

Modelling Method for Maintenance Decision-Making in Civil Aero Engines Based on Multiple State Parameters

Huiling Cao¹, Zemin Wu², and Jiahe Miao^{1,*}

¹School of Aeronautic Engineering, Civil Aviation University of China, Tianjin, 300300, P.R. China

²Uni-top Airlines Co., Ltd, Wuhan, 430302, P.R. China

Abstract. For airlines, a scientific and effective method for engine maintenance decision-making should be developed for the planning of aero engine maintenance and removal. The mathematical modelling method of maintenance decision-making for civil aero engines based on the currently widely used condition-based maintenance (CBM) strategy was mainly studied in this work, and the effects of multiple state parameters on the system operation were fully considered. Based on historical data for aero engine removal due to performance degradation, statistical regression modelling was used to establish a mathematical model of maintenance decision-making that can reflect a functional relationship between the engine state parameters and the time on wing. The model was based on the proportional hazards-proportional odds (PH-PO) model, combining two commonly used statistical regression models, the proportional hazards model (PHM) and the proportional odds model (POM), into a single new model form; as a result, the scope of application of the model was improved. Finally, the results of a case study of a specific example showed the high practical value of this method.

1 Introduction

With the increasing growth of airline fleet sizes, engine maintenance decision-making is becoming increasingly complex. To reduce maintenance costs, flight delays and cancellation rates, it is important, among other engine maintenance activities, to draw engine removal plans, for which it is necessary to establish a model of maintenance decision-making.

The condition-based maintenance (CBM) [1] presents a new maintenance strategy that became more commonly used with the wide application of condition monitoring technology. In this strategy, the operation time and the current operational state of the engine were used to determine whether maintenance should be conducted, and as a result, maintenance decision-making based on the state of the engine was realized.

For civil aviation engines, performance degradation is an important cause of engine removal. Initially, the widely applied CBM decision-making method was used to conduct a simple trend analysis of performance degradation for a single performance parameter, such as the exhaust gas temperature margin (EGTM), and then, a pre-set monitoring threshold was used to determine whether the engine needed to be repaired. While this method is simple, it provides a relatively rough estimate and has some difficulty in reflecting the comprehensive effect of multiple influential factors.

Therefore, based on the CBM strategy, a statistical regression modelling method of reliability theory was used in this article. Based on research and analyses of

historical maintenance data for an engine fleet, multiple state parameters were selected, and the time on wing of removal due to performance degradation was regarded as life data samples. Moreover, full consideration was simultaneously given to the effect of the necessary engine state information on the operation of the system. As a result, a comparatively accurate relationship was established among the system state parameters, the system reliability and the time on wing. Then, regarding the system reliability as a control objective, the maintenance decision-making model of the engine can be obtained and used to determine whether maintenance should be conducted based on the current state of the engine.

2 Mathematical form of the model

2.1. Typical statistical regression models

The PHM and the POM, as two main statistical regression models, are commonly applied in life management in the field of engineering maintenance.

Basic introductions of these two regression models are presented in Table 1.

However, when a severe deviation from the assumptions is present, it is difficult to obtain an accurate evaluation result. Later, based on the Aranda-Ordaz parameter family and its derivative models established by Huang, which extended the application of the model into the field of accelerated life testing, the proportional hazards-

* Corresponding author: miaojiahechn@gmail.com

proportional odds (PH-PO) [2] model was proposed. By using the transfer parameter, this model combined two kinds of commonly used statistical regression models, the PHM and the POM, into a single new model form, with these two kinds of models being two particular

cases of the general model. As a result, the application scope and accuracy of assessment were improved [2]. Therefore, in this work, a model of maintenance decision-making was developed based on the PH-PO model.

Table 1. Basic introduction to PHM and POM.

	Proportional Hazards Model (PHM)	Proportional Odds Model (POM)
Reliability of the model definition function	$h(t; Z(t)) = \frac{f(t; Z(t))}{R(t; Z(t))}$	$\theta(t; Z(t)) = \frac{F(t; Z(t))}{R(t; Z(t))}$
Model equations	$h(t; Z(t)) = h_0(t) \exp(\gamma \cdot Z(t))$	$\theta(t; Z(t)) = \theta_0(t) \exp(\gamma \cdot Z(t))$
Baseline function equation	$h_0(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$ (Weibull distribution)	$\theta_0(t) = (\lambda t)^p$ (log-logistic distribution)
Baseline equation and logarithmic form	$\ln h_0(t) = (\beta - 1) \ln t + \ln\left(\frac{\beta}{\eta^\beta}\right)$	$\ln \theta_0(t) = p \ln t + p \ln \lambda$
Basic assumptions of the model	$\frac{h(t; Z_1)}{h(t; Z_2)} = \frac{h_0(t) \exp(\gamma \cdot Z_1)}{h_0(t) \exp(\gamma \cdot Z_2)} = \exp[\gamma \cdot (Z_1 - Z_2)]$	$\frac{\theta(t; Z_1)}{\theta(t; Z_2)} = \frac{\theta_0(t) \exp(\gamma \cdot Z_1)}{\theta_0(t) \exp(\gamma \cdot Z_2)} = \exp[\gamma \cdot (Z_1 - Z_2)]$

2.2 PH-PO model

$g(t; Z(t))$ of the model is defined as follows [2]:

$$g(t; Z(t)) = \frac{1}{R^c(t; Z(t))} - 1 \quad (1)$$

where c is the transformation parameter.

When $c \rightarrow 0$, this function reduces to the PH model as $g_{c \rightarrow 0}(t; Z(t)) = \ln R(t; Z(t))$; and when $c = 1$, the model

is reduced to the PO model as $g_{c=1}(t; Z(t)) = \frac{F(t; Z(t))}{R(t; Z(t))}$.

Hence, the PH-PO model can be written as

$$g(t; Z(t)) = g_0(t) \exp(\gamma \cdot Z(t)) \quad (2)$$

where γ is an unknown parameter vector, $Z(t)$ is the covariate vector, and $g_0(t)$ is the baseline function of $g(t; Z(t))$ with an arbitrary form.

Inspection of Table 1 shows that the logarithmic forms of the baseline functions of the PHM and POM are similar. Therefore, we set the baseline function of the PH-PO model as follows :

$$\ln g_0(t) = \beta \ln t + \alpha \quad (3)$$

Rewriting Eq. (3), we obtain

$$g_0(t) = \exp(\alpha) t^\beta \quad (4)$$

where α and β are unknown parameters.

3 Establishment of the maintenance decision-making model

3.1 Modelling considerations

After confirmation of the model's mathematical form, the following issues must be considered in the establishment of the maintenance decision-making model.

(1) There are many reasons for engine removal, with the most common being performance degradation, lease period end, life limited parts (LLP), AD/SB requirements, foreign object damage (FOD) [3] and engine echeloned removal of fleet management. In this work, the establishment of the model was focused on engines removed due to performance degradation, so only removal caused by performance degradation was regarded as a real failure. However, considering that the complete data collected from current decision-making routes are still relatively deficient, in order to increase the sample size, multiclass data regarding removal can be considered for use in parameter estimation for the regression model. In cases of removal due to AD/SB, echeloned removal or LLP, the situation usually has not yet reached the level of performance degradation, so these cases can be regarded as censored data in the analysis of life data. Data for removal due to the end of the lease period are not considered in this article, because the time on wing of the engine is short in this case. FOD data are primarily caused by accidental factors and lack of regularity, so they are also ignored. Additionally, the model assumes that the faults occurring in the operation of the system can be remedied by maintenance, and then the system functionality after repair is as good as that of a new engine.

(2) For diverse types of engines, the structure and performance parameters are distinct; hence, the trends in monitoring parameters and the alert threshold are not the same. Therefore, in the model, statistical regression analysis must take the same type of engines as samples.

3.2 Covariate processing

There are many covariates that can be incorporated in the PH-PO model, but not all covariates have a significant correlation with life on wing; therefore, correlation analysis must first be performed between these covariates and life on wing, and then the most significant covariates are selected in the model. Additionally, the covariates are also required to be uncorrelated with each other in the model; in other words, the $Z(t)$ in the model equation must be independent. This requirement enables the use of principal component analysis (PCA) [4, 5]. This method can not only effectively eliminate correlations among covariates but can also achieve dimension reduction and can simplify the modelling.

For dimensionality reduction, the number of principal components can be obtained according to the cumulative contribution rate. Most often, when the cumulative contribution rate of the selected principal components exceeds 80%, the principal components include most of the information for the original variables.

3.3 Maximum likelihood estimation of parameters

The general form of the likelihood function is defined as

$$L(\alpha; \beta; \gamma) = \prod_{i=1}^n f(t_i; Z(t_i))^{\delta_i} S(t_i; Z(t_i))^{1-\delta_i} \quad (8)$$

where n is the overall number of samples. δ_i is the indicator of the censored data, where $\delta_i = 0$ indicates that the datum is a censored datum, and $\delta_i = 1$ indicates that the datum is regarded as a real failure. $f(t; Z(t))$ is the density function of the failure probability, and $S(t; Z(t))$ is the survival function and is equivalent to the reliability function $R(t; Z(t))$.

From the formulation of the PH-PO model, the reliability function and probability density function are respectively obtained as

$$R(t; Z(t)) = \exp\left\{\frac{1}{-c} \ln[c \exp(\alpha) t^\beta \exp(\gamma \cdot Z(t)) + 1]\right\} \quad (9)$$

$$f(t; Z(t)) = -\frac{dR}{dt} = \exp(\alpha) \beta t^{\beta-1} \exp(\gamma \cdot Z(t)) \cdot [c \exp(\alpha) t^\beta \exp(\gamma \cdot Z(t)) + 1]^{-\frac{1}{c}-1} \quad (10)$$

Substituting (9) and (10) into (8), the log-likelihood function can be written as

$$L(\alpha; \beta; \gamma) = \prod_{i=1}^q \exp(\alpha) \beta t_i^{\beta-1} \exp(\gamma \cdot Z(t_i)) \cdot [c \exp(\alpha) t_i^\beta \exp(\gamma \cdot Z(t_i)) + 1]^{-1} \quad (11)$$

$$\prod_{j=1}^n [c \exp(\alpha) t_j^\beta \exp(\gamma \cdot Z(t_j)) + 1]^{-\frac{1}{c}}$$

where q is the number of real failures.

From Eq. (11), the logarithmic form of the log-likelihood function can be obtained as

$$\ln L(\alpha; \beta; \gamma) = q\alpha + q \ln(\beta) + (\beta - 1) \sum_{i=1}^q \ln(t_i) + \sum_{i=1}^q \sum_{k=1}^m \hat{\gamma}_k Z_k(t_i) - \sum_{i=1}^q \ln\{c \exp(\alpha) t_i^\beta \exp[\sum_{k=1}^m \hat{\gamma}_k Z_k(t_i)] + 1\} - \frac{1}{c} \sum_{j=1}^n \ln\{c \exp(\alpha) t_j^\beta \exp[\sum_{k=1}^m \hat{\gamma}_k Z_k(t_j)] + 1\} \quad (12)$$

After the likelihood function is obtained, the unknown parameters can be determined by using a commonly used nonlinear optimization genetic algorithm (GA) approach [6, 7]. A GA is a probabilistic global optimization algorithm and shows a good global search ability. Hence, this algorithm can quickly search all solutions in the solution space and does not fall into the rapid decline locally optimal solution trap. Furthermore, this method takes full advantage of the GA's inherent parallelism, increasing the solution speed and obtaining the optimum solution. In addition, to increase the convergence efficiency and improve the calculation accuracy, the method in this paper is combined with the initial value estimates of the function, and then the search scope of the algorithm is restricted within a reasonable range close to the initial value. Finally, a favourable result was obtained in a case study, as described below.

3.4 Maintenance decision-making model

Based on the formulation of the PH-PO model, we have

$$R^{-c}(t; Z(t)) - 1 = c \exp(\alpha) t^\beta \exp(\gamma \cdot Z(t)) \quad (13)$$

Substituting the estimated parameters and rewriting Eq. (13), we obtain

$$\sum_{k=1}^m \hat{\gamma}_k Z_k = \ln[R^{-c}(t; Z(t)) - 1] - \ln(\hat{c}) - \hat{\alpha} - \hat{\beta} \ln(t) \quad (14)$$

Eq. (14) provides the function relating the system state parameters and the time on wing. Substituting a certain reliability value R^* , a logarithmic curve can be obtained with respect to time t as the X-axis and the comprehensive value of covariates $\sum_{k=1}^p \hat{\gamma}_k Z_k$ as the Y-

axis. This step provides the system maintenance threshold under the reliability mentioned above. This threshold is different from the general linear threshold, providing a synthetical statement about the reliability and state of the system. We then multiply the current values of the covariates by the coefficient of covariates and plot the multiplication results in the figure using dots. A dot at the bottom of the curve indicates that the system can still operate on wing; however, if the dot is above the curve, the system's performance is seriously degraded and the engine must be repaired [8]. Obviously, this maintenance threshold corresponds to the control limit of maintenance decision-making with reliability as the control target.

4 Case study

Table 2. Partial samples from historical data for PW4077D aero engine removal.

<i>SN</i>	<i>TSO/h</i>	δ	$\Delta\Delta EGT/C^\circ$	<i>EGTM/C°</i>	$\Delta\Delta FF /\%$	$\Delta\Delta NI/\%$	<i>AVMH /ips</i>	Removal reason
222082	5959	0	23.8	49	3.2	0.7	0.1	Engine echeloned removal
222095	10879	1	56	27	7.1	0.3	0.2	Compressor blade damage
...
222083	8365	1	44.3	37	7.7	1.1	0	Turbine stator blade ablation

(1) Principal component analysis

Using MATLAB programming and calculation, we obtained

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} -0.9290 & 0.6656 & -0.9287 & -0.7728 & 0.0004 \\ 0.0507 & -0.1413 & -0.0467 & -0.1261 & 0.9870 \\ -0.0268 & -0.6987 & -0.1358 & -0.4066 & -0.1570 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix} \quad (15)$$

where x_1, \dots, x_5 respectively represent $\Delta\Delta EGT$ 、*EGTM*、 $\Delta\Delta FF$ 、 $\Delta\Delta NI$ 、and *AVMH*. *Z* is the extracted principal component matrix and has three compressed principal components that account for 55.32%, 20.29%, and 13.95%, respectively.

(2) Solution of unknown parameters

Using the single objective GA nonlinear optimization and MATLAB programming and calculation, the estimated values of the parameters are obtained as

$$\alpha = -24.9993; \beta = 3.0996; c = 0.2013; \gamma_1 = 0.1743; \gamma_2 = 3.0751; \gamma_3 = -0.3033$$

Substituting these estimated values in the formulation of the PH-PO model, we obtain

$$g(t; Z(t)) = \exp(-24.9993)t^{3.0996} \cdot \exp(0.1743Z(t)_1 + 3.0751Z(t)_2 - 0.3033Z(t)_3) \quad (16)$$

Choosing historical data for the PW4077D engine fleet regarding removal from an airline during 1998–2013, a total of 70 data sets were obtained, as shown in Table 2. The data increases some state parameters at the moment of removal on the basis of historical records of the engine's life on wing. After the correlation analysis between covariates and life on wing, the selected parameters are $\Delta\Delta EGT$ (deviation of exhaust gas temperature), *EGTM* (exhaust gas temperature margin), $\Delta\Delta FF$ (deviation of fuel consumption), $\Delta\Delta NI$ (deviation of low pressure rotor speed), and *AVMH* (high pressure rotor vibration amplitude). Here, *TSO* refers to time on wing since overhaul, δ is an indicator of censored data, *SN* is the engine number, and “ $\Delta\Delta$ ” represents the D-value between the removal moment deviation and the installation moment deviation.

$$R^{-0.2013}(t; Z(t)) - 1 = 0.2013 \exp(-24.9993)t^{3.0996} \cdot \exp(0.1743Z(t)_1 + 3.0751Z(t)_2 - 0.3033Z(t)_3) \quad (17)$$

(3) Maintenance decision-making model

Substituting the historical engine removal data from Table 2 into Eq. (17), a series of reliability values can be calculated. For group incidents, in order to improve the accuracy of the reliability assessment, sample points that deviate strongly from the statistical regression area of the model should be eliminated. Using a conservative estimate, we set the R-value as 0.5 and 0.99 according to the maximum, minimum and distribution of the R-value. Then, on the basis of Eq. (14), the upper and lower control limits of the maintenance decision-making model can be obtained as follows:

The upper limit is

$$y_{\max} = \sum_{k=1}^p \hat{\gamma}_k Z_k = 39.3080 - 3.5101 \ln(t) \quad (18)$$

The lower limit is

$$y_{\min} = \sum_{k=1}^p \hat{\gamma}_k Z_k = 26.4665 - 3.5101 \ln(t) \quad (19)$$

According to the MATLAB simulation shown in Fig. 1, the following analysis results are obtained:

Decision-making 1: We multiply the current values of the covariates by the coefficient of covariates and then plot the multiplication results in the figure using dots. If a dot is at the bottom of the curve, the system can still operate on wing; however, if the dot is above the curve, the engine's performance is seriously degraded, and the

engine must be replaced and corrective maintenance (CM) activities must be conducted.

Decision-making 2: When the dot is in the middle transition region, the engine should be subject to either intensive monitoring or preventive maintenance (PM) activities.

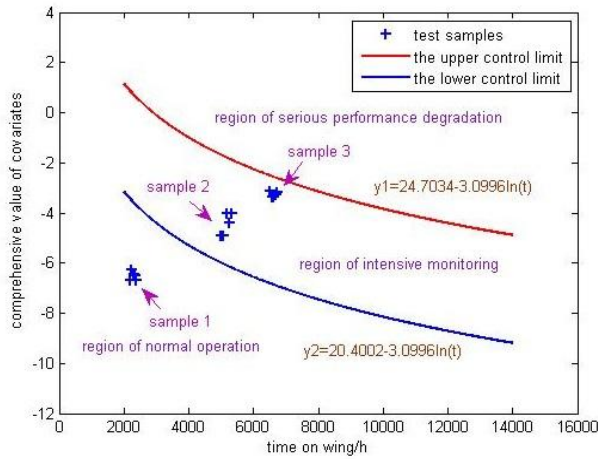


Fig. 1. Plot of maintenance decision-making model results.
 (4) Model demonstration

Using the SN222159 engine as an example, the actual life on wing is 6753 hours. Three groups of sample points are selected during the period from the time of engine installation to the time of engine removal, as shown in Table 3. Among these samples, Sample 1 is taken from the initial stage of the engine’s operation on wing, sample 2 is obtained from the mid-to-late stage, and sample 3 is extracted from the last stage prior to engine removal.

Table 3. Partial samples from historical data for SN22159 aero engine’s operation on wing.

	<i>TSO</i> /h	$\Delta\Delta EGT$ /C°	<i>EGTM</i> /C°	$\Delta\Delta F$ F/%	$\Delta\Delta NI$ /%	<i>AVM</i> <i>H</i> /ips
Sam ple 1	2192	17	61	2.4	0.5	0.2

	2377	22	57	3.3	0.2	0.1
Sam ple 2	4978	31	45	3.6	0.4	0.3

	5327	36	36	3.8	0.5	0.3
Sam ple 3	6517	46	27	4.6	0.9	0.4

	6726	51	25	5.2	1.2	0.4

We multiply the covariate values of each line from Table 3 by the coefficient of covariates and use MATLAB to plot the multiplication results in a figure using dots. The results are shown in Fig. 1. To a certain extent, the results can reflect the actual operational situation.

4 Conclusion

A mathematical modelling method for maintenance decision-making regarding civil aero engines based on multiple state parameters was studied in this paper. Based on historical data for engine removal due to performance degradation, the PH-PO model was used to establish a mathematical model of maintenance decision-making that can reflect the relationship between the engine state parameters and the time on wing. Then, regarding the system’s reliability as a control target, the maintenance decision-making model of the engine can be obtained and used to determine whether maintenance should be conducted based on the current state of the engine. Finally, in the case study of a specific example, the obtained results showed that the method has a high practical value and can provide a reliable basis for engine maintenance decision-making.

References

1. J. Moubray, *Reliability centred maintenance* (Butterworth-Heinemann, London, 1999)
2. T. Huang, E.A. Elsayed, T. Jiang, *The Proceedings of the 14th ISSAT International Conference on Reliability and Quality in Design*, 247-251 (2008)
3. G.R. Weckman, R.L. Shell, J.H. Marvel, *Computers & Industrial Engineering*, **40**, 51-63 (2001)
4. B. Schölkopf, A. Smola, K. Müller, *Neural Computation*, **10**, 1299-1319 (1998)
5. W.K. Härdle, L. Simar, *Applied Multivariate Statistical Analysis* (Springer-Verlag, Berlin, 2012)
6. Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs* (Springer-Verlag, Berlin, 1996)
7. D.E. Goldberg, *GENETIC algorithms in search, optimization, and machine learning* (Addison-Wesley Publishing Company, New York, 1989)
8. H.F. Zuo, H.J. Zhang, X. Rong, *Journal of Aerospace Power*, **21**, 716-721 (2006)