Investigation of failure times parameter through standard and mixture Weibull distribution

M N Ariffin¹,², A Shafie¹,³

¹ Fundamental of Applied Sciences, University of Technology PETRONAS, Malaysia
² muhammad_16000560@utp.edu.my
³ afza@utp.edu.my

Abstract. Reliability and survival analysis is one of the main concerns in the industry. The main problem in reliability study is to capture the information of failure behaviour of the components or the machines. This paper investigates the failure behaviour of the cutter which is the main component of leveller machine. There are three main stages of failure behaviour: early-failure stage, constant failure stage and wear-out failure stage. The failure behaviour is fitted with two distributions, the standard Weibull distribution and the mixture Weibull distribution. Mean Square Error and \(-2\ln(L(\theta))\) are used as the goodness-of-fit test. The results show that the combination shape parameters from both distributions is able to relate to the three failure stages of the cutter.

1. Introduction
Reliability analysis have become one of the key factors in industrial applications. The application of reliability helps in prediction and decision making. This is also supporting with the objective of Industrial 4.0. In the manufacturing industry, reliability has been used to determine the probability that a component manufactured will fail under a given environment. If the reliability is low, the machines or the components are expected to fail. Thus, changing the material, the process of manufacturing, or redesigning might be the alternatives that manufacturer might need to consider to improve reliability. Understanding the reliability of the components and machines helps the manufacturer to predict the next breakdown and take necessary action.

Numerous studies have been conducted to understand the failure behavior based on the failure data from the machines or the components. The reason is to identify the time-to-failure and improve the reliability of the machines or components. However, this study focuses on monitoring the reliability to predict the next downtime. Statistically, the behavior of failure can be measured by fitting with the parametric distribution, such as Weibull distribution. Chauhan and Malik derive the reliability model and mean time to system failure (MTSF) based on Weibull distribution [1]. The results show that the value of reliability and MTSF is high when the components follow Rayleigh distribution compared to Weibull distribution. However, the system shows to have more reliability when the components follow Weibull distribution.

Based on the discussion above, reliability study involves understanding the failure behavior through the distribution function. In real life application, there will be more than one causes of failure for the machine or the components. This shows that the failure data is a heterogeneous data failure. If standard Weibull distribution is applied, the result only shows one failure behavior. The main question arises, “Are the standard distribution enough to represent the failure data with more than one causes?” Therefore, the main problem in the reliability study is to identify the suitable distribution to represent
the failure data. Thus, the main objective of this paper is to investigate the failure behavior of the cutter from a leveller machine. The failure behavior will be fitted with the standard Weibull and mixture Weibull distributions using Maximum Likelihood Estimator (MLE). Cutter has been chosen because the cutter is the main component of the leveller machine. Downtime data was collected from 2007 until 2013. These data were divided into two groups, the planned and unplanned downtime. The reason was to compare the parameters affecting the cutter failure for the planned and unplanned downtime. Studies have shown that the unplanned downtime created more losses compared to the planned downtime [2, 3].

The organization of this paper is as follows. Section 1 is the introduction of the paper and Section 2 discuss the literature review of this paper with some theoretical explanations. Section 3 discuss the description of the data, the methodology and the process of analyzing the data with the application of Weibull and mixture Weibull distribution. The parameters are estimated by MLE. Results and discussions are show in Section 4. Finally, conclusions are given in Section 5.

2. Literature Review

Nowadays, the knowledge of the failure behavior of machines and components are crucial. These can be seen from the research that studies the failure of machines, such as the load haul dumper [4] and the failure of components, such as the wind turbine components [5] and the milling components [6]. In statistical analysis, one of the methods to investigate the behavior of data is by using lifetime distribution analysis. The most common lifetime distribution that involves failure data is Weibull distribution [7, 8]. Weibull distribution is a versatile distribution that can take on the characteristics of other type distributions, based on the shape parameter, $\beta$. The Weibull had wide applications of the lifetime of manufactured items and biological and medical applications [9]. Ossai et. al. used the Weibull distribution to develop six state Markov maintenance model [5]. The result shows that mechanical system have the highest maintenance downtime and unavailability rates with time. Besides that, Yuan et. al. develop accelerated Bayesian degradation model that follow Weibull distribution [6]. The proposed model was able to estimate the parameter of the Bayesian Weibull distribution with nine milling heads.

![Figure 1. The three failure stages; early-failure stage, constant failure stage and wear-out failure stage [10]](image)

Figure 1 shows the illustration of the three common failure stages that follow the Weibull distribution. If $\beta < 1$, the failure behavior is monotone decreasing also known as early-failure stage, if $\beta = 1$, the failure behavior is known as the constant failure stage and if $\beta > 1$, the failure behavior is monotone increasing that is also known as the wear-out failure [9]. Even though the Weibull distribution is widely used in the lifetime distribution, the failure behavior of the Weibull distribution is only monotone behavior [4].

This is the reason that standard Weibull distribution may not fit well with the data that have more than one failure behavior [11]. The information provided by the parameters of standard Weibull distribution may not be enough to explain the failure behavior. The used of mixture distributions are
widely discussed by researchers in a number of applications, statistical settings and also to solve the issues from the standard Weibull distribution [12, 13]. The mixture distribution arises in practical problems when the measurements of a random variable are taken under two or more different conditions [14] and the mixture distribution performed better than the conventional or standard distribution [7]. Castet and Saleh [12] found that the 2-Weibull mixture distribution were able to capture satellite early-failure stage and satellite wear-out failure stage. The authors also suggest that if the researchers intend to get precision result, mixture Weibull is more appropriate. Zhang et. al. [13] have proposed a mixture Weibull proportional hazard model to predict the failure time of a mechanical system with multiple failure modes. Elmahdy [7] also obtain the two failure stages from the failure data.

2.1. Standard Weibull distribution
The standard Weibull distribution developed by the Waloddi Weibull discussed the application of the statistics in wide field problems [8]. The probability density function (pdf) and reliability function of the standard Weibull distribution is shown below:

\[ f(x; \alpha, \beta) = \frac{\beta}{\alpha} \left( \frac{x}{\alpha} \right)^{\beta-1} e^{-\left( \frac{x}{\alpha} \right) \beta}, \quad x \geq 0, \alpha, \beta > 0 \]  \hspace{1cm} (1)

\[ R(x; \alpha, \beta) = e^{-\left( \frac{x}{\alpha} \right) \beta}, \quad x \geq 0, \alpha, \beta > 0 \]  \hspace{1cm} (2)

where the \( \alpha \) represents the scale parameter, \( \beta \) represents the shape parameter and \( x \) represents the failure data or time-to-failure data. In reliability studies, the interest parameter is the shape parameter. This is because the shape parameter describes the failure behavior of the components. There are three stages of failure, the early-life failure, constant failure and wear-out failure. The scale parameter is also known as the characteristics of life.

2.2. Mixture Weibull distribution
Mixture distribution is a distribution made of combining two or more component using mixing parameters (proportions) which are non-negative and if expressed as fractions of the mixture. The mixture Weibull distribution with the reliability function is described as below:

\[ f(x; \omega_j, \alpha_j, \beta_j) = \sum_{j=1}^{k} \omega_j \left( \frac{\beta_j x^{\beta_j-1}}{\alpha_j^{\beta_j}} e^{-\left( \frac{x}{\alpha_j} \right) \beta_j} \right), \quad x \geq 0, \alpha_j, \beta_j > 0, \omega_j > 0 \]  \hspace{1cm} (3)

\[ R(x; \omega_j, \alpha_j, \beta_j) = \sum_{j=1}^{k} \omega_j \left( e^{-\left( \frac{x}{\alpha_j} \right) \beta_j} \right), \quad x \geq 0, \alpha_j, \beta_j > 0, \omega_j > 0 \]  \hspace{1cm} (4)

where the \( \alpha_j \) represents the scale parameter, \( \beta_j \) represents the shape parameter \( \omega_j \) represents the mixing weight, \( x \) represents the failure data or time-to-failure data and \( \sum_{j=1}^{k} \omega_j = 1 \). The mixed Weibull distribution is the appropriate model when the machines or components has two or more failure causes. This occurs in many situations, such as both early failures stage and chances failures might be involved in a burn-in test and also in the case of quality control mode with an infant mortality followed by a wear out mode [7]. Castet and Saleh [12, 15], Zhang et. al. [13], Razali and Al-Wakeel [11] and others researcher discussed on the application of the mixture distribution.

2.3. Maximum Likelihood Estimator
The following maximum likelihood estimation is derived based on Lawless [9]. Suppose that the potentially observable data in a study is distributed according a probability distribution defined by the parameter vector \( \theta \) where \( \theta = \{ \alpha, \beta \} \). Then, calling \( X \) as the failure data actually observed, the likelihood function for \( \theta \) based on these data is
where \( \Pr \) represents the probability density or mass function from which the observed data are assumed to arise. When the probability density function has a parametric form \( f(x; \alpha, \beta) \), and the independent and identically distributed lifetimes \( x_1, \ldots, x_n \) for a random sample of \( n \) individuals are observed, equation (1) and equation (5) can be rewritten as:

\[
L(x_1, x_2, \ldots, x_n; \alpha, \beta) = \prod_{i=1}^{n} \left( \frac{\beta}{\alpha} \right)^{\frac{x_i}{\alpha}} \left( \frac{1}{\alpha} \right)^{\beta-1} e^{-\left( \frac{x_i}{\alpha} \right)^{\beta}}
\]

while the log-likelihood function is

\[
l(x_1, x_2, \ldots, x_n; \alpha, \beta) = n \log(\beta) - n \beta \log(\alpha) + (\beta - 1) \sum_{i=1}^{n} \log(x_i) - \sum_{i=1}^{n} \left( \frac{x_i}{\alpha} \right)^{\beta}
\]

Maximizing the log-likelihood function \( l(x_i, \alpha, \beta) \) or minimizing the \(-l(x_i, \alpha, \beta)\) from equation (7) is carried out to solve the equation.

3. Methodology

The data for this study was collected from 2007 until 2013. The downtime data is from the cutter that is the main component in the leveller machines. The downtime data represent time (in hour) between downtime and a complete dataset. A complete dataset means that the failure times of each downtime was recorded from all the samples.

The data is divided into two groups, the planned and unplanned downtime. The time to set up the cutter and the time to change the broken cutter are classified as planned downtime while the other types of downtime is classified as the unplanned downtime. The general cutter downtime, planned downtime and unplanned downtime datasets are being used in this study.

Next, the data is fitted by using the Kaplan-Meier non-parametric estimator. The Kaplan-Meier estimator will estimate the reliability calculation based on the data. Then, the standard Weibull distribution estimated by the MLE will be fitted. The process is repeated for the mixture Weibull distribution. The Mean Square Error (MSE) and -2ln(L(\(\theta\))) from the maximum likelihood model will be used to compare the performance of standard and mixture Weibull distribution. MATLAB software is used for all the calculations.

4. Results and discussion

Figure 2 shows the reliability behavior of the lifetime of the general cutter downtime, the planned and unplanned downtime using the Kaplan-Meier non-parametric technique. In general, the reliability of the three downtimes datasets seems to have the same behavior. The reliability decreases faster at the first 50 hours of operation.

![The Reliability Behavior](image)

**Figure 2.** The reliability behavior of cutter, planned and unplanned downtime based on the Kaplan-Meier estimator
Based on descriptive analysis in Table 1, there are 449 downtimes that involved the cutter, from the year 2007 until the year 2013. Based on the time-to-failure of general cutter, the average time-to-failure is 134.78 hours while the minimum and the maximum time-to-failure is 8.333 hours and 1406 hours, respectively.

Table 1. The descriptive statistics

|                          | General Downtime | Planned Downtime | Unplanned Downtime |
|--------------------------|------------------|------------------|---------------------|
| Sample size              | 449              | 311 (69.265%)    | 138 (30.735%)       |
| Mean (hours)             | 134.78           | 130.32           | 144.84              |
| Minimum (hours)          | 8.333            | 8.333            | 16                  |
| Maximum (hours)          | 1406             | 1406             | 1324.5              |
| Q_{0.025}(hours, Reliability) | 17.4 (83.88%)   | 15.2 (93.18%)   | 16.3 (67.31%)       |
| Q_{0.25}(hours, Reliability) | 83 (35.75%)     | 67 (40.84%)     | 57.5 (39.49%)       |
| Q_{0.5}(hours, Reliability) | 168.5 (20.71%)  | 140 (23.63%)    | 159 (24.28%)        |

Table 1 also shows the reliability estimates by the Kaplan-Meier estimator. The reliability of the cutter decrease to 83.88% with the time-to-failure 17.4 hours at the first 2.5% quartile of the data. As the quartile increase to 25% and 50%, the reliability of the cutter rapidly decreases to 35.75% and 20.71% with time-to-failure 83 hours and 168 hours, respectively. This support the graphical interpretation in Figure 2, where the reliability shows to decrease faster at the first 50 hours of operation. This implies that the general cutter lifetime has more than 80% reliability at the early 20 hours of operation and decreases to 40% reliability after 80 hours of operation.

From the descriptive analysis, 311 represent by the planned downtimes while the other 138 unplanned downtimes. The average of planned downtimes has lower time-to-failure with 130.32 hours compare to unplanned downtime with 144.84 hours. However, the lifespan of planned downtime shows to have longer hours from 8.333 hours to 1406 hours compare to unplanned downtime from 16 hours to 1324.5 hours. Furthermore, at the first 2.5% quartile, the planned downtime has the highest reliability with more than 90% compare to the unplanned downtime before decreasing to less than 70%. As the quartile increase to 25% and 50%, there is not much different in the reliability value between planned and unplanned downtime.

![Figure 3. The standard Weibull (blue line) and the mixture Weibull (brown line) are fitted based on the general downtime. The ECDF is the fitted Kaplan-Meier (red line).](image-url)
Figure 3 shows the fitted Kaplan-Meier estimator versus the standard and mixture Weibull distributions based on the general downtime. Figure 4 and Figure 5 show the fitted Kaplan-Meier estimator versus the standard and mixture Weibull distributions based on the planned and unplanned downtime, respectively. From the figures, the mixture Weibull shows to have a better fitting compared to the standard Weibull.

![Figure 4](image-url) **Figure 4.** The standard Weibull (blue line) and the mixture Weibull (brown line) are fitted based on the planned downtime. The ECDF is the fitted Kaplan-Meier (red line).

![Figure 5](image-url) **Figure 5.** The standard Weibull (blue line) and the mixture Weibull (brown line) are fitted based on the unplanned downtime. The ECDF is the fitted Kaplan-Meier (red line).

These results are supported by Castet and Saleh [12] and Elmahdy [7]. The most important is the parameters estimation for the standard and mixture Weibull distributions as shown in Table 2, where $\alpha$, as the scale parameter and $\beta$, as the shape parameter.

MSE and -2ln(L($\theta$)) were used as the goodness of fit tests. From the results, the MSE obtain by the standard Weibull distribution is slightly high compare to the mixture Weibull distribution for all three datasets. The MSE value for standard Weibull distribution is between 0.42% and 0.88% while the MSE value for mixture Weibull distribution is between 0.09% and 0.18%. The best model is the model with the lowest -2ln(L($\theta$)) value [7]. The results show that the -2ln(L($\theta$)) value of the mixture Weibull is lowest compare to the standard Weibull for the general, planned and unplanned downtime. This shows that the mixture Weibull is good distribution compared to standard Weibull distribution.

Based on the general downtime, the standard Weibull estimate that $\beta$ is 0.78914 while the mixture Weibull estimate that $\beta_1$ is 7.6227 and $\beta_2$ is 0.92279. Based on explanation in Figure 1 about the hazard plot, the combination information from the parameters from both standard and mixture Weibull distributions can explain all the three failure stages of the cutter. The standard Weibull capture $\beta<1$ and this explain the early-failure stage of the cutter. Meanwhile, the mixture Weibull capture $\beta_1>1$ and $\beta_2=1$ which explain the wear-out failure stage and constant failure stage of
the cutter, respectively. Based on the combinations information from both distributions, the scale parameters explain that 63.2% of the cutter population will have downtime after about 114 hours of operation, 180 hours of operation and 21 hours of operation if the cutter in early-failure, constant failure and wear-out failure stages, respectively.

Table 2. The parameter estimation based on the standard and mixture Weibull distribution

|                     | General Downtime | Planned Downtime | Unplanned Downtime |
|---------------------|------------------|------------------|---------------------|
| Standard Weibull    | MSE = 0.42%      | MSE = 0.47%      | MSE = 0.88%         |
|                     | -2ln(L(θ)) = 5245.06 | -2ln(L(θ)) = 3630.79 | -2ln(L(θ)) = 1607.36 |
|                     | α = 113.926      | α = 116.373      | α = 107.536         |
|                     | β = 0.78914      | β = 0.841274     | β = 0.697499        |
| Mixture Weibull     | MSE = 0.18%      | MSE = 0.09%      | MSE = 0.14%         |
|                     | -2ln(L(θ)) = 4973.00 | -2ln(L(θ)) = 3537.61 | -2ln(L(θ)) = 1518.96 |
|                     | α1 = 21.043      | α1 = 43.19       | α1 = 19.651         |
|                     | β1 = 7.6227      | β1 = 1.8837      | β1 = 6.0249         |
|                     | α2 = 180.45      | α2 = 244.18      | α2 = 256.89         |
|                     | β2 = 0.92279     | β2 = 1.0464      | β2 = 0.9459         |

Table 2 also shows the results for the planned and unplanned downtime for both standard and mixture Weibull distributions. The shape parameters from both distributions for planned and unplanned downtime shows the same information as the shape parameters from general cutter downtime. The standard Weibull distribution shows the $\beta$ that less than 1 with 0.841274 for planned downtime and 0.697499 for unplanned downtime. On the other hand, the mixture Weibull distribution gives $\beta_1 > 1$ and $\beta_2 < 1$ for both planned and unplanned downtimes. The $\beta_1$ that captured are 1.8837 for planned downtime and 1.0464 for unplanned downtime while $\beta_2$ that captured are 6.0249 for planned downtime and 0.9459 for unplanned downtime.

Based on the scale parameters, the cutter is expected to fail in the early-failure stages after about 116 hours of operation for the planned downtime and 107 hours of operation for unplanned downtime. On the other hand, the cutter is expected to be in constant failure stage, after about 244 and 257 hours of operation for the planned and unplanned downtime, respectively. However, if the cutter is in the wear-out failure stage, there are bigger difference in the expected time of downtime between planned and unplanned downtime. 63.2% of the cutter will have downtime after about 43 hours of operation while the unplanned downtime after 20 hours of operation. These results are constant with the previous studies [7, 12, 15].

5. Conclusion

This paper has investigated the failure behavior of general cutter downtime with planned and unplanned downtime. Based on the Kaplan-Meier estimator, the reliability of the cutter decreases to less than 70% due to the unplanned downtime when the machine is already operating more than 15 hours. The mixture Weibull distribution is a better distribution due the lowest MSE and -2ln(L(θ)). The main finding in this paper is the ability of standard and mixture Weibull distributions to relate with the early-failure stage, constant failure stage and wear-out failure stage of the cutter. These results can be used in planning the maintenance schedule for the cutter.

References

[1] S Chauhan and S Malik 2017 Evaluation of Reliability and MTSF of a Parallel System with Weibull Failure Laws Journal of Reliability and Statistical Studies; vol 10 137-48
[2] How Manufacturers Achieve Top Quality Performance. (Wall Street Journal); Available: http://partners.wsj.com/emerson/unlocking-performance/how-manufacturers-can-achieve-top-quartile-performance/
[3] 2014 Common Causes of Downtime. (Cognoscape); Available: https://www.cognoscape.com/common-downtime/
[4] B Samanta, B Sarkar and S K Mukherjee 2004 Reliability Modelling and Performance Analyses of an LHD System in Mining The Journal of The South African Institute of Mining and Metallurgy; vol 1-8
[5] C I Ossai, B Boswell and I J Davies 2016 A Markovian approach for modelling the effects of maintenance on downtime and failure risk of wind turbine components Renewable Energy; vol 96 775-83
[6] R Yuan, W Peng, H Li, L He and H-Z Huang 2011 Reliability estimation of milling head based on accelerated Bayesian degradation test. Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2011 International Conference on; (IEEE)
[7] E E Elmahdy 2015 A new approach for Weibull modeling for reliability life data analysis Applied Mathematics and Computation; vol 250 708-20
[8] W Weibull 1951 A Statistical Distribution Function of Wide Applicability Journal of Applied Mechanics; vol 293-7
[9] J F Lawless 2011 Statistical models and methods for lifetime data: (John Wiley & Sons);
[10] Wikipedia Wikipedia. Available: https://en.wikipedia.org/wiki/Bathtub_curve
[11] A M Razali and A A Al-Wakeel 2013 Mixture Weibull Distributions for Fitting Failure Times Data Applied Mathematics and Computation; vol 219 11358-64
[12] J-F Castet and J H Saleh 2010 Single versus Mixture Weibull Distributions for Nonparametric Satellite Reliability Reliability Engineering and System Safety; vol 95 295-300
[13] Q Zhang, C Hua and G Xu 2014 A Mixture Weibull Proportional Hazard Model for Mechanical System Failure Prediction Utilising Lifetime and Monitoring Data Mechanical Systems and Signal Processing; vol 43 103-12
[14] E E Elmahdy and A W Aboutahoun 2013 A new approach for parameter estimation of finite Weibull mixture distributions for reliability modeling Applied Mathematical Modelling; vol 37 1800-10
[15] J-F Castet and J H Saleh 2009 Satellite and satellite subsystems reliability: Statistical data analysis and modeling Reliability Engineering & System Safety; vol 94 1718-28