

## A Data Processing Method for CBM using PHM

BAIRONG WU, and ZHIGANG TIAN\*

*Concordia Institute for Information Systems Engineering, Concordia University  
1515 Ste-Catherine Street West EV-7.637, Montreal, H3G 2W1, Canada.*

*(Received on April 13, 2011, revised on June 16, 2011 and on April 02, 2012)*

**Abstract:** In condition based maintenance (CBM) using proportional hazards model (PHM), fitting PHM is a very important step because it has a great influence on the effectiveness of the optimal maintenance policy. Previously actual condition monitoring measurements are directly used to fit the PHM. However this may introduce external noise and the optimal maintenance policy obtained based on this model may not be really optimal. To resolve this problem, a data processing method, which is fitting the actual measurements using the Generalized Weibull-FR function, is proposed to remove the external noise and fit the data before using it as input to the PHM. Two case studies using real-world vibration monitoring data are used to demonstrate the proposed approach. The proposed approach is validated to be effective and will save the total average maintenance cost by increasing the average replacement interval and making better use of remaining useful life.

**Keywords:** *Condition based maintenance, proportional hazards model, generalized Weibull-FR function*

### 1. Introduction

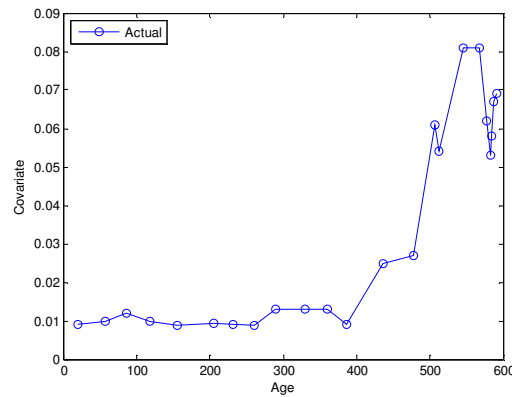
Condition based maintenance (CBM) is an advanced maintenance strategy which decides maintenance actions using the information collected via condition monitoring. The health conditions of a component or equipment can be monitored and predicted, and optimal maintenance actions can be scheduled to prevent component or equipment breakdown and minimize total maintenance costs [1]. Proportional hazards model (PHM) has been widely applied in many industries, such as mining industry, automobile industry, power generation industry, semiconductor industry, papermaking industry, petroleum industry and many other industries. These applications can be classified into two main categories: maintenance optimization and reliability analysis [2]. A key reason that PHM is more effective than previous approaches is that it considers not only age data but also condition monitoring data which influence the health of the component or equipment.

The CBM optimization methods using PHM have been developed and the main objective is to determine an optimal replacement policy to minimize long-run replacement cost [3, 4]. In these methods, the maintenance cost is calculated based on PHM and a risk threshold control limit policy. PHM based CBM methods can significantly decrease maintenance cost by reducing the number of unnecessary scheduled preventive maintenance operations [1]. In CBM using PHM, fitting the PHM is a vital step and the effectiveness of the optimal maintenance policy greatly depends on the accuracy of parameter estimation.

---

\* Corresponding author's. email: [tian@ciise.concordia.ca](mailto:tian@ciise.concordia.ca) (Z. Tian).

Usually actual condition monitoring measurement values are directly used as the inputs to the PHM model. However the actual measurements are often affected by external noise when they are collected as inspection points in the field. The example in Figure 1 shows an actual measurement series of a bearing failure history, which was collected from a pump in a Kraft mill. This history contains 23 inspection points and the bearing failed at the age of 591 days. In this figure, we can see the actual measurement values do not show a monotonic increasing trend. There are still a lot of fluctuations at various inspection points although its general trend is increasing. But as we all know, the deterioration of the health condition of a component or equipment is generally a monotonic process. Therefore, directly using the actual measurement values without any processing as input into the PHM model may introduce external noise into the model; thus the model built based on the actual measurement values may not represent the health condition of the component or equipment very accurately and the optimal maintenance policy obtained based on the PHM model may not be really optimal [5]. To resolve this problem, we propose an approach to remove the external noise and fit the data before feeding it into the PHM model. Compared to the actual measurement, the fitted measurements can better represent the deterioration of the component or equipment.



**Figure 1:** An Actual Inspection Measurements for a Sample Failure History (P1H\_Par5)

The remainder of this paper is organized as follows. Section 2 provides a synopsis of proportional hazards model. The proposed approach of fitting function is presented in Section 3, and two case studies are used in Section 4 to demonstrate the proposed approach. The conclusions are given in Section 5.

## 2. Synopsis of Proportional Hazards Model

### 2.1. PHM Basic Model

In CBM optimization process using PHM, the Weibull distribution function PHM is used to model the data. The PHM function combines the baseline hazard function and the covariates which affects the failure time. The age of the equipment is the main variable

while the condition monitoring measurements can be considered as a series of covariates. The basic model of PHM is described as follows [1]:

$$h(t, Z(t)) = \beta / \eta (t / \eta)^{\beta-1} \exp\left(\sum_{i=1}^m \gamma_i z_i(t)\right) \quad (1)$$

In this model,  $h(t, Z(t))$  is the conditional probability of failure at time  $t$ , given the values of  $z_1(t), z_2(t), \dots, z_m(t)$ . The first part of this model is a baseline hazard function  $\beta / \eta (t / \eta)^{\beta-1}$ , which takes into account the age of the equipment at time of inspection, given the values of parameters  $\beta$  and  $\eta$ . The second part  $\exp(\gamma_1 z_1(t) + \gamma_2 z_2(t) + \dots + \gamma_m z_m(t))$  takes into account the covariates which may be considered as the key factors influencing the health of equipment,  $z_1(t), z_2(t), \dots, z_m(t)$ , and their associated weights,  $\gamma_1, \gamma_2, \dots, \gamma_m$ .

## 2.2. Optimal Maintenance Policy

The method for calculating the cost and reliability values in the CBM optimization using PHM was described in [4]. A summary of this method is described as follows.

Let  $h(t, z(t))$  be the hazard rate at time  $t$  and  $K$  be the penalty cost. The basic theory of this approach can be described in the following way: if the observed risk  $K \times h(t, z(t))$  at the given inspection point of time is greater than a certain threshold value  $d$ , preventive replacement action should be taken; otherwise operation can continue. Failure replacement will be performed when a failure occurs between two inspection points of time. In this approach, the expected long run average cost per unit time is a function of the threshold risk level  $d$ , which is shown as follows:

$$\Phi(d) = \frac{C(1 - Q(d)) + (C + K)Q(d)}{W(d)} = \frac{C + KQ(d)}{W(d)} \quad (2)$$

where,  $\Phi(d)$  is the expected average cost per unit time and it is a function of the threshold risk level  $d$ ,  $C$  is the preventive replacement cost and  $C + K$  is the failure replacement cost.  $Q(d)$  is the probability that failure replacement will occur.  $W(d)$  denotes the expected time until replacement, regardless of whether it is a preventive action or a failure replacement. The objective of the CBM optimization using PHM is to find the optimal threshold value of the hazard rate to minimize maintenance cost. Once the optimal risk level,  $d^*$ , is determined, the component is replaced at the first moment  $t$  when

$$\beta / \eta (t / \eta)^{\beta-1} \exp\left(\sum_{i=1}^m \gamma_i z_i(t)\right) \geq d^* / K \quad (3)$$

## 3. Function for Fitting the Actual Inspection Measurement

To better represent the deterioration process of the equipment or component, an appropriate function is proposed to fit the actual measurements before they are used as input to the PHM. The fitting function is extracted from the Weibull distribution failure rate function. In reliability analysis, the health condition of equipment or component at a specified time is usually indicated by its failure rate at the given time. Weibull distribution is widely used in representing various practical lifetime distributions, and it is very

flexible to represent distributions with different scales and shapes [6]. Therefore, the following function generalized from the Weibull distribution failure rate function is used to fit the inspection measurements [5]:

$$\hat{Z}(t) = Y + K \frac{\beta}{\alpha^\beta} t^{\beta-1} \quad (4)$$

where  $t$  is the age of the unit,  $\hat{Z}(t)$  is the fitted measurement value,  $\alpha$  and  $\beta$  are the scale parameter and the shape parameter.  $(\beta/\alpha^\beta)t^{\beta-1}$  denotes the failure rate function for the 2-parameter Weibull distribution.  $K$  is introduced to scale the fitted measurement values to any ranges while  $Y$  indicates the covariate value when the age is 0.

The function in Eq. 4 is named as the “Generalized Weibull-FR function” [5]. There are four parameters in the Generalized Weibull-FR function to be determined:  $\alpha$ ,  $\beta$ ,  $K$  and  $Y$ . Genetic algorithm (GA) is used to find the optimal values for the four parameters because of the good global optimization performance of GA [7, 8]. After being tested using many actual inspection histories collected from the field, the Generalized Weibull-FR function is proved to have the capability of fitting all the tested measurement series very well. An example is given as follows to show how the Generalized Weibull-FR function works:

$$\hat{Z}(t) = 0.008 + 9.65 \frac{4.43}{610.2^{4.43}} t^{4.43-1} \quad (5)$$

The fitted results for the actual inspection measurements in Figure 1 are plotted in Figure 2, represented by “\*”. By removing the external noise from the actual inspection measurements, we can observe that the fitted measurements give a better indication of the degradation of the component.

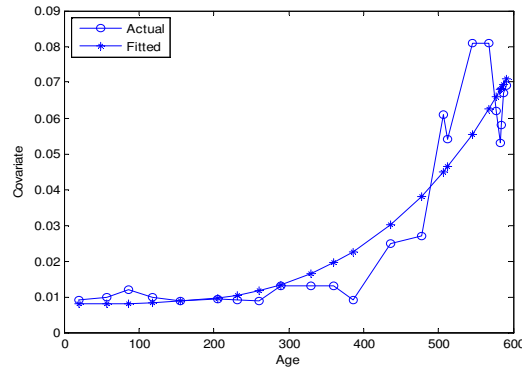


Figure 2: An Actual Inspection Measurements and the Fitted Measurement Series (PIH\_Par5)

#### 4. Case Studies

To validate the proposed approach, two case studies are conducted using real-world vibration monitoring data, which was collected from bearings on a group of Gould pumps at a Canadian Kraft Mill company [9] and from shear pump bearings in a food processing plant [4], respectively.

#### 4.1. Gould Pump Bearings Case

In this case study, two categories of data were obtained: event data and inspection data. There were three types of event data: beginning event, failure event and suspension event. For inspection data, 56(8\*5+8\*1+8\*1) vibration measurements were recorded. For each of the 8 pumps, seven measurements were analyzed at 5 different vibration frequency bands (8\*5), and the overall vibration reading (8\*1) plus the bearing's acceleration data (8\*1). In this case, 33 histories are used to demonstrate the proposed approach: 12 of them are failure replacements (ended with failure) and the other 21 histories are preventive replacements (ended with suspension). The actual inspection measurements and the fitted measurements obtained by the proposed approach will be used as input to the PHM respectively and their average total maintenance cost will be compared.

There are five steps to perform CBM based on PHM: significance analysis, parameter estimation, transition probability matrix development, cost data estimation and CBM optimization.

##### Step1 – Significance Analysis

Using the software EXAKT, we can perform the significance analysis for the 56 vibration measurements. Two covariates were identified to have significant influence on the health of bearings: P1H\_Par5 (band 5 vibration frequency in Pump location P1H), and P1V\_Par5 (band 5 vibration frequency in Pump location P1V).

##### Step 2- Parameter Estimation

In this case, there are four parameters to be estimated:  $\eta$  (scale parameter),  $\beta$  (shape parameter),  $\gamma_1$  (covariate weight for P1H\_Par5),  $\gamma_2$  (covariate weight for P1V\_Par5). Using the actual inspection measurements as input to the PHM, the four parameters are estimated as follows:

$$\begin{aligned} h(t, Z(t)) &= \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} e^{(\gamma_1 z_{P1H}(t) + \gamma_2 z_{P1V}(t))} \\ &= \frac{3.394}{2757} \left( \frac{t}{2757} \right)^{3.394-1} e^{(21.05 z_{P1H}(t) + 57.16 z_{P1V}(t))} \end{aligned} \quad (6)$$

Next, the actual inspection measurements are fitted using the proposed fitting function. The parameters estimated based on the fitted measurement value are given as follows:

$$h(t, Z(t)) = \frac{3.936}{2786} \left( \frac{t}{2786} \right)^{3.394-1} e^{(17.41 z_{P1H}(t) + 72.01 z_{P1V}(t))} \quad (7)$$

##### Step 3 – Transition Probability Matrix Development

To calculate the maintenance cost we need to specify the transition probability matrix. The transition probability matrix indicates the probabilities of a covariate in different ranges at the next inspection time given its current range. EXAKT can be used to estimate the transition probability matrices for the two covariates. Assuming the inspection interval is 28 days, the transition probability matrices obtained based on the actual inspection measurements and fitted measurements are given respectively as follows:

The transition probability matrices for actual inspection measurements are given in Table 1 and 2. The transition probability matrices for fitted measurements are shown in

Table 3 and 4, and they are intended to be used to represent the transition and degradation of the health of the component.

**Table 1:** Transition Probability Matrix for Covariate P1H\_Par5 based on Actual Measurements

P1H_Par5	0 to 0.00792	0.00792 to 0.014256	0.014256 to 0.05016	0.05016 to 0.136752	Above 0.136752
0 to 0.00792	0.784755	0.199967	0.0150435	0.000228608	5.48137e-006
0.00792 to 0.014256	0.0460993	0.8281	0.122933	0.00277821	8.95293e-005
0.014256 to 0.05016	0.00344127	0.121984	0.835499	0.0372439	0.00183203
0.05016 to 0.136752	0.000157207	0.00828724	0.111961	0.797492	0.0821025
Above 0.136752	0	0	0	0	1

**Table 2:** Transition Probability Matrix for Covariate P1V\_Par5 based on Actual Measurements

P1V_Par5	0 to 0.00741	0.00741 to 0.01404	0.01404 to 0.048165	0.048165 to 0.13299	Above 0.13299
0 to 0.00741	0.778123	0.209235	0.0124585	0.000180682	2.19389e-006
0.00741 to 0.01404	0.0453969	0.852736	0.0996783	0.00215347	3.49498e-005
0.01404 to 0.048165	0.00272389	0.100446	0.858878	0.037041	0.000911191
0.048165 to 0.13299	8.61624e-005	0.00473316	0.0807912	0.870673	0.0437164
Above 0.13299	0	0	0	0	1

**Table 3:** Transition Probability Matrix for Covariate P1H\_Par5 based on Fitted Measurements

P1H_Par5	0 to 0.0084952	0.0084952 to 0.0153832	0.0153832 to 0.047068	0.047068 to 0.135234	Above 0.135234
0 to 0.0084952	0.865102	0.132331	0.00253413	3.19206e-005	6.44605e-007
0.0084952 to 0.0153832	0	0.963069	0.0362363	0.000676333	1.81731e-005
0.0153832 to 0.047068	0	0	0.962856	0.0356956	0.00144835
0.047068 to 0.135234	0	0	0	0.923511	0.0764892
Above 0.135234	0	0	0	0	1

**Table 4** Transition probability matrix for covariate P1V\_Par5 based on fitted measurements

P1V_Par5	0 to 0.0081648	0.0081648 to 0.0151632	0.0151632 to 0.0452952	0.0452952 to 0.132386	Above 0.132386
0 to 0.0081648	0.879413	0.118227	0.00232983	3.00238e-005	3.37876e-007
0.0081648 to 0.0151632	0	0.961917	0.037356	0.000716413	1.07245e-005
0.0151632 to 0.0452952	0	0	0.962299	0.0368705	0.000830971
0.0452952 to 0.132386	0	0	0	0.956519	0.0434811
Above 0.132386	0	0	0	0	1

#### Step 4 – Maintenance Cost Determination

Based on the expertise and previous experiences, the preventive replacement cost  $C$  is estimated to be \$4,000, and the failure replacement cost  $C+K$  is \$12,000 for this case. Thus the penalty cost  $K$  equals to \$8,000.

#### Step 5 –Maintenance Policy Optimization

Now, the CBM optimal policy can be determined using the estimated parameters, transition probability matrices and cost data information. Using the parameters estimated based on the actual inspection measurements, which are:

$$\eta = 2757, \beta = 3.394, \gamma_1 = 21.05, \gamma_2 = 57.16 \quad (8)$$

The optimal maintenance policy is obtained as:

$$d^* = 7.23204\$ / day, C^* = 5.74075\$ / day \quad (9)$$

In this policy, the optimal risk threshold level  $d^*$  is obtained as 7.23204\$/day, which means it is time to perform preventive replacement when the observed risk  $K \times h(t, z(t))$  is greater than 7.23204\$/day. With this optimal policy, the total average maintenance cost is around 5.74075\$/day and the average preventive replacement interval is 867.234 days.

Now we calculate the optimal policy based on the parameters obtained using fitted measurements as input, which are

$$\eta = 2786, \beta = 3.936, \gamma_1 = 17.41, \gamma_2 = 72.01 \quad (10)$$

The optimal maintenance policy is determined as:

$$d^* = 5.06254\$ / day, C^* = 5.06254\$ / day \quad (11)$$

The optimal risk threshold level  $d^*$  is calculated as 5.06254\$/day and the total average maintenance cost is shown as around 5.06254\$/day. Based on the optimal policy, the average preventive replacement interval will be 949.648 days.

By comparing the optimized maintenance results before and after fitting the inspection measurements using the Generalized Weibull-FR function, we can see the total average maintenance cost based on the actual inspection measurements will be 5.74075\$/day while the average cost based on the fitted measurements will be 5.06254\$/day, as shown in Table 5. So using the proposed approach to fit the inspection measurements before applying to PHM will save the average maintenance cost around 11.81%. The average replacement interval is increased from 867.234 days to 949.648 days, which is around 9.5%. So we can conclude that by fitting the actual measurements before using them as input to the PHM will save the average maintenance cost and prolong the average replacement interval to make better use of remaining useful life.

**Table 5:** CBM Optimization Results Comparison before and after Fitting the Measurements

Method \ Results	Average Maintenance Cost (\$/day)	Average Replacement Interval (day)
Before	5.74075	867.234
After	5.06254	949.648
Changes	11.81%	9.5%

#### 4.2. Shear Pump Bearings Case

The second case is shear pump bearings in a food processing company. Totally 21 (3+3\*5+3) vibration measurements were collected using accelerometers, including vibration data in axial, horizontal and vertical directions for the overall velocity (3), velocities in 5 bands (3\*5=15) and acceleration in three directions (3). There are 25

histories in the recorded data, including 13 failure replacements and 12 preventive replacements.

In this case, three covariates are identified to be significant: VEL#1A (band 1 velocity in the axial direction), VEL#1V (band 1 velocity in the vertical direction), and VEL#2A (band 2 velocity in the axial direction). By plotting the 3 covariates we found out only two covariates (VEL#1A and VEL#2A) showing increasing pattern. Since the proposed approach only works well for those covariates which show increasing trend but not so well for decreasing covariates, we will only use covariate VEL#1A and VEL#2A to further demonstrate the proposed approach. So 4 parameters need to be estimated:  $\eta$  (scale parameter),  $\beta$  (shape parameter),  $\gamma_1$  (covariate weight for VEL#1A),  $\gamma_2$  (covariate weight for VEL#2A).

The parameters estimated for actual inspection measurements and fitted measurements are given in Eq. (12) and Eq. (13) respectively:

$$\eta = 739.9, \beta = 4.695, \gamma_1 = 6.358, \gamma_2 = 24.27 \quad (12)$$

$$\eta = 765.7, \beta = 5.226, \gamma_1 = 6.688, \gamma_2 = 28.19 \quad (13)$$

Transition probability matrices are also obtained based on actual inspection measurements and fitted measurements respectively.

In this case, the preventive replacement cost ( $C$ ) is estimated to be \$1,800, and the failure replacement cost ( $C + K$ ) is \$16,200.

Based on the actual inspection measurements, the optimal maintenance policy is obtained to be:

$$d^* = 16.9556\$ / day, C^* = 12.3408\$ / day \quad (14)$$

while the optimal policy for the fitted measurements is:

$$d^* = 10.3618\$ / day, C^* = 10.3618\$ / day \quad (15)$$

The comparison result given in Table 6 shows that the total average maintenance cost based on the actual inspection measurements is around 12.3408\$/day while the cost based on the fitted measurements is about 10.3618\$/day. So it will bring a saving of 16.36% in total average maintenance cost by using the proposed approach to fit the inspection measurements before applying to PHM. At the same time, the average replacement interval is increased from 168.305 days to 196.441 days, which is around 16.72%. So the proposed approach is further demonstrated to be able to save the total average maintenance cost and make better use of the remaining useful life.

**Table 6:** CBM Optimization Results Comparison before and after Fitting the Measurements

Results Method	Average Maintenance Cost (\$/day)	Average Replacement Interval (day)
Before	12.3408	168.305
After	10.3618	196.441
Changes	16.36%	16.72%



## 5. Conclusions

In CBM using PHM, generally actual condition monitoring measurement values are directly used to estimate the parameters for the PHM model. Nevertheless, the existing of external noise will change the monotonic increasing trend of inspection measurements and brings fluctuations in the deterioration process. Therefore the model built based on the actual measurement values may not accurately represent the health condition of the equipment or component and the optimal maintenance policy obtained based on the PHM model may not be really optimal. In this paper, an approach, which is fitting the actual measurements using the Generalized Weibull-FR function, is proposed to remove the external noise and fit the inspection measurements before feeding them into the PHM model.

Two case studies using real-world vibration monitoring data are used to demonstrate the proposed approach. Our studies show that the proposed approach will save the average maintenance cost and increase the average replacement interval to make better use of remaining useful life.

**Acknowledgments:** This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Le Fonds québécois de la recherche sur la nature et les technologies (FQRNT). We appreciate very much the help from OMDEC Inc. for providing the condition monitoring data used in the case study. Thanks are due to anonymous referees whose comments helped to improve the presentation

## References

- [1] Jardine, A., D. Lin and D. Banjevic. *A Review on Machinery Diagnostic and Prognostics Implementing Condition-based Maintenance*. Mechanical Systems and Signal Processing, 2006; 20(7): 1483-1510.
- [2] Tian, Z., D. Lin and B. Wu. *Condition based Maintenance Optimization Considering Multiple Objectives*. Journal of Intelligent Manufacturing, 2012; 23(2): 333-340.
- [3] Makis, V., and A. Jardine. *Optimal Replacement in the Proportional Hazards Model*. INFOR, 1992; 30(2): 172-183.
- [4] Banjevic, D., A. Jardine, V. Makis and M. Ennis. *A Control-Limit Policy and Software for Condition-based Maintenance Optimization*. INFOR, 2001; 39(1): 32-50.
- [5] Tian, Z. *An Artificial Neural Network Method for Remaining Useful Life Prediction of Equipment subject to Condition Monitoring*. Journal of Intelligent Manufacturing, 2012; 23(2): 227-237.
- [6] Kuo, W., and M. J. Zuo. *Optimal Reliability Modeling: Principles and Applications*. Wiley, Hoboken, 2003.
- [7] Levitin, G. *Universal Generating Function in Reliability Analysis and Optimization*. Springer-Verlag, London, 2005.
- [8] Levitin, G., A. Lisnianski, H. Ben Haim and D. Elmakis. *Redundancy Optimization for Series-Parallel Multistate Systems*. IEEE Transactions on Reliability. 1998; 47(2): 165-172.
- [9] Steven, B. EXAKT Reduces Failures at Canadian Kraft Mill. [www.omdec.com](http://www.omdec.com). 2006.

**Bairong Wu** is Ph.D. student in Mechanical Engineering at Concordia University, Montreal, Quebec, Canada from 2009 September. She received her M.S. degree in 2009 in Quality System Engineering at Concordia University, Montreal, Quebec, Canada. Her research focus on condition based maintenance optimization.

**Zhigang Tian** is currently an Assistant Professor at Concordia Institute for Information Systems Engineering at Concordia University, Montreal, Canada. He received his Ph.D. degree in 2007 in Mechanical Engineering at the University of Alberta, Canada; and his M.S. degree in 2003, and B.S. degree in 2000 both in Mechanical Engineering at Dalian University of Technology, China. His research interests focus on reliability analysis and optimization, prognostics, condition monitoring, and maintenance optimization. He is a member of IIE and INFORMS.