



ANALYSIS AND PREDICTION OF DEFECTS USING RELIABILITY GROWTH MODELS

Dr. Y. Prasanth

Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur, A.P, India

N. Vijaya Bharathi, B. Sai Tejaswini, P.Sai Krishna

Students, Department of Computer Science and Engineering,
Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur, A.P, India

ABSTRACT

In improvement of a software, software errors are unavoidable. For prediction of the defects in a software, we have to develop software reliability growth models to analyse the products and to predict them. The same growth model cannot be used for all the data. Hence various software reliability growth models are used for which the data is divided into several parts and different growth models are used for each data part. Later they are recombined. By combining all the traditional reliability growth models we get a multistage reliability growth model. At each stage, the finest Software Reliability Growth Model is to be applied. There are several criterions to evaluate the defects. The RSS ratio is used to select the best Reliability Growth Model. The Multi Stage Reliability Growth Model is a combination of many new-fangled models. These model shave to provide the software reliability and could support to solve problems which are hard to resolve by using Traditional software reliability growth models.

Key words: Software reliability, Software Reliability Growth Models, Defects, RSS Ratio, probit model.

Cite this Article: Dr. Y. Prasanth, N. Vijaya Bharathi, B. Sai Tejaswini, P. Sai Krishna, Analysis and Prediction of Defects Using Reliability Growth Models. *International Journal of Civil Engineering and Technology*, 9(1), 2018, pp. 22-27. <http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=9&IType=1>

1. INTRODUCTION

The vital part of software is Software Reliability because it guarantees quality. Software Reliability can be defined as the occurrence of no error functioning of a real system over a particular time in a particular environment for a particular reason. Because during software development errors are expected, it is of most significant to recommend a methodology to analyse and predict software and hardware reliability. For detecting the probable failures, reliability growth model is used as an indication and is the best solution. Since last several years, most companies used software growth models to improve software quality.

In certain cases defects get executed by the tester during testing then it results into a failure which is known as software failure. The software reliability growth models are divided into two classes, concave and S- shaped. The significance of both models is they have asymptotic behaviour i.e. the error prediction and detection rate reduces as the number of possible errors detected and predicted increases and the total number of errors and failures detected asymptotically reaches a finite value. Numbers of models were available for defect detection. One reliability growth will not fit the growth model well always.

Software Reliability Growth Models are used to detect and predict the number of failures before the software has been shipped to market. The reliability growth models use the data of system to detect and predict the number of failures in the software. The software reliability growth models to relate the failures in the set of code contained in the software use parameters. By knowing the parameter and the number of failures, we will know how many more defects are remaining in the code.

The residual defects helps us to decide whether the code is ready to be shipped and whether more testing is required if we decided the code is not ready to be shipped. The estimate of the number of defects that the software customers will encounter when operating and using the software.

To assess and forecast the software reliability of products, several software reliability growth models (SRGM's) are examined and used. Famous models are dependent on non-homogeneous Poisson process (NHPP), like the exponential Goel Okumoto model and s-shaped model. Gompertz model and a logistic model are also used as exponential reliability software growth models.

Presently OSS (Open Source Software) products are getting more accepted for both personal and business use. Since the use of OSS products, its developed body and process are expected to have more complexity. The pragmatic work determines an appliance of growth model curves to clarify and to foresee progress of open source system products. In spite of it, there will be a complexity left over to explain the total available of code modulation procedures with only one growth curve; it also tough to design the whole developing latest with one growth curve because of the difficult open source system developing method. So as to create a model that is casually and widely applicable, we use only growth curve data that are available from open code repositories and do not need any particular data method.

The development of software development will increase by developing a model curve. The technique here is to put a software reliability growth model to the actual data set that include error and bug information. Even though so many models are developed till now, a universal model not yet established.

2. LITERATURE SURVEY

Jieming Chi [1] said that a data can be fitted to both multistage model and software reliability growth models. They developed an evaluation and prediction tool for selecting the best reliability growth model. To evaluate the growth models they used AIC value and RSS value. To prove a good solution is provided by the multistage model AIC [2] [3] is used. This terminology provides a latest approach in software reliability growth models. The following is taken from [1] to show that RSS ratio is used to select the growth model. In [1] they compared the RSS ratio and Calculated ratio of Logistic and Gompertz Model. The RSS ratio of logistic model is 0.69 and that of gompertz model is 1.00. The calculated ratio of logistic is 85.67 and gompertz is 105.57.

The RSS value for Gompertz model is more compared to Logistic model. Hence Logistic model is the best reliability software growth model when evaluated with Gompertz model. This has been shown by Jieming Chi [1].

Atwood [4] mentioned that to specify the count of potential failures software reliability growth model is the excellent solution. Goel [5] observed that based on non-homogeneous poisson process (NHPP) a probabilistic model is presented for the software failure phenomenon. A better mean value function for NHPP is developed by analyzing the failure process and based on several performance measures expressions are given. Actual data of software failure is observed and compared. The wide range of software reliability growth models [5] are Goel model is the software reliability growth model which is represented as $p(t) = a(1-e^{-bt})$, where p is number of failures in time t . Many models have $p(t) = a F(t)$, where 'a' is the total number of defects in the code and $F(t)$ is cumulative distribution function. After a vigorous amount of software testing total number of defects 'a' is calculated. The defect finding rate decreases if the curve has concave shape. Many other software reliability growth models have S-shaped curves which means a large number of defects in the code. Different models are described in [5] and [6]. Logistic model is described by Tsoularis and J. Wallance [7]. They mentioned that to model the population growth the curve obtained is S-shaped curve. This model is widely applicable for scientific world and Gompertz curves. During the testing due S-shaped curve is obtained because of cumulative number of failures, a Logistic curve and/or Gompertz curve are fitted to actual data. This curve is used for predicting the number of remaining failures and for evaluating the testing process.

An S-shaped curve is obtained by using Gompertz model which is used to predict the number of potential failures. This growth curve model is applied to various areas of science and technology for developing different criterion [8]. A weibull curve is a curve used to model the growth of cumulative number of failure. Since weibull curve has a shape parameter this curve is more flexible than above two curves. Weibull curves are used for modelling a software reliability growth [9] [10]. Shigeru Yamada stated that as testing goes on the number of faults decreases which means there is a chance of decrease in software failure-occurrence so that time-interval becomes longer between the software failures and software reliability increases [11]. J. Mirosław Staron defined defect inflow as number of non-redundant defects. To create prediction models for defect inflow he used multivariate linear regression and also told that the accuracy of the prediction is as high as 72% of defects can be predicted [12].

Table 1 Reliability Models

Model	Formula
Gompertz curve	$f(T)=ae^{(-be)}$
Logistic curve	$dN/dT=rN(K-N/K)$
Weibull curve	$y(t)= a-be$

Table 1 [12] shows the comparison between various software reliability growth models.

3. WORK DONE

Here we had gone through all the growth models, their functions and the author who had proposed them and by which criteria they had got the curve in the form of either s-shaped or concave shaped and for log passion we had got infinite failure. From [1] here some growth models may be of 3 parameters and some would be of 4 parameters in the paper they had compared the relation between the 3 parameters and gave a conclusion that which growth model would better suit according to the given criteria. As of now we came to know that we

cannot perform task of comparing the 3 parameter and 4 parameter because we get the largest values. Here the point line is that “What if the largest value occurs?”. It means that in the software reliability, when we compare the values of growth models or individually we need to get the less value; so we can prove that the software that we had developed is an effective one. The term reliability means the probability of no failure software operation for a specified period of time in a specified environment. So that we need to prove, that the failure that would occur in a time and that to at low rate rather than if we get the more value that does not mean the growth model for this particular case is apt then we need to go through the other growth model. So as we thought of this and came to a conclusion that we need to compare the 4 parameter growth models and relation between them RSS ration and calculated value.

We had taken two growth models with 4 parameters Weibull and Probit model and calculated RSS ration and calculated data.

Table 2 Comparison of models based on RSS ratio

	Weibull	Probit
RSS Ratio	1.00	0.765
Calculated Data	121.844	119.972

Here going through Internet search we found the comparison between them and then we made it into the tabular form so that we understand the comparison between them. We need to tell what this probit growth model is and how to estimate it. In probit model, the value of $X\beta$ is put as the z-value of a normal distribution in the graph larger values of $X\beta$ is that the event is more likely to happen according to the graph. They have mentioned that to be cautious about the interpretation of estimation results herein the graph. A one single unit change in X_i causes to a β_i change in the z-score of Y. The estimated curve is an S-shaped cumulative normal distribution. For this we had got the graph of S-shaped.

Whereas for Weibull we had got the concave shaped graph here we need to know that what does these 4parameter and 3parameter means. It means that for every growth model we have formulae for it. There we have 3 parameter and 4 parameters used in the formula means 4 different values in the formula or 3. Based on it we need to compare the growth model. For the probit we had and got graph representation as Fig 1.

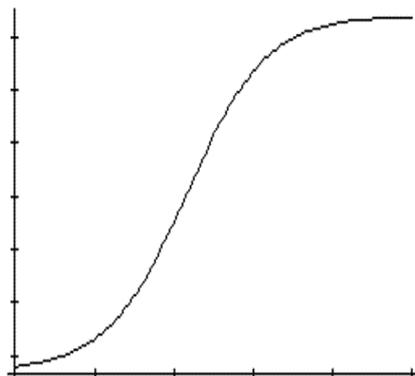


Figure 1 Graphical representation of weibull is s-shaped [13]

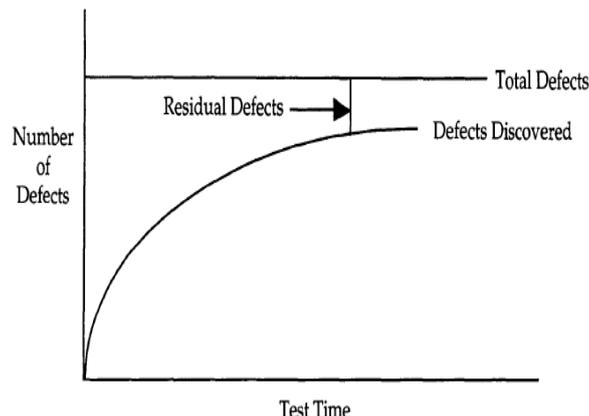


Figure 2 Graphical representation of the weibull and its shape is concave [13]

We take time on X-axis and No. of issues on Y-axis so that we get a graph as Fig 2.

4. OBSERVATION

	Weibull	Probit
RSS Ratio	From the given graphs (Fig 1 and Fig 2) we compared that weibull has the largest value.	Here in defects we get probit RSS value as the lowest value when compared to weibull.
Calculated value	The calculated data is more for weibull when compared to probit.	Here we get the less value when compares to weibull. In this comparison it is better to take probit.

5. CONCLUSIONS

Here we had taken the comparison on 4 parameter growth models they are WEIBULL and PROBIT and we came to know that PROBIT gives the less value and we can say that when compared both we can use this PROBIT growth models as it would be effective one when compared to the defects that would to be found/occurred.

In future we planned to find the RSS ratio for each it could be formed through the code written to it and so that we would get the comparison between them and according to what bases and which growth model would be better and according to defects, failure, risk and error we would form a tabular representation for all that occur.

REFERENCES

- [1] Jieming Chi” Defect Analysis and Prediction by applying multistage reliability growth Models s,” Empirical software Engineering in Practice (IWESEP), 8thInternational Workshop Mar-2017.
- [2] H. Akaike, “A new look at the statistical model identification,” IEEE Trans AutomaticControl, vol19, no. 6 pp716-723, 1974.
- [3] H.Akaike, “Information theory and an extension of the maximum likelihood principle.” In 2nd International Symposium on Information Theory, Petrov, B.N., and Caski, F. (eds), Budapest: Akadimiai Kiado, pp.267-281, 1973.
- [4] A. Wood, “Software reliability growth models,” Technical Report96.1, 1996, pp. 1.

- [5] A. Goel, "Software reliability models: Assumptions, limitations, and applicability," *Software Engineering, IEEE Transactions on*, vol. SE11, no. 12, pp. 1411-143, Dec 1985.
- [6] Musa, John, Anthony Iannino, and Kazuhira Okumoto, *Software Reliability*, McGraw-Hill, 1987.
- [7] A. Tsoularis and J. Wallace, "Analysis of logistic growth models," *Mathematical Biosciences*, vol.179, no. 1, pp21-55, 2002.
- [8] C. P. Winsor, "The Gompertz curve as a growth curve," in *Proc. National Academy of Sc.*, Vol. 18, no. 1, Jan.1932, pp. 1-8.
- [9] S. Yamada, J. Hishitani, and S. Osaki, "Software reliability growth with a Weibull test Effort: a model and application," *IEEE Trans. Reliability*, vol. 42, no. 1, pp.100-106, Mar. 1993.
- [10] N. Ahmad, M. Bokhari, S. Quadri, and M. G. Khan, "The exponentiated Weibull Software reliability growth model with various testing-efforts and optimal release Policy: a performance analysis," *Int'l Journal of Quality and Reliability Management* Vol. 25, no. 2, pp. 211-235, Feb. 2008.
- [11] Yamada, "Software Reliability Models and Their Applications: A Survey", August 17-18, 2000 Kyoto University, Kyoto, Japan.
- [12] Mirosław Stalón, "Predicting weekly defect inflow in large software projects", September 25-09-2007.
- [13] www.columbia.edu/~so33/SusDev/Lecture_9.pdf.
- [14] Geetika Batra and Kuntal Barua, "An Optimization of Effort and Cost Estimation by Code Reusability Concept", *International Journal of Computer Engineering & Technology (IJCET)*, Volume 4, Issue 5, 2013, pp. 251 - 257
- [15] Gulwatanpreet Singh, Surbhi Gupta and Baldeep Singh, "ACO Based Solution for TSP Model for Evaluation of Software Test Suite", *International Journal of Computer Engineering & Technology (IJCET)*, Volume 3, Issue 3, 2012, pp. 75 - 82