Prognostic and stochastic modeling of degradation

Khanh Le Son, Mitra Fouladirad, Anne Barros

ANR AMMSI

24/01/13
Introduction

Objectives

- Using a stochastic approach for prognostic in order to compare with the exciting non-stochastic methods applied on the 2008 Prognostic Health Management data.


- Using a stochastic process to model the deterioration of components (Remaining Useful Life estimation).
Experimental data

Two sub-data sets: the training data set and the testing data set.

The training data set is used to build the prediction model.

The testing data set is used to estimate the RUL for each testing unit.

<table>
<thead>
<tr>
<th>Unit</th>
<th>Cycle</th>
<th>OP1</th>
<th>OP2</th>
<th>OP3</th>
<th>SM1</th>
<th>SM2</th>
<th>...</th>
<th>SM21</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$T_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>218</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$T_{218}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Degradation indicator can not be directly deduced from the 21 sensor paths

All measurements are divided into 6 clusters corresponding to 6 operational modes

Selection of 7 sensors
Degradation indicator construction

Analyse of a failure times

- Select the measurements of 7 sensors only at the failure time
- Group the failure measurements according to their mode (6 groups)

Identification of a failure space and a failure place for each mode

- Create a projection space of dimension 2 with PCA (called failure space)
- Calculate the barycenter of the projected failure measurements in this space to create a failure place.
Principal Component Analysis

Results

- 6 plans of PCA $P_1, P_2, ..., P_6$, one for each mode.
- 6 failure places $L_1, L_2, ..., L_6$, one for each mode.

<table>
<thead>
<tr>
<th></th>
<th>Mode 1</th>
<th>Mode 2</th>
<th>Mode 3</th>
<th>Mode 4</th>
<th>Mode 5</th>
<th>Mode 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>60.85</td>
<td>72.64</td>
<td>61.45</td>
<td>54.41</td>
<td>58.95</td>
<td>79.65</td>
</tr>
<tr>
<td>PC2</td>
<td>38.04</td>
<td>26.75</td>
<td>37.85</td>
<td>44.55</td>
<td>40.07</td>
<td>19.11</td>
</tr>
<tr>
<td>PC3</td>
<td>0.66</td>
<td>0.28</td>
<td>0.34</td>
<td>0.56</td>
<td>0.41</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table: Contribution of principal components for each mode
Degradation indicator

- $N_k$ number of units in mode $k$
- $\bar{P}_k = (\bar{a}, \bar{b})$ the barycenter of the failure space $P_k$, $k = 1, \ldots, 6$
- $P^k_i = (a_i, b_i)$, $i = 1, \ldots, N_k$ is the $i^{th}$ failure place in the projection space $P_k$
- $P^k_{ij} = (a_{ij}, b_{ij})$ is the measure of the 7 selected sensors at time $j$ for component $i$ in the projection space $P_k$

The dispersion of the failure places in mode $k$ at time $j$ (noted $k(j)$) is defined by:

\[
Disper_{k(j)} = \sqrt{\frac{1}{N_{k(j)} - 1} \sum_{i=1}^{N_{k(j)}} ((a_i - \bar{a})^2 + (b_i - \bar{b})^2)}
\]

\[
D_{ij}^{k(j)} = \frac{\sqrt{(a_{ij}^{k(j)} - \bar{a})^2 + (b_{ij}^{k(j)} - \bar{b})^2}}{Disper_{k(j)}}
\]
Degradation indicator

Projections of failure on the 2-D space of PCs in mode $k$ ($P_k$)

$$D_{i,j}^k = \frac{\text{distance}(P_{i,j}^k; \bar{P}_k)}{\text{Disper}_k}$$

Barycenter of the failure projections $\bar{P}_k$
⇒ One component

Degradation indicator of component 1
Degradation indicator

⇒ All the components

- Non-linear and decreasing
- Significant dispersion in the beginning
- At the failure times, degradation tends to zero
Degradation model - Definition

- **Note:**
  - \( D_{i,j}^{k(j)} \) = degradation indicator of unit \( i \) at cycle \( j \).
  - \( Y(i) = (D_{i,1}^{k(1)}, ..., D_{i,n_i}^{k(n_i)}) \) : the observation vector for unit \( i \).
  - \( X(i) = (X_{i,1}^{k(1)}, ..., X_{i,n_i}^{k(n_i)}) \) : the non-observable actual random states of unit \( i \).

- **Our deterioration model:**

  \[
  D_{i,j}^{k(j)} = X_{i,j}^{k(j)} + \epsilon_{i,j}^{k(j)} \\
  Y(i) = X(i) + \epsilon(i)
  \]

  where:
  - \( \epsilon_{i,j}^{k(j)}, j = 1, ..., n_i \) : the independent gaussian random variables with standard deviation \( \sigma_{j}^{(i)} \) and mean equals to zero for unit \( i \).
  - Non-homogeneous Gamma process for \( X_{i,j}^{k(j)} \)
Degradation model

Definition of non-homogeneous Gamma process

- The initial state $X_0 = 0$.
- $(X_j)_{j \geq 0}$ is supposed to be monotone, increasing.
- The increments $X_j - X_{j-1}, j = 1, 2, ..., n$ are independent and have the Gamma density:

$$f_{\nu(t_j), \nu(t_{j-1}), \beta}(x) = \frac{\beta^{\nu(t_j)-\nu(t_{j-1})} e^{-\beta x}}{\Gamma(\nu(t_j) - \nu(t_{j-1}))} x^{\nu(t_j)-\nu(t_{j-1})-1} \mathbb{1}_{(0, \infty)}(x)$$

- $\Gamma(\nu) = \int_{z=0}^{\infty} z^{\nu-1} e^{-z} dz$ : Gamma function for $\nu > 0$.
- $\mathbb{1}_A(x) = 1$ for $x \in A$, $\mathbb{1}_A(x) = 0$ for $x \not\in A$.
- Shape function $\nu(t) = \alpha t^b$ and scale parameter $\beta$.

where: $\Gamma(\nu) = \int_{z=0}^{\infty} z^{\nu-1} e^{-z} dz$ is the gamma function,.

$\Rightarrow$ 4 parameters to be estimated: $\alpha, \beta, b, \sigma$. 

Mesure de dégradation

\[ \Gamma(v(t_i) - v(t_j), \beta) \]

Seuil de défaillance

\[ RUL(t_n) \]

Mesure de dégradation

\[ X_n \]

\[ X_i \]

\[ X_j \]
Remaining Useful Lifetime estimation

- Remaining Useful Lifetime (RUL) estimation is based on the failure probability at the next inspection given the \( n \) observations \( Y_1, \ldots, Y_n \).

- The distribution function of \( RUL(t_n) \) figured out the observations is defined as follows:

\[
F_{RUL}(t_n)(h) = P(X_{t_n+h} > L | X_n > L, Y_1, \ldots, Y_n)
= \int \int \bar{F}_{\alpha((t_n+h)^b-t_n^b),\beta}(l-x_n) \cdot f_L(l) \cdot \mu_{X_n/Y_1,\ldots,Y_n} dldx_n
\]

- \( \bar{F}_{\alpha((t_n+h)^b-t_n^b),\beta} \) : the reliability function of Gamma process with shape function \( \alpha((t_n + h)^b - t_n^b) \) and scale parameter \( \beta \).
- \( \mu_{X_n/Y_1,\ldots,Y_n} \) : the conditional density of \( X_n \).
- \( f_L(l) \) : the density function of the failure threshold.
Joint distribution of system state

For estimating the RUL, the joint conditional density of $X$ figured out the observation vector $Y$ is calculated as follows:

$$
\mu_{X/Y}(x_1, \ldots, x_n) = K_1 e^{-\beta x_n} \prod_{j=1}^{n} (x_j - x_{j-1})^{\alpha(t_j^b - t_{j-1}^b) - 1} e^{-\frac{g^2(x_j, Y_j)}{2\sigma^2}} |g'(x_j, Y_j)|
$$

where $g'(., y) = \frac{\partial g(., y)}{\partial y}$ and $K_1$ is the coefficient defined as follows:

$$
\frac{1}{K_1} = \int \cdots \int e^{-\beta x_n} \prod_{j=1}^{n} (x_j - x_{j-1})^{\alpha(t_j^b - t_{j-1}^b) - 1} e^{-\frac{g^2(x_j, Y_j)}{2\sigma^2}} |g'(x_j, Y_j)| \, dx_1 \ldots dx_n
$$

- It’s difficult to calculate the coefficient $K_1 \Rightarrow$ MCMC (Gibbs) algorithm.
Gibbs algorithm

- For $j = 1$,
  \[
  \mu_{X/Y}(x_1/x_2, \ldots, x_n) = K_{2,1} x_1^{(t_1^b)-1} (x_j - x_{j-1})^{(t_j^b-t_{j-1}^b)-1} e^{-\frac{g^2(x_1,Y_1)}{2\sigma^2}} |g'(x_1, Y_1)| 1(0 < x_1 < x_2)
  \]

- For $2 \leq j \leq n - 1$,
  \[
  \mu_{X/Y}(x_j/x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = K_{2,j} (x_j - x_{j-1})^{(t_j^b-t_{j-1}^b)-1} (x_{j+1} - x_j)^{a(t_{j+1}^b-t_j^b)-1} e^{-\frac{g^2(x_j,Y_j)}{2\sigma^2}} |g'(x_j, Y_j)| 1(x_{j-1} < x_j < x_{j+1})
  \]

- For $j = n$,
  \[
  \mu_{X/Y}(x_n/x_1, \ldots, x_{n-1}) = K_{2,n} e^{-\beta x_n} (x_n - x_{n-1})^{(t_n^b-t_{n-1}^b)-1} e^{-\frac{g^2(x_n,Y_n)}{2\sigma^2}} |g'(x_n, Y_n)| 1(x_{n-1} < x_n)
  \]

where $K_{2,j}$ are tractable constants dependent on $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ and $y_1, \ldots, y_n$
Parameters estimation

- Parameters of model are estimated based on the outputs of Gibbs algorithm and by using the Stochastic EM (SEM) method.
- The observations set \( Y^{(i)} = (Y_{i,j}^{k(1)}, \ldots, Y_{i,n_i}^{k(n_i)}) \), \( i = 1, \ldots, 218 \) with component independently observed at inspection times \( 0 < t_1 < \ldots < t_{n_i} \).
- Maximizing the likelihood:

\[
L(\alpha, b, \beta, \sigma) = \sum_{i=1}^{218} \sum_{j=1}^{n_i} \left[ (\alpha((t_j)^b-(t_{j-1})^b)-1) \ln(X_{i,j}^{k(j)}-X_{i,j-1}^{k(j-1)}) \right.
\]

\[
- \beta(X_{i,j}^{k(j)}-X_{i,j-1}^{k(j-1)}) - \frac{g^2(X_{i,j}^{k(j)}, Y_{i,j}^{k(j)})}{2\sigma^2} + \ln(|g'(X_{i,j}^{k(j)}, Y_{i,j}^{k(j)})|)
\]

\[
- \ln(\sigma\sqrt{2\pi}) + \alpha((t_j)^b-(t_{j-1})^b) \ln(\beta) - \ln(\Gamma(\alpha((t_j)^b-(t_{j-1})^b)))
\]
The conditional distribution \( F_{RUL}(t_n)(h) \):

\[
F_{RUL}(t_n)(h) = P(X_{t_n+h} > L | X_n > L, Y_1, ..., Y_n)
\]

\[
= \int \int \bar{F}_{\alpha((t_n+h)b_t^b),\beta(l-x_n)}(l) \cdot f_L(l) \cdot \mu_{X_n/Y_1, ..., Y_n} \, dl \, dx_n
\]

can be estimated by Gibbs algorithm as follows:

\[
\hat{F}_{RUL}(t_n)(h) = \frac{1}{Q} \sum_{q=Q_0+1}^{Q_0+Q} \int \bar{F}_{\hat{\alpha}((t_n+h)\hat{b}_t^b),\hat{\beta}((l-z_n(q))} \cdot f_L(l) \, dl
\]

\( Q_0 \) : the number of sequences to get the convergence state.

\( Q \) : the number of sequences to give sufficient precision to the empirical distribution of interest.
Performance assessment of the model

- Applying the stochastic degradation model to all 218 units of the testing data set, we obtained an estimated RULs set $RUL_{estimated}^{i'}$, for $i' = 1, ..., 218$.

- Penalty function criterion: provided by 2008 PHM Challenge, the penalty score for each testing unit is given by the following formula:

$$S_{i'} = \begin{cases} 
    e^{-d_{i'}/13} - 1, & d_{i'} \leq 0 \\
    e^{d_{i'}/10} - 1, & d_{i'} > 0 
\end{cases} \quad i' = 1, ..., 218$$

where $d_{i'} = RUL_{estimated}^{i'} - RUL_{actual}^{i'}$ and the total score $S = \sum_{i=1}^{218} S_{i'}$

- Root mean squared error: $RMSE = \sqrt{\sum_{i'=1}^{218} (d_{i'})^2}$
Performance assessment of the model

- Lifetime distribution model (Weibull distribution on the failure times): $S = 9870$

- Total score of the different models on our degradation indicator $D_{i,j}^{k(j)}$:
  - Similarity-based prognostic approach proposed by Wang in the 2008 PHM conference: $S = 6690$
  - Gamma + Noise model: $S = 4197$ and $RMSE = 420$

- The best results in the 2008 PHM Challenge:
  - Similarity-based prognostic model of Wang: $S = 5636$
The system is inspected periodically at inspection times \( T_1, T_2, \ldots \) where \( T_k = kT \) with \( k \in \mathbb{N} \) and \( T \in \mathbb{R} \) is the inspection interval.

Let be \( Q, (0 < Q < 1) \), a fixed percentile of the RUL distribution function, at each inspection time \( T_k \):

- If \( X_{kT} < L \) and \( P(RUL(T_k) < T) > Q \), the system is preventively replaced with a cost \( C_p \).
- If \( X_{kT} < L \) and \( P(RUL(kT) < T) < Q \), the decision is postponed until the next inspection.
- If \( X_{kT} \geq L \), a corrective replacement is carried out with a cost \( C_c \).
Durée d'indisponibilité
Remplacement préventif
Remplacement correctif
Cycle de renouvellement
Temps
Zone de défaillance
Point de renouvellement
Durée d'indisponibilité
Cycle de renouvellement
Cycle de renouvellement

Figure: RUL based maintenance policy
RUL based maintenance

\[ C^\infty = \lim_{t \to \infty} \frac{C(t)}{t}, \]

\( C(t) \) is the cumulated maintenance cost at time \( t \)

\[ C(t) = C_i N_i(t) + C_p N_p(t) + C_c N_c(t) + C_d d_d(t), \]

\( N_p(t) \) the number of preventive replacements before \( t \), \( N_c(t) \) the number of corrective replacements before \( t \), \( d_d(t) \) the cumulative unavailability duration of the system before \( t \) and \( N_i(t) \) the number of inspections before \( t \). Note that \( N_i(t) = \left\lfloor \frac{t}{T} \right\rfloor \) where \( \lfloor x \rfloor \) denotes the integer part of the real number \( x \).
RUL based maintenance

(T,Q) policy and maintenance cost

Figure: Iso-level curves and mean cost per time unit of (T, Q) policy
Thank you for your attention