
\]

Where \( n \) is the number of weeks of a software working for removing the faults, \( k \) is the number of independents variable in a data set and \( d \) is the number of dependent variable in a data set. It also ranges in value from 0 to 1. Small values indicate that the model does not fit the data well.

4. Sum of residual square error (SSE): This coefficients is a residual Sum of Squares(SS). Smaller the value of SSE is indicates the better goodness fit of data.

In other words, we evaluate the performance of the models using MSE, \( R^2 \), and ADJ \( R^2 \) metrics. The smaller the value of the metric the better the model fits relative to other models run on the same Data Set.

The proposed model is compared with the other models. The results are produced in table-. The proposed model is better for the given data sets under the above mentioned criteria.
5.2 Forecasted method

The forecast error is the difference between the actual value and the forecast value for the corresponding period.

\[ E_t = Y_t - F_t \]

where E is the forecast error at period t, Y is the actual value at period t, and F is the forecast for period t.

i) Mean Absolute Deviation (MAD):

\[ MAD = \frac{\sum_{t=1}^{N} |E_t|}{N} \]

ii) Root Mean Squared Error (RMSE):

\[ RMSE = \sqrt{\frac{\sum_{t=1}^{N} |E_t|^2}{N}} \]

we evaluate the performance of the models using MSE, SSE, R², and ADJ R² metrics for fitted data and and evaluate the performance of the models using MAD, RMSE of 20% data which is forecasted. The smaller the value of the metric the better the model fits relative to other models run on the same Data Set.
We have developed a three stage software reliability model. This model uniquely takes into account for removing the faults in three stage i.e. fault detection, fault isolation, and fault removal. Further, the model has been developed under the assumption that the fault removal process is governed not only by testing time but also by the effort expended. A relatively unexplored area of software reliability is investigated with this development structure. The proposed three stage model is applied to a real data sets. We have also presented numerical examples based on real data set obtained from real software development project the results are also formulated with forecasted method Results clearly show that the three stage model provides better goodnees of fit as well as it is also good for forecasted method. In future , the result estimated using SPSS is optimize and solved using a GA which minimizes the expected software cost subject to removing a minimum desired proportion of faults in the proposed model.

**TABLE 7**

<table>
<thead>
<tr>
<th>Models</th>
<th>a</th>
<th>b</th>
<th>r</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>108.00</td>
<td>0.294</td>
<td>0.502</td>
<td>3.712</td>
</tr>
<tr>
<td>2Stage</td>
<td>120.00</td>
<td>0.243</td>
<td>0.312</td>
<td>1.626</td>
</tr>
<tr>
<td>Proposed</td>
<td>180.00</td>
<td>0.195</td>
<td>0.112</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**TABLE 8**

<table>
<thead>
<tr>
<th>Models</th>
<th>Fitted data(80%)</th>
<th>Forecasted(20%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R2</td>
<td>R²</td>
</tr>
<tr>
<td>Basic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>.995</td>
<td>0.994</td>
</tr>
<tr>
<td>2Stage</td>
<td>.994</td>
<td>0.993</td>
</tr>
<tr>
<td>Proposed</td>
<td>.996</td>
<td>0.995</td>
</tr>
</tbody>
</table>

**REFERENCES**

13. S. Bittanti, P. Bolzern, E. Pedrotti, and R. Scattolini, “A flexible modeling approach for software reliability growth,” in *Software Reliability*


