Incorporating Burr Type XII Testing-efforts into Software Reliability Growth Modeling and Actual Data Analysis with Applications

Mohammad Ubaidullah Bokhari
Department of Computer Science, Aligarh Muslim University, Aligarh, India
Email: mubokhari@gmail.com

Nesar Ahmad *
University Department of Statistics and Computer Applications
T. M. Bhagalpur University, Bhagalpur, India
Email: nesar_bgp@yahoo.co.in

Abstract—Software reliability is the probability that the given software functions correctly under a given environment, during the specified period of time. During the software-testing phase, software reliability is highly related to the amount of development resources spent on detecting and correcting latent software errors, i.e. the amount of testing effort expenditures. This paper develops software reliability growth models (SRGM) based on non-homogeneous Poisson process (NHPP) which incorporates the Burr Type XII testing-effort functions (TEF). Numerous testing-effort functions for modeling software reliability growth based on NHPP have been proposed in the past decade. This paper shows that the Burr Type XII testing-effort function can be expressed as the actual testing-effort consumption during software development process. Its fault-prediction capability is evaluated through the numerical experiments. SRGM parameters are estimated by least square estimation (LSE) and maximum likelihood estimation (MLE) methods and computational experiments performed on actual software failure data set from various software projects. The results show that the proposed testing-efforts functions predict fault better than the other existing models. Thus, the proposed models evaluate software reliability more realistically. In addition, the optimal release policy based on reliability and cost criteria for software system are proposed.

Index Terms—SRGM, NHPP, Burr Type XII TEF, LSE, MLE, Testing effort consumptions

I. INTRODUCTION

In modern society, computer-controlled and computer-embedded systems are heavily dependent on the correct performance of software. So, it is quite natural to produce reliable software systems efficiently since the breakdown of the computer systems, which is caused by software errors, results in a tremendous loss and damage for social life. In the past years, several software reliability growth models (SRGM) based on NHPP which incorporates the testing-efforts have been proposed by many authors [2], [3], [6], [11], [13]-[17], [20]-[22], [37], [39], [40]. The testing-effort can be measured by the man power spent during the testing phase, the number of CPU hours, the number of executed test cases, and so on. Software reliability growth models proposed in the literature incorporating the effect of testing-effort expenditures described by the traditional Weibull type and Logistic type. However, it is difficult to represent the consumption curve only by these testing-effort consumption curves in various software development environments.

This paper describes the time dependent behavior of testing-effort expenditure by Burr Type XII model [9] as its curve is flexible having a wide variety of possible expenditure patterns in real software projects. This family includes exponential, Weibull and log-logistic as special cases. It also covers the curve shape characteristics of normal, log-normal, gamma, logistic and Pearson type X distributions as well as a significant portion of the curve shape characteristic for Pearson Type I (Beta), II, V, VII, IX and XII families [4], [29], [30], [33], [34]. Another advantage is that Burr XII has simple algebraic forms for reliability and hazard rate functions [4]. Thus Burr Type XII Provides a wide variety of density shapes along with functional simplicity. Currently there are few studies for the use of the Burr Type XII model in reliability and survival analysis [4], so this paper is to promote its use in software reliability analysis. The Burr Type XII failure model can be widely and effectively used in software reliability analysis, because it has a wide variety of shapes in its model and failure rate curves [4], making it useful for fitting many types of actual software failure data from various software projects.

Reference [1] has used of Burr Type XII distribution on software reliability growth modeling. This paper develops a realistic software reliability growth models based on NHHP which incorporates the Burr Type XII
testing–effort function [5]. It is assuming that the error detection rate in software testing is proportional to the current error content and the proportionality is the instantaneous software testing–effort expenditures at an arbitrary testing time. Its parameters are estimated by Least Square Estimation and Maximum Likelihood Estimation methods. Computational experiments are performed for three real software data and the results are compared with other existing model. It is shown that the proposed SRGM with Burr Type XII testing–effort function is wide and effective models for software reliability analysis. It can estimate the number of initial faults better as compare to other existing models. In addition, the optimal release policy of this model based on cost-reliability criterion is discussed.

II. BURR TYPE XII TESTING EFFORT FUNCTION

From the previous studies in [12], [13], and [20], we know that actual test effort data expressed various consumption pattern, sometimes the test effort consumption are difficult to describe only by Exponential, Rayleigh, Weibull or Logistic curve. Therefore, we try to incorporate a Burr Type XII test-effort function instead of above consumption function as the test effort function during the software development process in [7] and [8]. So, we proposed Burr Type XII curve as the test-effort function into SRGM.

The current testing effort consumption curve at testing time \( t \) is given as

\[
\omega(t) = \frac{\alpha \beta m \delta (\beta - t)^{\delta - 1}}{[1 + (\beta - t)^{\delta}]}^{n+1},
\]

where \( \alpha, \beta, m, \delta > 0, t > 0 \)

\[\alpha > 0, \beta > 0, m > 0, \delta > 0, t > 0 \quad (1)\]

The integral form of (1) is called the cumulative test-effort consumption of Burr Type XII in the time \([0, t]\) and is given by:

\[
W(t) = \int_0^t \omega(x)dx = \alpha \left[1 - (1 + (\beta - t)^{\delta})^{-n}\right]
\]

\[\alpha, \beta, m, \delta > 0, \quad t \geq 0 \quad (2)\]

The testing-effort function \( \omega(t) \) reaches its maximum value at the time \( t \)

\[
t_{\text{max}} = \frac{1}{\beta \left[\delta - 1 \right]} \left[\frac{\delta}{\beta \delta m + 1}\right]^{1/\delta}
\]

III. SOFTWARE RELIABILITY GROWTH MODEL

A. Model Description

A number of SRGMs have been proposed on the subject of software reliability. Among these models, Goel and Okumoto used an NHPP as the stochastic process to describe the fault process [11] and [23] modify the G-O model and incorporate the concept of testing-effort in an NHPP model to get a better description of the software fault detection phenomenon. We also propose a new SRGM with the Burr Type XII testing-effort function to predict the behavior of failure occurrences and the fault content of a software product

B. Assumptions

- The fault removal process is modeled by an NHPP.
- The software application is subject to failures at random times caused by the remaining faults in the system.
- The mean number of faults detected in the time interval \((t, t + \Delta t)\) by the current testing-effort is proportional to the mean number of remaining faults in the system at time \(t\), and the proportionality is a constant over time.
- Testing effort expenditures are described by the Burr Type XII testing-effort function.
- Each time a failure occurs, the corresponding fault is immediately removed and no new faults are introduced.
- The hazard rate for software occurring initially after the testing is proportional to the elapsed time \(r\) and the remaining faults.

An implemented software system is tested in the software development process. During the testing phase software errors remaining in the system cause software failures and the errors are detected and corrected by test personnel. A software failure is defined as an unacceptable departure of program operation. Following the usual assumptions in the area of software reliability growth modeling [10], we assume that the number of detected errors to the current test-effort expenditures is proportional to the current error content. Let \(m(t)\) represent the expected mean number of errors detected by testing calendar time \(t\) which is assumed to be a bounded non-decreasing function of \(t\) with \(m(0) = 0\). Then, using the Burr Type XII test-effort function in (1), we have the following differential equation [37]:

\[
\frac{dm(t)}{dt} = r[a - m(t)], \quad a > 0, \quad 0 < r < 1 \quad (3)
\]

where \(m(t)\) is the expected mean number of faults detected in time \((0, t)\), \(w(t)\) is the current testing-effort consumption at time \(t\), \(a\) is the expected number of initial faults, and \(r\) is the fault detection rate per unit testing-effort at testing time \(t\) and \(r > 0\).

Solving the differential equation (3) under the boundary condition \(m(0) = 0\) (i.e., the mean value function \(m(t)\) is equal to zero at time \(0\)), we have:

\[
m(t) = a \left[1 - e^{-r w(t)}\right] \quad (4)
\]

Substituting (2) for \(W(t)\) in (4) we get:
\[ m(t) = a \left[ 1 - e^{-r_\alpha \left[ t + (\alpha t)^{\beta} \right]^{-\delta}} \right] \]  

(5)

From (4), we have the following important relationship between \( m(t) \) and \( W(t) \):

\[ W(t) = \frac{1}{r} \ln \left( \frac{a}{a - m(t)} \right). \]  

(6)

For stochastic modeling of a software error detection phenomenon, let \( \{N(t), t > 0\} \) be a counting process representing the cumulative number of errors detected by testing time \( t \). Defining the expected value of \( N(t) \) by \( m(t) \) in (5), we can describe a software reliability growth model incorporating the Burr Type XII test-effort function \([13], [36]\) by an NHPP as:

\[ \text{Pr} \{ N(k) = n \} = \left[ m(t) \right]^n e^{-m(t)} \frac{n!}{n!}, \quad n = 0, 1, 2, \ldots \]  

(7)

where \( m(t) \) is called mean value function of the NHPP \([10], [38], [39]\) and \( \text{Poi}m \ m(n; m(t)) \) is a Poisson pmf with parameter \( m(t) \). The intensity function of the NHPP is given by:

\[ \lambda(t) = \frac{dm(t)}{dt} = \alpha r \cdot w(t) e^{-rW(t)} \]  

(8)

which means the instantaneous error detection rate. From (7) we can show that the limit distribution of \( N(t) \) is a Poisson distribution with the following mean:

\[ m(\infty) = a \left( 1 - e^{-r\alpha} \right) \]  

(9)

Equation (9) implies that even if a software system is tested during an infinitely long duration, all errors remaining in the system cannot be detected \([39], [40]\). Thus, the mean number of undetected errors \( d(t) \) if a test is applied for an infinite amount of time is:

\[ a - m(\infty) = a - a \left( 1 - e^{-r\alpha} \right) \]  

\[ d(t) = ae^{-r\alpha} \]

C. Software Reliability Measures

Let \( N(t) \) represent the number of errors remaining in the system of testing time \( t \). Based on the NHPP model with \( m(t) \), given by in equation (4), two quantitative measures for software reliability assessment can be derived \([10], [36]\). The expectation of \( N(t) \) and its variance are given by:

\[ r(t) = E \left[ N(t) \right] = E \left[ N(\infty) - N(t) \right] \]  

\[ = m(\infty) - m(t) = a \left[ e^{-rW(t)} - e^{-rW(\infty)} \right] \]  

\[ = Var \left[ N(t) \right], \]  

(10)

The software reliability representing the probability that a software failure does not occur in the time interval \((t, t + x)\) is given by:

\[ R = R \left( x | t \right) = e^{-\sum_{i=t+1}^{t+x} \left( m(i) - m(t) \right)} \]  

\[ = e^{-\sum_{i=t+1}^{t+x} \left[ e^{-rW(i)} - e^{-rW(t)} \right]} \]  

(11)

It can be easily seen that \( R \left( x | t \right) \) is a monotonic increasing function of \( t \). Taking log both sides in (11)

\[ \ln R = - \left[ m(t + x) - m(t) \right] \]

Solving the above equation with \( m(t) \), one can estimate derived reliability \( R \). The instantaneous mean time between failures (MTBF) at arbitrary testing can be defined as a reciprocal of error detection rate in equation (8). Then the instantaneous MTBF is given by:

\[ MTBF(t) = \frac{1}{\lambda(t)} = \frac{1}{a \cdot r \cdot w(t) e^{-rW(t)}} \]  

\[ = \frac{e^{-rW(t)}}{a \cdot r \cdot m(t) \cdot \beta \cdot \delta \cdot \left( \beta \cdot t \right)^{\delta} [1 + (\beta \cdot t)^{-\delta}]^{-1} \]  

(12)

IV. ESTIMATION METHODS OF PARAMETERS

A. Estimation of Testing-Effort Parameters

Two most popular estimation techniques are Maximum Likelihood Estimation (MLE) and Least Squares Estimation (LSE) \([23], [26]\). The parameters \( \alpha, \beta, m \) and \( \delta \) in the Burr Type XII testing-effort functions defined by the equation (1) can be estimated by least squares. The estimators for \( \alpha, \beta, m \) and \( \delta \) are investigated for testing-effort \( w_i \) spent during \((0, t_i) \) (\( k = 1, 2, \ldots, n \)). Then, based on the usual procedures, the least-squares estimators \( \hat{\alpha}, \hat{\beta}, \hat{m} \) and \( \hat{\delta} \) can be obtained by minimizing

\[ \text{Minimize } S(\alpha, \beta, m, \delta) = \sum_{k=1}^{n} \left( W_k - W(t_i) \right)^2 \]

Taking log in Equation (1), we get

\[ \ln w_k = \ln \alpha + \ln \beta + \ln m + \ln \delta + \left[ (\delta - 1) \ln (\beta \cdot t_i) \right] + \left[ (m + 1) \ln [1 + (\beta t_i)^\delta] \right] \]  

(13)

Then, the least–squares estimates \( \hat{\alpha}, \hat{\beta}, \hat{m} \) and \( \hat{\delta} \) of parameters \( \alpha, \beta, m \) and \( \delta \) can be obtained by minimizing the following sum of squares

\[ S(\alpha, \beta, m, \delta) = \sum_{k=1}^{n} \left( \ln w_k - \ln \alpha - \ln \beta - \ln m - \ln \delta \right)^2 \]  

(14)
Differentiate the above equation with respect to the \( \alpha, \beta, m \) and \( \delta \) and set the partial derivatives to zero, we get the following non-linear equations,

\[
\frac{dS}{d\alpha} = \sum_{k=1}^{n} \frac{n!}{(m+1)!} \left[ \frac{\ln w_k - \ln \alpha - \ln \beta - \ln m - \ln n \delta - (\delta - 1) \ln (\delta - 1) \ln (\beta_k^\delta)}{(m+1) \ln (\beta_k^\delta)} \right] = 0
\]

\[
\Rightarrow \sum_{k=1}^{n} \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \sum_{k=1}^{n} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) = 0 \tag{15}
\]

\[
\frac{dS}{d\beta} = \sum_{k=1}^{n} \left[ \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \right] \times \frac{1}{\beta} \frac{\delta - 1}{\beta_k^\delta + (m+1)} = 0
\]

\[
\Rightarrow \sum_{k=1}^{n} \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \sum_{k=1}^{n} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \times \frac{1}{\beta_k^\delta + (m+1)} = 0 \tag{16}
\]

\[
\frac{dS}{dm} = \sum_{k=1}^{n} \left[ \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \right] \times \frac{-1}{m} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) = 0
\]

\[
\Rightarrow \sum_{k=1}^{n} \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \sum_{k=1}^{n} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \times \frac{-1}{\beta_k^\delta + (m+1)} = 0 \tag{17}
\]

\[
\frac{dS}{d\delta} = \sum_{k=1}^{n} \left[ \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \right] \times \frac{-1}{\delta} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \times \frac{1}{(\beta_k^\delta + (m+1))} = 0
\]

\[
\Rightarrow \sum_{k=1}^{n} \ln w_k - n \ln \alpha - n \ln \beta - n \ln m - n \ln n \delta - (\delta - 1) \sum_{k=1}^{n} \ln (\beta_k^\delta) + (m+1) \ln (\beta_k^\delta) \times \frac{-1}{\beta_k^\delta + (m+1)} = 0 \tag{18}
\]

These non-linear equations can be solved numerically to get the estimate of \( \alpha, \beta, m \) and \( \delta \).

### C. Estimation of Reliability Growth Parameters

The reliability growth parameters \( a, r \) in the NHPP model with \( m(t) \) in (4) can be estimated by the method of maximum-likelihood [10]. Let the estimated parameters \( \hat{a}, \hat{r}, \hat{m} \) and \( \hat{\delta} \) in the Burr type XII test-effort function in (1) have been obtained by the method of least-squares. The \( \hat{a} \) and \( \hat{r} \) are determined for the \( n \) observed data pairs \( (t_k, y_k) \) \( (k=1,2,\ldots,n) \). Then, the joint p.m.f, the log-likelihood function, for the unknown parameters \( a, r \) and \( m(t) \) in the NHPP model with \( m(t) \) in (4), is:

\[
\ln L = \sum_{i=1}^{n} \ln (y_i - y_{i-1}) - \ln \left[ \exp \left[ -r W(t_i) \right] - \exp \left[ -r W(t_{i-1}) \right] \right] - a \left( 1 - \exp \left[ -r W(t_i) \right] \right) - \sum_{i=1}^{n} \ln \left( y_i - y_{i-1} \right)
\]

\[
t_0 = 0 \quad \text{and} \quad y_0 = 0.
\]

The usual calculus methods for an interior maximum result in

\[
y_0 = a \cdot f_s, \quad \Rightarrow \hat{a} = \frac{y_0}{f_s} \tag{20}
\]

and

\[
a \cdot g_s = \sum_{i=1}^{k} \frac{(y_i - y_{i-1})}{(f_i - f_{i-1})}, \tag{21}
\]

where,

\[
f_i = 1 - \exp \left[ -r W(t_i) \right],
\]

\[
g_i = W(t_i) \exp \left[ -r W(t_i) \right], \quad (k=1,2,\ldots,n), \tag{22}
\]

which can be solved numerically.

If the sample size \( n \) of the observed data is sufficient large, the maximum-likelihood estimates \( \hat{a} \) and \( \hat{r} \) asymptotically follow a bivariate normal distribution [28],

\[
\left( \hat{a}, \hat{r} \right) \sim \text{BIVN} \left( \left( \begin{array}{c} a \\ r \end{array} \right), \Sigma \right), \quad n \rightarrow \infty, \tag{23}
\]

The \( \Sigma \) in the asymptotic properties of (23) is useful in quantifying the variability of the estimated parameters \( \hat{a} \) and \( \hat{r} \), and is the inverse of \( F \):

\[
F = \begin{bmatrix}
E \left[ -\frac{\partial^2 \ln L}{\partial a^2} \right] & E \left[ -\frac{\partial^2 \ln L}{\partial a \partial r} \right] \\
E \left[ -\frac{\partial^2 \ln L}{\partial r \partial a} \right] & E \left[ -\frac{\partial^2 \ln L}{\partial r^2} \right]
\end{bmatrix}
\]

\[
\begin{bmatrix}
f_a & g_a \\
g_a & \sum_{k=1}^{n} (g_k - g_{k-1})^2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
f_a & g_a \\
g_a & \sum_{k=1}^{n} (g_k - g_{k-1})^2
\end{bmatrix}
\]

where, \( g_k = W(t_k) \cdot \exp \left[ -r W(t_k) \right] \), and \( f_1 = 1 - \exp \left[ -r W(t_k) \right] \) where \( k=1, \ldots, n \)

Substituting the value of \( a \) and \( r \) in (4.2.6) and calculate \( F^{-1} \). The estimated asymptotic variance-covariance matrix is:

\[
\hat{\Sigma} = F^{-1} = \begin{bmatrix}
\text{Var}(\hat{a}) & \text{Cov}(\hat{a}, \hat{r}) \\
\text{Cov}(\hat{a}, \hat{r}) & \text{Var}(\hat{r})
\end{bmatrix}
\]

### V. SOFTWARE FAILURE DATA ANALYSIS

The two performance comparison criteria are given here to check the performance of the proposed software reliability growth model and to make affair comparison with the other existing SRGM.

- **Mean square of fitting error (MSE):**

\[
\text{MSE} = \frac{\sum_{k=1}^{k} (m(t_k) - y_k)^2}{k}
\]

where \( k \) is the number of observation. A smaller MSE indicates a smaller fitting error and better performance [21], [24].

- **AE (Accuracy of Estimation)** is defined as:

\[
\text{AE} = \frac{M_o - M_a}{M_o}
\]

\( M_o \) is the actual cumulative number of detected faults after the test, and \( M_a \) is estimated number of initial faults [10], [22], [26].

### A. Performance Analysis

**First Data Set:** The first set of real data in this paper is the System T1 data of the Rome Air Development Center.
(RADC) projects and cited from [25] and [26]. The number of object instructions for the system T1 which is used for a real-time command and control application. In this case, the size of the software is approximately 21,700 object instructions. The software was tested for 21 weeks with 9 programmers. During the test phase, about 25.3 CPU Hours were used and 136 faults were detected. Similarly the MLE and LSE are used to estimate the parameters for the equation (1) and equation (4)

In order to estimate the parameters $\alpha$, $\beta$, $m$ and $\delta$ of the log-logistic test-effort function, the actual testing-effort data into equations (1) has been fitted and solve it by using the method of least squares. The estimated values of parameters of the Burr Type XII testing-effort function are:

$$\hat{\alpha} = 35.242, \hat{\beta} = 0.063, \hat{m} = 0.326 \text{ and } \hat{\delta} = 11.259$$

Fig. 1 and Fig. 2 shows the fitting of the estimated testing-effort by using Equation (1) and (2). The fitted curves are shown as a dotted line and solid line for actual software data in the graphs. Using the estimated parameters $\alpha$, $\beta$, $m$ and $\delta$ the other parameters $a$, $r$ in (4) can be solved by MLE method. The cumulative numbers of estimated failures by equation (4) are:

$$a = 133.7025, \quad r = 0.1553$$

For these estimates, the optimality was checked numerically. Table I summarizes the experimental results of estimated parameters with their standard errors and 95% confidence bound.

Following the same procedure, we plotted a fitted curve of the estimated mean value function with the actual software data in Fig. 3. Also a comparison table of the estimates of this model along with other SRGMs with initial faults $a$ and MSE is given in Table II. From Figs. 1, 2, and 3 and the comparison criteria in Table II, it is conceivable that the proposed SRGM has a better goodness of fit. Kolmogorov Smirnov goodness-of-fit test shows that this proposed SRGM described by an NHPP with $\hat{\lambda}(t)$ fits pretty well at the 5% level of significance.

Fig. 4 shows that the estimated intensity functions $\hat{\lambda}(t)$ from equation (8).

Substituting the estimated parameters $\alpha$, $\beta$, $m$ and $\delta$ in equation of $t_{\text{max}}$, the testing effort function reaches the maximum at time $t = 16.9143$ debug days which corresponds to $\nu(t) = 3.5986$ CPU hours and $W(t) = 11.1148$ CPU hours. Besides, the number of errors removed up to this time $t_{\text{max}}$ is 109.9073 and when $t$ goes to infinity, the numbers of errors removed is 133.14116.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>SUMMARY OF ESTIMATE OF NHPP MODEL PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Estimate</td>
</tr>
<tr>
<td>$a$</td>
<td>133.279</td>
</tr>
<tr>
<td>$r$</td>
<td>0.1553</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>COMPARISON RESULTS FOR THE FIRST DATA SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$a$</td>
</tr>
<tr>
<td>Burr Type XII Model</td>
<td>133.70</td>
</tr>
<tr>
<td>G-O Model [27]</td>
<td>142.32</td>
</tr>
<tr>
<td>Exponential Model [11]</td>
<td>137.2</td>
</tr>
<tr>
<td>Rayleigh Function [22]</td>
<td>866.94</td>
</tr>
<tr>
<td>Delayed s-shaped Model [13]</td>
<td>237.19</td>
</tr>
</tbody>
</table>

© 2014 ACADEMY PUBLISHER
Second Data Set: The second set of real data is the pattern of discovery of faults by [32]. The debugging time and the number of detected faults per day are reported. The cumulative number of discovered faults up to twenty two days is 86 and the total consumed debugging times is 93 CPU hours. All debugging data are used in this experiment. The testing-effort data are applied to estimate the parameters $\alpha$, $\beta$, $m$ and $\delta$ of the Burr Type XII distributed function described in equations (1) by using the method of least squares. Hence, we can find the estimates only through numerical procedures. We can estimate each parameter by the Maximum Likelihood Estimation and Least Square Estimation in the Burr Type XII Distribution Function (proposed SRGM). The estimated values of parameters are:

$$\hat{\alpha} = 121.4621, \hat{\beta} = 0.005657, \hat{\delta} = 1.908,$$
$$\hat{m} = 78.914, \hat{a} = 94.435, \hat{r} = 0.0255$$

Fig. 5 and Fig. 6 depict the fitting of the current estimated testing-effort by using Burr Type XII testing-effort function.

For these estimates, the optimality was checked numerically. Table III summarizes the experimental results of estimated parameters with their standard errors and 95% confidence bound. Similarly, we plotted a fitted curve of the estimated mean value function with the actual software data in Fig. 7. Table IV shows the estimated values of parameters by using different SRGMs and comparison criteria. Similarly, smaller AE and MSE indicate least fitting errors and better performance. From Figures 5, 6, and 7 and the comparison criteria in Table IV, we conclude that this proposed model is good enough a give more accurate description of resource consumption during the source development phase and gives better fit in this experiment. Kolmogorov Smirnov goodness-of-fit test shows that our proposed SRGM described by an NHPP with $\hat{m}(t)$ fits pretty well at the 5% level of significance. Figure 8 shows that the estimated intensity functions $\hat{\lambda}(t)$ from equation (8).

In addition, substituting the estimated parameters $\alpha$, $\beta$, $m$ and $\delta$ in equation of $t_{\text{max}}$, the testing effort function reaches the maximum at time $t_{\text{max}} = 65.0016$ and when $t$ goes to infinity, the numbers of errors removed is 90.1894.

**TABLE III**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>95% Confidence Lower</th>
<th>95% Confidence Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>94.435</td>
<td>2.556930</td>
<td>89.10086</td>
<td>99.7682</td>
</tr>
<tr>
<td>$r$</td>
<td>0.02554</td>
<td>0.0016003</td>
<td>0.022030</td>
<td>0.02888</td>
</tr>
</tbody>
</table>

**TABLE IV**

<table>
<thead>
<tr>
<th>Model</th>
<th>$a$</th>
<th>$r$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burr Type XII Model</td>
<td>94.435</td>
<td>0.02554</td>
<td>6.726</td>
</tr>
<tr>
<td>G-O Model [27]</td>
<td>137.072</td>
<td>0.0515445</td>
<td>25.33</td>
</tr>
<tr>
<td>Weibull Function [12]</td>
<td>87.0318</td>
<td>0.0345417</td>
<td>7.772</td>
</tr>
<tr>
<td>Delayed s-shaped Model</td>
<td>88.633</td>
<td>0.228148</td>
<td>6.3127</td>
</tr>
<tr>
<td>Logistic Function [22]</td>
<td>88.8931</td>
<td>0.0390591</td>
<td>25.228</td>
</tr>
</tbody>
</table>

$W(t) = 45.6458$ CPU hours. Besides, the number of errors removed up to this time $t_{\text{max}}$ is 65.0016 and when $t$ goes to infinity, the numbers of errors removed is 90.1894.
**Third Data Set:** The third set of real data is from the study by [27]. The system is PL/1 database application software, consisting of approximately 1,317,000 lines of code. During the nineteen weeks experiments, 47.65 CPU times were consumed and about 328 software errors were removed. The original data report gives that the total cumulative number of detected faults after a long period of testing is 358 faults [27]. In order to estimate the parameters $\alpha$, $\beta$, $m$ and $\delta$ of the Burr Type XII distributed function; we fit the actual testing-effort data into equations (1) and (2) and solve it by using the method of least squares. Hence, we can find the estimates only through numerical procedures. These estimated parameters are:

$$\hat{\alpha} = 675.20762, \quad \hat{\beta} = 0.000251, \quad \hat{\delta} = 1.11883, \quad \hat{m} = 29.1946$$

Fig. 9 and Fig. 10 show the fitting of the estimated testing-effort. Here, the fitted curves are shown as a dotted line and solid line is actual software data. Using the estimated parameters $\alpha$, $\beta$, $m$ and $\delta$, the other parameters $a$, $r$ in (4) can be solved by MLE method for these failure data:

$$a = 565.6973, \quad r = 0.01964$$

For these estimates, the optimality was checked numerically. Table V summarizes the experimental results of estimated parameters with their standard errors and 95% confidence bound.

Similarly, fitted curve of the estimated mean value function with the actual software data in Fig. 11 has been plotted. Also a comparison table of the estimates of this model along with other models with initial faults $a$ and MSE is given in Table VI. From Figures 9, 10, and 11 and the comparison criteria shows that this SRGM is better fit than the other models for PL/1 application program. Kolmogorov Smirnov goodness-of-fit test shows that our proposed SRGM described by an NHPP with $\hat{m}(t)$ fits pretty well at the 5% level of significance. Figure 12 shows that the estimated intensity functions $\hat{\lambda}(t)$ from equation (8).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>95% Confidence Lower</th>
<th>95% Confidence Upper</th>
<th>AE%</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>565.6973</td>
<td>0.01964</td>
<td>565.69734</td>
<td>544.9546</td>
<td>58.02</td>
<td>116.40</td>
</tr>
<tr>
<td>$r$</td>
<td>0.01964</td>
<td>0.002826</td>
<td>0.013677</td>
<td>0.0255998</td>
<td>58.02</td>
<td>116.40</td>
</tr>
</tbody>
</table>

**TABLE VI**

<table>
<thead>
<tr>
<th>Model</th>
<th>$a$</th>
<th>$r$</th>
<th>AE%</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burr Type XII Model</td>
<td>565.697</td>
<td>0.01964</td>
<td>58.02</td>
<td>116.40</td>
</tr>
<tr>
<td>Inflection s-shaped Model</td>
<td>389.1</td>
<td>0.0935493</td>
<td>8.69</td>
<td>133.53</td>
</tr>
<tr>
<td>Exponential Model</td>
<td>455.37</td>
<td>0.0267368</td>
<td>27.09</td>
<td>206.93</td>
</tr>
<tr>
<td>Weibull Function</td>
<td>565.35</td>
<td>0.0196597</td>
<td>57.91</td>
<td>122.09</td>
</tr>
<tr>
<td>Rayleigh Function</td>
<td>459.08</td>
<td>0.0273367</td>
<td>28.23</td>
<td>268.42</td>
</tr>
<tr>
<td>Exponential Function [12]</td>
<td>828.252</td>
<td>0.0117836</td>
<td>131.35</td>
<td>140.66</td>
</tr>
<tr>
<td>Delayed s-shaped Model [13]</td>
<td>374.05</td>
<td>0.197651</td>
<td>4.48</td>
<td>168.67</td>
</tr>
<tr>
<td>Delayed s-shaped Model with Rayleigh Function [12]</td>
<td>333.136</td>
<td>0.100415</td>
<td>6.93</td>
<td>798.49</td>
</tr>
<tr>
<td>S-Shaped Model with Logistic Function [22]</td>
<td>338.136</td>
<td>0.10004</td>
<td>5.54</td>
<td>242.79</td>
</tr>
</tbody>
</table>

**TABLE V**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>95% Confidence Lower</th>
<th>95% Confidence Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>338.136</td>
<td>0.10004</td>
<td>338.136</td>
<td>338.136</td>
</tr>
<tr>
<td>$r$</td>
<td>6.93</td>
<td>0.100415</td>
<td>6.93</td>
<td>798.49</td>
</tr>
<tr>
<td>AE%</td>
<td>5.54</td>
<td>0.100415</td>
<td>5.54</td>
<td>242.79</td>
</tr>
<tr>
<td>MSE</td>
<td>242.79</td>
<td>0.100415</td>
<td>242.79</td>
<td>242.79</td>
</tr>
</tbody>
</table>
δ = 0.005657, β = 0.000251, \hat{\alpha} = 0.0634, which corresponds to the function reaches the maximum at time \( t = 67.2605 \) CPU hours. Besides, the expected number of errors removed up to this time \( \max_m = 565.6993 \).

### A. Reliability Criteria

Reliability is maximized [28].

When to release software so that the cost incurred during operational phases) of the software is minimized or the reliability of a software system  is known to have reached an acceptable level, then we can obtain the right time to release this software. References [28] and [35] discussed the release problem by considering the software cost-benefit. The conditional reliability function after the last failure occurs at time \( t \) is:

\[
R = R(x | t) = e^{-(m(t+x) - m(t))} = e^{-e^{-\delta W(t)} - e^{-rW(t+x)}} \quad (25)
\]

Differentiate \( R(x | t) \) with respect to \( t \), then \( \frac{dR}{dt} \geq 0 \). Hence \( R \) is a monotonic increasing function of \( t \).

Taking the logarithm on both side of the above equation, we obtain.

\[
\log R = -[m(t+x) - m(t)] \quad (26)
\]

Solving (26) and (4) determines the testing time needed to reach a desired \( R \). \( R(t) \) is increasing in \( t (0 < t < T_{IC}) \). Using (26), one can get the required testing time needed to reach the reliability objective \( R \) or decide whether \( R \) is reached or not in a specified time interval.

- **Reliability Analysis For Real Data Sets**

  **First Data Set:** From the previous estimated parameters: we know that \( \hat{\alpha} = 35.2418, \hat{\beta} = 0.0634, \hat{\delta} = 11.2592, a = 133.7025, r = 0.1553 \)

  Suppose this software system is desired that this testing would be continued till the operational reliability is equal to 0.85 (at \( \Delta t = 0.1 \)), from equation (26) and equation (4), we get \( t = 20.3456 \) weeks. If the desired reliability is 0.90, then \( t = 21.1729 \) weeks. If the desired reliability is 0.95, then \( t = 22.7449 \) weeks. If the desired reliability is 0.99, then \( t = 27.2316 \) weeks.

  **Second Data Set:** In second data set, from equation (26)and equation (4), for \( \hat{\alpha} = 121.4621, \hat{\beta} = 0.005657, \hat{\delta} = 1.908, \hat{m} = 0.3261, \hat{\delta} = 78.9143, a = 94.4345, r = 0.02554 \) The testing time \( t = 17.8319 \) days is obtained, if we assume that the testing of this software system is desired to be continued till the operational reliability is equal to 0.85 (at \( \Delta t = 0.1 \)). If the desired reliability is 0.90, then \( t = 20.8985 \) days. If the desired reliability is 0.95 (0.99), then \( t = 24.3609 \) (33.7320) weeks.

  **Third Data Set:** From the previous estimated parameters: \( \hat{\alpha} = 675.20762, \hat{\beta} = 0.000251, \hat{\delta} = 1.11883, \hat{m} = 0.3261, \hat{\delta} = 29.1946, a = 565.6973, r = 0.0196 \), suppose this software system is desired that the testing would be continued till the operational reliability is equal to 0.8 (at \( \Delta t = 0.1 \)), from equation (26) and equation (4), we get testing time \( t = 10.3198 \) weeks. If the desired reliability is 0.85, then \( t = 10.9415 \) days. If the desired reliability is 0.92 (0.98), then \( t = 12.2399 \) (14.9614) weeks.

### B. Cost-Reliability Criteria

This section discusses the cost model and release policy based on the cost-reliability criterion we can
evaluate the total software cost by using cost criterion, 
the cost of testing-effort expenditures during software 
testing and development phase, and the cost of correcting 
errors before and after release as follows [18], [19], [38], 
[39]:

\[ C(T) = C_mT + C_{mT_e} - m(T) + C_{w(T)}dT \]  

(27)

Where \( C_i \) is the cost of correcting an error during 
testing, \( C_2 \) is the cost of correcting an error in operational 
use (\( C_2 > C_i \)), \( C_i \) is the cost testing per unit testing-effort 
expenditures and \( T_{L_i} \) is the software life-cycle length.

Differentiating the above equation w. r. t. \( T \) and setting 
\( C(T) \) to zero, we obtain

\[ \frac{dC(T)}{dT} = C_1 \frac{dm(T)}{dT} - C_2 \frac{dm(T)}{dT} + C_3 w(T) = 0 \]

Or, \( \frac{dC(T)}{dT} = w(T) \left[ -(C_2 - C_i) a r e^{-w(T)} + C_2 \right] \)  

(28)

Now

\[ \frac{w(T)C_i}{C_2 - C_i} = w(T) a r e^{-w(T)} \]

\[ m, \frac{C_2 - C_2}{C_2 - C_i} = a r e^{-w(T)} \]

\[ \lambda(T) = \frac{r(a - m(T))}{w(T)} \]

\[ \therefore \frac{\lambda(T)}{w(T)} = \frac{C_1}{C_2 - C_i} = r(a - m(T)) \]  

(29)

Case 1: If \( T = 0 \), then \( m(0) = 0 \), and

\[ \frac{\lambda(T)}{w(T)} = \lim_{T \to 0} a r \]

Case 2: If \( T \to \infty \), then \( W(\infty) = \infty \), \( m(\infty) = a(1 - e^{-\alpha}) \)

\[ \frac{\lambda(T)}{w(T)} = \lim_{T \to \infty} a r e^{-\alpha} \]

Therefore, \( \frac{\lambda(T)}{w(T)} \) is monotonically decreasing in \( T \).

If \( \frac{\lambda(0)}{w(0)} = a r \leq \frac{C_1}{C_2 - C_i} \)

Then,

\[ \frac{\lambda(T)}{w(T)} \leq \frac{C_1}{C_2 - C_i} \]  

for \( 0 < T < T_{L_i} \)

Hence for this case, the optimal software release time 
\( T' \) = 0,

since \( \frac{dC(T)}{dT} > 0 \) for \( 0 < T < T_{L_i} \).

If \( \frac{\lambda(0)}{w(0)} = a r > \frac{C_1}{C_2 - C_i} \), \( \frac{\lambda(T)}{w(T)} = a r e^{-\alpha} \),

Then, there exist a finite and unique solution. 
To satisfying equation (29) that is,

\[ \frac{\lambda(T)}{w(T)} = \frac{C_1}{C_2 - C_i} = r(a - m(T)) \]

\[ = a r e^{-\alpha} \]

\[ = a r e^{-\alpha} \left( 1 + (\beta T)^{\delta} \right) \]

Rearranging this equation gives,

\[ e^{\alpha(1+\beta T)^{-\delta}} = \frac{a r (C_2 - C_i)}{C_2} \]

\[ \text{or} \left[ 1 - (1 + (\beta T)^{-\delta}) \right] = \ln \left[ \frac{a r (C_2 - C_i)}{C_2} \right] \]

\[ \text{or} \left[ \frac{1}{(\beta T)^{-\delta}} \right] = \left[ \frac{r e^{-\alpha}}{a r (C_2 - C_i)} \right] \]

Minimizes \( C(T) \)

Because \( \frac{dC(T)}{dT} < 0 \) for \( 0 < T < T_0 \) and

\( \frac{dC(T)}{dT} > 0 \) for \( T_0 < T < T_{L_i} \).

The minimum of \( C(T) \) is at \( T = T_0 \) for \( T_0 < T \), because

\( \frac{d^2C(T)}{dT^2} > 0 \), then \( C(T) \) is a convex function.

Here our goal is to minimize the total software cost 
under the consideration of desired software reliability, 
and the optimal software release time is obtained. That is, 
the optimal software release problem can be formulated 
as follows.

Minimize \( C(T) \)  

Subject to \( R(x|t) \geq R_0, \)

\[ T \geq 0 \] for \( C_2 > C_1 > 0, C_2 > x \geq 0, \)

\[ 0 < R_0 < 1. \]

Then, we can obtain the solutions for the cost reliability 
optimum software release time:

\[ T' = \max \{ T_0, T_1 \} \]

Where \( T_0 \) is finite and the unique solution \( T \) of (31), \( T_1 \) is 
finite and unique \( T \) satisfying \( R(x|t) \geq R_0, 0 < R_0 < 1. \)

Theorem: We assume that;

\( C_1 > 0, C_2 > 0, C_1 > 0, C_2 > C_1, x > 0, 0 < R_0 < 1, \) then

- If \( \frac{\lambda(0)}{w(0)} > \frac{C_1}{C_2 - C_i} \) and \( \frac{\lambda(T)}{w(T)} = \frac{C_1}{C_2 - C_i} \) \( \therefore \) \( T' = \max \{ T_0, T_1 \} \) for \( R(x|t) < R_0 \) or \( T' = T_0 \) for \( 0 < R < R(x|t) = 0 \).
Exponential, Rayleigh, and Weibull type consumption testing-effort curve gives better estimates than compared to the other existing models. Burr Type XII that the proposed model has a better goodness of fit as three actual software failures data set. We also conclude the testing-effort function proposed here, gives a good reliability growth model. Computation results show that testing-effort function can be used to represent a software our framework. We conclude that the Burr Type XII incorporation Burr Type XII testing-effort expenditure.

ACKNOWLEDGMENT

The authors would like to thank the referee for his valuable comments and suggestions that helped us to improve the presentation of the paper.

REFERENCES


Mohammad Ubaidullah Bakohri is presently working as an Associate Professor as well as Chairman in the Department of Computer Science, AMU, Aligarh, (INDIA) and Principal Investigator (PI) of the ambitious project, NMEICT ERP Mission Project (Govt. of India).

He graduated from the Magadh University, Bodh Gaya, India in 1985. He obtained his M.C.A. and Ph. D. degree in computer science from the Aligarh Muslim University in 1989 and 2006, respectively. He received the CWC merit scholarship (1986-1988) and University Grant Commission Scholarship (1988-1989). He had also worked as Associate Professor and Director of Studies in Australian Institute of Engineering & Technology, Victoria, Melbourne (Australia).

Dr. Bakohri has a vast teaching experience of more than twenty three years. His research interests include Software Reliability, Security and Cryptography, Software Engineering, Database Management System and E-Learning. He has published more than 70 research papers in reputed National and International Journals and Conferences and authored 5 books on different areas of computer science. He is lifetime member of Computer Society of India (C.S.I) and member of IEEE.

Nesar Ahmad is an Associate Professor in University Department of Statistics and Computer Applications at Tilka Manjhi Bhagalpur University, Bhagalpur, India. He received the B. Sc. degree in Mathematics form Bihar niversity, India, in 1984 and the M. Sc., M. Phil, and Ph. D. degree in Statistics from Aligarh Muslim University, Aligarh, India, in 1987, 1990, and 1993, respectively.

From 1995 to 1996 he was a research associate of UGC/CSIR at Aligarh Muslim University, Aligarh. He has been a Lecturer in Statistics from January 2006 to December 2009 at the University of the South Pacific, Suva, Fiji Islands. After
working five months as a lecturer at Poona College, Pune, he joined the University Department of Statistics and Computer Applications at Tilka Manjhi Bhagalpur University, Bhagalpur, India in December 1996. He has about 16 years of experience in teaching and research. His research interests include life testing, statistical modeling, reliability analysis, software reliability, software engineering and optimization. He is an author and coauthor of more than 40 journal papers and conference papers in these areas.

Dr. Ahmad is life member of Indian Science Congress and Bihar Journal of Mathematical Society, and regular member of Aligarh Statistical Association.