République Algérienne Démocratique et Populaire

الجمهورية الجزائرية الديمقراطية الشعبية

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

وزارة التعليـــم العالـــي والبحـــث العلمـــي

UNIVERSITE MOHAMED SEDDIK BEN YAHIA - JIJEL

Faculté des Sciences et de Technologie

Département de Génie Mécanique

Projet de Fin d'Études

Pour l'obtention du diplôme de Master

En Génie Mécanique

Option Construction Mécanique

Thème

Utilisation et traitement des données incomplètes pour l'étude des actions de maintenance dans les systèmes industriels : modèles de régression et d'interpolation

Les membres du jury :

M. DELIOU Adel	***Président***
M. BELHADEF Rachid	***Encadreur***
M. FENNINECH Hocine	***Examinateur***

Présenté par :

AMARI Faris

SAMAH Mohammed Aymen

Promotion 2021

Université Mohamed Seddik Ben Yahia -Jijel

Soutenue le 09/09/2021

Remerciement

En préambule à ce mémoire nous remercions ﷲ qui nous aide et nous a donné la patience et le courage durant ces longues années d'études.

Nous tenons à remercier tout particulièrement notre encadreur Belhadef Rachid pour nous avoir permis de réaliser ce travail et pour son aide, son accueil et sa disponibilité.

Nous sincères remerciements s'adressent aux membres de jury qui ont accepté de lire et d'évaluer ce travail. Nous remercions chaleureusement nos familles et surtout nos parents pour leur soutien moral, leurs encouragements et leur patience durant les étapes difficiles de ce travail.

Nous tenons à remercier toute personne qui a participé de prés ou de loin à la réalisation de ce modeste travail.

Merci à tous nous collègues pour tous les bons moments passés, et les échanges scientifiques qu'on a partagés, leurs disponibilités et leurs compétences, merci du fond du cœur

Table des matières

Introduction Générale……..1

Chapitre I : Méthodes de traitement des données

I.1. Introduction 3
I.2. L'utilité de traitement des données 3
I.3. Méthodes de traitement des données manquantes 4
I.3.1. Données aberrantes 5
I.3.2. Données manquantes 5
I.3.3. Les différentes méthodes de traitement des données incomplètes 6
I.3.3.1. Procédures de suppression 7
I.3.3.1.a.Étude des cas complets 7
I.3.3.1.b. Étude des cas disponibles 8
I.3.3.2. Procédures de remplacement 9
I.3.3.3. Procédures basées sur un modèle 10
I.3.3.3.a. Maximum de vraisemblance 10
I.3.3.3.b.Maximisation d'espérance 10
I.4. Choix méthodologique 11
I.5. Conclusion 12

Chapitre II : Maintenance des systèmes industriels

II.1. Introduction 14
II.2. Maintenance de systèmes industriels 15
II.2.1. Présentation de la maintenance 15
II.2.1.1. Définition et objectifs de la maintenance 17
II.2.1.2. Les type de maintenance 18
II.2.1.2.a.La maintenance corrective 19
II.2.1.2.b.La maintenance préventive 19
II.2.1.3. Les approches industrielles 20
II.2.2. Modélisation et évaluation de la maintenance 22
II.2.2.1. Modèles de maintenance 22
II.2.2.2. Optimisation de la maintenance 24

II.2.3. Conclusion et problèmes posés 25
II.3. Maintenance basé sur les méthodes de fiabilité 26
II.3.1. Introduction 26
II.3.2. Méthodologie 26
II.3.2.1. Représentation du système étudié en schémas blocs 27
II.3.2.2. Collecte de données 28
II.3.2.3. Estimation de la loi de fiabilité de chaque composant 28
II.3.2.4. Optimisation de la maintenance 30
II.3.2.5. Incertitudes sur les résultats 30
II.4. Conclusion 30

Chapitre III : Méthode de maintenance basée sur données incomplète

III.1. Introduction 33
III.2.1. Modèles de fiabilité usuels 34
III.2.2. Fiabilité des systèmes complexes 36
III.2.3. Estimation des paramètres des modèles de fiabilité 37
III.2.3.2. Méthode du maximum de vraisemblance 38

chapitrev IV:application et resultas

IV.1 Introduction 42
IV.2 Modèle de régression (linéaire) 43
IV.3. Méthode d'interpolation linéaire 45
IV.4 Procédure de régression et interpolation pour la modélisation de la fiabilité 46
IV.4.1 Principales lois utilisées pour modéliser la fiabilité 46
IV.4.2 La loi de WEIBULL 46
IV.5 Résultats et interprétation de l'étude de la fiabilité de la turbine à gaz 46
IV.5.1 Situation du problème et objectif de l'étude 46
IV.6. Partie pratique 47
IV.6.1. premier cas d'étude 47
IV.6.2. Etude par la méthode de régression 48
IV.6.2.1. Comment effectuer une régression linéaire avec MATLAB 48
IV.6.2.2. Remplacement des données perdues 50
IV.6.2.3. Recherche aux indicateurs de la fiabilité 52
IV.6.2.4. Le test d'adéquation de modèle. 55

IV.6.3.Etude par la méthode d'interpolation 56
IV.6.3.1. Comment effectuer une interpolation linéaire avec Matlab 56
IV.6.3.2. Remplacement les données perdues 57
IV.6.3.3 Le test d'adéquation de modèle 61
IV.6.4.Interpretation des résultats 61
IV.6.5.Deuxième cas d'étude.....62
IV.6.6.Etude par la méthode de régression 64
IV.6.6.1. Remplacement les données perdues 65
IV.6.7.Etude par la méthode d'interpolation 66
IV.6.7.1 Remplacement les données perdues 67
IV.6.7.2 Détermination des paramètres de Weibull par MATLAB 68
IV.6.7.3 Le test d'adéquation de modèle 71
IV.6.8.Interpretation des résultats 73
Conclusion Générale…….....74
Références Bibliographique……..……..……76
Annexes.....78

Liste des tableaux

Tableau I.1 : Exemple avec valeur aberrante. 5
Tableau I.2 : Exemple sans valeur aberrante. 5
Tableau I.3 : Valeaur manquante et donnée incompléte 6
Tableau IV.1: des données incomplètes 47
Tableau IV.2: Remplacement les données perdues pour régression 50
Tableau IV.3: Tableau des résultats pour régression.....53
Tableau IV.4: La densité de défaillance et le taux défaillance instantané (régression).....56
Tableau IV.5: Remplacement les données perdues pour interpolation 57
Tableau IV.6: Tableau des résultats pour interpolation.....59
Tableau IV.7: Test d'adéquation de modèle (interpolation).....61
Tableau IV.8: des données incomplètes 63
Tableau IV.9: Remplacement les données perdues pour régression 66
Tableau IV.10: Remplacement les données perdues pour interpolation 68
Tableau IV.11: Tableau des résultats pour interpolation.....69
Tableau IV.12 : resultats obtenus avec interpolation……..……...73

Liste des figures

Figure I. 1:Les grandes catégories des méthodes pour le traitement des données manquantes 7
Figure II.1: Les différentes stratégies de maintenance.....19
Figure II.2: La nécessité de contrôler l'équilibre entre maintenance préventive et maintenance corrective.....25
Figure II.3: Schéma synoptique de la méthodologie 27
Figure II.4: Les trois approches pour l'estimation de la fiabilité et de la durée de vie 29
Figure III.1:systèmes complexes de fiabilité.....36
Figure IV. 1: la relation entre TBFi (h) et F (i) (%).....48
Figure IV. 2: la courbe de régression.....49
Figure IV. 3: Détermination graphique des paramètres de Weibull 51
Figure IV. 4: La densité de défaillance et le taux défaillance instantané 54
Figure IV. 5: La fonction de fiabilité et de répartition (régression) 54
Figure IV. 6: la courbe d'interpolation 57
Figure IV. 7: Détermination graphique des paramètres de Weibull 58
Figure IV. 8: La densité de défaillance et le taux défaillance instantané (interpolation) 60
Figure IV. 9: La fonction de fiabilité et de répartition (interpolation.....60
Figure IV. 10: la relation entre TBFi (h) et F (i) (%) 63
Figure IV. 11: la courbe de régression.....64
Figure IV. 12: la courbe d'interpolation 66
Figure IV. 13: Papier de Weibull par MATLAB.....68
Figure IV. 14: La fonction de fiabilité et de répartition.....70
Figure IV. 15: La densité de défaillance et le taux défaillance instantané 70

Introduction Générale

La maintenance joue un grand rôle dans le domaine de l'industrie dont ses buts sont d'assurer la disponibilité des machines. Etablir ou définir une politique ou un système de maintenance pour une machine ou système industriel nécéssite l'utilisation et la maitraise de plusieures techniques. Parmis ces techniques la maintenace basée sur la fiabilité, une technique qui utilise les données et les informations collectées durant le fonctionnement de ce système.

Dans le domaine de la maintenance conditionnelle, le problème ne réside plus dans l'accès à l'information, mais plutôt dans la qualité de l'information, son filtrage et son prétraitement, puisque la qualité du diagnostic de l'état de l'équipement est définie par la qualité des données utilisées à l'entrée du système du diagnostic.

Il est très fréquent de confronter le problème de données manquantes. Ce phénomène se manifeste lorsque les valeurs n'ont pas pu être observées, elles ont été perdues ou elles n'étaient pas enregistrées. La présence de ces dernières entraîne un dysfonctionnement du processus de traitement logique des données.

Pour palier à ce genres de problemes et pour la technique de statistique appelée la méthode de régression et la technique d'interpolation, en l'absence de méthode de traitement des données numériques manquantes, l'élaboration d'une nouvelle méthode de prétraitement des données, apparaît comme une alternative de grand intérêt pour substituer les données manquantes et par la suite, générer des modeles qui mènent à la classification de l'état de l'équipement.

C'est dans ce contexte que s'inscrit notre travail. Nous proposons dans ce mémoire deux méthodes de substitution des valeurs manquantes à savoir la méthode de regression et la méthode d'interpolation. Le but est d'utiliser ces deux methodes pour remplacer les valeurs manquantes puis faire les calculs de fiabilité ces parametres d'une turbine à gaz et essayer d'établir des actions de maintenance correspondantes. Une compairison entre les deux methodes et fait pour voir les points forts et faibles de chaque methode.

Ce travail est structuré de la manière suivante :

1. Une inroduction générale pour exposer le probleme et présenter notre travail
2. Un premier chapitre consacré à la présentation des méthodes de traitement des données
3. Un deuxième chapitre parle de la maintenance des systèmes industriels
4. Un troisiéme chapitre consacré pour expliquer la méthode de maintenance basée sur données incomplète
5. Un quatriéme chapitre reservé pour discuter les résultats obtenus d'application des deux méthodes la régression et l'intrpolation
6. Une conclusion générale et perpectives

Chapitre I : Méthodes de traitement des données

Chapitre I : Méthodes de traitement des données

I.1. Introduction

La qualité de l'information et le problème qui nous préoccupe en maintenance systématique car elle affecte la précision de notre diagnostic. Nous rencontrons beaucoup le problème des données perdues. Cette situation se produit lorsque nous ne pouvons pas observer les valeurs, perdues ou non enregistrées. La présence de ces éléments entraîne un dysfonctionnement du processus logique de traitement des données.

I.2. L'utilité de traitement des données

D'après [1], les données ne sont pas supposées être parfaites, mais précises, complètes, consistantes, cohérentes, opportunes, et flexibles et ceci dans le but de satisfaire des besoins.

- Précises : les données ne doivent pas contenir des erreurs ;
- Complètes : toutes les valeurs doivent être présentes ;
- Cohérentes : les données doivent satisfaire un ensemble de contraintes et doivent représenter le phénomène qu'elles mesurent ;
- Opportunes : les données doivent être disponibles quand elles sont réclamées ;
- Flexibles : les données doivent être décrites de façon à ce qu'elles peuvent être analysées de différentes méthodes.

Le traitement des données représente environ 60 à 80 % du temps impliqué dans le processus d'extraction de l'information ; la difficulté dépend de la nature des problèmes existants avec les données à savoir :

- Les données erronées : par exemple une valeur quantitative à la place d'une valeur qualitative.
- Les données manquantes : absence de données / valeurs.
- Les données aberrantes : les données qui s'écartant de la norme, du modèle normal
- Les données doublons : données faussement dupliquées.

Donc le traitement des données doit être réalisé pour les raisons suivantes : la résolution des problèmes qui peuvent nous empêcher d'appliquer une technique de diagnostic, la compréhension de la nature des données pour pouvoir les analyser et l'extraction plus efficace de l'information.

I.3. Méthodes de traitement des données manquantes

Les méthodes de traitement des données manquantes se distinguent selon deux approches, les méthodes supprimant les données manquantes et les méthodes utilisant toute l'information disponible. Dans la première catégorie, l'on retrouve les techniques connues sous l'appellation analyse des cas complets (listwise délétion) et analyse des cas complets par paires (pairwise délétion). Certains auteurs suggèrent également que le choix de ne pas inclure dans les analyses les variables qui comportent des données manquantes constitue une forme de suppression des données. Parmi les méthodes utilisant toute l'information disponible, notons l'ajustement par variable binaire, toutes les variantes de l'imputation, le maximum de vraisemblance, l'algorithme EM, le Markov Chain Monte Carlo (MCMC), la pondération ainsi que l'imputation multiple. Il est à noter que cette liste de méthodes n'est pas exhaustive. Les méthodes utilisant toute l'information disponible sont généralement préférées. Ce chapitre couvre d'abord certaines méthodes classiques et se termine par les méthodes récentes qui seront appliquées par la suite.

a. Méthodes classiques

-L'imputation simple :

- Imputation d'uneconstante
- Imputation d'un nombrealéatoire
- Imputation d'un nombre non aléatoire

b. Méthodes récentes

-Le maximum de vraisemblance ;

-L'imputation multiple ;

-La pondération.

I.3.1. Données aberrantes

Les données aberrantes sont définies comme des données qui ne sont pas en accord avec la majorité des données, selon [2], toutefois on peut dire que les données aberrantes se trouveront au périphérique du nuage formé par l'ensemble des données, Les données aberrantes pouvant être causées soit par une raison physique connue, exemple le mauvais étalonnage de l`appareil de mesure, erreur d'écriture de données, soit par une raison non connue. Pour repérer les données aberrantes, un contrôle de cohérence sur les données est nécessaire, l'usage du bon sens et de l'expérience est le plus sûr. Par exemple, des données de température d'un roulement, très élevé ou très faible, en tout cas très différent de la grande majorité des valeurs, doivent conduire à s'interroger sur la validité de ces données, c'est pour cette raison que la réflexion sur la cohérence des données est essentielle. L'exemple suivant (tableau 1.1 et 1.2) semble un peu extrême, mais cette situation peut être rencontrée au moment du traitement des données.

Les quatre premières valeurs dans chaque colonne contiennent les mêmes mesures. Toutefois, dans le tableau 1.2, la cinquième entrée de l'attribut (a_1*) a un grand écart par rapport à la valeur dans le tableau 1.1.

Observation	Attribut a_1
w_1	1
w_2	3
w_3	5
w_4	9
w_5	12

Tableau I.1 : Exemple sans valeur aberrante.

Observation	Attribut a_1*
w_1	1
w_2	3
w_3	5
w_4	9
w_5	**120**

Tableau I.2 : Exemple avec valeur aberrante.

À noter que dans la présence d'une mesure aberrante, la médiane des données ne change pas. La médiane est robuste (généralement, il ne varie pas beaucoup) en présence d'un petit nombre de valeurs aberrantes, par contre la moyenne change rapidement.

I.3.2. Données manquantes

Une donnée incomplète est une donnée pour laquelle la valeur de certain attribut est inconnue, ces valeurs sont dites manquantes.

Soit l'observation est un vecteur des valeurs de certains indicateurs ou attributs, les valeurs manquantes peuvent être de deux natures :

-Valeur manquante totale, c'est- à-dire que toute l'observation manque.

-Valeur manquante partielle, c'est-à-dire que l'observation est présente mais il manque certaines valeurs de cette observation.

Exemple 1 : Dans le tableau 1.3, la valeur de l'attribut (a_4) pour l'observation (w_2) est manquante, la valeur n'est pas présentée et l'observation (w_2) est dite incomplète.

	Attributs			
Observations	$\boldsymbol{a_1}$	$\boldsymbol{a_2}$	$\boldsymbol{a_3}$	$\boldsymbol{a_4}$
$\boldsymbol{w_1}$	**56**	**98**	**10**	**5**
$\boldsymbol{w_2}$	**40**	**87**	**21**	**?**

Tableau I.3 : Valeaur manquante et donnée incompléte

Des valeurs manquent parce qu'elles n'ont pas pu être observées ; elles ont été perdues ou elles n'étaient pas enregistrées.

I.3.3. Les différentes méthodes de traitement des données incomplètes

De nombreuses techniques de traitement des données manquantes ont été développées dans les années 90, selon [3], sans prétendre être exhaustifs, en identifiaient déjà plus d'une vingtaine, pour la plupart issues des recherches en statistique. Depuis, les chercheurs en intelligence artificielle et fouille de données (Data mining), se sont mis à étudier la question et à développer de nouvelles techniques. Recenser l'ensemble de ces techniques serait fastidieux. Aussi avons-nous opté pour une mise en évidence des principales caractéristiques des différentes méthodes. Nous pouvons alors présenter les techniques les plus usitées et avoir ainsi une vue d'ensemble du domaine ; ces méthodes s'appliquent selon la nature du processus et parfois compte tenu du nombre d'observations. Il y a trois stratégies possibles pour traiter des données manquantes :

(a) Utilisation des procédures de suppression.

(b) Utilisation des procédures de remplacement, (substitution) les données manquantes par les valeurs présentes.

(c) Utilisation des procédures de modélisation de la distribution des données manquantes et les estimés par certains paramètres.

Cette première introduction conduit à la typologie de la figure 1. Nous allons maintenant nous focaliser sur les techniques de substitution correspondant à la stratégie (b), en essayant de dégager les caractéristiques qui permettent de les différencier. La technique étiquetée CD, pour Case Deletion ou suppression de cas, correspond à la stratégie (a) dans laquelle on se ramène à une base de données complète par suppression de toutes les observations contenant au moins une valeur manquante.

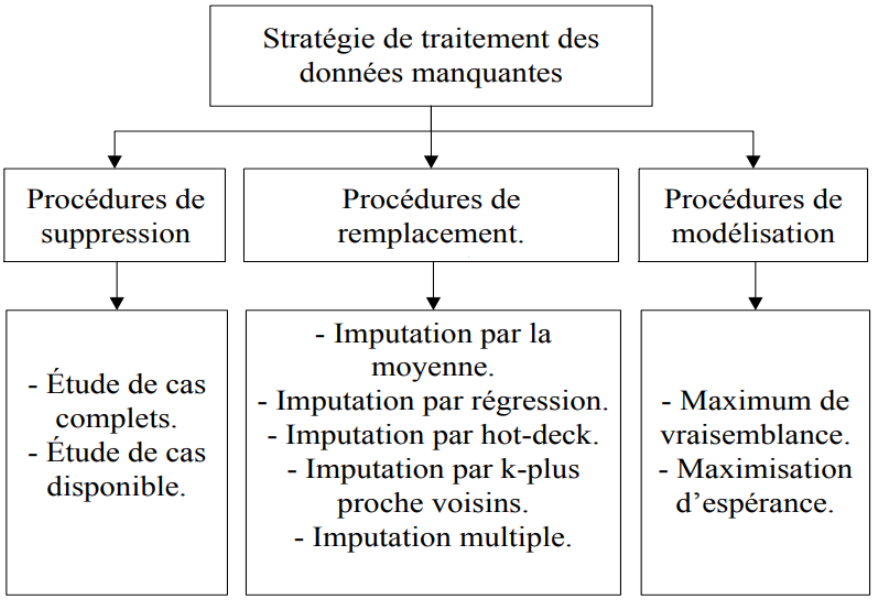


Figure I.1: Les grandes catégories des méthodes pour le traitement des données manquantes

I.3.3.1. Procédures de suppression

I.3.3.1.a.Étude des cas complets

Cette méthode permet de se ramener à une base de données complète par réduction de la dimension du problème. Pour cela, tous les exemples de la base contenant des valeurs manquantes sont supprimés (On peut également choisir de supprimer toutes les variables dont certaines observations manquent, mais il faut être prudent car certaines peuvent être essentielles pour l'analyse). Par conséquent, cette méthode sacrifie un grand nombre de données [4]. Selon [5], la suppression de 10 % des données de chaque variable dans une matrice de cinq variables peut facilement provoquer l'élimination de 59 % des observations de l'analyse. Rapporte qu'il a vu une baisse de dimension de l'échantillon de 624 à 201 avec l'utilisation de la méthode de suppression listwise.

Les techniques statistiques d'analyse des données ayant besoin d'un nombre suffisant d'observations pour que leurs résultats soient valides. Dans des cas qui ne sont pas rares, 16 où la quasi-totalité des exemples possède des valeurs manquantes, elle devient même inutilisable.

D'autre part, les statistiques, telles que la moyenne ou la variance, seront fortement biaisées, à moins que le mécanisme de génération des données ne soit complètement aléatoire (MCAR)

Malgré le fait que la grande perte de données réduit la puissance et l'exactitude statistiques, cette technique, du fait de sa simplicité est fréquemment l'option implicite pour l'analyse dans la plupart des progiciels statistiques. Cette méthode est aussi appelée par certains auteurs, la méthode d'analyse des données disponibles

I.3.3.1.b. Étude des cas disponibles

Dans cette méthode, on ne considère que les cas où ces variables sont complètement observées. Par exemple, si la valeur de l'attribut A est absente pour une observation, les autres valeurs pour le restes des attributs de la même observation pourraient encore être employées pour calculer des corrélations, telles que celle entre les attributs B et C.

Comparée à la première méthode (Étude des cas complets), cette méthode conserve beaucoup plus de données qui auraient été perdues si on employait la méthode d'étude des cas complets.

Il s'agit d'une autre méthode proposée par les logiciels statistiques, mais généralement problématique : le nombre d'observations (n) varie pour le calcul de chaque valeur de la nouvelle base de données, le risque d'obtenir une base de données réduite est grand et, encore une fois, la représentativité sera biaisée si les données manquantes ne sont pas distribuées de façon complètement aléatoire.

Les études de Monte Carlo ont montré que la suppression par la méthode « d'étude des cas disponibles » donne des évaluations moins précises des paramètres, La méthode Pairwisedeletion est uniformément plus précise, bien que les différences puissent parfois être minimes.

Toutefois, il existe certaines raisons, pour la considérer une bonne méthode CD doit être appliquée uniquement dans les cas où le nombre de valeurs manquantes est relativement faible.

De plus, il paraît que, même si les données sont manquantes selon un mécanisme complètement aléatoire, les méthodes qui utilisent l'ensemble de l'information contenue dans la matrice de données sont plus efficaces que les méthodes basées sur les données complètes.

I.3.3.2. Procédures de remplacement

Ces procédures visent à se ramener à une base complète en trouvant un moyen adéquat de remplacer les valeurs manquantes. On nomme ce procédé imputation, complétion ou substitution. De façon générale, des procédures de remplacement peuvent être employées dans certains cas, tant qu'on a une bonne raison pour remplacer. Généralement il est facile d'exécuter les procédures de remplacement, et certaines sont incluses comme options avec les logiciels statistiques. Les avantages les plus importants de ces procédures sont la conservation de la dimension de la base de données, par conséquent, de la puissance statistique d'analyses. Dans une plus ou moins large mesure, toutes les procédures de remplacement sont décentrées s'il y a une distribution non aléatoire des valeurs manquantes. Cependant, le remplacement des données manquantes est approprié quand les corrélations entre les variables sont faibles [6]. Différentes procédures de remplacement de données manquantes ont été élaborées au cours des années. Généralement on constate que les différences entre les diverses méthodes diminuent avec : (a) une plus grande dimension de la base de données, (b) un plus petit pourcentage des valeurs manquantes et (c) une diminution au niveau des corrélations entre les attributs, ont rapporté qu'il y a une différence entre les méthodes de remplacement des données manquantes si les effets des traitements sur les statistiques analytiques sont pris en considération. Avec de plus grandes dimensions 18 de la base de données, en fait, les différences entre les diverses procédures de remplacement augmentent ; ceci fournit davantage d'évidence qu'en évaluant l'efficacité des traitements des données manquantes, l'exactitude d'estimer la valeur des données manquantes et l'exactitude d'estimer les effets statistiques doivent être considérées. Trois types de procédures de remplacement peuvent être distingués : imputation basée sur la moyenne, basée sur la régression et le hot-deck imputation.

- ✓ **Imputation par la moyenne**
- ✓ **Imputation par régression**
- ✓ **Imputation par hot-deck**
- ✓ **Imputation par k-plus proches voisins**
- ✓ **Imputation multiple**

I.3.3.3. Procédures basées sur un modèle

I.3.3.3.a. Maximum de vraisemblance

Sous sa forme plus simple, l'approche de maximum de vraisemblance pour analyser des données manquantes, suppose que les données observées sont tirées d'une distribution normale multivariée [7]. Au lieu d'imputer des valeurs aux données manquantes, ces méthodes définissent un modèle à partir des données disponibles et basent les inférences de ce modèle sur la vraisemblance de la distribution des données sous ce modèle.

I.3.3.3.b.Maximisation d'espérance

Une dernière approche assez fréquente consiste à utiliser l'algorithme de maximisation d'espérance EM (Expectation-Maximisation) pour estimer les valeurs manquantes qui est un processus. Il est généralement utilisé pour estimer les paramètres d'une densité de probabilité. Il peut être appliqué sur des bases de données incomplètes, et présente l'avantage de procéder à l'estimation des valeurs manquantes en parallèle de l'estimation des paramètres. On suppose l'existence d'un modèle de génération des données, par exemple un mélange de gaussiennes pour les variables continues. Les paramètres du modèle sont calculés suivant la méthode du maximum de vraisemblance, de manière itérative ont présenté une nouvelle méthode de substitution basée sur une version simplifiée d'EM. À partir d'une estimation par défaut des valeurs manquantes, les paramètres du modèle sont ré-estimés, à chaque itération, à partir de la matrice complète, de manière à 23 accroître la vraisemblance des données. Le modèle avec ses nouveaux paramètres est alors utilisé pour ré-estimer les valeurs manquantes. Puis on recommence jusqu'à ce que la convergence soit atteinte (ou considérée comme telle). À la fin de l'exécution de l'algorithme, on dispose non seulement des paramètres de notre modèle, mais également d'une matrice de données complétée. Cette technique est très coûteuse en temps de calcul comme beaucoup d'approches itératives. De plus elle demande la spécification d'un modèle de génération des données. Cette tâche implique de faire un certain nombre d'hypothèses, ce qui est toujours délicat. Pour ces raisons, l'application d'EM pour remplacer les données manquantes n'est pas toujours envisageable. Un grand nombre de méthodes de traitement des valeurs manquantes sont explorées dans les diverses études relevées dans la littérature. Il est impensable de tenter d'inclure l'ensemble de ces diverses méthodes dans une seule recherche; les méthodes de traitement les plus souvent comparées ont donc été considérées, le tableau 1.5 présente un résumé des différentes méthodes discutées dans ce mémoire.

I.4. Choix méthodologique

Avant de réfléchir sur la méthode de traitement à appliquer aux données manquantes, un chercheur doit tout mettre en place pour ne pas avoir des valeurs manquantes [10], puisqu'il n'existe pas de méthode totalement efficace pour traiter le problème des données manquantes. La meilleure méthode de traitement c'est de ne pas avoir de valeurs manquantes ! Les méthodes permettent, au mieux, de réduire les biais induits par la présence de ces données manquantes. La théorie statistique prévoit plusieurs méthodes pour l'estimation des valeurs manquantes qui font défaut à l'observation. Ainsi, l'on rencontre plusieurs pratiques dans les études statistiques. Ces pratiques sont plus ou moins basées sur l'intuition et le bon sens, plutôt que sur une théorie proprement dite. Surtout quand il s'agit de différents domaines d'application. Dans ce sous-chapitre, l'accent sera mis sur les méthodes que nous utilisons dans notre étude, nous ne prétendons pas couvrir l'ensemble des méthodes, mais nous évoquons les plus courantes, celles que nous avons incluses dans notre travail. Toutes n'ont pas les mêmes propriétés, aussi est-il important de bien spécifier les objectifs que l'on s'assigne avant de choisir une méthode de substitution afin de pouvoir vérifier l'adéquation entre objectifs et propriétés de chaque méthode. Les principaux objectifs que l'on peut vouloir poursuivre sont les suivants :

1. Précision de la substitution : la valeur de remplacement doit être aussi proche que possible de la vraie valeur.
2. Précision de l'étape d'analyse : dans notre contexte, la phase d'analyse correspond à la construction d'un modèle de classification supervisée. Un des objectifs est alors de maximiser la performance du classificateur.
3. Complexité minimale. La présence de données aberrantes et manquantes est un problème important dans le contexte de notre travail, il faut les gérer avec précaution afin d'éviter de détériorer la performance de LAD.

Certaines méthodes de traitements spécifiques pour les données manquantes selon un mécanisme non aléatoire sont proposées, mais elles demeurent pour le moment expérimentales

Un grand nombre de méthodes de traitement des valeurs manquantes sont explorées dans les diverses études relevées dans la littérature. Il est impensable de tenter d'inclure l'ensemble de ces diverses méthodes dans une seule recherche ; les méthodes de traitement les plus souvent comparées ont donc été considérées.

I.5. Conclusion

Dans ce chapitre, nous avons pu voir dans un premier temps les différents concepts de la fonction maintenance, et l'impact économique de cette fonction sur le rendement global d'une entreprise. Nous avons ensuite exposé différents systèmes de surveillances et de diagnostic utilisés pour optimiser la fonction maintenance. Ces systèmes de surveillances doivent traiter un nombre important de données, et doivent pouvoir exploiter le savoir-faire des automaticiens. Après avoirprésenté les différentes méthodes de substitution des valeurs manquantes, on a présenté la méthodologie de choix de la technique de substitution.

Chapitre II : Maintenance des systèmes industriels

II.1. Introduction

Dans le contexte concurrentiel économique actuel, la réduction des coûts de production est devenue la clé du développement et de la survie de l'entreprise. Hormis, le maintien en condition opérationnelle aux meilleurs coûts d'un outil de production, la réduction des coûts de maintenance et l'augmentation des taux de disponibilité des équipements représentent un levier de productivité, voire de compétitivité. C'est l'enjeu vital pour toute entreprise. Les pannes et les incidents des systèmes de production sont l'un des fléaux majeurs de l'industrie. L'indisponibilité des machines vient amputer lourdement la capacité de production et donc accroître le prix de revient, étant donné que la machine industrielle est un équipement complexe, sujet à des multiples modes de défaillances ou les opportunités de pannes ou d'incidents sont multiples. Ces derniers sont l'un des fléaux majeurs de l'industrie. Un arrêt de machine non-prévu peut avoir de grandes conséquences financières. L'output de la machine suivante n'est plus assuré, voire le produit fini non plus.

Une panne est souvent très difficile à prévoir. Certains mécanismes sont parfois inutilement compliqués, ce qui les rend très fragiles. Il convient de noter aussi que les insuffisances des machines existantes nécessitent rarement leur remplacement par des matériels plus récents et le problème d'accessibilité des éléments des machines peut aussi être une cause dans la longueur de réparation de certaines pannes. De plus l'impact que peut laisser une fonction cachée dont la défaillance n'est pas évidente à l'équipage aux cours des rondes normales des opérateurs. En plus, les techniques d'entretien préventif et d'intervention régulière et systématique bien que coûteuses, ne garantissent pas l'absence de défaillance.

En effet, la gestion optimale d'un système industriel tout au long de sa durée de vie, de la conception au démantèlement, passe par la recherche d'un compromis entre des objectifs souvent conflictuels.

On distingue d'une part les performances économiques, coûts et bénéfices, et d'autre part les aspects de fiabilité, disponibilité, sécurité des personnes et sureté des installations.

Pour apporter des éléments d'aide à la décision face a ce problème, il est nécessaire de disposer d'outils et méthodes permettant d'analyser les systèmes et d'évaluer quantitativement leurs performances en termes de sureté de fonctionnement au sens large, c'est-a-dire de fiabilité et de maintenance, tout en respectant les contraintes économiques.

Les études récentes dans l'efficacité de gestion de la maintenance, indiquent que le un tiers de tous les coûts d'entretien est gaspillé pour un entretien inutile ou incorrectement effectué qui affecte de manière significative la capacité de fabriquer les produits de qualité et qui cause les pertes de temps de production. Quand vous considérez que l'industrie des États-Unis dépense plus de 200 milliards

dollars tous les ans sur l'entretien de l'équipement, l'impact sur la productivité et le bénéfice représenté par l'opération d'entretien deviennent clairs. Cela fait appel à des philosophies plus avancées de gestion de fabrication et de maintenance.

L'établissement du programme d'entretien a été, et est dans beaucoup de cas, toujours affirmé sur des données statistiques de tendance ou sur l'échec réel de l'équipement d'usine.

Jusqu'à récemment, la gestion a ignoré l'impact de l'opération d'entretien sur la qualité du produit, les coûts de production, et plus important, sur le bénéfice de ligne de fond. L'opinion générale a été: « la maintenance est un mal nécessaire ». [11]

La maintenance n'est pas une destination ou bien une fin en soi, mais c'est une des fonctions de l'entreprise. Les décideurs sous-estiment son impact et la considère comme un centre de Coût, pourtant elle devient une composante de plus en plus sensible de la performance de l'entreprise.

Comme toute activité industrielle, la maintenance évolue. Elle doit s'adapter aux contraintes de production et de service, elle bénéficie des développements technologiques et se doit d'être productive.

La variété de matériels a fourni le moyen de réduire ou d'éliminer les réparations inutiles, d'empêcher les pannes catastrophiques des machines, et de réduire l'impact négatif de l'opération de maintenance sur le rendement des entreprises et induit des politiques de maintenance variées et adaptées en fonction des besoins.

II.2. Maintenance de systèmes industriels

II.2.1. Présentation de la maintenance

La maintenance de ses débuts à nos jours : Bien avant la révolution industrielle dans la seconde moitié du dix-huitième (18ème) siècle, la maintenance consistait en des activités des artisans tels que les menuisiers, les forgerons, les soudeurs, les maçons, etc, dans la réparation des bâtiments, des machines primitives, et des moyens de transport de l'époque [12]. Comme il n'y avait pas de concepts tels que « pièces de rechange » ou autre concepts relatifs à la maintenance de jours, les défaillances étaient en général réparées par la fabrication de nouvelles pièces ou la réparation de l'ancienne. Par ailleurs, la réparation était très à l'ordre du jour et les structures de base étaient elles mêmes soit réparables, soit d'une très grande durabilité, ou les deux à la fois.

En l'absence de méthodes pour le calcul de la fatigue et autres concepts liés à la dégradation, la conception et la réparation était fortement intégrées, en ce sens que pour remplacer une pièce défaillante jugée pas assez efficace par le propriétaire, ce dernier exigeait une autre beaucoup plus « robuste ». Toutes ces pratiques étaient liées à l'évolution lente et au faible coût de ces services de l'époque.

Ces pratiques ont progressivement évolué à partir de la révolution industrielle. En 1785, Thomas Jefferson a noté que les pièces des canons étaient fabriquées de façon adéquate et assez rapidement pour qu'on envisage de les changer au besoin. Le concept, lent à l'époque, de pièces de rechange, s'est progressivement accéléré. Cette accélération est fortement accentuée par la complexité croissante des machines à « maintenir ». Le travail de l'agent de maintenance s'est alors progressivement dirigé vers plus d'aptitudes au diagnostic. Un autre facteur important de ce changement était le contrôle de qualité, puis l'automatisation dans les systèmes manufacturiers ; ceci a rendu le remplacement des pièces, et plus tard des unités entières, plus rapides et plus faciles que les réparations dans beaucoup de cas.

Les approches pour la gestion de la maintenance ont ainsi considérablement évolué au cours du vingtième siècle. Jusqu'en 1940, seule la maintenance corrective était pratiquée et elle était considérée comme un coût inévitable. Les agents spécialisés de maintenance (réparateurs) sont là pour la remise en état des équipements seulement une fois que ceux-ci sont tombés en panne.
La maintenance n'était pas envisagée à la conception, et ses impacts sur les performances des systèmes n'étaient pas non plus reconnus.

Les applications de la recherche opérationnelle à l'industrie (notamment l'industrie militaire) pendant la deuxième guerre mondiale ont conduit à une large pratique de la maintenance préventive

Les Allemands, dans leurs efforts de guerre étaient de ceux qui ont fortement pousser à l'exigence de plus en plus de fiabilité, et partant accordent de plus en plus d'importance stratégique à la maintenance. La maintenance périodique était initialement destinée à améliorer la sûreté plutôt qu'à accroître la disponibilité ou à réduire les coûts. Depuis les années 1950, les modèles de recherche opérationnelle sont apparus de façon croissante dans l'étude et l'optimisation des politiques de maintenance.

Mais l'impact réel des actions de maintenance sur la performance des activités n'est abordé que dans les années 1970 où des approches plus intégrées de la maintenance ont évolué aussi bien dans le domaine public que privé. L'acquisition de moyens de défense de plus en plus coûteux conduisit alors les autorités américaines à requérir des approches de « coût globale du cycle de vie » (LCC - life cycle costing), le coût de la maintenance devenant alors un composant très important de ce coût du cycle de vie.

Le lien direct entre la fiabilité et la maintenabilité est alors reconnu et le terme "R&M" (Reliability-Maintenability) devient très largement utilisé en parlant des systèmes de défense. Ce concept est aussi adopté dans les industries et services civiles. La maintenance basée sur la fiabilité (RCM - ReliabilityCentered Maintenance) apparaît alors dans l'aviation civile aux États-Unis. Au

même moment, les japonais développent la maintenance productive totale (TPM - Total Productive Maintenance).

Ces deux pratiques, sur lesquelles nous reviendrons plus tard, considèrent la maintenance dans le contexte plus large des activités économiques et prennent en compte les liens étroits entre les défaillances des composants et leurs impacts sur la performance des activités.

Ces deux pratiques opèrent sur des horizons à court et moyen terme. Les considérations sur ces horizons à court et moyen terme nécessitent des modèles de prédiction efficaces pour accéder aux conditions de divers éléments des systèmes et à leur comportement sous différentes stratégies de maintenance. Comme des insuffisances dans ces pratiques sont devenues apparentes dans les années 1990. Il apparut de plus en plus que la maintenance ne devrait pas être vue que dans un contexte opérationnel en considérant les défaillances des équipements et leurs conséquences. Elle devait tenir compte aussi bien des retombées économiques que des normes nationales et internationales de plus en plus exigeantes.

Ainsi, l'évolution de la maintenance et des outils pour son management ont donné à la maintenance diverses pratiques au sein de l'entreprise.

II.2.1.1. Définition et objectifs de la maintenance

D'après la norme française NF EN 13306 X 60-319 [13], la définition de la maintenance est "l'ensemble de toutes les actions techniques, administratives et de management durant le cycle de vie d'un bien, destinées à le maintenir ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise".

Le management de la maintenance concerne toutes les activités des instances de direction qui, d'une part, déterminent les objectifs, la stratégie et les responsabilités concernant la maintenance et, d'autre part, les mettent en application par des moyens tels que la planification, la maîtrise et le contrôle de la maintenance, l'amélioration des méthodes dans l'entreprise (y compris dans les aspects économiques).

L'activité actuelle de la maintenance s'inscrit dans un cadre de participation à la réalisation des objectifs de productivité, de rentabilité et de croissance de l'entreprise. Il est important de s'assurer donc du bon fonctionnement de la fonction maintenance. Toutes les activités d'un cycle d'exploitation doivent être réalisées au mieux de façon à maximiser les effets positifs de la maintenance, tout en minimisant les coûts. La réalisation des tâches de maintenance requiert à la fois des compétences "métiers" (mécanique, hydraulique, électricité, . . .) et des compétences organisationnelles (planification des tâches, gestion des ressources, soutien logistique intégré, . . .). Le choix et la mise en œuvre d'une stratégie de maintenance dépendent donc de nombreux paramètres. Il est nécessaire

de pouvoir mesurer a priori les conséquences de cette stratégie sur les performances globales du système qui peuvent n'apparaître qu'à long terme. En pratique, cependant, il est difficile d'évaluer et de comparer les gains et coûts de décisions de maintenance. Ainsi définir une approche rationnelle pour la maintenance représente un réel bénéfice. La modélisation des processus de maintenance peut constituer un support intéressant pour l'évaluation et la rationalisation des pratiques de maintenance. Dans la mesure où les activités de maintenance font intervenir une grande diversité de processus (depuis la conception du système jusqu'à la gestion des pièces détachées), une approche de modélisation doit nécessairement être limitée à un sous-ensemble de ces processus.

II.2.1.2. Les type de maintenance

Les stratégies de maintenance peuvent être répertoriées en deux grandes catégories : la maintenance corrective et la maintenance préventive.

La maintenance corrective est la maintenance qui intervient suite à la défaillance du système alors que la maintenance préventive est réalisée lorsque le système est encore en fonctionnement.

Le recours à l'une ou à l'autre de ces stratégies diffère suivant l'élément considéré mais aussi le type de structure, la politique d'exploitation et de suivi, les coûts, la disponibilité de l'information, etc.

Dans la figure 1, nous présentons les différentes stratégies suivant le type de maintenance étudiée. Alors que la mise en place d'opérations correctives ne dépend que de l'occurrence d'une panne, les maintenances préventives peuvent être programmées en fonction de différents paramètres.

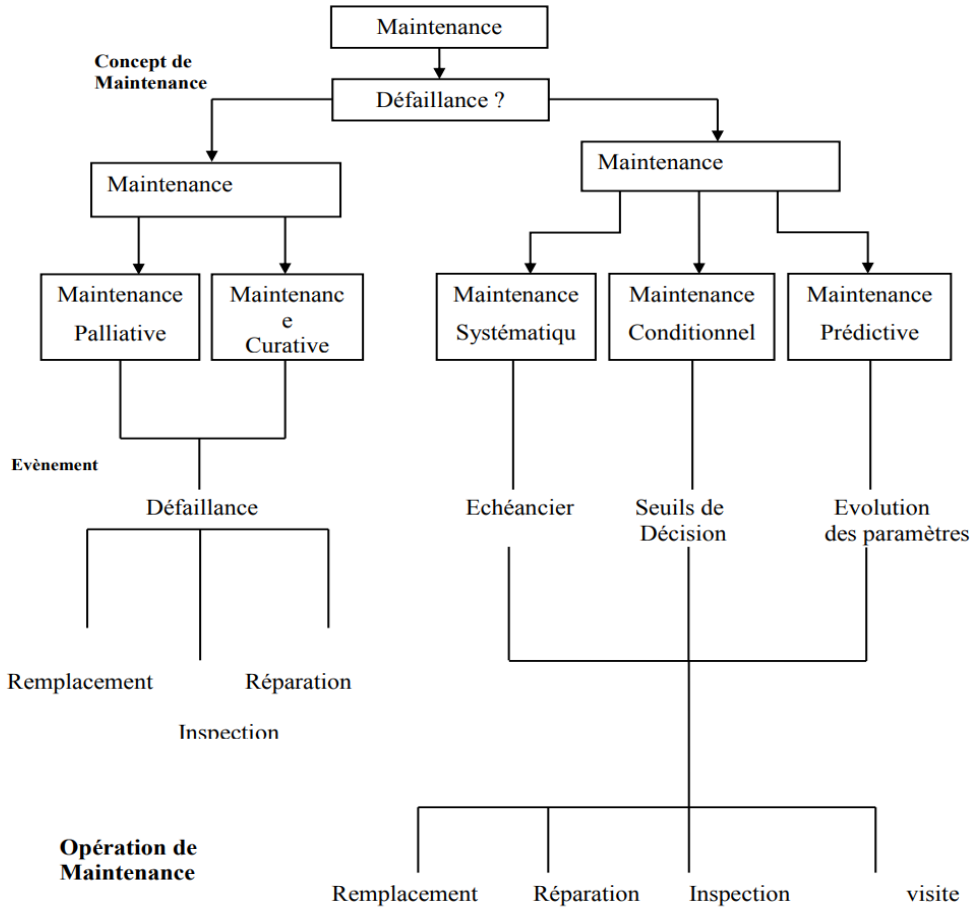


Figure II.1 : Les différentes stratégies de maintenance

II.2.1.2.a.La maintenance corrective

La maintenance corrective est l'ensemble des activités réalisées après la panne du système pouvant être liée à sa défaillance ou à la dégradation de sa fonction, elle a alors pour but de le remettre en état de marche. La maintenance corrective peut être :

– palliative : des réparations ou des remises en état à caractère provisoire sont effectuées.

– curative : des réparations, des modifications ou des remises en état à caractère permanent sont effectuées.

Cette maintenance est utilisée lorsque l'indisponibilité du système n'a pas de conséquences majeures ou quand les contraintes de sécurité sont faibles.

II.2.1.2.b.La maintenance préventive

La maintenance préventive a pour objet de réduire la probabilité de défaillance. Par ailleurs, une maintenance préventive est intéressante, en considérant le coût comme critère d'évaluation, si les coûts induits par la perte de performance sont élevés et si les coûts de réparation et d'inspection sont

relativement faibles par rapport aux premiers. La maintenance préventive peut être systématique, conditionnelle ou prévisionnelle.

Une synthèse des politiques de maintenance préventive est donnée par [14]. Ilexiste trois groupes de maintenance préventive:

- ✓ **Maintenance systématique** : lorsque la maintenance préventive est réalisée à des intervalles prédéterminés, on parle de maintenance systématique. L'opération de maintenance est effectuée conformément à un échéancier, un calendrier déterminé a priori. Aucune intervention n'a lieu avant l'échéance prédéterminée. L'optimisation d'une maintenance préventive systématique consiste à déterminer au mieux la périodicité des opérations de maintenance sur la base du temps, du nombre de cycles de fonctionnement, du nombre de pièces produites, etc.
- ✓ **Maintenance conditionnelle** : lorsque l'opération de maintenance préventive est subordonnée à l'analyse de l'évolution surveillée de paramètres significatifs de la dégradation ou de la baisse de performance d'une entité, on parle de maintenance conditionnelle. Les paramètres significatifs de la dégradation peuvent être soit des mesures de caractéristiques physiques du système (épaisseur d'un matériau, degré d'érosion, température, pression, ...), soit des informations sur la durée de vie résiduelle (on parle alors de maintenance prédictive). La planification des interventions repose sur l'existence et la détermination de seuils critiques pour ces paramètres de dégradation. On parlealors de seuils de décision.
- ✓ **Maintenance prévisionnelle** : lorsque la maintenance préventive est effectuée sur la base de l'estimation du temps de fonctionnement correct qui subsiste avant l'observation de l'événement redouté, on parle de maintenance prévisionnelle. Une maintenance prévisionnelle peut prendre en compte un âge du matériel qui n'est pas forcément calendaire mais par exemple le temps de fonctionnement mesuré depuis la dernière inspection.

Dans le milieu industriel, en général, une maintenance mixte est appliquée aux systèmes. En effet, la maintenance préventive est destinée à réduire la probabilité de défaillance mais il subsiste une part de maintenance corrective incompressible. Il est donc nécessaire de considérer des stratégies qui combinent les deux : maintenance corrective et maintenance préventive.

II.2.1.3. Les approches industrielles

La gestion de la maintenance est traditionnellement un domaine réservé à l'ingénieur et le processus de décision est fortement influencé par la formation technique de ces ingénieurs ainsi que par leur expérience industrielle. Dans la majeure partie des cas, la décision de maintenance repose essentiellement sur une contrainte budgétaire. En effet, la décision de maintenance prend en

considération un grand nombre de critères et elle est la réponse amenant au meilleur compromis (coût/disponibilité par exemple). Le coût de mise en place d'une politique de maintenance est alors important. Le choix d'une opération de maintenance devient très rapidement complexe (ne dépendant pas uniquement de l'exploitation technique du système) et les conséquences issues de ce choix sur l'exploitation future du système ne sont pas directement observables ou connues.

En pratique, les évaluations prévisionnelles et les démonstrations d'objectifs de maintenance et de fiabilité s'appuient sur des données quantitatives. Ces données proviennent de banques élaborées par des utilisateurs, des fabricants, des organismes qui ont en commun des connaissances de bases utilisables pour tous. En général, dans ces banques de données, plusieurs indications sont disponibles comme, par exemple, la dénomination du matériel ou du composant, la moyenne des temps de bon fonctionnement, le taux de défaillance moyen, etc. Il faut souligner que les éléments fournis par les banques de données électroniques ont une durée de vie courte (3 à 6 ans). Compte tenu de la rapidité de l'évolution technologique dans ce domaine, cela nécessite de les réactualiser régulièrement. Les données mécaniques ont une durée de vie plus longue (30 à 40 ans). Notons également l'hypothèse faite en mécanique d'un taux de défaillance constant correspond à une simplification qui n'est pas toujours justifiée. De plus, les patrimoines statistiques sont souvent plus faibles et moins précis qu'en électronique. Nous pouvons citer, par exemple, les banques de données suivantes :

– FIDES est une base de données française qui propose une méthodologie de calculs des taux de défaillance de cartes électroniques, à partir de contributeur d'environnement, de conception et de procédé résultant d'un groupe de travail.

– HDBK 217F est une banque de données militaire d'origine américaine, relativement élaborée, utilisée aussi dans le domaine civil. Elle concerne pour sa partie la plus importante des équipements électroniques et électrotechniques (moteurs électriques, actionneurs, capteurs incrémentaux). Elle propose une évaluation du taux de défaillance à partir de paramètres de fonctionnement et d'utilisation.

– La banque de données européenne EIReDA concerne les équipements mécaniques et électriques. Elle est destinée aux professionnels du secteur nucléaire.[15]

Par ailleurs, les industriels disposent de progiciels plus ou moins performants de type ERP (Entreprise Ressource Planning) ou GMAO (Gestion de Maintenance Assistée par Ordinateur) pour les accompagner dans l'organisation de leur plan de maintenance. L'avantage de ces progiciels est de pouvoir intégrer les informations propres à l'entreprise. Il faut toutefois noter que les résultats délivrés sont uniquement des informations statistiques. Ils ne donnent pas directement les actions à réaliser. Il est alors nécessaire à l'entreprise de disposer d'opérateurs qui puissent interpréter les résultats afin de prendre des décisions pertinentes.

Une approche qui semble prometteuse mais encore trop peu développée dans le secteur industriel est la construction de modèles de maintenance permettant d'estimer l'impact d'une décision de maintenance sur l'exploitation future du système, Il est alors possible de "simuler" des stratégies de maintenance et de noter les effets et les coûts de la maintenance afin de déterminer au mieux la politique à mettre en oeuvre sur le système. La modélisation de la maintenance offre un bon compromis pour l'évaluation, la validation ou la comparaison des performances de politiques de maintenance différentes qui ne sont pas toujours aisées à mettre en oeuvre en situation réelle : nécessité de données de retour d'expérience et risques industriels ou financiers liés à des politiques "expérimentales" inadéquates.

Pour rester pertinent dans l'analyse des gains apportés ou la "pré-validation" d'une politique, le modèle de maintenance doit intégrer un maximum d'information sur le système ainsi que l'expérience technique accumulée au cours de son exploitation. Toutefois, on souligne qu'une modélisation mathématique peut caractériser une solution idéale à un problème de maintenance (pour des conditions idéales) et les résultats et les analyses issues de cette modélisation ne doivent être qu'un support, une aide à la décision.

De plus, il faut noter que dans un grand nombre de domaines (aéronautique, nucléaire par exemple), il est nécessaire de prouver en amont la performance de la politique de maintenance d'où la nécessité de disposer d'un modèle de maintenance permettant d'estimer l'impact d'une décision de maintenance sur l'exploitation future du système.

II.2.2. Modélisation et évaluation de la maintenance

II.2.2.1. Modèles de maintenance

Lorsqu'on doit faire un choix vis-à-vis d'une politique de maintenance, l'efficacité des décisions qui sont prises va dépendre de la qualité du modèle qui représente le système. La modélisation du système permet en effet d'avoir accès au comportement "a priori" du système en termes de probabilité d'occurrence : probabilité de tomber en panne à chaque instant, probabilité d'être dans un état de dégradation donné en fonction du temps, etc. Cette représentation est indispensable pour pouvoir évaluer les coûts de maintenance sur une période donnée et prévoir les interventions. Mais elle peut être complétée par des informations de données en ligne ou ponctuellement sur l'état réel du système. C'est le cas par exemple lorsqu'un dispositif de surveillance signale la panne totale du système dès son occurrence, lorsqu'il indique en continu le niveau de dégradation du système ou lorsqu'il rend compte ponctuellement de l'état "courant" d'un système après inspection. On oppose donc à l'état "courant" donné par la surveillance, l'état "probable" donné par le modèle. Les

informations sur l'état courant permettent d'élargir considérablement l'ensemble des actions de maintenance possibles : on peut envisager d'intervenir à un moment précis. Les actions de maintenance qui s'appuient à la fois sur la modélisation du système et les informations données par le dispositif de surveillance seront dites conditionnelles. Le terme "conditionnel" indique ici que l'on agit en fonction de la condition, de l'état du système (panne, dépassement d'un seuil de dégradation). Les actions de maintenance qui se fondent uniquement sur le modèle de durée de vie du système seront dites systématiques : elles consistent à prévoir des échéances fixes sur le système indépendamment de l'état d'usure, de marche ou de panne.

Les différentes étapes pour l'élaboration d'un modèle sont les suivantes. Tout d'abord, l'espace d'états, qui correspond à l'ensemble des mesures qui permettent de rendre compte de l'état de fonctionnement du système à chaque instant, doit être défini. L'ensemble de tous les états possibles peut être un ensemble de valeurs (discrètes ou continues). Les états menant à la défaillance du système sont généralement modélisés de l'une des manières suivantes :[16]

– soit comme une "boîte blanche" qui correspond à un modèle conceptuel dont l'idée est de décrire à un niveau de réduction donné le comportement du système. L'ensemble des relations correspondent alors à des phénomènes connus et identifiés, dont la signification est pertinente dans la problématique posée. Toutefois, une relation entre variables d'état inclut souvent des phénomènes qui se produisent à des niveaux inférieurs au niveau de réduction retenu.

– soit comme une "boîte noire" qui est à l'opposé du modèle conceptuel ("boîte blanche"). La "boîte noire" est un modèle empirique où on ne cherche pas à comprendre pourquoi le système fonctionne comme on l'observe mais à décrire les relations entre variables de sortie et variables d'entrée, généralement à partir d'une analyse statistique des entrées et des sorties. Ce modèle correspond à une approche statistique (modèle de durée de vie) ;

– soit comme une "boîte grise" qui est un bon intermédiaire entre la "boîte blanche" et la "boîte noire". La "boîte grise" correspond à une approche stochastique : la dégradation du système peut être modélisée par un processus stochastique basé sur des quantités mesurables corrélées au niveau de dégradation. Lorsque la transparence du modèle dans le cas de la "boîte blanche" devient moins nette à cause des relations masquées par des processus non décrits, on parle alors de "boîte grise". Dans la "boîte noire" qui est un modèle déterministe, on cherche à associer une sortie à une entrée, et l'on affecte à cette sortie une probabilité d'occurrence. Alors que dans la "boîte grise" qui est un modèle stochastique, on associe, pour une entrée donnée une distribution de probabilité des sorties. Ce type de modèle est très utile en mode prédictif où l'on peut évaluer l'évolution du système en fonction d'un comportement aléatoire des variables

Ensuite, il faut déterminer les différentes actions de maintenance disponibles - qui peuvent être de deux types : de type surveillance (inspections) ou curative (remplacements, réparations) - et la qualité de ces actions :

– inspections parfaites ou non ;
– remplacements et réparations permettant de remettre le système dans un état aussi bon que neuf (AGAN : As Good As New) ou non et, dans ce cas, il faut arriver à déterminer l'efficacité des actions de maintenance.

L'étape suivante consiste à définir la classe des règles de décision à étudier (markovienne ou dépendante de l'histoire, déterministe ou stochastique). On s'intéressera au cas où la politique de maintenance peut être décrite au travers d'une structure, il faudra alors chercher les paramètres de décision (des seuils, la période d'inspection, par exemple) correspondants. Dans ce cas, l'optimisation de la règle de décision se ramène à un problème d'optimisation de paramètres.

II.2.2.2. Optimisation de la maintenance

L'optimisation de la maintenance consiste à trouver la balance optimale entre maintenance préventive et corrective tout en respectant les objectifs fixés. Il faut alors déterminer les instants de maintenance et les actions à effectuer de manière à optimiser un critère de décision fixé. Ce critère de décision peut aussi bien reposer sur le coût, par exemple un coût moyen de maintenance à long terme ou le coût d'opération par unité de temps, que sur la disponibilité du système ou encore la sécurité [17].

La figure 2 est une illustration des effets de la fréquence des opérations de maintenance sur les coûts liés soit à la maintenance corrective, soit à la maintenance préventive. L'augmentation du nombre d'interventions sur le système permet de réduire les effets indésirables engendrés par une panne mais pénalise le fonctionnement du système et peut entraîner une augmentation du coût global d'exploitation du système dans le mesure où chaque opération de maintenance engendre un coût.

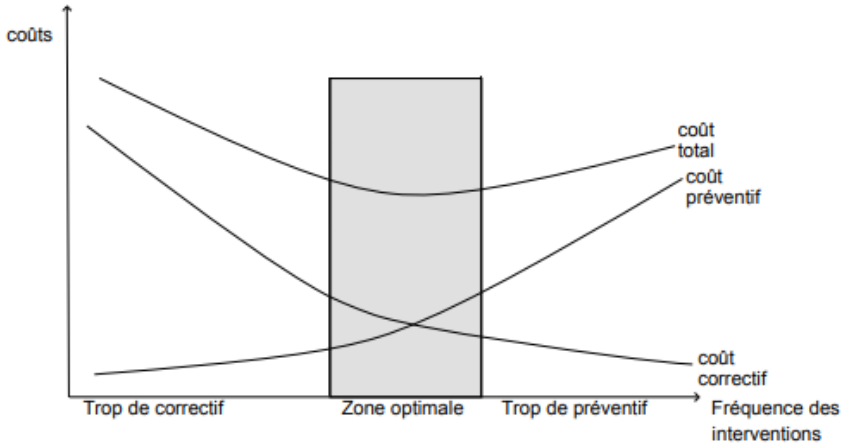


Figure II.2 : La nécessité de contrôler l'équilibre entre maintenance préventive et maintenance corrective

De plus, d'une manière générale, les performances d'un système sont fortement liées à la durée d'exploitation qui peut être considérée sur un horizon de temps fini ou infini (dans ce cas on étudie le critère moyen à long terme). L'inconvénient majeur de l'hypothèse d'horizon fini est le manque de robustesse de cette approche vis-à-vis de la longueur de l'intervalle de temps considéré. Par ailleurs, il est à noter que la modélisation asymptotique s'éloigne du problème industriel du fait de l'hypothèse d'une exploitation sur un intervalle de temps infini. De plus, cette approche ne permet pas de tenir compte de phénomènes ponctuels, de possibilité de changement de stratégie au cours de l'exploitation du système qui peut être dû, par exemple, à un changement d'une propriété du système, etc. Cependant la modélisation asymptotique permet d'élaborer des règles de décision stables qui peuvent mettre en évidence des structures de maintenance faciles à implanter. Les paramètres de décision associés sont déterminés pour optimiser un critère moyen de performance à long terme.

II.2.3. Conclusion et problèmes posés

Dans ce chapitre nous avons souligné la nécessité de bien organiser les tâches de maintenance dans le temps pour répondre au mieux à des exigences d'utilisation d'un système prédéfinies. Nous nous sommes affranchis de certaines hypothèses telles que le management des compétences, la disponibilité des ressources liées à l'activité de maintenance, les aspects techniques, la mutualisation des ressources, etc., et nous nous sommes focalisés sur un unique système pour lequel nous avons cherché à déterminer les dates et les actions de maintenance à effectuer afin de minimiser un critère choisi. Nous avons axé notre discussion aux vues de la principale difficulté liée à cette organisation qui est la caractérisation du comportement de maintenance entre l'augmentation de la disponibilité

d'un matériel et la réduction des coûts de maintenance. En effet, cette caractérisation requiert la connaissance et donc l'évaluation de l'impact d'une décision de maintenance sur l'exploitation du système à long terme.

Pour obtenir des réponses satisfaisantes aux questions initialement posées sur les instants de réalisations d'inspections et des actions de maintenance, il est nécessaire d'élaborer des modèles qui prennent en considération les caractéristiques des systèmes industriels à examiner. Cependant il faut bien distinguer les modèles qui prennent en compte la dégradation, qui ont donc une "signification" physique du mécanisme de défaillance, des modèles purement statistiques qui, d'une part, nécessitent un très grand nombre de données et, d'autre part, ne sont pas forcément compris par l'opérateur de maintenance qui doit interpréter les résultats pour pouvoir prendre une décision.

II.3. Maintenance basé sur les méthodes de fiabilité

II.3.1. Introduction

Dans le monde industriel actuel, les équipements sujets à des pannes et/ou détériorations sont nombreux. Souvent critiques, ces équipements doivent être maintenus afin de continuer à remplir les missions pour lesquelles ils ont été conçus. La maintenance joue ainsi un rôle primordial permettant de garantir la disponibilité pour la production. Toutefois, cette fonction a souvent été négligée car trop fréquemment perçue comme une source de dépenses. Mais cet état d'esprit tend à changer avec l'évolution des équipements et des techniques de production. Qui plus est, les machines sont de plus en plus complexes et les industriels cherchent à les exploiter à leur plein régime dans un souci de compétitivité et de respect des délais tout en cherchant à garantir les exigences de qualité et de sécurité requises. Ces objectifs peuvent difficilement être atteints sans une maintenance adéquate. Une politique de maintenance se définit par un ensemble d'actions exécutées et ordonnées selon une certaine stratégie.

II.3.2. Méthodologie

Les différentes étapes de la méthodologie envisagée sont reprises à la figure 3.

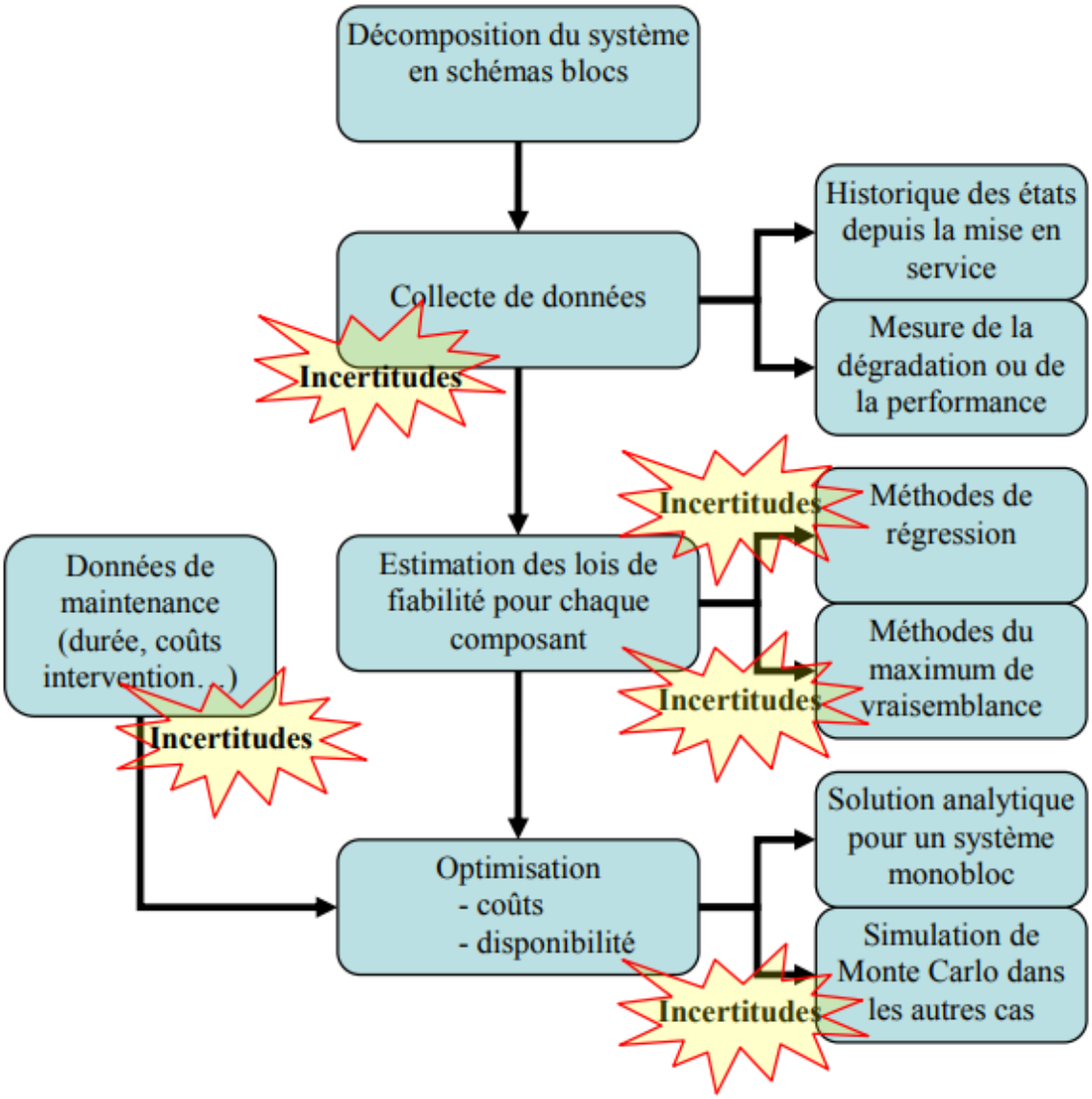


Figure II.3 : Schéma synoptique de la méthodologie

II.3.2.1. Représentation du système étudié en schémas blocs

De manière générale, un système est modélisé comme un ensemble d'éléments dont l'état de bon fonctionnement conditionne celui du système. Pour ce faire, on a recours à une analyse des modes de défaillance qui va conduire à un modèle topologique du système où chaque élément est représenté par un bloc. Ceux-ci sont associés en série lorsque la défaillance d'un seul élément provoque la défaillance de l'ensemble. Par contre, pour qu'une association en parallèle soit défaillante, il faut que tous ses éléments le soient également. D'autres types de configurations sont envisageables, notamment pour modéliser une redondance ou un élément de secours.

II.3.2.2. Collecte de données

L'étape suivante consiste à collecter des données pour chaque machine du système. Ces données peuvent soit provenir d'un historique d'événements ou de mesures de dégradation. Dans le premier cas, les états de la machine sont enregistrés depuis sa première mise en service (mise en marche, première panne, en maintenance, remise en marche, arrêt pour inspection, remise en marche, deuxième panne…). Il peut parfois arriver que l'historique soit incomplet (données manquantes entre deux événements), Décomposition du système en schémas blocs Collecte de données Estimation des lois de fiabilité pour chaque composant Optimisation - coûts - disponibilité Historique des états depuis la mise en service Mesure de la dégradation ou de la performance Méthodes de régression Méthodes du maximum de vraisemblance Simulation de Monte Carlo dans les autres cas Solution analytique pour un système monobloc Données de maintenance (durée, coûts intervention…) Incertitudes

Figure 3. Schéma synoptique de la méthodologie Christophe Letot, Guillaume Fleurquin, Pierre Dehombreux les méthodes mises en œuvre doivent donc être capables de prendre en compte les données censurées. Dans le deuxième cas, les données de dégradation mesurées peuvent soit provenir d'inspections ponctuelles réalisées manuellement ou être enregistrées en continu à l'aide d'un dispositif de monitoring. On distingue également les cas où la dégradation est directement mesurable des cas où elle est mesurée de façon indirecte par l'intermédiaire de variables explicites de la dégradation (co-variables). Par exemple, l'évolution de la fissuration d'un arbre de machine peut être mesurée directement (taille de fissure) si ce dernier est accessible, ou indirectement à l'aide de mesures de variables explicatives. A ce stade, on distingue les variables actives (les causes) des variables passives (les conséquences) qui sont corrélées à la dégradation. Dans le cas de l'arbre fissuré, la variable active est le niveau de charge et la variable passive la vibration mesurée. Par ailleurs, la notion de dégradation peut s'étendre aux notions de rentabilité, productivité, qualité. Le principe est d'avoir un seuil de performance minimal en dessous duquel on considère que le système est défaillant.

II.3.2.3. Estimation de la loi de fiabilité de chaque composant

La caractérisation de la fiabilité d'un équipement peut être obtenue selon trois approches qui se distinguent par le type de données dont on dispose (figure 4). Ces données peuvent correspondre à des temps de défaillance issus d'un historique de panne (modèles statistique sur base d'un historique d'événements), à des mesures de la dégradation physique prélevées au cours du temps dont on connait la loi de dégradation (modèles physiques de défaillance), ou bien à l'évolution de variables explicites de la dégradation (suivi d'un indicateur de dégradation). (Les méthodes d'estimation des lois de fiabilité à partir de ces données sont soit des méthodes de régression soit des méthodes de maximum

de vraisemblance. Dans le premier cas, le principe consiste à linéariser la loi de fiabilité paramétrique que l'on souhaite ajuster et d'identifier les paramètres par simple régression linéaire. Dans le second cas, l'idée est d'estimer les paramètres de la loi de fiabilité qui maximisent la vraisemblance de retrouver les temps de défaillances observés. Ces deux méthodes seront expliquées et illustrées en détails dans le chapitre 3. L'approche basée sur la fiabilité permet également de la calculer la durée de vie résiduelle moyenne de l'équipement, c'est-à-dire l'espérance du temps de bon fonctionnement à un instant donné. Cette information est une donné essentielle pour choisir une politique de maintenance appropriée.

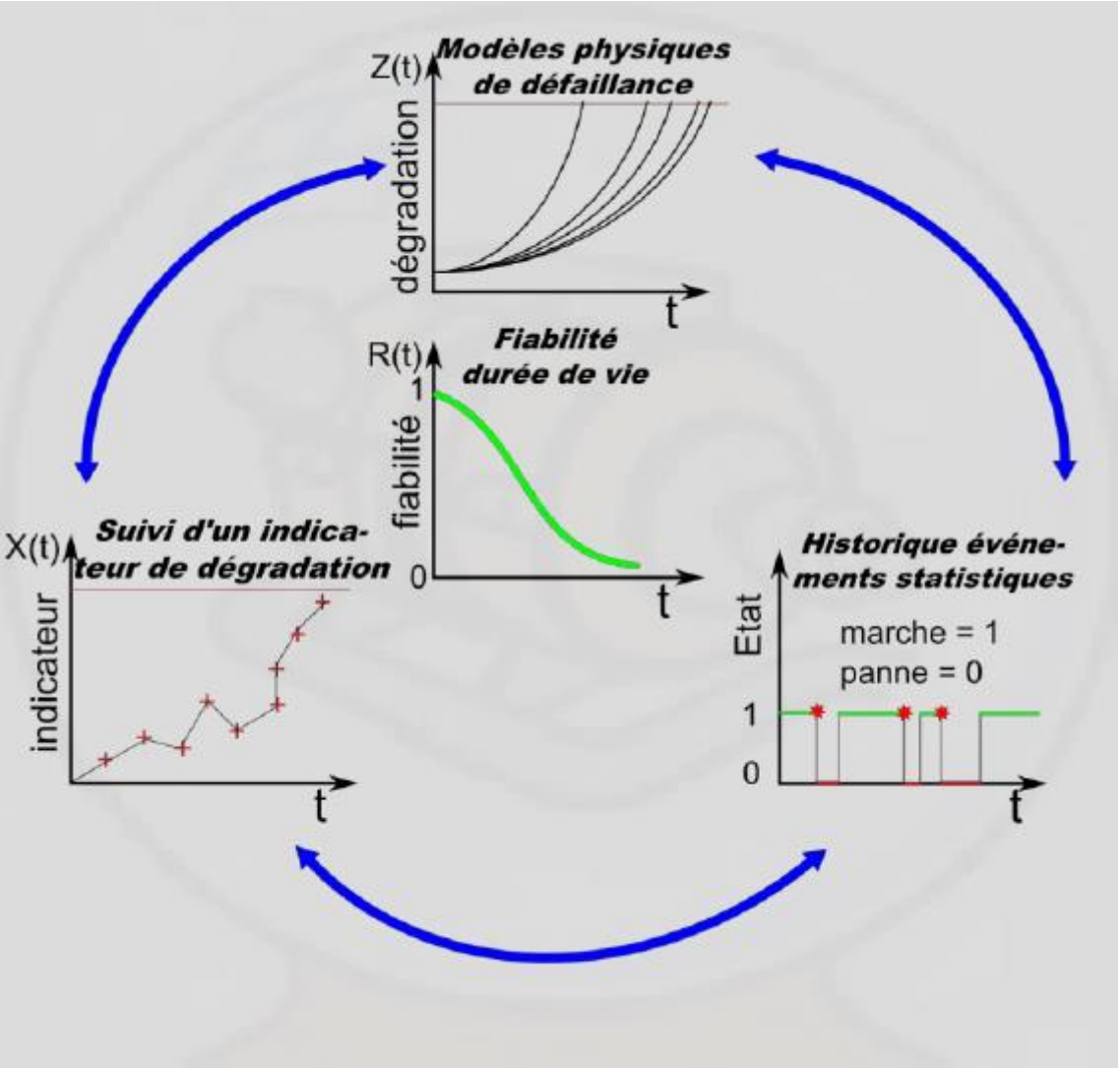


Figure II.4: Les trois approches pour l'estimation de la fiabilité et de la durée de vie

II.3.2.4. Optimisation de la maintenance

Lorsque la fiabilité de chaque composant a été identifiée, la dernière étape consiste à déterminer la périodicité de maintenance préventive optimale. Pour cela, des données de maintenance sont nécessaires. Parmi ces données on recense : - la possibilité de faire du correctif ou uniquement du préventif ; - les coûts fixes du préventif et du correctif ; - les coûts variables du préventif et du correctif, c'est-à-dire les coûts indirects liés à la non production ; - la durée des interventions de maintenance préventive et corrective ; - la fréquence d'inspection de chaque composant ; - le degré de réparation après une intervention de maintenance. Si le système étudié est composé d'un seul élément bloc, une solution analytique du problème d'optimisation existe. Dans le cas contraire, il est généralement nécessaire d'avoir recours à des méthodes de simulation de Monte Carlo. Christophe Letot, Guillaume Fleurquin, Pierre Dehombreux.

II.3.2.5. Incertitudes sur les résultats

L'ensemble des résultats obtenus peuvent être quantifiés en termes d'incertitudes sur la valeur espérée. Cet aspect est important pour le décideur. En effet, est-il préférable de choisir une politique de maintenance préventive dont l'espérance du coût estimé est faible mais avec un écart type élevé ? Ou bien au contraire n'est-il pas préférable de choisir une autre politique de maintenance dont l'espérance du coût est certes légèrement plus élevée mais présentant un écart type plus faible afin de maîtriser les incertitudes ? De la même façon, prédire la fiabilité d'un équipement à un instant donné à d'autant plus de crédibilité si on peut lui associer des intervalles de confiance. Cela permet au décideur de choisir la stratégie la plus adaptée au risque qu'il accepte de prendre. Pour obtenir ces incertitudes, des méthodes de simulation de Monte Carlo sont principalement utilisées. Connaissant la loi de fiabilité, des temps de défaillance sont générés aléatoirement. A partir de ces temps, le modèle de fiabilité est réajusté à chaque simulation. Nous obtenons ainsi des valeurs différentes des paramètres de la loi sur base desquelles un ajustement statistique est réalisé.

II.4. Conclusion

L'objectif de cet article est de proposer une méthodologie générale permettant aux personnes chargées de la maintenance de déterminer la périodicité de maintenance préventive optimale. A cette fin, il est nécessaire de disposer de données. Ces données peuvent soit provenir d'un historique de bon fonctionnement de la machine, du suivi d'un indicateur de dégradation ou bien d'une mesure directe de la dégradation dont l'évolution est connue. A partir de ces données, des modèles de fiabilité sont ajustés soit par des méthodes de régression ou par des méthodes de maximum de vraisemblance. Ces

lois de fiabilité sont bornées par des intervalles de confiance ce qui permet de quantifier une certaine dispersion sur la valeur annoncée de la fiabilité. La fiabilité du système est alors obtenue à l'aide d'une combinaison en série ou parallèle de la fiabilité de ses composants élémentaires. Une estimation de la durée de vie résiduelle du système est alors possible, celle-ci peut être une fois encore bornée par des intervalles de confiance. Finalement, la phase d'optimisation proprement dite est réalisée par simulation en prenant en compte les données issues de la maintenance. Il est évident que les résultats obtenus vont fortement dépendre du ratio entre les coûts/durées d'indisponibilité de la maintenance préventive et les coûts/durées d'indisponibilité de la maintenance corrective. L'exemple utilisé a montré l'importance de calculer les intervalles de confiance sur les indicateurs afin de permettre au gestionnaire de la maintenance de choisir une périodicité de maintenance préventive la plus adaptée à sa stratégie. La méthode proposée peut être utilisée sur n'importe quel système ou équipement du moment que l'on dispose de données mesurables sur celui-ci.

Chapitre III : Méthode de maintenance basée sur données incomplète

III.1. Introduction

La fiabilité est définie comme la probabilité qu'un système accomplisse sa fonction pendant une période donnée et dans des conditions de fonctionnement données. En pratique, les systèmes peuvent être complexes et les conditions de fonctionnement variables. Afin d'estimer la fiabilité des systèmes complexes, nous avons développé Reliabilitix, une librairie MATLAB. La méthodologie suivie consiste en une modélisation fonctionnelle du système sous forme d'un schéma-bloc, chaque bloc représente un composant caractérisé par une fonction de fiabilité. Parmi les structures caractérisant les systèmes complexes, citons les associations série et parallèle, les structures m parmi n et les redondances passives. Les principales lois statistiques utilisées pour modéliser la fiabilité sont la loi exponentielle, la loi de Weibull et la loi lognormale. Afin d'estimer les paramètres des lois, la méthode de régression et la méthode du maximum de vraisemblance sont le plus couramment utilisées ; Meeker et Escobar ont détaillé ces méthodes dans [18]. Ces deux méthodes ont été programmées dans la librairie MATLAB Simulatrix. Par ailleurs, Simulatrix a été doté d'un simulateur utilisant la méthode de Monte Carlo afin de réaliser des études statistiques des paramètres de fiabilité. La littérature dégage deux méthodes pour caractériser les mécanismes de dégradation. La première s'intéresse à caractériser les phénomènes qui sont à l'origine des modes de défaillance comme les propagations de fissures. D'une manière générale, ces phénomènes physiques vont provoquer un dommage au système étudié. La deuxième méthode consiste à élaborer, sur base de tests en laboratoire, des modèles statistiques en fonction des modes de dégradation et des sollicitations ; le modèle de Wöhler et les modèles utilisés dans les tests accélérés appartiennent à cette famille. On a deux modèles de fiabilité qui prennent en compte les conditions opératoires. Le premier modèle consiste à exprimer les paramètres des distributions en fonction des variables opératoires. Le deuxième modèle étudié est le modèle à risques proportionnels ou PHM (Proportionalhazards model). Dans le domaine mécanique, appliquent le PHM à des équipements soumis à vibrations. Plus particulièrement, cette étude s'intéresse à la fatigue. Pour une contrainte donnée, la fonction de fiabilité d'un système soumis à la fatigue peut être établie en considérant la distribution de la limite d'endurance σD, qui peut être déduite d'essais de fatigue, et le modèle de Wöhler.

III.2. Modélisation de la fiabilité

Les méthodes les plus utilisées pour estimer la durée de vie des équipements sont généralement la revue de la littérature, la méthode Delphi et l'analyse de l'historique d'un échantillon

La première étape est de collecter des informations concernant les équipements considérés obsolètes. Une fois ces données historiques prêtes, on peut procéder à l'évaluation de la fonction de durée de vie de cet équipement. Les quatre types de fonctions de distribution statistique les plus utilisées sont : 1) distribution de Weibull, 2) distribution normale, 3) distribution log-normale et 4) distribution exponentielle.

III.2.1. Modèles de fiabilité usuels

La fiabilité, notée R, est définie comme la probabilité qu'un système accomplisse sa fonction pendant un temps donné et dans des conditions de fonctionnement données. En termes mathématiques, la loi de fiabilité est définie par :

$$R(t)= Pr(T >t) \tag{3.1}$$

Avec T, la variable aléatoire représentant les durées de vie.

La loi de défaillance indique la probabilité qu'une défaillance survienne avant un temps donné, il s'agit de la fonction de répartition des temps de défaillance :

$$F(t)= Pr(T \leq t)= 1- R(t) \tag{3.2}$$

La densité de probabilité est définie par la relation :

$$f(t) = \frac{dF(t)}{dt} \tag{3.3}$$

Le taux de défaillance indique la probabilité qu'un système soit défaillant entre t et t+dt sachant qu'il fonctionnait à t :

$$\lambda(t)dt = \frac{f(t)dt}{R(t)} \tag{3.4}$$

De cette expression, on déduit l'expression générale de la loi de fiabilité en fonction du taux de défaillance :

$$R(t) = \exp[-\int_0^t \lambda(u)du] \tag{3.5}$$

Si le taux de défaillance est constant $\lambda(t)=\lambda$, la loi de fiabilité est la loi exponentielle :

$$R(t) = \exp(-\lambda t) \tag{3.6}$$

Si le taux de défaillance dépend du temps, les lois le plus souvent utilisées sont la loi de Weibull et la loi lognormale. La loi de Weibull a pour expression :

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)\right] \tag{3.7}$$

oùβ est le paramètre de forme et η est le paramètre d'échelle. Si β1, le taux de défaillance est croissant et si β=1 le système n'affiche aucune tendance, le taux de défaillance est constant et la loi de fiabilité est la loi exponentielle.

Une expression plus générale de la loi de Weibull est obtenue en posant η=exp(μ) et σ=1/β. Dans ce cas l'équation 7 devient :

$$F(t) = 1 - \exp\left[-\exp\left(\frac{lnt-u}{\sigma}\right)\right] \tag{3.8}$$

$$F(t) = \Phi_{sev}\left(\frac{lnt-u}{\sigma}\right) \tag{3.9}$$

oùΦ_{sev}est la fonction de répartition relative à la distribution des valeurs extrêmes minimales.

L'expression de la loi lognormale est donnée par :

$$F(t) = \Phi_{nor}\left(\frac{lnt-u}{\sigma}\right) \tag{3.10}$$

oùΦnor est la fonction de répartition de la distribution normale centrée réduite :

$$\Phi_{nor} = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) dw \tag{3.11}$$

Nous remarquons que l'équation [9] et [10] ont la même forme :

$$F(t) = \Phi\left(\frac{lnt-u}{\sigma}\right) \tag{3.12}$$

Il s'agit de l'expression générale caractérisant la famille des distributions log fonction d'un paramètre de localisation μ et d'un paramètre d'échelle σ comme décrit par Meeker et Escobar dans [18].

III.2.2. Fiabilité des systèmes complexes

D'une manière générale, les systèmes réels sont constitués de plusieurs composants et présentent plusieurs modes de défaillance ; de tels systèmes sont dits complexes. Une modélisation fonctionnelle sous forme de schéma-bloc de ces systèmes permet d'en déterminer la fiabilité. Chaque bloc représente un composant ou un mode de défaillance. Par ailleurs, les blocs sont considérés indépendants les uns des autres. Les associations fonctionnelles de blocs les plus couramment rencontrées sont les associations série, parallèle, série-parallèle, les redondances actives m parmi n et les redondances passives.

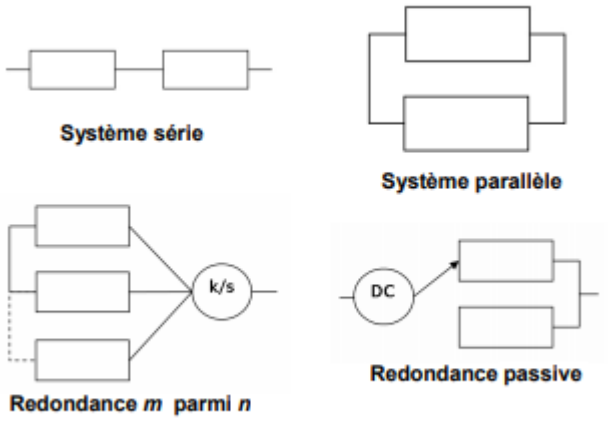


Figure III.2:systèmes complexes de fiabilité

- **Fiabilité d'un système série :**

Un système série fonctionne si et seulement si tous les composants fonctionnent. La fiabilité est calculée par la relation suivante :

$$\begin{aligned} R(t) &= \Pr(T > t) \\ &= \Pr[(T_1 > t) \cap (T_2 > t) \cap \dots] \\ &= \prod_i \Pr(T_i > t) = \prod_i R\,(t_i) \end{aligned} \tag{3.13}$$

- **Fiabilité d'un systèmeparallèle**:

Un système parallèle fonctionne si au moins un de ses composants fonctionne. La fiabilité pour ce système est donnée par :

$$\begin{aligned} R(t) &= 1 - F(t) \\ &= 1 - Pr[(T_1 \leq t)\ \cap (T_2 \leq t) \cap \dots] \end{aligned} \tag{3.14}$$

$$= 1 - \prod_i \Pr(T_i \leq t) = \prod_i R\,(t_i)$$

- **Redondance active m parmi n :**

Un système en redondance active m parmi n fonctionne si au moins m parmi les n composants redondants fonctionnent. Si les composants sont indépendants et identiquement distribués (R=Ri, i=1,2,…,n), alors la fiabilité du système est donnée par :

$$R(t) = \sum_{j=m}^{n} \binom{n}{j} R^j (1-R)^{n-j} \tag{3.15}$$

Avec

$$\binom{n}{j} = \frac{n!}{j!\,(n-j)}$$

- **Redondance passive :**

Dans une redondance passive les composants redondants ne sont mis en service que lorsque les composants du système sont défaillants. Dans l'exemple on à deux composants en redondance passive, le deuxième composant est mis en service lorsque le composant principal est défaillant. Si les composants ont la même loi de fiabilité R(t), la fiabilité de ce système est donnée par :

$$R_{sys}(t) = R(t) + \int_0^t f(\tau) R(t-\tau) d\tau \tag{3.16}$$

III.2.3. Estimation des paramètres des modèles de fiabilité

Dans ce paragraphe, nous présentons deux méthodes utilisées pour estimer les paramètres des fonctions de fiabilité : la méthode de régression et la méthode du maximum de vraisemblance.

III.2.3.2. Méthode de régression

La méthode de régression consiste à linéariser le modèle de fiabilité et à calculer ses paramètres à partir de la droite de régression qui s'ajuste au mieux aux points estimés (x_i, y_i).

Par exemple, la linéarisation de la loi exponentielle est obtenue en prenant le logarithme de la fonction de distribution :

$$R(t) = 1 - F(t) = \exp(-\lambda . t) \tag{3.17}$$

$$\ln[1 - F(t)] = -\lambda . t \tag{3.18}$$

Si on pose y=ln [1-F(t)] et x=t, l'expression (17) devient y=-λx, le paramètre λ est donné par la pente de la droite de régression qui s'ajuste au mieux aux points (t_i, F(t_i)). La fonction F(t) peut être estimée par la méthode des rangs moyens ou des rangs médians.

La forme linéaire des distributions appartenant à la famille des distributions log fonction d'un paramètre de localisation et d'un paramètre d'échelle est donnée par :

$$\Phi^{-1}(t) = \frac{\ln(t)}{\sigma} - \frac{\mu}{\sigma} \tag{3.19}$$

Si on pose y= Φ^{-1} (t), x=ln(t), σ=1/a et μ=-bσ, la relation (19) devient : y=ax+b, soit l'équation d'une droite.

III.2.3.2. Méthode du maximum de vraisemblance

La fonction de vraisemblance est égale ou proportionnelle à la probabilité d'observer les événements (défaillances). Pour un échantillon et un modèle donné, la fonction de vraisemblance peut être vue comme une fonction de paramètres inconnus p, les paramètres du modèle. La fonction totale de vraisemblance peut être écrite comme l'intersection d'observer les événements.

Supposons n événements indépendants, la fonction de vraisemblance est donnée par :

$$L(p) = L(p; DATA) = \prod_{i=1}^{n} L_i(p; data_i) \tag{3.20}$$

Où L_i(p ; $data_i$) est la probabilité d'observer l'événement i, datai regroupe les données de l'événement i. Les valeurs recherchées de p sont celles qui maximalisent la fonction de vraisemblance L(p).

Cette méthode peut être étendue quand certains événements sont censurés. Par exemple, les observations peuvent être censurées dans un intervalle ou à droite.

Si un temps est censuré dans l'intervalle [t_i-1, t_i[, cela signifie que la défaillance est survenue entre ti-1 et t_i, la fonction de vraisemblance correspondante est calculée par:

$$L_i(p) = \int_{t_{i-1}}^{t_i} f(t)dt = F(t_i) - F(t_{i-1}) \tag{3.21}$$

Si le temps t_i est censuré à droite, c'est-à-dire que la défaillance est survenue après t_i, l'expression de la fonction de vraisemblance est donnée par :

$$L_i(p) = \int_{t_i}^{\infty} f(t)dt = F(\infty) - F(t_i) = 1 - F(t_i) = R(t_i)) \quad (3.22)$$

Considérons m+1 intervalles d'observation $[t_0, t_1[$,$[t_1, t_2[$,…,$[t_m, t_{m+1}[$, la fonction de vraisemblance totale a pour expression :

$$L(p) = \prod_{i=1}^{m+1}[F(t_i) - F(t_{i-1})]^{d_i} \times [1 - F(t_i]^{r_i}) \quad (3.23)$$

avec di le nombre d'événements censurés dans l'intervalle $[t_i$-1, $t_i[$ et r_i le nombre d'événements censurés à droite de t_i. En pratique, si la défaillance est survenue à ti (donnée exacte), la fonction de vraisemblance est calculée à partir de la fonction de densité et non de la fonction de répartition. En effet, nous avons :

$$f(t) = \frac{dF(t)}{dt} = \frac{F(t_i) - F(t_i - \Delta i)}{\Delta i} \quad (3.24)$$

avec Δi=(t_i-t_{i+1}) ◊ 0. Etant donné que les Δi sont indépendants des paramètres p, la fonction de vraisemblance peut être estimée par :

$$L_i(p) \approx f(t_i, p) \quad (3.25)$$

Dans le cas des distributions log fonction d'un paramètre de localisation et d'un paramètre d'échelle, pour des données exactes et censurées à droite, la fonction de vraisemblance a pour équation :

$$L(\mu, \sigma) = \prod_{i=1}^{n} \left[\frac{1}{\sigma t_i} \phi\left(\frac{lnt_i - \mu}{\sigma}\right)\right]^{\delta_i} \left[1 - \Phi\left(\frac{lnt_i - \mu}{\sigma}\right)\right]^{1-\delta_i} \quad (3.26)$$

Avec δi =1 si ti correspond à une donnée exacte et δi =0 si la donnée est censurée à droite.

Chapitre IV : Application et Résultats

IV.1 Introduction

La fiabilité est née avec les câbles sous-marins dont les réparations même élémentaires étaient fort coûteuses à cause du déplacement d'un navire, mais souvent impossible selon les saisons. Des composants pouvant fonctionnés très longtemps sans tomber en panne ont donc été mis au point.

La fiabilité a ensuite subi un deuxième essor avec l'activité aéronautique et spatiale, les interventions étant impossibles en vol.

Enfin, elle a subi un troisième essor avec l'électronique mais dans un but de maintenance, donc différent. Dans ce cas, le matériel reste accessible mais muet, et la fiabilité permet de supprimer une partie des pannes, préventivement, ce que la maintenance ne pourrait faire.

La fiabilité a pour objectif de :

- Mesurer une garantie dans le temps ;
- Evaluer rigoureusement un degré de confiance ;
- Déchiffrer une durée de vie ;
- Evaluer avec précision un temps de fonctionnement ;
- Déterminer la stratégie de l'entretien ;
- Choisir le stock.

C'est un modèle probabiliste particulièrement bien adapté à l'étude statistique des pannes, (domaine mécanique et électrique).

Pour estimer la fiabilité nous avons appliqué la loi de Weibull parce qu'il est très souple, car la loi de probabilité a trois paramètres qui permettent d'ajuster correctement toutes sortes de résultats.

Contrairement au modèle exponentiel, la loi de probabilité de Weibull couvre les cas où le taux de défaillance est variable et permet donc de s'ajuster aux périodes de jeunesse et d'obsolescence.

Son utilisation implique la saisie des résultats des temps de bon fonctionnement entre deux défaillances.

La détermination des paramètres permettra d'évaluer la MTBF.

IV.2 Modèle de régression (linéaire)

En statistique, la régression linéaire est une mesure de la relation existant entre deux ou plusieurs variables. Cette relation peut également se définir par une interdépendance des éléments observés sur le marché.

Essentiellement, la technique de régression linéaire repose sur l'hypothèse qu'à défaut de pouvoir établir une relation fonctionnelle exacte entre une variable dépendante (aussi appelée « endogène » ou « expliquée ») et une ou plusieurs variables indépendantes (aussi appelées « exogènes » ou « explicatives ») reflétant des phénomènes observables, il est possible d'estimer une telle relation statistiquement, à partir d'une série d'observations, grâce à l'introduction d'un terme lié au hasard (terme stochastique), permettant de prendre en considération les erreurs d'estimation. C'est ce terme qui fait la différence entre une relation mathématique exacte et l'estimation probabiliste des paramètres de l'équation de régression.

Dans le cas de la régression simple, dans laquelle se retrouve une seule variable indépendante, la relation peut s'écrire de la façon suivante :

$Y = m X + b$ où :

Y = variable dépendante

X = variable indépendante

b= ordonnée à l'origine (constante)

m= coefficient de régression (paramètre estimé)

La droite de régression obtenue par la méthode des moindres carrés ordinaires appliquée à deux variables peut être représentée par un graphique à deux dimensions dans lequel l'axe des ordonnées supporte les valeurs de « Y » alors que les valeurs de « X » sont en abscisse.

Le tracé de la meilleure droite au milieu de points expérimentaux est un problème courant au laboratoire. Une méthode numérique objective appelée analyse de régression permet d'établir cette droite et de calculer les incertitudes associées. On ne considère ici que la méthode de régression la plus simple, la méthode des moindres carrés.

- Hypothèses:

Soit une relation linéaire $Y = m X + b$.

On considère que x est affecté d'incertitude négligeable mais que y est entaché d'erreurs expérimentales aléatoires. Ce cas simple s'applique facilement dans le cas d'une courbe d'étalonnage.

-Calcul de la droite des moindres carrés:

L'écart vertical de chaque point à la droite est appelé un résidu. La droite calculée par la méthode des moindres carrés est celle qui minimise la somme des carrés des résidus de tous les points (voir ouvrage spécialisé).

En pratique on définit trois grandeurs Sxx, Syy, Sxy:

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{N}$$

$$S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{N} \qquad (4.1)$$

$$S_{xx} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{N}$$

Où $x_i$et y_i sont des paires de données, N le nombre de paires et x et y sont les valeurs moyennes des variables.

On peut en déduire:

- la pente de la droite: m = S_{xy}/S_{xx}

- l'ordonnée à l'origine: b = $\bar{y}$ -m$\bar{x}$

- l'écart type des résidus: $S_r = \sqrt{\frac{S_{yy} - m^2 S_{xx}}{(N-2)}}$

- l'écart type de la pente: $S_m = \sqrt{\frac{S_r^2}{S_{xx}}}$

- l'écart type de l'ordonnée à l'origine: $S_b = S_r \sqrt{\frac{\sum {x_i}^2}{N \sum {x_i}^2 - (\sum x_i)^2}}$

Ces calculs sont facilement accessibles sur calculette. Par ailleurs il existe de nombreux logiciels sur micro-ordinateur, tableurs ou graphiques, intégrant ces calculs avec le tracé automatique des points et de la droite ce qui permet un contrôle visuel du résultat. Les écarts types obtenus sont analogues à l'écart type de données unidimensionnelles. On peut en déduire d'une manière similaire des limites de confiance comme dans 1).

-Le coefficient de corrélation : r ou coefficient de Pearson permet d'établir le niveau de validité de la relation linéaire entre les deux variables x et y :

$$r = \frac{N\sum x_i y_i - \sum x_i y_i}{\sqrt{[N\sum {x_i}^2 - (\sum x_i)^2][N\sum {y_i}^2 - (\sum y_i)^2]}} \quad (4.2)$$

Plus la valeur de r est proche de 1 plus grande est la probabilité d'une corrélation de linéarité entre les deux variables x et y. Ainsi un jugement sur la valeur de r peut être porté à partir de tables de valeurs critiques de r (voir bibliographie).

Néanmoins, le seul calcul de r ne peut suffire, il faut absolument vérifier graphiquement comment la droite passe au milieu des points. Ce contrôle est d'autant plus important que le nombre de points est faible.

IV.3. Méthode d'interpolation linéaire

L'interpolation linéaire est la méthode la plus simple pour estimer la valeur prise par une fonction continue entre deux points déterminés (interpolation). Elle consiste à utiliser pour cela la fonction affine (de la forme f(x) = m.x + b) passant par les deux points déterminés. Cette technique était d'un emploi systématique lorsque l'on ne disposait que de tables numériques pour le calcul avec les fonctions transcendantes : les tables comportaient d'ailleurs à cet effet en marge les « différences tabulaires », auxiliaire de calcul servant à l'interpolation linéaire.

Enfin l'interpolation linéaire est la base de la technique de quadrature numérique par la méthode des trapèzes.

Principe :

Supposons que l'on connaisse les valeurs prises par une fonction f(x) en deux points x_a et x_b :

$f(x_a) = y_a$

$f(x_b) = y_b$

La méthode consiste à approcher la fonction f par $\bar{f}$ telle que $\bar{f}\,(x_a) = y_a$ et $\bar{f}\,(x_b) = y_b$

Cette fonction a pour équation (trois formulations équivalentes) :

$$\bar{f}\,(x) = \frac{y_a - y_b}{x_a - x_b}x + \frac{x_a.y_b - x_b.y_a}{x_a - x_b} \quad (4.3)$$

On peut aussi écrire (formule de Taylor-Young au premier ordre)

$$\bar{f}(x) = y_a + (x - x_a)\frac{y_b - y_a}{x_b - x_a} \quad (4.4)$$

Aussi :

$$\bar{f}(x) = \frac{x_b - x}{x_b - x_a} y_a + \frac{x - x_a}{x_b - x_a} y_b \quad (4.5)$$

IV.4 Procédure de régression et interpolation pour la modélisation de la fiabilité

IV.4.1 Principales lois utilisées pour modéliser la fiabilité

Dans les études de fiabilité des différents équipements, une variable aléatoire continue ou discrète peut être distribuée suivant diverses lois qui sont principalement on site: La loi exponentielle, la loi de Weibull, la loi normale, loi log-normale, la loi binomiale et la loi de Poisson. Alors dans notre cas nous utilisons la loi de Weibull.

Elle est la plus couramment utilisée en fiabilité électronique pour décrire la période durant laquelle le taux de défaillance des équipements est considéré comme constant. Elle décrit le temps écoulé jusqu'à une défaillance, ou l'intervalle de temps entre deux défaillances successives.

IV.4.2 La loi de WEIBULL

C'est une loi continue à trois paramètres, donc d'un emploi très souple. En fonction de la valeur de ses paramètres, elle peut s'ajuster à toutes sortes de résultats expérimentaux. Cette loi a été retenue pour représenter la durée de vie des pièces mécaniques. Son utilisation implique la saisie des résultats des temps de bon fonctionnement entre deux défaillances.

IV.5 Résultats et interprétation de l'étude de la fiabilité de la turbine à gaz

IV.5.1 Situation du problème et objectif de l'étude

L'exploitation normale d'une turbine à gaz dans un environnement rude tel que le Sahara nécessite une étude et un diagnostic rigoureux afin d'assurer un bon fonctionnement de l'unité. Dans cette étude, on s'est intéressé à des données techniques de terrain, recueillis de l'historique d'une installation de turbine à gaz afin d'évaluer au mieux sa fiabilité et sa sûreté de fonctionnement ainsi que de proposer un planning scientifique des actions de maintenance préventive.

Afin de bien pouvoir mieux estimer la fiabilité d'une turbine à gaz, on prise comme modèle la turbine GEMS5002B de la station de traitement et compression située aux champs de TFT inamens.

L'allure de l'histogramme de pannes nous révèle déjà que la distribution de pannes de turbine suit la loi de Weibull.

IV.6. Partie pratique

Dans cette partie de notre travail nous allons étudier deux cas pour définera a la fin les paramètres de la fiabilité d'une turbine à gaz qui a enregistré un manque de données. Dans les deux cas, nous allons utiliser les deux méthodes citées ci desus à savoir la méthode de régression et d'interpolation.

IV.6.1. Premier cas d'étude

Les données

Il est possible de déterminer les valeurs de : β, η et γ graphiquement sur un papier gradué spécial dit papiers d'Allan Palit.

Pour cela il faut:

- classer les TBFi par ordre croissant.

- Calculer la fréquence relative pour chaque TBFi par la méthode des rangs médians:

- Si N>50 ; $F_{(i)} = \frac{i}{N}$

- Si 20<N ; $F_{(i)} = \frac{i}{N+1}$

- Si N<20 : $F_{(i)} = \frac{i-0.3}{N+0.4}$

Ordre i	TBFi (h)	F (i) (%)
1	2993	6
2	?	14.91
3	5036	24
4	6150	32
5	6919	41
6	?	50
7	8862	59
8	?	67.54
9	11488	76
10	?	85.08
11	13545	94

Tableau IV.1: des données de départ incomplètes

IV.6.2. la méthode de régression

IV.6.2.1. Etapes de la méthode de régression linéaire avec MATLAB

L'analyse de régression linéaire utilise la méthode des « moindres carrés » pour tracer une droite sur l'ensemble des observations, et analyse l'incidence des variables indépendantes sur la variable dépendante unique.

La prochaine étape avant de faire la régression est de tracer un graphique de la relation entre X et Y. On voit que la relation entre X et Y est bien de forme linéaire (d'après le graphique ci-dessous).

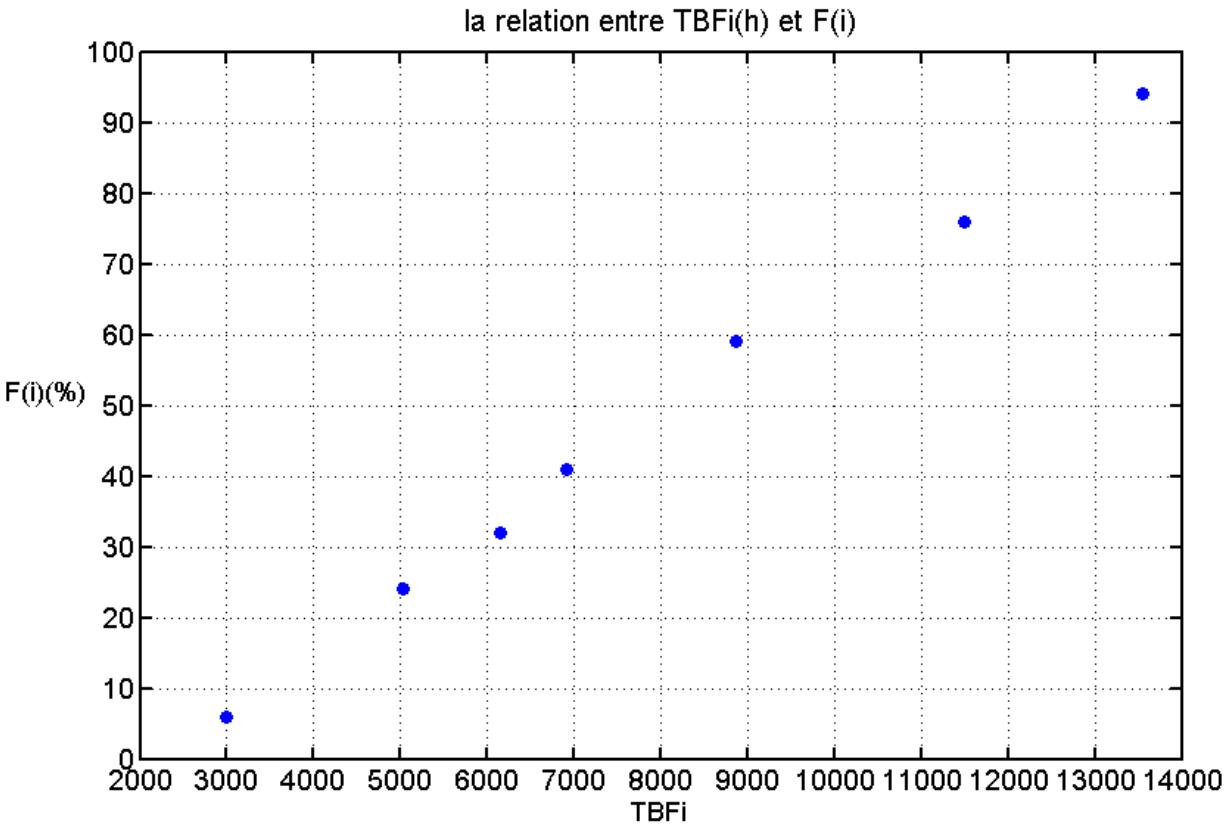


Figure IV.1:la relation entre TBFi (h) et F (i) (%)

Effectuons maintenant la régression entre TBFi (h) et F (i) (%). On peut la faire par deux méthodes.

-**Premiére méthode** : on écrit un code utilisant les commandes ' **polyfit et polyval** ' qui nous donne la relation linéaire et graphiquement droite de régression.

-**Deuxiéme méthode** : utilisons la figure de la relation entre TBFi (h) et F (i) (%) on va à '**tools**', on choisit '**basic fitting**' et après on coche '**linear**', '**show equations'**et ça nous donne la courbe de régression et l'équation linéaire.

La figure ci-dessous représente le résultat obtenu :

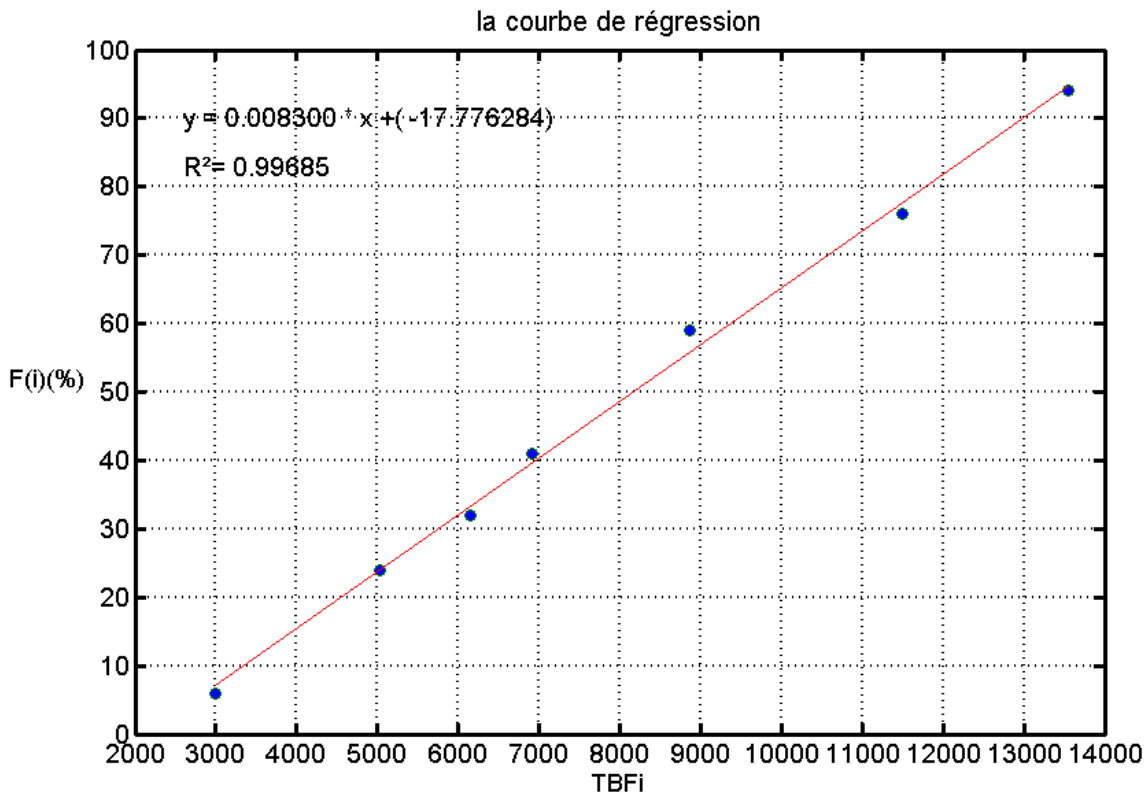


Figure IV.2:la courbe de régression

F (i) (%) = -17.7763 + 0.0083 * TBFi (h)

Les coefficients : sont les coefficients de la droite de régression

$\hat{Y} = b0 + b1*X$

b0 = -17.7763 (ordonnée à l'origine)

b1 = 0.0083 (pente de la droite)

- le coefficient de détermination R² : il donne une idée du % de variabilité de la variable à modéliser, et plus le coefficient R² est proche de 1, plus il y a une corrélation et meilleur est le modèle. Dans notre exemple, 99.68% de la variabilité de F (i) (%) est expliquée par TBFi (h).

En effet, d'après le graphique de la courbe de régression, il est clair qu'il y a une relation significative entre TBFi (h) et F(i) (%).

IV.6.2.2. Remplacement des données perdues

Ordre i	TBFi (h)	F (i) (%)
1	2993	6
2	3938.1	14.91
3	5063	24
4	6150	32
5	6919	41
6	8165.8	50
7	8862	59
8	10279	67.54
9	11488	76
10	12392.3	85.08
11	13545	94

Tableau IV.2: Remplacement les données perdues pour régression

On porte sur le papier d'Allan palit : (voir annexe 1)

Le papier du Weibull sert à déterminer graphiquement les paramètres d'une loi deWeibull, dans le cas où γ = 0.

F(t)=1-exp $\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \Leftrightarrow \ln\left(1-F(t)\right) = -\left(\frac{t}{\eta}\right)^{\beta}$

$\Leftrightarrow \ln(-\ln\left(1-F(t)\right) = \beta ln\left(\frac{t}{\eta}\right) \Leftrightarrow \ln(-ln\left(1-F(t)\right) = \beta\ln(t) - \beta\ln(\eta)$

$Y = \beta X - \beta\ln(\eta)$on pose : X=ln(t) et $Y = \ln(\ln\left(\frac{1}{1-F(t)}\right))$

D'où Y=aX+b est une équation d'une droite.

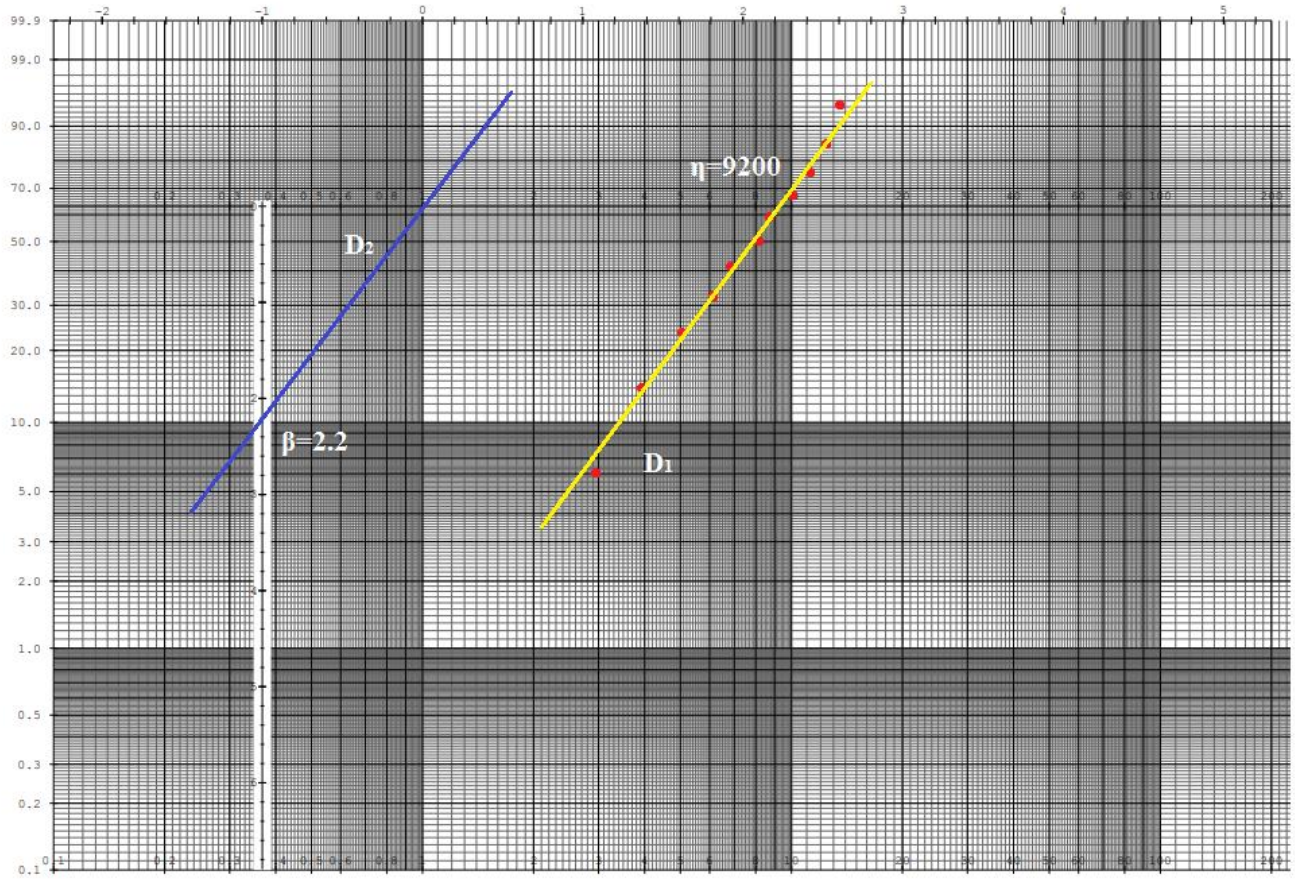


Figure IV.3: Détermination graphique des paramètres de Weibull

Détermination de β et η

La droite D_1 de régression coupe l'axe A (t, η) à l'abscisse t = η

Si le nuage de point est une droite donc

(justification : quand γ = 0 ln (t- γ) = ln(t) = ln(η) donc t = η).

β est la pente de D1 et pour obtenir sa valeur, nous traçons une droite D_2 parallèle à D_1 passant par le point η =1 (origine de X, Y). La droite D_2 coupe l'axe β en un point qui donnesa valeur.

Sur l'axe **A** les valeurs **TBFi**.

Sur l'axe **B** les valeurs de F (i) associées.

On obtient un nuage de point, si ce dernier peut être représenté approximativement par une droite D_1.

Donc : γ =0

La droite D_1 coupe l'axe (t ; η) en η

η=9200

Nous traçons la droite D_2 parallèle à D_1, passant par le point η =1. On detémermine ensuite β sur

l'axe B

Cette droite D_2 coupe l'axe (β, b) en β

β=2.2

Exploitation

1. Si β < 1 alors λ (t) décroît, c'est la période de jeunesse (rodage).

2. Si β = 1 alors λ (t) = $\frac{1}{\eta}$ (constant) on a l'indépendance du processus et du temps.

3. Si β > 1 alors λ (t) croit, c'est la période d'obsolescence que l'on peut analyser plus finement pour orienter un diagnostic.

Si 1,5 < β < 2,5 : phénomène de fatigue.

Si 3 < β <4 : phénomène d'usure, (débute au temps t = γ).

Si β ≈ 3,5 : la densité de probabilité est symétrique, la distribution deprobabilité est normale, alors que le matériel électronique monte une longue phase de vie àtaux de défaillance constant. Le matériel mécanique, en phénomènes d'usure, ne montrepas de palier dans la courbe en baignoire et doit être modélisé par la loi de Weibull.

IV.6.2.3. Recherche aux indicateurs de la fiabilité

On utilise les tables donnant Aet B (voir Annexe2) telles que :

-$MTBF=A\boldsymbol{\eta}+\gamma$

-L'écart type des temps de fonctionnement :=$B\boldsymbol{\eta}$

Nous pouvons connaître la variance : V=σ^2

Tracés et applications numériques des lois de probabilité dont les équations sont définies par les trois paramètres de Weibull.

A chaque instant, nous pouvons ainsi, graphiquement ou analytiquement déterminer :

- La fiabilité : $$R(t) = e^{-(\frac{t-\gamma}{\eta})^{\beta}}$$

- La fonction de répartition : $$F(t) = 1 - e^{-(\frac{t-\gamma}{\eta})^{\beta}}$$

- La densité de défaillance : $$f(t) = \frac{\beta}{\eta}(\frac{t-\gamma}{n})^{\beta-1}e^{-(\frac{t-\gamma}{\eta})^{\beta}}$$

- Le taux de défaillance instantané : $$\lambda(t) = \frac{\beta}{\eta}(\frac{t-\gamma}{\eta})^{\beta-1}$$

Les relations réciproques, en particulier l'instant t, associé à un seuil de fiabilité :

$$t = \gamma + [Ln\frac{1}{R(t)}]^{1/\beta}$$

La durée de vie nominale : L_{10}=γ+ η(0.105) $^{1/\beta}$

$\boldsymbol{\gamma}$	0
$\boldsymbol{\eta}$	9200
B	2.2
La fiabilité	$\mathbf{R(t) = e^{-\ (\frac{t}{9200})^{2.2}}}$
La fonction de repartition	$\mathbf{F(t) = 1 - e^{-(\frac{t}{9200})^{2.2}}}$
La densité de défaillance	$\mathbf{f(t) = \frac{2.2}{9200}(\frac{t}{9200})^{1.2}e^{-(\frac{t}{9200})^{1.2}}}$
Le taux de défaillance instantané	$\boldsymbol{\lambda = \frac{2.2}{9200}(\frac{t}{9200})^{1.2}}$
A	0.8856
B	0.425
$\boldsymbol{\sigma}$	3910 h
MTBF	8147.52 h
$\boldsymbol{L_{10}}$	3302 h

Tableau IV.3: Tableau des résultats obtenus par régression

A la pochaine étape nous trace les graphes donnants les parametres et les fonctions de la fiabilité de la turbine.

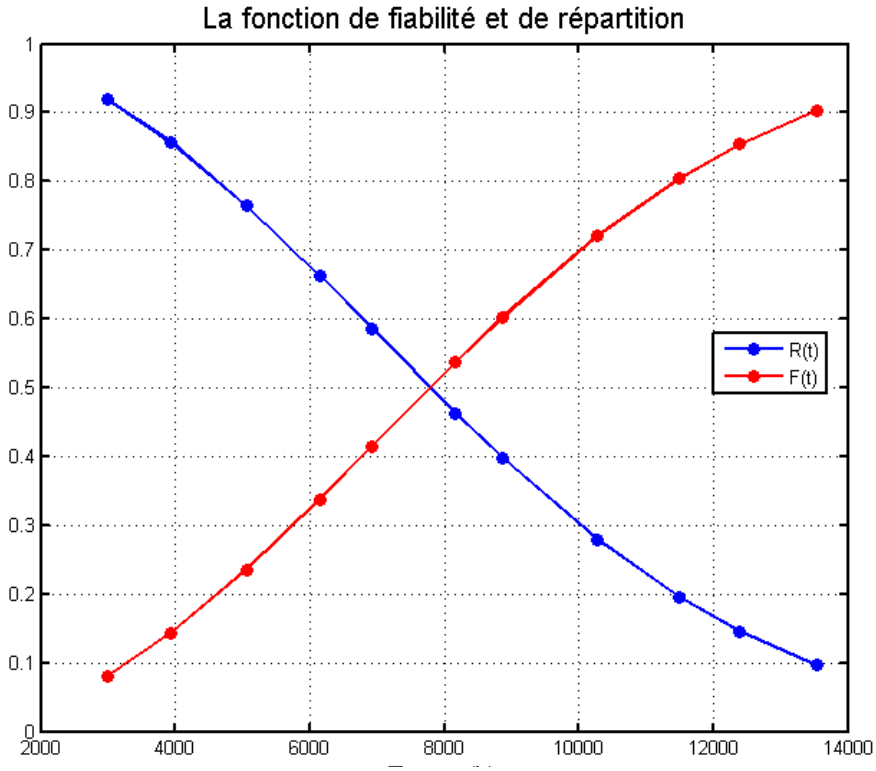


Figure IV.4: La fonction de fiabilité et de répartition (régression)

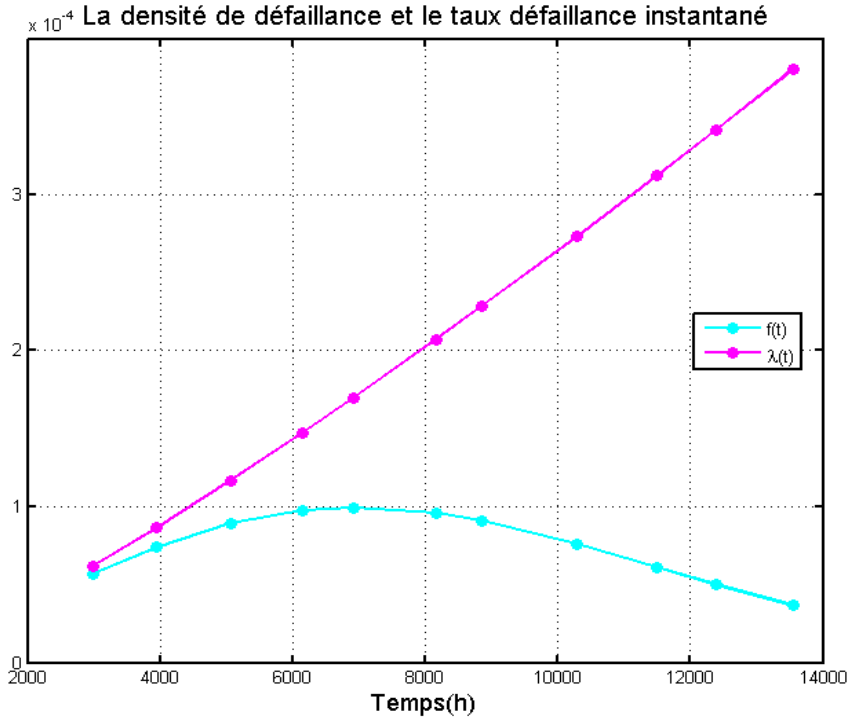


Figure IV.5: La densité de défaillance et le taux défaillance instantané

IV.6.2.4. Le test d'adéquation de modèle

Pour la validation du modèle élaboré, il est recommandé d'appliquer le test de Kolmogorov-Smirnov (anexxe 3), pour lequel :

On pose les conditions suivantes :

Si D max < D n,α ; Le modèle est accepté.

Si D max > D n,α ; Le modèle est rejeté.

où D exprime la plus grande valeur absolue tirée du tableau 3, cette valeur absolue est égale à la fonction F (t) pratique moins la fonction F (t) théorique**.**

N^0	TBF	f(ti)*10^{-4}	R(t)	Fth(ti) = 1-R(t)	f(ti)- Fth(ti)
1	**2993**	0.36569694	0.918927864	0.081072135	0.284624805
2	**3938.1**	0.49834150	0.856730934	0.143269065	0.355072437
3	**5063**	0.57108261	0.764326867	0.235673132	0.335409484
4	**6150**	0.61158270	0.662138997	0.337861002	0.273721703
5	**6919**	0.74007885	0.586096034	0.413903965	0.326174891
6	**8165.8**	0.76231225	0.463363330	0.536636669	0.225675586
7	**8862**	0.89260250	0.398140148	0.601859851	0.290742649
8	**10279**	0.91025595	0.279063092	0.720936907	0.189319047
9	**11488**	0.96030592	0.195917135	0.804082864	0.156223062

10	**12392.3**	0.97653714	0.145765683	0.854234316	0.122302827
11	**13545**	0.99565812	0.096138436	0.903861563	0.091796556

Tableau IV.4 : La densité de défaillance et le taux défaillance instantané (régression)

Détermination de la valeur absolue :

α : niveau significatif Pour nombre de panne = 11 (notre cas), α généralement pris est 5% (Selon la loi de Kolmogorov-Smirnov) (voir annexe 3).

D n, α =$\boldsymbol{D_{11.\ 0.05}}$= **0.391**

Donc dans notre cas :

$\boldsymbol{D_{max}}$ = **0.355**⇨ Le modèle est accepté

IV.6.3.Etude par la méthode d'interpolation

Toujours avec les memes données utilisées precedemment nous utilisons maitenant une autre technique à savoir l'interpolation afin de trouver toujours les données manquantes et refaire le meme travail et faire une comparaison entre les deux techniques.

IV.6.3.1. Comment effectuer une interpolation linéaire avec Matlab

Pour effectuer une interpolation sur MATLAB, on utilise un code qui nous permet de calculer les données manquantes.

Ou on utilise la figure de la relation entre TBFi (h) et F (i) (%) on va à '**tools**',on choisit '**basic fitting**' et après on fait coché '**shape-preserving interpolant**' et ça nous donne la courbe d'interpolation.

Le resultat obtenu est présenté sur la figure suivante :

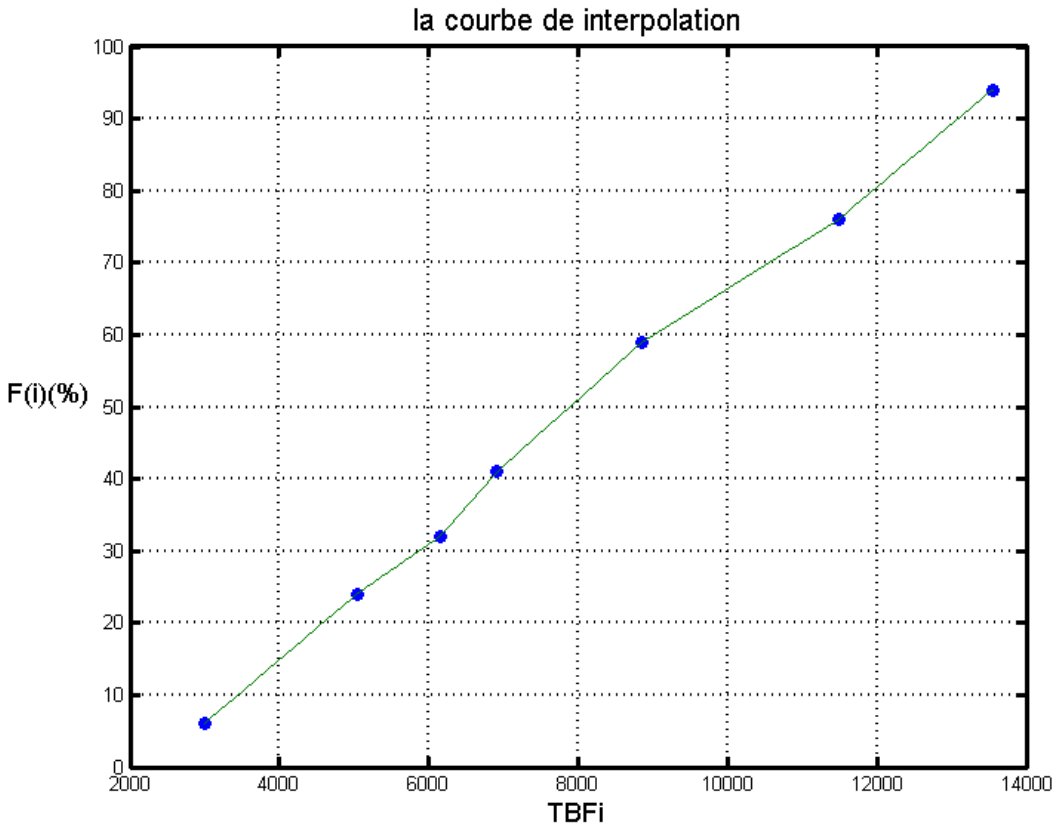


Figure IV.6: la courbe d'interpolation

IV.6.3.2. Remplacement les données perdues

Ordre i	TBFi (h)	F (i) (%)
1	2993	6
2	4014.5	14.91
3	5063	25
4	6150	32
5	6919	41
6	7890.5	50
7	8862	59
8	10175	67.54
9	11488	76
10	12517	85.08
11	13545	94

Tableau IV.5: Remplacement les données perdues pour interpolation

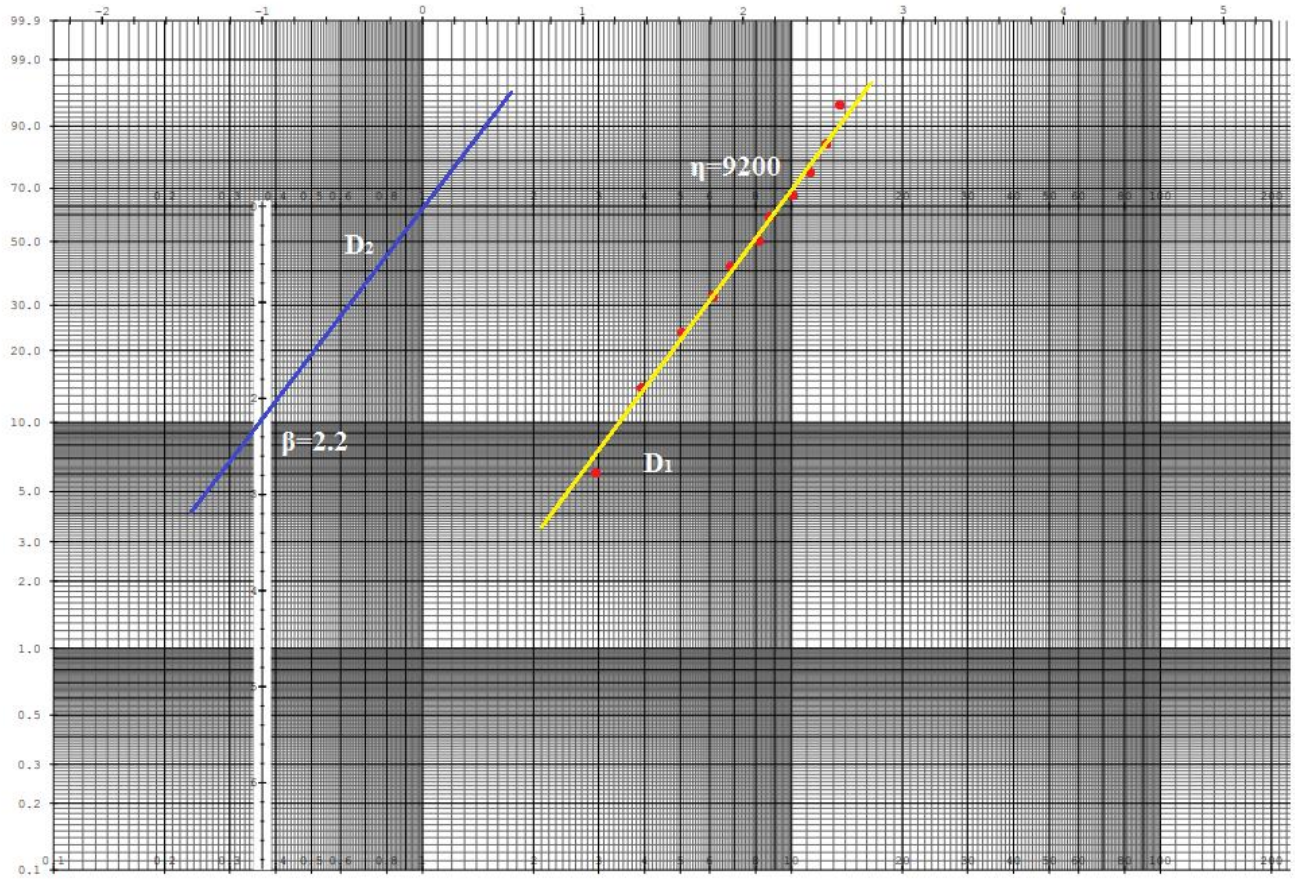


Figure IV.7 : Détermination graphique des paramètres de Weibull

Remarque : pour les deux cas (interpolation et régression) on obtient les mêmes paramètres de Weibull

Le tableau suivant résume ces résultats :

γ	**0**
η	9200
B	2.2
La fiabilité	$R(t) = e^{-\left(\frac{t}{9200}\right)^{2.2}}$
La fonction de repartition	$F(t) = 1 - e^{-\left(\frac{t}{9200}\right)^{2.2}}$
La densité de défaillance	$f(t) = \frac{2.2}{9200}\left(\frac{t}{9200}\right)^{1.2} e^{-\left(\frac{t}{9200}\right)^{1.2}}$
Le taux de défaillance instantané	$\lambda = \frac{2.2}{9200}\left(\frac{t}{9200}\right)^{1.2}$

A	0.8856
B	0.425
$\boldsymbol{\sigma}$	3910 h
MTBF	8147.52 h
$\boldsymbol{L_{10}}$	3302 h

Tableau IV.6: Tableau des résultats pour interpolation

Nous tracons a la fin les fonctions de fiabilité de la turbine aorés avoir determiner ces parametres et celon on remlacant les données manquantes bien on utisant la technique d'interpolation pour déduire ces données.

Les resultats obtenus sont présentés sur les figures ci-dessous :

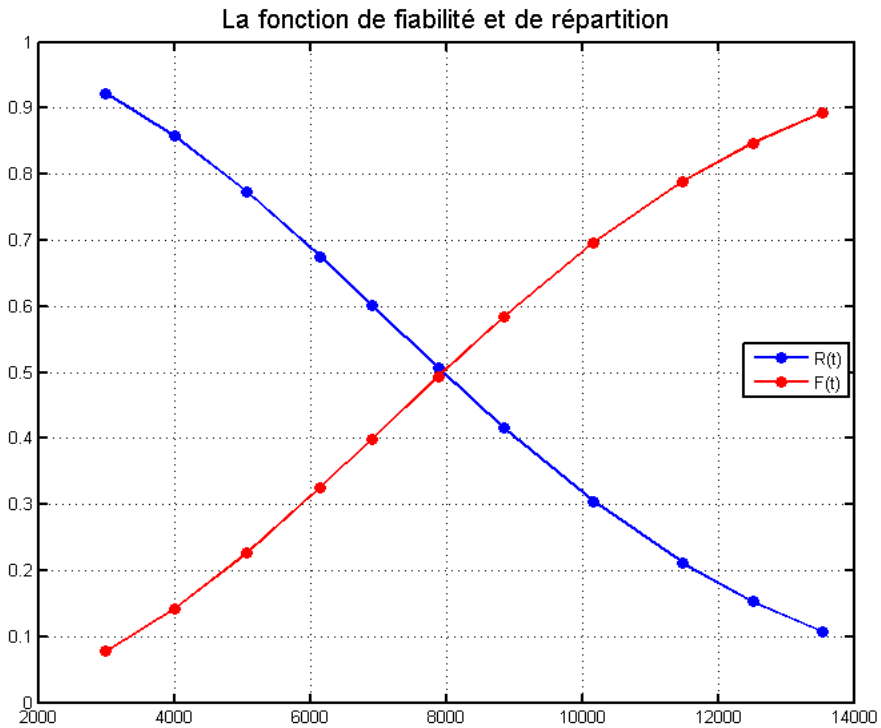


Figure IV.8: La fonction de fiabilité et de répartition (interpolation

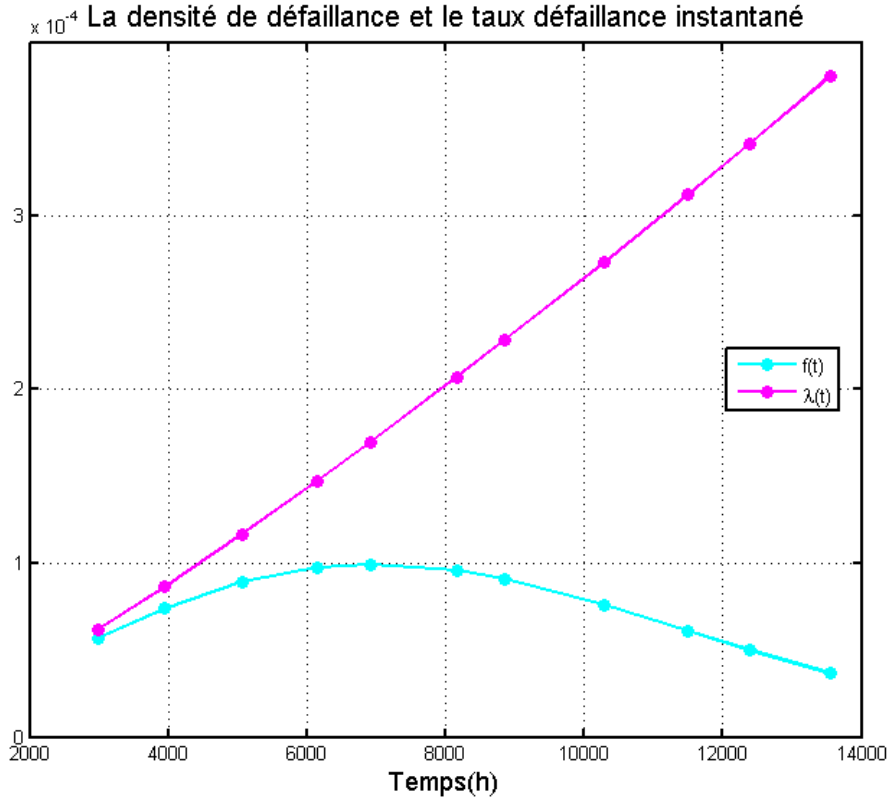


Figure IV.9: La densité de défaillance et le taux défaillance instantané (interpolation)

IV.6.3.3 Le test d'adéquation de modèle

Pour valider le modele obtenu nous fesons un test de d'adéquation. Il existe plusiures types de test ; pour notre cas nous utilisons le test Kolmogorov-Smirnov**.**

N°	TBF	f(ti)*10^{-4}	R(t)	Fth(ti) = 1-R(t)	f(ti)- Fth(ti)
1	2993	0.36569694	0.918927864	0.081072135	0.284624805
2	4014.5	0.48319034	0.851029885	0.148970114	0.334220229
3	5063	0.57108261	0.764326867	0.235673132	0.335409484
4	6150	0.61158270	0.662138997	0.337861002	0.273721703
5	6919	0.75230167	0.586096034	0.413903965	0.338397705
6	7890.5	0.77463374	0.490004719	0.509995280	0.264638463
7	8862	0.89260250	0.398140148	0.601859851	0.290742649
8	10175	0.91025595	0.287055352	0.712944647	0.197311308
9	11488	0.97457474	0.195917135	0.804082864	0.170491880
10	12517	0.97653714	0.139645994	0.860354005	0.116183137
11	13545	0.99565812	0.096138436	0.903861563	0.091796556

Tableau IV.7: Test d'adéquation de modèle (interpolation)

D n, α = $\boldsymbol{D_{11\ 0.05}}$= **0.391**

Donc dans notre cas :

$\boldsymbol{D_{max}}$= **0.338** ⇨ Le modèle est accepté.

IV.6.4. Interpretation des résultats

A vrais dire les résultats que nous avons obtenus par les deux techniques sont satisfait et se rapprochent de la realité et la preuve que les valeurs obtenues par les deux techniques sont presque les mêmes.

IV.6.5.Deuxième cas d'étude

Pour donner plus de crédibilité à notre étude nous augmontons la taille de l'echantillon de donnés **(42 TBFI dont 7 TBFI manquants**) et nous faisons le meme travail éffectué dans la premiere partie de notre travail ; c'est à dire essayer de déduire les valeurs manquantes et cela on utilisant toujours les deux techniques la regression et l'interpolation. Comme precedemment nous calculons les parametres de fiabilité et nous tracons ces fonctions.

Les données : les données de départ sont présentées dans le tableau ci dessous

Calculer la fréquence relative pour chaque TBFi par la méthode des rangs médians :

- Si 20<N ; $F_{(i)} = \frac{i}{N+1}$

Ordre i	TBFi (h)	F (i) (%)	Order i	TBFi (h)	F (i) (%)
1	267	2.3	**22**	4573	51.16
2	1110	4.6	**23**	4591	53.48
3	1746	6.9	**24**	4808	55.81
4	1767	9.3	**25**	4826	58.13
5	1810	11.6	**26**	4940	60.46
6	1978	13.9	**27**	?	62.79
7	?	16.2	**28**	5198	65.11
8	2744	18.6	**29**	5212	67.44
9	2996	20.9	**30**	5222	69.76
10	3129	23.25	**31**	5234	72.09
11	3183	25.58	**32**	5575	74.41
12	?	27.9	**33**	5587	76.74

13	3507	30.23	**34**	5693	79.06
14	3510	32.55	**35**	?	81.39
15	3840	34.88	**36**	6477	83.72
16	3955	37.2	**37**	6483	86.04
17	4223	39.53	**38**	6983	88.37
18	?	41.86	**39**	?	90.7
19	4367	44.18	**40**	7959	93.02
20	4410	46.51	**41**	8408	95.34
21	4562	48.83	**42**	8908	97.67

Tableau IV.8: des données de départ incomplètes

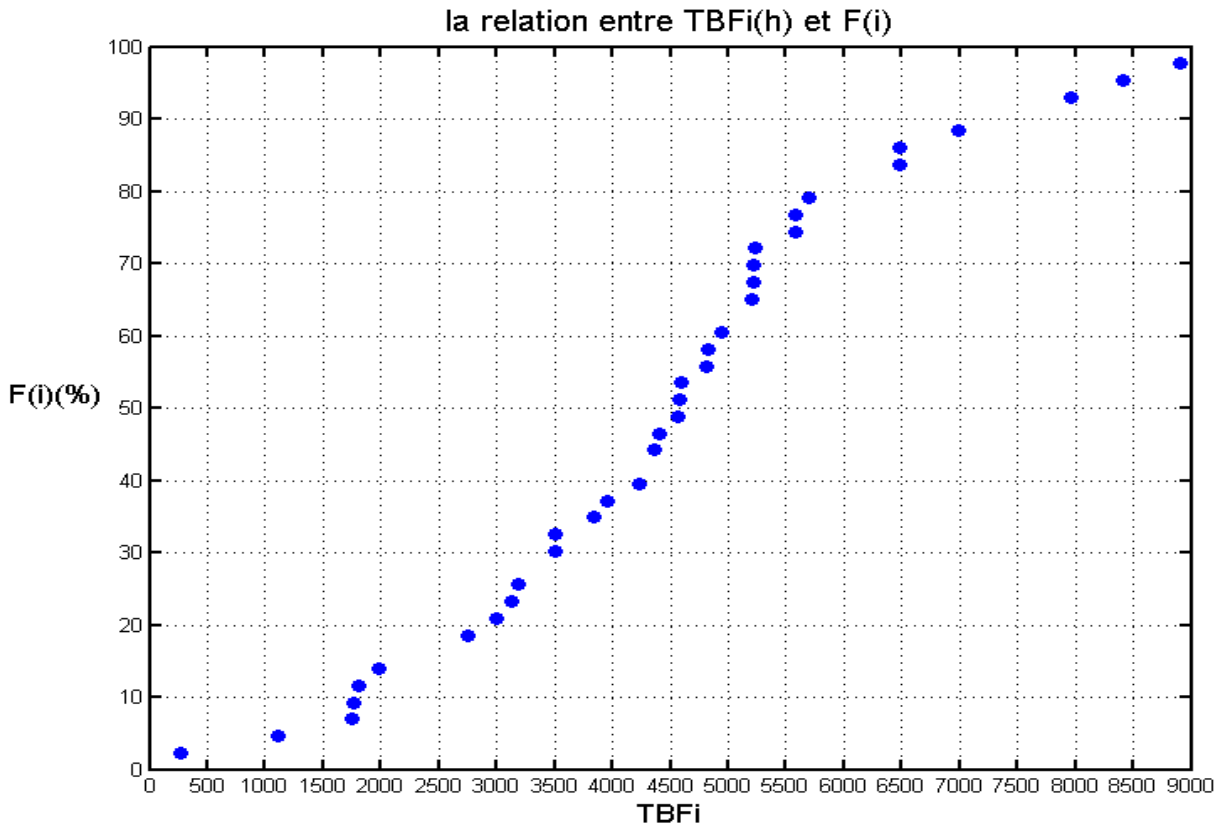


Figure IV.10:la relation entre TBFi (h) et F (i) (%)

IV.6.6. Etude par la méthode de régression

Nous faisons le meme travail et avce la même façon que la première partie.

Après La régression sur MATLAB, nous obtenons la courbe suivante.

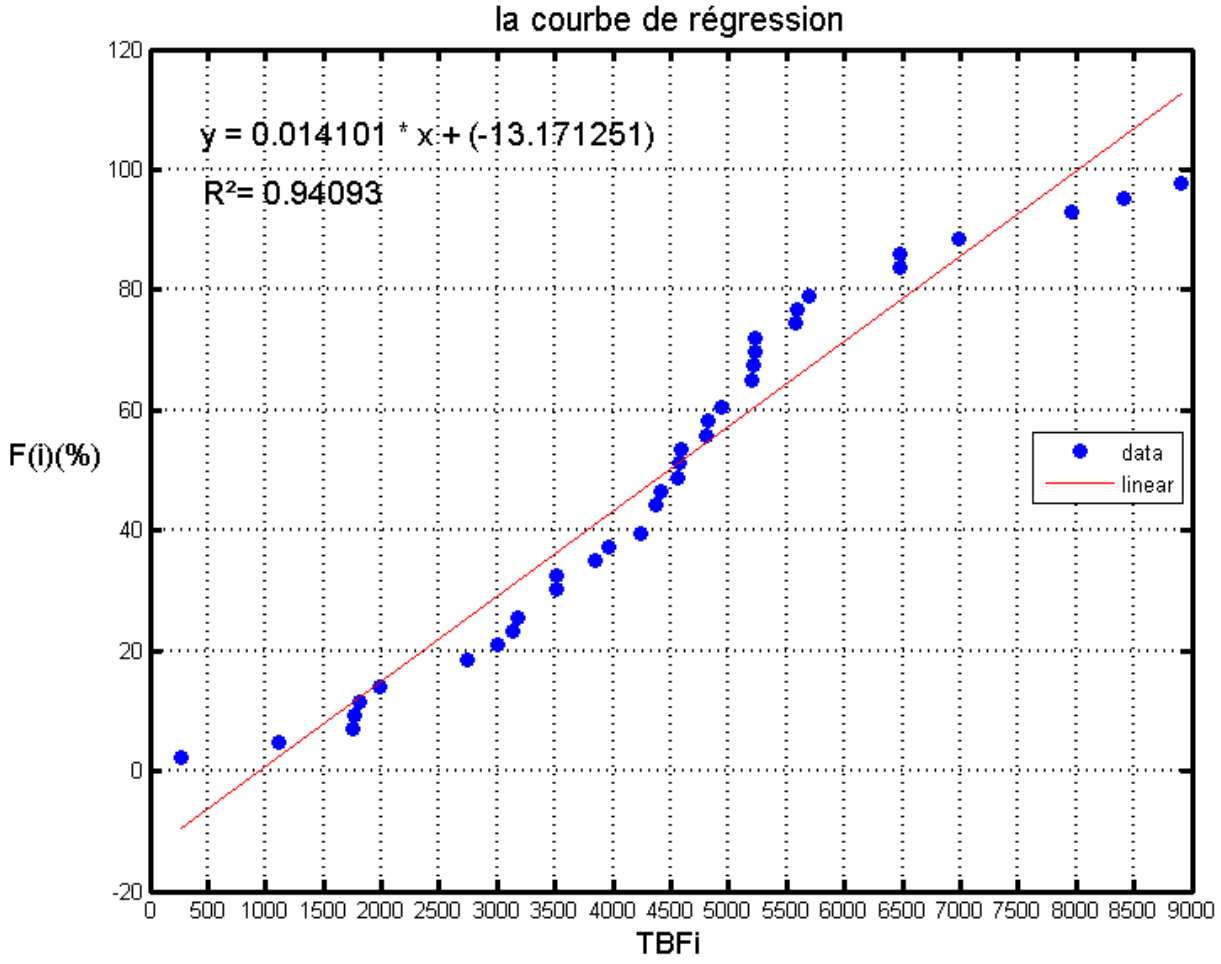


Figure IV.11:la courbe de régression

F (i) (%) = -13.171251 + 0.0014101 * TBFi (h)

Les coefficients : sont les coefficients de la droite de régression

$\hat{Y} = b_0 + b_1 * X$

b_0 = -13.171251 (ordonnée à l'origine)

b_1 = 0.0014101 (pente de la droite)

IV.6.6.1. Remplacement les données perdues

Ordre i	TBFi (h)	F (i) (%)	Order i	TBFi (h)	F (i) (%)
1	267	2.3	**22**	4573	51.16
2	1110	4.6	**23**	4591	53.48
3	1746	6.9	**24**	4808	55.81
4	1767	9.3	**25**	4826	58.13
5	1810	11.6	**26**	4940	60.46
6	1978	13.9	**27**	5386	62.79
7	2088	16.2	**28**	5198	65.11
8	2744	18.6	**29**	5212	67.44
9	2996	20.9	**30**	5222	69.76
10	3129	23.25	**31**	5234	72.09
11	3183	25.58	**32**	5575	74.41
12	2913	27.9	**33**	5587	76.74
13	3507	30.23	**34**	5693	79.06
14	3510	32.55	**35**	6706	81.39
15	3840	34.88	**36**	6477	83.72
16	3955	37.2	**37**	6483	86.04
17	4223	39.53	**38**	6983	88.37

18	3902	41.86	**39**	7365	90.7
19	4367	44.18	**40**	7959	93.02
20	4410	46.51	**41**	8408	95.34
21	4562	48.83	**42**	8908	97.67

Tableau IV.9: Remplacement les données perdues pour régression

Remarque : les données obtenues par la méthode de régression sont aberrantes et pour cela on ne peut pas les utiliser ensuite pour faire des calculs et nous nous contontons des méthodes d'interpolation.

IV.6.7. Etude par la méthode d'interpolation

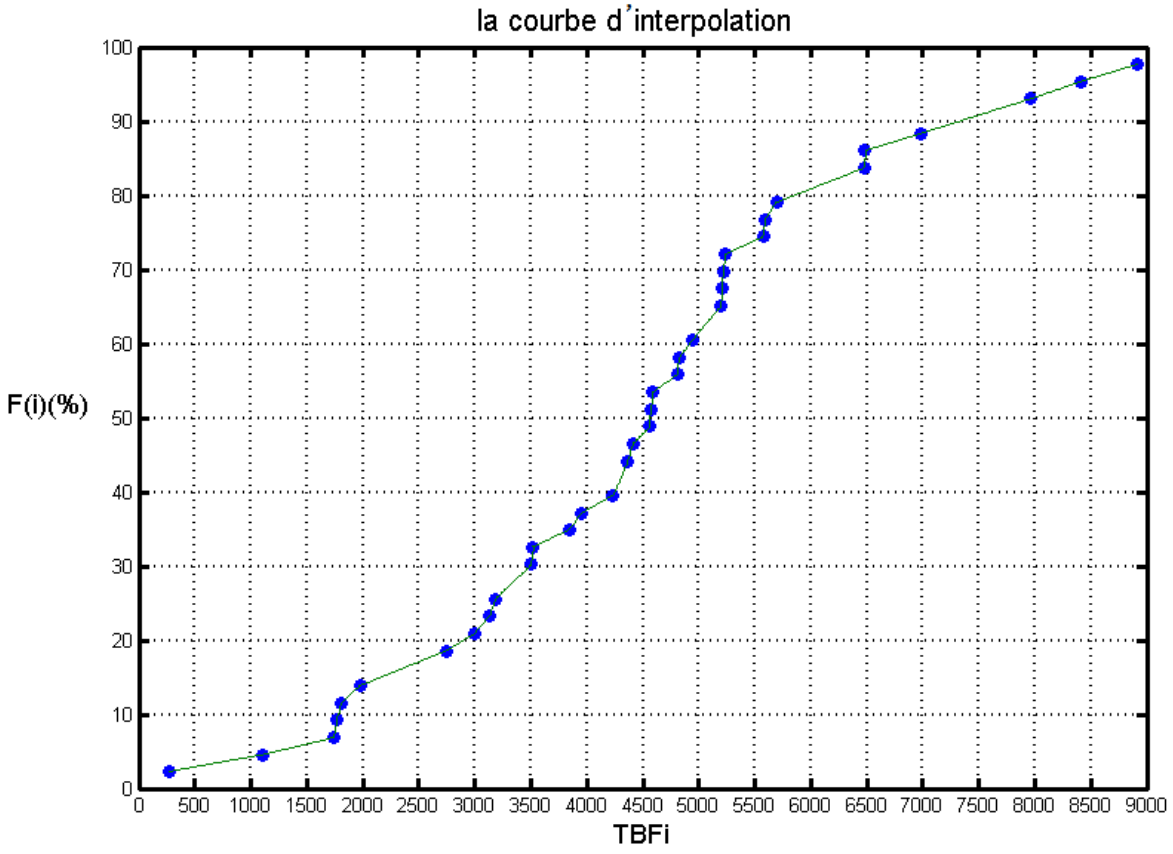


Figure IV.12 : La courbe obtenue après l'interpolation

IV.6.7.1.Remplacement les données perdues

Ordre i	TBFi (h)	F (i) (%)	Order i	TBFi (h)	F (i) (%)
1	267	2.3	**22**	4573	51.16
2	1110	4.6	**23**	4591	53.48
3	1746	6.9	**24**	4808	55.81
4	1767	9.3	**25**	4826	58.13
5	1810	11.6	**26**	4940	60.46
6	1978	13.9	**27**	5069	62.79
7	2361	16.2	**28**	5198	65.11
8	2744	18.6	**29**	5212	67.44
9	2996	20.9	**30**	5222	69.76
10	3129	23.25	**31**	5234	72.09
11	3183	25.58	**32**	5575	74.41
12	3345	27.9	**33**	5587	76.74
13	3507	30.23	**34**	5693	79.06
14	3510	32.55	**35**	6085	81.39
15	3840	34.88	**36**	6477	83.72
16	3955	37.2	**37**	6483	86.04
17	4223	39.53	**38**	6983	88.37
18	4300	41.86	**39**	7471	90.7

19	4367	44.18	**40**	7959	93.02
20	4410	46.51	**41**	8408	95.34
21	4562	48.83	**42**	8908	97.67

Tableau IV.10: Remplacement les données perdues après interpolation

IV.6.7.2 Détermination des paramètres de Weibull par MATLAB

Pour la deuxième partie, on utilise les commandes "**wblplot**" et "**wblfit**" pour determiner graphiquement le papier d'Allan Plait et numériquement les paramétres de Weibull

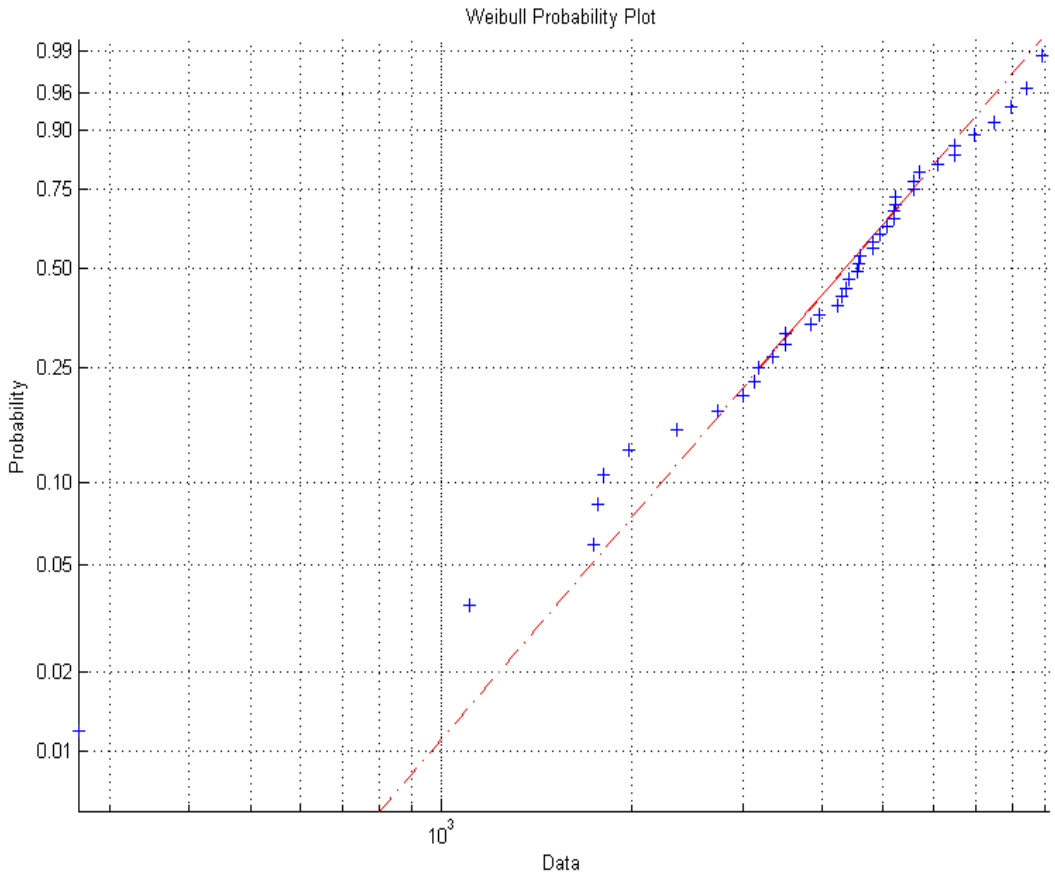


Figure IV.13: Papier de Weibull par MATLAB

Resultats du matlab :

D'aprés le graphe obtenu nous deduisons les valeurs de parametres de la fiabilité qui sont :

η=5000

β=2.4

$\boldsymbol{\gamma}$	**0**
$\boldsymbol{\eta}$	5000
B	2.4
La fiabilité	$\mathbf{R(t) = e^{-\left(\frac{t}{5000}\right)^{2.4}}}$
La fonction de repartition	$\mathbf{F(t) = 1 - e^{-\left(\frac{t}{5000}\right)^{2.4}}}$
La densité de défaillance	$\mathbf{f(t) = \frac{2.4}{5000}\left(\frac{t}{5000}\right)^{1.4} e^{-\left(\frac{t}{5000}\right)^{2.4}}}$
Le taux de défaillance instantané	$\boldsymbol{\lambda = \frac{2.4}{5000}\left(\frac{t}{5000}\right)^{1.4}}$
A	0.8865
B	0.393
$\boldsymbol{\sigma}$	1965 h
MTBF	4432.5 h
$\boldsymbol{L_{10}}$	1955 h

Tableau IV.11: Tableau des résultats obtenu aprés interpolation

Nous traçons ensuite les fonctions de fiabilité ; les graphes ci-dessous résument les résultats :

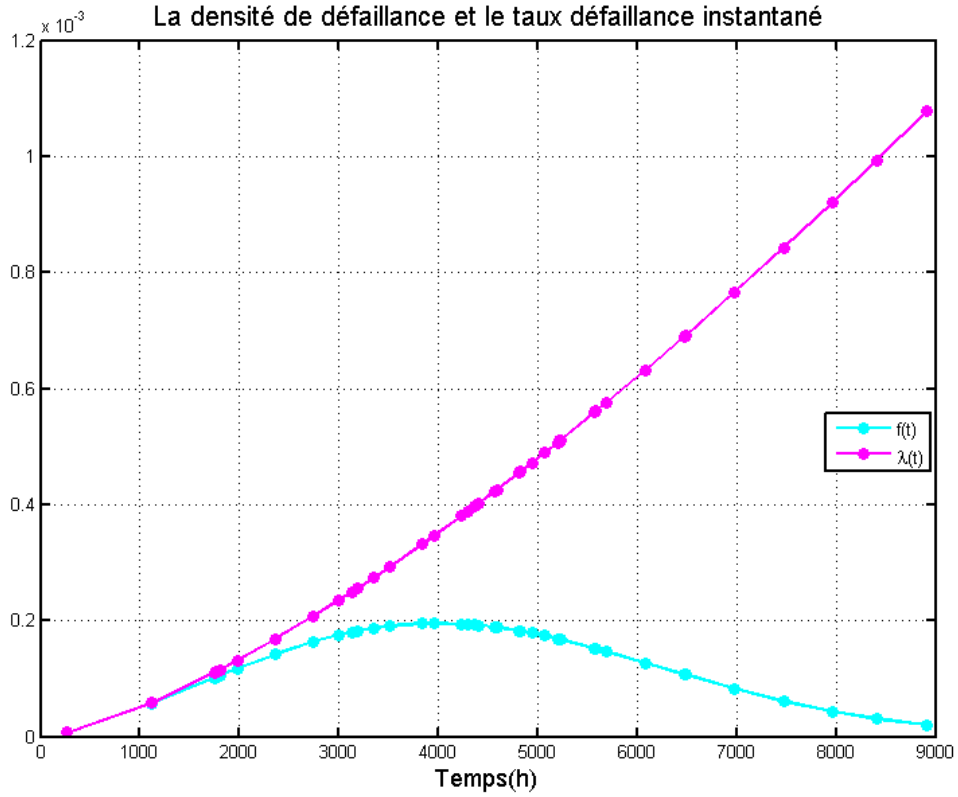


Figure IV.14: La densité de défaillance et le taux défaillance instantané

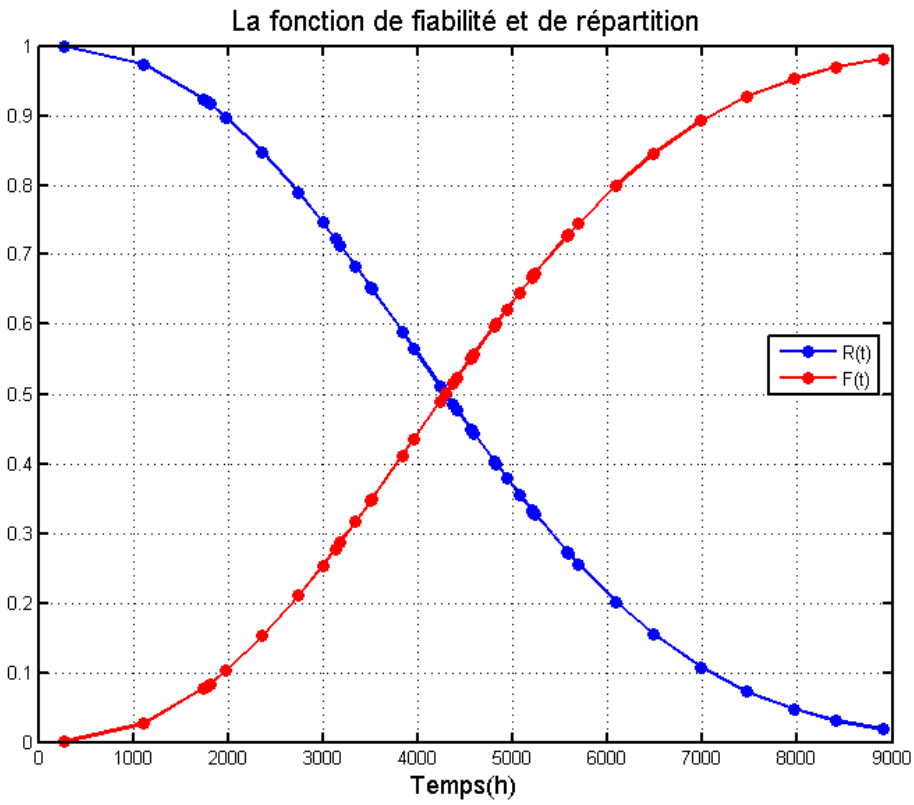


Figure IV.15: La fonction de fiabilité et de répartition

IV.6.7.3 Le test d'adéquation de modèle

N°	TBF	f(ti)*10^{-3}	R(t)	Fth(ti) = 1-R(t)
1	**267**	0.07932597	0.999117108	0.000882891
2	**1110**	0.0197542	0.973368298	0.026631701
3	**1746**	0.0305751	0.923067328	0.076932671
4	**1767**	0.0435088	0.920918811	0.079081188
5	**1810**	0.0568084	0.916423071	0.083576928
6	**1978**	0.0612109	0.897629533	0.102370466
7	**2088**	0.0824441	0.847758862	0.152241137
8	**2744**	0.1015730	0.789057800	0.210942199
9	**2996**	0.1030470	0.746371888	0.253628111
10	**3129**	0.1060545	0.722764108	0.277235891
11	**3183**	0.1069317	0.712993621	0.287006378
12	**2913**	0.1072363	0.683119325	0.316880674
13	**3507**	0.1176243	0.652534142	0.347465857
14	**3510**	0.1273073	0.651962154	0.348037845
15	**3840**	0.1423278	0.588180267	0.411819732
16	**3955**	0.1469376	0.565711936	0.434288063
17	**4223**	0.1520061	0.511429847	0.488570152
18	**3902**	0.1525708	0.498426162	0.501573837

19	**4367**	0.1635045	0.485477282	0.514522717
20	**4410**	0.1676362	0.477200993	0.522799006
21	**4562**	0.1681261	0.448207357	0.551792642
22	**4573**	0.1685318	0.446127204	0.553872795
23	**4591**	0.1690962	0.442729133	0.557270866
24	**4808**	0.1740832	0.402401672	0.597598327
25	**4826**	0.1749029	0.399115249	0.608847501
26	**4940**	0.1786525	0.378536971	0.621463028
27	**5386**	0.1799896	0.355780779	0.644621922
28	**5198**	0.1818612	0.333641459	0.66635854
29	**5212**	0.1823071	0.331278057	0.668721942
30	**5222**	0.1828492	0.329594765	0.670405234
31	**5234**	0.1867817	0.327580187	0.672419812
32	**5575**	0.1885782	0.272926430	0.727073569
33	**5587**	0.1889833	0.271098966	0.728901033
34	**5693**	0.1892254	0.255253718	0.744746281
35	**6706**	0.1906323	0.201467207	0.798532792
36	**6477**	0.1906933	0.155501728	0.844498271
37	**6483**	01.192131	0.154859224	0.845140775

38	**6983**	0.1928007	0.107601310	0.893986898
39	**7365**	0.1937045	0.072680096	0.927319903
40	**7959**	0.1944360	0.047281617	0.952718382
41	**8408**	0.1951000	0.307691077	0.969230892
42	**8908**	0.1955616	0.018335131	0.981664868

Tableau IV.12 : Résultats obtenus avec interpolation

IV.6.8. Interpretation des résultats

Pour le deuxieme cas de notre application nous avons constaté que la methode de regression n'a pas donné de bons résultats où les valeurs manquantes trouvées étaient un peu loin de la realité.

Par contre la technique d'interpolation a présenté toujours son efficacité où les valeurs manquantes trouvées par cette technique étaient prés de la réalité et nous ont permet de calculer et tracé les fonctions de fiabilité.

Le calcul de la fiabilité mécanique présente des limitations et cela pour plusieurs raisons, la notion du taux de défaillance constant n'existe pas, l'absence de modèles satisfaisants pour décrire certains processus complexes de dégradation et de leur interaction, le quasi absence de normalisation et de standardisation internationales et le manque de données.

À travers la revue de littérature, on constate que les méthodes de substitution des valeurs manquantes, sont nombreuses et qu'il n'existe pas de recettes définitives, et on doit agir de cas par cas. Cependant, nous sommes à même d'affirmer que la démarche que nous proposons conduit à des substitutions satisfaisantes tant d'un point de vue théorique qu'empirique. En outre, la méthode statistique de régression et d'interpolation permettent de remplacer les données manquantes à partir d'un échantillon disponible, ainsi une voie intéressante dans le traitement des données manquantes pour la maintenance conditionnelle.

Les principales lois statistiques utilisées pour modéliser la fiabilité sont la loi exponentielle, la loi de Weibull et la loi lognormale. Afin d'estimer les paramètres des lois, la méthode de régression et la méthode du maximum de vraisemblance et l'interpolation sont le plus couramment utilisées.

Dans ce travail, nous avons procédé au remplacement des données incomplètes d'une turbine à gaz et nous avons utilisé le MATLAB et crée des programmes qui nous ont permet d'effectuer une régression linéaire et interpolation linéaire a fin de remplacer ces données perdues.

Deux cas d'études ont été abordés dans le présent travail :

Dans le premier cas d'études et avec un échantillon de de 11 données de TBFI on a réussie à remplacer les données manquantes avec les deux méthodes la régression et l'interpolation. Nous avons pu déterminer les paramètres de la fiabilité par la loi de weibull sur le papier de allan plait pour calculer ensuite la fiabilité et tracer les graphes de ces fonctions.

Dans le deuxième cas d'études et avec un échantillon de taille plus grande et exactement 44 données TBFI dont 7 manquantes, la méthode de régression nous a donné des résults un peu loin de la realité une chose qui nous poussé à travailler seulement avec la technique d'interpolation pour déduire les parametres et et tracer les fonctions de la fiabilité de notre turbine.

Donc après remplacement des données par les deux méthodes on a pu estimer la fiabilité par la loi de Weibull une loi très souple, à trois paramètres qui permettent d'ajuster correctement toutes sortes de résultats. Cette loi a été retenue pour représenter la durée de vie des pièces mécaniques. Son utilisation implique la saisie des résultats des temps de bon fonctionnement entre deux défaillances.

Conclusion Générale

Les méthodes statistiques de régression ou d'interpolation permettent de remplacer les données manquantes à partir d'un échantillon disponible, ainsi une voie intéressante dans le traitement des données manquantes pour la maintenance conditionnelle.

.

[1] E. Vannan « Quality Data: An Improbable Dream. A process for reviewing and improving d ata quality makes for reliable and usable results », EDUCAUSE QUARTERLY, CUMREC conferences, Centre for Education Information, 2001.

[2] Chiang et al, L.H. Chiang, R.J. Pell et M.B. Seasholtz, Exploring process data with the use of robust outlier detection algorithms, Journal of Process Control 13, pp. 437–449, 2003.

[3] S.M. Hu, M.P. Cohen: Evaluation of some popular imputation algorithms, in Section on Survey Research Methods, American Statistical Association, pages 309-313, 2000.

[4] N.K. Malhotra, Analyzing marketing research data with incomplete information on the dependent variable. Journal of Marketing Research 24, 74–84, 1987.

[5] J.O. Kim, J. Curry, The treatment of missing data in multivariate analysis, Sociological Methods and Research 6, 215–241, 1977.

[6] Q.A.W. Raaijmakers, Effectiveness of different missing data treatments in surveys with Likert-type data: Introducing the relative mean substitution approach. Educational and Psychological Measurement, 59, 725-748, 1999.

[7] S.Y. Lee, Y.M. Chiu, Analysis of multivariate polychoric correlation models with incomplete data. British Journal of Mathematical and Statistical Psychology 43, 145–154, 1990.

[8] A. Dempster, N. Laird et D. Rubin: Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, 39(1):1-38, 1977.

[9] Z. Ghahramani et M.I. Jordan : Supervised learning from incomplete data via an EM approach. In J.D. Cowan, G. Tesauro et J. Alspector, éditeurs: Advances in Neural Information Processing Systems 6, pages 120-127. Morgan Kaufman, 1994.

[10] M. Fichman. et J.N.Cummings, Multiple imputation for missing data: Making the most of what you know. OrganizationalResearchMethods, 6: 282-308, 2003.

[11] S. DENIAUD, N. ZERHOUNI., A. EL MOUDNI, F. MOREL, « Sur une modélisation de maintenance préventive à l'Alstom », International Journal of Mechanical production systems Engineering. Nombre 2, V-53 à V-65, 1999.

[12] D. Sherwin, A review of overall models for maintenance management, Journal of Quality in Maintenance Engineering, 6 (3), p. 138–164, 2000.

[13] NF-EN-13306-X-60-319. Terminologie de la maintenance. Norme AFNOR, 2001.

[14] R.E. Barlow et F. Proschan. Mathematical Theory of Reliability. Wiley, New York, 1965.

[15] P. Lyonnet, Ingénierie de la fiabilité, Lavoisier, 2006.

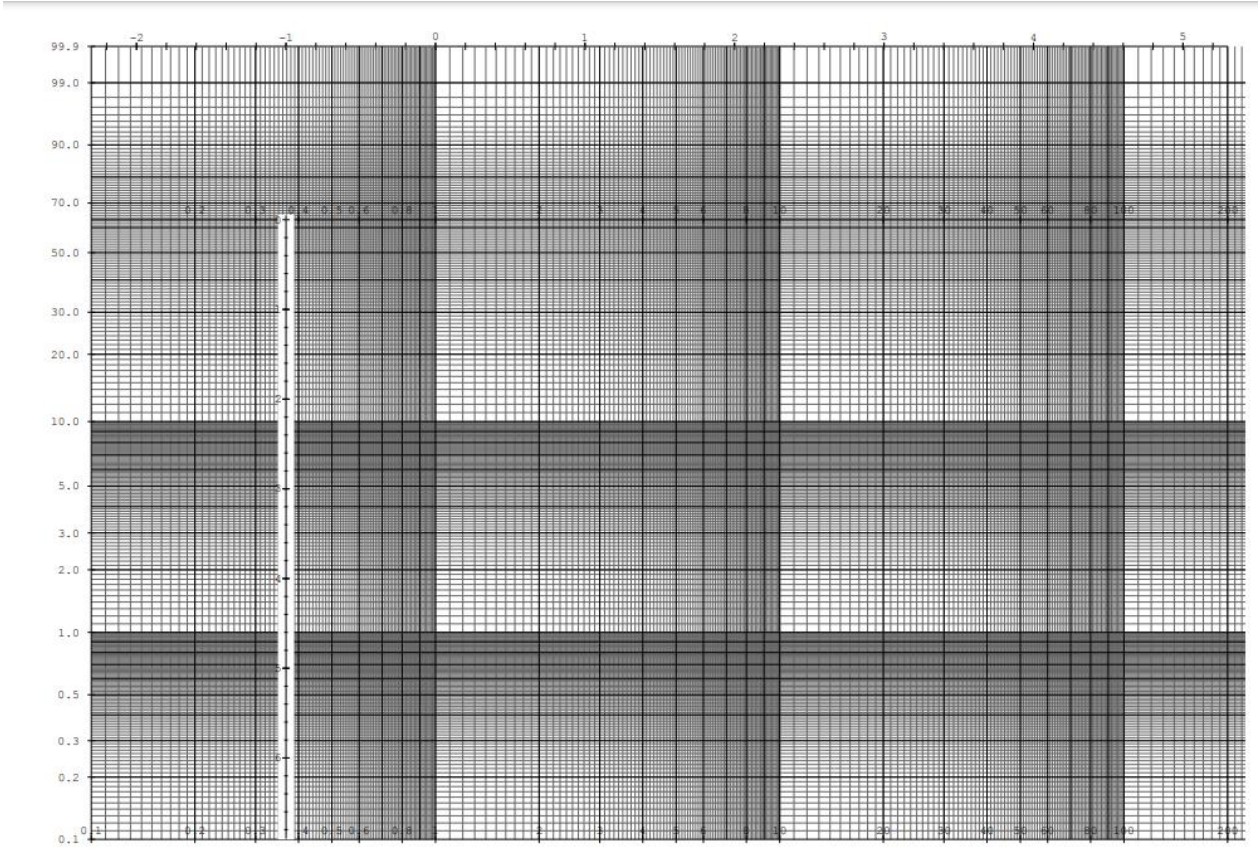
[16] R.P. Nicolai, Maintenance Models for Systems subject to Measurable Deterioration.PhD thesis, Erasmus University Rotterdam, 2008.

[17] D. Sherwin, A review of overall models for maintenance management. Journal of Quality in Maintenance Engineering, 6 (3), p. 138–164, 2000.

[18] Q. William et L.A. Escobar, Statistical Methods for Reliability Data, Wiley series in probability and statistics, 1999.

Annexes

Annexe1 : papier d'allan plait



Annexe 2 :

Distribution de Weibull : valeurs des coefficients A et B en fonction du paramètre de forme :

β	A	B
0,20	120	1901
0,25	24	199
0,30	9,2605	50,08
0,35	5,0291	19,98
0,40	3,3234	10,44
0,45	2,4786	6,46
0,50	2	4,47
0,55	1,7024	3,35
0,60	1,5046	2,65
0,65	1,3663	2,18
0,70	1,2638	1,85
0,75	1,1906	1,61
0,80	1,1330	1,43
0,85	1,0880	1,29
0,90	1,0522	1,17
0,95	1,0234	1,08
1	1	1
1,05	0,9603	0,934
1,10	0,9649	0,878
1,15	0,9517	0,830
1,20	0,9407	0,787
1,25	0,9314	0,750
1,30	0,9236	0,716
1,35	0,9170	0,687
1,40	0,9114	0,660
1,45	0,9067	0,635

β	A	B
1,50	0,9027	0,613
1,55	0,8994	0,593
1,60	0,8966	0,574
1,65	0,8942	0,556
1,70	0,8922	0,540
1,75	0,8906	0,525
1,80	0,8893	0,511
1,85	0,8882	0,498
1,90	0,8874	0,486
1,95	0,8867	0,474
2	0,8862	0,463
2,1	0,8857	0,443
2,2	0,8856	0,425
2,3	0,8859	0,409
2,4	0,8865	0,393
2,5	0,8873	0,380
2,6	0,8882	0,367
2,7	0,8893	0,355
2,8	0,8905	0,344
2,9	0,8917	0,334
3	0,8930	0,325
3,1	0,8943	0,316
3,2	0,8957	0,307
3,3	0,8970	0,299
3,4	0,8984	0,292
3,5	0,8997	0,285
3,6	0,9011	0,278
3,7	0,9025	0,272
3,8	0,9038	0,266
3,9	0,9051	0,260

β	A	B
4	0,9064	0,254
4,1	0,9077	0,249
4,2	0,9089	0,244
4,3	0,9102	0,239
4,4	0,9114	0,235
4,5	0,9126	0,230
4,6	0,9137	0,226
4,7	0,9149	0,222
4,8	0,9160	0,218
4,9	0,9171	0,214
5	0,9182	0,210
5,1	0,9192	0,207
5,2	0,9202	0,203
5,3	0,9213	0,200
5,4	0,9222	0,197
5,5	0,9232	0,194
5,6	0,9241	0,191
5,7	0,9251	0,188
5,8	0,9260	0,185
5,9	0,9269	0,183
6	0,9277	0,180
6,1	0,9286	0,177
6,2	0,9294	0,175
6,3	0,9302	0,172
6,4	0,9310	0,170
6,5	0,9318	0,168
6,6	0,9325	0,166
6,7	0,9333	0,163
6,8	0,9340	0,161
6,9	0,9347	0,160

Annexe 3: Tableau de loi Kolmogorov-Smirnov

N/p	.01	.05	.10	.15	.20
6	0,618	0,521	0,470	0,436	0,410
7	0,577	0,486	0,438	0,405	0,381
8	0,543	0,457	0,411	0,381	0,358
9	0,514	0,432	0,388	0,360	0,339
10	0,490	0,410	0,368	0,342	0,322
11	0,468	0,391	0,352	0,326	0,307
12	0,450	0,375	0,338	0,313	0,295
13	0,433	0,361	0,325	0,302	0,284
14	0,418	0,349	0,314	0,292	0,274
15	0,404	0,338	0,304	0,283	0,266
16	0,392	0,328	0,295	0,274	0,258
17	0,381	0,318	0,286	0,266	0,250
18	0,371	0,309	0,278	0,259	0,244
19	0,363	0,301	0,272	0,252	0,237
20	0,356	0,294	0,264	0,246	0,231
25	0,317	0,265	0,238	0,221	0,208
30	0,290	0,242	0,218	0,203	0,191
35	0,269	0,225	0,202	0,188	0,177
> 35	$\frac{1,63}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,22}{\sqrt{N}}$	$\frac{1,14}{\sqrt{N}}$	$\frac{1,07}{\sqrt{N}}$

Résumé

La méthode de régression et la méthode d'interpolation nous permettent de remplacer les données incomplètes de TBF (temps de bon fonctionnement) pour faire l'étude de fiabilité en façon requise.

La turbine à gaz étudiées suivant le modèle appliqué de la loi de Weibull ou à partir des données complet et juste avec une analyse probabiliste détaillée des indicateurs de fiabilité, alors on peut prévoir un système de recherche permanente des défaillances imprévues possible et de leur criticité et faire appel à des techniques prédictives de test et d'inspection qui permet de vérifier l'état de la turbine.

Les mots clés

Méthode de régression, Méthode d'interpolation, fiabilité, temps de bon fonctionnement, turbine à gaz, loi de Weibull

Abstract

The regression methode and the interpolation methode allow us to replace incomplete TBF (uptime) data to do the reliability study as required.

The gas turbine studied according to the model applied to the law of Weibull or from complete and fair data with a detailed probabilistic analysis of reliability indicators, and then we can provide a system of permanent search for possible unforeseen failures and their criticality and use predictive testing and inspection techniques to verify the condition of the turbine.

Keywords

Regression method, Interpolation method, reliability, uptime, gas turbine, Weibull's law

ملخص

تسمح لنا طريقة الانحدار وطريقة الاستيفاء باستبدال بيانات TBF (الجهوزية) غير المكتملة لإجراء دراسة الموثوقية كما هو مطلوب.

التوربينات الغازية التي تمت دراستها وفقًا للنموذج المطبق على قانون Weibull أو من بيانات كاملة وعادلة مع تحليل احتمالي مفصل لمؤشرات الموثوقية، ثم يمكننا توفير نظام بحث دائم عن الأعطال المحتملة غير المتوقعة وأهميتها واستخدام الاختبار التنبئي وتقنيات الفحص للتحقق من حالة التوربينات.

الكلمات الدالة

طريقة الانحدار، طريقة الاستيفاء، الموثوقية، الجهوزية، التوربينات الغازية، قانون Weibull