International Journal of Statistics and Probability; Vol. 8, No. 1; January 2019
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Characteristics and Application of the NHPP Log-Logistic Reliability Model

Lutfiah Ismail Al turk

Correspondence: Lutfiah Ismail Al turk, Statistics Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia. E-mail: lturk@kau.edu.sa

Received: November 1, 2018 Accepted: November 23, 2018 Online Published: December 4, 2018

doi:10.5539/ijsp.v8n1p44 URL: https://doi.org/10.5539/ijsp.v8n1p44

Abstract

In this paper, a Nonhomogeneous Poisson Process (NHPP) reliability model based on the two-parameter Log-Logistic (LL) distribution is considered. The essential model's characteristics are derived and represented graphically. The parameters of the model are estimated by the Maximum Likelihood (ML) and Non-linear Least Square (NLS) estimation methods for the case of time domain data. An application to show the flexibility of the considered model are conducted based on five real data sets and using three evaluation criteria. We hope this model will help as an alternative model to other useful reliability models for describing real data in reliability engineering area.

Keywords: NHPP log-logistic model, maximum likelihood estimation, non-linear least square estimation, reliability assessment

1. Introduction

The Log-Logistic (LL) distribution that results from a simple transformation of the familiar logistic distribution has been found useful in many areas such as engineering, reliability data analysis, economics and hydrology. In the literature, it is well-known as the Fisk distribution due to (Fisk, 1961). In some cases, the LL distribution is proved to be a good alternative to the log-normal distribution since it characterizes increasing hazard rate function. Further, its use is well appreciated in case of censored data that usually common in reliability and life-testing experiments. The Cumulative Distribution Function (CDF) and Probability Density Function (PDF) of the two-parameter LL distribution can be defined respectively as follows:

$$\mathrm{F}(\mathrm{t_i};\boldsymbol{\gamma},\beta)=\frac{\gamma \mathrm{t_i}^{\beta}}{1+\gamma \mathrm{t_i}^{\beta}}, \tag{1}$$

and

$$f(\mathrm{t_i};\boldsymbol{\gamma},\beta)=\frac{\gamma\beta \mathrm{t_i}^{\beta-1}}{\left(1+\gamma \mathrm{t_i}^{\beta}\right)^{2}}, \tag{2}$$

where $\gamma > 0$ is the scale parameter, and $\beta > 0$ is the shape parameter. The log-logistic reliability growth model is quite flexible to analyze reliability data since it can capture increasing/decreasing nature of the failure occurrence rate per fault. This property has attracted more attention of researchers. (Gokhale and Trivedi, 1998) considered the log-logistic reliability growth model. The Maximum Likelihood (ML) estimation method of several existing finite-failure NHPP models, as well as the log-logistic model was conducted based on inter-failure times data. They presented analysis using two real data sets which encouraged the development of the log-logistic model. (Harishchandra, 2016) considered a software reliability model in which time between two successive failures is assumed to follow the log-logistic distribution. The parameters of their model were estimated using the ML method in the cases of interval domain data and time domain data. A simulation study and real data were used to examine their model. The results showed that their considered model performs better compared to previously suggested four NHPP models. In this paper, Nonhomogeneous Poisson Process Log-Logistic (NHPP LL) model is considered. Reliability characteristics of this model including: intensity function, number of remaining errors function, error detection rate function, instantaneous and cumulative mean time between failure function, and conditional reliability function are provided and represented graphically. The estimation of the model parameters is performed by the ML and Nonlinear Least Square (NLS) estimation methods. The flexibility of the new model is illustrated by means of an application to five real data sets. We hope that the new model will help as an alternative model to other useful models for representing positive real data in many areas.

The rest of the paper is organized as follows. In Section 2, we define the NHPP LL model and provide mathematical formulas and plots of its reliability characteristics. Estimation by the method of the ML and NLS methods is presented in Section 3. Evaluation criteria is presented in Section 4. An application to a real data set illustrates the flexibility of the NHPP LL model is given in Section 5. Conclusions are presented in Section 6.

2. NHPP Log-Logistic Reliability Growth Model

One way to model software failure phenomena is Non-Homogeneous Poisson process (NHPP) family of models with Mean Value Function (MVF) at time t_i, $m(t_i)$. The derivative of the MVF is the failure intensity, $h(t_i)$, of the software which ordinarily decreases as faults are detected and removed. If F(t) is the distribution function that denotes the expected number of faults that would be detected in a given infinite testing time, then the MVF as presented in (Lyu, 2002) is as follows:

$$m(t_i) = NF(t_i) \tag{3}$$

By inserting Eq.(1) in Eq.(3), we obtain the MVF of the NHPP LL model as follows:

$$m(t_i) = \frac{N\gamma {t_i}^{\beta}}{1+\gamma {t_i}^{\beta}}, \quad N, \gamma, \beta > 0, \tag{4}$$

where t_i, $i = (1, 2, \dots, n)$ is the failure times, N is the number of initial errors, γ is positive scale parameter, and β is shape parameter.

The failure intensity function corresponding to (4) is defined as:

$$h(t_i) = \frac{\partial m(t_i)}{\partial t}$$

$$= \frac{N\gamma\beta {t_i}^{\beta-1}}{\left(1+\gamma {t_i}^{\beta}\right)^2}, \tag{5}$$

while the constructed model's number of remaining errors function is given by:

$$n(t_i) = N - m(t_i)$$

$$= \frac{N}{1+\gamma {t_i}^{\beta}}, \tag{6}$$

also, its error detection rate function is given as follows:

$$d(t_i) = \frac{h(t_i)}{n(t_i)}$$

$$= \frac{\gamma\beta {t_i}^{\beta-1}}{1+\gamma {t_i}^{\beta}}, \tag{7}$$

Additionally, the instantaneous mean time between failures (MTBF) can be found by the inverse of the intensity function:

$$\text{MTBF}_I(t_i) = \frac{1}{h(t_i)}$$

$$= \frac{\left(1+\gamma {t_i}^{\beta}\right)^2}{N\gamma\beta {t_i}^{\beta-1}} \tag{8}$$

while the cumulative MTBF can be calculated by:

$$\text{MTBF}_c(t_i) = \frac{t_i}{m(t_i)}$$

$$= \frac{t_i\left(1+\gamma {t_i}^{\beta}\right)}{N\gamma {t_i}^{\beta}}, \tag{9}$$

lastly, we have the conditional reliability function as follows:

$$R(t_i|x_n) = \exp\left\{-\left(\mu(t_i + x) - \mu(t_I)\right)\right\}$$

$$= exp\left\{-N\gamma\left(\frac{(t_i+x_n)^{\beta}-x_n^{\beta}}{\left(1+\gamma x_n^{\beta}\right)\left(1+\gamma(t+x_n)^{\beta}\right)}\right)\right\} \tag{10}$$

All the above reliability characteristics of the NHPP LL model are summarized in Table 1. While, Figures [1-7] show plots of the NHPP LL model's characteristic for different selected values of parameters. Figure 1 displays that the intensity function varies in shape over the different selected values of the shape parameter, it reaches a larger peak level with the larger value of the parameter N. Figure 2 illustrates the MVF which represents the variation of number of faults detected with respect to time. From this figure we can see that, initially the faults detected during testing are very high but later on become constant, also larger value of the parameter N gives higher MVF form. The number of remaining errors function in Figure 3 decreases as the testing time increases, smaller value of the parameter N gives lower form of the number of remaining errors function. Figure 4 shows the effect of different values of the parameters on the error detection rate function, when the shape parameter is less than or equal to 1 the error detection rate function is declining exponentially, while the error detection rate function is increasing at the beginning before start declining when the shape parameter is greater than1. In Figure 5, the conditional reliability function shows a decrease form with the progress of time, the sharpness of the decreasing varies according to the variation in the selected parameters' values, larger value of the parameter N gives lower reliability form. The instantaneous and cumulative MTBF functions in Figure 6 and Figure 7, respectively, either increase rapidly with the progress of testing time or show an initial decrease before start increasing, in both cases larger value of the parameter N gives lower MTBF form.

Table 1. Listing of the NHPP LL model's characteristics.

Characteristic name	Characteristic function
Mean value function (MVF).	$m(\mathrm{t_i}) = \frac{N\gamma \mathrm{t_i}^{\beta}}{1+\gamma \mathrm{t_i}^{\beta}}.$
Intensity function.	$\mathrm{h}(\mathrm{t_i}) = \frac{N\gamma\beta \mathrm{t_i}^{\beta-1}}{\left(1+\gamma \mathrm{t_i}^{\beta}\right)^2}.$
Number of remaining errors function (NRE).	$\mathrm{n}(\mathrm{t_i}) = \frac{N}{1+\gamma \mathrm{t_i}^{\beta}}.$
Error detection rate (EDR).	$\mathrm{d}(\mathrm{t_i}) = \frac{\gamma\beta \mathrm{t_i}^{\beta-1}}{1+\gamma \mathrm{t_i}^{\beta}}.$
Instantaneous mean time between failures (I-MTBF).	$\mathrm{MTBF_I}(\mathrm{t_i}) = \frac{\left(1+\gamma \mathrm{t_i}^{\beta}\right)^2}{N\gamma\beta \mathrm{t_i}^{\beta-1}}.$
Cumulative mean time between failures (C-MTBF).	$\mathrm{MTBF_c}(\mathrm{t_i}) = \frac{\mathrm{t_i}\left(1+\gamma \mathrm{t_i}^{\beta}\right)}{N\gamma \mathrm{t_i}^{\beta}}.$
Conditional reliability function.	$\mathrm{R}(\mathrm{t_i}\vert \mathrm{x_n}) = \exp\left\{-N\gamma\left(\frac{(\mathrm{t_i}+\mathrm{x_n})^{\beta}-\mathrm{x_n^{\beta}}}{\left(1+\gamma \mathrm{x_n^{\beta}}\right)(1+\gamma(\mathrm{t}+\mathrm{x_n})^{\beta})}\right)\right\}.$

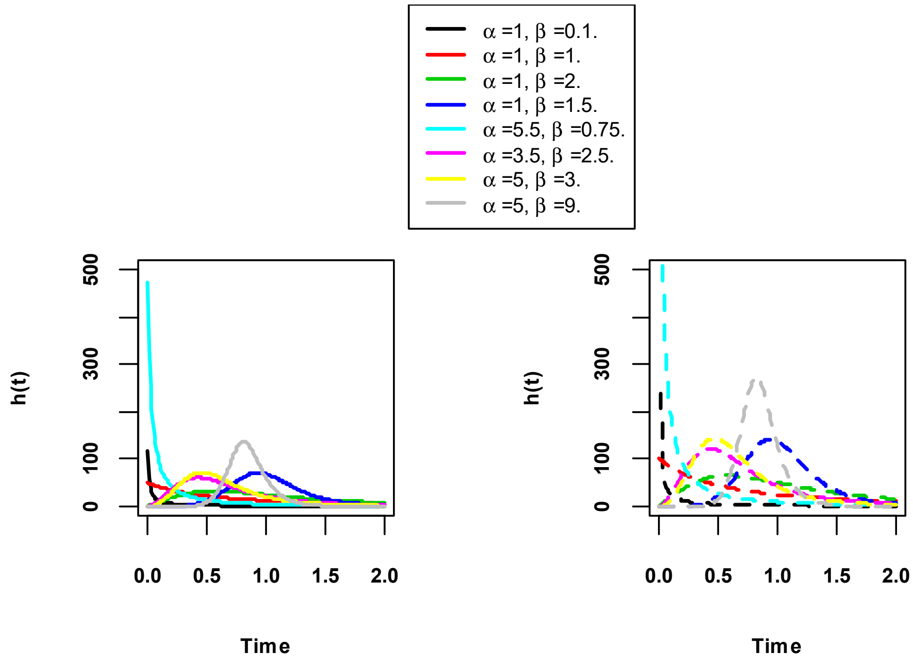

Figure 1. Plots of the NHPP LL model's intensity function for some selected values of parameters (Solid lines indicate N=50 and dashed lines indicate N=100)

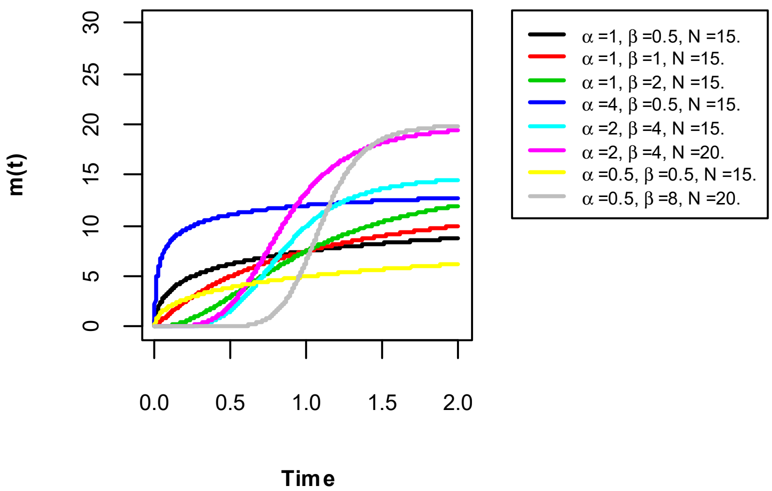

Figure 2. Plots of the NHPP LL model's MVF for some selected values of parameters

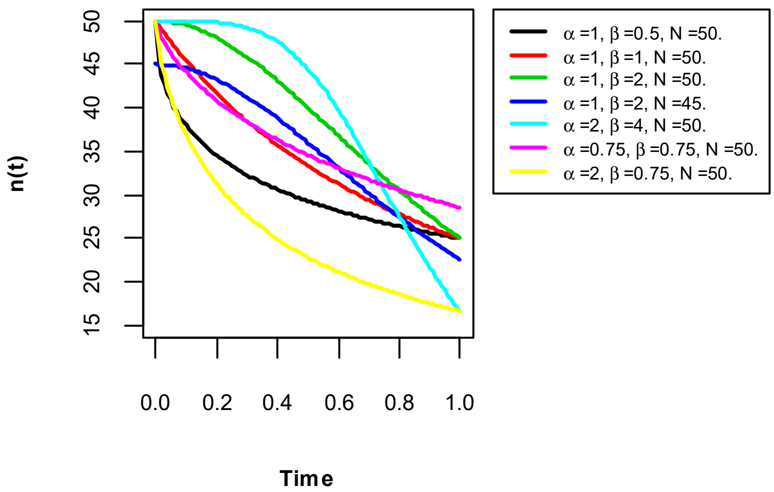

Figure 3. Plots of the NHPP LL model's NRE function for some selected values of parameters

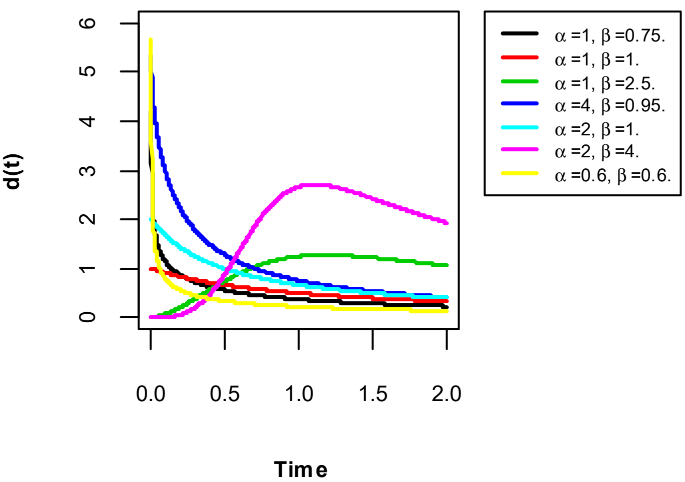

Figure 4. Plots of the NHPP LL model's EDR function for some selected values of parameters

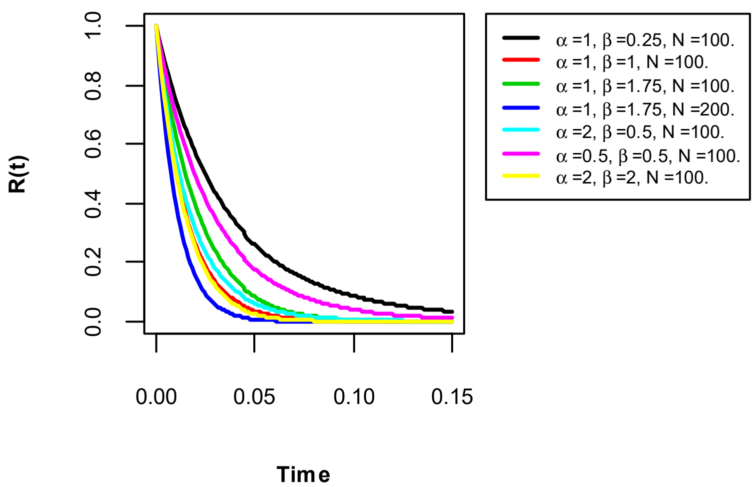

Figure 5. Plots of the NHPP LL model's reliability function for some selected values of parameters

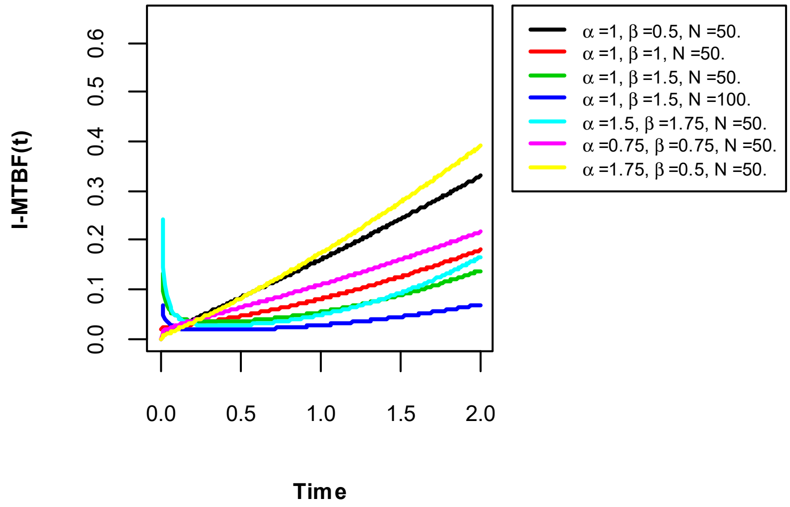

Figure 6. Plots of the NHPP LL model's I-MTBF function for some selected values of parameters

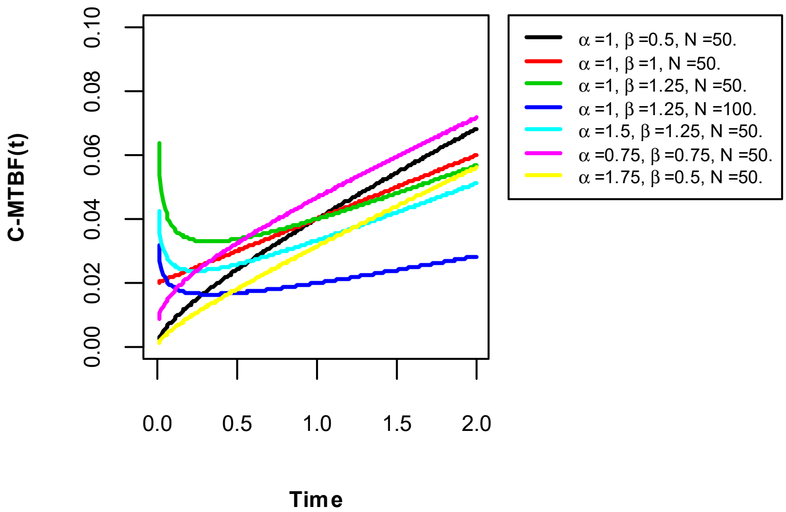

Figure 7. Plots of the NHPP LL model's C-MTBF function for some selected values of parameters

3. Estimation of the NHPP LL Model's Parameters

Two commonly used methods for parameter estimation in a NHPP model are the Maximum Likelihood (ML) and Least Squares (LS) estimation methods (Knafl, 1992; Lyu, 1996; Zhao and Xie, 1996; Musa, 1999; Chang, 2001; Prasad et al., 2011; Rana et al., 2013; Zeephongsekul et al., 2016). In this section, the ML and NLS estimation methods will be applied to the NHPP LL model.

3.1 Maximum Likelihood Estimation

Suppose that we have n observations, denoted by $t_1, t_2, \dots, t_n$, then the likelihood function of the NHPP model can be written as follows:

$$L(\Theta|\underline{S}) = e^{-m(t_i)} \prod_{i=1}^{n} \lambda(t_i) \tag{11}$$

where Θ is the NHPP model's parameters, $m(t_i)$ and $\lambda(t_i)$ are, respectively, the NHPP model's mean value and intensity functions.

To simplify the mathematical computations, we take the natural logarithm of both sides of Eq.(11):

$$\ln L(\Theta|\underline{S}) = -m(t_i) + \sum_{i=1}^{n} \ln \lambda(t_i). \tag{12}$$

By substituting Eqs.(4) and (5) in Eq.(12), the log-likelihood function of the NHPP LL model can be written as:

$$\ln \mathrm{L}\left(\mathrm{N}_0, \boldsymbol{\gamma}, \beta \middle| \underline{\mathrm{S}}\right) = -\frac{N\gamma S_n^{\beta}}{1+\gamma S_n^{\beta}} + \sum_{i=1}^{n} \ln\left(\frac{\mathrm{N}\gamma\beta S_i^{\beta-1}}{\left(1+\gamma S_i^{\beta}\right)^2}\right) \tag{13}$$

$$= -\frac{N\gamma S_n^{\beta}}{1+\gamma S_n^{\beta}} + n\,ln\,\gamma + n\,ln\,\beta + n\,ln\,N + \beta\sum_{i=1}^{n} \ln s_i - \sum_{i=1}^{n} \ln s_i - 2\sum_{i=1}^{n} \ln\left(1+\gamma S_i^{\beta}\right) \tag{14}$$

Differentiating Eq. (14) with respect to N, γ and β, we have:

$$\begin{cases} \dfrac{\partial \ln \mathrm{L}(\mathrm{N}_0,\boldsymbol{\gamma},\beta|\underline{\mathrm{S}})}{\partial N} = -\dfrac{\gamma S_n^{\beta}}{1+\gamma S_n^{\beta}} + \dfrac{n}{N} \\ \dfrac{\partial \ln \mathrm{L}(\mathrm{N}_0,\boldsymbol{\gamma},\beta|\underline{\mathrm{S}})}{\partial \gamma} = \dfrac{n}{\gamma} - \dfrac{N S_n^{\beta}}{1+\gamma S_n^{\beta}} + 2\sum_{i=1}^{n} \dfrac{S_i^{\beta}}{1+\gamma S_i^{\beta}} \\ \dfrac{\partial \ln \mathrm{L}(\mathrm{N}_0,\boldsymbol{\gamma},\beta|\underline{\mathrm{S}})}{\partial \beta} = \dfrac{n}{\beta} + \sum_{i=1}^{n} \ln s_i - \dfrac{N\gamma S_n^{\beta} \ln S_n}{\left(1+\gamma S_n^{\beta}\right)^2} + 2\sum_{i=1}^{n} \dfrac{\gamma S_i^{\beta} \ln S_i}{1+\gamma S_i^{\beta}} \end{cases} \tag{15}$$

Setting the three expressions of Eq.(15) to zero we get the following system of equations:

$$\begin{cases} N = n\left(\dfrac{1+\gamma S_n^{\beta}}{\gamma S_n^{\beta}}\right) \\ \dfrac{n}{\gamma} - \dfrac{n}{\gamma\left(1+\gamma S_n^{\beta}\right)} + 2\sum_{i=1}^{n} \dfrac{S_i^{\beta}}{1+\gamma S_i^{\beta}} = 0 \\ \dfrac{n}{\beta} + \sum_{i=1}^{n} \ln s_i - \dfrac{n \ln S_n}{1+\gamma S_n^{\beta}} + 2\sum_{i=1}^{n} \dfrac{\gamma S_i^{\beta} \ln S_i}{1+\gamma S_i^{\beta}} = 0 \end{cases} \tag{16}$$

The second and third expressions of Eq.(16) do not have a closed-form so we need numerical methods to obtain the ML estimates of the parameters γ and β, then by substituting $\hat{\gamma}$ and $\hat{\beta}$ in the first expression, $\widehat{N}$ can be obtained.

3.2 Nonlinear Least Squares Estimation

Assuming $(\mathrm{t}_1, \mathrm{y}_1), (\mathrm{t}_2, \mathrm{y}_2), \ldots, (\mathrm{t}_\mathrm{n}, \mathrm{y}_\mathrm{n})$ are n pairs of observations where $\mathrm{i} = 1, \ldots, \mathrm{n}$. The model to be fitted to these data is:

$$y_i = f(t_i, \theta) + \varepsilon_i, \tag{17}$$

where θ is the parameter vector, and ε_i is the error term. In statistics theory ε_i is assumed as independent variables of normal distribution $\mathrm{N}(0, \sigma^2)$, where σ^2: is the variance of the normal distribution. The NLS estimation method involves in determining the value of the unknown parameters that minimizes:

$$\psi_{NLS} = \sum_{i=1}^{n} [y_i - f(t_i, \theta)]^2 . \tag{18}$$

By substituting Eq.(4), our considered fitting function, in Eq.(18), the NLS estimates of the NHPP LL model's parameters are obtained by minimizing:

$$\psi_{NLS}(N, \boldsymbol{\gamma}, \beta) = \sum_{i=1}^{n} \left[\mathrm{y_i} - \frac{N\gamma t_i^{\beta}}{1+\gamma t_i^{\beta}}\right]^2 . \tag{19}$$

Differentiating Eq. (19) with respect to $N, \gamma, and\ \beta$ then equating the resulted equations to zero subsequently yields the following system of equations:

$$\mathrm{N} = \sum_{i=1}^{n} \frac{\mathrm{y_i}\, t_i^{\beta}}{1+\gamma t_i^{\beta}} \Big/ \gamma \sum_{i=1}^{n} \left(\frac{\mathrm{y_i}\, t_i^{\beta}}{1+\gamma t_i^{\beta}}\right)^2 . \tag{20}$$

$$\frac{\partial s_{NLS}(N_0, \boldsymbol{\gamma}, \beta)}{\partial \gamma} = -\sum_{i=1}^{n} \frac{y_i\, t_i^{\beta}}{\left(1+\gamma t_i^{\beta}\right)^2} + N\gamma \sum_{i=1}^{n} \frac{t_i^{2\beta}}{\left(1+\gamma t_i^{\beta}\right)^3} = 0. \tag{21}$$

$$\frac{\partial s_{NLS}(N_0, \boldsymbol{\gamma}, \beta)}{\partial \beta} = -\sum_{i=1}^{n} \frac{y_i\, t_i^{\beta}\, ln\, t_i}{\left(1+\gamma t_i^{\beta}\right)^2} + N\gamma \sum_{i=1}^{n} \frac{t_i^{2\beta}}{\left(1+\gamma t_i^{\beta}\right)^3} ln\, t_i = 0. \tag{22}$$

The estimates of the parameters $\gamma\ and\ \beta$ can be obtained by solving the nonlinear Eqs.(21) and (22) numerically, then by substituting these estimates in Eq.(20) the estimate of the parameter N can be obtained.

4. Numerical Application

To illustrate the estimation procedures and examine the considered model, data analysis of real software data set is carried out in this section.

4.1 Description of Datasets

The dataset used in our data analysis was that developed by (Musa, 1980) of Bell Telephone Laboratories, Cyber Security and Information Systems Information Analysis Centre (CSIAC). Tables [2-6] present the selected five data sets.

Table 2. Failure time data Ds-1.

Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)
1	115	11	50	21	15	31	10571
2	0	12	71	22	390	32	563
3	83	13	606	23	1863	33	2770
4	178	14	1189	24	1337	34	652
5	194	15	40	25	4508	35	5593
6	136	16	788	26	834	36	11696
7	1077	17	222	27	3400	37	6724
8	15	18	72	28	6	38	2546
9	15	19	615	29	4561		
10	92	20	589	30	3186		

Table 3. Failure time data Ds-2.

Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)
1	5	15	424	29	283	43	887
2	73	16	92	30	50	44	149
3	141	17	520	31	308	45	469
4	491	18	1424	32	279	46	716
5	5	19	0	33	140	47	604
6	5	20	92	34	678	48	0
7	28	21	183	35	183	49	774
8	138	22	10	36	2462	50	256
9	478	23	115	37	104	51	14637
10	325	24	17	38	2178	52	18740
11	147	25	284	39	285	53	1526
12	198	26	296	40	171		
13	22	27	215	41	0		
14	56	28	116	42	643		

Table 4. Failure time data Ds-3.

Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)
1	3	35	227	69	529	103	108
2	30	36	65	70	379	104	0
3	113	37	176	71	44	105	3110
4	81	38	58	72	129	106	1247
5	115	39	457	73	810	107	943
6	9	40	300	74	290	108	700
7	2	41	97	75	300	109	875
8	91	42	263	76	529	110	245
9	112	43	452	77	281	111	729
10	15	44	255	78	160	112	1897
11	138	45	197	79	828	113	447
12	50	46	193	80	1011	114	386
13	77	47	6	81	445	115	446
14	24	48	79	82	296	116	122
15	108	49	816	83	1755	117	990
16	88	50	1351	84	1064	118	948
17	670	51	148	85	1783	119	1082
18	120	52	21	86	860	120	22
19	26	53	233	87	983	121	75
20	114	54	134	88	707	122	482
21	325	55	357	89	33	123	5509
22	55	56	193	90	868	124	100
23	242	57	236	91	724	125	10
24	68	58	31	92	2323	126	1071
25	422	59	369	93	2930	127	371
26	180	60	748	94	1461	128	790
27	10	61	0	95	843	129	6150
28	1146	62	232	96	12	130	3321
29	600	63	330	97	261	131	1045
30	15	64	365	98	1800	132	648
31	36	65	1222	99	865	133	5485
32	4	66	543	100	1435	134	1160
33	0	67	10	101	30	135	1864
34	8	68	16	102	143	136	4116

Table 5. Failure time data Ds-4.

Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)
1	191	15	50	29	3910	43	0
2	222	16	660	30	6900	44	0
3	280	17	1507	31	3300	45	300
4	290	18	625	32	1510	46	9021
5	385	19	912	33	195	47	2519
6	570	20	638	34	1956	48	6890
7	610	21	293	35	135	49	3348
8	365	22	1212	36	661	50	2750
9	390	23	612	37	50	51	6675
10	275	24	675	38	729	52	6945
11	360	25	1215	39	900	53	7899
12	800	26	2715	40	180	54	
13	1210	27	3551	41	4225		
14	407	28	800	42	15600		

Table 6. Failure time data Ds-5.

Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)	Failure Number	Failure Interval Length (in CPU seconds)
1	3	20	8	39	41	58	27
2	14	21	1	40	7	59	140
3	59	22	12	41	43	60	33
4	32	23	36	42	1	61	5
5	8	24	38	43	4	62	36
6	52	25	74	44	5	63	74
7	2	26	43	45	1	64	40
8	25	27	236	46	16	65	2
9	2	28	121	47	70	66	86
10	3	29	18	48	60	67	221
11	4	30	9	49	2	68	6
12	1	31	23	50	2	69	891
13	30	32	1	51	3	70	23
14	21	33	672	52	169	71	4
15	196	34	189	53	29	72	437
16	265	35	83	54	88	73	66
17	6	36	52	55	55		
18	3	37	8	56	27		
19	8	38	1	57	24		

4.2 Goodness of Fit Tests

Three evaluation criteria are used in the application. The variation between the predicted and actual values of observations is calculated by the Mean Square Error (MSE) as follows (Hwang and Pham, 2009):

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}\left(\mathrm{y_i} - \hat{\mathrm{m}}(\mathrm{t_i})\right)^2}{n-k}, \tag{23}$$

where n is the number of observations and k is the number of model's unknown parameters, $\mathrm{y_i}$ denotes the number of faults observed to the moment $\mathrm{t_i}$, and $\hat{\mathrm{m}}(\mathrm{t_i})$ denotes the estimated number of faults detected to the time $\mathrm{t_i}$ according to the considered model; for i=1, 2, …, n. The lower MSE indicates less fitting error, thus better performance. The Theil Statistic (TS) is the average deviation percentage over all periods with regard to the actual values. The closer, TS is to zero, the better the prediction capability of model. It is defined as (Li et al., 2005):

$$\mathrm{TS} = 100 * \sqrt{\frac{\sum_{i=1}^{n}\left(\mathrm{y_i} - \hat{\mathrm{m}}(\mathrm{t_i})\right)^2}{\sum_{i=1}^{n}\mathrm{y}_i^2}}\%. \tag{24}$$

The coefficient of multiple determinations R^2 value indicates the predictive measure of the difference among the forecasting values. It is defined as follows (Xie and Yang, 2003):

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}\left(\mathrm{y_i} - \hat{\mathrm{m}}(\mathrm{t_i})\right)^2}{\sum_{i=1}^{n}\left(\mathrm{y_i} - \sum_{k=1}^{n}{}^{\mathrm{y_k}}/_{n}\right)^2}. \tag{25}$$

It ranges from 0 to1. The larger R^2 is the better the model fits data.

4.3 Numerical Results and Analysis

The parameter estimation and evaluation criteria results of the NHPP LL model for the five considered data sets using the ML and NLS estimation methods are respectively shown in Table [7] and Table [8]. By comparing the results in these two tables, it is clear that the NHPP LL model provides better values of the MSE, R^2, and TS criteria when using the NLS estimation method for all cases. Based on the two studied methods of estimation, it is observed that the ranking of the data sets varies with respect to the selection of evaluation criteria as follows: According to MSE criteria the NHPP LL model's performance is the best for Ds-1. While, according to TS and R^2 the NHPP LL model's performance is the best for Ds-3. According to all considered criteria the NHPP LL model's performance is the worst for Ds-5.

Table 7. Estimated parameters values and comparison criteria results using MLE method.

Data sets	The NHPP LL Model					
	Parameters estimates			Model selection criteria		
	$\widehat{N}_{MLE}$	$\widehat{\gamma}_{MLE}$	$\widehat{\beta}_{MLE}$	MSE	TS	R^2
Ds-1	50.2091	0.0009	0.7298	2.4263	6.7768	0.9809
Ds-2	58.7724	0.0001	1.0684	6.1654	7.849	0.9746
Ds-3	239.8247	0.0004	0.7108	6.297	3.1549	0.996
Ds-4	86.625	0.0002	0.7746	2.9425	5.3253	0.9883
Ds-5	346.946	0.0004	0.7650	17.1994	9.6056	0.9623

Table 8. Estimated parameters values and comparison criteria results using NLSE method.

Data sets	The NHPP LL Model					
	Parameters estimates			Model selection criteria		
	$\widehat{N}_{NLsE}$	$\widehat{\gamma}_{NLsE}$	$\widehat{\beta}_{NLsE}$	MSE	TS	R^2
Ds-1	43.7384	0.0009	0.7731	1.349	5.0532	0.9894
Ds-2	60.1523	0.00003	1.193	4.4649	6.6794	0.9816
Ds-3	244.4601	0.0004	0.7037	6.0325	3.0879	0.9961
Ds-4	102.3934	0.0005	0.6621	2.0405	4.4346	0.9919
Ds-5	3156.474	0.00005	0.7244	15.8935	9.2337	0.9652

5. Conclusion

As software has become more diverse and spread, software reliability has also become a key concern in software development process. During the last 47 years numerous reliability models have been proposed (see; Yamada et al., 1983; Goel, 1985; Cai and Lyu, 2007; Yamada, 2013). These models are used to measure the software reliability through several characteristics such as: number of remaining errors, error detection rate, and mean time between failures. In this paper, we have considered a NHPP model that based on the log-logistic distribution which can capture increasing/decreasing nature of hazard function. Several essential characteristics of our studied model, the NHPP LL model, have been obtained and represented graphically. The considered model's parameters have been estimated using the ML, and NLS estimation methods. An application has been conducted using five real data sets and three different evaluation criteria. The considered model displays acceptable performance for the studied real data sets, particularly in the case of Ds-1 and Ds-3. The findings reveal that that the NHPP LL model gives a reasonable predictive capability for the studied real failure data.

References

Cai, X., & Lyu, M. R. (2007, November). Software reliability modeling with test coverage: Experimentation and measurement with a fault-tolerant software project. In *null* (pp. 17-26). IEEE.

Chang, Y. P. (2001). Estimation of parameters for nonhomogeneous Poisson process: Software reliability with change-point model. *Communications in Statistics-Simulation and Computation*, *30*(3), 623-635. https://doi.org/10.1081/SAC-100105083

Fisk, P. R. (1961). The graduation of income distributions. *Econometrica, 29*(2), 171-185. http://dx.doi.org/10.2307/1909287.

Goel, A. L. (1985). Software reliability models: Assumptions, limitations, and applicability. *IEEE Transactions on software engineering*, (12), 1411-1423. https://doi.org/10.1109/TSE.1985.232177

Gokhale, S. S., & Trivedi, K. S. (1998, November). Log-logistic software reliability growth model. In *High-Assurance Systems Engineering Symposium, 1998. Proceedings. Third IEEE International* (pp. 34-41). IEEE. https://doi.org/10.1109/HASE.1998.731593

Harishchandra, B. K. (2016). On a Software Reliability Growth Model with Log Logistic failure time distribution. *International Journal of Engineering and Computer Science, 5*(12), 19553- 9560, Retrieved from http://www.ijecs.in/index.php/ijecs/article/view/3473

Hwang, S., & Pham, H. (2009). Quasi-renewal time-delay fault-removal consideration in software reliability modeling. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *39*(1), 200-209. https://doi.org/10.1109/TSMCA.2008.2007982

Knafl, G. J. (1992, October). Solving maximum likelihood equations for two-parameter software reliability models using grouped data. In *Proceedings Third International Symposium on Software Reliability Engineering* (pp. 205-213). IEEE. https://doi.org/10.1109/ISSRE.1992.285844

Li, P. L., Herbsleb, J., & Shaw, M. (2005, November). Forecasting field defect rates using a combined time-based and metrics-based approach: a case study of OpenBSD. In *Software Reliability Engineering, 2005. ISSRE 2005. 16th IEEE International Symposium on* (pp. 10-pp). IEEE.

Lyu, M. R. (1996). Handbook of Software Reliability Engineering, McGraw- Hill.

Lyu, M. R. (2002). Software Reliability Theory. Encyclopedia of Software Engineering, Wiley, 1611-1630. https://doi.org/10.1002/0471028959.sof329

Musa J. (1980). Data analysis center for software: An information analysis center, Western Michigan University Library, Kalamazoo, Michigan.

Musa, J. D. (1999). Software Reliability Engineering: More Reliable Software, Faster Development and Testing, McGraw-Hill.

Prasad, R. S., Rao, K. H., & Kantha, R. L. (2011). Software Reliability Measuring Using Modified Maximum Likelihood Estimation and SPC, *International Journal of Computer Applications, 21*(7), 1-5.

Rana, R., Staron, M., Berger, C., Hansson, J., Nilsson, M., & Törner, F. (2013). Comparing between Maximum Likelihood Estimator and Non-linear Regression Estimation Procedures for NHPP Software Reliability Growth Modelling, Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement Ankara. pp. 213-218.

Xie, M., & Yang, B. (2003). A study of the effect of imperfect debugging on software development cost, IEEE *Transactions on Software Engineering, 29*(5), 471–473. https://doi.org/10.1109/TSE.2003.1199075

Yamada, S. (2013). Recent developments in software reliability modeling and its applications, in Stochastic Reliability and Maintenance Modeling. Springer, pp. 251–284. https://doi.org/10.1007/978-1-4471-4971-2_12

Yamada, S., Ohba, M., & Osaki, S. (Dec 1983). S-shaped reliability growth modeling for software error detection, Reliability, *IEEE Transactions on, 32*(5), 475–484. https://doi.org/10.1109/TR.1983.5221735

Zeephongsekul, P., Jayasinghe, C. L., Fiondella, L., & Nagaraju, V. (2016). MaximumLikelihood Estimation of Parameters of NHPP Software Reliability Models Using Expectation Conditional Maximization Algorithm. *IEEE Transactions on Reliability, 65,* 1571-1583. https://doi.org/10.1109/TR.2016.2570557

Zhao, M., & Xie, M. (1996). On maximum likelihood estimation for a general nonhomogeneous Poisson process. *Scandinavian Journal of Statistics, 23,* 597- 607.

Author Information

Lutfiah Ismail Al turk is currently working as associate professor of mathematical statistics in Statistics Department at Faculty of Sciences, King AbdulAziz University, Jeddah, Kingdom of Saudi Arabia. Lutfiah Ismail Al turk obtained her B.Sc degree in statistics and computer science from Faculty of Sciences, King AbdulAziz University in 1993 and M.Sc (mathematical statistics) degree from Statistics Department, Faculty of Sciences, King AbdulAziz University in 1999. She received her Ph.D in mathematical statistics from university of Surrey, UK in 2007. Her current research interests include software reliability modeling and statistical machine learning.

Email: lturk@kau.edu.sa

URL: http://lturk.kau.edu.sa

Address: P.O. Box 42713 Jeddah 21551. Kingdom of Saudi Arabia.

Copyrights