Reliability Analysis Based on Field Data

Jang-wook Hur, Dhominick Pagulayan, Bo-Hyeon Kim, Kyeong-Hwan Hwang

Abstract—Modeling reliability can be very hard especially when several methods have already been established. Most methods are very specific to what type of component is being analyzed which is not very convenient. In this paper, the failure data is used to simulate the reliability of components to come up with a function which describes the reliability of the components. The system reliability can be computed by combining the components' reliability. Finally, the reliability function of a field data can be predicted using component data tested under standard conditions and an environment derived using cox proportional hazards model.

Keywords—Field Data, Reliability, Weibull Distribution

I. INTRODUCTION

Reliability, a measure of the probability of a component in performing its desired function, is important because it predicts the time that failures will occur thus reducing the probability of having a failed component during critical times. This measure must include confidence intervals as well as environment factors that can affect the system's reliability. This paper also discusses how censored data can be manipulated in getting the reliability function of the component with censored data.

First, the method of modeling reliability is described which can be applied to several probability distributions. Next, the Weibull distribution is applied to the equations described in the general model because this distribution has characteristics that make the modeling easier. An example calculation is shown at the end to show how the equations can be used and what functions can be obtained using the model described.

II. RELIABILITY

Different methods have been established in calculating the reliability of different components or systems but most of them are designed only for certain types of components. These methods aim to include every factor affecting the reliability of a component and try simulating the effects of each factor to produce a model describing the reliability of components. This

method is useful for critical components but it can be very expensive and time-consuming

A. Data Classification

Before modeling, the data should first be classified to determine if it is complete or has a type or combination of censorship. A complete data must have the exact failure time of all components that were tested. If some components survived after testing time, then it is classified under right censored data. If there is no exact time of failure but the interval of where the failure happened is known it is considered to have an interval censorship [1]. A special case of interval censorship is left censorship wherein the start of the interval is zero but these will be considered as interval censorship.
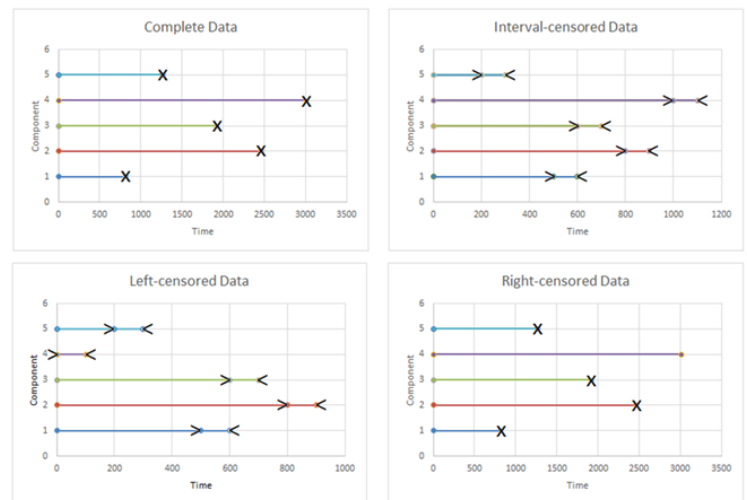

Fig. 1 Different types of censorships

B. General Reliability Model

Since reliability can predict the times at which failures will occur, it can also be said that the reliability of a component is computed based on its failure data. This means that after testing a few number of components, the failure of the same component can already be predicted. To determine the reliability of a system, the failure data from tests should first be analyzed by matching the failure data to a distribution. The data can be presented differently depending on what kind of test the component experienced and sometimes the data can be censored in such a way that the exact time of failure is not known. These are the things to consider but the important thing is that the nature of the data matches the nature of the equation of the distribution. If the data is presented as a probability distribution, then, the probability density equation should be used. The cumulative

Jang-wook Hur, Kumoh National Institute of Technology, South Korea hhjw88@gmail.com

Dhominick Pagulayan, Kumoh National Institute of Technology, South Korea. dapagulayan17@yahoo.com

Bo-Hyeon Kim, Kumoh National Institute of Technology, South Korea koc343434@gmail.com

Kyeong-Hwan Hwang , Kumoh National Institute of Technology, South Korea hkhlove89@gmail.com

distribution equation is used if the data is presented in a cumulative manner. Below is a table of some distributions used in reliability with β and η as the shape and scale parameters of the distribution respectively. The cumulative distribution function F(t) is used match the failure data while the probability distribution while the probability distribution function f(t) is used to know how many components will fail at a certain time.

TABLE I
EQUATIONS OF FEW DISTRIBUTIONS

Distribution	f(t)	F(t)
Weibull	$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}}$	$1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}}$
Exponential	$\beta e^{-\beta(t-\eta)}$	$1 - e^{-\beta(t-\eta)}$
Gamma	$\frac{1}{\Gamma(\beta)\eta^{\beta}} t^{\beta-1} e^{-\frac{x}{\eta}}$	$\frac{1}{\Gamma(\beta)}\gamma\left(\beta, \frac{t}{\eta}\right)$

The type of distribution that is used depends mainly on the type of component that is tested because some components follow a certain distribution but any type of distribution can be used as long as there is a way to calculate the scale and shape parameters because approximating them can be very tedious if not using the right software. There are many methods is calculating for parameters but the basic idea is that the error between the failure data and the distribution is minimized [2]. Finally, the predicted reliability and failure rate of the component can be calculated using the following equations:

$$R(t) = 1 - F(t) \tag{1}$$

$$\lambda(t) = \frac{f(t)}{R(t)} \tag{2}$$

where R(t) is the reliability of the component with respect to time, f(t) and F(t) are the probability and cumulative distribution function respectively.

C. Finding Confidence Interval

In the previous section, the parameters of the reliability function were only an approximation and the resulting curve does not take into account every point on the failure data, thus, it is needed to have a confidence interval that will project the predicted data from the failure data. One method is the likelihood ratio confidence bounds which can compute the upper and lower confidence levels of a given failure data. This method is more tedious than the previous section since two equations of probability distribution equations are needed to be approximated using the equation:

$$L(\theta) = \prod_{i=1}^{N} f(x_i, \theta) \tag{3}$$

$$L(\theta) = L(\hat{\theta}) * e^{\frac{-x_{\alpha;1}^2}{2}} \tag{4}$$

where L(θ) is the likelihood function, θˆ is the estimated parameter vector obtained from the previous section, θ is the new parameter vector that is to be estimated that will describe the different possible boundaries of the corresponding distribution, and $X^2_{\alpha;1}$ is the chi-square statistic value corresponding to the level of confidence that is to be used.

D. Calculating System Level Reliability

Reliability of a system can be calculated when the reliability of its components is known [3]. Two components that are used in series has reduced reliability than both components since that configuration requires both components to function in order to work. Parallel configurations, on the other hand, requires only one component to work which is why this configuration greatly increases the reliability of the system as a whole. The system reliability is also approximated numerically using the equations:

$$R_{series}(t) = \prod_i^N R_i(t) \tag{5}$$

$$R_{parallel}(t) = 1 - \prod_i^N\left(1 - R_i(t)\right) \tag{6}$$

Other systems that have a combination of parallel and series configurations can be decomposed into different subsystems, solved separately, and assembled until the system level reliability is computed. The confidence bounds are also computed using the equations above which is shown in the example.

E. Integrating Environment Factors

A potential problem with this method in calculating the reliability of a component is that when the component is used in different environment factors such as extreme temperatures, pollution, and etc. The reliability of the component can change significantly which makes the prediction of reliability harder. Various methods for computing for the effect of the environment has been used by references [4-8]. A common method for integrating the environment factor to the reliability equation is by using cox proportional hazards model:

$$\lambda(t) = \lambda_0(t)e^{\sum \alpha_i x_i} = \lambda_0(t)E(t) \tag{7}$$

where λ0(t) is the failure rate at standard environment conditions, λ(t) is the failure rate at the new environment, and eαX is the environment factor which transforms the reliability function from the standard conditions to the new reliability function that predicts the reliability of the component at the environment conditions. The purpose of getting the environment factor is that a different component with known reliability function at standard conditions can be used to predict the reliability of the said component in the new environment.

III. MODELLING USING WEIBULL DISTRIBUTION

The most common distribution used in modeling reliability is Weibull distribution because it can model components with increasing failure rate as well decreasing failure rate. The other reason is that if Beta is set to one, it has the same form as the exponential distribution. The general form of the Weibull distribution has three parameters but since it can be assumed that there are no failures initially, the two-parameter Weibull is a good estimation for the failure data.

A. Weibull Distribution Characteristics and Data Modeling

Based on equations (1) and (2), the reliability and failure rate

function of the Weibull distribution is given by:

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad (8)$$

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \qquad (9)$$

where β is the shape parameter and η is the scale parameter of the Weibull distribution. If there is an interval censorship in the data, the failure time t_i will be the average of the time interval. The left-censored data which is the special case of the interval censored data will have the failure time of component said component to be half of the time the failure was discovered. For right censored data, termination time will be the failure time of the components that did not fail at the end of the test. There is no need to modify the complete data since the failure time of all components are known.

The cumulative distribution function (CDF) given in Table 1 will be used in modeling the failure rate thus the failure data must be arranged in increasing time of failure. At time zero, the cumulative failure is zero and at the time equal to infinity, the cumulative failure is equal to one. This means that increment between the values of the cumulative failure should be 1/(n+1) where n is the total number of components tested. To get the parameters of the Weibull distribution, the CDF should be first be linearized using mathematical manipulations:

$$F(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad (10)$$

$$\ln(1 - F(t)) = -\left(\frac{t}{\eta}\right)^{\beta} \qquad (11)$$

$$\ln\left[\ln\left(\frac{1}{1-F(t)}\right)\right] = \beta \ln(t) - \beta \ln(\eta) \qquad (12)$$

where:

$$x = \ln(t) \qquad (13)$$

$$y = \ln\left[\ln\left(\frac{1}{1-F(t)}\right)\right] \qquad (14)$$

$$m = \beta \qquad (15)$$

$$b = -\beta \ln(\eta) \qquad (16)$$

where x and y corresponds to the x and y-axis of the Cartesian coordinate, m is the slope of the line, and b is the x-intercept of the graph. The failure time and the cumulative failure is then converted into its x and y-values based on equations above. Figure 2 shows an example of linear plot of the cumulative distribution function of the Weibull distribution. The slope of the trend line is the value of shape parameter while scale parameter can be obtained using the x-intercept of the trend line.
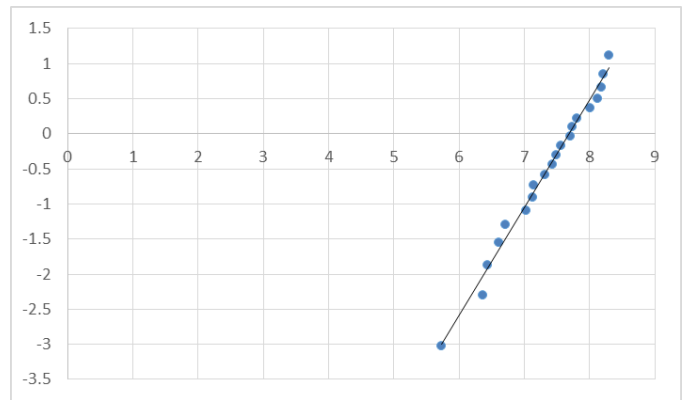

Fig. 2 Linearized cumulative distribution

B. *Weibull Distribution Confidence Interval*

In getting the confidence interval for Weibull distributions, the likelihood function based on equation (3) will be:

$$L(\beta, \eta) = \prod_{i=1}^{N} \frac{\beta}{\eta}\left(\frac{t_i}{\eta}\right)^{\beta-1} e^{-\left(\frac{t_i}{\eta}\right)^{\beta}} \qquad (17)$$

where β and η are the shape and scale parameters, ti is the failure time for each component. When using equation (4), there are a lot of combinations of beta and eta which is shown in figure 3. The two values of the scale parameters that has the same shape obtained from the previous section. The reason behind this is that when the shape parameters are not the same, the failure function will overlap to each other.
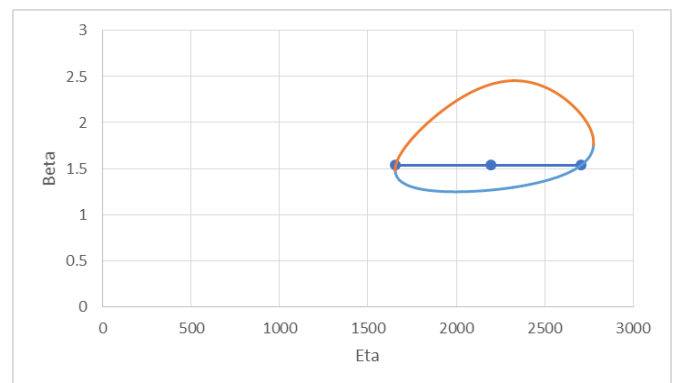

Fig. 3 Different combinations of parameters using equation (4)

Mathematical manipulations of the Weibull's cumulative distribution function can't obtain the exact parameters of the system level reliability function. It can only be approximated using the method in subsection A.

C. *Environment Factor in Weibull Distribution*

Using equation (7) and (9), the following equation below can be obtained. As can be seen from the equations, it is impossible to the equation to be true unless β_0 and β are equal.

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} = \frac{\beta_0}{\eta_0}\left(\frac{t}{\eta_0}\right)^{\beta_0-1} E(t) \qquad (18)$$

In order to make the equation true, the environment factor must be modified like this:

$$E(t) = t^x e^y \qquad (19)$$

where:

$$x = \beta - \beta_0 \tag{20}$$

$$y = \ln(\beta) - \ln(\beta_0) + \beta_0 \ln(\eta_0) - \beta \ln(\eta) \tag{21}$$

IV. EXAMPLE

In illustrating the method above, a fictional data is used to demonstrate on how to do a Reliability analysis which will show its effectiveness in tackling different practical applications in the field. The first set of data is a complete data which tells the exact time in which 20 different components of the same type failed during its tests.
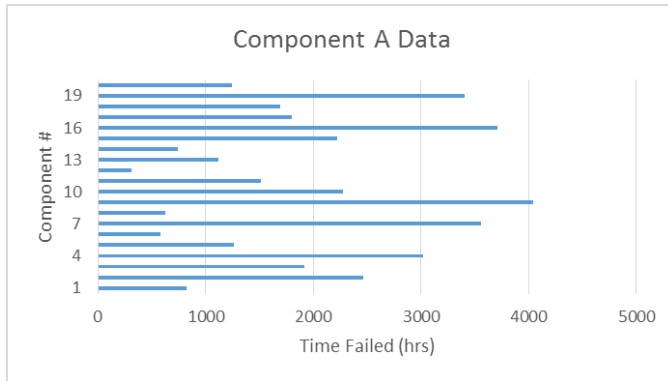
Fig. 4 Testing of component A

Figure 4 describes how each of the 20 items survived with component 12 failing very early and component 9 which failed last. The data were arranged from lowest to highest and the cumulative failures were calculated so that the data will match the cumulative distribution function of the Weibull distribution. A code in C++ was created in order to calculate the optimal values for the parameters of the Weibull distribution which were $\beta = 1.537905411$ and $\eta = 2193.819371$. The confidence bounds were computed for 90% confidence level which has a chi-square value of 2.705543. There are many possible combinations of β and η that can satisfy equation (4) but the trick here is that the lowest and highest possible η value should be calculated first and substituting them back to equation (4) to get their corresponding beta.
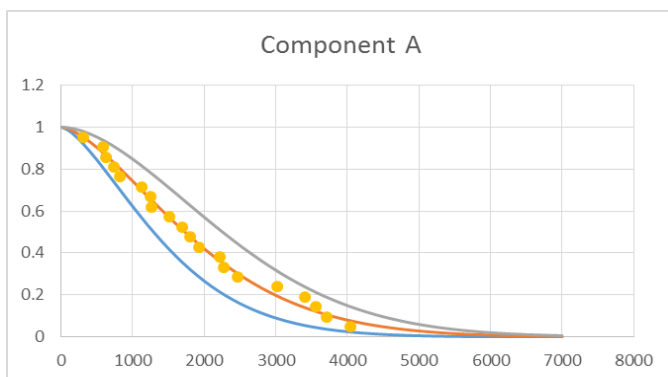
Fig. 5 Reliability of component A with 90% confidence interval

Using the same method as above, the distribution parameters of an interval censored for component B is computed which were

approximated as $\beta = 2.302464541$ and $\eta = 2855.723269$ which is also shown in figure 6.
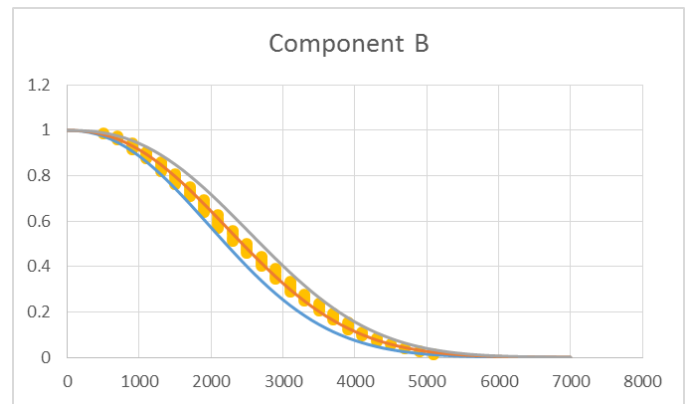
Fig. 6 Reliability of component B with 90% confidence interval

The lower confidence level scale parameter for component A was computed to be equal to 1657.22420664 while the lower confidence level scale parameter component B was 2706.58576219. The higher confidence level scale parameter $\eta = 2600.94358869$ for component A and $\eta = 3086.259706$. These values are only and approximate and are not the exact. The graph also shows a comparison of both components which is very helpful when choosing what components to use for the system.

Figure 7 shows a system composed of component A and B in series. The approximate beta and eta computed using equation (5) are 1.70179837999 and 1545.89431 respectively. The confidence interval was also approximated using equation (5) which $\eta_{low} = 1347.70982466$ for the lower confidence level and $\eta_{high} = 2143.15852795$ for the upper confidence level.
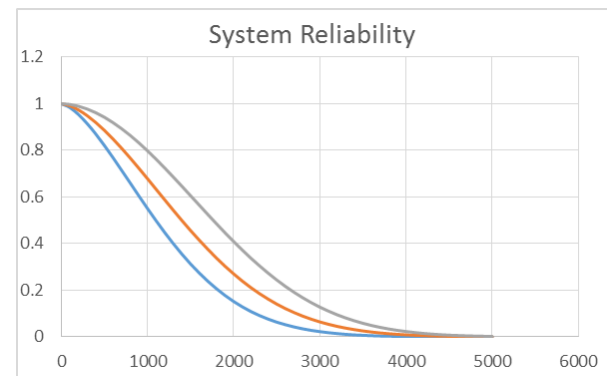
Fig. 7 System level reliability

A field data for component A was created and is shown in Figure 8. The new data has a higher failure function and the parameters were approximated with values of $\beta = 1.465874822$ and $\eta = 1292.743913$. Using equation (17), the environment factor parameters were approximated to be $x = -0.07203059239$ and $y = 1.2814584236615$.
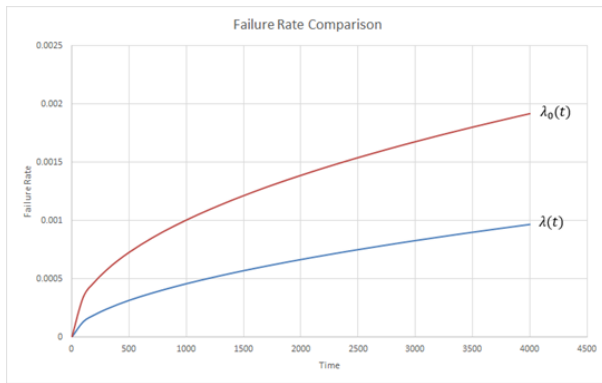
Fig. 8 Component A subjected with and without environment factors

V. CONCLUSION

All computations were very tedious especially when several iterations were needed. The approximated values were not perfect specifically in approximating the cumulative distribution of the system but with the help of a good computer program, the above method can be applied actual field data to test whether the predictions will be accurate or not.

APPENDIX

TABLE II
COMPONENT A DATA

Component #	Time Failed (hrs.)	Component #	Time Failed (hrs.)
1	821	11	1514
2	2460	12	310
3	1921	13	1122
4	3017	14	744
5	1263	15	2222
6	580	16	3707
7	3558	17	1804
8	627	18	1696
9	4045	19	3403
10	2273	20	1246

TABLE III
COMPONENT B DATA

Time(hours)	Failures	Time(hours)	Failures
400-600	1	2800-3000	4
600-800	2	3000-3200	4
800-1000	3	3200-3400	3
1000-1200	3	3400-3600	3
1200-1400	4	3600-3800	3
1400-1600	4	3800-4000	3
1600-1800	4	4000-4200	2
1800-2000	5	4200-4400	2
2000-2200	5	4400-4600	1
2200-2400	4	4600-4800	1
2400-2600	4	4800-5000	1
2600-2800	4	5000-5200	1

TABLE IV
COMPONENT A WITH ENVIRONMENT FACTORS

Component #	Time Failed (hrs.)	Component #	Time Failed (hrs.)
1	147	11	1138
2	279	12	1244
3	369	13	1316
4	455	14	1355
5	542	15	1337
6	627	16	1388
7	714	17	1890
8	800	18	2041
9	902	19	2193
10	1012	20	2777

REFERENCES

[1] ReliaSoft Corporation. (2015). Life Data Analysis Reference Book [PDF].Retrieved August 7, 2016, from ReliaWiki.org

[2] Walck, C. (2007). Hand-book on STATISTICAL DISTRIBUTIONS for experimentalists. University of Stockholm.

[3] Guo, H., Honecker, S., Mettas, A., & Ogden, D. (2010). Reliability estimation for one-shot systems with zero component test failures. 2010 Proceedings - Annual Reliability and Maintainability Symposium (RAMS). https://doi.org/10.1109/RAMS.2010.5448016

[4] Teng, X., & Pham, H. (2006). A New Methodology for Predicting Software Reliability in the Random Field Environments. IEEE Transactions on Reliability IEEE Trans. Rel., 55(3), 458-468. doi:10.1109/tr.2006.879611 https://doi.org/10.1109/TR.2006.879611

[5] Zhenyu, L., Xiaobing, M., & Yu, Z. (2013). Field reliability prediction based on degradation data and environmental data. 2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS). doi:10.1109/rams.2013.6517755 https://doi.org/10.1109/RAMS.2013.6517755

[6] Johnson, D., Coit, D. W., Kosaka, R., & Megow, K. (2013). System level reliability analyses and predictions in a varying stress environment. 2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS). doi:10.1109/rams.2013.6517617 https://doi.org/10.1109/RAMS.2013.6517617

[7] Kumar, E. S., & Sarkar, B. (2013). Improved modeling of failure rate of photovoltaic modules due to operational environment. 2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT). doi:10.1109/iccpct.2013.6528850 https://doi.org/10.1109/ICCPCT.2013.6528850

[8] Ghodrati, B., & Ahmadi, A. (2013). Product support logistics based on system reliability characteristics and operating environment. 2013 IEEE International Conference on Industrial Engineering and Engineering Management. doi:10.1109/ieem.2013.6962453 Ross, S. M. (2014). Introduction to Probability Models (11th ed.). https://doi.org/10.1109/IEEM.2013.6962453 Elsevier Korea LLC.