UNIVERSITÉ DE MONTRÉAL

STUDIES IN CONDITION BASED MAINTENANCE USING PROPORTIONAL HAZARDS MODELS WITH IMPERFECT OBSERVATIONS

ALIREZA GHASEMI
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR (Ph.D.)
(GÉNIE INDUSTRIEL)
AVRIL 2009

Your file Votre référence
ISBN: 978-0-494-49415-8
Our file Notre référence
ISBN: 978-0-494-49415-8

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

STUDIES IN CONDITION BASED MAINTENANCE USING PROPORTIONAL HAZARDS MODELS WITH IMPERFECT OBSERVATIONS

présentée par : GHASEMI Alireza
en vue de l'obtention du diplôme de : Philosophiae Doctor (Ph.D)
a été dûment acceptée par le jury d'examen constitué de :

M. MALHAMÉ , Roland, ing. Ph.D., président;
Mme YACOUT Soumaya, D.Sc., directrice de recherche;
M. OUALI Mohamed-Salah, ing. Doct., codirecteur;
M. ADJENGUE Luc, Ph.D., membre;
M. KLIM Zdzislaw, ing., Ph.D., membre externe.

ACKNOWLEDGEMENTS

My greatest debt is to Professor Soumaya Yacout and Dr. M-Salah Oulai. Their careful reading of my original manuscripts and their detailed, perceptive comments contributed in no mean way to the shape and clarity of the finished thesis. I would like to thank them for their constant supervision and invaluable help throughout all the phases of this research. Also, I would like to take the opportunity to thank Dr Bernard Clément, Dr. Luc Adjengue, Dr. Zdzislaw Klim, and Dr. Marek Balazinski for their contributions and advices as the president and members of the jury.

Also, I would like to thank deeply Mdm Marie-Hélène Leroux, who kindly helped me with the *Condensé en Français*.

On a general level, I see this thesis as a modest tribute to all the people that have had a formative influence on me over the years; teachers, professors and research assistants, and of course my parents, and friends which each in their own way, have helped to unlock some of the mysteries of this world.

RÉSUMÉ

La maintenance conditionnelle (CBM), est une stratégie d'intervention en maintenance basée sur l'observation à des intervalles réguliers d'éléments indiquant l'état de dégradation d'un équipement. Le principal problème est de prendre les meilleures décisions pour effectuer l'inspection de l'équipement, établir le lien entre les éléments observés lors de l'inspection et l'état de dégradation effectif de l'équipement et d'évaluer la fonction de fiabilité et la durée de vie résiduelle comme critères de décision. Plus particulièrement, cette thèse propose une démarche cohérente afin: (a) de déterminer la politique de remplacement optimale ainsi que l'intervalle d'inspection optimal des équipements lorsque le processus de détérioration n'est pas directement observable (non visible), (b) de déterminer la fonction de fiabilité (RF) ainsi que l'espérance de vie résiduelle (MRL) de tel équipement à chaque période d'observation afin d'évaluer le pouvoir de prédiction du modèle de remplacement proposé, (c) d'introduire des méthodes d'estimation des paramètres du modèle dans un contexte où la relation entre les symptômes de dégradation (indicateurs) et l'état réel de l'équipement n'est pas déterministe.

Un modèle taux de défaillance proportionnel de Cox (PHM) est utilisé pour modéliser le taux de défaillance de l'équipement. Un Modèle de Markov Caché (HMM) est proposé pour modéliser la dégradation non visible. Nous présentons un politique CBM optimale et un intervalle d'inspection optimal, RF et MRL de l'équipement en plus de l'estimation des paramètres permettant d'adapter le modèle en situation réelle, lorsqu'il

existe une relation stochastique entre la dégradation non visible de l'équipement et la valeur de l'indicateur d'inspection.

La programmation dynamique (DP), le Processus de décision markovien partiellement observable (POMDP) et les probabilités appliquées sont utilisés afin de résoudre les problèmes étudiés dans cette thèse. Des exemples numériques sont donnés pour illustrer les modèles proposés. Des simulations ont été effectuées afin de tester la robustesse et la convergence des méthodes d'estimation des paramètres proposés.

Mots-clés : Maintenance conditionnelle, Maintenance prédictive, Modèle taux de défaillance proportionnel de Cox, Espérance de vie résiduelle, Fonction de fiabilité, Processus de décision markovien partiellement observable, Simulation, Programmation dynamique, Optimisation stochastique.

ABSTRACT

Condition Based Maintenance (CBM) or predictive maintenance is based on observing an indicator of the degradation state of the equipment at different intervals of time in order to make an informed decision concerning the maintenance of this equipment. The objectives of this thesis are: (a) to determine the optimal replacement policy and optimal inspection interval for a piece of equipment when the degradation process is not outwardly visible, the indicator does not directly indicate the equipment state, and the inspections are costly; (b) to determine the Reliability Function (RF) and the Mean Residual Life (MRL) of such equipment at each observation moment; (c) to introduce a method for estimating the parameters of the models introduced in previous objectives.

Throughout this thesis, we assume that the equipment's unobservable degradation state transition follows a Markov Chain and we model it by a Hidden Markov Model. Bayes' rule is used to determine the probability of being in a certain degradation state at each observation moment. Cox's time-dependent Proportional Hazards Model (PHM) is considered to model the equipment's failure rate.

The first part of this thesis introduces a model to find the optimal inspection period for Condition Based Maintenance (CBM) of a system when the information obtained from the gathered data on the system does not reveal the system's exact degradation state and the collection of data is costly. By using dynamic programming, the system's optimal replacement policy and its total long run average operating maintenance cost are found.

Based on the long run average cost, the optimal inspection interval and the corresponding replacement criterion are specified. A numerical example shows the behaviour of the CBM model when the inspection is costly, and finds the optimal inspection period and the maintenance cost.

In the second part of this thesis, a model to calculate the Reliability Function (RF) and the Mean Residual Life (MRL) of a piece of equipment when its degradation state is not directly observable is introduced. At each observation moment, an indicator of the underlying unobservable degradation state is observed and the monitoring information is collected. The conditional reliability is derived from the PHM and it is used to calculate the RF and the MRL. Two examples are presented. The MRL is calculated at all possible state probabilities for four observation moments. It is shown that the MRL can be used as a supplementary decision tool, in particular when the cost elements of preventive replacement are unknown, or when there are criteria other than the cost to respect.

The third part of this thesis proposes a method to estimate the parameters of the models that were introduced in the previous parts. The parameters of the PHM, the Markov process transition matrix, and the stochastic matrix of observations/states are estimated based on the Maximum Likelihood Estimation (MLE) method. By using a Monte Carlo simulation approach, it is shown that the method used gives estimation results that converge to the real values of the parameters as the sample size increases. In addition, the behavior of the method has been examined when censored data exist.

Keywords: CBM, predictive maintenance, PHM, MRL, Reliability Function, POMDP, Monte Carlo Simulation, Dynamic Programming, Stochastic Optimization.

CONDENSÉ EN FRANÇAIS

Introduction

Dans cette thèse, nous nous penchons sur le modèle de maintenance préventive conditionnelle basé sur le modèle de taux de défaillance proportionnel (PHM) proposé par Cox [1972]. Nous supposons que les informations recueillies, θ, lors des inspections ne révèlent pas l'état de dégradation du système. Les information recueillies lors des inspections, θ, sont représentées par un nombre fini M de valeur possibles ($\theta = \{1, ..., M\}$). Les informations sont recueillies à intervalle Δ régulier (ou pratiquement régulier). Dans cette étude, Z représente l'état de dégradation de l'équipement et servira de variable diagnostique dans le PHM. Les conditions du système sont donc décrites de la façon suivante :

- L'équipement a un nombre fini et connu d'états de dégradations N. $J = \{1, ..., N\}$ représente l'ensemble des états de dégradation possibles.
- Le changement d'état de l'équipement suit un processus Markovien caché (HMM) où l'état de l'équipement n'est pas directement observé. Une matrice de transition P entre les différents états de dégradation possibles de l'équipement est introduite dans le modèle de base du taux de défaillance proportionnel;
- Les informations collectées lors des inspections de l'équipement sont stochastiquement représentatives de son état de dégradation. Une matrice Q donne les probabilités $q_{j\theta}$ d'obtenir une certaine information (indicateur) θ quand l'équipement est dans l'état j. Cette matrice est également introduite dans le modèle de base;

- La collection d'informations s'effectue lors des inspections périodiques (à intervalles fixes Δ);
- La défaillance ne fait pas partie des états de dégradation possibles. La défaillance pour survenir à tout moment, à n'importe quel état de dégradation, et entraîne l'arrêt de l'équipement. La défaillance sera donc immédiatement remarquée.
- Le taux de défaillance de l'équipement suit un PHM qui est fonction du temps. Le taux de défaillance $h(s,Z_k)=h_0(s)\psi(Z_k)$ est le produit de deux fonctions indépendantes : $h_0(.)$ est une fonction représentant uniquement l'âge de l'équipement et $\psi(.)$ est une fonction représentant uniquement l'état de dégradation.

Puisque l'état de dégradation de l'équipement ne peut être observé, nous avons utilisé la probabilité conditionnelle π_i^k d'être à l'état i à la fin de la période d'observation k, tel qu'introduit par Ghasemi et al. [2007]. π_i^k est donné par:

$$\pi^k = \left\{ \pi_i^k;\ \ 0 \le \pi_i^k \le 1 \text{ for } i=1,\ldots,N, \sum_{i=1}^{N} \pi_i^k = 1 \right\},\ k = 0,1,2,\ldots \qquad \mathbf{(1)}$$

Avec :

$$\pi_i^0 = \begin{cases} 1 & i=1 \\ 0 & 1<i\le N \end{cases}.$$

π_i^0 signifie que l'équipement sera dans l'état 1 initialement ou s'il est remplacé. Après qu'une observation θ soit réalisée, la probabilité conditionnelle π^{k+1} est recalculée. En utilisant la formule de Bayes, et sachant que l'observation s'est produite à k+1, alors $\pi_j^{k+1}(\theta)$ est donné par :

$$\pi_j^{k+1}(\theta) = \sum_{i=1}^{N} \pi_i^k p_{ij} q_{j\theta} \Big/ \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_i^k p_{ij} q_{j\theta}, \ j = 1,\ldots,N \quad \textbf{(2)}$$

Dans les sections suivantes, nous expliquerons les objectifs de la thèse, les méthodologies utilisées et les solutions développées.

Premier objectif: Intervalle d'inspection optimal et politique de remplacement optimale

Le premier objectif consiste à déterminer, dans le cas où les inspections représentent des frais considérables, l'intervalle d'inspection optimal et la politique de remplacement optimale. Le coût d'un remplacement préventif est noté C, alors que le coût d'un remplacement après défaillance est noté $K + C$ où K, $C > 0$. Les deux actions, le remplacement préventif et le remplacement après défaillance, sont instantanées. Le coût de l'inspection sont représentées par C_I et son indépendants de l'intervalle d'inspection. Nous définissons $V(k, \pi^k)$ comme le coût minimum *de la maintenance et de l'inspection* au cours de la période de renouvellement, tandis que l'équipement est au k-ième point d'inspection avec les probabilités conditionnelles $\pi_i^k, i = 1,\ldots,n$. La période de renouvellement est définie par l'intervalle de temps compris entre deux remplacements consécutifs, que ces remplacements soient préventifs ou fassent suite à une défaillance. $V(k, \pi^k)$ est donné par:

$$V(k, \pi^k) = \min\left\{kC_I + C + V(0, \pi^0), W(k, \pi^k, g)\right\} \quad \textbf{(3)}$$

où $kC_I + C + V\left(0,\pi^0\right)$ est coût total au cours de la période de renouvellement, tandis que l'équipement est au k-ième point d'inspection, si la décision est d'effectuer le remplacement préventif. De plus, $W\left(k,\pi^k,g\right)$ est défini de la façon suivante:

$$\begin{aligned} W\left(k,\pi^k,g\right) = {} & \left[kC_I + K + C + V\left(0,\pi^0\right)\right]\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right] - g\overline{\tau}\left(k,\pi^k,\Delta\right) \\ & + \left[\sum_{\theta=1}^{M} V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta \mid k,\pi^k\right)\right]\overline{R}\left(k,\pi^k,\Delta\right) \end{aligned} \qquad \textbf{(4)}$$

où $\left[kC_I + K + C + V\left(0,\pi^0\right)\right]$ représente le coût total au cours de la période de renouvellement, tandis que l'équipement est au k-ième point d'inspection, si la décision est de ne pas effectuer de remplacement préventif (ne rien faire).

et $\left[\sum_{\theta=1}^{M} V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta \mid k,\pi^k\right)\right]$ représente le coût minimum estimé de la maintenance et de l'inspection, tandis que l'équipement est au k+1-ième point d'inspection. $\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]$ et $\overline{R}\left(k,\pi^k,\Delta\right)$ représentent respectivement la probabilité de défaillance Durant la k-ième période et la probabilité que l'équipement survive jusqu'au début de la période k+1 quand les probabilités conditionnelles à la période k sont $\pi_i^k, i=1,\ldots,n$. $\overline{\tau}\left(k,\pi^k,\Delta\right)$ est la durée moyenne de séjour de l'équipement à la période k+1 lorsque les probabilités conditionnelles à la k-ième période $\pi_i^k, i=1,\ldots,n$ sont disponibles. g représente le coût moyen par unité donnée de temps (jour, semaine, etc.) pour un horizon de durée infini. g inclus seulement le coût de remplacement : les coûts d'inspection sont exclus.

$\overline{R}(k,\pi^k,\Delta)$ et $\overline{\tau}(k,\pi^k,\Delta)$ sont ainsi calculés:

$$\overline{R}\left(k,\pi^k,\Delta\right)=\sum_{i=1}^{N} R\left(k,i,\Delta\right)\ \pi_i^k \tag{5}$$

$$\overline{\tau}\left(k,\pi^k,\Delta\right)=\int_0^{\Delta} \overline{R}\left(k,\pi^k,s\right)\ ds \tag{6}$$

où $R\left(k,i,t\right)=\exp\left(-\psi\left(i\right)\int_{k\Delta}^{k\Delta+t} h_0(s)ds\right)$.

Selon ces hypothèses et, en utilisant le modèle développé, la politique de remplacement sera exprimée par :

$$a\left(k,\pi^k\right)=\begin{cases}\infty & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]<g\overline{\tau}\left(k,\pi^k,\Delta\right)\\ 0 & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]\geq g\overline{\tau}\left(k,\pi^k,\Delta\right)\end{cases} \tag{7}$$

Cette politique est donc fonction de K (le coût de défaillance), $\overline{R}\left(k,\pi^k,\Delta\right)$ (la fiabilité conditionnelle à la k-ième période) et $\overline{\tau}\left(k,\pi^k,\Delta\right)$ (la durée de séjour à la k-ième période). Afin de déterminer la valeur de g^*, nous utilisons une méthode récursive. En remplaçant g par g^*, la politique optimale de remplacement est obtenue à partir de l'équation 7.

L'intervalle d'inspection optimal est choisi à partir d'un ensemble fini de L possibilités ($\Delta_l; l=1,2,...,L$). L'intervalle d'inspection optimal est celle qui, parmi les L possibilités, minimise le coût total à long terme G^* où $G_l^*=g_l^*+\dfrac{C_I}{\Delta_l}$. En pratique, après avoir déterminé la valeur de G^* et après avoir choisi le Δ^* (Δ correspondent à G^*), la politique de remplacement optimale est déterminée en utilisant l'équation 7.

Deuxième objectif : Fonction de fiabilité et Durée de vie résiduelle moyenne

Dans le troisième chapitre de cette thèse, nous avons calculé la fonction de fiabilité (RF) and la durée de vie résiduelle moyenne (MRL) de l'équipement. Le MRL et le RF peuvent être utilisé comme outil additionnel d'aide à la décision, en particulier dans le cas où les coûts remplacement préventifs sont inconnus, ou dans le cas où le coût n'est pas le seul critère en jeu. En connaissant le MRL et le RF un industriel pourra tirer avantage des évènements à venir (par exemple : un arrêt de production planifié) qui ne sont pas normalement pris en compte dans la politique de remplacement à coût optimal, afin de faire du remplacement préventif.

La fiabilité conditionnelle à (k, Z_k), i.e. à la k-ième période d'observation lorsque l'état de l'équipement est Z_k et que $t > \Delta$, est ainsi formulé :

$$\begin{aligned} R(k,Z_k,t) &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_1, Z_2, \ldots, Z_k\right), t > \Delta \\ &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_k\right), t > \Delta \end{aligned} \quad \textbf{(8)}$$

Dans le cas d'une observation directe, sous l'hypothèse $Z_k = i$, nous avons démontré que:

$$R(k,i,t) = \begin{cases} \exp\left(-\psi(i)\int_{k\Delta}^{k\Delta+t} h_0(s)ds\right) & 0 < t \leq \Delta \\ R(k,i,\Delta)\sum_{j=1}^{N} p_{ij} R\left(k+1, j, (t-\Delta)\right) & t > \Delta \end{cases} \quad \textbf{(9)}$$

Dans le cas d'une observation indirecte, nous définissons $\overline{R}\left(k,\pi^k,t\right)$ comme étant la fiabilité conditionnelle de l'équipement à la k-ième période d'observation, quand les probabilités conditionnelles à la période k sont $\pi_i^k, i=1,\dots,n$. $\overline{R}\left(k,\pi^k,t\right)$ est donc calculé de la façon suivante:

$$\overline{R}(k,\pi^k,t)=\begin{cases} \displaystyle\sum_{i=1}^{N}\pi_i^k \exp\left(-\psi(i)\int_{k\Delta}^{k\Delta+t}h_0(s)ds\right) & 0<t\leq\Delta \\ \displaystyle\sum_{i=1}^{N}\pi_i^k R(k,i,\Delta)\sum_{j=1}^{N}p_{ij}\times R\left(k+1,j,(t-\Delta)\right) & t>\Delta \end{cases} \qquad \textbf{(10)}$$

Dans le cas d'une observation directe, le MRL est défini de la façon suivante :

$$e\left(k,i\right)=\int_{k\Delta}^{\infty}R\left(k,i,t\right)dt$$

Nous calculons le MRL, $\overline{e}\left(k,\pi^k\right)$, à la k-ième période d'observation, quand les probabilités conditionnelles à la période k sont $\pi_i^k, i=1,\dots,n$. et lorsqu'il peut être représenté par:

$$\overline{e}\left(k,\pi^k\right)=\int_{k\Delta}^{\infty}\overline{R}\left(k,\pi^k,t\right)dt \qquad \textbf{(11)}$$

Troisième objectif: Estimation des Paramètres

Tous les modèles mathématiques, incluant les modèles introduits au cours des deux premières sections, sont basés sur un ensemble de paramètres devant être estimés pour pouvoir être appliqués en situation réelle. Ces paramètres sont estimés à partir de l'information disponible sur le système à l'étude. Dans le chapitre 4, nous introduisons des méthodes permettant d'estimer les paramètres des modèles mathématiques utilisés dans cette thèse. Dans notre cas, afin d'appliquer ces modèles à une situation réelle, nous devons estimer les paramètres du PHM, les probabilités de la chaîne de Markov et les probabilités la matrice des états/indicateurs.

Nous avons défini *T*, la durée de vie de l'équipement, une variable aléatoire positive continue et i.i.d. « *independent identically distributed* ». Nous avons également défini $\theta(s)=\left\{\theta^1,\theta^2,...,\theta^k\right\};\quad s\le T; k=1,2,...;\quad k\Delta\le s$ l'historique des valeurs prises par l'indicateur jusqu'au temps *s*, où Δ est l'intervalle d'observation. Les valeurs prise par l'indicateur jusqu'au temps *s*, peuvent être représentée par la distribution de probabilité conditionnel de l'état de l'équipement par $\pi(s)=\left\{\pi^1,\pi^2,...,\pi^k\right\};\ s\le T;\ k=1,2,...;$ $k\Delta\le s$, où les élément de $\pi(s)$ sont calculés à partir des équation 1 et 2 à chaque période d'observation.

Nous avons démontré que la fonction de survie est exprimée par :

$$R\left(t,\theta(t)\right)=$$
$$\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1}\mid T>(l+1)\Delta,\pi^l\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^l\exp\left(-\int_{l\Delta}^{(l+1)\Delta}h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^k\exp\left(-\int_{k\Delta}^{t}h(\tau,i)\,d\tau\right)$$

(12)

et puisque $f\left(t,\theta(t)\right)=\lim_{\Delta t\to 0^+}\dfrac{\Pr\left(t\le T<t+\Delta t\mid T>t,\theta(t)\right)}{\Delta t}=-\dfrac{dR\left(t,\theta(t)\right)}{dt}$:

$$f\left(t,\theta(t)\right)=$$
$$\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1}\mid T>(l+1)\Delta,\pi^l\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^l\exp\left(-\int_{l\Delta}^{(l+1)\Delta}h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^k\exp\left(-\int_{k\Delta}^{t}h(\tau,i)\,d\tau\right)h(t,i)$$

(13)

L'estimation des paramètres se fait de deux façons différentes selon que les données soient censurées ou non.

Dans le cas de données non-censurées, pour un ensemble de *n* expériences indépendantes, nous assumons que T_r est le temps de défaillance de la r-ième expérience. La vraisemblance de l'ensemble de paramètres inconnus Ω, peut être ainsi calculée à l'aide des données disponibles:

$$L(\Omega)=\prod_{r=1}^{n} f\left(T_r,\theta\left(T_r\right);\Omega\right) \tag{14}$$

Dans le cas de données censures, nous avons définis un indicateur de censure :

$$\delta_r=\begin{cases}1 & \text{s'il y a eu défaillance}\\ 0 & \text{si l'équipement a été retiré}\end{cases}$$

Cet indicateur de censure indique si la valeur de T_r est : 1) une défaillance ou 2) l'instant auquel l'équipement à été retiré du service, i.e. T_r représente une censure. La vraisemblance de l'ensemble de paramètres inconnus Ω, peut être ainsi calculée à l'aide des données disponibles:

$$L(\Omega)=\prod_{r=1}^{n} f\left(T_r,\theta;\Omega\right)^{\delta_r} R\left(T_r,\theta;\Omega\right)^{1-\delta_r} \tag{15}$$

Dans les deux cas, avec données censures ou non, utilisant une technique d'optimisation telle que la « *line search method* », la valeur maximale de la fonction de vraisemblance est obtenue.

La convergence et la robustesse des méthodes d'estimation des paramètres introduites dans cette thèse ont été évaluées à l'aide de simulation de Monte Carlo. Nos résultats montrent que les paramètres estimés par les méthodes introduites convergent vers les valeurs réelles lorsque la taille de l'échantillon augmente. Dans le cas de données censurées, la méthode donne d'excellents résultats même lorsque le taux de censure atteint les 50%.

Conclusion

Cette thèse offre plusieurs outils relatifs à la maintenance conditionnelle. Elle introduit une politique de remplacement optimale de l'équipement ainsi qu'une technique permettant de déterminer l'intervalle d'inspection optimal. La thèse offre également d'utiliser le MRL et le RF comme mesure de la performance future de l'équipement. Le MRL et le RF permettent aux industriels de prendre des décisions éclairées relatives à la maintenance de l'équipement. Finalement, les méthodes d'estimation des paramètres des différents modèles proposés dans la thèse permettent l'application de ces outils en situation réelle. De plus, cette thèse innove en adressant le problème du choix des paramètres qui a rarement été soumis à l'étude jusqu'ici.

Des travaux additionnels pourraient être effectués afin d'élargir le champ d'étude de cette thèse et inclure les cas où l'intervalle d'inspection est variable. Dans ce cas, la date de la prochaine inspection et la politique de remplacement minimisant les coûts devront être déterminés à chaque inspection. À chaque fois, on devra décider si l'équipement doit être remplacé ou s'il demeure en place jusqu'à la prochaine inspection. Si l'équipement demeure en place, la date de la prochaine inspection sera déterminée à l'aide des donnés historiques. Voilà pourquoi, dans un tel cas, l'intervalle d'inspection est variable tel que mentionné plus haut.

Des travaux additionnels pourraient également être effectués afin d'introduire une politique optimale de remplacement dans le cas où plusieurs types de réparations peuvent être effectuées (on ne remplace pas systématiquement l'équipement). En situation réelle, il est possible d'effectuer différents travaux de réparations auxquels sont associés des coûts différents. Chaque réparation modifiera l'état de dégradation de l'équipement afin de le ramener à un état acceptable connu.

Finalement, rappelons que dans cette thèse, nous avons fait l'hypothèse que le modèle de Markov est homogène. Le cas contraire (non homogène) pourrait également être une piste de recherche intéressante.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

CBM .... Condition Based Maintenance

CDF .... Cumulative Distribution Function

DP .... Dynamic Programming

EM .... Expectation Maximization

IFR .... Increasing Failure Rate

LR .... Likelihood Ratio

MC .... Markov Chain

MDP .... Markov Decision Processes

MLE .... Maximum Likelihood Estimation

MP .... Markov Process

MRL .... Mean Residual Life

MTTF .... Mean Time To Failure

PDF .... Probability Distribution Function

PHM .... Proportional Hazards Model

POMDP .... Partially Observed Markov Decision Processes

RF .... Reliability Function

ST .... Stochastically

TP2 .... Totally Positive of order 2

CHAPTER 1 : INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Depending on the specific industry, maintenance costs can represent between 15 and 40 percent of the costs of goods produced. For example in food related industries, the average maintenance cost represents about 15 percent of the production cost; while in iron and steel, pulp and paper and other heavy industries maintenance cost represents up to 40 percent of the total production cost [Mobley, 2002]. Since maintenance cost is a major part of the total operating costs of manufacturing and production, one of the tools of securing the productivity and decreasing the production cost is to have a well functioning maintenance system and strategy. The maintenance system has the role of looking after the equipment and keeping track of it in order to secure the functional requirements of productivity, safety, and quality. Without a performing maintenance system money will be lost due to lost production capacity, excessive amount of spare parts, and lack of quality, late deliveries and loss of safety.

Traditionally the maintenance systems are categorized either as *preventive* or *corrective*. The preventive maintenance aims at preventing the components, the sub-systems or the equipment from deteriorating or failing by performing repair, overhaul, service or component's replacement. The corrective maintenance is performed after system or equipment's failure or breakdown. While preventive maintenance is *Age Based*, i.e. the equipment's maintenance is based on its age, since few decades some industries have started to perform maintenance actions in a *Condition Based* or *predictive* approach. In the latter, the equipment's condition is the key parameter in triggering the appropriate maintenance actions. This approach is called *Condition Based Maintenance* (CBM) and/or sometimes referred to as *predictive maintenance*. Figure 1-1 depicts a schematic of different types of maintenance systems. The equipment's condition may be obtained through different levels of automation, from human visual inspection, to on-line highly sophisticated condition monitoring equipment.

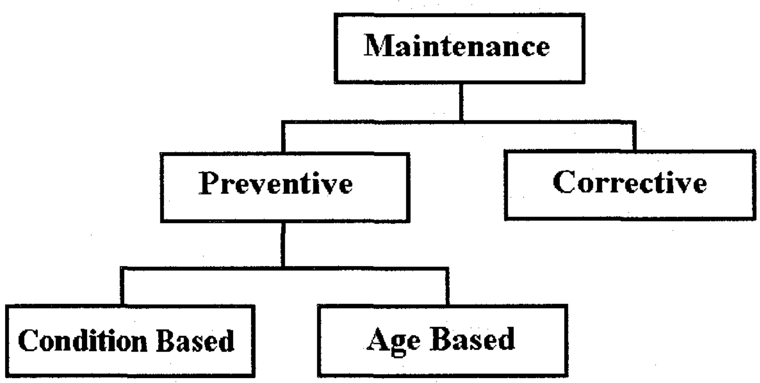

Figure 1-1: Schematic of different types of maintenance

1.1.1 Condition Based Maintenance

Condition Based Maintenance has been defined as "Maintenance actions based on actual condition obtained from in-situ, non-invasive tests, operating and condition measurement." [Mitchell, 1998] or "CBM is a set of maintenance actions based on real-time or near-real time assessment of equipment degradation state which is obtained from embedded sensors and/or external tests and measurements taken by portable equipment." [Butcher, 2000]. In these definitions and many similar ones that can be found in the literature as well as in the internet, the common idea is that the maintenance actions are not considered until there is an obvious need. This should increase the availability of equipment and decrease maintenance cost, including labor and spare parts costs.

The purpose of CBM is to eliminate breakdowns and protract the preventive maintenance intervals which should result in an increase in the availability of equipment. Using CBM technology, the condition monitoring data are analyzed in depth to determine whether the equipment is running at a normal operating condition or not. If the preset limits for normal condition are exceeded, the maintenance actions are performed. With this information, it is easier to plan the maintenance actions more effectively [Marcus *et al.*, 2002].

CBM systems may need several components and level of automation in order to give the required information to make the right maintenance decision. Some companies may use hand-held devices out in the field and then make analysis of the data later in laboratories, while others may use more complicated on-line systems that give the results right away. In both cases, the way that the companies use the information determines whether they are having a CBM strategy or they are just inspecting their equipments. When having a CBM program, the results of the analysis are taken into account and the maintenance actions are planned accordingly.

A variety of technologies may be used as part of a CBM program. Since mechanical equipment is part of most industries' equipment, vibration monitoring is generally the most used technique in CBM programs. This technique is limited to monitoring the mechanical condition. For this reason a CBM program may include one or more of the following monitoring and diagnostic techniques:

- Vibration Monitoring
- Thermography
- Tribology
- Ultrasonic monitoring
- Other nondestructive testing techniques
- Process Parameters
- Visual Inspection

In the next part, a general description of each of these techniques is provided [Mobley 2002].

Vibration Monitoring

Since most of the typical industry equipment is mechanical, this technique has the widest application. This technique uses the noise or vibration created by mechanical equipment to determine its actual condition. The degradation of the mechanical condition can be detected using vibration-monitoring techniques.

Thermography

Thermal anomalies of equipment, i.e. areas that are hotter or colder than they should be, can be used to monitor the conditions of the equipment. Thermography uses instrumentation designed to monitor the emission of infrared energy, i.e. temperature, by the equipment to determine its degradation state. Infrared technology is based on the fact that; all the objects having a temperature above absolute zero emit energy or radiation.

Tribology

Tribology refers to design and operating dynamics of the bearing-lubrication rotor support structure of machinery. Several tribology techniques like: lubricating oils analysis, spectrographic analysis, and ferrography and wear particle analysis can be used for predictive maintenance. For instance, some forms of lubricating oil analysis will provide an accurate quantitative breakdown of individual chemical elements, both oil additive and contaminates, contained in the oil. A comparison of the amount of trace of metals in successive oil samples can indicate wear patterns of oil wetted parts in the equipment and will provide indication of impending machine failure.

Ultrasonic Monitoring

This predictive maintenance technique uses principles similar to vibration analysis. Both monitor the noise generated by machines or equipment to determine their actual degradation state. Unlike vibration monitoring, ultrasonic monitoring monitors the higher frequencies, i.e. ultrasound, produced by unique dynamics in process systems or machines. The normal monitoring range for vibration analysis is from less than 1 Hertz to 20,000 Hertz. Ultrasonic techniques monitor the frequency range between 20,000 Hertz and 100 kHz. This technique is ideal for detecting leaks in valves, steam traps, piping and similar process systems.

Process Parameters

Machinery that is not operating within acceptable efficiency parameters can severely limit the productivity of many types of equipment. As an example of the importance of process parameters monitoring, consider a process pump that may be critical to industry operation. The pump can be operating at less than 50% efficiency and the predictive maintenance program which does not consider the efficiency, will not detect the problem.

Visual Inspection

Regular visual inspection of the machinery and equipment is a necessary part of any predictive maintenance program. In many cases, visual inspection will detect potential problems that will be missed using the other predictive maintenance techniques. Routine visual inspection of critical equipment will augment the other techniques and insure that potential problems are detected before serious damage can occur.

1.1.2 Proportional Hazard Model

Introduced by D. R. Cox, the Proportional Hazards Model (PHM) was developed in order to take into account the effects of equipment's condition that influences its times-to-failure. The model has been broadly used in the biomedical field [Leemis 1995] and recently there has been an increasing application in reliability engineering [Makis and Jardine 1992].

According to the PHM, the failure rate of a piece of equipment is affected not only by its operating time, i.e. its age, but also by the degradation state under which it operates. It is clear that this factor affects the failure rate of the equipment. Equipment in a better degradation state has less chance to fail than a worn one even if they both have the same age. The proportional hazards model assumes that the failure rate of a piece of equipment is the product of a baseline failure rate, $h_0(t)$, which is a function of the equipment's age t only, and a positive function $\psi(Z_t)$, that is independent of age and incorporates the effects of the equipment's degradation state Z_t. The failure rate of a unit is then given by $h(t,Z_t)=h_0(t)\psi(Z_t)$. Z_t is the random variable representing the degradation state of the equipment at time t. PHM also assumes that the form of $\psi(Z_t)$ is known and is the exponential form and is given by $\psi(Z_t)=e^{\gamma Z_t}$, where γ is the degradation state coefficient.

The Weibull distribution is one of the most commonly used distributions in reliability engineering because of the several shapes it can take for different values of its parameters. Hence it can model a great range of data and life characteristics. By considering a two parameters Weibull distribution to formulate the baseline failure rate, the parametric proportional hazard model is introduced. In this case, the baseline failure rate is given by: [Cox 1984]

$$h_0(t)=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}$$

where η and β are the scale and shape parameters of the Weibull distribution function, respectively. The PHM failure rate then becomes:

$$h(t, Z_t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma Z_t}$$

More details about PHM and Weibull distribution are presented later.

1.1.3 CBM with Indirect Observations

In many real cases, the degradation state of the equipment is not outwardly visible while the failure of the equipment is immediately obvious and causes the equipment to cease functioning. This is known as equipment with *obvious failures* as opposed to *silent failures*, which are not immediately discovered. Ideally, inspection of a piece of equipment reveals the degradation state of the equipment with certainty. This type of observation system is known as *Perfect Observation* or *Direct Observation* as opposed to *Imperfect Observation*, *Partially Observed system*, or *Indirect Observation* [Lin *et al.*, 2003; 2004, Fernandez-Gaucherand, 1993, Wang and Christer, 2000]. If the observations are taken in selected periods rather than all periods, the system is *Partly Observed* in opposed to *Completely Observed* which is not considered in this thesis.

In the case of the direct observation, it is assumed that the information collected regarding the equipment's condition (indicator) θ, is a direct pointer to the equipment's degradation state Z. The indicator is assumed to be in a *some-to-one* or a *one-to-one* relationship with the degradation state that influences the time-to-failure of the equipment. In a some-to-one approach, each possible indicator's value, from a predefined interval, refers to one degradation state. In this approach, there is no indicator value that can refer to more than one degradation state. If the condition monitoring reading is of value θ_i where $i=1,2,...$, the state will be a certain value Z_j, where $j=1,2,....$ In this approach, each indicator value θ_i, refers to only one state Z_j. At the same time, any state Z_j, may be referred to by several possible values of the indicator in a predefined interval [e.g. Makis and Jardine, 1992]. Figure 1-2 demonstrates the some-to-one characteristic. It can be seen in the figure that any indicator value in the interval $[a,b)$, e.g. θ_1 and/or θ_3, refers to the same state value Z_1. There is no possibility to have more than one state referred to by a single indicator value.

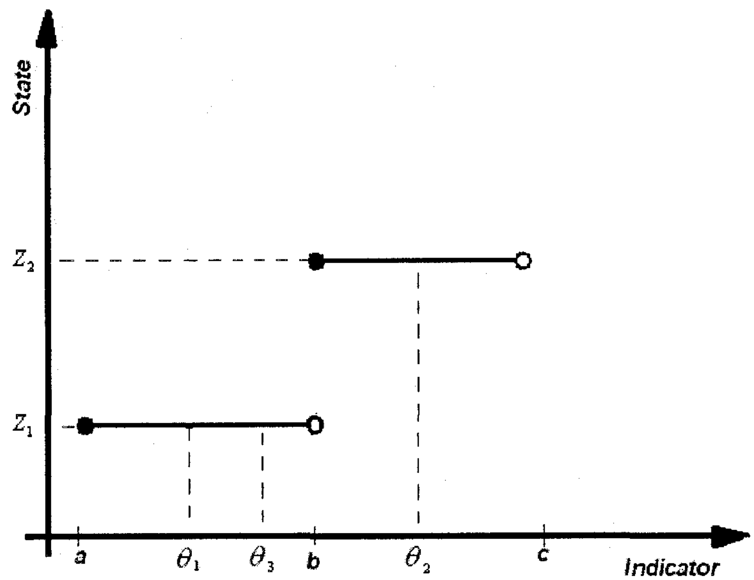

Figure 1-2: Direct observation for equipment with 2 degradation states (some-to-one relationship)

In the one-to-one approach, each indicator value is assumed to be a direct pointer to and only to one degradation state of the equipment, and the indicator's value is used directly as the diagnostic covariate in the PHM (e.g. Kumar *et al.,* 1996). Nevertheless, in both cases, a certain value of the indicator refers to a certain degradation state deterministically.

Realistically, information may contain noise due to errors of measurement, interpretations, accuracy of measurement instruments, etc. and may not reveal the exact degradation state of the equipment. The information is, however, stochastically correlated with the underlying state. In this case, information collected may refer to more than one possible state. For example, a certain level of vibration (indicator) θ_1, may be read while the equipment is in any of two different levels of degradation states Z_1 and Z_2.

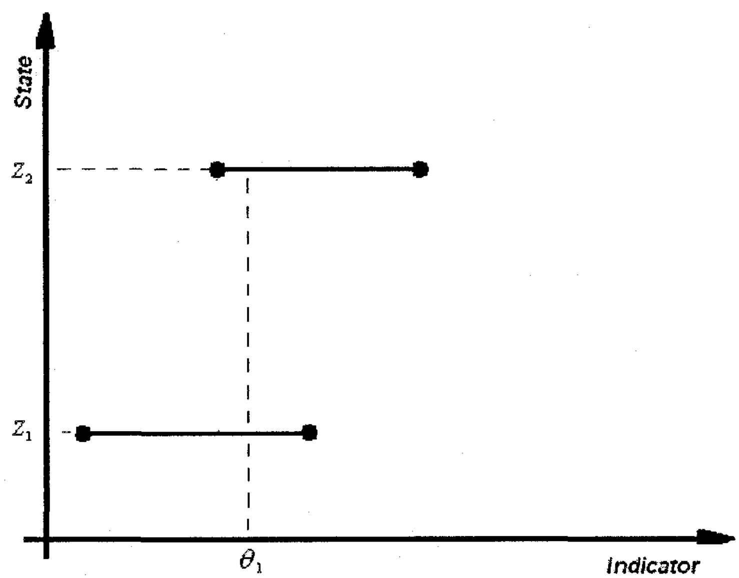

Figure 1-3: Indirect or imperfect observation (some-to-some relationship)

This situation is represented by a probability distribution function or a stochastic matrix. In the latter case, the relationship between the collected indicator and the degradation state is *some-to-some*. One collected indicator value may refer to several degradation states and vice versa. Figure 1-3 illustrates the stochastic relationship between the indicator and the state in this case. As shown in this figure, the indicator value θ_1, may refer to either state Z_1 or state Z_2. An indicator value may be a sign of several possible degradation states and equipment in certain degradation state may demonstrate different indicator values.

The relationship between the indicator's value and the state is introduced via an observation probability matrix or a probability distribution. For example, in Figure 1-4, if the state is $Z_i; i=1,2$, the probability of observing different values of the indicator follows a normal distribution $N\left(\mu_{X_i}, \sigma_{X_i}^2\right)$.

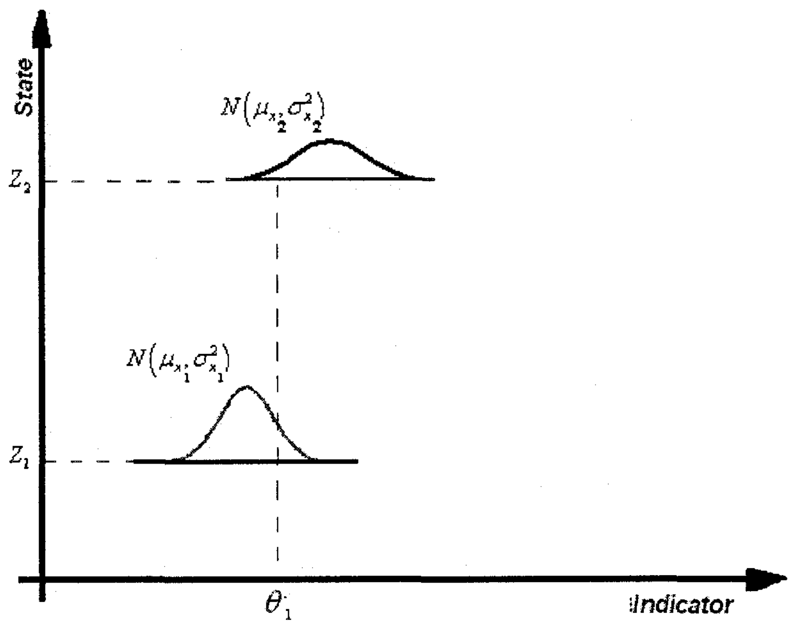

Figure 1-4: Probabilistic relationship between the indicator and state

The latter case, combined with PHM and Markov Process, was originally introduced by Ghasemi *et al.,* [2007] to address the indirect observation problem and to propose a solution to the main drawback of the time-dependent PHM, i.e. the inclusion of only the latest condition monitoring information in the model. In this thesis, we consider a some-to-some indicator-state relationship.

To demonstrate the impact of not considering all the history of information, consider Figure1-5 which depicts an explanatory example of a piece of equipment that follows a time-dependent PHM. In this example, the equipment demonstrated by the bold line has been in the state $Z=1$ from time zero to time t_2. It can be seen that the failure rate of the equipment at times t_1 and t_2 are equal to h_1 and h_2, respectively.

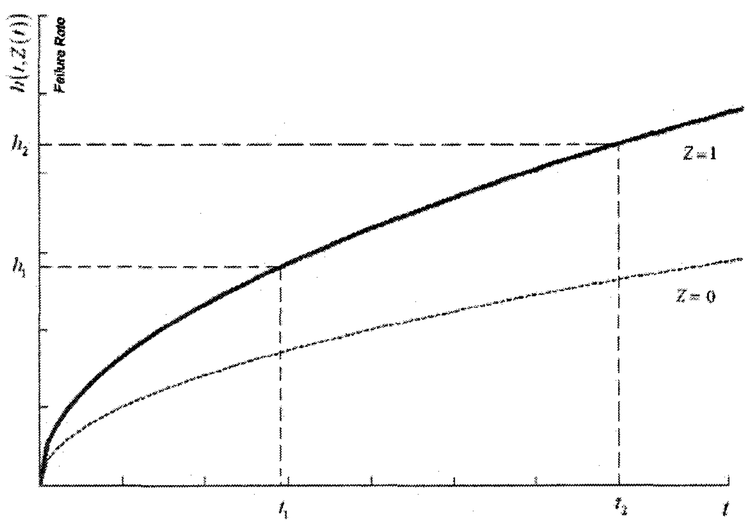

Figure1-5: An explanatory example of a time-dependent PHM without a state change

Now we consider the case demonstrated in Figure 1-6. From time zero to time t_1, the equipment state is $Z=0$ and right after, from time t_1 to time t_2, the equipment's state is $Z=1$. According to the PHM, at time t_2 while the equipment is at state $Z=1$, the failure rate of the equipment is again h_2 and the fact that the equipment has been in state $Z=0$ from time zero to time t_1, has no effect on the value of the failure rate h_2, at time t_2. Obviously, this is not realistic.

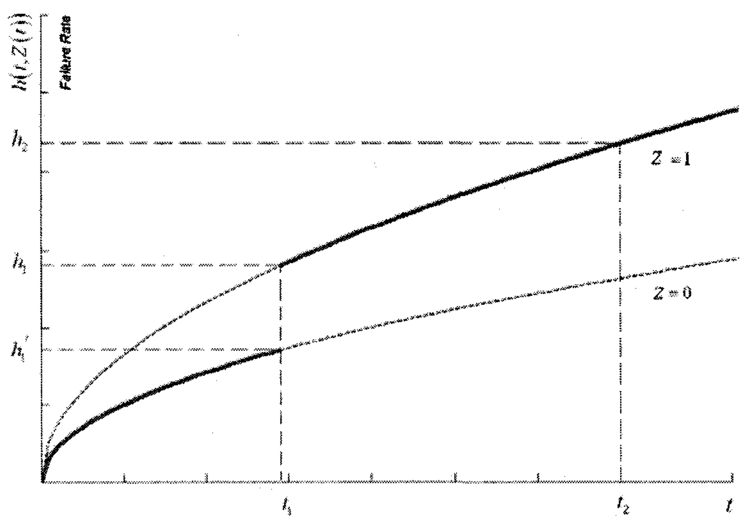

Figure 1-6: An explanatory example of a time-dependent PHM with state change

Conceptually, this drawback may be addressed as shown in Figure 1-7. In this figure, if the equipment is in state $Z = 0$ until time t_1, the failure rate at that time is h_1. If at that time the state changes to $Z = 1$, then the increase in the failure rate will follow the path of the bold line from h_1' on the curve of $Z = 1$. This is equivalent to assuming that the equipment was in state $Z = 1$ from time zero but its age is $t_1 - t$. t is demonstrated in the figure. Also, after $t_2 - t_1$, i.e. at age t_2, the failure rate will be h_2' and not h_2. These examples clearly show how the original approach of the time-dependent PHM gives a misleading value of the failure rate by ignoring the degradation history.

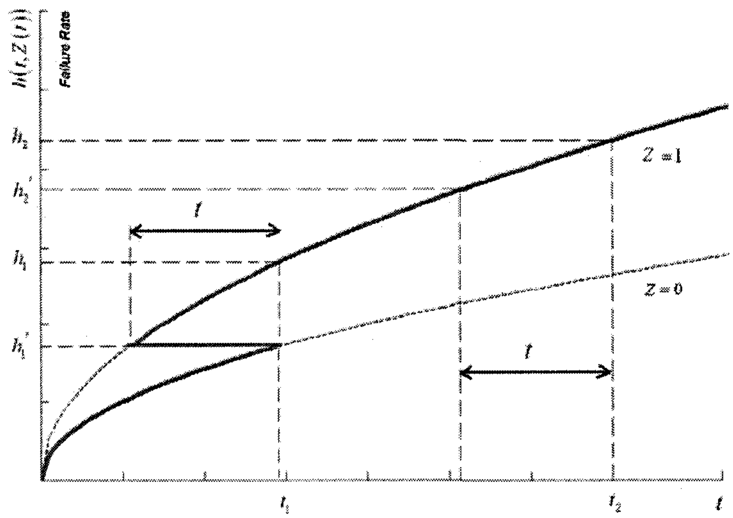

Figure 1-7: Conceptual solution for PHM drawback

In what follows, the notations and basic definitions and methods used in this thesis are explained.

1.2 Preliminary Notations

We define a continuous random variable T, as the time-to-failure of the equipment which can take any value in $[0,\infty)$.

1.2.1 The Probability and Cumulative Density Functions

The ***Probability Density Function*** (PDF), $f(t)$, and ***Cumulative Distribution Function*** (CDF), $F(t)$, of time-to-failure, T, are such that:

$\Pr(a \le t \le b) = \int_a^b f(t)dt$ and $F(t) = \Pr(T \le t) = \int_0^t f(s)\,ds = 1 - R(t)$.

The CDF is used to measure the probability that the equipment in question will fail before the associated time value t, and is also called *unreliability* where $R(t)$ is the ***R****eliability Function* (RF).

1.2.2 Reliability Function (RF)

When a piece of equipment is subjected to condition monitoring, data concerning one or more indicators of the degradation state are collected periodically. The information obtained from this data is used to establish a diagnosis of equipment's condition and a prognosis for future performance. Two measures of this performance are the failure rate (or the hazard function), and the Mean Residual Life (MRL). These two measures are calculated from the reliability function.

In reliability analysis, two reliability functions are of interest. The first is the unconditional reliability function given by the probability $P(T > t)$, which is the probability that the failure time T, of a piece of equipment that has not yet been put into operation, is bigger than a certain time t. The second is the conditional reliability function calculated by $P(T > t \mid T > \tau)$, which is the probability that the time-to-failure T is bigger than t, knowing that the equipment has already survived until time τ, where $\tau < t$. In some reliability analysis, it is assumed that every piece of equipment is used in the same environment and under the same conditions. This assumption allows the calculation of the MRL and the hazard function prior to the actual use of the equipment. In real-life, the environments in which the equipment is performing and the conditions of utilization affect the process of degradation. Consequently, the failure rate, the conditional reliability, and the residual life of the equipment are affected. Taking this fact into consideration improves the diagnosis of the equipment's degradation state and the prognosis for future performance.

1.2.3 The Failure Rate Function

The *failure rate function* enables the determination of the number of failures occurring per unit of time. The failure rate function is mathematically calculated as:

$$h(t) = \frac{f(t)}{R(t)}$$

This gives the instantaneous failure rate. The cumulative of failure rate is called *hazard function.*

$$H(t) = \int_0^t h(s)\,ds$$

These functions are useful in characterizing the failure behavior of the equipment [Ross 1997].

1.2.4 Mean Life or Mean Time-to-failure (MTTF)

The *mean life* function, which provides a measure of the average time of the equipment's life, is given by:

$$MTTF = \overline{T} = \int_0^\infty t f(t)\,dt$$

This is the expected or average time-to-failure for a piece of equipment with instantaneous replacement and is denoted as the *MTTF*, ***Mean Time To Failure***. The

MTTF, even though an index of reliability performance, does not give much information on the failure distribution of the equipment in question when dealing with most probability distributions.

1.2.5 Mean Residual Life (MRL)

In biomedical science, researchers analyze survivorship of patients by MRL. Actuaries apply MRL to set the rates and benefits for life insurance [Ghai and Mi, 1999]. In general, MRL provides a more descriptive measure of an aging process than the hazard rate. The hazard rate takes just the instantaneous present into account while MRL considers the whole future [Bradley and Gupta, 2003; Siddiqui and Caglar, 1994].

MRL or *Remaining Useful Life* (RUL) is defined as "the expected time interval between the point of gathering the information to the point of future failure based on the history of the condition monitoring and the performed preventive actions" [Wang and Zhang, 2005]. Wang and Christer [2000] consider a similar concept; the *Conditional Residual Time* (CRT), defined as "the time lapse from any time point that monitoring information is obtained to the time that it may fail given no other preventive maintenance action".

Based on the definitions, MRL can be determined by $E\{T-t \mid T>t, Z(t)\}$ where T is a random variable indicating the equipment's time-to-failure, t is the current observation time and $Z(t)$ is the state of the equipment at time t.

There are two general methods to calculate the targeted equipment's MRL. First; the

MRL is calculated based on the RF of the equipment. Second, the probability density function of the residual life is modeled, and then expected value of the residual life, i.e. MRL is evaluated.

1.2.6 Weibull Distribution

The *Weibull distribution* is one of the most commonly used distributions in reliability engineering because of the several shapes it attains for different values of its parameters. Hence it can model a great range of data and life characteristics. The most general expression of the Weibull PDF is given by the *three-parameter Weibull distribution* expression, or:

$$f(t)=\frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1}e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

Where, $f(t)\geq 0, t\geq \gamma,\ \beta>0,\ \eta>0, -\infty<\gamma<\infty$ and:

- $\eta =$ Scale parameter
- $\beta =$ Shape parameter or Slope.
- $\gamma =$ Location parameter

Usually, the location parameter is not used, and the value for this parameter is set to zero. When this is the case, the PDF expression reduces to that of the *two-parameter Weibull distribution*. In this thesis, we use the two-parameter Weibull distribution. Figure 1-8 shows the effect of different values of β value on the Weibull failure rate. As shown by the figure, Weibull distributions with $\beta<1$ have a failure rate that decreases with time, also known as *infantile* or *early-life* failures. Weibull distributions with β close to or equal to one have a fairly constant failure rate, indicative of *useful life* or *random* failures.

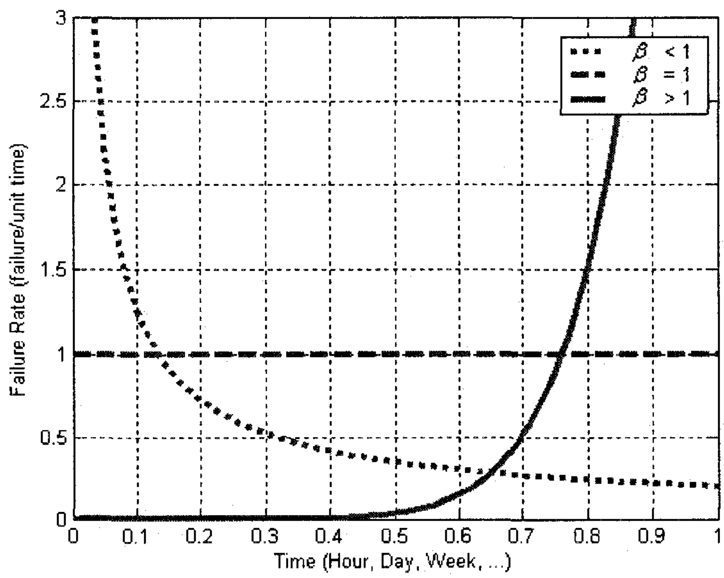

Figure 1-8: Weibull failure rate with $0<\beta<1, \beta=1$ **and** $\beta>1$

Weibull distributions with $\beta>1$ have a failure rate that increases with time, also known as *wear-out* failures. These include the three sections of the *bathtub curve*. Figure 1-9 exhibits an example of a bathtub curve.

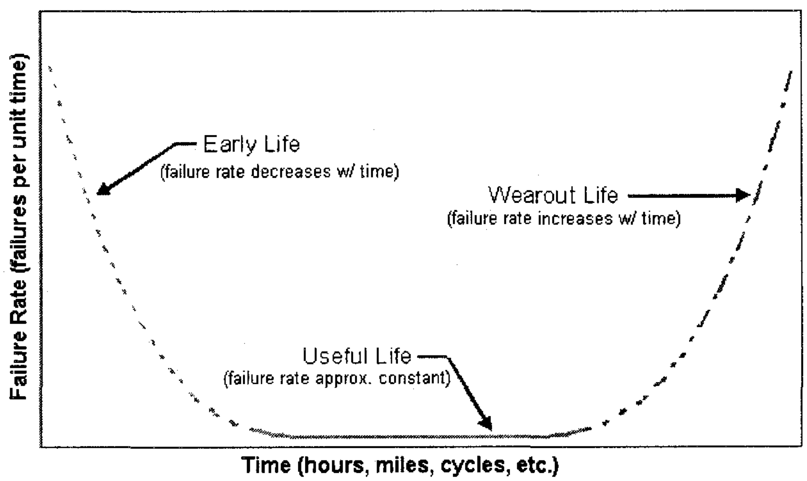

Figure 1-9: An example of a bathtub curve

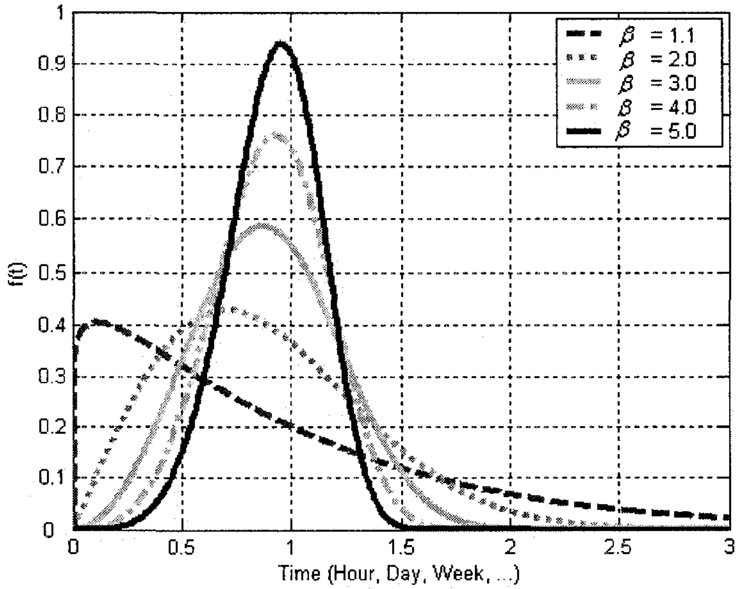

Figure 1-10: Weibull PDF plot with varying the value of β while $\beta > 1$

Increasing the value of β while $\beta > 1$ and holding η constant, stretches the PDF while moving the mass of PDF to the right. Since the area under a PDF curve is a constant value of one, the peak value of the PDF curve will also increase with the increase of β, i.e. the distribution gets stretched-in to the right and its height increases. If β is decreased, while η is constant, the distribution mass gets pushed in toward the left i.e. toward 0, and its height decreases.

Increasing the value of η while holding β constant, has the effect of stretching out the PDF. Since the area under a PDF curve is a constant value of one, the peak of the PDF curve will also decrease with the increase of η, i.e. the distribution gets stretched out to the right and its height decreases, while maintaining its shape and location. If η is decreased, while β is kept the same, the distribution gets pushed in toward the left i.e.

toward 0, and its height increases. η has the same unit as t, such as hours.

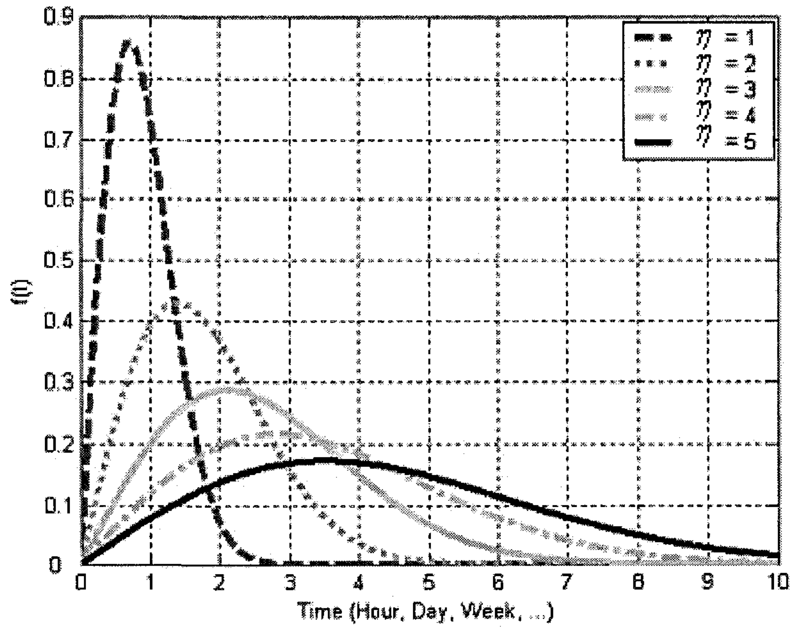

Figure 1-11: Weibull PDF plot with varying the value of η

1.2.7 Stochastic Processes

In many situations, we need to study the interaction of chance with time e.g. the number of failures in a certain period of time. To model this we need a family of random variables, all defined on the same probability space, $\{Z(t);t\geq 0\}$ where $Z(t)$ represents the degradation state of the equipment at time t. $\{Z(t);t\geq 0\}$ is a continuous time stochastic process or random process. For many studies, both theoretical and practical, we discretize the time and replace the continuous interval $[0,\infty)$ with the discrete set $\mathbb{N}$ or sometimes $\mathbb{N}\cup\{0\}$. We then have a discrete time stochastic process, $\{Z_k;k=0,1,2,\ldots\}$. $Z(t)$ and Z_t, and similarly $Z(k)$ and Z_k, are used interchangeably in the literature on this subject.

For stochastic processes, all the component random variables take values in a given set J, called the *state space*. Typically this will be a subset of $\mathbb{N}$, $\mathbb{N} \cup \{0\}$, $\mathbb{Z}$ or $\mathbb{R}$. When the random variable at time t takes an amount i from the state space, i.e. $Z(t) = i$, we say that the *state* of the equipment at time t is i.[Ross 1997]

1.2.8 Markov Process

In general, a stochastic process has the *Markov property,* if given the present state; the future state is conditionally independent of the past states. Many of the most popular stochastic processes used in both practical and theoretical work are supposed to have this property. If the states take on value in $\mathbb{R}$, we have a *Markov process*, if they take amount from $\mathbb{N}$ or $\mathbb{Z}$, we are dealing with a *Markov chain.* If we discretize time to be $\{0,1,2,\ldots\}$, we'll be working with *discrete time* Markov chains or process. For more details see [Ross 1997].

The probability that $Z(k+1) = j$ given that $Z(k) = i$ is called the *one-step transition probability* and we write:

$$p_{ij}^{k,k+1} = \Pr\left(Z(k+1) = j \mid Z(k) = i\right)$$

We have *stationary transition probabilities*, if $p_{ij}^{k,k+1}$ is the same for all k i.e. $p_{ij}^{k,k+1} = p_{ij}$ for any k, so the probability of going from i to j in one transition at any time is the same.

We can collect together all the transition probabilities into a matrix $P = \left[p_{ij} \right]$ called the *transition probability matrix* or sometimes *transition matrix* for short. [Ross 1997]

1.2.9 Dynamic Programming

An approach for solving dynamic optimization problems was pioneered by Richard Bellman in the late 1950s. This approach has been applied to problems in both continuous and discrete times. It is developed to solve sequential, or multi-stage, decision problems; hence, the name *dynamic programming.* [Bertsekas 1976] Dynamic Programming principle is given by Bellman [1957]:

> *"An optimal policy has the property that whatever the initial state and the initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."*

To show general idea of dynamic programming approach, consider the following diagram:

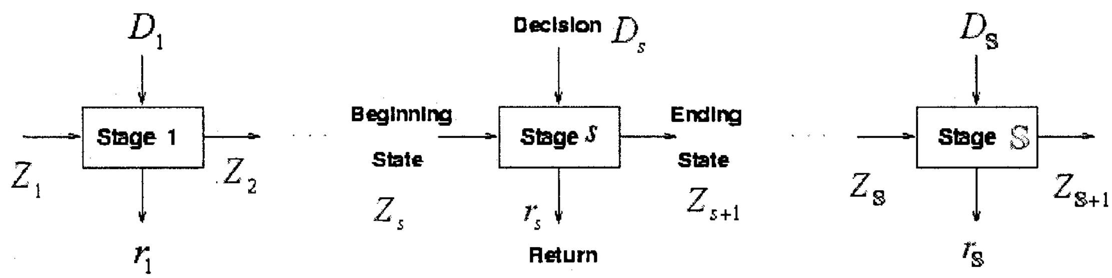

Figure 1-12: Schematic representation of dynamic programming idea

Knowing the state of the process at the beginning of a stage (periods), say stage s, we make a decision which results in a specific return benefit/cost and changes the state to the ending state at the end of the stage. The objective is to maximize/minimize the total return over all the stages. The conceptual framework is as follows:

- We observe a system (a piece of equipment) whose primary state, Z_1, is known.
- We make a decision (action), D_1, which makes the system to change its state to a state Z_2 by the transition function t_1 i.e. $Z_2 = t_1(Z_1, D_1)$. The transition's return is $r_1 = r_1(Z_1, D_1)$.
- We make a second decision, D_2, upon which the system changes its state to $Z_3 = t_2(Z_2, D_2)$. The stage's return is $r_2 = r_2(Z_2, D_2)$.

This process continues:

- After a number of iterations the system will be in state Z_s and we make the s-th decision, D_s, by which the system will change its state to $Z_{s+1} = t_s(Z_s, D_s)$ and the stage returns $r_s = r_s(Z_s, D_s)$.

There are finite number of possible states and decisions available. This will ultimately result in the fundamental deterministic recursion equations as follows: [Nemhauser 1966]

$$f_s(Z_s) = \max_{D_s} Q_s(Z_s, D_s), \quad s = 1, 2, \ldots, \mathbb{S}$$

$$Q_s(Z_s, D_s) = \begin{cases} r_{\mathbb{S}}(Z_{\mathbb{S}}, D_{\mathbb{S}}) & \text{for } s = \mathbb{S} \\ r_s(Z_s, D_s) + f_{s+1}(t_s(Z_s, D_s)) & \text{for } s = 1, \ldots, \mathbb{S} - 1 \end{cases}$$

1.2.10 Stochastic Dynamic Programming

An $\mathbb{S}$-stage stochastic system is similar to an $\mathbb{S}$-stage deterministic system except that at each stage there is a random variable, k_s, that affects the stage transformation and return.

$$r_s = r_s(Z_s, D_s, k_s)$$

$$Z_{s+1} = t_s(Z_s, D_s, k_s), \; s = 1, \ldots, \mathbb{S}$$

The random variables $k_1, \ldots, k_{\mathbb{S}}$ are assumed to be independently distributed with probability distributions $p_1(k_1), \ldots, p_{\mathbb{S}}(k_{\mathbb{S}})$ respectively. By defining the expected value return from stage s through $\mathbb{S}$, $\bar{f}_s$, and applying probability rules, (e.g. see [Nemhauser 1966]) fundamental stochastic recursion equations will be as follows:

$$\bar{f}_s(Z_s) = \max_{D_s} \sum_{k_s} p_s(k_s) Q_s(Z_s, D_s, k_s), \quad s = 1, 2, \ldots, \mathbb{S}$$

$$Q_s(Z_s, D_s, k_s) = \begin{cases} r_{\mathbb{S}}(Z_{\mathbb{S}}, D_{\mathbb{S}}, k_{\mathbb{S}}) & \text{for } s = \mathbb{S} \\ r_s(Z_s, D_s, k_s) + \bar{f}_{s+1}(t_s(Z_s, D_s, k_s)) & \text{for } s = 1, \ldots, \mathbb{S} - 1 \end{cases}$$

Introduction of uncertainty causes no increase in the amount of the state variables. Since Q_s is a function of only one random variable k_s, some difficulties of optimizing functions with several random variables have been eliminated. The optimal decision policy, resulting from stochastic multistage optimization, is itself stochastic, except for the first optimal decision, $D_1^*(Z_1)$. The rest, $D_2^*(Z_2),\dots,D_{\mathbb{S}}^*(Z_{\mathbb{S}})$, can not be expressed deterministically in terms of Z_s, until the stochastic elements that precede them are revealed. This is not a deficiency of dynamic programming, but a property of stochastic multistage system. For more details see [Bertsekas 1976].

1.2.11 Markovian Decision Processes

Markovian Decision Processes (MDP) represent a class of stochastic optimization problems. *MDP* is based on the Markov Process. It is assumed that there are a finite number of states at each stage, and a finite number of stages. Each state at stage $s=1,\dots,\mathbb{S}$ is denoted by i, $i=1,\dots,N$. As explained above, the probability of transition from state i at stage s to state j at stage $s+1$ is denoted by p_{ij} and is independent from s. These probabilities can be represented by transition matrix P:

$$P=\begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1N} \\ \vdots & & \vdots & & \vdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{iN} \\ \vdots & & \vdots & & \vdots \\ p_{N1} & \cdots & p_{Nj} & \cdots & p_{NN} \end{pmatrix}$$

The probability of being in state j at stage $s+1$, denoted by π_j^{s+1}, is determined from:

$$\pi_j^{s+1} = \sum_{i=1}^{N} \pi_i^s p_{ij}, j = 1,\ldots,N, s = 1,\ldots,\mathbb{S}$$

Corresponding to the transition matrix P, there is a return matrix R:

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1j} & \cdots & r_{1N} \\ \vdots & & \vdots & & \vdots \\ r_{i1} & \cdots & r_{ij} & \cdots & r_{iN} \\ \vdots & & \vdots & & \vdots \\ r_{N1} & \cdots & r_{Nj} & \cdots & r_{NN} \end{pmatrix}$$

which gives the return r_{ij} for transition from state i to state j. A decision variable $d_s = k$, $k = 1,\ldots,K$, designates the choice of the k-th transition matrix and k-th return matrix at s-th stage, in other words, if the system is in state i, $d_s = k$ means that the relevant transition probabilities and returns at stage s are the i-th row of the k-th transition and return matrices. The probability of transition is denoted by $p_{ij}(d_s)$ and the return by $r_{ij}(d_s)$.

This is simply a variation of the multistage stochastic optimization model given in the previous section where $p_{ij}(d_s) = t_s(Z_s, D_s, k_s)$ and $r_{ij}(d_s) = r_s(Z_s, D_s, k_s)$. The state variable, Z_s, is presented by i and decision variable, D_s, by d_s. The random variable k_s is hidden in the new notation. We denote the excepted total return from stage s through stage $\mathbb{S}$, starting in state i, by $\overline{f}_s(i)$. The stochastic recursion equations will be as follows:

$$\bar{f}_s(i) = \max_{d_s=1,\dots,K} \sum_{j=1}^{N} p_{ij}(d_s)\left[r_{ij}(d_s) + \bar{f}_{s+1}(j)\right], \quad s = 1,2,\dots,\mathbb{S}; i = 1,\dots,N$$

The term $\sum_{j=1}^{N} p_{ij}(d_s) r_{ij}(d_s)$ is the expected return from stage s. We denote it by $q_i(d_s)$, so:

$$\bar{f}_s(i) = \max_{d_s=1,\dots,K}\left[q_i(d_s) + \sum_{j=1}^{N} p_{ij}(d_s)\bar{f}_{s+1}(j)\right], \quad s = 1,2,\dots,\mathbb{S}; i = 1,\dots,N$$

These stochastic recursion equations remain to be solved by the usual computational methods used in DP. For more details see [Howard 1960].

1.2.12 Infinite stage MDP

Problems containing an infinite number of decisions arise in two fundamental different ways. First, is the case where there are a very large number of stages remaining and there is regularity in the stage returns and transformations in a way that we expect the optimal decision to be independent of the particular stage number. In the second case, the horizon is infinite, or the time periods (stage) are very small and negligible in comparison with the horizon. In the limit, as the size of the time periods approaches zero, we assume that decisions are made continuously. The former is called a *discrete infinite-stage process* and the latter is a *continuous infinite-stage decision process.* Under certain circumstances, the solution to the infinite-stage problem is stated as:

$$\begin{aligned} f(Z) &= \lim_{\mathbb{S}\to\infty} f_s(Z_s) = \lim_{\mathbb{S}\to\infty}\left\{\max_{D_s}\left[r_s(Z_s, D_s) + f_{s+1}\left(t_s(Z_s, D_s)\right)\right]\right\} \\ &= \max_{D_s}\left[r(Z,D) + f\left(t(Z,D)\right)\right] \end{aligned}$$

For more details see [Howard 1960].

1.2.13 Maximum Likelihood Estimation

Parameter estimation is a branch of statistics which deals with estimating the parameters of a mathematical model to fit to a set of historical/experimental data. The objective is to find the parameters of the system in a way that best describes the available data of the equipment. The Maximum Likelihood Estimation (MLE) is considered the most robust parameter estimation method with some exceptions. The method was pioneered by geneticist and statistician Sir R. A. Fisher between 1912 and 1922.

Assume that the parameters set Ω, of a probability distribution function has to be estimated. Also assume that the PDF of the distribution is denoted as f_Ω. We draw a sample $(x_1,...,x_n)$ of n values from this distribution, and then using f_Ω we compute the (multivariate) probability density associated with our observed data, $f_\Omega(x_1,...,x_n)$.

As a function of Ω with fixed $x_1,...,x_n$, the likelihood function is:

$$L(\Omega) = f_\Omega(x_1,...,x_n)$$

The method of maximum likelihood estimates Ω by finding the value of Ω that maximizes $L(\Omega)$. This is the maximum likelihood estimator (MLE) of Ω:

$$\hat{\Omega} = \arg \max_{\Omega} L(\Omega)$$

Usually, one assumes that the data drawn from a particular distribution are independent, identically distributed (i.i.d) with unknown parameters. This simplifies the problem because the likelihood can then be written as a product of n probability densities:

$$L(\Omega)=\prod_{i=1}^{n} f_{\Omega}(x_i)$$

and the Logarithm likelihood of n probability densities will be:

$$LL(\Omega)=\sum_{i=1}^{n}\log f_{\Omega}(x_i)$$

which its maximum can be found by various optimization methods [e.g see Box *et al.* 1969, Press *et al.* 2007].

1.3 Literature Review

Under an age replacement policy, a device is replaced or overhauled at failure or at a predetermined age. See [McCall 1965] and [Valdez-Flores and Feldman 1989] for some examples. Modeling the lifetime of a device whose failure depends upon the effects of time and usage has also received a great attention in the past decade. Scott *et al.* [2003] have considered a piece of equipment whose age is measured by two scales e.g. automobiles in the parallel scales of calendar time since purchase, and number of miles driven. Lawless *et al.* [1995], Murthy *et al.* [1995] and also Singpurwalla and Wilson [1998], among others, have considered this case. Some literature reviews on maintenance optimization in general can be found in [McCall 1965], [Valdez-Flores 1989] and [Dekker 1996].

A condition-based maintenance policy performs generally better than an age based one, e.g. see [Rao 1996] and [Gertsbakh 2000]. Scarf [1997] states that "the increase in the use of condition monitoring techniques within industry has been so extensive that it perhaps marks the beginning of a new era in maintenance management". Condition-based maintenance has been addressed in several papers, for some examples see [Christer and Wang 1992], [Scarf 1997], and [Wang 2000]. In most of the papers, either the critical threshold for replacement, or the inspection interval is a decision variable. In [Wang 2000] a renewal theory is used to calculate the cost criterion in terms of both these decision variables. Also Dieulle *et al.* [2003], using a Gamma process, developed a model which allows to investigate the joint influence of the critical threshold value and the choice of the inspection dates on the total cost of the maintained system.

In general, many existing models of CBM policies are based on a continuous-time discrete-state Markovian deterioration process and focus on determining the states in which the equipment should be replaced to get the minimum expected cost. Mostly the inspection period and/or the critical states are optimized by applying the Markovian decision process. Coolen and Dekker [1995] optimized the interval between successive condition measurements (inspections), where measurements are expensive and cannot be made continuously. Lam and Yeh [1994] determined an optimal inspection & replacement policy such that the mean long-run average cost is minimized. For more instances on this approach see: [Mine and Kawai 1975], [Ohnishi *et al.* 1986], [Tijms and Schouten 1984], and [Wijnmalen and Hontelez 1992].

A second group of researches concern with continuous state processes. Hontelez *et al.* [1996] considered a continuous-time, continuous-state deterioration process. In the model, it was assumed that the relationship between the state of the equipment and its age is more or less known and, additionally, that the state can be observed. Park [1998, 1998a] considered a piece of equipment failed when it wears beyond a certain breakdown threshold and the wear accumulates continuously, but the wear is difficult to monitor continuously. Chelbi and Ait-Kadi [1999] addressed the problem of generating optimal inspection strategies for equipment with failure which is obvious only through inspection. A situation where it is possible to identify one parameter well correlated with the equipment deterioration state is considered. Generally, in this group, the aim is either to calculate the optimal inspection period while the critical threshold is given or to find the optimal threshold when the inspection period is prefixed. For more instances see [Hopp and Kuo 1998], [Christer and Wang 1992, 1995], [Barbera *et al.* 1996], [Wang 2000], [Wang and Christer 2000], [Christer *et al.* 1997] and [Aven 1996].

The more the information on the equipment is close to its real degradation state, the more the policy is efficient [Barros *et al.* 2002b]. The ideal case, widely studied in the literature, is when the information is perfect (direct monitoring), i.e. the state of the equipment, like degradation level, is perfectly known, e.g. see [Cho and Parlar 1991]. Christer and Wang [1992, 1995] considered particular problems of directly monitored systems. Grall *et al.* [2002] found the optimum threshold and inspection schedule jointly for a piece of equipment releasing perfect information.

Condition based maintenance decisions in practice are largely based upon measurements of the state of the equipment obtained at monitoring moments. These measures can likely contain noises, and in general, may not tell directly the exact condition of the monitored equipment. They are, however, assumed to be stochastically related with the actual state. This type of condition monitoring is called indirect condition monitoring which provides imperfect information or partial information in contrast to direct monitoring which provides perfect information or complete information. In many realistic situations, the observation is imperfect. For instance Rosqvist [2002] formulated a stopping time model, using experts' judgments on the residual operating time of the equipment. The judgment is based on an indication of the equipment's state which releases imperfect information about equipment's state. The objective is to maximize expected utility. Experts are asked to provide percentage information on residual lifetime of the equipment, given the indicator of the equipment's state.

There are two general approaches regarding the use of observations' information in Condition Based Maintenances. The first approach, considers just the current information obtained from the observation. For instance, Christer *et al.* [1997] presented a case study of furnace erosion prediction and replacement. A state space model is used to predict the erosion condition of the inductors in an induction furnace in which a measure of the conductance ratio is used to indirectly assess the relative state of the inductors, and to guide replacement decisions. Campodonico and Singpurwalla [1994] used a Bayesian approach considering the vibration of the equipment as the covariant of the equipment degradation state. Zilla [1993] considered the number of the defective items as the observable covariant of the equipment degradation state. The history of the process at each period contains the number of the defective items for each of the previous periods, and the decision made in each previous periods. Christer and Wang [1995] addressed the problem of condition monitoring of a component which has available wear which is considered as a measure of the state. Supposing that the past measurements of the wear are available up to the present, and the component is still working, the decision problem is to choose an appropriate time for the next inspection based upon the condition information obtained to date.

The Proportional Hazards Model (PHM) was developed in order to take into account the effects of equipment state influencing the times-to-failure of that equipment. The model has been broadly used in the biomedical field and recently there has been an increasing application in reliability engineering. Kumar and Westburg [1997] used PHM to identify the importance of monitored variables and estimate the reliability function using the values of monitored variables. Then, the reliability function is used to estimate the optimum maintenance time interval or threshold values of monitored variables for the equipment. Jardine *et al.* [1987], and Makis and Jardine [1992] used the Proportional Hazard Model to model deterioration behavior of the equipment in condition based replacement problems and to find the optimal replacement policy to minimize the expected average cost in long-run. Ghasemi *et al.,* 2007 addressed the optimal replacement problem of equipments with indirect observations while its unobservable degradation state follows a Markov model. They used a Hidden Markov Model to model the degradation state and assumed that the equipment's degradation state follows the PHM.

In presence of condition monitoring systems, practitioners are not only interested in the optimal replacement solution but also in optimal inspection interval. In what follows we consider the literature review more specifically related to optimal inspection period. Lam and Yen [1994] investigated the maintenance policy for a system whose exact degree of degradation is known through inspections. Their objective was to find the optimal replacement criteria and inspection period that minimizes the long-run average cost of the maintenance plan. They assumed that there is a fixed inspection cost M, that the mean inspection time is q and the cost rate per unit of time when the system is under inspection is m. This results in an average inspection cost equivalent to $mq+M$. Dynamic Programming is used to solve the problem. Christer and Wang [1995] considered a system with perfect information where inspections reveal the system's exact degradation state. The objective is to determine the next inspection schedule, based on the inspections' observations up to date. The next inspection point is selected in a manner that minimizes the average maintenance cost per unit time between current and the next inspection. A constant predetermined threshold on the degradation state which determines the failure is considered and there is a cost related to the inspections. In this model, the time between the inspections is not constant and next inspection point is always determined at the current inspection point by taking into account the available information from condition monitoring system. Possible actions are "inspection and replacement" or "inspection and no replacement". A closed form solution to calculate the cost per unit time is given. By minimizing the cost, the optimal inspection date is obtained.

Hontelez *et al.* [1996] considered a system revealing perfect information with $N+1$ possible degradation states, where states N and $N+1$ are failure states. State N is detected through inspection but state $N+1$ is an obvious failure and is detected as soon as it happens. There is a cost associated with the inspections. Applying dynamic programming, a control limit rule such as $\Pi = [\pi; \rho_1, \rho_2, ..., \rho_{\pi-1}]$ is obtained. It means that at an inspection point; replace if and only if the system is in degradation state $i \geq \pi$ OR: perform an inspection after ρ_i, if the degradation state is $i < \pi$. In this approach the inspections will not take place at every inspection period unless there is an evident need to do that.

Chelbi and Ait-Kadi [1999] considered an optimal inspection time with a hidden failure being detected through inspection. A pre-defined threshold on the system's degradation state is set to identify the failure and associated costs are considered for the inspections, repairs and replacements. The average long-run cost of the maintenance plan $E(AC) = E(C)/E(T)$, is minimized. $E(C)$ and $E(T)$ present the renewal period's expected cost and the renewal period's expected length, respectively. The renewal period is the time between two consecutive replacements, whether due to a failure or to a preventive replacement [Cox, 1962]. By minimizing the long-run expected cost over renewal period, the optimal inspection period is found.

Grall *et al.* [2002] modeled a CBM policy where both the replacement threshold and the inspection schedule are decision variables. It is allowed to have irregular inspection periods, i.e., the next inspection date is dynamically updated on the basis of the present system degradation state revealed by the current inspection. $N(>1)$ threshold values $0<\xi_1,\ldots,\xi_N<L$, are set in the system degradation state range. L is the predefined degradation threshold assumed for the failure and ξ_N is the replacement threshold. At any inspection point t_k, where the observation value (degradation state) is Z_k, if $\xi_l \le Z_k < \xi_{l+1}$ for some $0 \le l < N$, then the next inspection will be after $N-l$ period(s). Otherwise if $\xi_N \le Z_k < L$ then a preventive replacement is performed while $Z_k \ge L$ results in a failure replacement. The long-run expected cost per unit of time is minimized and the optimal value of the decision variables which are the number of thresholds N, and the different thresholds' values $\xi_i, i=1,\ldots,N$ are found.

The information obtained from condition monitoring is used to establish a diagnosis of the equipment's condition and a prognosis for future performance. Two measures of this future performance are the Reliability Function (RF), and the Mean Residual Life (MRL). Many researchers have studied the mathematical structure of the MRL based on reliability analysis without considering information concerning the actual use and the state of the equipment. Tang *et al.* [1999] considered the residual life as a random variable and studied its asymptotic behaviour when the reliability function is represented by various discrete and continuous distribution functions. Lim and Park [1995] studied the monotonic behaviour of the residual life. They tested the null hypothesis that the residual life is not monotone, against the alternative hypothesis that it is indeed monotone. Siddiqui and Caglar [1994] treated the residual life as a random variable and gave a representation of its distribution function. When the distribution is Gamma or Weibull, the authors calculated the mean and the variance of the variable. Bradley and Gupta [2003] also studied the asymptotic behaviour of the residual life.

Researchers that consider the presence of condition monitoring information use two main approaches to calculate the MRL; *recursive filtering* and PHM [Jardine *et al.*, 2006]. Recursive filtering is an approach in signal processing that extracts information (MRL) based on available signals (indicators), and previously extracted information [Byrne, 2005]. Wang and Christer [2000], Wang and Zhang [2005, 2008] and Wang [2002], among others, determined the MRL by applying a recursive filtering model. The MRL given the condition monitoring history up to date is obtained. The recursive filtering technique includes the entire observation history.

Wang and Christer [2000] assume that the observed condition monitoring indicator is a function of the underlying residual life and not vice versa, and use the indicator as the covariate. This assumption may not be realistic in many cases. For instance, the wear of a rotating shaft (which is reflected in oil particles as the observable indicator) affects its residual life, not vice versa. They used a recursive filter in order to calculate the MRL, and added to the existing MRL models the possibility of including all past information.

[Wang and Christer, 2000] Wang and Zhang [2005] introduced a new methodology that uses the difference between two consecutive observed indicators as the covariate and uses recursive filtering to determine the MRL. They define the MRL=$e(t, Z(t))$, as the expected time interval between the last inspection, when the most recent information was gathered, and the expected time of failure, given that no maintenance action is taken in this interval. $e(t, Z(t))$ is thus equal to $E((T - t) \mid T > t,\ Z(t))$, where $Z(t)$ is the covariate (the difference of two consecutive readings of the indicator) at time t. Wang and Zhang [2008] modeled the MRL of the asset by considering the expert's judgment based on the equipment's observed indicator. In this case, the judgment is assumed to be a function of the residual life, which may include some noise.

The PHM may be more suitable in many cases, like oil analysis, since it assumes that the failure rate, and so the MRL, is a function of the degradation state or the observed indicator, which is representing the degradation state. But the PHM's drawback is that, it uses only the latest information of the condition monitoring system. Kumar and Westberg [1996] calculated the MRL using PHM when only the most recent information is available. Maguluri and Zhang [1994] are inspired by the PHM and calculate the proportional MRL by using the equation $e(t \mid Z) = \exp(-\beta' Z) e_0(t)$, where $e_0(t) = E(T - \tau \mid T > \tau)$ is the MRL calculated without including the covariate. Z is the vector of indicators used as the covariates in the model and β' is the vector of the covariates' coefficients. Sen [2004] calculates the conditional MRL given by

$e(t \mid z) = \int_0^{\infty} e^{-\beta' z(\Lambda_0(t+\mu)-\Lambda_0(t))} d\mu$, where $\Lambda_0(t) = \int_0^t \lambda_0(u) du$, $\lambda_0(t)$ is the hazard function, and $\lambda(t \mid z) = \lambda_0(t) e^{\beta' z}$ is the proportional hazard function. Banjevic and Jardine [2006] calculated the joint distribution of time-to-failure and the diagnostic covariate $z(t)$ at time t, and the probability of transition between states $L_{ij}(\tau, t) = P(T > t, z(t) = j \mid T > \tau, z(\tau) = i)$, $\tau \leq t$. $z(t)$ is the descritized observed information from the indicator into a new state space $\{0, 1, \ldots, N\}$, which represents the degradation state of the equipment (some-to-one relationship). The conditional reliability is thus given by the equation $R(t \mid \tau, z(\tau)) = \sum_j L_{ij}(\tau, t)$, $\tau \leq t$, and the MRL is $e(t, z(\tau)) = \int_t^{\infty} R(\tau \mid t, z(t)) d\tau$. All these models assumed that the information gathered from the indicator likely reveals the equipment's exact state, or used the collected information directly as a diagnostic covariate that affects the failure rate. Moreover, some of them include only the most recently collected information.

The diagnosis and prognosis processes are based on mathematical models which are in turn constructed using several parameters. In order to apply any diagnosis and/or prognosis method to a real world problem, the parameters have to be estimated considering the historical behaviour of the equipment. Parameter estimation of condition monitoring models using PHM has been considered by the researchers in two categories: perfect and imperfect observations. Considering perfect observation systems, Jardine *et* *al.* [1987] incorporated indicators (diagnostic covariates) affecting the equipment's time-to-failure into a fully parametric Weibull PHM and estimates the model's parameters based on Maximum Likelihood Estimation (MLE). Elsayed *et al.* [1990] developed PHM to estimate thin-oxide dielectric reliability by applying the partial likelihood in the analysis. Banjevic *et al.* [2001] estimated the parameters of a PHM used in analysis of equipment following Markovian degradation. A parametric PHM with Weibull baseline hazard function was considered and its parameters were estimated by MLE method. The method of MLE is also used to estimate the transition probabilities of the Markovian process.

Cox [1972] introduced the conditional likelihood, later called partial likelihood [Cox, 1975], to estimate the parameters of a semi-parametric PHM, supposing that the base line hazard function in the PHM, $\lambda_0(.)$, is arbitrary and the covariates are time-dependent. It was assumed that the exponential function incorporated the effect of the covariates into the equipment's time-to-failure. For r^{th} failure time $t_{(r)}$, the probability of observing the failure on the equipment that has actually failed given the *risk set* of $R\left(t_{(r)}\right)$ is $\exp\left(\gamma Z_{(r)}\right) \Big/ \sum_{l \in R\left(t_{(r)}\right)} \exp\left(\gamma Z_{(l)}\right)$, where the risk set $R(t)$ is the set of all equipments that have not yet failed until time t, Z represents the diagnostic covariate of the equipment and γ is the coefficient which represents a weight factor for the covariates. r is the index counter of the sample data of a set of n independently observed histories. Consequently, the log partial likelihood function is:

$$\log\left(L(\gamma)\right)=\sum_{r=1}^{n}Z_{(i)}\gamma-\sum_{r=1}^{n}\log\left[\sum_{l\in R\left(t_{(r)}\right)}\exp\left(\gamma Z_{(l)}\right)\right].$$

Several methods are suggested to estimate the base line hazard function $\lambda_0(t)$. One of them is to assume that $\lambda_0(t)$ is zero except at failure points $t_{(r)}$. The estimator of $\lambda_0(t)$ is thus given by $\hat{\lambda}_0\left(t_{(r)}\right)=\left\{\sum_{l\in R\left(t_{(r)}\right)}\exp\left(\hat{\gamma}Z_{(l)}\right)\right\}^{-1}$, where $\hat{\gamma}$ is the MLE of γ [Kay, 1984].

If ties exist and the number of ties is small in comparison to the number of available information, then the log partial likelihood is calculated by the following equation $\log\left(L(\gamma)\right)=\sum_{r=1}^{n}S_{(r)}\gamma-\sum_{r=1}^{n}\log\left[\sum_{l\in R\left(t_{(r)}\right)}\exp\left(Z_{(l)}\gamma\right)\right]^{d_r}$, where d_r denotes the number of ties for failure time $t_{(r)}$ and $S_{(r)}$ is the sum of the failed items' covariate at time $t_{(r)}$ [Kalbfleisch and Prentice, 1980]. Also the estimator of $\lambda_0(t)$ is given by:

$$\hat{\lambda}_0\left(t_{(r)}\right)=d_r\Big/\left\{\sum_{l\in R\left(t_{(r)}\right)}\exp\left(\hat{\gamma}z_l\right)\right\}.$$

Cox [1972] proposes that the covariates of the PHM can be allowed to be time-dependent, that is to say; their values may vary in equipment's lifetime. In this case the equation $h\left(t,Z(t)\right)=\lambda_0(t)\exp\left(\gamma Z(t)\right)$ indicates the PHM with time-dependent covariates, where γ stands for the covariate's coefficients and $Z(t)$ is the time-dependent covariate at time t.

In a semi-parametric PHM where there is no assumption about the form of $\lambda_0(t)$, γ is estimated by maximizing partial likelihood that does not depend on $\lambda_0(t)$. In a parametric function of a certain form, such as Weibull, the model parameters can be estimated by full likelihood [Lin *et al.*, 2005].

For calculating the full likelihood, the complete covariate realization $\{Z_r(t), 0 \le t < T_r\}$, where T_r is the failure or the censoring time of the r-th equipment, should be known. Practically, it is not possible to have the covariate recorded continuously. Instead, it is known in discrete times of observations. An approach to deal with this problem is to assume that the covariate $Z_r(t)$ is constant between the observations. For the time-dependent PHM, the partial likelihood function that estimates the parameter γ is given by [Kalbfleisch and Prentice, 1980] as follows:

$$L(\gamma) = \sum_{r=1}^{n} \gamma Z_r\left(t_{(r)}\right) - \sum_{r=1}^{n} \log\left[\sum_{l \in R\left(t_{(r)}\right)} \exp\left(\gamma Z_l\left(t_{(r)}\right)\right)\right]$$

It is assumed that the hazard at time t depends only on the current covariate vector. The introduced partial likelihood has almost the same form as time-independent covariates, except that the covariates are time-dependent now.

Banjevic *et al.* [2001] showed that the likelihood of the set of n independently observed histories $\left(T_r, C_r, \left(Z_r(s); s \le T_r\right)\right)$, $r = 1, 2, \ldots, n$ is:

$$L(\theta) \propto \prod_{r:C_r=1} h\left(T_r, Z_r(T_r)\right) \prod_j S\left(T_j, Z_j\right)$$

where T_r is the failure or censoring time of the r-th experiment, C_r is the censoring indicator that indicates whether the equipment has failed or has been censored. It takes the following values $C_r = \begin{cases} 0 & \text{Censored} \\ 1 & \text{Failed} \end{cases}$, and $S(t;Z) = S\left(t; Z(s), s \le t\right) = \exp\left\{-\int_0^t h\left(\tau, Z(\tau)\right) d\tau\right\}$, j is the risk set at T_r. If the value of Z at the failure or the censoring moment is not known, which might often be the case, the value of the latest covariate is used.

Estimation of the transition matrix of Markov chain can also be obtained by MLE. By considering constant observation times, the estimator of transition probability, $p_{ij}(k) = \Pr\left(Z_{k+1} = j \mid T > (k+1)\Delta, Z_k = i\right)$ is given by $\hat{p}_{ij}(k) = \dfrac{n_{ij}(k)}{\sum_j n_{ij}(k)}$, where $n_{ij}(k)$ is the amount of one-step transitions from state i to state j at $k-$th observation point, $k = 1, 2, \ldots$ [Basawa and Rao, 1980].

We are considering the imperfect observation system to illustrate the hidden degradation process in the model. The model will consist of two separate stochastic processes: a Hidden Markov Model with finite state space describing the state transition and an observation process. When the observations are not perfect, some researchers used Expectation Maximization (EM) to estimate the parameters. Fernandez-Gaucherand [1993] considered a finite state Markov Chain for equipment with partial information. He assumed that a maintenance action resets the state of the equipment to a known value, and consequently, its future evolution becomes independent of the past. It is shown that the parameter estimator converges to the true parameter.

When the observations are imperfect, the EM method is used to avoid modeling the Probability Density Function (PDF) of the observed imperfect information. Lin *et al.* [2003, 2004] considered a CBM problem while the equipment state is partially observable and the failure is obvious. Parameters are estimated using a recursive EM algorithm. Adjengue and Yacout [2005] used an EM algorithm for estimating the parameters of CBM with imperfect information.

In the next section, we represent the problem statement and the three main objectives of this thesis.

1.4 Problem Statement

In this thesis we concentrate on a Condition Based Maintenance with indirect observations for a piece of equipment which is operating continuously. We consider the time-dependent PHM proposed by [Cox, 1972], and we assume that the condition monitoring is indirect. Instead, an indirect indicator's value, θ, of the underlying degradation state is available at each observation moment. Observations are collected at a constant (or a near constant) interval Δ. In this study, Z represents the degradation state of the equipment which will be used as the diagnostic covariate in the PHM, and θ is a value from the set of all the possible indicator values $\Theta = \{1,...,M\}$. The whole set of the indicator values is discretized into a finite set of M possible values. This assumption does not limit the scope of this work since its relaxation only entails the replacement of the observation probability matrix Q by a continuous probability distribution such as the normal distribution as shown in Figure 1-4. The equipment condition is described as follows:

- The equipment has a finite and known number of degradation states N. $J = \{1,...,N\}$ is the set of all possible degradation states;
- The degradation state transition follows a Markov Chain with unobservable states and is modeled by a Hidden Markov Model (HMM). The transition matrix is $P = \left[p_{ij}\right]$, where p_{ij} is the probability of going from state i to state j, $i, j \in J$ during one observation interval, knowing that the equipment does not fail before the end of the interval;
- The value of the indicator is stochastically related to the equipment's state through the observation probability matrix $Q = \left[q_{j\theta}\right]$, $j \in J$, $\theta \in \Theta$. $q_{j\theta}$ is the probability of getting the indicator value θ, while the equipment is in state j;
- The indicator is collected periodically at a fixed (or a near fixed) interval Δ;

- Failure is not a degradation state. It is a non-working condition of the equipment that can happen at any time and while in any degradation state, and is known immediately (obvious failure).

Figure 1-13 depicts the process of degradation and the transition from one degradation state to another, and from the degradation state to failure. The circles represent the states. State 1 is the best state (new or as new equipment). State N is the worst state, but it is not failure and the equipment is still working and fulfilling part of its mission. It should be noted that failure can happen at any time and while the equipment is in any degradation state. T is a random variable showing the time-to-failure and $1-r_i$ is the probability of failure before the end of the interval, while the equipment state is i. $r_i = R(k,i,\Delta)$, which is calculated for each observation moment k, is the conditional reliability of the equipment for a period of time Δ, while the equipment state is i. The equation of $R(k,i,\Delta)$ will be presented later in this thesis.

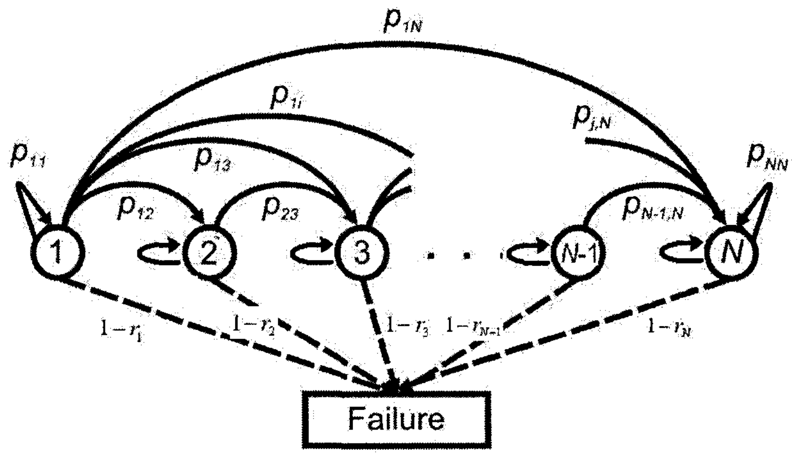

Figure 1-13 : The process of degradation and failure

At time $s=0$, the equipment is always in state 1, which indicates that the equipment is in its best state. At fixed interval Δ, i.e. at $s=\Delta,2\Delta,...$ an indicator of the equipment's state is observed. The indicator's value θ, is collected with a probability of $q_{j\theta}$ when the equipment is in degradation state j, $j \in J, \theta \in \Theta$. The transition matrix $P=\left[p_{ij}\right]$ $i,j \in J$ is assumed to be an upper triangular matrix, i.e. $p_{ij}=0$ for $j<i$, meaning that the degradation state cannot improve by itself, which is the case in most practical problems.

Ghasemi *et al.,* [2007] has addressed the optimal replacement policy of this problem while considering costly failure replacements and *non-costly* pre-fixed inspections. They have found the minimum long-run average cost of the replacement and the optimal replacement criteria that guarantees this minimum cost. In this thesis we will address three other objectives concerning the same problem assumptions.

The first objective considers the problem of optimum inspection period. The unrealistic assumption of non-costly inspection will be relaxed and corresponding optimal replacement policy and long-run average cost will be found. In the CBM modeling, if the inspections are non-costly, the optimal inspection period is zero i.e. the best choice is to monitor and analyze the system continuously. That's because the higher frequency of inspections will provide more frequent information about the system's degradation state with no extra cost. Consequently, this will reduce the likelihood of performing unnecessary preventive replacements, hence, will result in a more cost effective maintenance system. When there is considerable cost for collecting and analyzing the observations, an optimal inspection period that minimizes the *maintenance and inspection cost* should be applied. In reality, in many cases, inspections require personnel and equipment, and sometimes it is necessary to stop or suspend the operations when performing the inspections [Lam and Yen, 1994]. Also some tests for analysis and extraction of useful information may be needed which may be destructive; therefore some costs are associated to the collection and analysis of the observations. The total optimal long-run average cost of the maintenance and inspections leads to selection of the optimal inspection period between several possible inspection periods.

The second objective of this thesis deals with evaluating and modeling the Reliability Function (RF) and Mean Residual Life (MRL) of the equipment described earlier. When a condition monitoring system is used, the information obtained from the monitoring system is used to establish a diagnosis of the equipment's condition and a prognosis of the future performance. Two measures of the future performance are the RF and the

MRL. We will also show how these two measures are helpful for the reliability and maintenance practitioners and will compare the output of these utilities with the results of optimal replacement policy introduced by Ghasemi *et al.* [2007].

The third objective of this thesis is to estimate all the parameters of the introduced condition monitoring system. In all previous objectives, it is assumed that all the parameters of the system are known, but in order to apply any of these methods on a real world problem, the parameters have to be estimated by considering the historical behaviour of the equipment. Algorithms for data with and without censoring will be presented. In this work, while the observations are indirect, we will directly model the PDF of the observed information (indicator) and will use the MLE method to estimate the model parameters. Also the convergence behaviour and the robustness of the introduced methods will be studied by simulation.

The rest of this thesis is organized as follows: Chapter 2 is a book chapter produced based on the results of the first objective of the thesis and published by the American Institute of Physics. This work was originally presented as a conference paper at The International Conference on Systems Engineering and Engineering Management 2007, and received the best paper award of the conference. Later, the authors were invited to submit an extended version of the article to be considered for publication as the book chapter. Chapter 3 is the second revision of an article submitted to the IEEE Transactions on Reliability (TR2008-056) which represents the second objective of the thesis. Chapter 4 is an article on the parameter estimation problem, i.e. third objective of this thesis, which has also been submitted to the IEEE Transactions on Reliability (TR2008-240). Finally, chapter 5 is the summary, conclusion and future research discussions.

CHAPTER 2 :

OPTIMAL INSPECTION PERIOD AND REPLACEMENT POLICY FOR CBM WITH IMPERFECT INFORMATION USING PHM

Alireza Ghasemi, Soumay Yacout, M-Salah Ouali

CP1007, Current Themes in Engineering Technologies, Edited by: Sio-long Ao, Mahyar A. Amouzegar, Su-Shing Chen, AMERICAN INSTITUTE of PHYSICS, ISBN: 978-0-7354-0526-4, ISSN: 0094-243X

2.1 Abstract

This paper introduces a model to find the optimal inspection period for Condition Based Maintenance (CBM) of a system when the information obtained from the gathered data on the system does not reveal the system's exact degradation state and the collection of data is costly. The proposed model uses the Proportional Hazards Model (PHM) introduced by [Cox 1972] to model the failure rate of the system. The PHM takes into consideration the system's degradation state as well as its age. Since the acquired information is imperfect, the degradation state of the system is not precisely known. Bayes' rule is used to estimate the probability of being in any of the possible states. The system's degradation process follows a Hidden Markov Model (HMM). By using dynamic programming, the system's optimal replacement policy and its total long run average operating maintenance cost are found. Based on the long run average cost, the optimal inspection interval and the corresponding replacement criterion are specified. A numerical example shows the behaviour of the CBM model when the inspection is costly, and finds the optimal inspection period and maintenance cost.

Keywords: Condition Based Maintenance (CBM), Imperfect Information, Proportional Hazard Model (PHM), Hidden Markov Model (HMM), Costly Inspections.

PACS: 89.20.Bb, 45.10.Db, 46.15.Cc, 87.55.de

2.2 Introduction

For a system subject to a Condition Based Maintenance (CBM) program, inspections are performed to obtain information (observation) about the degradation state of the system. In this paper, the information acquired during the inspections does not reveal the exact degradation state of the system but represents some data which is stochastically related to the system's degradation state [Maillart, 2004, Ohnishi *et al.*, 1986]. This data is used to calculate the probability of being in a certain degradation state. The hidden degradation state of the system is modeled by a Markov Chain. In CBM studies, several models have been used to take into account the system's degradation state. One of these models is the Proportional Hazards Model (PHM), introduced by [Cox 1972], which has an increasing application in the CBM recently [Lin *et al.* 2005, Banjevic and Jardin 2001]. According to the PHM, the system's failure rate (also called hazard rate) is calculated based on its age as well as its degradation state. In this paper the PHM is used to calculate the optimal replacement policy and long-run average cost for a system with imperfect information. The unrealistic assumption of non-costly inspection is relaxed and corresponding optimal replacement policy and long-run average cost are found.

In the CBM modeling, if the inspections are done at no cost, the optimal inspection period is zero i.e. the best choice is to monitor and analyze the system continuously. That's because the higher frequency of inspections will provide more frequent information about the system's degradation state with no extra cost. Consequently, this will reduce the likelihood of performing unnecessary preventive replacements, hence, will result in a more cost effective maintenance system. When there is considerable cost for collecting and analyzing the inspections' observations, an optimal inspection period that minimizes the *total maintenance cost* which includes the inspections cost should be applied. In reality, in many cases, inspections require personnel and equipment, and sometimes it is necessary to stop or suspend the operations when performing the inspections [Lam and Yen, 1994]. Also some tests for analysis and extraction of useful information may be needed which may be destructive; therefore some costs are associated to the collection and analysis of the observations. The total optimal long-run average cost of the maintenance plan with costly inspections leads to comparison and selection of the optimal inspection period between several possible inspection periods.

The rest of this paper consists of following sections; in imminent section a brief literature review of the principle models in replacement optimization is presented. Next section deals with the assumptions and the details of the proposed model and the optimal solution. Then a numerical example is presented and finally the conclusion and the areas of further researches are presented in last section.

2.3 Literature Review

Lam and Yen [1994] investigated the maintenance policy for a system whose exact degree of degradation is known through inspections. Their objective was to find the optimal replacement criteria and inspection period that minimizes the long-run average cost of the maintenance plan. They assumed that there is a fixed inspection cost M, that the mean inspection time is q and the cost rate per unit of time when the system is under inspection is m. This results in an average inspection cost equivalent to $mq + M$. Dynamic Programming is used to solve the problem.

Christer and Wang [1995] considered a system with perfect information where inspections reveal the system's exact degradation state. The objective is to determine the next inspection schedule, based on the inspections' observations up to date. The next inspection point is selected in a manner that minimizes the average maintenance cost per unit time between current and the next inspection. A constant predetermined threshold on the degradation which determines the failure is considered and there is a cost related to the inspections. In this model, the time between the inspections is not constant and the next inspection point is always determined at the current inspection point by taking into account the available information from condition monitoring system. Possible actions are "inspection and replacement" or "inspection and no replacement". A closed form solution to calculate the cost per unit time is given. By minimizing the cost, the optimal inspection date is obtained.

Hontelez *et al.* [1996] considered a system revealing perfect information with $N+1$ possible degradation states, where states N and $N+1$ are failure states. State N is detected through inspection but state $N+1$ is an obvious failure and is detected as soon as it happens. There is a cost associated with the inspections. Applying dynamic programming, a control limit rule such as $\Pi = [\pi; \rho_1, \rho_2, ..., \rho_{\pi-1}]$ is obtained. It means that at an inspection point; replace if and only if the system is in degradation state $i \geq \pi$

OR: perform an inspection after ρ_i periods if the degradation state is $i < \pi$. In this approach the inspections will not take place at every inspection period unless there is an evident need to do that.

Chelbi and Ait-Kadi [1999] considered an optimal inspection time with a hidden failure being detected through inspection. A pre-defined threshold on the system's degradation state is set to identify the failure and associated costs are considered for the inspections, repairs and replacements. The average long-run cost of the maintenance plan $E(AC) = E(C)/E(T)$, is minimized. $E(C)$ and $E(T)$ represent the renewal period's expected cost and the renewal period's expected length, respectively. The renewal period is the time between two consecutive replacements, whether due to a failure or to a preventive replacement [Cox, 1962]. By minimizing the long-run expected cost over renewal period, the optimal inspection period is found.

Grall *et al.* [2002] modeled a CBM policy where both the replacement threshold and the inspection schedule are decision variables. It is allowed to have irregular inspection periods, i.e., the next inspection date is dynamically updated on the basis of the present system degradation state revealed by the current inspection. $N(>1)$ threshold values $0 < \xi_1, \ldots, \xi_N < L$, are set in the system degradation state range. L is the predefined degradation threshold assumed for the failure and ξ_N is the replacement threshold. At any inspection point t_k, where the observation value (degradation state) is Z_k, if $\xi_l \le Z_k < \xi_{l+1}$ for some $0 \le l < N$, then the next inspection will be after $N - l$ period(s).

Otherwise if $\xi_N \le Z_k < L$ then a preventive replacement is performed while $Z_k \ge L$ results in a failure replacement. The long-run expected cost per unit of time is minimized and the optimal value of the decision variables which are the number of thresholds N, and the different thresholds' values $\xi_i, i = 1,...,N$ are found.

[Ghasemi *et al.* 2007] considered a CBM based on imperfect information when there is no considerable cost associated with the inspections

2.4 Problem Formulation

This paper presents a deteriorating system subject to random failure. While the degradation state of the system can be continuous we have discretized the degradation state set. The degradation state of the system is illustrated by a finite set of non-negative integers, i.e. by state the space $S = \{1, 2,, ..., N\}$. The circles represent the states. State 1 indicates the best possible state for the system which means that the system is new or like new. The degradation state process $\{Z(t) = 1, 2, ..., N\}$, is a discrete time homogeneous Markov chain with N unobservable states. All N degradation states are working states and do not include the failure state which is a non-working state. If a failure happens it is known instantaneously. Figure 2-1 shows the Markov transition process between degradation states along with the transitions from each degradation state to the failure state. p_{ij} is the probability of going from degradation state i to the degradation state j during one discrete period given that the system has not failed yet, while f_i is the probability of going from degradation state i to the failure state.

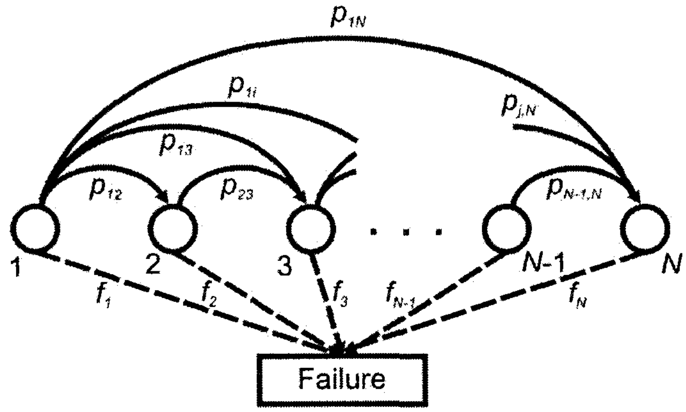

Figure 2-1: Markov process transition and transition to failure

The degradation states of the system are not observable except at time $t = 0$ when the state of the system is certainly 1. The transition matrix P is an upper triangular matrix, i.e. $p_{ij} = 0$ for $j < i$ and $p_{ij} = \Pr\left(Z(t+\Delta) = j \mid Z(t) = i, T > t+\Delta\right)$, $t = 0, \Delta, 2\Delta, \ldots$ otherwise. $p_{ij} = 0$ for $j < i$ means that system degradation state does not improve by itself which is true in most cases. T is a random variable representing the system's failure time. The system indicator is inspected at times; $t = \Delta, 2\Delta, \ldots$. The value of the indicator (observation) is assumed to take a value in a finite set of M non-negative integers, i.e. $\theta \in \Theta = \{1, 2, \ldots, M\}$. It is also supposed that a value of indicator θ is observed with a known probability $q_{j\theta}$, when the degradation state of the system is j. Q represents the stochastic matrix which specifies these probabilities, i.e. $Q = \left[q_{j\theta}\right]$, $j \in S$, $\theta \in \Theta$.

The failure is not considered a degradation state. It is a condition that causes the system to cease functioning and is outwardly obvious. If the failure happens, it is immediately recognized and the only possible action is *Failure Replacement* (FR). Otherwise, at any inspection point, we can decide whether to perform *Preventive Replacement* (PR) or *Do-Nothing*. The FR and the PR renew the system and return it to state 1 and period zero i.e. new or like new. The cost for the PR is C, while a FR costs $K+C, K, C>0$. Both actions, FR and PR, are instantaneous. Performing the inspections costs C_I per inspection, independent of the inspection's interval. More frequent inspections will cost more while provides information in a higher frequency that subsequently results in a more efficient replacement policy.

The system failure rate follows the PHM where the failure rate $h(t,Z_k) = h_0(t)\psi(Z_k)$ is a product of two independent functions. $h_0(.)$ is a function of the system's age only and $\psi(.)$ is a function of the system's degradation state only. $Z_k = Z(k\Delta)$ is the degradation state of the system at period k and Δ is the fixed inspection period. We assume that the degradation state of the system remains unchanged during each period (between two consecutive inspections) and each degradation state transfer, if any, is assumed to take place at the end of each period, just before the inspection point.

The objective is to find the optimal inspection period and corresponding replacement policy that minimizes the total long-run average cost per unit time.

Considering the failure rate of the system, the reliability of the system conditional on $Z_k = Z(k\Delta)$ is:

$$\begin{aligned} R(k,Z_k,t) &= P\left(T > k\Delta + t \mid T > k\Delta, Z_1, Z_2, \ldots, Z_k\right) \\ &= \exp\left(-\psi\left(Z_k\right)\int_{k\Delta}^{k\Delta+t} h_0(s)ds)\right); t < \Delta \end{aligned} \quad \textbf{(2-1)}$$

where T is the random variable indicating the failure time of the system. The conditional reliability of the system indicates the probability of not having the failure in time t after $k\Delta$, given that the failure has not happened until time $k\Delta$ and the degradation states of the system have been $Z_1, Z_2, \ldots, Z_k$ at times; $\Delta, 2\Delta, \ldots k\Delta$. The conditional mean sojourn time of the system, if no preventive action is performed while the system is in degradation state Z_k at period k, is: [Makis and Jardine 1992]

$$\tau\left(k, Z_k, \Delta\right) = \int_0^{\Delta} R\left(k, Z_k, t\right) dt \quad \textbf{(2-2)}$$

The conditional mean sojourn time of the system is defined as the expected remaining life (time to failure) of the system until between the current and the next inspection point. In what follows, to address the problem, the theory of Partially Observed Markov Decision Process (POMDP) is used. Since the degradation state of the system is not perfectly observable through the inspections, an alternative state space for the POMDP indicating the "conditional probability distribution of the system's degradation state" and then an "alternative state's transition" complying with the alternative state space are introduced.

2.5 Formulation of the POMDP

2.5.1 Alternative state space

We adapt the definition of π^k as defined in [Ghasemi *et al.* 2007] to indicate the *Conditional Probability Distribution of the system's degradation State* (CPDS) at period k:

$$\pi^k = \left\{ \pi_i^k;\ \ 0 \le \pi_i^k \le 1 \text{ for } i = 1,\dots,N, \sum_{i=1}^{N} \pi_i^k = 1 \right\},\ k = 0,1,2,\dots \qquad \textbf{(2-3)}$$

π_i^k represents the probability of being at state i at the k-th inspection point. The initial value of the CPDS for a new system is defined as:

$$\pi_i^0 = \begin{cases} 1 & i = 1 \\ 0 & 1 < i \le N \end{cases} \qquad \textbf{(2-4)}$$

which means that a new or renewed system is at state 1.

2.5.2 Alternative state transition

After each collection of observation θ at an inspection point, the CPDS is updated considering the latest observation θ. Using the Bayes' formula and assuming that the observation θ has occurred at the $k+1^{st}$ inspection point, the updated CPDS $\pi_j^{k+1}(\theta)$ is calculated as [Ghasemi *et al.* 2007]:

$$\pi_j^{k+1}(\theta) = \sum_{i=1}^{N} \pi_i^k p_{ij} q_{j\theta} \Big/ \Pr(\theta \mid k, \pi^k),\ j = 1,\dots,N \qquad \textbf{(2-5)}$$

where $\Pr(\theta \mid k, \pi^k) = \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_i^k p_{ij} q_{j\theta}$ is the probability of observing a certain value θ at $k+1^{st}$ inspection period (just before the inspection point) when the CPDS at the k-th period is π^k. This updated CPDS carries all the observations and actions history from the last replacement point. After any PR or FR the period counter k will be reset to zero and the CPDS will be set to π^0 as introduced in Eq. (2-2).

2.5.3 Decision space

$\{0, \infty\}$ is the decision space of the POMDP, where 0 means "replace the system immediately (PR)" and ∞ means "Do-Nothing". If "Do-nothing" is selected at a decision point and a failure happens before the next inspection (decision) point the system will be replaced immediately. The occurring cost for this event is $K + C$, $K, C > 0$. A Preventive Replacement costs C.

2.5.4 Dynamic Programming Formulation

Let $V(k, \pi^k)$ denote minimum *total cost of maintenance over the renewal period*, while the system is in the k-th inspection point and the CPDS is π^k. The *renewal period* is the time between two consecutive replacements, whether FR or PR. The *total cost* is defined as replacement cost of maintenance plus the inspections' cost.

$$V\left(k, \pi^k\right) = \min\left\{kC_I + C + V\left(0, \pi^0\right), W\left(k, \pi^k, g\right)\right\} \quad \textbf{(2-6)}$$

where $kC_I + C + V\left(0,\pi^0\right)$ is the renewal period's total cost at k-th inspection point, if one decides to perform the PR (decision=0). $W\left(k,\pi^k,g,C_I\right)$ is the optimal renewal period's total cost when at the k^{th} inspection point, one decides that no action takes place (decision=∞).

$$\begin{aligned} W\left(k,\pi^k,g\right) = {} & \left[kC_I + K + C + V\left(0,\pi^0\right)\right]\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right] - g\overline{\tau}\left(k,\pi^k,\Delta\right) \\ & + \left[\sum_{\theta=1}^{M} V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta \mid k,\pi^k\right)\right]\overline{R}\left(k,\pi^k,\Delta\right) \end{aligned} \tag{2-7}$$

$\left[kC_I + K + C + V\left(0,\pi^0\right)\right]$ represents the renewal period's total cost if the decision is "Do-nothing" (decision=∞) and the system fails during the next inspection period.

$\left[\sum_{\theta=1}^{M} V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta \mid k,\pi^k\right)\right]$ is the expected total future cost of the system at the $k+1$ inspection point, provided that the failure has not happened during the k-th period. $\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]$ and $\overline{R}\left(k,\pi^k,\Delta\right)$ are the probability of having the failure during the k-th period and the probability that the system is still working at the beginning of the $k+1^{\text{st}}$ period consecutively while the CPDS at period k is π^k. $\overline{\tau}\left(k,\pi^k,\Delta\right)$ and g are the mean sojourn time of the system at the $k+1^{\text{st}}$ period when the CPDS at the k-th period π^k, is available and the average replacement cost per unit of time over infinite horizon respectively [Ghasemi *et al.* 2007]. g includes the cost of replacements only and excludes the inspections cost.

$$\overline{R}\left(k,\pi^{k},t\right)=\sum_{i=1}^{N}R\left(k,i,t\right)\,\pi_{i}^{k} \tag{2-8}$$

$$\overline{\tau}\left(k,\pi^{k},\Delta\right)=\int_{0}^{\Delta}\overline{R}\left(k,\pi^{k},t\right)dt \tag{2-9}$$

If a failure happens, the system is replaced and begins the service immediately anyhow next inspection will continue as scheduled. The term $g\overline{\tau}\left(k,\pi^{k},\Delta\right)$ is the expected long-run cost of the overlapped time of two consecutive replacements of the system when the system fails. For more details please refer to [Ghasemi *et al.* 2007].

2.5.5 Optimal Policy

To establish the optimal maintenance policy of the described problem, one needs an optimal decision criterion to apply at each decision (inspection) point. This criterion is a function of the observed indicator, the age of the system, the system's cost parameters and finally g, the long-run average cost of the system. In the following parts, first we introduce a decision criterion depending on g. The decision criterion and the *minimum long-run average cost* of the system g^{*} together give the *optimal* decision criterion for the introduced problem.

2.5.6 Decision Criterion

Considering Eq. (2-7) above one can write:

$$W\left(k,\pi^k,g\right)=kC_I+C+V\left(0,\pi^0\right)-\left(kC_I+C+V\left(0,\pi^0\right)\right)\overline{R}\left(k,\pi^k,\Delta\right)+K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]$$
$$-g\overline{\tau}\left(k,\pi^k,\Delta\right)+\left[\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\,|\,k,\pi^k\right)\right]\overline{R}\left(k,\pi^k,\Delta\right)$$

$$\begin{aligned}&W\left(k,\pi^k,g\right)-\left[kC_I+C+V\left(0,\pi^0\right)\right]=K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]\\&-\left(kC_I+C+V\left(0,\pi^0\right)\right)\overline{R}\left(k,\pi^k,\Delta\right)\\&-g\overline{\tau}\left(k,\pi^k,\Delta\right)+\left[\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\,|\,k,\pi^k\right)\right]\overline{R}\left(k,\pi^k,\Delta\right)\end{aligned} \quad \textbf{(2-10)}$$

Since $V\left(k+1,\pi^{k+1}\right)$ is the minimum expected renewal period cost at the $k+1^{st}$ period, then:

$$V\left(k+1,\pi^{k+1}\right)\le kC_I+C+V\left(0,\pi^0\right)$$
$$V\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\,|\,k,\pi^k\right)\le\left[kC_I+C+V\left(0,\pi^0\right)\right]\Pr\left(\theta\,|\,k,\pi^k\right)$$
$$\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\,|\,k,\pi^k\right)\le\sum_{\theta=1}^{M}\left[kC_I+C+V\left(0,\pi^0\right)\right]\Pr\left(\theta\,|\,k,\pi^k\right)$$
$$\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\,|\,k,\pi^k\right)\le kC_I+C+V\left(0,\pi^0\right)$$

This means that:

$$\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\,|\,k,\pi^k\right)-\left[kC_I+C+V\left(0,\pi^0\right)\right]\le 0$$
$$\therefore\left[\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\,|\,k,\pi^k\right)-kC_I-C-V\left(0,\pi^0\right)\right]\overline{R}\left(k,\pi^k,\Delta\right)\le 0 \quad \textbf{(2-11)}$$

If $K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]<g\overline{\tau}\left(k,\pi^k,\Delta\right)$ then the sum of all the terms in right hand side of Eq. (2-8) will be negative or zero i.e. $W\left(k,\pi^k,g\right)\le\left[kC_I+C+V\left(0,\pi^0\right)\right]$. This final equation means that the cost of leaving the system and doing no preventive action is less than the cost of the PR so the optimal decision at k-th inspection point i.e. optimal decision$=\infty$.

In the case that $K\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\geq g\bar{\tau}\left(k,\pi^{k},\Delta\right)$ we show that the best solution is to replace the system immediately i.e. Decision=0. To continue assume the contrary, that is to say; assume while $K\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\geq g\bar{\tau}\left(k,\pi^{k},\Delta\right)$, the optimal action is "Do-Nothing", from Eq. (2-4) based on this assumption:

$$V\left(k,\pi^{k}\right)=W\left(k,\pi^{k},g\right)<C+V\left(0,\pi^{0}\right)+C_{I} \tag{2-12}$$

Also we can write:

$$\begin{aligned}V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)=&V\left(k+1,\pi^{k}\right)\bar{R}\left(k,\pi^{k},\Delta\right)+V\left(k+1,\pi^{k}\right)\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\\&-V\left(k,\pi^{k}\right)\end{aligned}$$

$$\begin{aligned}V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)=&V\left(k+1,\pi^{k}\right)\bar{R}\left(k,\pi^{k},\Delta\right)+V\left(k+1,\pi^{k}\right)\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\\&-W\left(k,\pi^{k},g\right)\end{aligned}$$

Replacing $W\left(k,\pi^{k},g\right)$ with right hand side of (2-5):

$$\begin{aligned}V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)=&\left[V\left(k+1,\pi^{k}\right)-K-C-kC_{I}-V\left(0,\pi^{0}\right)\right]\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\\&+\left[V\left(k+1,\pi^{k}\right)-\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\mid k,\pi^{k}\right)\right]\bar{R}\left(k,\pi^{k},\Delta\right)\\&+g\bar{\tau}\left(k,\pi^{k},\Delta\right)\end{aligned}$$

$$\begin{aligned}V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)=&\left[\underbrace{V\left(k+1,\pi^{k}\right)-C-kC_{I}-V\left(0,\pi^{0}\right)}_{1}\right]\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\\&+\left[\underbrace{V\left(k+1,\pi^{k}\right)-\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\mid k,\pi^{k}\right)}_{2}\right]\bar{R}\left(k,\pi^{k},\Delta\right)\\&+g\bar{\tau}\left(k,\pi^{k},\Delta\right)-K\left[1-\bar{R}\left(k,\pi^{k},\Delta\right)\right]\end{aligned}$$

By definition $V\left(k+1,\pi^{k}\right)$ is the minimum renewal period cost then $V\left(k+1,\pi^{k}\right)\leq C+kC_{I}+V\left(0,\pi^{0}\right)$ so term 1 is not positive. We have proved in appendix that $V\left(k,\pi^{k}\right)$ is non-decreasing in $\left(k,\pi^{k}\right)$ then:

$$
\begin{aligned}
&V\left(k+1,\pi^{k}\right)\leq V\left(k+1,\pi^{k+1}(\theta)\right)\\
&\Pr\left(\theta\mid k,\pi^{k}\right)V\left(k+1,\pi^{k}\right)\leq \Pr\left(\theta\mid k,\pi^{k}\right)V\left(k+1,\pi^{k+1}(\theta)\right)\\
&\sum_{\theta=1}^{M}\Pr\left(\theta\mid k,\pi^{k}\right)V\left(k+1,\pi^{k}\right)\leq \sum_{\theta=1}^{M}\Pr\left(\theta\mid k,\pi^{k}\right)V\left(k+1,\pi^{k+1}(\theta)\right)\\
&V\left(k+1,\pi^{k}\right)\leq \sum_{\theta=1}^{M}\Pr\left(\theta\mid k,\pi^{k}\right)V\left(k+1,\pi^{k+1}(\theta)\right)
\end{aligned}
$$

where θ is the indicator observed at $k+1^{\text{st}}$ inspection point. Then the term 2 is not positive. Terms $\overline{R}\left(k,\pi^{k},\Delta\right)$ and $\left[1-\overline{R}\left(k,\pi^{k},\Delta\right)\right]$ are not negative by definition, so:

$$
V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)<-K\left[1-\overline{R}\left(k,\pi^{k},\Delta\right)\right]+g\overline{\tau}\left(k,\pi^{k},\Delta\right)
$$

Since we are considering the case where $K\left[1-\overline{R}\left(k,\pi^{k},\Delta\right)\right]\geq g\overline{\tau}\left(k,\pi^{k},\Delta\right)$:

$$
\therefore V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)<0
$$

In the other hand we have proved that $V\left(k,\pi^{k}\right)$ is non-decreasing in $\left(k,\pi^{k}\right)$ then $V\left(k+1,\pi^{k}\right)-V\left(k,\pi^{k}\right)\geq 0$, which is a contradiction. This means that the optimal decision is to replace the system immediately (Decision=0).

The decision criteria can be briefly written as:

$$
a\left(k,\pi^{k}\right)=\begin{cases}\infty & \text{if}\quad K\left[1-\overline{R}\left(k,\pi^{k},\Delta\right)\right]<g\overline{\tau}\left(k,\pi^{k},\Delta\right)\\ 0 & \text{if}\quad K\left[1-\overline{R}\left(k,\pi^{k},\Delta\right)\right]\geq g\overline{\tau}\left(k,\pi^{k},\Delta\right)\end{cases} \qquad \textbf{(2-13)}
$$

where $a(k,\pi^k)$ indicates the decision at period k while the CPDS is π^k. The optimal decision is dependent on the long run expected cost per unit time of replacement system (excluding inspection cost) i.e. g. The optimal decision based on minimum long-run average cost g^*, can be calculated based on (2-11) by replacing g by g^*.

$$a^*\left(k,\pi^k\right)=\begin{cases}\infty & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]<g^*\overline{\tau}\left(k,\pi^k,\Delta\right)\\ 0 & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]\geq g^*\overline{\tau}\left(k,\pi^k,\Delta\right)\end{cases} \tag{2-14}$$

The calculation detail of g^* is given in the next section.

We note that, if the inspection period can be changed, on one hand, there is a constant cost that is paid at every inspection epoch i.e. more frequent inspection costs more. On the other hand, more frequent inspections provide more information that can lead to a more cost effective replacement policy. This means that the optimal inspection period can be selected between several feasible inspection periods. The criterion that helps us to select the optimal inspection period is the *minimum total long-run average cost* G^*, which is the long-run average cost of replacements and inspections. In next section we calculate this measure as well.

2.6 Long-run average cost and total long-run average cost

In this part we introduce a method to calculate the minimum long-run average cost g^* and the minimum total long-run average cost G^*. The iterative method introduced by [Ghasemi *et al.* 2007] can be used to calculate the optimal value of g, i.e. g^*. The minimum long run average cost per unit of time where stopping-time is T_g, is the unique solution of: $g = \dfrac{C + K\overline{\Pr}(T_g > T)}{\bar{E}_{\min}(T_g, T)}$ where T is the time to failure, $\overline{\Pr}(T_g > T)$ is the probability of a FR, and $\bar{E}_{\min}(T_g, T)$ is the expected average length of a replacement cycle. Stopping time is defined as $T_g = \Delta.\inf\{n \geq 0 \mid K[1 - \bar{R}(n, \pi^n, \Delta)] \geq g\bar{\tau}(n, \pi^n, \Delta)\}$ [Ghasemi *et al.* 2007].

By letting $C^{T_{g^*}}$ and $P^{T_{g^*}}$ represent the expected cost and expected length over the renewal period associated with a replacement policy in which the optimal time to replacement is T_{g^*} and g^* represents the minimum long-run average cost of replacement (excluding the inspection cost). The min. total long-run average cost per unit of time is:

$$G^* = \frac{C^{T_{g^*}}}{P^{T_{g^*}}} + \frac{C_I}{\Delta} = \frac{C\left[1 - \overline{\Pr}(T_{g^*} > T)\right] + (C + K)\overline{\Pr}(T_{g^*} > T)}{\bar{E}_{\min}(T_{g^*}, T)} + \frac{C_I}{\Delta}$$

$$G^* = \frac{C + K\overline{\Pr}(T_{g^*} > T)}{\bar{E}_{\min}(T_{g^*}, T)} + \frac{C_I}{\Delta}$$

where C, K and C_I are the replacement cost, failure cost and inspection cost respectively.

The following equations can be used to calculate $\overline{\Pr}(T_{g^*} > T)$ and $\bar{E}_{\min}(T_{g^*}, T)$.

$\overline{\Pr}(T_g > T) = Q(0, \pi^0)$ and $\bar{E}_{\min}(T_g, T) = W(0, \pi^0)$ where:

$$W\left(j,\pi^j\right)=\begin{cases}0 & j\geq k\\ \displaystyle\int_0^{t_g\left(\pi^j\right)-j\Delta}\overline{R}\left(j,\pi^j,s\right)ds & j=k-1\\ \displaystyle\int_0^{\Delta}\overline{R}\left(j,\pi^j,s\right)ds+\sum_{\theta=1}^{M}W\left(j+1,\pi^{j+1}(\theta)\right)\overline{R}\left(j,\pi^j,\Delta\right)\Pr\left(\theta\mid j,\pi^j\right) & j<k-1\end{cases}$$

also:

$$Q(j,\pi^j)=\begin{cases}0 & j\geq k\\ 1-\overline{R}(j,\pi^j,t_g\left(\pi^j\right)-j\Delta) & j=k-1\\ 1-\overline{R}(j,\pi^j,\Delta)+\displaystyle\sum_{\theta=1}^{M}Q(j+1,\pi^{j+1}(\theta))\Pr(\theta\mid j,\pi^j)\overline{R}(j,\pi^j,\Delta) & j<k-1\end{cases}$$

where $t_g(\pi)=\Delta\left\{r\in R^+\mid K\left[1-\overline{R}(r,\pi,\Delta)\right]=g\overline{\tau}(r,\pi,\Delta)\right\}$. For more details please refer to [Ghasemi *et al.* 2007].

2.7 Optimal inspection period

We assume that the optimal inspection period is to be chosen from a finite set of L possible inspection periods $\Delta_l, l=1,2,...,L$. The optimal inspection period is the one with the minimum total long-run average cost G^* calculated in previous section. First, the optimal inspection period, based on the previous sections results is found and then based on the result of "Optimum Policy" section, the optimal replacement policy is fixed.

In the following section we solve a replacement example without any considerable inspection cost and a prefixed inspection period; latter we add a considerable cost for the inspections and assume two possible inspection periods and we find the optimal inspection period, long-run average cost and the corresponding optimal replacement criteria.

2.8 Numerical Example

We use the example presented by [Ghasemi *et al.* 2007] and adapt it to our case of costly inspections. In this example, it is assumed that system has a two parameter Weibull like behaviour with baseline distribution hazard function having the following parameters. $h_0(t) = \frac{\beta t^{\beta-1}}{\alpha^{\beta}}, t \geq 0$, $\alpha = 1, \beta = 2$ and $\psi(X_t) = e^{0.5(X_t - 1)}$. The system's two possible degradation states are $\{1,2\}$ with the transition probability matrix $P_1 = \begin{bmatrix} 0.4 & 0.6 \\ 0 & 1 \end{bmatrix}$ when the inspection period is $\Delta_1 = 0.5$. θ, the observed value of the system's indicator, can take three possible values. The indicator value and the system's degradation state are related by the probability distribution $Q = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$. $C = 5$ and $K = 2$ represent the replacement cost and the failure cost of the system consecutively. The long-run average cost of replacement, based on the provided method, is found to be $g_1^* = 8.67$ and the optimum stopping time of the system is $T_{g_1^*} = \inf\left\{k \geq 0; 2 \times \left[1 - \overline{R}\left(k, \pi^k, 0.5\right)\right] \geq 8.67 \times \overline{\tau}\left(k, \pi^k, 0.5\right)\right\}$.

Now assume that inspection cost $C_I = 1$ applies for each inspection to obtain the system's indicator value. We also assume that the there is another possible inspection period $\Delta_2 = 0.6$ with corresponding degradation state transition matrix $P_2 = \begin{bmatrix} 0.3 & 0.7 \\ 0 & 1 \end{bmatrix}$.

We are interested in finding the optimal replacement interval and corresponding replacement criteria. The following table shows the final result of the method applied on the data.

Table 2-1: Numerical example (optimal inspection interval)

l	Δ_l	g_l^*	G^*_l	**Stopping Time (Replacement Policy)**
1	0.5	8.67	10.67	$T_{g_1^*} = \inf\left\{k \ge 0; 2\times\left[1-\overline{R}\left(k,\pi^k,0.5\right)\right] \ge 8.67\times\overline{\tau}\left(k,\pi^k,0.5\right)\right\}$
2	0.6	8.73	10.39	$T_{g_1^*} = \inf\left\{k \ge 0; 2\times\left[1-\overline{R}\left(k,\pi^k,0.6\right)\right] \ge 8.73\times\overline{\tau}\left(k,\pi^k,0.6\right)\right\}$

While the long-run average cost of replacement for the shorter inspection period $\Delta_1 = 0.5$, is smaller; the total long-run average cost G^*_2, corresponding to $\Delta_2 = 0.6$, is the optimal one. It means that we will totally pay less, if we observe the system by inspection period equal to 0.6 and applying the corresponding stopping-time.

2.9 Conclusion

For a system which is subject to a CBM program, inspections are performed to obtain proper indicators concerning the degradation state of the system and decide on an optimal replacement policy. In many practical cases, the inspections do not reveal the exact system degradation state. In this work we have relaxed the assumption of non-costly inspection and found the optimal replacement policy and the total long-run average cost of the system replacement and inspections. More frequent inspections while cost more, can lead to a less costly inspection and replacement policy due to higher frequency of information provided regarding the system's degradation state. So an optimal inspection period minimizing the total long-run average cost of the system can be identified. The numerical example shows the application of the model.

2.10 References

1. Banjevic, D.; Jardine, A. K. S.; Makis, V., and Ennis, M. A control-limit policy and software for condition-based maintenance optimization. INFOR Journal. 2001; 39(1):pp 32-50.

2. Chelbi, Anis and Ait-Kadi, Daoud. An optimal inspection strategy for randomly failing equipment . Reliability Engineering and System Safety . 1999; 63:pp127-131.

3. Christer, A. H.; Wang, W., and Sharp, J. A state space condition monitoring model for furnace erosion prediction and replacement. European Journal of Operational Research. 1997; 101:1-14.

4. Cox, David Roxbee. Regression models and life tables (with discussion). Journal of the Royal Statistical Society - Series B. 1972; 26 pp 187-220.

5. Cox, David Roxbee. Renewal Theory. London: Chapman and Hall; 1962.

6. Ghasemi, Alireza; Yacout S. , and Ouali, M. S. Optimal condition based maintenance with imperfect information and the proportional hazards model. International Journal of Production Research. 2007 Feb; 45(4):pp 989-1012.

7. Grall, A.; Berenguer, C., and Dieulle L. A condition-based maintenance policy for stochastically deteriorating systems. Reliability Engineering and System Safety. 2002; 76:pp 167-180.

8. Hontelez, A. M.; Burger, Helen H., and Wijnmalen, Diederik J. D. Optimum condition-based maintenance policies for deteriorating systems with partial information. Reliability Engineering and System Safety. 1996; 51:pp267-274.

9. Jardine, A. K. S.; Anderson, P. M., and Mann D. S. Application of the Weibull Proportional Hazards Model to Aircraft and Marine Engine Failure Data. Quality and Reliability Engineering International. 1987; 3:pp 77-82.

10. Lam, Teresa C. and Yen, R. H. Optimal Maintenance-Policies For Deteriorating Systems Under Various Maintenance Strategies. IEEE Transactions on Reliability . 1994; 43(3):pp 423-430.

11. Leemis, L. M. Reliability- Probabilistic Models and Strategical Methods. New Jersey: Prentice all, Inc., Engelwood Cliffs; 1995.

12. Lin, D.; Banjevic, D., and Jardine, A. K. S. Using principal components in a proportional hazards model with application in condition-based maintenance. Journal of the Operational Research Society. 2005; pp 1-10.

13. Maillart, L. M. Optimal condition-Monitoring Schedules for Multi-State Deterioration Systems with Obvious Failure. 2004; Technical Memorandum Number 778.

14. Makis, V. and Jardine, A. K. S. Computation of optimal policies in replacement models. IMA Journal of Mathematics Applied in Business & Industry. 1992; 3 pp 169-175.

15. Makis, V. and Jardine, A. K. S. Optimal replacement in the proportional hazards model. INFOR Journal . 1992a; 30(2):172-183.

16. Ohnishi, M.; Kawai, H., and Mine, H. An optimal inspection and replacement policy under incomplete state information. European Journal of Operational Research. 1986; 27: pp 117-128.

2.11 Appendix

2.11.1 Monotone behavior

In this part, the condition under which the cost function introduced in the previous part has a monotonic behavior is established. Some definitions and propositions presented by [Ohnishi *et al.* 1994], [Rosenfield 1975] and [Kurano 1985] are adopted.

Definition 1: [Ohnishi 1994] An N-dimensional vector x, is said to be *stochastically less* than an N-dimensional vector y, if and only if $\sum_{i=k}^{N} x_i \leq \sum_{i=k}^{N} y_i$ for any $k; 1 \leq k \leq N$ and is denoted by $x \underset{ST}{\leq} y$.

Definition 2: [Ohnishi 1994] An N-dimensional vector x, is said to be *less* than an N-dimensional vector y in *Likelihood ratio*, if and only if $\begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} \geq 0$ for $1 \leq \mathrm{i} \leq \mathrm{j} \leq \mathrm{N}$ or equivalently $x_i y_j \geq x_j y_i$ for $1 \leq \mathrm{i} \leq \mathrm{j} \leq \mathrm{N}$ and is denoted by $x \underset{LR}{\leq} y$.

Definition 3: [Rosenfield 1995] An N-dimensional probability transition matrix P is said to be *Increasing Failure Rate* (IFR) if its rows are stochastically increasing i.e. $\sum_{j=k}^{N} p_{i\,j} \leq \sum_{j=k}^{N} p_{\hat{i}\,j}$ and $1 \leq k \leq N$ and $\hat{i} > i$. In other words, the $\sum_{j=k}^{N} p_{i\,j}$ is non-decreasing in i for $1 \leq k \leq N$.

Definition 4: [Rosenfield 1975] An N-dimensional probability transition matrix P, is said to be *Totally Positive of Order* 2 (TP2), if its rows are increasing in likelihood ratio, i.e. [Ohnishi 1994] $\begin{vmatrix} p_{im} & p_{in} \\ p_{jm} & p_{jn} \end{vmatrix} \geq 0$ for $\begin{cases} 1 \leq \mathrm{i} \leq \mathrm{j} \leq \mathrm{N} \\ 1 \leq \mathrm{m} \leq \mathrm{n} \leq \mathrm{N} \end{cases}$.

In this research, it is supposed that $P = \left[p_{ij} \right]$ $i, j \in S$, the Markovian probability transition matrix, and $Q = \left[q_{j\theta} \right]$ $j \in S, \theta \in \Theta$ are *TP2*.

Proposition 1: [Ohnishi 1994] Having $[a_i, 1 \leq i \leq N]$, an N-dimensional vector with non-decreasing elements, if $x \underset{ST}{\leq} y$ then $\sum_{i=k}^{N} a_i x_i \leq \sum_{i=k}^{N} a_i y_i$ for $1 \leq \mathrm{k} \leq \mathrm{N}$.

Proposition 2: [Rosenfield 1976] If P is *TP2* then P is *IFR*.

Proposition 3: [Ohnishi 1994] $\pi \underset{LR}{\leq} \hat{\pi} \Rightarrow \pi \underset{ST}{\leq} \hat{\pi}$, where π, $\hat{\pi}$ are two N-dimensional vectors.

The following lemmas are adopted without the proofs from [Ghasemi *et al.* 2007] where π^k and $\hat{\pi}^k$ are two hypothetical CPDSs at the k-th period.

Lemma 1: If $\pi^k \underset{ST}{\leq} \hat{\pi}^k$ then for any a, $\overline{R}(k, \pi^k, a) \geq \overline{R}(k, \hat{\pi}^k, a)$ where $\overline{R}(k, \pi^k, a)$ is the probability that the system is still working at $k\Delta + a$ while the CPDS at period k is π^k

Lemma 2: $\overline{R}(k, \pi^k, a)$ is non-increasing in k for any a.

Lemma 3: If $\pi^k \underset{ST}{\leq} \hat{\pi}^k$, then $\bar{\tau}\left(k,\pi^k,a\right) \geq \bar{\tau}\left(k,\hat{\pi}^k,a\right)$ for any a where $\bar{\tau}\left(k,\pi^k,a\right)$ is the mean sojourn time of the system between $k\Delta$ and $k\Delta+a$ when the CPDS at the k-th period is π^k.

Lemma 4: $\bar{\tau}\left(k,\pi^k,a\right)$ is non-increasing in k for any a.

Theorem 1: Assuming that assumptions 1 through 5 as stated in [Makis and Jardine 1992] are satisfied, function V introduced by Eq. (2-6) defined on $\bar{S}$, where $\bar{S}$ is the set of all possible variations of the pair $\left(k,\pi^k\right)$, with a constant $g \geq 0$, is a bounded measurable *non-decreasing* function.

Proof:

We consider the restricted action space to be defined as $A_\varepsilon = \{\varepsilon,\infty\}$ where ε means taking the action in a short time. [Ghasemi *et al.* 2007] have shown that under this condition, there exist a non-negative real valued function $v_\varepsilon(k,\pi^k) \in D$ for the restricted action space A_ε, such that $v_\varepsilon\left(k,\pi^k\right) = U^\varepsilon v_\varepsilon\left(k,\pi^k\right)$ where D is a Borel subset of $\bar{S}$. The map U^ε is defined as:

$U^\varepsilon u\left(k,\pi^k\right) = \min\left\{kC_l + C + U\left(0,\pi^0,+\infty,u\right), U\left(k,\pi^k,+\infty,u\right)\right\}, u \in D$. By letting $k=0$, since $C>0$, $U^\varepsilon u\left(0,\pi^0\right) = \min\left\{kC_l + C + U\left(0,\pi^0,+\infty,u\right), U\left(0,\pi^0,+\infty,u\right)\right\} = U\left(0,\pi^0,+\infty,u\right)$ and then we can write:

$$U^{\varepsilon}u\left(k,\pi^{k}\right)=\min\left\{\ C+U^{\varepsilon}u\left(0,\pi^{k}\right),\ U\left(k,\pi^{k},+\infty,u\right)\right\} \tag{2-15}$$

where we define:

$$U\left(k,\pi^{k},a,u\right)=\left[kC_{I}+K+C+u\left(0,\pi^{0}\right)\right]-g\bar{\tau}\left(k,\pi^{k},a\right)$$
$$+\sum_{\theta=1}^{M}\left[\left[u\left(k+1,\pi^{k+1}(\theta)\right)-\left(kC_{I}+K+C+u\left(0,\pi^{0}\right)\right)\right]\Pr\left(\theta\mid k,\pi^{k}\right)\right]\bar{R}\left(k,\pi^{k},a\right)$$

$$\begin{aligned}U\left(k,\pi^{k},a,u\right)=\left[kC_{I}+K+C+u\left(0,\pi^{0}\right)\right]\left(1-\bar{R}\left(k,\pi^{k},a\right)\right)-g\bar{\tau}\left(k,\pi^{k},a\right)\\+\sum_{\theta=1}^{M}\left[u\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\mid k,\pi^{k}\right)\right]\bar{R}\left(k,\pi^{k},a\right)\end{aligned} \tag{2-16}$$

for each $u\in D$ and any constant $g>0$.

Corollary1:

For any non-decreasing function $u(k,\pi)$, where $u\left(k,\pi\right)\le kC_{I}+K+C+u(0,\pi^{0})$ for any (k,π), $U(k,\pi,\Delta,u)$ is non-decreasing in k i.e. $U(k,\pi,\Delta,u)\ge U(k',\pi,\Delta,u)$ where $k\le k'$.

Proof:

Using Eq. (2-16), we can write:

$$\begin{aligned}\frac{\delta U\left(k,\pi,\Delta,u\right)}{\delta k}=C_{I}\left(1-\bar{R}\left(k,\pi,\Delta\right)\right)-\left[kC_{I}+K+C+u\left(0,\pi^{0}\right)\right]\frac{\delta\bar{R}\left(k,\pi,\Delta\right)}{\delta k}-g\frac{\delta\bar{\tau}\left(k,\pi,\Delta\right)}{\delta k}\\+\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr\left(\theta\mid k,\pi\right)\right]\frac{\delta\bar{R}\left(k,\pi,\Delta\right)}{\delta k}\\+\frac{\delta\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr\left(\theta\mid k,\pi\right)\right]}{\delta k}\bar{R}\left(k,\pi,\Delta\right)\end{aligned}$$

$$\frac{\delta U(k,\pi,\Delta,u)}{\delta k}=\underbrace{\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)-\left(kC_I+K+C+u\left(0,\pi^0\right)\right)\right]}_{1}\frac{\delta\overline{R}(k,\pi,\Delta)}{\delta k}$$

$$-g\frac{\delta\overline{\tau}(k,\pi,\Delta)}{\delta k}+C_I\left(1-\overline{R}(k,\pi,\Delta)\right)$$

$$+\underbrace{\frac{\delta\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]}{\delta k}}_{2}\overline{R}(k,\pi,\Delta)$$

(2-17)

Now we show that term 1 is negative. Following the assumption of the corollary which states that $u\left(k+1,\pi^{k+1}\right)<kC_I+K+C+u\left(0,\pi^0\right)$ by multiplying both side with $\Pr\left(\theta\mid k,\pi^k\right)$ and summing up on all possible amounts of θ,we can write:

$$u\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\mid k,\pi^k\right)<\left[kC_I+K+C+u\left(0,\pi^0\right)\right]\Pr\left(\theta\mid k,\pi^k\right)$$

$$\sum_{\theta=1}^{M}u\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\mid k,\pi^k\right)<\sum_{\theta=1}^{M}\left[kC_I+K+C+u\left(0,\pi^0\right)\right]\Pr\left(\theta\mid k,\pi^k\right)$$

$$\sum_{\theta=1}^{M}u\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\mid k,\pi^k\right)<kC_I+K+C+u\left(0,\pi^0\right)$$

which means that $\sum_{\theta=1}^{M}u\left(k+1,\pi^{k+1}\right)\Pr\left(\theta\mid k,\pi^k\right)-kC_I-K-C-u\left(0,\pi^0\right)<0$

Now we consider term 2:

$$\frac{\delta\left[\sum_{\theta=1}^{M} u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]}{\delta k}=\frac{\sum_{\theta=1}^{M}\delta\left[u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]}{\delta k}$$
$$=\frac{\sum_{\theta=1}^{M}\Pr(\theta\mid k,\pi)\,\delta u\left(k+1,\pi(\theta)\right)}{\delta k}$$
$$=\sum_{\theta=1}^{M}\Pr(\theta\mid k,\pi)\frac{\delta u\left(k+1,\pi(\theta)\right)}{\delta k}$$

Since $u(k,\pi)$ is non-decreasing in k the summation will be positive or zero.

From lemma 2, it follows that $\frac{\delta \overline{R}(k,\pi,\Delta)}{\delta k}\leq 0$, and from lemma 4, it follows that $\frac{\delta \overline{\tau}(k,\pi,\Delta)}{\delta k}\leq 0$ and also the term $C_I\left(1-\overline{R}(k,\pi,\Delta)\right)$ is always positive. So that in EQ. (2-17), $\frac{\delta U(k,\pi,\Delta,u)}{\delta k}\geq 0$ i.e. $U(k,\pi,\Delta,u)$ defined by Eq. (2-16) is non-decreasing in k.

Corollary 2:

For any non-decreasing function $u(k,\pi)$, where $u(k,\pi)\leq kC_I+K+C+u(0,\pi^0)$, for any (k,π), $U(k,\pi^k,\Delta,u)$ is non-decreasing in π, i.e. $U(k,\pi,\Delta,u)\geq U(k,\pi',\Delta,u)$ if $\pi \underset{LR}{\leq} \pi'$.

Proof:

$$\frac{\delta U(k,\pi,\Delta,u)}{\delta\pi}=-\left[kC_I+K+C+u\left(0,\pi^0\right)\right]\frac{\delta\bar{R}(k,\pi,\Delta)}{\delta\pi}-g\frac{\delta\bar{\tau}(k,\pi,\Delta)}{\delta\pi}$$
$$+\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]\frac{\delta\bar{R}(k,\pi,\Delta)}{\delta\pi}$$
$$+\frac{\delta\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]}{\delta\pi}\bar{R}(k,\pi,\Delta)$$

$$\frac{\delta U(k,\pi,\Delta,u)}{\delta\pi}=\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)-kC_I-K+C+u\left(0,\pi^0\right)\right]\frac{\delta\bar{R}(k,\pi,\Delta)}{\delta\pi}$$
$$-g\frac{\delta\bar{\tau}(k,\pi,\Delta)}{\delta\pi}+\frac{\delta\left[\sum_{\theta=1}^{M}u\left(k+1,\pi(\theta)\right)\Pr(\theta\mid k,\pi)\right]}{\delta\pi}\bar{R}(k,\pi,\Delta)$$

By using lemmas 1 and 3, $\frac{\delta U(k,\pi,\Delta,u)}{\delta\pi}\geq 0$ i.e. $U(k,\pi,\Delta,u)$ defined by Eq. (2-16) is non-decreasing in π.

Suppose $u_0(k,\pi^k)=0$ for any (k,π^k) in (2-15), so u_0 is non-decreasing in (k,π^k). By corollaries 1 and 2, and also by considering that $u_n=U^{\varepsilon}u_{n-1}$, as given by [Kurano 1985], then u_1 is non-decreasing. By induction $u_n\left(k,\pi^k\right)$ is non-decreasing in $\left(k,\pi^k\right)$ for any n.

Since $v_n\rightarrow v_{\varepsilon}$ when $n\rightarrow\infty$ [Kurano 1985], v_{ε} is non-decreasing as well. It can be seen that $v_{\varepsilon}\left(k,\pi^k\right)\leq v_{\bar{\varepsilon}}\left(k,\pi^k\right)$ if $\varepsilon\leq\bar{\varepsilon}$. Suppose $v\left(k,\pi^k\right)=\lim_{\varepsilon\rightarrow 0}v_{\varepsilon}\left(k,\pi^k\right)$ for any $\left(k,\pi^k\right)$. This implies that $v\left(k,\pi^k\right)$ is non-decreasing in $\left(k,\pi^k\right)$. We note that $A_{\varepsilon}\rightarrow A$

while $\varepsilon \to 0$. Since map U is monotone [Kurano 1985], and by the monotone convergence theorem [Capinski 2004] we get:

$$\lim_{\varepsilon \to 0} U\left(k, \pi^k, +\infty, v_\varepsilon\right) = U\left(k, \pi^k, +\infty, v\right) \tag{2-18}$$

From equations (2-15) and (2-16), when $n \to \infty$ we get:

$$\lim_{n \to \infty} U^\varepsilon u_n\left(k, \pi^k\right) = \min\left\{ \lim_{n \to \infty}\left(kC_I + C + U^\varepsilon u_n\left(0, \pi^k\right)\right),\ \lim_{n \to \infty} U\left(k, \pi^k, +\infty, u_n\right)\right\}$$
$$U^\varepsilon v_\varepsilon\left(k, \pi^k\right) = \min\left\{ kC_I + C + U^\varepsilon v_\varepsilon\left(0, \pi^k\right),\ U\left(k, \pi^k, +\infty, v_\varepsilon\right)\right\}$$

and when $\varepsilon \to 0$ we get:

$$\lim_{\varepsilon \to \infty} U^\varepsilon v_\varepsilon\left(k, \pi^k\right) = \min\left\{\lim_{\varepsilon \to \infty}\left(kC_I + C + U^\varepsilon v_\varepsilon\left(0, \pi^k\right)\right), \lim_{\varepsilon \to \infty} U\left(k, \pi^k, +\infty, v_\varepsilon\right)\right\}$$

$$v\left(k, \pi^k\right) = \min\left\{kC_I + C + v\left(0, \pi^k\right), U\left(k, \pi^k, +\infty, v\right)\right\} \tag{2-19}$$

U and v defined by (2-18) and (2-19) respectively, represent W and V defined in (2-6) and (2-7) respectively which finishes the proof of theorem 1.

CHAPTER 3 :

EVALUATING THE REMAINING LIFE FOR EQUIPMENT WITH UNOBSERVABLE STATES

Alireza Ghasemi, Soumay Yacout, M-Salah Ouali

Submitted to: IEEE Transactions on Reliability, TR number: TR2008-056

3.1 Abstract

This article proposes a model to calculate the Reliability Function (RF) and the Mean Residual (Remaining) Life (MRL) of a piece of equipment, when its degradation state is not directly observable. At each observation moment, an indicator of the underlying unobservable degradation state is observed and the monitoring information is collected. The observation process is due to a condition monitoring system where the obtained information may contain noise due to errors of measurement, interpretation, accuracy of measuring devices, etc. For that reason, the observation process is not perfect and does not directly reveal the exact degradation state. In order to match an indicator's value and the unobservable degradation state, a stochastic relation between them is given by an observation probability matrix.

We assume that the equipment's unobservable degradation state transition follows a Markov Chain and we model it by a Hidden Markov Model. Bayes' rule is used to determine the probability of being in a certain degradation state at each observation moment. Cox's time-dependent Proportional Hazards Model (PHM) is considered to model the equipment's failure rate. This paper addresses two main problems: entire problem of imperfect observations and the problem of taking into account the whole history of observations. Two numerical examples are presented.

Keywords: Condition Based Maintenance, Condition Monitoring, Mean Residual Life, Hidden Markov Model, Time-dependent Proportional Hazards Model.

3.2 Practical Implications

The conditional reliability function is a measure of equipment's performance indicating the probability of survival during a period of time, knowing that the equipment has not yet failed. This probability can be used to calculate the MRL which can be used in finding the optimal replacement policy for the equipment. If the probability is calculated while assuming that the equipment has not yet been put to work, it indicates the unconditional reliability of the equipment. When a condition monitoring system is available, analysts are interested in knowing the reliability based on the latest available information on the equipment's degradation state, i.e. the conditional reliability, while taking into consideration the information obtained from the condition monitoring system. This paper introduces a model that calculates the conditional reliability function and the MRL of a piece of equipment in the presence of condition monitoring data, where this data does not directly reveal the degradation state of the equipment, but discloses some information which is stochastically related to the equipment's degradation state.

3.3 Introduction

Condition Based Maintenance (CBM) is based on observing and collecting information concerning the condition of equipment, in order to prevent its failure and to determine maintenance actions. When a piece of equipment is subjected to CBM, data concerning one or more indicators of degradation are collected periodically. The information obtained from this data is used to establish a diagnosis of the equipment's condition and a prognosis for future performance. Two measures of this performance are the failure rate (or the hazard function), and the MRL. These two measures are calculated from the reliability function.

In reliability analysis, two reliability functions are of interest. The first is the unconditional reliability function given by the probability $P(T>t)$, which is the probability that the failure time *T*, of a piece of equipment that has not yet been put into operation, is bigger than a certain time *t*. The second is the conditional reliability function calculated by $P(T>t \mid T>\tau)$, which is the probability that the time to failure T is bigger than *t*, knowing that the equipment has already survived until time τ, where $\tau<t$. In this later case, the MRL is $E(T - \tau \mid T>\tau)$ (Jardine *et al.* 2006), which is equal to $\int_{\tau}^{\infty} \frac{1-F(t)}{1-F(\tau)} dt, \tau>0$, where F is the cumulative distribution function and T is having distribution function F. The hazard function $\lambda(\tau)$, is obtained from the equation $\lambda(\tau)\Delta\tau = P(\tau<T<\tau+\Delta\tau \mid T>\tau)$. In some reliability analysis, it is assumed that every piece of equipment is used in the same environment and under the same conditions. This assumption allows the calculation of the MRL and the hazard function prior to the actual use of the equipment. In real-life, the environment in which the equipment is performing and the conditions of utilization affect the process of degradation. Consequently, the conditional reliability, the residual life and the failure rate of the equipment are affected. Taking this fact into consideration improves the diagnosis of the equipment's degradation state and the prognosis for future performance.

Many researchers have proposed different reliability models incorporating the information gathered periodically regarding the equipment's observed condition. These models are used to calculate an adjusted hazard function and the corresponding MRL. One of these models is the Proportional Hazards Model (PHM), proposed by [Cox 1972]. This model has been widely used in the medical field [Crowley and Hu., 1977; Leemis, 1995], and in the field of CBM [Jardine *et al.,* 1985, 1987, 2001; Kumar and Westberg, 1996; Ansell and Phillips, 1997; Jozwaik, 1997]. The PHM has the advantage of improving maintenance decisions since it is based on a more accurate estimate of the hazard function and the MRL [Banjevic and Jardin, 2006].

In all previous applications of the PHM, it was assumed that the information collected regarding the equipment's condition, indicator θ, was a direct pointer to the equipment's degradation state Z, and that the indicator was in a *some-to-one* or *one-to-one* relation with the degradation state that influences the time to failure. In a some-to-one approach, each possible value of the indicator, in a predefined interval, refers to one degradation state. In this approach, there is no indicator value that can refer to more than one degradation state. If the condition monitoring reading is of value θ_i, the state will be a certain value Z_j. In this approach, each indicator value θ_i refers to only one state Z_j. At the same time, any state Z_j, may be referred to by several possible values of the indicator in a predefined interval [Makis and Jardine, 1992]. Figure 3-1 demonstrates the some-to-one characteristic used in this discussion. It can be seen in Figure 3-1 that any indicator value in the interval $[a,b\ )$, e.g. θ_1 and/or θ_3, refers to the same state value Z_1. It is not possible to have more than one state referred to by a certain indicator value.

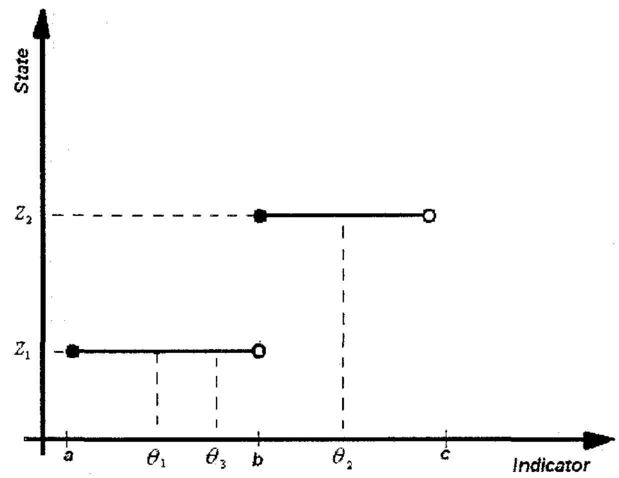

Figure 3-1: Direct observation for a system with 2 degradation states (some-to-one relationship)

In the one-to-one approach, the indicator value is assumed to be a direct pointer to the equipment's degradation state, and is used directly as the diagnostic covariate in the PHM [Kumar *et al.,* 1996]. In all the previous cases that have used the PHM, either the indicator reading was used directly as the diagnostic covariate in the PHM (one-to-one), or a transformation of the indicator into a new state space (some-to-one) is considered as the diagnostic covariate.

Realistically, information may contain noise due to errors of measurement, interpretation, accuracy of measurement instruments, etc, and may not reveal the exact degradation state of the equipment. The information is, however, stochastically correlated with the underlying state. In this case, information collected may be referring to more than one possible state. For example, a certain level of vibration (indicator) θ_1, may be read while the equipment is in any of two different levels of degradation states, Z_1 and Z_2. This situation is represented by a probability distribution function. In the latter case, the relationship between the collected condition monitoring information (indicator) and the state is *some-to-some*. One collected indicator value may refer to several degradation states and vice versa. This category of condition monitoring is referred to as Indirect Monitoring [Wang and Christer, 2000] or Partial Observation [Makis and Jiang, 2003]. Figure 3-2 illustrates the stochastic relationship between the indicator and the state in this case. As shown in Figure 3-2, the indicator value θ_1 may refer to either state Z_1 or state Z_2. An indicator value may be a sign of several possible degradation states, and a system in certain degradation states may demonstrate different indicator values.

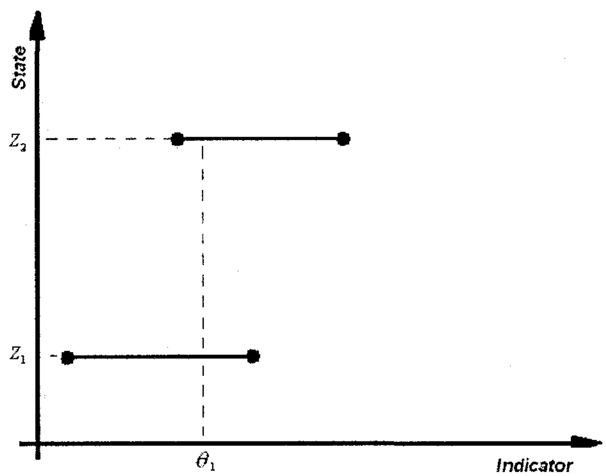

Figure 3-2: Indirect or imperfect observation (some-to-some relation)

The relationship between the indicator's values and the states is introduced via an observation probability matrix or a probability distribution. For example, in Figure 3-3, if the state is, $Z_i, i=1,2$, the probability of observing different values of the indicator follows a normal distribution $N\left(\mu_{X_i}, \sigma^2_{X_i}\right)$. In this paper, we consider a *some-to-some* indicator-state relation. The model is then used in order to propose a solution to the main drawback of the time-dependent PHM which is the inclusion of only the latest condition monitoring information in the model. In this paper, the proposed model considers the entire history of information obtained from the observation process.

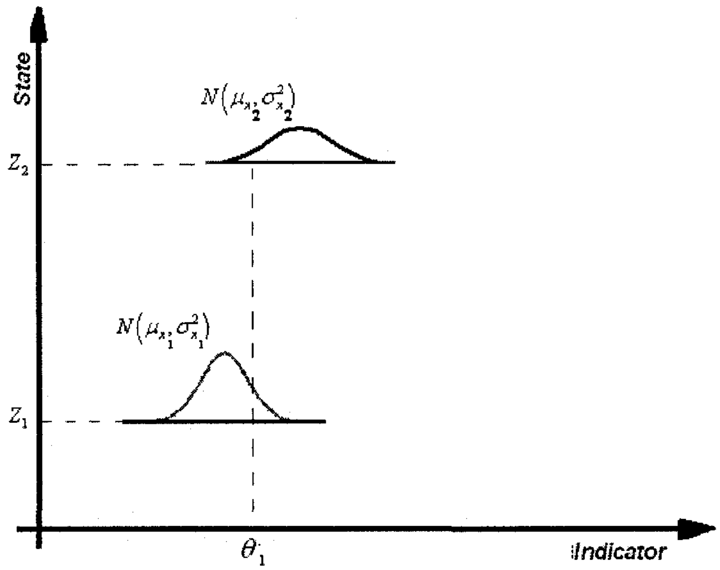

Figure 3-3: Probabilistic relationship between the indicator and state

To demonstrate the impact of not considering the entire history of information, Figure 3-4 depicts an explanatory example of a piece of equipment that follows the time-dependent PHM. In this example, the equipment represented by the bold line has been in the state $Z=1$ from time zero to time t_2. It can be seen that the failure rate of the equipment at times t_1 and t_2 are equal to h_1 and h_2 respectively. Now we consider the case demonstrated in Figure 3-5. From time zero to time t_1, the equipment state is $Z=0$ and right after, from time t_1 to time t_2, the state becomes $Z=1$. It can be seen that at time t_2, while the equipment is at state $Z=1$, the failure rate of the system is again h_2. The fact that the equipment was in state $Z=0$ from time zero to time t_1, had no effect on the value of the failure rate h_2 at t_2, which is obviously wrong.

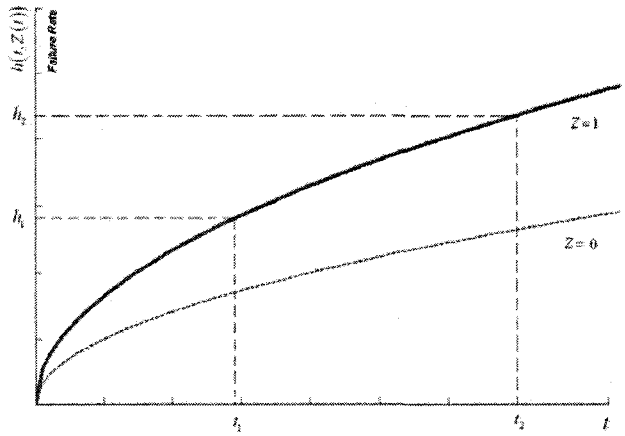

Figure 3-4: An explanatory example of a time-dependent PHM without a state change

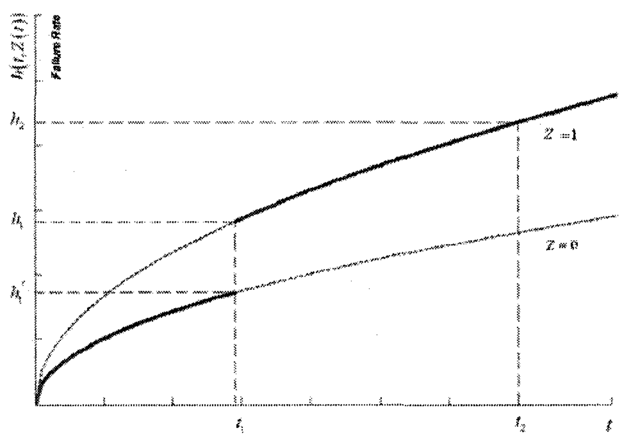

Figure 3-5: An explanatory example of a time-dependent PHM with state change

Conceptually, this drawback may be addressed as shown in Figure 3-6. In this figure, if the equipment is in state $Z=0$ until time t_1, the failure rate at that time is h_1. If at that time the state changes to $Z=1$, then the increase in the failure rate will follow the bold line from $h_1^{'}$ on the curve of $Z=1$. This is equivalent to assuming that the system was in state $Z=1$ from time zero but its age is t_1-t. Also after t_2-t_1, i.e. at age t_2, the failure rate will be $h_2^{'}$ and not h_2. These examples clearly show how the original approach of the time-dependent PHM gives a misleading value of the failure rate by ignoring the information history.

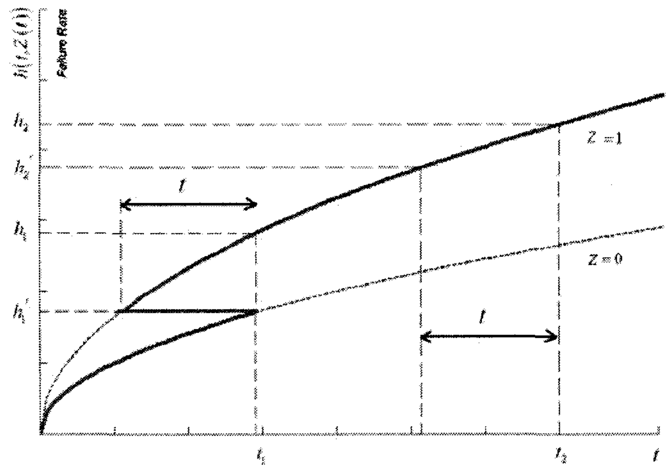

Figure 3-6: Conceptual solution for PHM drawback

This paper considers the case of indirect monitoring, i.e. the obtained indicator does not reveal the equipment's underlying degradation state. Instead, a probability of being in a certain state is calculated by using Bayes' rule. This rule considers all the previous history of information and incorporates it into the PHM. The proposed model thus overcomes the main drawback of the previously applied PHM. Based on this modification the conditional reliability and the MRL are calculated. Throughout this work "state" and "degradation state" are used interchangeably.

This paper is organized into four sections. Section 2 presents a literature review of the principal models used in the evaluation of the residual life. Section 3 introduces the proposed model that assumes the existence of imperfect observations. In section 4, numerical examples are presented. Conclusions and future researche are presented in section 5.

3.4 Literature review

Many researchers have studied the mathematical structure of the MRL based on reliability analysis without considering information concerning the actual use and the state of equipment. Tang *et al.* [1999] consider the residual life as a random variable and study its asymptotic behaviour when the reliability function is represented by various discrete and continuous distribution functions. Lim and Park [1995] study the monotonic behaviour of the residual life. They test the null hypothesis that the residual life is not monotone, against the alternative hypothesis that it is indeed monotone. Siddiqui and Caglar [1994] treat the residual life as a random variable and give a representation of its distribution function. When the distribution is Gamma or Weibull, the authors calculate the mean and the variance of the variable. Bradley and Gupta [2003] also study the asymptotic behaviour of the residual life.

Researchers that consider the presence of condition monitoring information use two main approaches to calculate the MRL; *recursive filtering* and PHM [Jardine *et al.*, 2006]. Recursive filtering is an approach in signal processing that extracts information (MRL) based on available signals (indicators), and previously extracted information [Byrne, 2005]. Wang and Christer [2000], Wang and Zhang [2005, 2008] and Wang [2002], among others, determined the MRL by applying a recursive filtering model. The MRL, given the condition monitoring history up to date, is obtained. The recursive filtering technique includes the entire observation history.

Wang and Christer [2000] assume that the observed condition monitoring indicator is a function of the underlying residual life and not vice versa, and use the indicator as the covariate. This assumption may not be realistic in many cases. For instance, the wear of a rotating shaft (which is reflected in oil particles as the observable indicator) affects its residual life, not vice versa. They use a recursive filter in order to calculate the MRL, and add to the existing MRL models the possibility of including all past information. Addressing the drawback of the approach in [Wang and Christer, 2000], Wang and Zhang [2005] introduce a new methodology that uses the difference between two consecutive observed indicators as the covariate and uses recursive filtering to determine the MRL. They define the MRL $e(t, Z(t))$ as the expected time interval between the last inspection, when the most recent information was gathered, and the expected time of failure, given that no maintenance action is taken in this interval. $e(t, Z(t))$ is thus equal to $E((T - t) \mid T > t,\ Z(t))$, where $Z(t)$ is the covariate (the difference of two consecutive readings of the indicator) at time t. Wang and Zhang [2008] model the MRL of the asset by considering the expert's judgment based on the equipment's observed indicator. In this case, the judgment is assumed to be a function of the residual life, which may include some noise.

The PHM may be more suitable in many cases, like oil analysis, since it assumes that the failure rate, and so the MRL, is a function of the degradation state or the observed indicator, which is representing the degradation state. But the PHM's drawback is that, it uses only the latest information of the condition monitoring system. Kumar and Westberg [1996] calculate the MRL using PHM when only the most recent information is available. Maguluri and Zhang [1994] are inspired by the PHM and calculate the proportional MRL by using the equation $e(t\mid Z)=\exp(-\beta' Z)e_0(t)$, where $e_0(t)= E(T-\tau\mid T>\tau)$ is the MRL calculated without including the covariate. Z is the vector of indicators used as the covariates in the model and β' is the vector of the covariates' coefficients. Sen [2004] calculates the conditional MRL given by $e(t\mid z)=\int_0^{\infty} e^{-\beta' z(\Lambda_0(t+\mu)-\Lambda_0(t))}d\mu$, where $\Lambda_0(t)=\int_0^t \lambda_0(u)du$, $\lambda_0(t)$ is the hazard function, and $\lambda(t\mid z)=\lambda_0(t)e^{\beta' z}$ is the proportional hazard function. Banjavic and Jardine [2006] calculate the joint distribution of time to failure and the diagnostic covariate $z(t)$ at time t, and the probability of transition between states $L_{ij}(\tau,t)= P\left(T>t, z(t)=j\mid T>\tau, z(\tau)=i\right)$, $\tau\le t$. $z(t)$ is the descritized observed information from the indicator into a new state space $\{0,1,\ldots,N\}$, which represents the degradation state of the equipment (some-to-one). The conditional reliability is thus given by the equation $R(t\mid\tau,z(\tau))=\sum_j L_{ij}(\tau,t)$, $\tau\le t$, and the MRL is $e(t,z(\tau))=\int_t^{\infty} R(\tau\mid t,z(t))d\tau$.

All of these models assume that the information gathered from the indicator likely reveals the equipment's exact state, or use the collected information directly as a diagnostic covariate that affects the failure rate. Moreover, some of them include only the most recently collected information.

In this paper, the residual life is modelled using the PHM, in the case of indirect condition monitoring, i.e. the equipment state is not deterministically known. We present a modified PHM model which takes into consideration the whole observations' history.

We assume that MRL is related to the equipment's degradation state, and the condition monitoring indicator is stochastically related to the equipment's underlying degradation state. This work not only addresses the problem of indirect observations, but also eliminates the main drawback of using the PHM model, which is the inclusion of only the latest information in the calculation of the MRL. Another drawback of the traditional time-dependent PHM is that the discretization of the indicator's possible values into classes of states (covariates) is very sensible to the lower and upper bounds selected. In this paper, since probability distribution is used to relate the indicator's values to each class of states, this problem is moderated. The model takes into consideration all previous information as well as the equipment's age and assumes that residual life is affected by the equipment's degradation state.

3.5 Model assumptions

Consider the following notations:

T	: Failure time of the equipment;
Δ	: Observation interval;
$Z(s)$	: Equipment's degradation state at time s;
Z_k	: Equipment's degradation state after k-th observation interval; Z_k = $Z(k\Delta)$;
θ	: Current value of the indirect indicator of the system's degradation state;
p_{ij}	: Probability of going from state i to state j during one inspection interval, knowing that the equipment has not failed during that interval. It is an element of the transition matrix *P;*
$q_{j\theta}$	: Probability of getting condition indicator value θ while the equipment is in state *j*. It is an element of the observation probability matrix *Q;*
$h(s,z)$	: The hazard function of the PHM at time s while the system state is

z;

$h_0(.)$: Baseline hazard function;

$\psi(.)$: State effect function;

$R(k, Z_k, t)$: Conditional reliability at period k for a period of t, knowing that the state is Z_k;

$\tau(k, Z_k, t)$: Conditional mean sojourn time at period k knowing that the state is Z_k;

π^k : Conditional probability distribution of the equipment's state at observation moment k, $k=0,1,...;$

π_i^k : Probability of being in state i at observation moment k, $k=0,1,...;$

$\overline{R}(k, \pi^k, t)$: Conditional reliability of the equipment for a period of t, at observation moment k while the conditional probability distribution of equipment's state is π^k;

$e(k, z)$: Mean residual life at observation moment k while the state is z;

$\overline{e}(k, \pi^k)$: Mean residual life at observation moment k while the conditional probability distribution of the equipment's state is π^k.

We consider the PHM proposed by Cox [1972], and we assume that the condition monitoring is indirect. Instead, an indirect indicator's value θ, of the underlying degradation state is available at each observation moment. Observations are collected at constant (or near constant) interval Δ. In this study, Z represents the degradation state of the equipment which will be used as the diagnostic covariate in the PHM, and θ is a value from the set of possible indicator values $\Theta = \{1, ..., M\}$. The whole set of indicator values is discretized into a finite set of M possible values. This assumption does not limit the scope of this work since its relaxation only entails the replacement of the observation probability matrix Q by a continuous probability distribution such as the normal distribution shown in Figure 3-3. The equipment's condition is described as follows:

- The equipment has a finite and known number of degradation states N. $J = \{1, ..., N\}$ is the set of all possible degradation states;
- Degradation state transition follows a Markov Chain with unobservable states and is modeled by a Hidden Markov Model (HMM). The transition matrix $P = \left[p_{ij}\right]$, is known or can be calculated. p_{ij} is the probability of going from state i to state j, $i, j \in J$ during one observation interval, knowing that the equipment has not failed before the end of the interval;
- The value of the indicator is stochastically related to the equipment's state through the observation probability matrix $Q = \left[q_{j\theta}\right]$, $j \in J$, $\theta \in \Theta$. $q_{j\theta}$ is the probability of getting indicator value θ, while the equipment is in state j;
- Failure is not a degradation state. It is a non-working condition of the equipment that can happen at any time and while in any degradation state, and is known immediately (obvious failure).

Figure 3-7 depicts the process of degradation and the transition from one degradation state to another, and from the degradation states to failure. The circles represent the states. State 1 is the best state (new or as new equipment). State N is the worst state, but it is not failure; the equipment is still working and fulfilling part of its role. It should be noted that failure can happen at any time and while the equipment is in any degradation state. T is a random variable showing the failure time and $1-r_i$ is the probability of going from state i to failure before the end of the interval. $r_i = R(k, i, \Delta)$, which is calculated for each observation moment k, is the conditional reliability of the equipment for a period of time Δ, while the equipment state is i, and is calculated by Equation (3-2) below .

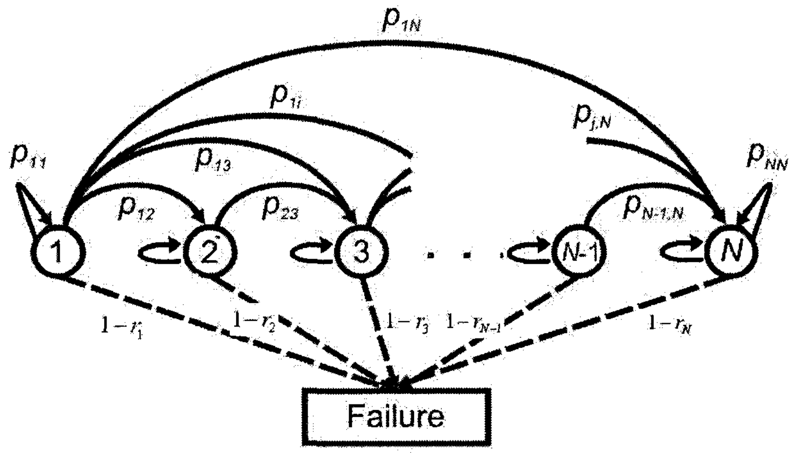

Figure 3-7: The process of degradation and failure

At time $s=0$, the equipment is always in state 1, which indicates that the equipment is in its best state. At fixed inspection interval Δ, i.e. at $s=\Delta, 2\Delta, \ldots$ an indicator of the equipment's state is observed. The indicator's value θ, is collected with a probability of $q_{j\theta}$ when the equipment is in degradation state j, $j \in J, \theta \in \Theta$. The values of $q_{j\theta}$ are assumed to be known. The transition matrix $P=\left[p_{ij}\right]$ $i, j \in J$ is assumed to be an upper triangular matrix, i.e. $p_{ij}=0$ for $j<i$, meaning that the degradation state cannot improve by itself, which is the case in most practical problems.

The parameter estimation problem of this model has been addressed in Ghasemi *et al.*, [2008]. So, it is assumed that all model parameters are known. In this paper, the objective is to derive the conditional reliability and the MRL when the condition monitoring is indirect.

3.5.1 Modeling the residual life

In the proposed model, the hazard function $h(s,Z_k)$ follows the PHM and is represented by the equation:

$$h(s,Z_k)=h_0(s)\psi(Z_k),\ k=0,1,2,\dots,k\Delta\le s<(k+1)\Delta \qquad \textbf{(3-1)}$$

where $h_0(.)$ is the hazard function of a Weibull distribution and represents the aging process, and $\psi(.)$ is a function of the equipment degradation state Z_k. The most used function is usually exponential in the form $\psi(Z_k)=\exp(\gamma Z_k)$. This means that the hazard function depends on the equipment's age and its state.

Since the observations are gathered at fixed intervals $\Delta, 2\Delta, \dots$ and the state is assumed to be invariable during each interval, the notation $Z_k = Z(k\Delta)$ is used. Each change of state is assumed to take place at the end of the interval, exactly before the observation moment. This assumption requires the observation interval to be short enough to include at most one transition during each interval. Having short enough intervals also supports the assumption of having the transition at the end of the interval just before the next observation moment. The choice of Δ depends on the nature of the equipment and the historical knowledge of its performance.

In the initial PHM, the conditional reliability is given by [Makis and Jardine, 1992]:

$$\begin{aligned}R(k,Z_k,t)&=P(T>k\Delta+t\mid T>k\Delta,Z_1,Z_2,\dots,Z_k),0<t\le\Delta\\&=P(T>k\Delta+t\mid T>k\Delta,Z_k),0<t\le\Delta\\&=\exp\left(-\psi(Z_k)\int_{k\Delta}^{k\Delta+t}h_0(s)ds)\right),0<t\le\Delta\end{aligned} \qquad \textbf{(3-2)}$$

The conditional reliability indicates the probability of survival until time $k\Delta + t$, $(0 < t \leq \Delta)$, knowing that the failure has not happened until time $k\Delta$, and the states of the equipment have been $Z_1, Z_2, ..., Z_k$, at $\Delta, 2\Delta, \ldots k\Delta$. T is the random variable indicating the time to failure. Also, the conditional mean sojourn time, if no action is performed before time $k\Delta + t$, while the equipment is in state Z_k at interval $k\Delta$, is [Makis and Jardine, 1992]:

$$\tau(k, Z_k, t) = \int_0^t R(k, Z_k, s)\, ds, 0 < t \leq \Delta \qquad \textbf{(3-3)}$$

Equations (3-2) and (3-3) are not valid for $t \geq \Delta$, since Z_k may change at any of the subsequent intervals. The conditional reliability at (k, Z_k), i.e. at the k-th observation moment while the state is Z_k and for $t > \Delta$, is formulated by the following equation:

$$\begin{aligned} R(k, Z_k, t) &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_1, Z_2, ..., Z_k\right), t > \Delta \\ &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_k\right), t > \Delta \end{aligned} \qquad \textbf{(3-4)}$$

Since $t > \Delta$, we continue the calculation of the reliability function by conditioning it on the survival until $(k+1)\Delta$, i.e. until the next observation moment (see Figure 3-8).

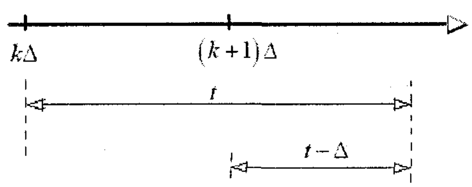

Figure 3-8: Demonstration of survival for $t > \Delta$

If the equipment survives for a period of $t; t > \Delta$ after $k\Delta$, it has to have survived until the next observation moment $(k+1)\Delta$, which may happen with a probability of

$\Pr\left(T>(k+1)\Delta \mid T>k\Delta, Z_k\right) = R(k, Z_k, \Delta)$. Just before the moment of $(k+1)\Delta$, the equipment state transfers from state Z_k to state Z_{k+1}. At $(k+1)\Delta$, the equipment has to survive for another period of $t-\Delta$ (equivalent of surviving for a period of t after $k\Delta$) which may happen with a probability of $\Pr\left(T>(k+1)\Delta+(t-\Delta) \mid T>(k+1)\Delta, Z_{k+1}\right)$. Assuming that the equipment will survive until $k+1^{\text{st}}$ observation moment and its state at $k+1^{\text{st}}$ is Z_{k+1}, we can conclude that its conditional reliability is:

$R(k, Z_k, \Delta)\Pr\left(T>(k+1)\Delta+(t-\Delta) \mid T>(k+1)\Delta, Z_{k+1}\right)$.

But, since Z_{k+1} can take any value $j=1,\ldots,N$ (N is the number of the possible states), with corresponding probability $p_{Z_k,j}$ in the Markov transition matrix, then:

$$R(k, Z_k, t) = \sum_{j=1}^{N} R(k, Z_k, \Delta) p_{Z_k,j} \Pr\left(T>(k+1)\Delta+(t-\Delta) \mid T>(k+1)\Delta, Z_{k+1}=j\right), t>\Delta$$

$$R(k, Z_k, t) = \sum_{j=1}^{N} \underbrace{R(k, Z_k, \Delta)}_{1} \underbrace{p_{Z_k,j}}_{2} \underbrace{\Pr\left(T>k\Delta+t \mid T>(k+1)\Delta, Z_{k+1}=j\right)}_{3}, t>\Delta \qquad \textbf{(3-5)}$$

"1" is the probability of survival until Δ, "2" is the probability of transition from state Z_k to state j at the next observation moment, and "3" is the probability of survival until $k\Delta+t$ while at the $k+1^{\text{st}}$ observation moment the state is j, which is equal to $R\left(k+1, j, (t-\Delta)\right)$. Equation (3-5) can be written as follows:

$$R(k, Z_k, t) = R(k, Z_k, \Delta)\sum_{j=1}^{N} p_{Z_k,j} R\left(k+1, j, (t-\Delta)\right), t>\Delta \qquad \textbf{(3-6)}$$

Considering equations (3-2) and (3-6), the conditional reliability at the k^{th} observation moment, while the state is $Z_k = i$ and no action is taken is:

$$R(k,i,t)=\begin{cases}\exp\left(-\psi(i)\int_{k\Delta}^{k\Delta+t}h_0(s)ds\right) & 0<t\le\Delta\\ R(k,i,\Delta)\sum_{j=1}^{N}p_{ij}R\left(k+1,j,(t-\Delta)\right) & t>\Delta\end{cases} \qquad \textbf{(3-7)}$$

In this paper, the degradation state is not directly observable, but an indicator of the degradation state is observed, and its value θ is recorded. A new state space (conditional probability distribution of the equipment's degradation state at period k, π^k), and a new transition rule are introduced in equations 8 to 10. π^k includes all the indicator's observations from the last renewal point up to the k-th observation moment, and provides a methodology to deal with unobservable states by calculating the conditional probability π_i^k, the probability of being in state i at time $k\Delta$. π^k, the conditional probability distribution of the equipment's degradation state at period k, is defined as follows:

$$\pi^k=\left\{\pi_i^k;\ \ 0\le\pi_i^k\le 1 \text{ for } i=1,\ldots,N;\sum_{i=1}^{N}\pi_i^k=1\right\},\ k=0,1,2,\ldots \qquad \textbf{(3-8)}$$

and since new or as new equipment is always in state1:

$$\pi_i^0=\begin{cases}1 & i=1\\ 0 & 1<i\le N\end{cases} \qquad \textbf{(3-9)}$$

After an indicator value θ is collected at the $k+1^{\text{st}}$ observation moment, the prior conditional probability distribution π^k, is updated to π^{k+1}. By using Bayes' formula, and knowing that the indicator value θ has been read at the $k+1^{\text{st}}$ observation moment, $\pi_j^{k+1}(\theta)$, the probability of being in state j at $k+1^{\text{st}}$ observation moment is updated:

$$\pi_j^{k+1}(\theta) = \frac{\sum_{i=1}^{N} \pi_i^k p_{ij} q_{j\theta}}{\sum_{i=1}^{N} \sum_{l=1}^{N} \pi_i^k p_{il} q_{l\theta}}, \; j = 1,\ldots,N \quad \textbf{(3-10)}$$

Since at observation moment $k+1$, the calculation of π^{k+1} is based on π^k and the latest value of the θ observed at the $k+1$ st observation moment, the updated conditional distribution π^{k+1}, carries the history of all the indicator's values from the last replacement point. After any preventive or failure replacement, the period counter will be reset to zero and the conditional probability distribution of the equipment state will be set to π^0 using equation (3-9).

In the case of indirect information, we define $\overline{R}(k,\pi^k,t)$ as the conditional reliability of the equipment at the k-th observation moment, while the state conditional probability distribution is π^k. It is calculated as follows:

$$\overline{R}(k,\pi^k,t) = \Pr(T > k\Delta + t \mid T > k\Delta, (k,\pi^k)) = \sum_{i=1}^{N} R(k,i,t)\pi_i^k \quad \textbf{(3-11)}$$

By substituting equation (3-7) into equation (3-11) we get:

$$\overline{R}(k,\pi^k,t) = \begin{cases} \sum_{i=1}^{N} \pi_i^k \exp\left(-\psi(i) \int_{k\Delta}^{k\Delta+t} h_0(s)ds\right) & 0 < t \le \Delta \\ \sum_{i=1}^{N} \pi_i^k R(k,i,\Delta) \sum_{j=1}^{N} p_{ij} \times R\left(k+1, j, (t-\Delta)\right) & t > \Delta \end{cases} \quad \textbf{(3-12)}$$

In the case of direct observations, using PHM, the MRL is given by [Banjevic and Jardine, 2006] as follows:

$$
\begin{aligned}
e(k,i) &= E\left(T - k\Delta \mid T > k\Delta, X_k = i\right) \\
&= \int_0^{\infty} R\left(k,i,k\Delta + t\right) dt \\
&= \int_{k\Delta}^{\infty} R\left(k,i,t\right) dt
\end{aligned}
\tag{3-13}
$$

For indirect observations, we define the MRL, $\bar{e}\left(k,\pi^k\right)$, calculated at the k^{th} observation moment, while the state conditional probability is π^k, as follows:

$$
\begin{aligned}
\bar{e}\left(k,\pi^k\right) &= E\left(T - k\Delta \mid \left(k,\pi^k\right)\right) \\
&= \sum_{i=1}^{N} \pi_i^k e\left(k,i\right) \\
&= \sum_{i=1}^{N} \pi_i^k \int_{k\Delta}^{\infty} R\left(k,i,t\right) dt
\end{aligned}
$$

$$
\begin{aligned}
\bar{e}\left(k,\pi^k\right) &= \sum_{i=1}^{N} \int_{k\Delta}^{\infty} \pi_i^k R\left(k,i,t\right) dt \\
&= \int_{k\Delta}^{\infty} \left(\sum_{i=1}^{N} \pi_i^k R\left(k,i,t\right) \right) dt
\end{aligned}
$$

$$
\bar{e}\left(k,\pi^k\right) = \int_{k\Delta}^{\infty} \bar{R}\left(k,\pi^k,t\right) dt \tag{3-14}
$$

where $\bar{R}\left(k,\pi^k,t\right)$ is calculated by equation (3-12).

The steps for calculating the MRL at each observation moment k, where the indicator obtained is θ, are as follows:

- At any observation moment k, when an indicator value θ is obtained; calculate the conditional probability distribution π^k, at period k by using equations (3-10);
- Calculate the conditional reliability of the equipment, $\bar{R}(k,\pi^k,t)$, at the k-th observation moment by using equation (3-12);

- Calculate the MRL of the equipment, $\bar{e}\left(k,\pi^{k}\right)$, by applying equation (3-14).

3.6 Numerical examples

The first example is adopted from [Ghasemi *et al.*, 2007]. The hazard function $h_0(s)$, representing the aging process, follows a Weibull distribution, and the equipment condition $\psi(Z_k)$, is given in an exponential form as follows:

$$h_0(s) = \frac{\beta s^{\beta-1}}{\alpha^{\beta}}, s \geq 0\ , \alpha = 1, \beta = 2$$

$$\psi\left(Z_k\right) = e^{0.5\left(Z_k-1\right)}$$

For $\Delta = 1$, the equipment's hazard function and its conditional reliability from equations (3-1) and (3-2) are as follows:

$$h(s, Z_k) = 2se^{0.5\left(Z_k-1\right)}, k\Delta \leq s < \left(k+1\right)\Delta$$

$$R(k, Z_k, t) = \exp[-(t^2 + 2tk)e^{0.5\left(Z_k-1\right)}]$$

In addition to the obvious failure, the equipment can be in any of two unobservable states $\{1, 2\}$. 1 is the new or as new state. The transition matrix P is given as follows:

$$P = \begin{array}{c} \\ \text{State1} \\ \text{State2} \end{array} \begin{array}{c} \begin{array}{cc} \text{State1} & \text{State2} \end{array} \\ \begin{bmatrix} 0.4 & 0.6 \\ 0 & 1 \end{bmatrix} \end{array}$$

The indicator value θ can take the values of: Excellent (1), Normal (2), or Bad (3). The probabilities of observing one of these three values, while the equipment is in state 1 or 2 are given by the matrix Q as follows:

$$\begin{array}{rccc} & \text{Excellent} & \text{Normal} & \text{Bad} \\ Q = \text{State1} & & & \\ \text{State2} & & & \end{array}\begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$$

For example, the probability of finding the indicator *Excellent*, while the equipment is actually in state 1, is 0.6. Based on the developed model, and by assuming that the equipment's state is known, the MRL at different observation moments $k = 0,1,2,3,4$ is calculated from equation (3-13) and the results are shown in Table 3-1. The value of the MRL at $k = 0$, while the equipment is in state 2, is not applicable (N/A), since we assume that new equipment is always in state 1. According to the calculations, new equipment has an MRL equal to 1.62.

Table 3-1: Mean Residual Life at different observation moments (direct observation)

Period	State	MRL (time unit)
0	1	1.62
	2	N/A
1	1	0.84
	2	0.53
2	1	0.40
	2	0.23
3	1	0.18
	2	0.10
4	1	0.08
	2	0.05

If the equipment's state is unobservable, after the collection of the indicator's value θ, the probability of being in state $i = 1,2$ is calculated from equation (3-10), then used in equations (3-11) to (3-14) to obtain the MRL. Since we are dealing with an example and we do not have in hand the values of the indicator, we calculate the MRL for all possible values of θ, and consequently of π_2^k at each of the next 4 observation moments. For example, at $k=1$, the corresponding values of π_2^1 is $0.3333, 0.6667$ or 0.8571 if the readings of θ is 1, 2, or 3, respectively. π_2^k is the probability of being in state 2 while the equipment is at the k[th] observation moment. This means that in a real situation, at $k=1$, if we receive an observation $\theta=2$ from the equipment, and based on equation (3-10), we get a value of $\pi_2^k=0.6667$, then the MRL of the equipment is 0.63 (See Figure 3-9)

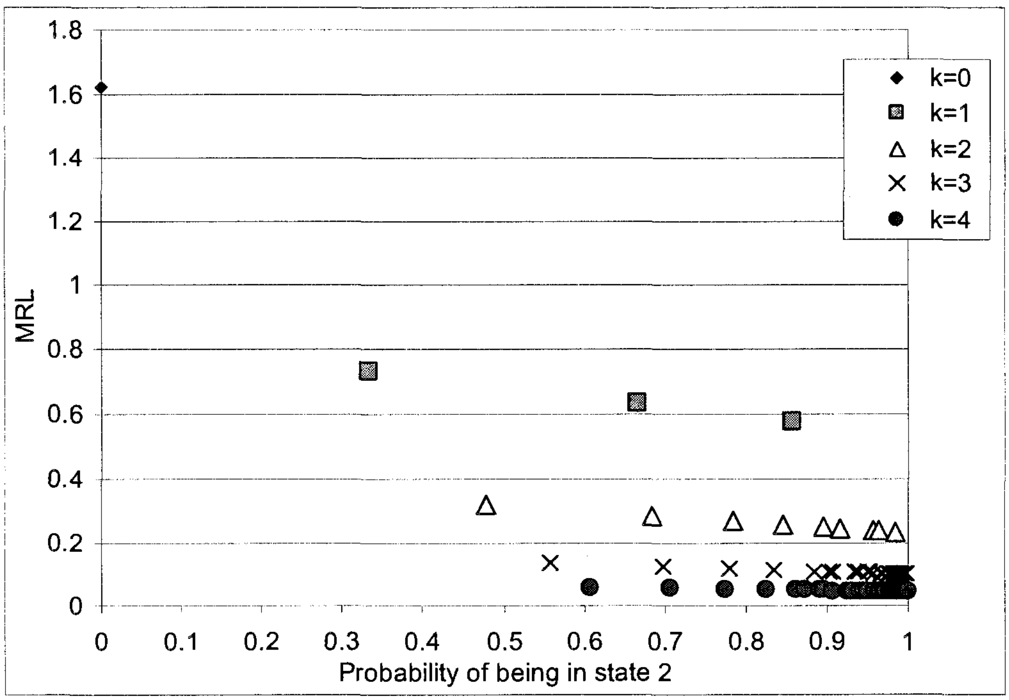

Figure 3-9: Mean Residual Life for all possible π_2^k at $k=0,1,\ldots,4$ observation moments

To explain the relationship between the optimal replacement policy that was obtained in [Ghasemi *et al.,* 2007], and the MRL, we recall that in that paper, the cost of preventive replacement is $C=5$, while the cost of replacement after failure increases by $K=2$.

The optimal replacement criterion is given as a function of the reliability and the conditional mean sojourn in the following observation period as follows:

$$T_{g^*} = \inf\left\{\ k \geq 0\ ;\ \ 2\times\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right] \geq 8.1704\times\overline{\tau}\left(k,\pi^k,\Delta\right)\ \right\} \qquad \textbf{(3-15)}$$

where $g^* = 8.1704$ is the long run average cost of replacement. Figure 3-10 shows the decision criterion for this example. The straight line indicates the threshold value of $g^*/K = 8.1704/2$ found in [Ghasemi *et al.,* 2007]. It can be seen that independent of the value of π_2^k, it is never cost optimal to replace after the first interval. Similarly, after the second interval, the equipment should optimally be replaced regardless of the value of π_2^k.

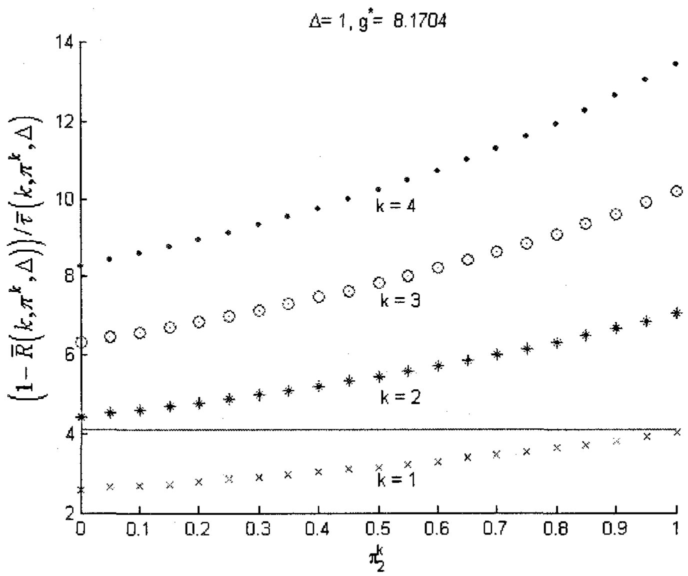

Figure 3-10: Decision Criterion for $\Delta = 1$, K=2 and C=5

This means that after two periods of utilization, the equipment should be replaced regardless of its state. This decision takes into consideration the replacement costs, as well as the value of the conditional reliability. Another way of decision-making will be to consider the value of the MRL. For example, in Figure 3-10, the decision based on the cost and the value of the conditional reliability is to never replace at $k=1$ i.e. at the first observation moment. However, by considering the value of the MRL given in Figure 3-9, which depends on the observation collected at the end of this period, a practitioner may decide differently. For example, if the updated π_2^1 is equal to 0.8571, then the MRL is 0.57, which is about half the length of an interval. The decision maker may then decide to replace the equipment, although it is not the cost optimal decision. The decision can aim to prevent a potential interruption during the next period, or may consider scheduling a replacement at a convenient moment before the next observation moment in a manner that has minimum influence on the equipment's mission. This decision is not based on cost considerations and is basically considering the availability of the equipment.

To further explain this criterion, we assume that $C=5$ and $K=4$, which will result in $g^*=10.17$ and the threshold line will shift to $g^*/K=2.54$, as shown in Figure 3-11. This shift means that if at the first observation moment the calculated π_2^1 is larger than or equal to 0.35, the equipment should be replaced; otherwise it should be replaced at the next observation moment. From Figure 3-9, if the indicator's reading at $k=1$ is normal (2) or bad (3) then π_2^1 is equal to 0.6667 or 0.8571 respectively and the MRL is either 0.63 or 0.57. Considering that the cost of a failure in this case is twice as much as in the previous example, it is obvious why the optimum replacement policy is to replace the equipment, even if in these cases, where $K=4$, the MRLs are higher or equal to that in the previous example ($K=2$), where the optimum decision was not to replace the equipment at the first observation moment. Again, here a practitioner may decide to go with a different decision than what the optimal replacement criteria suggests, in order to address other priorities in the organization, and not just to consider the cost. This gives another indication to the decision-makers as to whether they should replace the equipment or not.

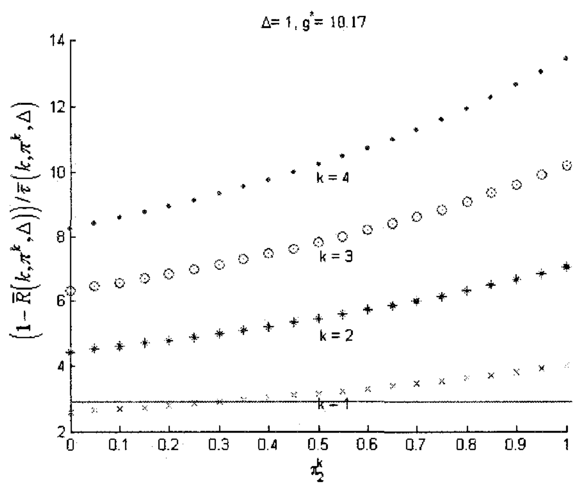

Figure 3-11: Decision Criterion for $\Delta = 1$, *K*=4 and *C*=5

The results of another example is shown in Figure 3-12. We consider a piece of equipment which may be in one of three degradation states $i = 1,2,3$ (the failure is not included), and the observation can again take any of three values, excellent (1), normal (2), or bad (3). The transition matrix is $P = \begin{bmatrix} 0.9 & 0.1 & 0.0 \\ 0.0 & 0.9 & 0.1 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$, the information matrix is $Q = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.5 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$, the Weibull distribution parameters are $\beta = 3$, $\alpha = 3$, $a = 0.5$, and $\Delta = 1, C = 5, K = 2$. As for the previous example, based on the optimality condition that minimizes the cost of replacement in a renewal cycle, the decision is not to replace after the first interval and to replace after the second interval $k = 2$, if the corresponding point defined by $\left(\pi_1^k,\ \pi_2^k\right)$ on the surface is above the replacement threshold surface as shown in Figure 3-12. Another criterion for decision-making is the value of the MRL given in Table 3-2.

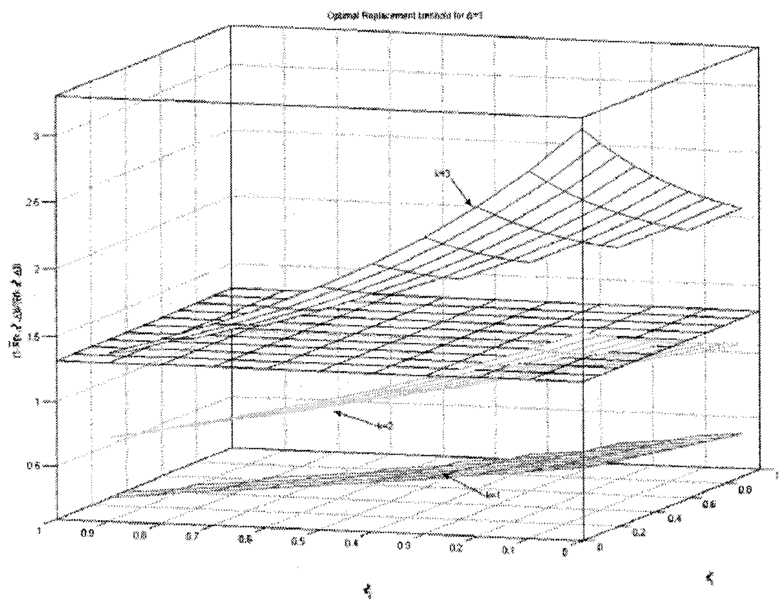

Figure 3-12: Decision criterion for equipment with three states

Table 3-2: Mean residual Life of the equipment for all possible $\left(\pi_1^k, \pi_2^k, \pi_3^k\right)$ at $k = 0,1,2,3$

k	0	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3
π_1^k	1.00	0.96	0.90	0.78	0.94	0.92	0.87	0.82	0.72	0.71	0.62	0.47	0.94	0.93	0.90	0.88
π_2^k	0.00	0.04	0.10	0.22	0.06	0.08	0.13	0.18	0.28	0.28	0.34	0.46	0.06	0.07	0.10	0.12
π_3^k	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.04	0.07	0.00	0.00	0.00	0.00
MRL	**1.88**	**1.07**	**1.04**	**0.97**	**0.61**	**0.60**	**0.57**	**0.55**	**0.51**	**0.50**	**0.46**	**0.39**	**0.34**	**0.34**	**0.33**	**0.32**

k	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
π_1^k	0.85	0.83	0.79	0.75	0.70	0.68	0.67	0.65	0.60	0.57	0.51	0.48	0.41	0.38	0.30	0.20
π_2^k	0.15	0.17	0.21	0.25	0.30	0.30	0.33	0.32	0.40	0.38	0.43	0.52	0.51	0.48	0.50	0.53
π_3^k	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.03	0.00	0.05	0.06	0.00	0.09	0.14	0.20	0.27
MRL	**0.32**	**0.31**	**0.30**	**0.29**	**0.27**	**0.27**	**0.27**	**0.26**	**0.25**	**0.24**	**0.22**	**0.21**	**0.19**	**0.18**	**0.15**	**0.12**

In Table 3-2, the values of $\left(\pi_1^k, \pi_2^k, \pi_3^k\right)$ at $k = 0,1,2,3$ indicate all possible values for the $\pi^k = \left(\pi_1^k, \pi_2^k, \pi_3^k\right)$. For example, since at $k = 0$, there is only one possibility for π^0, i.e. $\pi^0 = (1,0,0)$, then at $k = 1$, for each possible observation (Excellent (1), Normal (2), Bad (3)), we will have one possible value of $\left(\pi_1^k, \pi_2^k, \pi_3^k\right)$ as shown in Table 3-2. There are less than 9 (3×3) incidents for $k = 2$. That is because different values of π^1 have produced the same values of π^2 when taking into account the different indicator's readings. This is also true for $k = 3$.

3.7 Summary & Conclusion

In most published papers that use the PHM and Markov model in CBM, the MRL is calculated when the observations are direct or when they are a transformation (some-to-one) of the indicator's value that is used as a diagnostic covariate that influences the time to failure directly. However, in this paper, the MRL is modelled and calculated for equipment with indirect observations and obvious failure. The model is based on the PHM with time to failure following a Weibull distribution and the equipment's condition represented by an exponential function. It is assumed that the failure rate is a function of the equipment's degradation state, and we have taken into account that the observed indicator is an indirect pointer to the equipment degradation state and that it reveals some stochastic information about the underlying state.

The conditional reliability is derived from the PHM and used to calculate the MRL. Two examples are presented. The cost optimal replacement policy and the MRL are calculated at all possible state probabilities for four observation moments. It has been shown that the MRL can be used as a supplementary decision tool, in particular when the cost elements of preventive replacement are unknown, or there are criteria other than the cost to respect.

A practitioner equipped with an MRL result, may take advantage of the upcoming events (like an upcoming shutdown of a production line) that are not usually considered in cost optimal replacement criteria, to perform a CBM and to improve the availability of the equipment.

3.8 References

1. Ansell, J. I. and Phillips, M. J. Practical aspects of modeling of repairable systems data using proportional hazard models. Reliability Engineering and System Safety. 1997; 58:165-71.

2. Banjevic, D. and Jardine, A. K. S. Calculation of reliability function and remaining useful life. IMA Journal of Management Mathematics. 2006; 17:115-130.

3. Bradley, David M. and Gupta, Ramesh C. Limiting behavior of The Mean Residual Life. Ann. Inst. Statist. Math. 2003; 55(1):217-226.

4. Byrne, Charles L. Signal Processing : a mathematical approach. Wellesley, Mass.: A.K. Peters; 2005.

5. Cox, David Roxbee. Regression models and life tables (with discussion). Journal of the Royal Statistical Society - Series B. 1972; 26(187-220).

6. Crowley, J and Hu, M. Covariance analysis of heart transplant survival data. Journal of American Statistical Association; 72(357):27-36.

7. Ghasemi, A.; Yacout, S., and Ouali, M. S. Optimal condition based maintenance with imperfect information and the proportional hazards model . International Journal of Production Research. 2007; 45(4):989-1112.

8. Ghasemi, A.; Yacout S. , and Ouali M.S. Optimal inspection period and replacement policy for CBM with imperfect information using PHM. Current Themes in Engineering Technologies. Melville, NY: American Institute of Physics; 2008; pp. 247-266.

9. Ghasemi, A.; Yacout S. , and Ouali M.S. Parameter Estimation algorithms of a Condition Based Maintenance with Imperfect Information and Proportional Hazard Model. IIE Annual Conference and Expo ; , Vancouver, Canada. 2008 May.

10. Hontelez, A. M.; Burger, Helen H., and Wijnmalen, Diederik J. D. Optimum condition-based maintenance policies for deteriorating systems with partial information. Reliability Engineering and System Safety. 1996; 51:267-274.

11. Jardine A. K. S. and Anderson M. Use of concomitant variables for reliability estimation. Maintenance Management International. 1985; 5(2):135-140.

12. Jardine, A. K. S.; Anderson, P. M., and Mann D. S. Application of the Weibull

Proportional Hazards Model to Aircraft and Marine Engine Failure Data. Quality and Reliability Engineering International. 1987; 3:77-82.

13. Jardine, A. K. S.; Banjevic, D.; Wiseman, M.; Buck, S., and Joseph, T. Optimizing a mine haul truck wheel motors' condition monitoring program. Journal of Quality in Maintenance Engineering. 2001; 7:286-301.

14. Jardine A.K.S. ; Lin, Daming, and Banjvic, Dargan. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing. 2006; 20:1483-1510.

15. Jozwiak, I. J. introduction to the studies of reliability of systems using the Weibull proportional hazards model. Microelectronics and Reliability. 1997; 37:915-18.

16. Kumar, Dhananjay and Klefsjo, Bengt. Proportional hazards model : a review. Reliability Engineering and System Safety. 1994; 44:177-188.

17. Kumar, Dhananjay and Westberg, Ulf. Proportional Hazards Modeling of Time-Dependent Covariates Using Linear Regression: A Case Study. IEEE Transactions on Reliability . 1996; 45(3):386-392.

18. Leemis, L. M. Reliability- Probabilistic Models and Strategical Methods. New Jersey: Prentice all, Inc., Engelwood Cliffs; 1995.

19. Lim, Jea-Hak and Park, Dong Ho. Trend Change in Mean Residual Life. IEEE Transactions on Reliability . 1995; 44(2):291-296.

20. Maguluri, G and Zhang, C. H. Estimation in mean residual life regression model. J. R. Statist. Ser. B. 1994; 56:477-489.

21. Maillart, L. M. Optimal condition-Monitoring Schedules for Multi-State Deterioration Systems with Obvious Failure. 2004; Technical Memorandum Number 778.

22. Makis, V. and Jardine, A. K. S. Computation of optimal policies in replacement models. IMA Journal of Mathematics Applied in Business & Industry. 1992; 3(169-175).

23. Makis, V. and Jiang X. Optimal Replacement Under Partial Observations. Mathematics of Operations Research. 2003; 28(2):382-394.

24. Sen, P. K. HQ|RQoL and concomitant adjusted mean residual life analysis. Nikulin, M. S and Balakrishnan, N. Parametric and semiparametric models with applications to reliability, survival analysis, and quality of life. Boston, Birkhauser; 2004; pp. 349-362.

25. Siddiqui, M. M. and Caglar, M. Residual Lifetime Distribution and Its Applications. Microelectronics Reliability. 1994; 34(2):211-227.

26. Tang, L. C; Lu, Y., and Chew, E. P. Mean Residual Life of Lifetime Distributions. IEEE Transactions on Reliability . 1999; 48(1):73-78.

27. Wang, w and Christer, A. H. Towards a general condition based maintenance model for a stochastic dynamic system. Journal of the Operational Research Society. 2000; 51:145-155.

28. Wang W and Zhang W. An asset residual life prediction model based on expert judgment. European Journal of Operational Research. 2008; 188:496-505.

29. Wang, Wenbin . A model to predict the residual life of rolling element bearing given monitored condition information to date. IMA Journal of Management Mathematics. 2002; 13:3-16.

30. Wang, Wenbin and Zhang, Wenjuan. A model to predict the residual life of aircraft engines based upon oil analysis data. Naval Research Logistics. 2005; 52(276-284).

CHAPTER 4 :

DEVELOPMENT OF PARAMETER ESTIMATION METHODS FOR A CONDITION BASED MAINTENANCE WITH INDIRECT OBSERVATIONS

Alireza Ghasemi, Soumay Yacout, M-Salah Ouali

Submitted to: IEEE Transactions on Reliability, TR number: TR2008-240

4.1 Abstract:

This article proposes methods to estimate the parameters of a Condition Based Maintenance model when the equipment's failure rate follows the Cox's time-dependent Proportional Hazards Model. Due to errors of measurement, interpretation, accuracy of measurement instruments, etc., the observation process is not perfect and doesn't directly reveal the exact degradation state. At each observation moment, an indicator of the underlying unobservable degradation state is observed and the monitoring information is collected. In order to match indicator's value to the unobservable degradation state, the stochastic relation between them is given by an observation probability matrix. In this study we consider the case of imperfect observations and also we assume that the equipment's unobservable degradation state transition follows a Hidden Markov Model. We determine the Probability Density Function of the time to failure and use the Maximum Likelihood Estimation to estimate the model's parameters. The cases of censored and uncensored data are studied. Simulation studies are carried out to test the accuracy and the convergence of the methods.

Keywords: Parameter Estimation, Maximum Likelihood Estimation, Condition Based Maintenance, Condition Monitoring, Hidden Markov Model, Time-dependent Proportional Hazards Model.

4.2 Introduction

Condition Based Maintenance (CBM) is based on observing and collecting information concerning the condition of an equipment, in order to prevent its failure and to determine maintenance actions. When a piece of equipment is subjected to CBM, data concerning one or more indicators of degradation are collected periodically. The information obtained from this data is used to establish a diagnosis of equipment's condition and a prognosis for future performance. The diagnosis and prognosis processes are based on mathematical models which contain several parameters. In order to apply any diagnosis and/or prognosis method on a real world problem, the parameters must be estimated from the available data. In this paper we address the parameter estimation problem of a CBM system where the failure rate of the equipment is assumed to follow a time-dependent Proportional Hazards Model (PHM) and its unobservable degradation state is modeled by a Hidden Markov Model (HMM). The relation between the indicator's values and the degradation states is assumed to be modeled by a stochastic matrix.

A wide range of parameter estimation methods for CBM models incorporating the information gathered periodically regarding the equipment's observed condition exist in the literature. Some of these models uses the (PHM), proposed by D. R. Cox [1972]. This model has been widely used in the medical field [Crowley and Hu., 1977; Leemis, 1995], and in the field of CBM [Jardine *et al.,* 1985, 1987, 2001; Kumar and Westberg, 1996; Ansell and Phillips, 1997; Jozwaik, 1997]. In all previous applications of the PHM, it was assumed that the information collected regarding the equipment's condition, that is the indicator θ, is a direct pointer to the equipment's degradation state Z. They assumed that the indicator is in a *some-to-one* or *one-to-one* relationship with the degradation state. In a *some-to-one* relationship, any state Z_i may be referred to by several possible values of the indicator in a predefined interval [e.g. Makis and Jardine, 1992]. Figure 4-1 demonstrates the some-to-one relationship. It can be seen that any indicator value in the interval $[a,b)$, e.g. θ_1 and/or θ_3, refers to the same state value Z_1. It is not possible to have more than one state referred to by the same indicator's value.

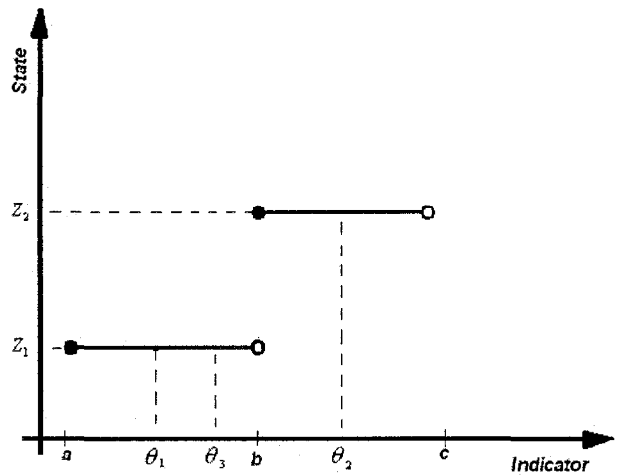

Figure 4-1: Direct observation for equipment with 2 degradation states (some-to-one relationship)

In a *one-to-one* approach, each possible indicator's value θ_i refers to one degradation state Z_i. The indicator value is assumed to be a direct pointer to the equipment's degradation state, and it is used directly as the diagnostic covariate in the PHM [e.g. Kumar *et al.,* 1996]. In both cases, there is no indicator's value that can refer to more than one degradation state. Either the indicator reading has been used directly as the diagnostic covariate in the PHM (one-to-one) or a transformation of the indicator value into a new state space (some-to-one) is considered as the diagnostic covariate. Nevertheless, in both cases, a certain value of the indicator refers deterministically to a certain degradation state.

Realistically, information may contain noise due to errors of measurement, interpretations, accuracy of measurement instruments, etc and may not reveal the exact degradation state of the equipment. The information is, however, stochastically correlated with the degradation state. In this case, information collected may be referring to more than one possible state with different probabilities. For example, a certain level of vibration (indicator) θ_1, may be read while the equipment is in any of two different levels of degradation states Z_1 and Z_2. This situation is represented by a probability distribution function or a stochastic matrix. In the latter case, the relation between the collected condition monitoring information (indicator) and the state is *some-to-some*. One collected indicator value may refer to several degradation states and vice versa. This category of condition monitoring is referred to as Indirect Monitoring [Wang and Christer, 2000] or Partial Observation [Makis and Jiang, 2003]. Figure 4-2 illustrates the stochastic relationship between the indicator and the state for this case. As shown in Figure 4-2, the indicator value θ_1 may refer to either state Z_1 or state Z_2. An indicator value may be a sign of several possible degradation states, and a piece of equipment, in certain degradation state, may demonstrate different indicator's values.

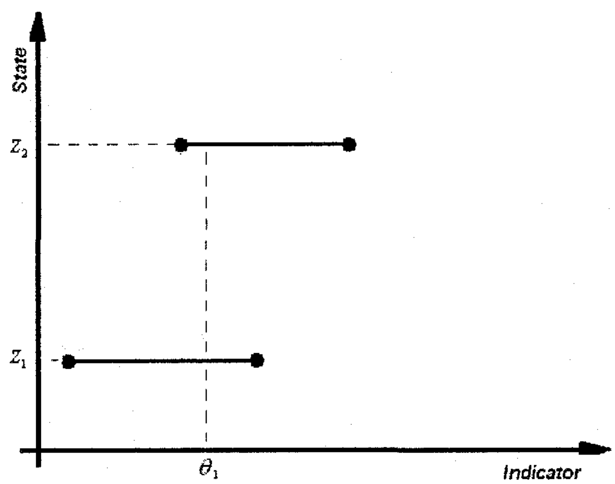

Figure 4-2: Indirect or imperfect observation (some-to-some relation)

The probabilistic relationship between the indicator's values and the states may be introduced via an observation probability matrix or a probability distribution. For example, in Figure 4-3, if the state is $Z_i, i=1,2$, the probability of observing different values of the indicator follows a normal distribution $N\left(\mu_{X_i}, \sigma^2_{X_i}\right)$. In this paper, we consider a some-to-some indicator-state relation.

Combination of PHM and HMM was originally proposed by Ghasemi *et al.,* [2007] to address the imperfect observation problem and to propose a solution to the main drawback of the time-dependent PHM, which is the inclusion of only the latest condition monitoring information in the model. For more details on the impact of not considering all the history of information, please refer to [Ghasemi *et al. 2008*].

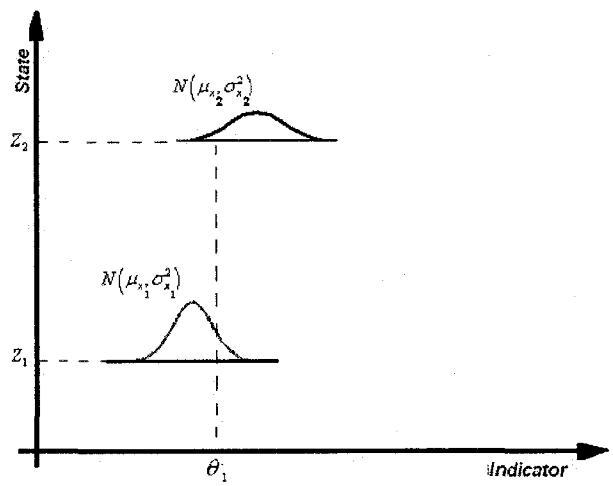

Figure 4-3: Probabilistic relation between the indicator and state

We consider the parameter estimation of a CBM system where the degradation state of the equipment is unobservable and modeled by a HMM and its failure rate follows Cox's time-dependent PHM.

This paper is organized into four sections. Section 2 presents a literature review of the principal methods used in parameter estimation of a PHM, and more specifically PHM used in CBM. Section 3 introduces the proposed model that assumes the existence of imperfect observations. In section 4, the parameter estimation algorithms and simulation studies are represented. Conclusions and future researches are presented in section 5.

4.3 Literature review

Parameter estimation of CBM models using PHM has been considered by the researchers in two categories: perfect and imperfect observations.

For perfect observations, Jardine *et al.* [1987] incorporated indicators (diagnostic covariates) affecting the equipment's time-to-failure into a fully parametric Weibull PHM and estimates the model's parameters based on Maximum Likelihood Estimation (MLE). Elsayed *et al.* [1990] developed PHM to estimate thin-oxide dielectric reliability by applying the partial likelihood method. Banjevic *et al.* [2001] estimated the parameters of a PHM used in the analysis of a piece of equipment that follows a Markovian degradation. A parametric PHM with Weibull baseline hazard function was considered and its parameters were estimated by MLE method. The method of MLE is also used to estimate the transition probabilities of the Markovian process.

Cox [1972] introduced the conditional likelihood, later called partial likelihood [Cox, 1975], to estimate the parameters of a semi-parametric PHM, supposing that the base line hazard function in the PHM, $\lambda_0(.)$, is arbitrary and the covariates are time-dependent. It was assumed that the exponential function incorporated the effect of the covariates into the equipment's time-to-failure. For r^{th} failure time $t_{(r)}$, the probability of observing the failure on the equipment that has actually failed given the *risk set* of $R\left(t_{(r)}\right)$ is:

$$\exp\left(\gamma Z_{(r)}\right) \Big/ \sum_{l \in R\left(t_{(r)}\right)} \exp\left(\gamma Z_{(l)}\right) \qquad \textbf{(4-1)}$$

where the risk set $R(t)$ is the set of all equipments that have not yet failed until time t, Z represents the diagnostic covariate of the equipment and γ is the coefficient which represents a weight factor for the covariates. r is the index counter of the sample data of a set of n independently observed histories. Consequently, the log partial likelihood function is:

$$\log\left(L(\gamma)\right) = \sum_{r=1}^{n} Z_{(i)}\gamma - \sum_{r=1}^{n} \log\left[\sum_{l \in R\left(t_{(r)}\right)} \exp\left(\gamma Z_{(l)}\right)\right] \qquad \textbf{(4-2)}$$

Several methods are suggested to estimate the base line hazard function $\lambda_0(t)$. One of them is to assume that $\lambda_0(t)$ is zero except at failure points $t_{(r)}$. The estimator of $\lambda_0(t)$ is thus given by:

$$\hat{\lambda}_0\left(t_{(r)}\right) = \left\{\sum_{l \in R\left(t_{(r)}\right)} \exp\left(\hat{\gamma} Z_{(l)}\right)\right\}^{-1} \qquad \textbf{(4-3)}$$

where $\hat{\gamma}$ is the MLE of γ [Kay, 1984].

If ties exist in the data and the number of ties is small in comparison to the number of available information, then the log partial likelihood is calculated by the following equation:

$$\log\left(L(\gamma)\right)=\sum_{r=1}^{n}S_{(r)}\gamma-\sum_{r=1}^{n}\log\left[\sum_{l\in R\left(t_{(r)}\right)}\exp\left(Z_{(l)}\gamma\right)\right]^{d_r} \tag{4-4}$$

where d_r denotes the number of ties for failure time $t_{(r)}$ and $S_{(r)}$ is the sum of the failed items' covariate at time $t_{(r)}$ [Kalbfleisch and Prentice, 1980]. Also the estimator of $\lambda_0(t)$ is given by:

$$\hat{\lambda}_0\left(t_{(r)}\right)=d_r\bigg/\left\{\sum_{l\in R\left(t_{(r)}\right)}\exp\left(\hat{\gamma}z_l\right)\right\} \tag{4-5}$$

Cox [1972] proposes that the covariates of the PHM can be allowed to be time-dependent, that is to say; their values may vary in equipment's lifetime. In this case the equation $h\left(t,Z(t)\right)=\lambda_0(t)\exp\left(\gamma Z(t)\right)$ indicates the PHM with time-dependent covariates, where γ stands for the covariate's coefficients and $Z(t)$ is the time-dependent covariate at time t.

In a semi-parametric PHM where there is no assumption about the form of $\lambda_0(t)$, γ is estimated by maximizing partial likelihood that does not depend on $\lambda_0(t)$. In a parametric function of a certain form, such as Weibull, the model parameters can be estimated by full likelihood [Lin *et al.*, 2005].

For calculating the full likelihood, the complete covariate realization $\left\{Z_r(t),0\le t<T_r\right\}$, where T_r is the failure or the censoring time of the r-th equipment, should be known. Practically, it is not possible to have the covariate recorded continuously. Instead, it is known in discrete times of observations. An approach to deal with this problem is to

assume that the covariate $Z_r(t)$ is constant between the observations. For the time-dependent PHM, the log partial likelihood function that estimates the parameter γ is given by [Kalbfleisch and Prentice, 1980] as follows:

$$\log\left(L(\gamma)\right)=\sum_{r=1}^{n}\gamma Z_r\left(t_{(r)}\right)-\sum_{r=1}^{n}\log\left[\sum_{l\in R\left(t_{(r)}\right)}\exp\left(\gamma Z_l\left(t_{(r)}\right)\right)\right] \tag{4-6}$$

It is assumed that the hazard at time t depends only on the current covariate vector. The introduced partial likelihood has almost the same form as time-dependent covariates, except that the covariates are now time-dependent.

Banjevic *et al.* [2001] showed that the likelihood of the set of n independently observed histories $\left(T_r,C_r,\left(Z_r(s);s\le T_r\right)\right)$, $r=1,2,\ldots,n$ is:

$$L(\theta)\propto\prod_{r:C_r=1}h\left(T_r,Z_r\left(T_r\right)\right)\prod_{j}S\left(T_j,Z_j\right) \tag{4-7}$$

where T_r is the failure or censoring time of the r-th experiment, C_r is the censoring indication that indicates whether the equipment has failed or has been censored. It takes the following values:

$$C_r=\begin{cases}0 & \text{Censored}\\ 1 & \text{Failed}\end{cases} \tag{4-8}$$

and:

$$S(t;Z)=S\left(t;Z(s),s\le t\right)=\exp\left\{-\int_0^t h\left(\tau,Z(\tau)\right)d\tau\right\} \tag{4-9}$$

j is the risk set at T_i. If the value of Z at the failure or the censoring moment is not known, which might often be the case, the value of the latest covariate is used.

Estimation of the transition matrix of Markov chain can also be obtained by MLE. By considering constant observation times, the estimator of transition probability, $p_{ij}(k) = \Pr(Z_{k+1} = j \mid T > (k+1)\Delta, Z_k = i)$ is given by the following equation:

$$\hat{p}_{ij}(k) = \frac{n_{ij}(k)}{\sum_j n_{ij}(k)} \quad \textbf{(4-10)}$$

where $n_{ij}(k)$ is the amount of one-step transitions from state i to state j at $k-$th observation point, $k = 1, 2, \ldots$ [Basawa and Rao, 1980].

For the imperfect observations, with hidden degradation process, the CBM model consists of two separate stochastic processes: a Hidden Markov Model with finite state space describing the state transition and an observation process. When the observations are not perfect, some researchers used the Expectation Maximization (EM) technique in order to estimate the parameters. Fernandez-Gaucherand [1993] considered a finite state Markov Chain for equipment with partial information. He assumed that a maintenance action resets the state of the equipment to a known value, and consequently, its future evolution becomes independent of the past. He showed that the parameters' estimators converge to their true values.

Lin *et al.* [2003, 2004] considered a CBM problem while the equipment state is partially observable and the failure is obvious. The model's parameters are estimated using a recursive EM algorithm. Adjengue and Yacout [2005] used an EM algorithm for estimating the parameters of CBM with imperfect information.

When the observations are imperfect, the EM method is used to avoid modeling the Probability Density Function (PDF) of the time to failure. In this work, while the observations are imperfect, we have directly modeled the PDF of the observed information (indicator) and used the MLE method to estimate the model's parameters. In what follows, we introduce the proposed model.

4.4 Proposed Model

We consider the PHM proposed by Cox [1972], and we assume that the condition monitoring is indirect i.e. an indirect indicator's value θ, of the underlying degradation state is available at each observation moment. Observations are collected at constant (or near constant) interval Δ. In this study, Z represents the degradation state of the equipment which will be used as the diagnostic covariate in the PHM, and θ is a value from the set of all the possible indicator's values $\Theta = \{1,...,M\}$, where the whole set of the indicator's values is descritized into a finite set of M possible values. The equipment's condition is described as follows:

- The equipment has a finite and known number of degradation states N. $J = \{1,...,N\}$ is the set of all possible degradation states;
- Degradation state transition follows a Markov Chain with unobservable states and is modeled by a Hidden Markov Model (HMM). The transition matrix is $P = \left[p_{ij}\right]$, where p_{ij} is the probability of going from state i to state j, $i, j \in J$ during one observation interval, knowing that the equipment does not failed before the end of the interval;

- The value of the indicator is stochastically related to the equipment's state through the observation probability matrix $Q=\left[q_{j\theta}\right]$, $j\in J$, $\theta\in\Theta$. $q_{j\theta}$ is the probability of getting indicator's value θ, while the equipment is in state j;
- The indicator is collected periodically at fixed intervals Δ;
- Failure is not a degradation state. It is a non-working condition of the equipment that can happen at any time and while the system is in any degradation state, and it is known immediately (obvious failure).

Figure 4-4 depicts the process of degradation and the transition from one degradation state to another, and from each degradation states to the failure. The circles represent the states. State 1 is the best state (new or as new equipment). State N is the worst state, but it is not the failure and the equipment is still working and partially fulfilling its mission. It should be noted that failure can happen at any time and while the equipment is in any degradation state. T is a random variable denoting the failure time and $(1-r_i^k)$ is the probability of going from state i to the failure before the end of the observation interval, while $r_i^k=R\left(k,i,\Delta\right)$, which can be calculated at each observation instant k, is the conditional reliability of the equipment for a period of time Δ, in equipment degradation state is i. For more details about the calculation of $R\left(k,i,\Delta\right)$, please refer to [Ghasemi *et al.,* 2007].

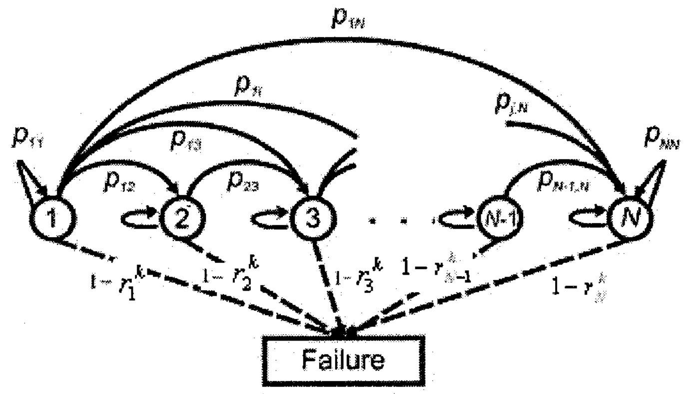

Figure 4-4: The process of degradation and failure

At time $t=0$, the equipment is always in state 1, which indicates that the equipment is in its best state. At fixed interval Δ, i.e. at $t=\Delta, 2\Delta, \ldots$ an indicator of equipment's degradation state is observed. The indicator's value, θ, is observed with a probability of $q_{j\theta}$ when the equipment is in degradation state j, $j \in J, \theta \in \Theta$.

$Z(t) \in \{1,2,\ldots N\}$, the equipment's unobservable degradation state at time t, follows a discrete homogeneous Markov process. We assume that the equipment transition from one degradation state to another takes place just before the observation moment. This assumption requires the observation interval to be short enough to include at most one transition during each interval. Having short enough intervals, also, supports the assumption of having the transition at the end of the interval just before the next observation moment. So $Z(t)$ can be denoted by Z_k where $k\Delta \le t < (k+1)\Delta$, $k=0,1,2,\ldots$. We will use Z_k all through this work.

In this model, the hazard function $h(t, Z_k)$ follows the PHM and is represented by the equation $h(t, Z_k) = h_0(t)\psi(Z_k)$, $k\Delta \le t < (k+1)\Delta, k=0,1,2,\ldots$, where $h_0(.)$ is the base hazard function of a Weibull distribution and represents the aging process, and $\psi(.)$ is a function of the equipment's degradation state. The most used function is exponential in the form $\psi(Z_k)=\exp(\gamma Z_k)$, where γ is the indicator's coefficient.

Ghasemi *et al.* [2007] introduced a new state space and transition rule for this problem. The transition rule includes all the observations from the last renewal point, and provide a methodology to deal with unobservable degradation states by calculating the conditional probability of being in degradation state i at the k^{th} observation moment, π_i^k. π^k is the conditional probability distribution of the equipment's degradation state at period k, and is defined as follows:

$$\pi^k=\left\{\pi_i^k;\ \ 0\le\pi_i^k\le 1 \text{ for } i=1,\dots,N, \sum_{i=1}^{N}\pi_i^k=1\right\},\ k=0,1,2,\dots \tag{4-11}$$

We have also assumed that $\pi_i^0=\begin{cases}1 & i=1\\ 0 & \text{o.w.}\end{cases}$, meaning that the equipment is in its best possible state at period zero. After obtaining an indicator value θ via an inspection at an observation moment, the prior conditional probability π^k, is updated to π^{k+1}. By using Bayes' formula, and knowing that the indicator θ has occurred at the $k+1$-st observation moment, $\pi_j^{k+1}(\theta)$ is determined as follows: [Ghasemi *et al.,* 2007]

$$\pi_j^{k+1}(\theta)=\frac{\sum_{i=1}^{N}\pi_i^k p_{ij}q_{j\theta}}{\sum_{i=1}^{N}\sum_{l=1}^{N}\pi_i^k p_{il}q_{l\theta}},\ j=1,\dots,N \tag{4-12}$$

In the next section, we develop the PDF of the time to failure, and introduce the MLE of the model's parameters, with uncensored and censored data.

4.5 Parameters' Estimation

Let T be the lifetime of the equipment, which is an i.i.d., non-negative continuous random variable. $\theta(s)=\{\theta^1,\theta^2,...,\theta^k\};s\le T;k=1,2,...;k\Delta\le s$ is the history of the indicator values up to time s, where Δ is the observation interval. At any observation moment, the observed indicator value θ, is stochastically related to the underlying degradation state of equipment j, through the probability matrix $Q=\left[q_{j\theta}\right]$. The indicator's values history up to time s, can be mapped into the state conditional probability distribution up to time s as $\pi(s)=\{\pi^1,\pi^2,...,\pi^k\}$; $s\le T$; $k=1,2,...;k\Delta\le s$, where the elements of $\pi(s)$ are calculated from equations (4-11) and (4-12) at corresponding observation moment when a new indicator value is available. This history carries the conditional probability distributions from time zero up to time s, at all observation epochs.

We have assumed that the equipment's unobservable degradation states transition follows a time homogeneous Markov Process. Then, the transition probability at time $t=k\Delta$, from state i to state j, knowing that the equipment has survived at least until next observation moment, can be expressed as:

$$p_{ij}(k)=p_{ij}=\Pr\left(Z_{k+1}=j\mid Z_k=i,T>(k+1)\Delta\right),\ k=1,2,3,....$$

The survival function of the assumed model is:

$$\begin{aligned}R(t,\theta(t))&=\Pr\left(T>t,\theta^1,\theta^2,...,\theta^k\right);\ k\Delta\le t<(k+1)\Delta\\&=\Pr\left(T>t,\pi^0,\pi^1,...,\pi^k\right);\ k\Delta\le t<(k+1)\Delta\end{aligned}$$

$$R(t,\theta(t)) = \Pr(T>0,\pi^0)\Pr(\theta^1,T>\Delta \mid T>0,\pi^0)\Pr(\theta^2,T>2\Delta \mid T>\Delta,\pi^1)$$
$$\ldots\Pr(\theta^k,T>k\Delta \mid T>(k-1)\Delta,\pi^{k-1})\Pr(T>t \mid T>k\Delta,\pi^k)$$
$$\times\Pr(T>t \mid T>k\Delta,\pi^k)$$
$$R(t,\theta(t)) = \Pr(T>0,\pi^0)\left[\prod_{l=0}^{k-1}\Pr(\theta^{l+1},T>(l+1)\Delta \mid T>l\Delta,\pi^l)\right]$$

We have also assumed that $\pi_i^0 = \begin{cases} 1 & i=1 \\ 0 & \text{o.w.} \end{cases}$, meaning that the equipment is in its best possible state at period zero. For the sake of calculation and without losing generality, we also assume that $\Pr(T>0,\pi^0)=1$. This means that no failure happens at time zero when new or as new equipment is put to performance, which is an acceptable assumption.

$$R(t,\theta(t))$$
$$=\left[\prod_{l=0}^{k-1}\Pr(\theta^{l+1} \mid T>(l+1)\Delta,T>l\Delta,\pi^l)\Pr(T>(l+1)\Delta \mid T>l\Delta,\pi^l)\right]\Pr(T>t \mid T>k\Delta,\pi^k)$$
$$=\left[\prod_{l=0}^{k-1}\Pr(\theta^{l+1} \mid T>(l+1)\Delta,\pi^l)\Pr(T>(l+1)\Delta \mid T>l\Delta,\pi^l)\right]\Pr(T>t \mid T>k\Delta,\pi^k)$$
$$=\prod_{l=0}^{k-1}\Pr(\theta^{l+1} \mid T>(l+1)\Delta,\pi^l)\prod_{l=0}^{k-1}\Pr(T>(l+1)\Delta \mid T>l\Delta,\pi^l)\Pr(T>t \mid T>k\Delta,\pi^k)$$

The probability of observing an observation $\theta^{l+1}=\theta$, at the $l+1^{\text{st}}$ observation moment, knowing that the conditional probability distribution of the equipment's degradation state at the l^{th} observation moment was π^l and it will survive until the $l+1^{\text{st}}$ observation moment, is [Ghasemi *et al.*,2008]:

$$\Pr(\theta \mid T>(l+1)\Delta,\pi^l) = \sum_{i=1}^{N}\sum_{j=1}^{N}\pi_i^l p_{ij} q_{j\theta} \qquad \textbf{(4-13)}$$

$\Pr\left(T>(l+1)\Delta \mid T>l\Delta,\pi^{l}\right)$, the conditional probability of survival from l^{th} observation moment until next observation moment, while the conditional probability distribution of the equipment's degradation state at the l^{th} observation moment is π^{l}, can be calculated by [Ghasemi *et al.* 2005]:

$$\begin{aligned}\Pr\left(T>(l+1)\Delta \mid T>l\Delta,\pi^{l}\right) &= \sum_{i=1}^{N}\pi_i^{l}\Pr\left(T>(l+1)\Delta \mid T>l\Delta, Z_l=i\right)\\ &= \sum_{i=1}^{N}\pi_i^{l}\exp\left(-\int_{l\Delta}^{(l+1)\Delta} h(\tau,i)\,d\tau\right)\end{aligned}$$

where $h(\tau,i)$ is the PHM's hazard function at time τ while the equipment's degradation state is i. Also, we have:

$$\begin{aligned}\Pr\left(T>t \mid T>k\Delta,\pi^{k}\right) &= \sum_{i=1}^{N}\pi_i^{l}\Pr\left(T>t \mid T>k\Delta, Z_k=i\right)\\ &= \sum_{i=1}^{N}\pi_i^{k}\exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right)\end{aligned}$$

then:

$$R\left(t,\theta(t)\right)=$$

$$\underbrace{\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1} \mid T>(l+1)\Delta,\pi^{l}\right)\right]}_{A}\underbrace{\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^{l}\exp\left(-\int_{l\Delta}^{(l+1)\Delta} h(\tau,i)\,d\tau\right)\right]\left[\sum_{i=1}^{N}\pi_i^{k}\exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right)\right]}_{B} \tag{4-14}$$

Based on equation (5-13), term A is function of the probabilities of Markov transition p_{ij}'s, and of the probabilities of stochastic relation between the equipment's degradation states and the indicator's observations, $q_{j\theta}$'s. Term B is a function of p_{ij}'s, $q_{j\theta}$'s and the set of all parameters of the hazard function $h(\tau,i)$, as well. From now on, we will optionally point to the set of all p_{ij}'s, $q_{j\theta}$'s and the set of all parameters of the hazard function as Ω.

According to failure analysis methods [Kalbfleisch and Prentice, 1980]:

$$f(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t \le T < t + \Delta t \mid T > t)}{\Delta t} = -\frac{dR(t)}{dt}$$

where $R(t)$ is the survival function of the equipment. In the case of existence of condition monitoring data, $f(t,\theta(t))$, the PDF of T, the time to failure, is calculated by:

$$f(t,\theta(t)) = -\frac{dR(t,\theta(t))}{dt}$$

In equation (5-14), since two first terms of $R(t,\theta(t))$ do not depend on t, we can write:

$$\frac{d}{dt}R(t,\theta(t)) =$$

$$\left[\prod_{l=0}^{k-1} \Pr\left(\theta^{l+1} \mid T > (l+1)\Delta, \pi^l\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N} \pi_i^l \exp\left(-\int_{l\Delta}^{(l+1)\Delta} h(\tau,i)\,d\tau\right)\right]\frac{d}{dt}\sum_{i=1}^{N}\pi_i^k \exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right)$$

and since $\frac{d}{dt}\sum_{i=1}^{N}\pi_i^k \exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right) = \sum_{i=1}^{N}\pi_i^k \exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right)h(t,i)$ then:

$$f(t,\theta(t)) =$$

$$\left[\prod_{l=0}^{k-1} \Pr\left(\theta^{l+1} \mid T > (l+1)\Delta, \pi^l\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N} \pi_i^l \exp\left(-\int_{l\Delta}^{(l+1)\Delta} h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^k \exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right)h(t,i)$$

In the next section, we introduce the Maximum Likelihood Estimator of Ω, the set of parameters of interest in the model.

4.5.1 Maximum Likelihood Estimation

We have considered the problem of parameter estimation in two cases of with and without censoring data. First, we consider the cases of data without censoring in what follows.

a) With uncensored data

For a set of n independent experiments, we assume that T_r is the time to failure of the r^{th} experiment. Also we assume that $\theta(T_r)=\{\theta_r^1,\theta_r^2,...,\theta_r^k\}, k\Delta \le T_r < (k+1)\Delta$ is the history of observations of the r^{th} experiment up to T_r. The likelihood of the set of parameters Ω, based on the available data can be calculated by: [Kalbfleisch and Prentice 1980]

$$L(\Omega)=\prod_{r=1}^{n} f\left(T_r,\theta(T_r);\Omega\right)$$

We assume that $\left(T_r,\theta(T_r)\right)$, $r=1,...,n$ are independent. In this paper, we consider a parametric PHM with a baseline Weibull hazard function as the hazard function of the equipment, which is known as Weibull parametric regression model [Banjevic *et al.*, 2001].

$$h\left(t,Z_k;\beta,\eta,\gamma\right)=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\exp\left(\gamma Z_k\right);k\Delta \le t<(k+1)\Delta, k=0,1,... \tag{4-15}$$

Since we have assumed that the equipment degradation state does not change during an observation period, and any change takes place only at the end of the observation period, just before the next observation, then:

$$\int_{k\Delta}^{t} h(\tau,i)\,d\tau = \int_{k\Delta}^{t} \frac{\beta}{\eta}\left(\frac{\tau}{\eta}\right)^{\beta-1} \exp(\gamma i)\,d\tau; t \le (k+1)\Delta$$

$$= \exp(\gamma i) \int_{k\Delta}^{t} \frac{\beta}{\eta}\left(\frac{\tau}{\eta}\right)^{\beta-1} d\tau$$

$$= \exp(\gamma i) \left[\left(\frac{\tau}{\eta}\right)^{\beta}\right]_{k\Delta}^{t}$$

$$\int_{k\Delta}^{t} h(\tau,i)\,d\tau = \exp(\gamma i)\left[\left(\frac{t}{\eta}\right)^{\beta} - \left(\frac{k\Delta}{\eta}\right)^{\beta}\right]$$

By applying a maximization technique such as the line search method [Box *et al.,* 1969, Press *et al.,* 2007], the maximum of the log likelihood function can be calculated. This will result in finding the model's set of parameters Ω, which contains β, η and γ as well as the p_{ij}'s and $q_{j\theta}$'s for all possible values of i, j.

Simulation

The accuracy of the provided model is evaluated based on a simulation study. In what follows, we simulate the equipment's behaviors based on the assumptions of the model and some pre-set parameters. Then by applying the estimation method provided earlier, the parameters are estimated and compared to the original pre-set parameters. The simulation procedure follows the following steps:

Overall Initialization:

We assume a piece of equipment with two possible degradation states $\{1,2\}$, and the related Markov transition matrix as $P = \begin{bmatrix} p & 1-p \\ 0 & 1 \end{bmatrix}$. There are three possible values for the observations indicator, for example $\{1,2,3\}$ which can be interpreted as level 1, 2 or 3. These are not the underlying degradation state, but the indicator's values that are stochastically related to the equipment's underlying degradation state. The matrix Q represents the relation between the equipment's degradation state and the observation as

$$Q=\begin{bmatrix}1-q_1 & q_1 & 0\\ 0 & q_2 & 1-q_2\end{bmatrix}.$$

According to the equation (5-15), there are also three parameters for the PHM, β,η and γ. The observation interval is Δ, which is usually set based on the system's constraints or expert's opinion and is not a decision variable or a model's parameter. Ghasemi *et al.* [2008] have addressed the problem of finding the optimum observation interval. We set $\Delta=1$, and we estimate six parameters: p,q_1,q_2,β,η and γ. In order to generate n simulated experiments, we introduce $r=1,...,n$ to represent the experiment's index.

General Initialization:

Let k represent the observation period counter and let z_r^k represent the r-th experiment's real degradation state at the k-th observation moment. It is important to notice that the real degradation state is not needed for the parameter estimation method, we just keep record of this information for the simulation purpose, and it will not be recorded for the experiments at the end of the simulation. This will comply with the indirect observation assumption of the model.

Set $r=1$ to generate the first experiment.

Experiment Initialization:

New or as-new equipment is assumed to be in its best degradation state. This means that for new or as-new equipment, i.e. when $k=0$,we set $z_r^k=1$, for any $r=1,\ldots,n$.

Survival Simulation:

The conditional probability of survival until the next observation moment for r-th experiment at k-th observation moment, while its underlying degradation state is z_r^k, will be calculated by [Ghasemi *et al.,* 2007]:

$$\begin{aligned}R(k,z_r^k,\Delta) &= P\left(T_r > k\Delta+\Delta \mid T_r > k\Delta, z_r^0, z_r^1,\ldots,z_r^k\right)\\ &= 1-P\left(T_r \le k\Delta+\Delta \mid T_r > k\Delta, z_r^0, z_r^1,\ldots,z_r^k\right)\\ &= \exp\left(-\psi\left(z_r^k\right)\int_{k\Delta}^{(k+1)\Delta} h_0(\tau)d\tau)\right)\end{aligned}$$

The probability of survival until the next observation moment can be simulated by generating a random number u from a uniform distribution $U(0,1)$, and comparing it with the probability of survival. If $u \le R(k,z_i^k,\Delta)$ the simulation goes to the *Transition and Observation Simulation,* otherwise it goes to *Failure Simulation.* For example $R(k,z_i^k,\Delta)=0.25$ means that the probability of survival until $k\Delta+\Delta$ is 25% . If a uniform random number u between zero and one is generated, it will have 25% chance of being less than 0.25 . So any $u \le 0.25$ can be interpreted as the equipment will survive until $k\Delta+\Delta$.

Transition and Observation Simulation:

According to the Markov transition matrix P, which is a stochastic matrix, degradation state of a piece of equipment that has not failed during k-th period, will change from current degradation state, i, to a future degradation state j, with a probability p_{ij}. We determine the experiment's degradation state during $k+1$ period, z, by generating a uniform random number between zero and one, and comparing it with corresponding *cumulative probabilities* of Markov transition matrix to the current degradation state. For example assume that the corresponding line of matrix P to state i is $[0.4 \quad 0.6]$. It means that the equipment will stay at degradation state 1 by 40% chance or move to degradation state 2 by 60% chance. By comparing the generated random variable $u \sim U(0,1)$, with cumulative probabilities i.e. 0.4 and 1.0 we are able to simulate the Markov transition behavior. If $u \leq 0.4$ then $z=1$, otherwise $z=2$.

By considering the experiment's real degradation state during $k+1$ period, z, we simulate the indicator's observation at $k+1$ observation moment ,θ. We generate a random number from a uniform distribution, $U(0,1)$ and compare it with the *cumulative probabilities* of matrix Q, corresponding to z, the experiment's degradation state during $k+1$ period. For example assume that $Q=\begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.4 & 0.6 \end{bmatrix}$ and the real degradation state is $z=1$, so we have to consider the first row of matrix Q. Also assume that, the random number generator has generated 0.65, which is greater than $q_{11}=0.5$. Then we compare the random number with $q_{11}+q_{12}$; $0.65 < q_{11}+q_{12}=1.00$ which means that the simulated observation $\theta=2$. q_{ij} in this example refers to the element in i^{th} line and j^{th} column of the matrix Q.

We set $k = k+1$, $z_r^k = z$ and $\theta_r^k = \theta$, and continue from *Survival Simulation* step.

Failure Simulation:

This procedure simulates the failure time of an experiment, during $k+1$-st period. Time of failure of an experiment, knowing that it has failed before $k+1$ observation moment, can be simulated using conditional survival function [Ghasemi *et al.*, 2008].

$$F(k, z_r^k, t) = \Pr\left(T_r \le k\Delta + t \mid T_r > k\Delta, z_r^0, z_r^1, \ldots, z_r^k\right), t \le \Delta$$

then:

$$\begin{aligned}
&\Pr\left(T_r \le k\Delta + t \mid k\Delta \le T_r < k\Delta + \Delta, z_r^0, z_r^1, \ldots, z_r^k\right) \\
&\qquad = \frac{\Pr\left(T_r \le k\Delta + t, T_r < k\Delta + \Delta \mid T_r > k\Delta, z_r^0, z_r^1, \ldots, z_r^k\right)}{\Pr\left(k\Delta \le T_r < k\Delta + \Delta, z_r^0, z_r^1, \ldots, z_r^k\right)} \\
&\qquad = \frac{\Pr\left(T_r \le k\Delta + t \mid T_r > k\Delta, z_r^0, z_r^1, \ldots, z_r^k\right)}{\int_{k\Delta}^{k\Delta+\Delta} f\left(k, z_r^k, \tau\right) d\tau} \\
&\qquad = \frac{F\left(k, z_r^k, t\right)}{F\left(k, z_r^k, \Delta\right) - F\left(k, z_r^k, 0\right)}, t \le \Delta
\end{aligned}$$

$$\Pr\left(T_r \le k\Delta + t \mid k\Delta \le T_r < k\Delta + \Delta, z_r^0, z_r^1, \ldots, z_r^k\right) = \frac{1 - R\left(k, z_r^k, t\right)}{R\left(k, z_r^k, 0\right) - R\left(k, z_r^k, \Delta\right)}, t \le \Delta \qquad \textbf{(4-16)}$$

where $R(k, z_r^k, t) = \exp\left(-\psi\left(z_r^k\right) \int_{k\Delta}^{t} h_0(\tau) d\tau)\right); t \le \Delta$. According to its CDF, the random variable $U =$ $\Pr\left(T_r \le k\Delta + t \mid k\Delta \le T_r < k\Delta + \Delta, z_r^0, z_r^1, \ldots, z_r^k\right)$, a transformation of continuous random variable T_r, is uniformly distributed random variable between zero and one [Mirham, 1972]. The inverse of function U is not easy to determine, so the reverse CDF method is used here. To simulate the failure time of an experiment T_r, given that it has failed between $k\Delta$ and $k\Delta+\Delta$, we discretize the observation interval Δ, in to m (an optional number chosen based on the systems characteristics and the desired precision) equal parts and calculate $\Pr(l) = \Pr\left(T_r \le k\Delta + \frac{l\Delta}{m} \mid k\Delta \le T_r < k\Delta+\Delta, z_r^0, z_r^1, \ldots, z_r^k\right)$, for all $l = 0,\ldots,m-1$. This approach makes it possible to use the discrete random variable generation method, similar to what we used in previous parts, to simulate the time to failure of the experiment. A uniform random $u \sim U(0,1)$, is generated and compared with *cumulative probabilities* $\sum_0^l \Pr(l)$, $l = 0,\ldots,m-1$. The simulated failure time of the experiment, then will be:

$$T_r = k\Delta + \inf\left\{\frac{l\Delta}{m} : \sum_0^l \Pr(l) > u, l = 0,\ldots,m-1\right\}.$$

Record the experiment:

The parameter estimation method is based on the experiments' failure time T_r and the experiments indicator values set $\{\theta_r^1,\ldots,\theta_r^k\}; k\Delta \le T_r < k\Delta+\Delta$. We save these two piece of information for the current experiment and continue to *Experiment Initialization,* by $r = r+1$ while $r \le n$, where n is the desired number of experiments to generate.

Numerical Example

We have run the simulation by considering $\Delta = 1$, for $n = 10, 30, 100, 300, 500, 1000$ and 5000. Each simulation has been run for 100 samples. For instance, for $n = 30$, we have simulated 30 experiments and estimated the parameters using the introduced method.

Each simulation has been repeated for 100 times. The starting parameters of the experiments are $\beta = 1.50$, $\eta = 2.5$, $\gamma = 1$, $p = 0.95$, $q_1 = 0.5$ and $q_2 = 0.6$. The results of the simulations are shown in the following figures and tables.

Figure 4-5, shows that all the estimated parameters converge to the real value of the parameter while the number of the experiment increases.

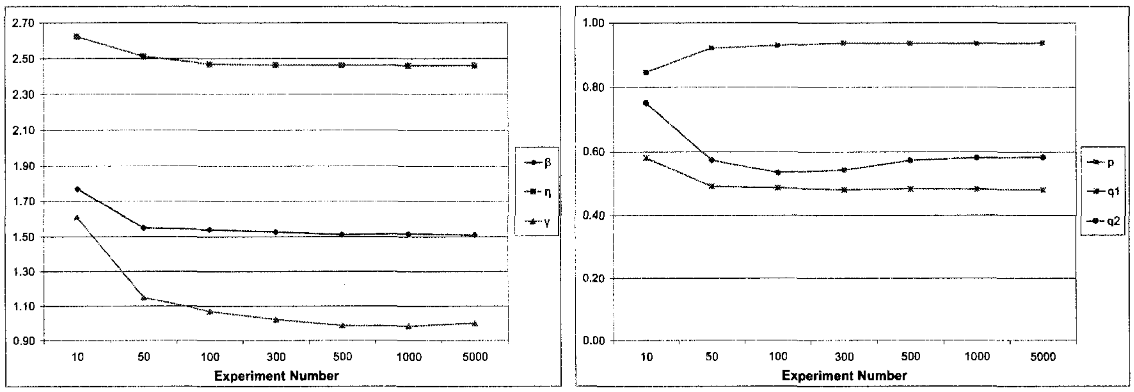

Figure 4-5: Mean of the estimated parameters with different experiment number

Table 4-1 demonstrates the variance and the standard error of the estimated parameters. Again, all the cases demonstrate decrease in the variance and standard error by increasing the number of the experiments. Generally, it can be concluded from Figure 4-5 and Table 4-1 that by 50 instances of experiments and higher, the methods gives very reasonable results with acceptable variance and standard error.

Table 4-1: Variance and standard error of the estimated parameters by different experiment size

	Variance						Standard Error					
n	β	η	γ	p	q1	q2	β	η	γ	p	q1	q2
10	0.554	0.751	1.978	0.026	0.039	0.131	0.074	0.087	0.141	0.016	0.020	0.036
50	0.033	0.086	0.614	0.003	0.006	0.117	0.018	0.029	0.078	0.005	0.008	0.034
100	0.016	0.046	0.309	0.001	0.003	0.064	0.013	0.022	0.056	0.004	0.006	0.025
300	0.005	0.012	0.060	0.000	0.001	0.025	0.007	0.011	0.024	0.002	0.003	0.016
500	0.003	0.008	0.036	0.000	0.001	0.014	0.005	0.009	0.019	0.001	0.002	0.012
1000	0.002	0.004	0.016	0.000	0.000	0.006	0.004	0.006	0.012	0.001	0.002	0.008
5000	0.000	0.001	0.004	0.000	0.000	0.002	0.002	0.003	0.007	0.001	0.001	0.004

Table 4-2 includes the Mean Squared Error (MSE) of the estimated parameters with different experiment sizes. Decrease in MSE is noticeable while the number of experiments increases.

Table 4-2: Mean Squared Error of the estimated parameters by different experiment numbers

MSE						
n	β	η	γ	p	q1	q2
10	0.072	0.016	0.374	0.011	0.007	0.023
50	0.003	0.000	0.022	0.001	0.000	0.001
100	0.002	0.001	0.005	0.000	0.000	0.004
300	0.001	0.001	0.000	0.000	0.000	0.003
500	0.000	0.001	0.000	0.000	0.000	0.001
1000	0.000	0.002	0.000	0.000	0.000	0.000
5000	0.000	0.001	0.000	0.000	0.000	0.000

Figure 4-6 shows the confidence interval for $1-\alpha=95\%$, for all the estimated parameters. It can be seen that all the intervals are narrowing by higher number of experiments, and as discussed before, at the same time, the mean value of the estimated value converges to the real value of the parameter.

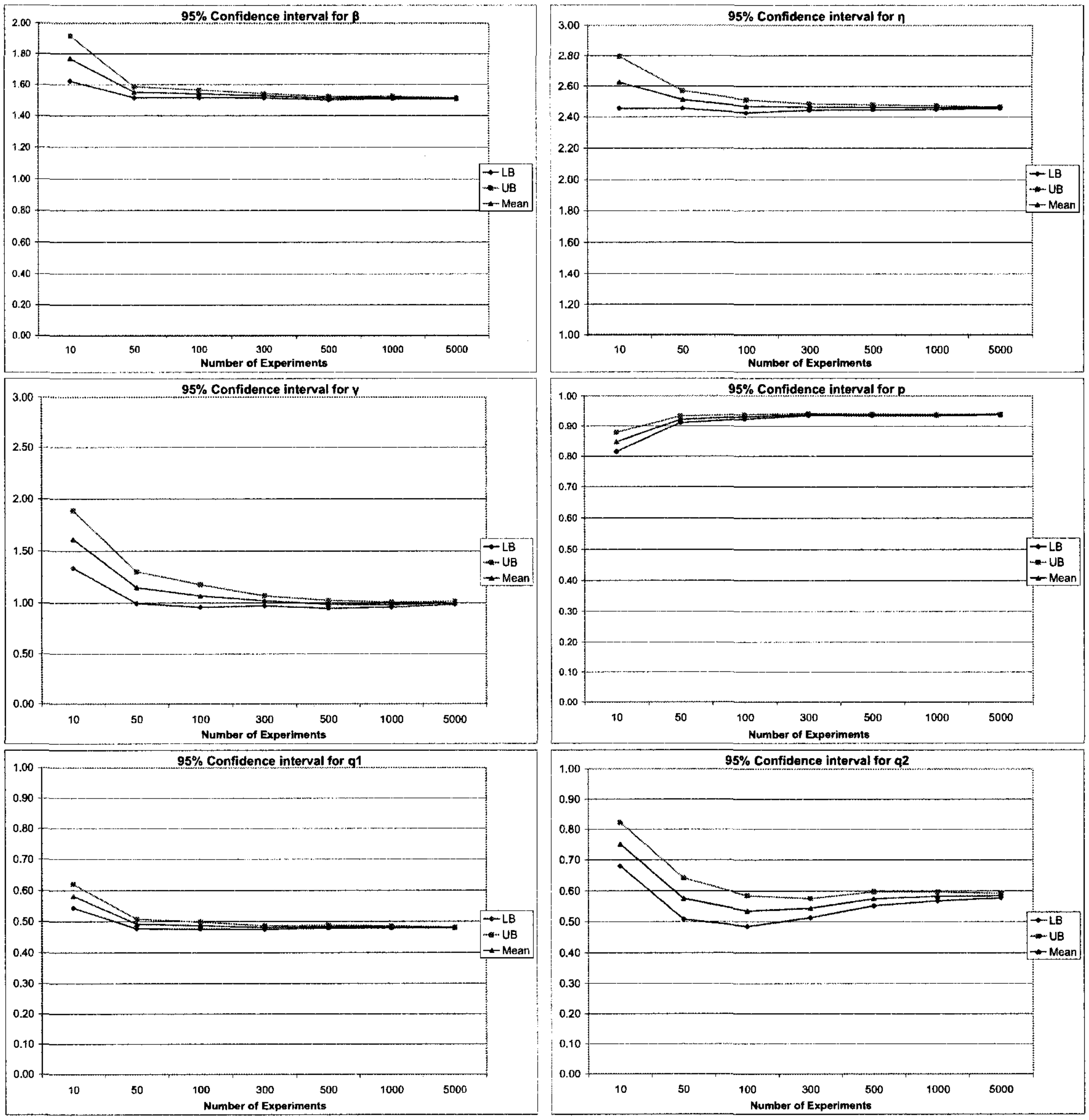

Figure 4-6: Confidence interval of 95% for estimated parameters with different sample sizes

We conclude that the introduced parameter estimation method is very effective and consistent. In what follows, we will introduce a method for estimating the parameters of the same problem when random censoring exists in the data. Similarly, we will study the consistency and accuracy of the method.

b) With censored data

For a set of n independent experiments, assume that T_r is the stopping time of the r^{th} item. Also assume that $\theta(T_r)$ is the history of the indicator's observations up to T_r. We also assume an censoring indicator $\delta_r = \begin{cases} 1 & \text{if the item has failed} \\ 0 & \text{if the item is censored} \end{cases}$, which indicates whether the value of T_r is an actual failure time or it is the moment at which we have randomly stopped the experiment i.e. T_r is the censoring time. We are assuming random censorship, and that the censoring times are independent of each other and of the failure times, as well. We assume also that the censoring is non-informative [Kalbfleisch and Prentice, 1980]. In short, non-informative censoring means that the censoring distribution does not depend on the unknown parameters in the model. Since we condition on observed values of censorings in Cox's regression model, the censoring distributions do not enter the partial likelihood. Then the likelihood of the set of parameters Ω based on the available data can be given by: [Kalbfleisch and Prentice, 1980]

$$L(\Omega) = \prod_{r=1}^{n} f(T_r, \theta; \Omega)^{\delta_r} R(T_r, \theta; \Omega)^{1-\delta_r}$$

Earlier we have shown that:

$$f(t,\theta) =$$
$$\left[\prod_{l=0}^{k-1} \Pr\left(\theta^{l+1} \mid T > (l+1)\Delta, \pi^l\right)\right]\left[\prod_{l=0}^{k-1} \sum_{i=1}^{N} \pi_i^l \exp\left(-\int_{l\Delta}^{(l+1)\Delta} h(\tau,i)\,d\tau\right)\right]$$
$$\times \sum_{i=1}^{N} \pi_i^k \exp\left(-\int_{k\Delta}^{t} h(\tau,i)\,d\tau\right) h(t,i),$$

and

$$R\left(t,\theta(t)\right)=$$

$$\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1}\mid T>(l+1)\Delta,\pi^{l}\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^l\exp\left(-\int_{l\Delta}^{(l+1)\Delta}h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^k\exp\left(-\int_{k\Delta}^{t}h(\tau,i)\,d\tau\right).$$

By applying the same maximization technique that was used for the likelihood of uncensored data, the maximum likelihood function of $L(\Omega)$ can be calculated. In what follows, we have analyzed the effectiveness of the estimation method by a simulation study.

Simulation

The simulation procedure with censored data is similar to what was presented in the previous part. One difference is at the step of *Overall Initialization*, where we define I, as the censoring percentage. I is the ratio of the experiments to be censored randomly by the simulation. Also we introduce $\delta_r, r=1,\ldots,n$, as the censoring indicator which will be recorded at *Record the experiment* step , along with other information. The step *Censoring Simulation* is an extra step in the simulation just before *Record the experiment*. The simulation is censored as follows:

Censoring Simulation:

The value of I, the percent of censored experiments in each simulation study, is chosen at the beginning of each simulation. For each experiment a random number $U(0,1)$ is generated and compared to the value of I. If the value of the random number is smaller than I, this means that the current experiment should be censored. In this case, we update δ_r to 1.

Next step for a censored experiment is to determine at what point of time it has been censored. We assume that any experiment is censored at a uniformly distributed time between zero and the experiment's failure time. Obviously, the censoring time must happen before the equipment failure time; otherwise the equipment would have failed before being censored. To apply this step, we generate a random number $u \sim U(0,1)$ and replace T_r with uT_r for the experiments with $\delta_r = 1$, $r = 1,\ldots,n$.

Numerical Example

To run the simulation with censored data a sample size of 300 was considered. . Figure 4-7 demonstrates the mean of estimated parameters with different censoring percentage, beginning with 0%, i.e. no censoring data. It can be seen that, parameters β,η and γ consistently increase by increasing the censoring percentage. Nevertheless, the estimations are very close to the case without censoring for β and γ even for 80% of censoring data. The average estimated value for η also shows little increase up to 30% of censoring. The increase in average estimated η increases more significantly after 50% of censoring data.

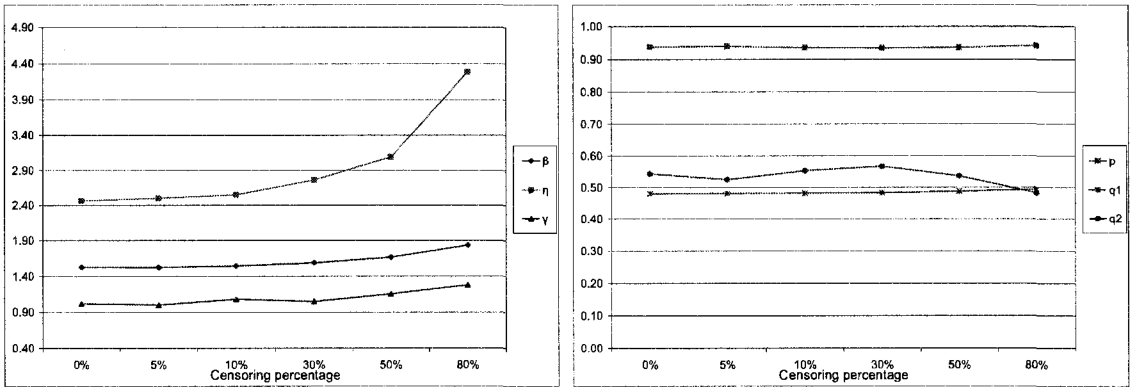

Figure 4-7: Mean of the estimated parameters with different censoring percentage and sample size 300

The average estimated value of the parameters p and q_1 are almost the same values with and without censored data. Average estimated value of q_2 fluctuates more significantly around the original value estimated without censoring. The MSE of average estimated parameter q_2 is less than 0.004 for all the cases (see Table 4-4). Insensibility of these three parameters to the percentage of the censored data can be justified by the fact that; even when the equipment is stopped before its failure, the relationship between the equipment's real state and the observation indicator still exists. Their stochastic relationship is not affected by the fact that there has been a failure or censoring. Based on the simulation results, we have also calculated the variance and the standard error of the estimated parameters shown in Table 4-3. The value of the variances and the standard errors show small increase with the increase in the censoring percentage. Yet, they are almost constant for parameters p and q_1, and slightly increasing for q_2.

Table 4-3: Variance and standard error of the estimated parameters with different censoring percentage

	Variance						Standard Error					
I	β	η	γ	p	q1	q2	β	η	γ	p	q1	q2
0%	0.005	0.012	0.060	0.000	0.001	0.025	0.007	0.011	0.024	0.002	0.003	0.016
5%	0.005	0.012	0.074	0.001	0.001	0.036	0.007	0.011	0.027	0.002	0.004	0.019
10%	0.005	0.013	0.083	0.001	0.001	0.028	0.007	0.012	0.029	0.002	0.003	0.017
30%	0.010	0.022	0.111	0.001	0.001	0.033	0.010	0.015	0.033	0.003	0.003	0.018
50%	0.010	0.029	0.148	0.001	0.001	0.048	0.010	0.017	0.038	0.003	0.004	0.022
80%	0.035	0.241	0.418	0.001	0.001	0.081	0.019	0.049	0.065	0.003	0.004	0.028

Table 4-4: Mean Squared Error of the estimated parameters with different censoring percentage

MSE						
I	β	η	γ	p	q1	q2
5%	0.000	0.001	0.000	0.000	0.000	0.000
10%	0.000	0.008	0.004	0.000	0.000	0.000
30%	0.004	0.089	0.001	0.000	0.000	0.001
50%	0.019	0.392	0.018	0.000	0.000	0.000
80%	0.094	3.309	0.067	0.000	0.000	0.004

The MSE of the estimated parameters are shown in Table 4-4. As explained before, the estimation error for PHM parameters, i.e. β, η and γ increases by increasing the censoring percentage and it is almost constant for rest of the parameters. The MSE values in Table 4-4 are calculated by considering the average estimated values for the parameters without censoring with 300 replications, as the reference point. Figure 4-8 demonstrates the confidence interval of 95% for estimated parameters while the censoring percentage increases. As expected, the confidence interval gets broader by increasing the censoring percentage for PHM parameters and is almost constant for the rest of the parameters.

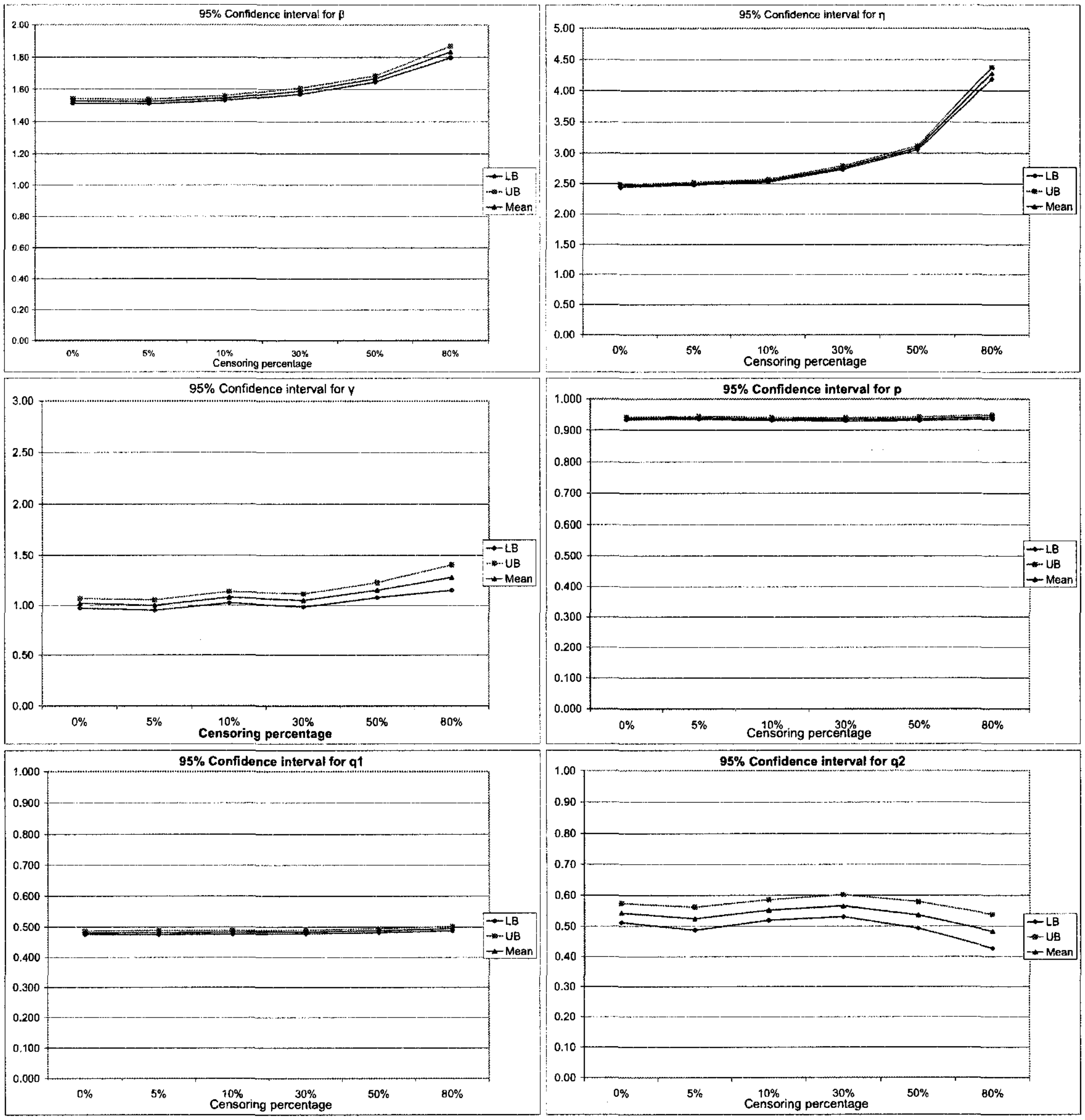

Figure 4-8: Confidence interval of 95% for estimated parameters with different censoring percentage

4.6 Conclusion

In this research, we have addressed the parameter estimation problem for a condition monitoring system, where the degradation state of the equipment is not directly observable and is modeled by a Hidden Markov Model. The observed indicator on equipment's degradation state has a stochastic relation with the degradation state of the equipment via a stochastic matrix and does not reveal the real degradation state of the equipment. The failure rate of the equipment is assumed to follow the Cox's PHM. We have introduced an approach to estimate the parameters of the PHM, Markov process transition matrix and the stochastic matrix of observations/state using Maximum Likelihood Estimation method. By a simulation approach, we have shown that the method converges to the real value of the parameters for bigger sample size. In addition, the behavior of the method has been examined when there exist censoring in data. The parameters of PHM show higher level of sensitivity to censoring data. The higher the percentage of censoring data, farther the amount of the estimated parameter to the real value. The existence of censoring data results in higher value of the PHM parameters. Based on the same study, the parameters of the Markov process and the stochastic matrix of observation/state are not very sensitive to the percentage of the censoring data.

4.7 References

1. Adjengue L. and Yacout S. Parameters estimation for condition based maintenance with Uncorrelated and Corelated Observations. Quality Engineering. 2005; 19:197-206.

2. Ansell, J. I. and Phillips, M. J. Practical aspects of modeling of repairable systems data using proportional hazard models. Reliability Engineering and System Safety. 1997; 58:165-71.

3. Banjevic, D.; Jardine, A. K. S.; Makis, V., and Ennis, M. A control-limit policy and software for condition-based maintenance optimization. INFOR Journal. 2001; 39(1):32-50.

4. Basawa, I. V. and Rao, P. Statistical Inference for Stocastic Processes. London: Academic Press; 1980.

5. Box M.J.; Davies D., and Swann W.H. Non-Linear Optimisation Techniques. London: Oliver & Boyd; 1969.

6. Cox, David Roxbee. Regression models and life tables (with discussion). Journal of the Royal Statistical Society - Series B. 1972; 26(187-220).

7. Crowley, J and Hu, M. Covariance analysis of heart transplant survival data. Journal of American Statistical Association; 72(357):27-36.

8. Elsayed, E. A. and Chan, C. K. Estimation of Thin-Oxid Reliability Using Proportional Hazards Models. IEEE Transactions on Reliability . 1990; 39(3):329-335.

9. Fernandez-Gaucherand, Emmanuel. Analysis of an Adaptive Control Scheme for a Partially Observed Controlled Markov Chain. IEEE Transactions on Automatic Control. 1993; 38(6):987-993.

10. Ghasemi, A.; Yacout, S., and Ouali, M. S. Evaluating the remaining life for equipment with unobservable states. IEEE Transactions on Reliability. 2008; (submitted).

11. Optimal condition based maintenance with imperfect information and the proportional hazards model. International Journal of Production Research. 2007; 45(4):989-1112.

12. Ghasemi, Alireza. Condition Based Maintenance Using the Proportional Hazards Model with Imperfect Information: Ecole Polytechnique de Montreal; 2005; c2005.

13. Jardine A. K. S. and Anderson M. Use of concomitant variables for reliability estimation. Maintenance Management International. 1985; 5(2):135-140.

14. Jardine, A. K. S.; Anderson, P. M., and Mann D. S. Application of the Weibull Proportional Hazards Model to Aircraft and Marine Engine Failure Data. Quality and Reliability Engineering International. 1987; 3:77-82.

15. Jardine, A. K. S.; Banjevic, D.; Wiseman, M.; Buck, S., and Joseph, T. Optimizing a mine haul truck wheel motors' condition monitoring program. Journal of Quality in Maintenance Engineering. 2001; 7:286-301.

16. Jozwiak, I. J. introduction to the studies of reliability of systems using the Weibull proportional hazards model. Microelectronics and Reliability. 1997; 37:915-18.

17. Kalbfleisch, John D. and Prentice, Ross L. The Statistical Analysis of Failure Time Data. New York: Wiley; 1980.

18. Kay, R. Goodness of fit methods for the proportional Hazarsds Regression model : A review. Revue D'Épidémiologie Et De Santé Publique. 1984; 32:185-198.

19. Kumar, Dhananjay and Westberg, Ulf. Proportional Hazards Modeling of Time-Dependent Covariates Using Linear Regression: A Case Study. IEEE Transactions on Reliability . 1996; 45(3):386-392.

20. Leemis, L. M. Reliability- Probabilistic Models and Strategical Methods. New Jersey: Prentice all, Inc., Engelwood Cliffs; 1995.

21. Lin, D.; Banjevic, D., and Jardine, A. K. S. Using principal components in a proportional hazards model with application in condition-based maintenance. Journal of the Operational Research Society. 2005; (1-10).

22. Makis, V. and Jardine, A. K. S. Computation of optimal policies in replacement models. IMA Journal of Mathematics Applied in Business & Industry. 1992; 3(169-175).

23. Makis, V. and Jiang X. Optimal Replacement Under Partial Observations. Mathematics of Operations Research. 2003; 28(2):382-394.

24. Mirham G. A. Simulation Statistical Foundations and Methodology.: Academic Press; 1972.

25. Press W. H. ; Teukolsky W. ; Vetterling W.T., and Flannery B.P. Numerical Recipes Third Edition. New York: Cambridge University Press; 2007.

26. Wang, w and Christer, A. H. Towards a general condition based maintenance model for a stochastic dynamic system. Journal of the Operational Research Society. 2000; 51:145-155.

CHAPTER 5 : SUMMARY AND CONCLUSION

In this thesis, we consider a CBM model based on the PHM proposed by Cox [1972], and we assumed that the condition monitoring is indirect i.e. an indicator's value θ, of the unobservable degradation state is available at each observation moment. Observations are collected at constant (or near constant) interval Δ. In this study, Z represents the degradation state of the equipment which will be used as the diagnostic covariate in the PHM, and θ is a value from the set of all the possible indicator's values $\Theta = \{1,...,M\}$. The whole set of the indicator's values is descritized into a finite set of M possible values. The equipment's condition is described as follows:

- The equipment has a finite and known number of degradation states N. $J = \{1,...,N\}$ is the set of all possible degradation states;
- Degradation state transition follows a Markov Chain with unobservable states and is modeled by a Hidden Markov Model (HMM). The Markovian transition matrix is $P = \left[p_{ij}\right]$, where p_{ij} is the probability of going from state i to state j, $i, j \in J$ during one observation interval, knowing that the equipment has not failed before the end of the interval;
- The value of the indicator is stochastically related to the equipment's state through the observation probability matrix $Q = \left[q_{j\theta}\right]$, $j \in J$, $\theta \in \Theta$. $q_{j\theta}$ is the probability of getting indicator's value θ, while the equipment is in state j;

- The indicator is collected periodically at intervals Δ;
- Failure is not a degradation state. It is a non-working condition of the equipment that can happen at any time and while the system is in any degradation state, and is known immediately (obvious failure).
- The system's failure rate follows a time-dependent PHM where the failure rate $h(s,Z_k)= h_0(s)\psi(Z_k)$ is a product of two independent functions. $h_0(.)$ is a function of the system's age only and $\psi(.)$ is a function of the system's degradation state only.

Since the degradation state of the equipment is not observable, we adapted an alternative state space π^k, the *Conditional Probability Distribution of the system's degradation State* (CPDS) at period k, as introduced by Ghasemi *et al.* [2007]:

$$\pi^k = \left\{\pi_i^k;\ \ 0 \le \pi_i^k \le 1 \text{ for } i=1,\dots,N, \sum_{i=1}^{N}\pi_i^k = 1\right\},\ k=0,1,2,\dots \qquad \textbf{(5-1)}$$

π_i^k represents the probability of being at state i at the k-th inspection moment. The initial value of the CPDS for a new system is defined as $\pi_i^0 = \begin{cases} 1 & i=1 \\ 0 & 1<i\le N \end{cases}$.

At each observation moment, after collection of an observation θ, the CPDS is updated by considering the latest observation θ as follows:

$$\pi_j^{k+1}(\theta) = \sum_{i=1}^{N}\pi_i^k p_{ij} q_{j\theta} \Big/ \sum_{i=1}^{N}\sum_{j=1}^{N}\pi_i^k p_{ij} q_{j\theta},\ j=1,\dots,N \qquad \textbf{(5-2)}$$

In the following subsections, we explain the objectives of the thesis and the corresponding methods and solutions that were developed for each objective.

5.1 Objective 1: Optimal inspection interval and optimal replacement policy

As the first objective of this thesis, we addressed the problem of optimal inspection interval and introduced the corresponding optimal replacement policy when the inspections are costly. The cost for a preventive replacement is C, while a failure replacement costs $K+C, K, C>0$. Both actions, failure replacement and preventive replacement, are instantaneous. The inspection costs C_I, independently of the inspection's interval.

We defined $V(k,\pi^k)$ as the minimum cost of *maintenance and inspection* over the renewal period, while the system is in the k-th inspection point and the CPDS is π^k. The *renewal period* is the time between two consecutive replacements, whether failure or preventive replacements.

$$V\left(k,\pi^k\right)=\min\left\{kC_I+C+V\left(0,\pi^0\right),W\left(k,\pi^k,g\right)\right\} \tag{5-3}$$

where $kC_I+C+V\left(0,\pi^0\right)$ is the total cost over the renewal period at the k-th inspection moment, if the decision is to replace preventively and:

$$\begin{aligned}W\left(k,\pi^k,g\right)=&\left[kC_I+K+C+V\left(0,\pi^0\right)\right]\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]-g\overline{\tau}\left(k,\pi^k,\Delta\right)\\&+\left[\sum_{\theta=1}^{M}V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta\mid k,\pi^k\right)\right]\overline{R}\left(k,\pi^k,\Delta\right)\end{aligned} \tag{5-4}$$

$\left[kC_I + K + C + V\left(0,\pi^0\right)\right]$ represents the renewal period's total cost if the decision is "Do-nothing" (decision= ∞) and the system fails during the next inspection period.

$\left[\sum_{\theta=1}^{M} V\left(k+1,\pi^{k+1}(\theta)\right)\Pr\left(\theta \mid k,\pi^k\right)\right]$ is the expected total future cost of the model at the $k+1$ inspection moment, provided that the failure has not happened during the k-th period. $\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right]$ and $\overline{R}\left(k,\pi^k,\Delta\right)$ are the probability of having the failure during the k-th period and the probability that the system is still working at the beginning of the $k+1^{\text{st}}$ period consecutively while the CPDS at period k is π^k. $\overline{\tau}\left(k,\pi^k,\Delta\right)$ and g are the mean sojourn time of the system at the $k+1^{\text{st}}$ period when the CPDS at the k-th period π^k, is available and, the average cost per unit of time over infinite horizon respectively. g includes the cost of replacements only and excludes the inspections cost.

$\overline{R}\left(k,\pi^k,\Delta\right)$ and $\overline{\tau}\left(k,\pi^k,\Delta\right)$ are calculated by:

$$\overline{R}\left(k,\pi^k,\Delta\right) = \sum_{i=1}^{N} R\left(k,i,\Delta\right)\ \pi_i^k \tag{5-5}$$

$$\overline{\tau}\left(k,\pi^k,\Delta\right) = \int_0^{\Delta} \overline{R}\left(k,\pi^k,s\right)\ ds \tag{5-6}$$

where $R\left(k,i,t\right) = \exp\left(-\psi\left(i\right)\int_{k\Delta}^{k\Delta+t} h_0(s)ds\right)$.

Under these assumptions and by using the developed model, the replacement criterion is found and is given as follows:

$$a\left(k,\pi^k\right) = \begin{cases} \infty & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right] < g\overline{\tau}\left(k,\pi^k,\Delta\right) \\ 0 & \text{if} \quad K\left[1-\overline{R}\left(k,\pi^k,\Delta\right)\right] \geq g\overline{\tau}\left(k,\pi^k,\Delta\right) \end{cases} \tag{5-7}$$

This criterion is a function of K, the failure cost; $\overline{R}\left(k,\pi^k,\Delta\right)$ and $\overline{\tau}\left(k,\pi^k,\Delta\right)$, the conditional reliability and conditional sojourn at the k-th observation moment, consequently; and g, the long run average cost of replacement.

A recursive method to find g^*, the minimum long-run average cost of replacement is introduced. By replacing g with g^*, the optimal replacement policy is obtained from equation (5-7).

The optimal inspection period is chosen from a finite set of L possible inspection intervals $\Delta_l; l = 1,2,...,L$. The optimal inspection period is the one with the minimum total long-run average cost G^* where $G_l^* = g_l^* + \frac{C_I}{\Delta_l}$. In practice, after finding the G^* and by fixing Δ to the corresponding Δ^*, the optimal replacement policy based on the earlier results in this study is found.

5.2 *Objective 2: Reliability Function and Mean Residual Life*

In chapter three, we have calculated the Reliability Function (RF) and Mean Residual Life (MRL) of the assumed equipment. The MRL and the RF can be used as a supplementary decision tool, in particular when the cost elements of preventive replacement are unknown, or there are criteria other than the cost to respect. By knowing the MRL and the RF a practitioner can take advantage of the upcoming maintenance events (like a scheduled shutdown of production line), that are not usually considered in cost optimal replacement criteria, to perform a CBM and to improve the availability of the equipment.

The conditional reliability at (k, Z_k), i.e. at the k-th observation moment while the state is Z_k and for $t > \Delta$, is formulated by the following equation:

$$\begin{aligned} R(k, Z_k, t) &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_1, Z_2, \dots, Z_k\right), t > \Delta \\ &= \Pr\left(T > k\Delta + t \mid T > k\Delta, Z_k\right), t > \Delta \end{aligned} \tag{5-8}$$

In the case of direct observation assuming $Z_k = i$, we have shown that:

$$R(k,i,t) = \begin{cases} \exp\left(-\psi(i) \int_{k\Delta}^{k\Delta+t} h_0(s)ds\right) & 0 < t \le \Delta \\ R(k,i,\Delta)\sum_{j=1}^{N} p_{ij} R\left(k+1, j, (t-\Delta)\right) & t > \Delta \end{cases} \tag{5-9}$$

In the case of indirect information, we defined $\overline{R}\left(k, \pi^k, t\right)$ as the conditional reliability of the equipment at the k-th observation moment, while the state conditional probability distribution is π^k. $\overline{R}\left(k, \pi^k, t\right)$ is then calculated as follows:

$$\overline{R}(k, \pi^k, t) = \begin{cases} \sum_{i=1}^{N} \pi_i^k \exp\left(-\psi(i) \int_{k\Delta}^{k\Delta+t} h_0(s)ds\right) & 0 < t \le \Delta \\ \sum_{i=1}^{N} \pi_i^k R(k,i,\Delta) \sum_{j=1}^{N} p_{ij} \times R\left(k+1, j, (t-\Delta)\right) & t > \Delta \end{cases} \tag{5-10}$$

In the case of direct observations, the MRL is given as follows:

$$e(k,i) = \int_{k\Delta}^{\infty} R(k,i,t)\, dt$$

We calculated the MRL, $\overline{e}\left(k, \pi^k\right)$, at the k^{th} observation moment, while the state conditional probability is π^k, and proved that it can be represented as follows:

$$\overline{e}\left(k, \pi^k\right) = \int_{k\Delta}^{\infty} \overline{R}\left(k, \pi^k, t\right) dt \tag{5-11}$$

5.3 *Objective 3: Estimation of the model's Parameter*

Any mathematical model, including the models introduced in the two previously mentioned objectives, is based on a set of unknown parameters that needs to be estimated in order to apply the model in a real life problem. This estimation is based on the available data of the system under study. In the chapter four, we introduced methods to estimate the parameters of the models under study. To apply any of these models into a real life problem, one needs to estimate the parameters of the PHM, the probabilities of the Markov transition matrix, and the probabilities of the state-indicator matrix.

We defined T to be the lifetime of the equipment, which is an i.i.d., non-negative continuous random variable, and $\theta(s)=\left\{\theta^1,\theta^2,...,\theta^k\right\};s\le T;k=1,2,...;k\Delta\le s$ the history of the indicator's values up to time s, where Δ is the observation interval. The indicator's values up to time s, can be mapped into the state conditional probability distribution up to time s by $\pi(s)=\left\{\pi^1,\pi^2,...,\pi^k\right\};\ s\le T;\ k=1,2,...;k\Delta\le s$, where the elements of $\pi(s)$ are calculated from equations (5-1) and (5-2) at the corresponding observation moment, when a new indicator's value is available.

We proved that the lifetime's survival function is:

$$R(t,\theta(t))=$$

$$\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1}\,|\,T>(l+1)\Delta,\pi^l\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^l\exp\left(-\int_{l\Delta}^{(l+1)\Delta}h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^k\exp\left(-\int_{k\Delta}^{t}h(\tau,i)\,d\tau\right)$$

(5-12)

and since $f(t,\theta(t))=\lim\limits_{\Delta t\to 0^+}\dfrac{\Pr\left(t\le T<t+\Delta t\,|\,T>t,\theta(t)\right)}{\Delta t}=-\dfrac{dR(t,\theta(t))}{dt}$:

$$f\left(t,\theta(t)\right)=$$

$$\left[\prod_{l=0}^{k-1}\Pr\left(\theta^{l+1}\,|\,T>(l+1)\Delta,\pi^{l}\right)\right]\left[\prod_{l=0}^{k-1}\sum_{i=1}^{N}\pi_i^{l}\exp\left(-\int_{l\Delta}^{(l+1)\Delta}h(\tau,i)\,d\tau\right)\right]\sum_{i=1}^{N}\pi_i^{k}\exp\left(-\int_{k\Delta}^{t}h(\tau,i)\,d\tau\right)h(t,i)$$

(5-13)

We addressed the parameter estimation problem in two categories: with and without censoring.

In the case of uncensored data, for a set of n independent experiments, we assume that T_r is the time to failure of r^{th} experiment. The likelihood of the set of unknown parameters Ω, based on the available data can be calculated by:

$$L(\Omega)=\prod_{r=1}^{n}f\left(T_r,\theta(T_r);\Omega\right) \quad \textbf{(5-14)}$$

In the case of censored data, we have defined a censoring indicator $\delta_r=\begin{cases}1 & \text{if the item has failed}\\ 0 & \text{if the item is censored}\end{cases}$, which indicates whether the value of T_r is an actual failure time or it is the moment at which we have randomly stopped the experiment i.e. T_r is the censoring time. Then the likelihood of the set of unknown parameters Ω based on the available data can be given by:

$$L(\Omega)=\prod_{r=1}^{n}f\left(T_r,\theta;\Omega\right)^{\delta_r}R\left(T_r,\theta;\Omega\right)^{1-\delta_r} \quad \textbf{(5-15)}$$

In both cases, with censored and uncensored data, by applying a maximization technique such as the line search method, the maximum of the log likelihood functions is calculated.

Convergence and robustness of the introduced estimation methods have been studied by using a Monte Carlo simulation technique. Based on this study, the parameters' values obtained by the developed methods converge to the real values when the size of the data sample increases. The estimation method gives very good results even with 50% censoring data.

This thesis provides a package of useful utilities in condition monitoring studies. It introduces an optimal criterion to replace the equipment as well as a method to find the optimal inspection interval. It also offers two measures of the future performance of the equipment, i.e. the MRL and the RF. These measures help the practitioners in taking more accurate decisions concerning the equipment's maintenance. Finally, the parameters' estimation techniques allow the practitioners to use the proposed models in real cases and also, answer to the problem of finding the best parameters' value, a problem that was not addressed in many published researches. Several numerical examples are solved and the results are discussed.

Areas of future studies are to expand the results obtained in this thesis to the case with non fixed inspection intervals. In this case, the next inspection moment and the replacement policy that minimizes the maintenance cost have to be determined at each inspection moment. Based on the replacement policy at each inspection moment, one will decide whether to replace the system or leave it work until next inspection point. In the latter case, the next observation point will be set at current inspection point based on the available condition monitoring data up to date. The time between inspections in this case may be non-equal.

The case in which, a decision can be made among several repair possibilities can be considered as another realistic case of future research. In many practical cases, different partial repairs for the equipments which cause different costs may take place. Each specific partial repair will changes the degradation state of the system from a current state to a probable pre-known state.

In this thesis, we assumed a homogeneous Markov Model. The case of the non homogeneous Markov Model is also an interesting area of prospective research.

REFERENCES:

1. Adjengue L. and Yacout S. (2005) Parameters estimation for condition based maintenance with Uncorrelated and Corelated Observations. *Quality Engineering* **19**, 197-206.

2. Ansell, J.I. and Phillips, M.J. (1997) Practical aspects of modeling of repairable systems data using proportional hazard models. *Reliability Engineering and System Safety* **58**, 165-71.

3. Aven, T. (1996) Condition based replacement policies A counting process approach. *Reliability Engineering and System Safety* **51**, 275-442.

4. Banjevic, D. and Jardine, A.K.S. (2006) Calculation of Reliability Function and Remaining Useful Life. *IMA Journal of Management Mathematics* **17**, 115-130.

5. Banjevic, D., Jardine, A.K.S., Makis, V. and Ennis, M. (2001) A control-limit policy and software for condition-based maintenance optimization. *INFOR Journal* **39**, 32-50.

6. Barbera, F., Schneider, H. and Kelle, P. (1996) A condition based maintenance model for a two-unit series system. *European Journal of Operational Research* **47**, 1037-1045.

7. Barros, A., B'Erenguer, C. and Grall, A. (2002) Optimization of replacement times using imperfect monitoring information. *IEEE Transactions on Reliability.*

8. Basawa, I.V. and Rao, P. (1980) *Statistical Inference for Stocastic Processes*, edn. London: Academic Press.

9. Bellman R E (1957) Dynamic Programming. NJ: Princeton.

10. Bertsekas D. P. (1976) *Dynamic programming and stochastic control*, New York: Academic Press.

11. Box M.J., Davies D. and Swann W.H. (1969) *Non-Linear Optimisation Techniques*, edn. London: Oliver & Boyd.

12. Bradley, D.M. and Gupta, R.C. (2003) Limiting Behavior of The Mean Residual Life. *Ann. Inst. Statist. Math.* **55**, 217-226.

13. Butcher, S. W. Assessment of Condition-Based Maintenance in the Department of Defense. 2000. 2005.

14. Byrne, C.L. (2005) *Signal Processing : a mathematical approach*, Wellesley, Mass.: A.K. Peters.

15. Campodonico, S. and Singpurwalla, N.D. (1994) *The signature as a covariate in reliability and biometry*, London: JohnWiley.

16. Chelbi, A. and Ait-Kadi, D. (1999) An optimal inspection strategy for randomly failing equipment . *Reliability Engineering and System Safety* **63**, 127-131.

17. Cho, D.I. and Parlar, M. (1991) A Survey of Maintenance Models for Multi-Unit Systems. *European Journal of Operational Research* **51**, 1-23.

18. Christer A. H. and Wang W. (1992) A model of condition monitoring of production plant. *International Journal of Production Research* **30**, 2199-2211.

19. Christer, A.H. and Wang, W. (1995) A simple condition monitoring model for a direct monitoring process. *European Journal of Operational Research* **82**, 258-269.

20. Christer, A.H., Wang, W. and Sharp, J. (1997) A state space condition monitoring model for furnace erosion prediction and replacement. *European Journal of Operational Research* **101**, 1-14.

21. Coolen, P.A.F. and Dekker, R. (1995) Analysis of a 2-phase model for optimization of condition-monitoring intervals. *IEEE Transactions on Reliability* **44**, 505-511.

22. Cox, D.R. (1972) Regression models and life tables (with discussion). *Journal of the Royal Statistical Society - Series B* **26**,

23. Cox, D.R. (1962) *Renewal Theory*, London: Chapman and Hall.

24. Crowley, J. and Hu, M. (1977) Covariance analysis of heart transplant survival data. *Journal of American Statistical Association* **72**, 27-36.

25. Dekker, R. (1996) Applications of maintenance optimization models: A review and analysis. *Reliability Engineering and System Safety* **51**, 229-240.

26. Dieulle, L., Berenguer, C., Grall, A. and Roussignol, M. (2003) Sequential condition-based maintenance scheduling for a deteriorating system. *European Journal of Operational Research* **150**, 451-461.

27. Elsayed, E.A. and Chan, C.K. (1990) Estimation of Thin-Oxid Reliability Using Proportional Hazards Models. *IEEE Transactions on Reliability* **39**, 329-335.

28. Fernandez-Gaucherand, E. (1993) Analysis of an Adaptive Control Scheme for a Partially Observed Controlled Markov Chain. *IEEE Transactions on Automatic Control* **38**, 987-993.

29. Gertsbakh, I. (2000) *Reliability Theory - With Applications to Preventive Mainten ance*, Springer.

30. Ghai, G.L. and Mi, J. (1999) Mean Residual Life and Its Association With Failure Rate. *IEEE Transactions on Reliability* **48**, 262-266.

31. Ghasemi, A., Yacout, S. and Ouali, M.S. (2008) Evaluating the remaining life for equipment with unobservable states. *IEEE Transactions on Reliability* **(submitted)**,

32. Ghasemi, A., Yacout, S. and Ouali, M.S. (2007) Optimal condition based maintenance with imperfect information and the proportional hazards model. *International Journal of Production Research* **45**, 989-1112.

33. Ghasemi, A., Yacout S. and Ouali M.S. (2008) Optimal inspection period and replacement policy for CBM with imperfect information using PHM. In: *Current Themes in Engineering Technologies*, pp. 247-266. Melville, NY: American Institute of Physics]

34. Parameter Estimation algorithms of a Condition Based Maintenance with Imperfect Information and Proportional Hazard Model. edn. (2008)

35. Ghasemi, A. (2005) Condition Based Maintenance Using the Proportional Hazards Model with Imperfect Information. Ecole Polytechnique de Montreal.

36. Grall, A., Berenguer, C. and Dieulle L. (2002) A condition-based maintenance policy for stochastically deteriorating systems. *Reliability Engineering and System Safety* **76**, 167-180.

37. Hontelez, A.M., Burger, H.H. and Wijnmalen, D.J.D. (1996) Optimum condition-based maintenance policies for deteriorating systems with partial information. *Reliability Engineering and System Safety* **51**, 267-274.

38. Hopp W. and Kuo Y. L. (1998) An optimal structured policy for maintenance of partially observed aircraft engine components. *Naval Res Logist* **45**,

39. Howard Ronald A. (1960) *Dynamic Programming and Markov Processes*, edn. The M.I.T. Press.

40. Jardine A. K. S. and Anderson M (1985) Use of concomitant variables for reliability estimation. *Maintenance Management International* **5**, 135-140.

41. Jardine, A.K.S., Anderson, P.M. and Mann D. S. (1987) Application of the Weibull Proportional Hazards Model to Aircraft and Marine Engine Failure Data. *Quality and Reliability Engineering International* **3**, 77-82.

42. Jardine, A.K.S., Banjevic, D., Wiseman, M., Buck, S. and Joseph, T. (2001) Optimizing a mine haul truck wheel motors' condition monitoring program. *Journal of Quality in Maintenance Engineering* **7**, 286-301.

43. Jardine A.K.S. , Lin, D. and Banjvic, D. (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing* **20**, 1483-1510.

44. Jozwiak, I.J. (1997) introduction to the studies of reliability of systems using the Weibull proportional hazards model. *Microelectronics and Reliability* **37**, 915-18.

45. Kalbfleisch, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*, New York: Wiley.

46. Kay, R. (1984) Goodness of fit methods for the proportional Hazarsds Regression model : A review. *Revue d'épidémiologie et de santé publique* **32**, 185-198.

47. Kumar, D. and Klefsjo, B. (1994) Proportional hazards model : a review. *Reliability Engineering and System Safety* **44**, 177-188.

48. Kumar, D. and Westberg, U. (1997) Maintenance scheduling under age replacement policy using proportional hazards model and TTT-plotting. *European Journal of Operational Research* **99**, 507-515.

49. Kumar, D. and Westberg, U. (1996) Proportional Hazards Modeling of Time-Dependent Covariates Using Linear Regression: A Case Study. *IEEE Transactions on Reliability* **45**, 386-392.

50. Lam, T.C. and Yeh, R.H. (1994) Optimal Maintenance-Policies For Deteriorating Systems Under Various Maintenance Strategies. *IEEE Transactions on Reliability* **43**, 423-430.

51. Lawless, J., Hu, J. and Cao, L. (1995) Methods for estimation of failure distributions and rates from automobile warranty data. *Lifetime Data Anal* **1**, 227-240.

52. Leemis, L.M. (1995) *Reliability- Probabilistic Models and Strategical Methods*, New Jersey: Prentice all, Inc., Engelwood Cliffs.

53. Lim, J.-H. and Park, D.H. (1995) Trend Change in Mean Residual Life. *IEEE Transactions on Reliability* **44**, 291-296.

54. Lin, D., Banjevic, D. and Jardine, A.K.S. (2005) Using principal components in a proportional hazards model with application in condition-based maintenance. *Journal of the Operational Research Society*

55. Lin, D. and Makis, V. (2004) On-Line Parameter Estimation For a Failure-Prone System Subject to Condition Monitoring. *Journal of Applied Probability* **41**, 1-10.

56. Lin, D. and Makis, V. (2003) Recursive Filters For A Partially Observable System Subject to Random Failure. *Adv. Appl. Prob.* **35**, 207-227.

57. Maguluri, G. and Zhang, C.H. (1994) Estimation in mean residual life regression model. *J. R. Statist. Ser. B* **56**, 477-489.

58. Maillart, L.M. (2004) Technical Memorandum Number 778,

59. Makis, V. and Jardine, A.K.S. (1992) Computation of optimal policies in replacement models. *IMA Journal of Mathematics Applied in Business & Industry* **3**,

60. Makis, V. and Jardine, A.K.S. (1992) Optimal replacement in the proportional hazards model. *INFOR Journal* **30**, 172

61. Makis, V. and Jiang X. (2003) Optimal Replacement Under Partial Observations. *Mathematics of Operations Research* **28**, 382-394.

62. Marcus, B., Olsson, E., Funk, P., and Jackson, M. Technical Design of Condition Based Maintenance System -A Case Study using Sound Analysis and Case-Based Reasoning. 2002. 2005.

63. McCall, J.J. (1965) Maintenance policies for stochastically failing equipment: A survey. *Management Science* **11**, 493-524.

64. Mine H. and Kawai H. (1975) An optimal inspection and replacement policy. *IEEE Trans. Reliability* **R-24**, 305-309.

65. Mirham G. A. (1972) Simulation Statistical Foundations and Methodology. Academic Press.

66. Mitchell, J. S. Five to ten year vision for CBM, ATP Fall Meeting Condition Based Maintenance Workshop. 98. 2005.

67. Mobley K. R. (2002) *Introduction to Predictive Maintenance*, 2 edn. Van Nostrand Reinhold's plant engineering series:

68. Murthy, D.N.P., Iskander, B.P. and Wilson, R.J. (1995) Two-dimensional failure-free warranty policies: two-dimensional point process models. *Operations Research* **43**, 356-366.

69. Nemhauser G. L. (1966) Introduction To Dyanamic Programming. NY: John Wiley & Sons.

70. Ohnishi, M., Kawai, H. and Mine, H. (1986a) Optimal inspection and replacement policy for a deteriorating system. *Journal of Applied Probability* **23**, 973-988.

71. Ohnishi, M., Kawai, H. and Mine, H. (1986) An optimal inspection and replacement policy under incomplete state information. *European Journal of Operational Research* **27**, 117-128.

72. Park, K.S. (1988) Optimal Continuous-Wear Limit Replacement under Periodic Inspections. *IEEE Transactions on Reliability* **37**, 97-102.

73. Park, K.S. (1988a) Optimal Wear-Limit Replacement with Wear-Dependent Failure. *IEEE Transactions on Reliability* **37**, 293-294.

74. Press W. H. , Teukolsky W. , Vetterling W.T. and Flannery B.P. (2007) *Numerical Recipes Third Edition*, edn. New York: Cambridge University Press.

75. Rao, B.K.N. (1996) *Handbook of Condition Monitoring*, Elsevier.

76. Rosqvist, T. (2002) Stopping time optimisation in condition monitoring. *Reliability Engineering and System Safety* **76**, 319-325.

77. Ross, S.M. (1997) *Introduction to probability models*, 6 edn. San Diego, Calif; Toronto: Academic Press.

78. Scarf, P.A. (1997) On the application of mathematical models in maintenance. *European Journal of Operational Research* **99**, 493-506.

79. Scott, G., Frickenstein, L. and Whitaker, R. (2003) Age Replacement Policies in Two Time Scales. *Naval Research Logistics* **50**, 592-613.

80. Sen, P.K. (2004) HQ|RQoL and concomitant adjusted mean residual life analysis. In: Nikulin, M.S. and Balakrishnan, N., (Eds.) *Parametric and semiparametric models with applications to reliability, survival analysis, and quality of life*, pp. 349-362. Boston, Birkhauser:

81. Siddiqui, M.M. and Caglar, M. (1994) Residual Lifetime Distribution and Its Applications. *Microelectronics reliability* **34**, 211-227.

82. Singpurwalla, N.D. and Wilson, S.P. (1998) Failure models indexed by two scales. *Adv. Appl. Prob.* **30**,

83. Tang, L.C., Lu, Y. and Chew, E.P. (1999) Mean Residual Life of Lifetime Distributions. *IEEE Transactions on Reliability* **48**, 73-78.

84. Tijms H. and Schouten F. (1984) A Markov decision algorithm for optimal inspection and revisions in a maintenance system with partial information. *European Journal of Operation Research* **21**, 245-253.

85. Valdez-Flores, C. and Feldman, R.M. (1989) A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics* **36**, 419-446.

86. Wang, w. and Christer, A.H. (2000) Towards a general condition based maintenance model for a stochastic dynamic system. *Journal of the Operational Research Society* **51**, 145-155.

87. Wang W and Zhang W (2008) An asset residual life prediction model based on expert judgment. *European Journal of Operational Research* 188, 496-505.

88. Wang, W. (2002) A model to predict the residual life of rolling element bearing given monitored condition information to date. *IMA Journal of Management Mathematics* 13, 3-16.

89. Wang, W. and Zhang, W. (2005) A model To Predict the Residual Life of Aircraft Engines Based upon Oil Analysis Data. *Naval Research Logistics* **52**,

90. Wijnmalen D. and Hontelez J. (1992) Review of a Markov decision algorithm for optimal inspections and revisions in a maintenance system with partial information. *European Journal of Operation Research* **62**, 96-104.

91. Zilla S (1993) Replacement policy under partially observed Markov process. *Int. J. of Prod. Econo* **29**.