

Lifetime Distribution With a Limited Failure in Condition Based Maintenance: A Case Study

Burairah Hussin

I. INTRODUCTION

Today, the value of condition monitoring information is recognised by most of plant personnel. The main idea is that, the observed information can act as maintenance indicators, which could be used to describe the key relationship between equipment condition and a maintenance decision. The deteriorating of equipment condition can be modelled using indicator such as cumulative wear and residual time, (Hussin 2007). If the condition or the state of the equipment can be predicted, maintenance actions which include manpower, equipment and tools and spare parts can be planned and scheduled, (Duffua, 1997).

In this paper, we seek to solve the predicting problem of residual time of an item associated with their observed condition-monitoring data. According to Wang (2002), the residual time x_t at time t_t is defined as the remaining time at t_{i-1} minus the interval between t_i and t_{i-1} provided that the item has survived to t_i and no maintenance action has been taken since. In mathematical notation it can be written as follows

$$x_t = \begin{cases} x_{i-1} - (t_i - t_{i-1}) & \text{if } x_{i-1} > t_i - t_{i-1}, \\ \text{otherwise not defined} & \end{cases} \quad (1)$$

The probability distribution for the residual time given their condition monitoring history is calculated using filtering approach, (Wang and Christer, 2000) as

$$p(x_t | \mathfrak{S}_t) = \frac{p(y_t | x_t) p(x_t | \mathfrak{S}_{t-1})}{\int_0^\infty p(y_t | x_t) p(x_t | \mathfrak{S}_{t-1}) dx_t} \quad (2)$$

where,

y_t = observed condition monitoring data

$\mathfrak{S}_t = y_1, y_2, \dots, y_t$ i.e, history of the condition monitoring data

$p(y_t | x_t)$ is a probability distribution which established the relationship between the observed information y_t given the residual time x_t .

$$p(x_t | \mathfrak{S}_{t-1}) = p(x_t | x_{t-1}) p(x_{t-1} | \mathfrak{S}_{t-1})$$

Here, it is assumed that x_0 is the residual time measured since new and all measurements of the condition monitoring parameter y_0 used are also since new. Since \mathfrak{S}_0 is not available at t_0 in most cases, we could set $p(x_0 | \mathfrak{S}_0) = p(x_0)$, which is the probability distribution of the engine life. It can be shown that equation (2) can be determined recursively if $p(x_0 | \mathfrak{S}_0) = p(x_0)$ and $p(y_t | x_t)$ are known. The forms of $p(x_0)$ and $p(y_t | x_t)$ can be chosen from any distribution and finally assessed against the goodness of fit test.

In the case of $p(x_0)$, its estimated parameter is obtained from a failure data or by subjective assessment of experienced experts who are familiar with the equipment. If the equipment failure data is used, it poses the question of how we can obtain the parameters of the distribution given limited failure information, which are the main question addressed in this paper.

II. DATA ANALYSIS

This section describes the data collection process of 54 sets of maintenance data used by marine diesel engines. The dataset consists of three types of data, as follows

1. Lifetime data - the age in operating hours of the engine life before it has been replaced because of failure or preventive replacement.

The author is with Universiti Teknikal Malaysia, Melaka.
Email : burairah@utem.edu.my

2. Condition indicators – obtained from observed Spectrometric Oil Analysis Programme (SOAP) to assess the deterioration of the engine and the oil.
3. Maintenance events – indicates how the maintenance events is performed

At glance, the obtained data was unsorted, contained missing values, was incomplete and inconsistent which needed further explanations, as it is common in practice. In order to correlate the monitoring data and residual life, we had to model data the relationship between them, where it is assumed that a shorter residual life will produce a high value of monitoring indicators.

III. PARAMETER ESTIMATIONS

In this task, maximum likelihood estimator (MLE) is used to estimate the model parameters. MLE exhibits several criteria to be a good estimator (John, 1997) and attractive to use as well, (E-magazine, 2001). It is asymptotically consistent, which means that as the sample size increased, the estimates converge to the true values. It is asymptotically efficient, which means that for large samples, it produces the most precise estimates. It is also asymptotically unbiased, which means that for large samples, one expects to get the true value on average.

Furthermore, as the sample size increased, the distribution of the MLE tends to the Gaussian distribution with mean $\hat{\theta}$, where θ are the parameters to be estimated. This enables us to calculate the variance and covariance of the estimated parameters by calculating the inverse of the Fisher information matrix.

Unfortunately, it should be noted that the size of the sample data need to be sufficient in order to achieve these properties that are varied depending on the application. If limited sample data is used, the methods can be biased which can cause discrepancies in further analysis. Another way to carry out parameter estimation is by using the complete data with censored data as well. The likelihood function for parameter $\hat{\theta}$ is given by

$$L(\theta, t) \propto \prod_{i \in U} f(\theta, t) \prod_{i \in C} S(\theta, t) \quad (3)$$

where $f(\theta, t)$ and $S(\theta, t)$ are the probability density and survival probability of the chosen distribution, U is the set of items which are observed to fail and C is the set of right-censored observations.

Applying this approach to the available data, we estimate the distribution parameters and compute the expected life of an engine. However, it is noticed that the result was far (38,064 hours) from the recommended engine life (20,150 hours) given by the personnel dealing with this data. Here, we could argue about what is the definition of an engine life: it either fail completely or fail to perform its intended functions. It became clear at this stage that the censored data we had in terms of preventive replacement was not random but purposely censored because of the definition of functional failure.

Due to limited failure data and purposely-censored data, we proposed two approaches that could work to estimate the lifetime distribution given the incomplete life data as follows:

1. Using failure and interval-censored information.
2. Predicting failures of those censored using a regression model based on existing failure data.

Interval-censored data arises when a failure time cannot be observed, but can be estimated to lie within an interval. This assumes that the life of a censored engine can only survive up to the maximum point, t_{max} , hence the likelihood function can be written as

$$L(\theta, t) \propto \prod_{i \in U} f(\theta, t) \prod_{i \in C} S(\theta, t) - S(\theta, t_{max}) \quad (4)$$

In our case, we could choose $t_{max} = \max(t_{i \in C}) + K$, where $K = 100$ is an arbitrary constant value. The next method that we believed would give us a better result is predicting the engine failures from the incomplete data by using the relationship of total metal concentration values and failure time of complete life data. Using the 7 datasets of failure data that we have, we need to have their failure times and condition monitoring values. We computed seven datasets that established a linear regression from this relationship, as shown in **Figure 1** below.

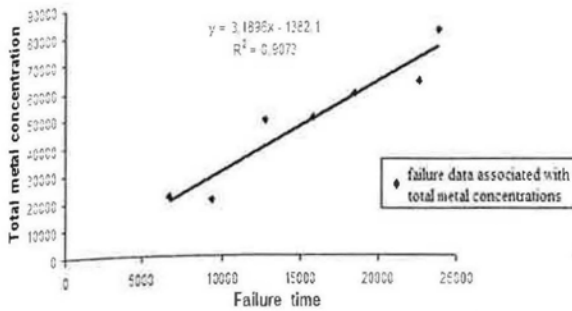


Fig. 1: Regression of total metal concentration (condition monitoring indicator) and failure time

Using this regression technique, we calculated the 95% prediction level for total metal concentration values, which we expected would yield the predicted value for the time to failure. Next, we mapped the total metal concentration from the incomplete replacement dataset into this 95% prediction level, shown in Figure 2 below.

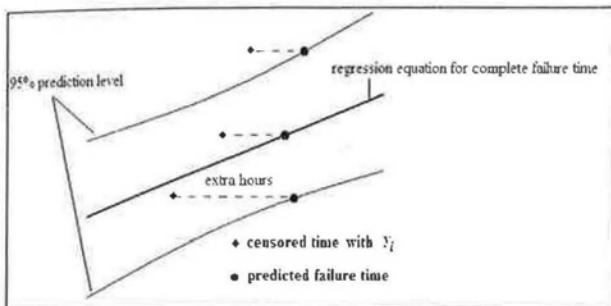


Fig. 2: 95% prediction interval from regression of failure data

As we assume that a shorter residual life will produce higher monitoring indicator, (metal concentration) y_i , we add some extra hours to those censored lives based on the y_i at their last check, which must be in the 95% prediction interval. To do this, we draw a straight line from the censored time until it touches either the regression line or one of 95% prediction intervals. A lower y_i will produce a longer residual life. Using this approach provided more failure data, allowing the estimates of the parameters within $p(x_0)$ to be performed using the likelihood function written above.

IV. NUMERICAL RESULT

The estimation result for both approaches is shown in TABLE 1.

TABLE 1: Estimated parameters for failure and interval-censored information

Method	$\hat{\alpha}$	$\hat{\beta}$	$E(t)$
Interval censored	0.0009478	6.256	23,542 hours
Predicted failure time	0.0009762	5.333	22,657 hours

The result shows that the expected failure time for each method is not far from the recommended mean value, which provides evidence of the validity of this approach. In addition, the variance for $\hat{\alpha}$ and $\hat{\beta}$ is relatively small compared with the values of estimated parameters, which shows the accuracy of the estimation. It can be observed that both approaches produce similar result. This scenario takes place because both approaches are seeking a method for adding some extra hours to predict when the failure will occur. It is not our aim to choose which of this approach is better than others, but to enrich the techniques if similar problem is occurred.

V. CONCLUSION

This paper has presented techniques which could be used in determining the parameter of the distribution where the failure data is limited. Two approaches have been discussed namely interval censored and predicted failure time. With available dataset, each approach is tested and their numerical result is shown.

VI. REFERENCES

1. Al-Sultan, K. S. and S. O. Duffua (1995). "Maintenance Control Via Mathematical Programming" *Journal of Quality in Maintenance Engineering* 1(3):36-46.
2. Aldrich, J. (1997). "R.A Fisher and the Making of Maximum Likelihood 1912-1922". *Statistical Science* 12(3):162-176.
3. Wang, W. (2002). "A Model to Predict the Residual Life of Rolling Element Bearings Given Monitored Condition Information to Date." *IMA Journal of Management Mathematics* 13:3-16
4. Wang, W. and A. H. Christer (2000). "Towards a General Condition-based Maintenance Model for a Stochastic Dynamic System." *Journal of the Operational Research Society* 51(2):145-155.

5. Hussin, B (2007). Development of a State Prediction Model to Aid Decision Making in Condition Based Maintenance. PhD Thesis, University of Salford.
6. Reliability Hotwire, The E-magazine for the Reliability Professional, Issue 9, November 2001.