Piecewise NHPP Models with Maximum Likelihood Estimation for Repairable Systems

Huairui Guo, PhD, ReliaSoft Corporation

Adamantios Mettas, ReliaSoft Corporation

Georgios Sarakakis, ReliaSoft Corporation

Pengying Niu, ReliaSoft Corporation

Key Words: piecewise NHPP model, maximum likelihood estimation, repairable system

SUMMARY & CONCLUSIONS

Non-homogeneous Poisson process (NHPP) models are widely used for repairable system analysis. Different NHPP models have been developed for different applications. It has been noticed that almost all the existing models apply only a single model for the entire system development or operation period. However, in some circumstances, such as when the system design or the system operation environment experiences major changes, a single model will not be appropriate to describe the failure behavior for the entire timeline.

In this paper, we proposed a piecewise NHPP model for repairable systems with multiple stages. The maximum likelihood estimation (MLE) for the model parameters is also provided.

1 INTRODUCTION

Consider a repairable system, such as a vehicle or an air conditioner. Its failure process usually can be modeled as a non-homogeneous Poisson process. Ascher and Feingold's book [1], one of the most influential books on repairable system analysis, discussed many research results published before the 1990s. In the past several decades, different NHPP models have been developed for modeling different applications. For example, the Crow-AMSAA used by Crow [2] and the log-linear model proposed by Cox and Lewis [3] are two of the most popular ones. These two models are used for the cases when the failure intensity monotonically decreases, increases or keeps constant with time. The bounded failure intensity models [4, 5] are designed for the situations when the failure intensity has an upper bound when time goes to infinity. S-shaped models such as the Yamada model [6] and the generalized GO-model [7] are used to analyze software reliability where the failure intensity vs. time plot usually is S-shaped.

All of the aforementioned models apply one single model for the entire system development or operation period. In some situations, however, such as when the system under development experiences major changes or the system operates under different environments, a single model may not be accurate enough. To overcome this drawback, piecewise HPP (homogeneous Poisson process) models were introduced for an NHPP process and the nonparametric method was used for the parameter estimation [8-10]. The piecewise HPP models assume that the failure intensity is constant within a given time period and would be different for different periods. However, this assumption is not always true. For a given time interval, especially for systems deployed in the field, the failure intensity can change with time.

In this paper, we propose a piecewise NHPP model for repairable systems with multiple development or operation stages. If the separation or change points between stages cannot be identified based on engineering knowledge and have to be found by analyzing the data, the iteration method proposed in this paper can be used to estimate them. In this paper, the Crow-AMSAA model is applied as the base model for each stage. Unlike the piecewise HPP model, which ignores the damages from the previous stages, the damages accumulated from the previous stages are considered in the proposed piecewise NHPP model. This is different from the method that analyzes each stage separately.

Two well-known methods have been used for model parameter estimation in repairable system modeling. One is the least squares estimation (LSE) and another is the maximum likelihood estimation (MLE). For a given NHPP model, for instance for the Crow-AMSAA model, it is found that the confidence intervals for the model parameters estimated from LSE are usually tighter than the confidence intervals estimated from MLE. One of the reasons is that in the application of the least squares method, each observation (cumulative number of failures) needs to be independent and with constant variance. However, on page 85 of Ascher and Feingold's book [1], it is stated that for a NHPP the independency assumption usually does not hold true from a statistical viewpoint. The dependency between the observed values results in the underestimation of the variability of the model parameters when the LSE is used. This also results in narrow confidence intervals for the parameters and the functions of these parameters such as the number of failures. We use the MLE method in this paper. For the MLE method, the conditional probability will be used in constructing the

978-1-4244-5103-6/10/$26.00 ©2010 IEEE

likelihood function. By using the conditional probability, the dependency between the observations is considered in the model.

However, there are some barriers in the use of MLE. The piecewise NHPP model is a discrete function. It consists of different NHPP models for distinct time regions. This causes the maximum likelihood function for the overall model to be a discrete function too. The traditional method that uses the first order derivatives to get the solution for the model parameters cannot be applied directly when the change points are not pre-defined. Limited research has been done on how to obtain the MLE solution for a piecewise function. For example, Tishler and Zang [11] employed a continuous function to approximate the piecewise linear function, then obtained analytical derivatives of the likelihood function from the approximated function to get the MLE solutions. However, getting a sufficiently accurate approximated function is not an easy task. In this paper, a practical method is proposed for obtaining the change points first. Once the change points are obtained, a continuous likelihood function can be formed. Solutions for other model parameters such as the scale and shape parameters in the Crow-AMSAA model can then be directly estimated using the derivatives.

This paper is organized as follows. In section 2, an example is provided to illustrate why the piecewise NHPP model is needed. Section 3 discusses how to build the likelihood function. In Section 4, parameter estimation and confidence intervals are discussed. Section 5 provides the solutions for an illustration example. Finally, we provide conclusions in Section 6.

2 PROBLEM STATEMENT

2.1 Example

Consider a repairable system with the following failure data.

Cum. Num. of Failures	Failure Time	Cum. Num. of Failures	Failure Time
1	15.70	12	141.41
2	29.39	13	143.67
3	41.14	14	144.63
4	56.47	15	144.95
5	75.61	16	145.16
6	98.83	17	146.25
7	112.42	18	146.70
8	125.61	19	147.26
9	129.39	20	148.15
10	133.45	21	152.40
11	138.94		

Table 1 – Failure Data for the Illustration Example

For the data in Table 1, the cumulative number of failures vs. time plot is shown in Figure 1.

In Figure 1, the points are the observed number of failures and the curve represents the predicted values using the Crow-

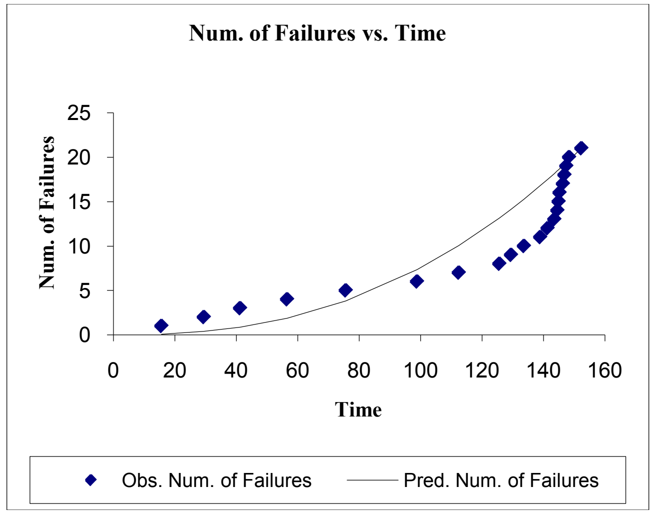

Figure 1 – Failure Number vs. Time Plot

AMSAA model, which will be discussed in section 2.2. Clearly, from Figure 1, one can see that the single model cannot fit the data well. The Cramér-von Mises (CVM) test [2, 12] shows that at a significance level of 0.1, the model fails the goodness-of-fit test. The calculated CVM is 0.534, which is larger than the critical value of 0.172. By examining the curve, we can see that there are two segments with the change point at around 120. It will be possible to fit an individual NHPP model for each segment. However, these two individual models are not independent. They are connected by the change point. In the following section, we will discuss the Crow-AMSAA model and the piecewise NHPP model.

2.2 Piecewise Power Law NHPP Model

The failure intensity function of a NHPP can be described by a power function:

$$\lambda(t) = \lambda \beta t^{\beta - 1} \tag{1}$$

where λ is the scale parameter and β is the shape parameter. This model is also known as the Crow-AMSAA model [2]. Under this model, the cumulative number of failures is calculated by:

$$N(t) = \lambda t^{\beta} \tag{2}$$

By taking the logarithmic transformation of both sides, equation (2) can be linearized as:

$$\ln[N(t)] = \ln(\lambda) + \beta \ln(t) \tag{3}$$

Equation (3) is a standard simple linear regression model. Therefore, the least squares method can be applied if the cumulative numbers of failures are assumed to be independent and with constant variance. This is the basic assumption of using least squares estimation.

For the data given in Table 1, the individual Crow-AMSAA model will be applied to each of the segments. Assume the change point is *C*. The piecewise NHPP model will be:

$$\begin{cases} N_1(t) = \lambda_1 t^{\beta_1} & t <= C \\ N_2(t) = \lambda_2 t^{\beta_2} & t > C \end{cases} \tag{4}$$

The two curves intercept at time *C*. Therefore, $N_1(C) = N_2(C)$

and the model parameters have the following relationship:

$$\lambda_2 = \lambda_1 C^{\beta_1 - \beta_2} \tag{5}$$

From equation (5), it can be seen that if C is known, there are only three independent parameters. For example, we can use λ_1, β_1 and β_2. If C is unknown and needs to be estimated, C will be an additional parameter. This is similar to the strategy used by the traditional piecewise linear regression [13]. Once we have the piecewise NHPP model, the next step is to estimate the model parameters. Section 3 will give the MLE solution.

3 PARAMETER ESTIMATION AND STATISTICAL INFERENCES

In order to use MLE, the likelihood function should be established first. The example in section 2 is for only one system. In some cases, failure data from multiple systems will be available. For example, several prototypes might be built in the development stage, or several systems might be deployed in the field at the same time. In the following sections, we will give the likelihood function for the case of multiple systems and then provide the closed form solutions for the case of a single system.

3.1 Likelihood Function

To make the discussion simple, let's assume there is only one change point and the change point C is known. For a system, t_i is used to denote the ith failure time. The probability of the ith failure is conditional on the (i-1)th failure time. The conditional probability of the ith failure in segment 1 is:

$$F_1(t_i \mid t_{i-1}) = 1 - \frac{R_1(t_i)}{R_1(t_{i-1})} = 1 - \exp\left[\lambda_1 t_{i-1}^{\beta_1} - \lambda_1 t_i^{\beta_1}\right] \tag{6}$$

where $F_1(\cdot)$ and $R_1(\cdot)$ are the probability of failure and the reliability at segment 1. From equation (6), the conditional probability density function can be obtained as:

$$f_1(t_i \mid t_{i-1}) = \lambda_1 \beta_1 t_i^{\beta_1 - 1} \exp\left[\lambda_1 t_{i-1}^{\beta_1} - \lambda_1 t_i^{\beta_1}\right] = \frac{f_1(t_i)}{R_1(t_{i-1})} \tag{7}$$

For the failures at segment 2, we only need to use the model for segment 2 to replace the model in segment 1 in equation (6) and (7). Assuming that the system starts at time S and ends at time T, the likelihood function will be:

$$\begin{aligned} LKV &= f_1(t_1 \mid S) \prod_{i=2}^{N_1} f_1(t_i \mid t_{i-1}) R_1(C \mid t_{N_1}) \\ &\times f_2(t_{N_1+1} \mid C) \prod_{i=N_1+2}^{N} f_2(t_i \mid t_{i-1}) R_2(T \mid t_{N_1}) \end{aligned} \tag{8}$$

where:

- $f_i(\cdot)$ is the probability density function for segment i.
- N is the total number of failures.
- N_1 is the total number of failures in segment 1.

Equation (8) can be expanded as:

$$\begin{aligned} LKV &= \frac{f_1(t_1)}{R_1(S)} \cdot \frac{f_1(t_2)}{R_1(t_1)} \cdots \frac{R_1(C)}{R_1(t_{N_1})} \\ &\times \frac{f_2(t_{N_1+1})}{R_2(C)} \frac{f_2(t_{N_1+2})}{R_2(t_{N_1+1})} \cdots \frac{R_2(T)}{R_2(t_N)} \end{aligned} \tag{9}$$

At time C, the system reliability can only have one value. So $R_1(C) = R_2(C)$, thus they cancel-out in equation (9). If the system end time is the same as the last failure time, the last term in equation (9) will equal to 1. Therefore, equation (9) is for both time and failure terminated situations.

If we take the logarithmic transform of equation (9), and by canceling some terms, we get the log likelihood function as:

$$\begin{aligned} \ln(LKV) &= N_1\left[\ln(\lambda_1) + \ln(\beta_1)\right] + (\beta_1 - 1)\sum_{i=1}^{N_1} \ln(t_i) + \lambda_1 S^{\beta_1} \\ &+ N_2\left[\ln(\lambda_2) + \ln(\beta_2)\right] + (\beta_2 - 1)\sum_{i=N_1+1}^{N} \ln(t_i) - \lambda_2 T^{\beta_1} \end{aligned} \tag{10}$$

Generalizing equation (10) to multiple systems, the log likelihood function is:

$$\begin{aligned} \Lambda &= \sum_{q=1}^{K}\left[N_{q,1}\left[\ln(\lambda_1) + \ln(\beta_1)\right] + (\beta_1 - 1)\sum_{i=1}^{N_{q,1}} \ln(t_{q,i}) + \lambda_1 S_q^{\beta_1}\right] \\ &+ \sum_{q=1}^{K}\left[N_{q,2}\left[\ln(\lambda_2) + \ln(\beta_2)\right] + (\beta_2 - 1)\sum_{i=N_{q,1}+1}^{N_q} \ln(t_{q,i}) - \lambda_2 T_q^{\beta_1}\right] \end{aligned} \tag{11}$$

where:

- K is the total number of systems. $q = 1...K$.
- N_q is the total number of failures for the qth system.
- $N_{q,1}$ is the number of failures in segment 1 for the qth system.
- $N_{q,2}$ is the number of failures in segment 2 for the qth system.
- S_q is the start time of the qth system.
- T_q is the end time of the qth system.

Recall that equation (5) shows that there are only three independent parameters, if the change point C is given. Therefore, we replace λ_2 using the other parameters. Equation (11) then becomes:

$$\begin{aligned} \Lambda &= \sum_{q=1}^{K}\left[N_{q,1}\left[\ln(\lambda_1) + \ln(\beta_1)\right] + (\beta_1 - 1)\sum_{i=1}^{N_{q,1}} \ln(t_{q,i}) + \lambda_1 S_q^{\beta_1}\right] \\ &+ \sum_{q=1}^{K}\left[N_{q,2}\left[\ln(\lambda_1) + \ln(\beta_2)\right] + N_{q,2}(\beta_1 - \beta_2)\ln(C)\right] \\ &+ \sum_{q=1}^{K}\left[(\beta_2 - 1)\sum_{i=N_{q,1}+1}^{N_q} \ln(t_{q,i}) - \lambda_1 C^{\beta_1}\left(\frac{T_q}{C}\right)^{\beta_2}\right] \end{aligned} \tag{12}$$

When C is unknown, it can be seen from equation (12) that it is not possible to get the likelihood function. This is because the number of failures for each segment, which is required in the likelihood function, is not available without knowing the change point. Therefore, it is not possible to construct the likelihood function without having C as given. In section 3.3, a heuristic method will be proposed for iteratively solving for C.

Equation (12) is a continuous function for λ_1, β_1 and β_2. So we can get the first order derivative for each parameter. By setting the derivatives equal to 0, the following nonlinear equations are obtained:

$$\lambda_1 = \frac{\sum_{q=1}^{k} N_q}{C^{\beta_1-\beta_2}\sum_{q=1}^{k} T_q^{\beta_2} - \sum_{q=1}^{k} S_q^{\beta_1}} \quad (13)$$

$$\beta_1 = \frac{\sum_{q=1}^{k} N_{q,1}}{\ln(C)\left[\sum_{q=1}^{k} N_{q,1} + \lambda_1 \sum_{q=1}^{k} S_q^{\beta_1}\right] - \sum_{q=1}^{k}\sum_{i=1}^{N_{q,1}} \ln(t_{q,i}) - \lambda_1 \sum_{q=1}^{k} S_q^{\beta_1}\ln(S_q)} \quad (14)$$

$$\beta_2 = \frac{\sum_{q=1}^{k} N_{q,2}}{\ln(C)\sum_{q=1}^{k} N_{q,2} + \lambda_1 C^{\beta_1-\beta_2}\sum_{q=1}^{k}\left[T_q^{\beta_2}\ln(\frac{T_q}{C})\right] - \sum_{q=1}^{k}\sum_{i=N_{q,1}+1}^{N_q} \ln(t_{q,i})} \quad (15)$$

By solving the nonlinear equations (13-15), we can get the ML estimators for the model parameters.

When all the systems start from 0 and end at the same time, T, closed form solutions exist. These are:

$$\lambda_1 = \frac{\sum_{q=1}^{k} N_q}{C^{\beta_1-\beta_2} K T^{\beta_2}} \quad (16)$$

$$\beta_1 = \frac{\sum_{q=1}^{k} N_{q,1}}{\ln(C)\left[\sum_{q=1}^{k} N_{q,1}\right] - \sum_{q=1}^{k}\sum_{i=1}^{N_{q,1}} \ln(t_{q,i})} \quad (17)$$

$$\beta_2 = \frac{\sum_{q=1}^{k} N_{q,2}}{\ln(C)\sum_{q=1}^{k} N_{q,2} + \ln(\frac{T}{C})\sum_{q=1}^{k} N_q - \sum_{q=1}^{k}\sum_{i=N_{q,1}+1}^{N_q} \ln(t_{q,i})} \quad (18)$$

So far the discussion is based on the case of one change point. If there are multiple change points, the same methodology is still valid.

Once the model parameters are estimated, the next step is to obtain the variance and covariance matrix for them. To get the variance and covariance matrix, we first need to calculate the Fisher information matrix [14], which is:

$$F = -\begin{bmatrix} \frac{\partial^2 \Lambda}{\partial \lambda_1^2} & \frac{\partial^2 \Lambda}{\partial \lambda_1 \partial \beta_1} & \frac{\partial^2 \Lambda}{\partial \lambda_1 \partial \beta_2} \\ & \frac{\partial^2 \Lambda}{\partial \beta_1^2} & \frac{\partial^2 \Lambda}{\partial \beta_1 \partial \beta_2} \\ & & \frac{\partial^2 \Lambda}{\partial \beta_2^2} \end{bmatrix} \quad (19)$$

F is a symmetric matrix. The variance and covariance matrix is obtained by:

$$\Sigma = F^{-1} \quad (20)$$

If there is only one system and it starts from time 0, the calculation will be even simpler. In section 3.2, we will provide the MLE solutions for a single system and also the confidence bounds of the model parameters.

3.2 MLE Solutions for a Single System

For a system that starts from time 0 and has one change point C, the log likelihood function in equation (12) can be simplified as:

$$\begin{aligned} L &= N\ln(\lambda_1) + N_1\ln(\beta_1) + N_2\ln(\beta_2) + N_2(\beta_1-\beta_2)\ln(C) \\ &+ (\beta_1-1)\sum_{i=1}^{N_1}\ln(t_i) + (\beta_2-1)\sum_{i=N_1+1}^{N}\ln(t_i) - \lambda_1 c^{\beta_1-\beta_2}T^{\beta_2} \end{aligned} \quad (21)$$

The closed form solutions for the model parameters are:

$$\lambda_1 = \frac{N}{C^{\beta_1-\beta_2}T^{\beta_2}} \quad (22)$$

$$\beta_1 = \frac{N_1}{N_1\ln(C) - \sum_{i=1}^{N_1}\ln(t_i)} \quad (23)$$

$$\beta_2 = \frac{N_2}{N\ln(T) - \sum_{i=N_1+1}^{N}(\ln t_i) - N_1\ln(C)} \quad (24)$$

The Fisher information matrix is:

$$F = \begin{bmatrix} \frac{N}{\lambda_1^2} & C^{\beta_1}\left(\frac{T}{C}\right)^{\beta_2}\ln(C) & C^{\beta_1}\left(\frac{T}{C}\right)^{\beta_2}\ln(\frac{T}{C}) \\ & \frac{N_1}{\beta_1^2} + N\ln^2(C) & N\ln(C)\ln(\frac{T}{C}) \\ & & \frac{N_2}{\beta_2^2} + N\ln^2(\frac{T}{C}) \end{bmatrix} \quad (25)$$

By taking the inverse of matrix F, we can get the variance and covariance matrix for λ_1, β_1 and β_2. The 2-sided $(1-\alpha)100\%$ confidence bounds on model parameters can be obtained by:

$$CB_\theta = \hat{\theta}\exp\left(\pm z_{\alpha/2}\sqrt{Var(\hat{\theta})}/\hat{\theta}\right) \quad (26)$$

Where:

- θ is a given parameter.
- $Var(\hat{\theta})$ is the variance of that parameter.
- $\hat{\theta}$ is the ML estimator.
- $z_{\alpha/2}$ is the $(1-\alpha/2)$ percentile of a standard normal distribution.

The confidence bounds for a function of the model parameters can also be calculated. For example for λ_2 which is a function of λ_1, β_1 and β_2, its variance is:

$$\begin{aligned} Var(\lambda_2) &= \left[\frac{d\lambda_2}{d\lambda_1}\right]^2 Var(\lambda_1) + \left[\frac{d\lambda_2}{d\beta_1}\right]^2 Var(\beta_1) + \left[\frac{d\lambda_2}{d\beta_2}\right]^2 Var(\beta_2) \\ &+2\left[\frac{d\lambda_2}{d\lambda_1}\right]\left[\frac{d\lambda_2}{d\beta_1}\right]Cov(\lambda_1,\beta_1) + 2\left[\frac{d\lambda_2}{d\lambda_1}\right]\left[\frac{d\lambda_2}{d\beta_2}\right]Cov(\lambda_1,\beta_2) \\ &+2\left[\frac{d\lambda_2}{d\beta_1}\right]\left[\frac{d\lambda_2}{d\beta_2}\right]Cov(\beta_1,\beta_2) \end{aligned} \quad (27)$$

Using the variance from equation (27) in equation (26), we can get the bounds for λ_2.

The method in equations (26) and (27) is also valid for other functions of the model parameters such as the cumulative number of failures, MTBF and failure intensity. Their expected values can be calculated using the ML estimators of λ_1, β_1 and β_2 and their confidence bounds can be calculated using the method in equations (26) and (27).

In the above discussion, we assume that the change point *C* is known. Sometime, it is difficult to know the exact value of *C*. In section 3.3, we will discuss how to estimate the *C* value.

3.3 Discussion on Change Point C

If the change point *C* cannot be determined upfront, the following procedure can be used to estimate it.

Step 1: Plot the failure number vs. time plot in both linear and logarithmic scale.

Step 2: From the plots, identify the range for *C*, denoted as [C_Min, C_Max].

Step 3: Set $C_i = C\_Min + i \times \Delta C$. Calculate the MLE solution for $\hat{\lambda}_1$, $\hat{\beta}_1$ and $\hat{\beta}_2$ using C_i.

Step 4: Calculate the log likelihood value using $\hat{\lambda}_1$, $\hat{\beta}_1$ and $\hat{\beta}_2$.

Step 5: Set $i = i + 1$ and repeat steps 3 and 4 until *C* reaches C_Max.

The solution of $\hat{\lambda}_1$, $\hat{\beta}_1$ and $\hat{\beta}_2$ and the value of *C* that provide the largest likelihood value will be the ML solution.

It has been found that several local maximal points may exist in the search range. If it is necessary, further refined search with smaller ΔC should be conducted at the vicinity of the local maximal points. Searching from C_Min to C_Max will ensure that the optimum solution in this range will be found. In section 4, we will use an example to illustrate how the *C* value affects the likelihood values.

In the discussion of equation (12), we pointed that it is not possible to get the likelihood function because the values for number of failures $N_{q,1}$ and $N_{q,2}$ at each stage cannot be determined if *C* is unknown. Now, let's assume that we know that the range for *C* is from C_Min to C_Max and there is no failure between C_Min and C_Max. This means we know $N_{q,1}$ and $N_{q,2}$. If this is the case, can we get the ML estimator for *C* by taking its derivative from the likelihood function? The answer is no.

Recall that the log likelihood function for one system is equation (21). The first order derivative for *C* from this equation is:

$$\frac{\partial \Lambda}{\partial C} = \frac{N_2(\beta_1 - \beta_2)}{C} - \frac{\lambda_1 T^{\beta_2} C^{\beta_1 - \beta_2}(\beta_1 - \beta_2)}{C} \tag{28}$$

From equation (22), we know:

$$N = \lambda_1 C^{\beta_1 - \beta_2} T^{\beta_2}$$

So equation (28) is simplified as:

$$\frac{\partial \Lambda}{\partial C} = \frac{N_2(\beta_1 - \beta_2)}{C} - \frac{N(\beta_1 - \beta_2)}{C} = \frac{N_1(\beta_2 - \beta_1)}{C} \tag{29}$$

By setting equation (29) equal to 0, we will have $\beta_1 = \beta_2$ and there is no solution for *C*. This proves that using the derivative to get the solution for change point *C* is impossible. The heuristic method given in this section should be used to estimate *C*.

4 SOLUTION FOR THE EXAMPLE

In this section, we will provide the results for the example given in section 2.

4.1 Estimate the Range for Change Point C

From Figure 1, we can estimate that *C* is within 120 and 140. Using the method described in section 3.3, the global optimum of *C* is found to be 125.5999.

To check how the likelihood value changes with *C*, we calculated the likelihood values for C values from 50 to 145. The results are given in Figure 2.

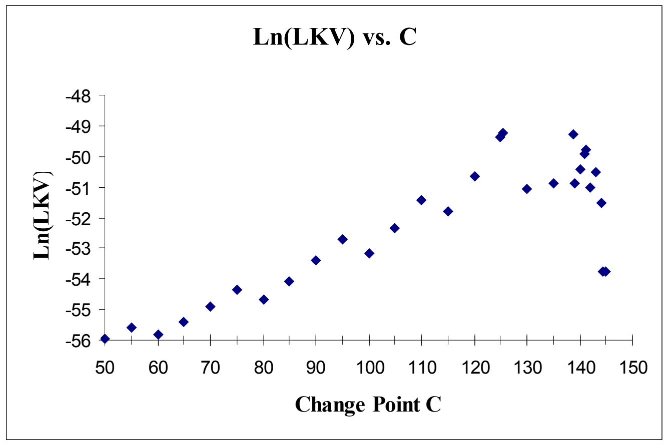

Figure 2 – Ln(LKV) vs. Change Point C

From Figure 2, we can see that there are several local maximal points.

4.2 ML Estimators and Statistical Inferences

Based on the estimated *C* value, using the equations given in section 3.2, the MLE results for the parameters and their standard deviations are given in Table 2.

	MLE		LSE	
Parameter	Mean	Std. Deviation	Mean	Std. Deviation
λ_1	0.0312	0.0644	0.0675	0.0099
β_1	1.1093	0.4193	0.9987	0.0343
β_2	5.9601	1.5929	6.7706	0.9723

Table 2 – MLE and LSE Solutions

For comparison, Table 2 also provides the LSE solutions. As noted in the introduction section, the LSE method usually provides smaller standard deviations for model parameters in the NHPP model. This can be seen in Table 2.

Using the information in Table 2, the confidence bounds for each of the model parameters can be calculated. Table 2 also can be used to get statistical inferences for the functions such as the number of failures, MTBF and the failure intensity, based on the model parameters. For example, we can calculate the cumulative number of failures and its confidence intervals using the procedure given in section 3. The results are given in Figure 3.

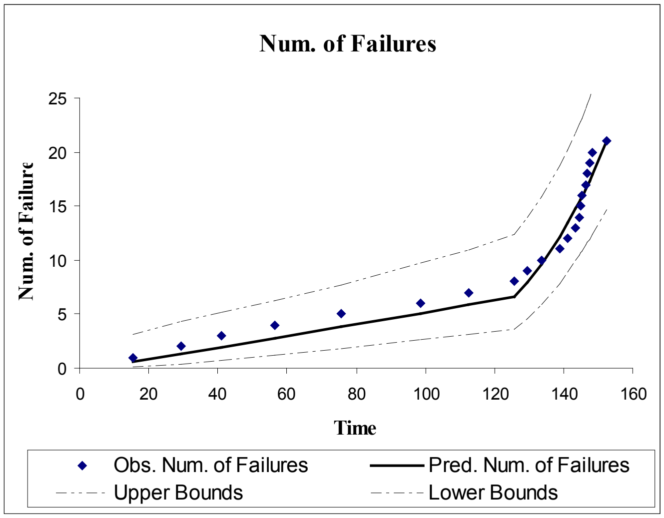

Figure 3– MLE Solution: Cumulative Num. of Failures

In Figure 3, the dashed lines are the upper and lower bounds for the cumulative number of failures. For comparison, the predicted number of failures and their bounds from the LSE are given in Figure 4. Using LSE, the estimated value for the change point is 133.878.

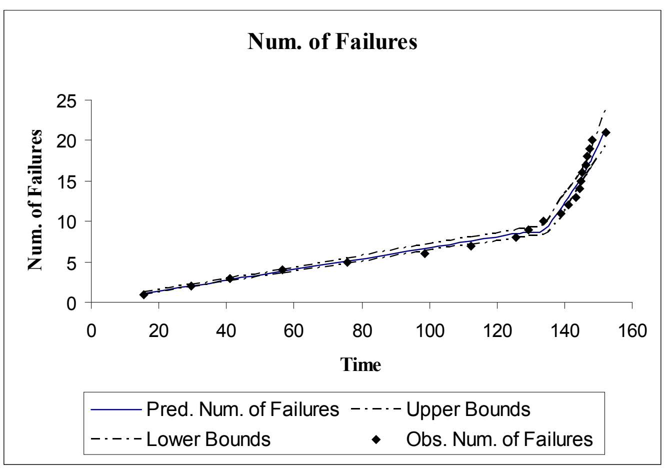

Figure 4– LSE Solution: Cumulative Num. of Failures

4.3 Goodness-of-Fit Test

Since the individual failure times are known, the Cramér-von Mises statistic is used to test the null hypothesis that the piecewise NHPP model with power law failure intensity functions can describe the failure behavior well. It is known that for a single NHPP model, the CVM goodness-of-fit is given by [12]:

$$C_M^2 = \frac{1}{12M} + \sum_{i=1}^{M}\left[\left(\frac{t_i}{T}\right)^{\beta} - \frac{i-0.5}{M}\right]^2 \tag{30}$$

where $M = N$ if the test is time terminated; $M = N - 1$ if the test is failure terminated. N is the total number of failures. For the piecewise NHPP model of this example, we modified the CVM statistic and calculated it as:

$$\begin{aligned} C_M^2 = & \frac{1}{12M} + \sum_{i=1}^{M_1}\left[\left(\frac{t_i}{C}\right)^{\beta_1} - \frac{i-0.5}{M_1}\right]^2 \\ & + \sum_{i=N_1+1}^{M}\left[\left(\frac{t_i^{\beta_2} - C^{\beta_2}}{T^{\beta_2} - C^{\beta_2}}\right) - \frac{i - N_1 - 0.5}{M_2}\right]^2 \end{aligned} \tag{31}$$

where $M_1 = N_1$. $M_2 = N_2$ if the test is time terminated; $M_2 = N_2 - 1$ if the test is failure terminated. N_i is the total number of failures at the *i*th segment.

For this example, $M_1 = 7$ and $M_2 = 13$. Using equation (31), the C_M^2 is calculated to be 0.150. The critical value at a significant level of 0.1 is 0.172. Since C_M^2 is less than the critical value, we cannot reject the null hypothesis. The piecewise NHPP model passes the goodness-of-fit test. In fact, by examining Figure 3, we can see that the proposed model can fit the data very well which is a big improvement that than the model in Figure 1.

5 CONCLUSIONS

In this paper, we examined the case of repairable systems with multiple stages for which a single model is not adequate to describe the failure behavior for the entire timeline. We proposed a piecewise NHPP model to address this. Such cases often occur when there is a design change, natural operator learning, or when there is a change in the system's operational environment.

We provided the maximum likelihood function for the general case of multiple repairable systems. Closed form solutions for the model parameters for single system were also given. We proposed an iteration method to identify the change point in the cases that the separation between stages cannot be identified based on engineering knowledge. Parameter bounds and an integrated test statistic for the goodness of fit of the proposed model were also discussed.

The proposed methodology allows the data analyst to apply a piecewise NHPP model for repairable systems with multiple stages and uses a robust statistical methodology to assure that the model provides a good fit even if the change points are unknown.

REFERENCES

1. H. Ascher and H. Feingold, *Repairable Systems Reliability*, Marcel Dekker, Inc. 1984.
2. L. H. Crow, "Reliability Analysis for Complex Repairable Systems in Reliability and Biometry," *Society for Industrial and Applied Mathematics*, ed. by F. Proschan & R.J. Serfling, Philadelphia, 1974, pp 379–410.
3. D. R. Cox and P. A. Lewis, *The Statistical Analysis of Series of Events*, Methouen, London, 1966.

4. L. Attardi and G. Pulcini, “A New Model for Repairable Systems with Bounded Failure Intensity,” *IEEE Trans. Reliability*, vol. 54, no. 4, 2005, pp 572-582.
5. G. Pulcini, “A Bounded Intensity Process for the Reliability of Repairable Equipment,” *J. Qual. Technol*. vol. 33. no. 4, 2001, pp 480-492.
6. S. Yamada M. Ohba, and S. Osaki, “S-Shaped Reliability Growth Modeling for Software Error Detection,” *IEEE Trans. Reliability*, vol. R-32, no. 12, 1983, pp 475-478.
7. L. Goel and K. Okumoto, “A Time Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures,” *IEEE Trans. on Reliability*, vol. R-28. no. 3, 1979, pp 206—211.
8. L.M. Leemis, “Nonparametric Estimation and Variate Generation for a Nonhomogeneous Poisson Process from Event Count Data,” *IIE Trans*. vol. 36, 2004, pp 1155-1160.
9. L.M. Leemis, “Nonparametric Estimation of the Intensity Function for a Nonhomogeneous Poisson Process,” *Management Science*, vol. 37, no. 7, 1991, pp 886-900.
10. L. C. Tang and M. Xie, “A Simple Graphical Approach for Comparing Reliability Trends of Different Units in a Fleet,” *Proc. Ann. Reliability & Maintainability Symp.*, 2002, pp 280-285.
11. A. Tishler and I. Zang, “A Maximum Likelihood Method for Piecewise Regression Models with a Continuous Dependent Variable,” *Appl. Statist*. vol. 30, no. 2, 1981, pp 116-124.
12. D. Kececioglu, *Reliability & Life Testing Handbook*, vol. 1, DEStech Publications, Inc, 2002.
13. S. E, Ryan and L. S. Porth, *A Tutorial on the Piecewise Regression Approach Applied to Bedload Transport Data,* Gen. Tech. Rep. RMRS-GTP-189. Fort Collins, CO. U. S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 2007.
14. W. Q. Meeker, and L. A. Escobar, *Statistical Method for Reliability Data*, John Wiley & Sons, Inc., 1998.

BIOGRAPHIES

Huairui Guo, Ph.D.
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, Arizona 85710, USA

e-mail: Harry.Guo@ReliaSoft.com

Dr. Huairui Guo is the Director of Theoretical Development at ReliaSoft Corporation. He received his PhD in Systems and Industrial Engineering from the University of Arizona. His dissertation focuses on process modeling, diagnosis and control for complex manufacturing processes. He has published numerous papers in the areas of quality engineering including SPC, ANOVA and DOE and reliability engineering. His current research interests include repairable system modeling, accelerated life/fegradation testing, warranty data analysis and robust optimization. Dr. Guo is a member of SRE, IIE and IEEE.

Adamantios Mettas
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, Arizona 85710, USA

e-mail: Adamantios.Mettas@ReliaSoft.com

Mr. Mettas is the Vice President of product development at ReliaSoft Corporation. He fills a critical role in the advancement of ReliaSoft's theoretical research efforts and formulations in the subjects of life data analysis, accelerated life testing, and system reliability and maintainability. He has played a key role in the development of ReliaSoft's software including, Weibull++, ALTA and BlockSim, and has published numerous papers on various reliability methods. Mr. Mettas holds a B.S. degree in Mechanical Engineering and an M.S. degree in Reliability Engineering from the University of Arizona.

Georgios Sarakakis
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, Arizona 85710, USA

e-mail: Georgios.Sarakakis@Reliasoft.com

Georgios Sarakakis is a research scientist at ReliaSoft Corporation, where he is involved with the development of ReliaSoft's software products, publications and consulting projects. His research interests include reliability and quality management and the development of reliability program plans. Mr. Sarakakis received his M.S. in Reliability Engineering from the University of Arizona. He is a Certified Reliability Engineer, Certified Project Management Professional and Certified Six Sigma Green Belt. Prior to joining ReliaSoft, he worked as a reliability engineer at Hewlett-Packard.

Pengying Niu
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, Arizona 85710, USA

e-mail: Pengying.Niu@Reliasoft.com

Pengying Niu is a research scientist at ReliaSoft Corporation. She played a key role in the development of Lambda Predict. Before joining ReliaSoft, she worked at Texas Instruments where she was involved in IC design and testing. She received her Masters degree from the National University of Singapore and M. E in Electrical and Computer Engineering from the University of Arizona. She has done extensive work on AC/DC and DC/AC converters. Her current research interests include reliability prediction and physics of failure for electronic components such as MOSFET, IGBT and electronic systems.