Integrated Design and Process Technology, **IDPT-2002**
Printed in the United States of America, June, 2002

"SAHINOGLU" SOFTWARE RELIABILITY MODEL- JPL DATA

http://www.areslimited.com

ABSTRACT

The proposed Compound Poisson Non-Linear Regression (CPNLR) estimation model, henceforth entitled "Sahinoglu" Software Reliability Model for failure-count prediction in clustered data sets is revisited in this research paper. The JPL software failure data are used to illustrate the benefits of this prediction model. Diagnostic checks are also outlined for a data set to qualify or not to qualify for this model.

INTRODUCTION

Clustered data are frequently encountered in software testing practice, such as in the telecommunications world or aeronautical / space research, where testing is carried on for example, in units of days, weeks or months. Such results have been observed in the Bellcore and Jet Propulsion software testing laboratories. (Dallal and MacIntosh, 1994; Lyu, 1996). The CPNLR prediction model was earlier published in greater detail (Sahinoglu and Can, 1997). Due to non-influencing phenomenon of the multiple failures found clumped within each sampling time-unit, the compounding distribution for the Poisson process was elected to be the geometric distribution following a series of goodness-of-fit studies. The resulting distribution is thus called Poisson ^ Geometric. Note that the discrete geometric distribution is a discrete analogue of the continuous negative-exponential probability density function. Five varying JPL data are studied in terms of their compatibility with the proposed Sahinoglu failure-count prediction model. After the diagnostic checks, the Sahinoglu model is compared favorably to Musa-Okumoto in terms of Average Relative Error(ARE), Mean Square Error (MSE) and Kolmogorov-Smirnov (K-S). Note that the conventional non-homogeneous Poisson process models (NHPP) do not permit the option of multiple counts and Sahinoglu's CPNLR model is superior when clumping exists (Sahinoglu, 1992: Xie, 1993)

MATHEMATICAL FORMULATION

The proposed "Sahinoglu" Compound Poisson (CP) reliability model suggested that the expected number of remaining software failures occurring within the next remaining time interval $[t, t_{rem}]$ was as follows (Sahinoglu, 1992; Randolph and Sahinoglu, 1995):

$$E\{ X (t_{rem)} \}=\{ \hat{\beta} /(1-\hat{r}) \}t_{rem} + \text{error} \quad (1)$$

Namely, the estimate of the quotient in (1) multiplied by the remaining time units in CPU hours, days or weeks will estimate the expected number of remaining future failures. Note, β of the Poisson process is the average number of (failure) arrivals per unit time and r is the probability of finding the next failure in a batch or clump or cluster (e.g. week) following that arrival. Then, p=1-r is the probability of starting the Poisson process for the next failure arrival. If r=0, or p=1, then the CP defaults to a purely Poisson process. The remaining failure-count prediction is further added to the number of failures in the past to finally estimate the number of total failures at the end of a mission time. Consequently, (2) is converted to a non-linear regression equation in (3).

$$X_{total}= X_{past} + \{ \hat{\beta} / (1-\hat{r}) \}t_{rem} + \text{error} \quad (2)$$

$$X_{past}= X_{total} - \{ \hat{\beta} / (1-\hat{r}) \}t_{rem} + \text{error} \quad (3)$$

The Levenberg-Marquardt (L-M) algorithm is the most popularly used algorithm employed to solve for the unknown parameter vector (β , r) in the nonlinear regression equation by means of least squares estimation (Kennedy and Gentle, 1980). The JPL data sets to be used for the application by Sahinoglu model will be described in the following subsection.

RESULTS AND CONCLUSIONS

It is appropriate to describe briefly the weekly data (WD) sets, renamed WD1 To WD5 for simplicity. These sets were actually time-based simulated at the Jet Propulsion Laboratory (JPL) under the code-names GALILEO and ALASKA etc. (M. Lyu, 1996, Sahinoglu, 1997) The WD1 corresponds to a 60 week-long software and has a total of 131 accumulated failures at the termination of the testing activity. WD2: 213 failures in 223 weeks, WD3: 340 failures in 41 weeks, WD4: 197 failures in 114 weeks and WD5: 366 failures in 50 weeks. Additionally, when a sequence of goodness of fit tests are applied, the frequency histogram of number of weeks vs. clump size displays a negative-exponential like plot for WD1, WD2, WD4 and WD5. This is a sign of geometric distribution hypothesis not being rejected (come true), as Geometric is discrete analogue of negative exponential. Concerning WD3, which depicts a quasi-uniform frequency plot, clearly non-exponential and very different (Sahinoglu and Can; 1977). The diagnostic check reports that Sahinoglu's CPNLR method is compatible with the WD sets except for WD3, which is best compatible with a logarithmic-Poisson estimation method as in Musa-Okumoto (Sahinoglu and Can, 1997; Lyu, 1996).

Due to space limitation, rather than five different particular analyses on five data sets, a summary of results will be presented in terms of their goodness of fit measures, such as ARE and MSE over 19 data inspection points from 5 to 100% respectively. However, WD1 is studied in detail to apply the Sahinoglu's CPNLR prediction model, and a detailed analysis is given in the Appendix including the diagnostic checks.

Table 1. Comparisons of Models for WD1-WD5

	CPNLR(Sahinoglu)			*M-O(Musa-Okumoto)*	
	ARE	**MSE**	X_{true}	**ARE**	**MSE**
WD1	.069	160	131	.320	3247
WD2	.316	6465	213	.539	18902
WD3	.454	31869	340	.213	13289
WD4	.144	1795	197	.449	11906
WD5	.160	6200	366	.276	27245

where,

$$ARE = (1/n) \sum \{ABS(X_{total} - X_{true})\} / X_{true} \quad (4)$$

$$MSE= [1/(n-1)] \sum (X_{total} - X_{true})^2 / (n-1) \quad (5)$$

One can easily observe from the Table 1 above that CPNLR (Sahinoglu) is better fitted for WD1, WD2, WD4 and WD5 as it was clear from our diagnostic checks. M-O (Musa-Okumoto) is better for WD3, whose clump size frequency plot does not display a negative-exponential (or geometric) shape, but a uniform.

Frequency plots for the clump sizes and the comparisons of AREs for WD1are given in the Appendix. Table II in the Appendix is tabulated by using an advanced statistical package.

REFERENCES

Dallal S. R., and MacIntosh, A. A. 1994, "When to Stop Testing for Large Software Systems with Changing Code," .IEEE-TSE, Vol. 20(4), 318-323

Randolph P. and Sahinoglu M., 1995, "A Stopping Rule for a Compound Poisson Random Variable," Applied Stochastic Models and Data Analysis, Vol. 11, 135-143

Xie M., 1993, "Software reliability Models,- A Selected Annotated Bibliography", Journal of STVR, Vol. 3, No.1, 3-28

Kennedy, W. J. and Gentle J. E., 1980, Statistical Computing, Marcel Dekker, New York, U.S.A.

Lyu M., Handbook of Software Reliability Engineering, Mc Graw Hill, 1996

Sahinoglu M., 1992, "Compound Poisson Software Reliability Model," IEEE Trans. Software Engineering, Vol. 18, 624-630

Sahinoglu M. and Can U., 1997, "Alternative Parameter Estimation Methods for the Compound Poisson Software Reliability Model with Clustered Failure Data," Software Testing Verification and Reliability, Vol. 7, 35-57

Sahinoglu M., Deely J., Capar S., 2001," Stochastic Bayes Measures to Compare Forecast Accuracy of Software Reliability Models," IEEE Trans.Reliability, Vol. 18, No.7, 92-97

APPENDIX

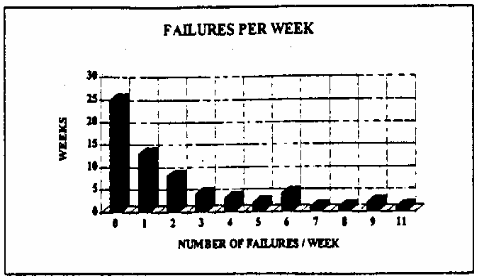


Figure 6. Failures per week for data set WD1.

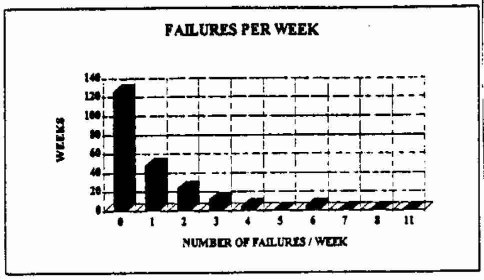


Figure 7. Failures per week for data set WD2.

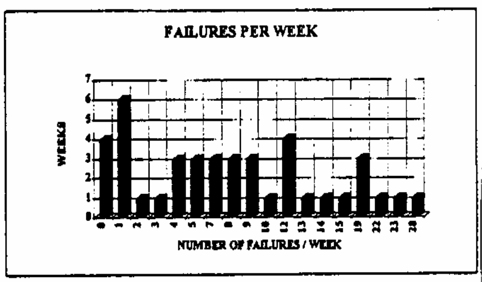


Figure 8. Failures per week for data set WD3.

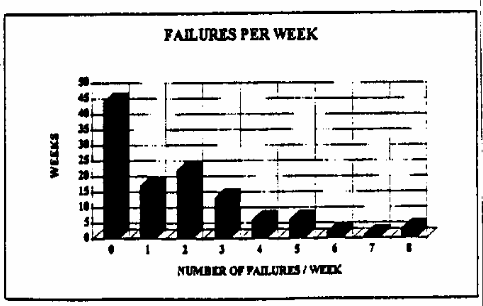
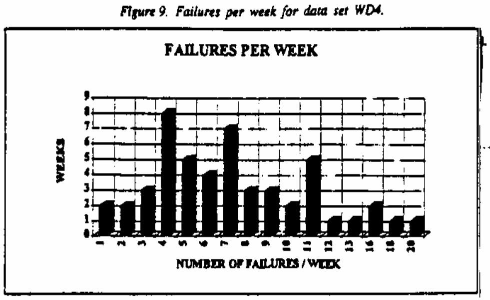


Figure 9. Failures per week for data set WD4.



Fieure 10. Failures per week for data set WD5.

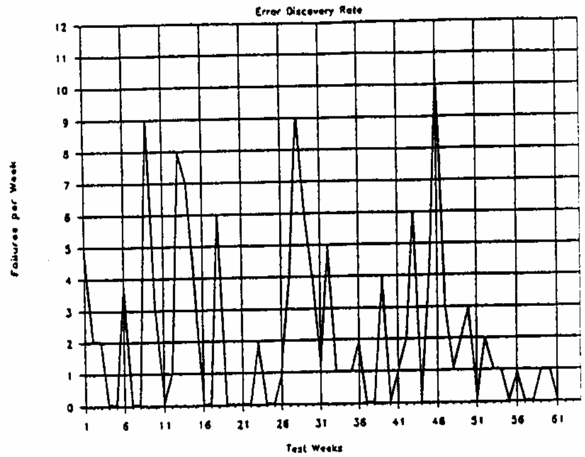
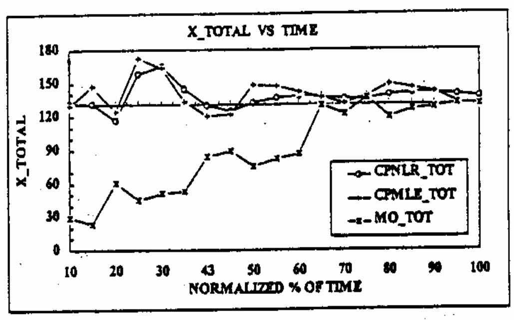


Figure I. Data set WD1: (a) failures per calendar week (continued)



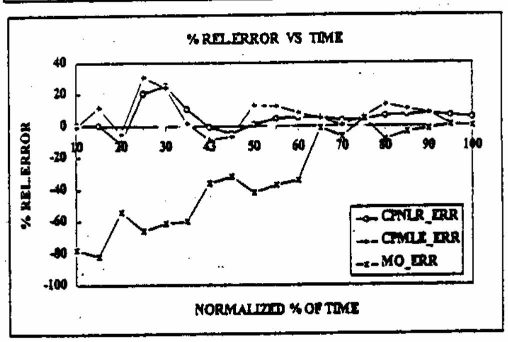


Figure I. Data set WD1: (b) X_{tot} versus time; (c) percentage relative error versus time.

Table II. CPNLR parameter estimation results for data set WD1 (in weeks)

Time (%)	X_{past}	Y_{past}	t_{past}	t_{rem}	$1-r$	β	X_{tot}	rel. error (%)
10	13	4	6	54	0·038806	0·084080	131·67	0·51
15	22	5	9	51	0·038806	0·084080	131·67	0·51
20	25	7	12	48	0·004193	0·007966	117·	−10·69
25	43	10	15	45	0·002365	0·006307	158·4	20·92
30	49	11	18	42	0·000101	0·000281	164·96	25·93
38·3	51	12	23	37	0·000281	0·000673	144·64	10·41
43·3	52	13	26	34	0·030508	0·064161	129·74	−0·96
45	55	14	27	33	0·113466	0·227077	124·61	−4·88
50	74	17	30	30	0·000050	0·000107	132·23	0·94
55	81	20	33	27	0·086373	0·194792	136·99	4·57
60	85	23	36	24	0·004136	0·009436	138·22	5·51
65	89	24	39	21	0·049780	0·112873	137·58	5·02
70	92	26	42	18	0·156972	0·348929	135·58	3·5
75	104	28	45	15	0·048107	0·107635	136·22	3·98
80	119	31	48	12	0·002712	0·006260	139·23	6·28
86·7	126	34	52	8	0·001313	0·003076	140·68	7·39
90	128	36	54	6	0·008521	0·020031	141·01	7·64
98·3	130	38	59	1	0·000014	0·000031	139·15	6·22
100	131	39	60	0	0·002828	0·006406	137·68	5·10