



HAL
open science

Fiabilité des systèmes embarqués

Sorin Voiculescu

► **To cite this version:**

Sorin Voiculescu. Fiabilité des systèmes embarqués. Sciences de l'ingénieur [physics]. Université d'Angers, 2009. Français. NNT: . tel-00468219

HAL Id: tel-00468219

<https://theses.hal.science/tel-00468219>

Submitted on 30 Mar 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FIABILITE DES SYSTEMES EMBARQUEES
THESE DE DOCTORAT

Spécialité : Sciences pour l'ingénieur

**ECOLE DOCTORALE Sciences et Techniques de l'Information et de
Mathématiques**

Présentée et soutenue publiquement
Le Juin-Juillet 2009
A l'Institut des Sciences et Techniques de l'Ingénieur d'Angers

Par Sorin VOICULESCU

Devant le jury ci-dessous :

Ioan BACIVAROV	Examineur	Professeur à l'Université Polytechnique de Bucarest
Mihaela BARREAU	Examineur	Maître de Conférences à l'Université d'Angers
Zorha CHERFI	Rapporteur	Professeur à l'UTC
Pierre DERSIN	Examineur	Ingénieur Docteur à ALSTOM
Léo GERVILLE-REACHE	Rapporteur	Maître de Conférences HdR à l'Université de Bordeaux 2
Fabrice GUERIN	Examineur	Professeur à l'Université d'Angers
François PERES	Examineur	Maître de Conférences HdR à l'ENIT

Directeur de thèse : **Fabrice GUERIN**
Co-encadrante : **Mihaela BARREAU**

Laboratoire : **Laboratoire en Sécurité de fonctionnement, Qualité et Organisation**
62, avenue Notre Dame du Lac
49000 ANGERS

Résumé

Les travaux présentés dans cette thèse abordent la problématique de l'estimation de la fiabilité des systèmes qui est devenue un enjeu majeur pour les entreprises et ce dès les premières phases de développement. En effet, plus tôt les caractéristiques d'un produit ou d'un système sont appréhendées dans son cycle de vie et moins les risques (financiers ou liés à la sécurité des installations), dus à la non réalisation des performances attendues sont élevés. Dans un contexte exigeant des systèmes de plus en plus fiables et sûrs, et de durées de garanties croissantes, il est impératif de vérifier le plus rapidement possible que les performances des systèmes soient conformes au cahier des charges. La démarche idéale pour identifier la fiabilité d'un produit ou d'un système avant même sa fabrication en série, est de procéder à des séries d'essais sur des prototypes, quand ils existent. Cette démarche nécessite un investissement trop important en temps et en nombre de prototypes, car ces derniers étant de plus en plus fiables, l'observation de défaillances est de moins en moins probable. Ainsi, pour réduire ces durées, on peut réaliser des essais accélérés. Pour cela, le produit est soumis à des sollicitations d'utilisation ou d'environnement amplifiées, afin d'accélérer les mécanismes d'endommagement (les mécanismes de défaillances provoqués doivent être représentatifs des conditions normales d'emploi) et de réduire la durée nécessaire pour l'estimation de certaines caractéristiques comportementales du produit dans les conditions normales d'emploi. Pour ce faire, nous avons proposé d'utiliser les modèles SVA (Standard de Vie Accélérée) paramétrique et semi-paramétrique pour analyser les données d'essais. Afin d'obtenir des estimations de bonne qualité, nous avons introduit la démarche Bayésienne permettant d'intégrer toute la connaissance disponible sur la fiabilité du produit étudié. Ainsi, nous avons proposé des méthodes d'estimation, une technique de construction de la distribution pseudo conjuguée a priori et une méthodologie d'analyse de compatibilité entre la connaissance a priori et les résultats d'essai. Pour finir, nous avons développé des méthodes d'estimation de la fiabilité intégrant les variabilités d'usage et d'environnement.

I.4.1.3	Comparaison des modèles paramétrique et non paramétrique	36
I.4.2	Influence de la censure	37
I.4.3	Analyse globale	39
I.5	Les essais accélérés	40
I.5.1	Le plan d'essais	41
I.6	Le modèle de vie accélérée (MVA)	44
I.6.1	Principe	44
I.6.2	Lois d'accélération courantes	48
I.6.2.1	Loi d'Arrhenius	48
I.6.2.2	Loi de Peck	48
I.6.2.3	Loi puissance inverse	48
I.6.2.4	Loi d'Eyring	49
I.6.3	La construction d'un modèle paramétrique	49
I.6.4	La construction d'un modèle semi-paramétrique	52
I.7	Application numérique	55
I.8	Conclusion	59
II	Estimation Bayésienne des modèles SVA	61
II.1	Introduction	63
II.2	L'inférence Bayésienne dans le cadre de MSVA	66
II.2.1	Application de l'inférence Bayésienne au modèle SVA paramétrique	67
II.2.1.1	Formulation du problème	67
II.2.1.2	Estimation bayésienne des paramètres de modèles	68
II.2.1.2.1	méthode classique	68
II.2.1.2.2	méthode MAP (Maximum of A Priori	69
II.2.2	L'inférence Bayésienne du modèle SVA semi-paramétrique	73
II.3	L'information a priori	79
II.3.1	Sources d'information	79
II.3.1.1	Recueils de données	80
II.3.1.2	L'avis d'experts	81
II.3.2	Le choix de la loi a priori	82
II.3.2.1	Lois conjuguées	83
II.3.2.2	Détermination de la loi a priori pseudo conjuguée par simulation	84

II.3.3	Construction de la loi a priori	86
II.4	Application numérique	89
II.4.1	A priori cohérent et peu informatif	89
II.4.2	A priori cohérent et très informatif	91
II.4.3	A priori très informatif et incohérent	93
II.5	Compatibilité de l'a priori avec les données	97
II.5.1	Le principe de calcul d'un facteur de compatibilité normalisé	100
II.5.2	L'utilisation du facteur de compatibilité comme facteur de pondé- ration de l'a priori	102
II.5.3	Application au modèle SVA	106
II.6	Exemple numérique	109
II.6.1	Résultats	109
II.6.1.1	A priori peu informatif et cohérent	109
II.6.1.2	A priori très informatif et cohérent	109
II.6.1.3	A priori incohérent et informatif	110
II.6.1.3.1	Les résultats de la méthode Bayésienne classique	112
II.6.1.3.2	Les résultats de la méthode Bayésienne pondérée	112
II.6.2	Synthèse	112
II.7	Conclusion	114
III	Estimation de la fiabilité en environnement aléatoire	115
III.1	Introduction	116
III.2	Définition des profils de mission	120
III.3	La variabilité d'usage	122
III.3.1	Estimation directe de la fonction de fiabilité	122
III.3.2	Exemple d'illustration	123
III.3.3	Estimation de la fiabilité par la méthode contrainte-résistance	125
III.4	Variabilité temporelle des sollicitations	127
III.4.1	introduction	127
III.4.2	Stress défini par une variable aléatoire indépendante du temps	128
III.4.2.1	Principe de l'analyse	128
III.4.2.2	Caractérisation de la variable aléatoire	130
III.4.2.3	Histogramme ou temps de maintien dans une classe d'am- plitudes	134
III.4.2.4	Comptage du nombre de dépassements d'un niveau donné	134

III.4.3 Estimation de la fiabilité	136
III.4.4 Séries temporelles	136
III.4.5 Discrétisation du profil de stress	139
III.4.5.1 Principe de la discrétisation	139
III.4.5.2 Modèle d'endommagement cumulé (MEC)	139
III.5 Exemple d'application	142
III.6 Conclusion	147
IV Conclusions et perspectives	149
Bibliographie	153

Introduction générale

Quand on veut savoir si un système est fiable, la première chose à faire, quand on peut, c'est de l'essayer au préalable. On veut voir que ça marche. Cet essai préliminaire est même parfois la seule raison de croire un peu à la fiabilité du système. Quelle est sa valeur prédictive ? Est-ce que cet essai prouve vraiment quelque chose quant à la fiabilité du système ?

Une façon ironique de poser le problème consiste à appliquer la règle des cinq minutes. On la cite souvent sur l'exemple d'une cathédrale : une fois que l'édifice est terminé, on retire tous les échafaudages puis on attend cinq minutes. Si dans ce laps de temps la cathédrale ne s'est pas effondrée, on estime alors qu'elle pourra tenir pendant des siècles. On pourrait aussi appliquer la règle de la première fois : si un système fonctionne bien la première fois, c'est qu'il fonctionnera bien toujours. Il est évident que ces règles ne fournissent pas toujours de bonnes prédictions : certains édifices s'effondrent quelque temps après avoir été construits, des machines, qui fonctionnaient bien, cessent de fonctionner. Quel sens alors faut-il donner à la règle de l'essai préalable ? Doit-on conclure qu'un essai ne sert à rien ? Bien sûr que non. Pourquoi ?

Quand des vies humaines sont en jeu on pourrait croire que la prise de risques est le problème majeur des études de fiabilité. Dans certains cas particuliers oui mais pas en général. Ce serait nous demander beaucoup trop. Même quand des vies humaines sont en jeu, on ne peut que rarement faire des prédictions infaillibles. Tout le monde le sait bien. On prend un risque à chaque fois qu'on fait un pas sur un trottoir. Il serait absurde d'exiger l'infaillibilité des prédictions. Ce serait nous condamner à ne plus rien faire ou presque. Et ne plus rien faire c'est la mort assurée. Qui ne risque rien n'a rien. Vivre c'est toujours prendre des risques.

Le problème général des études de fiabilité ce n'est pas l'infaillibilité, c'est de réduire les risques. On sait qu'il y a des risques mais on veut faire les bons choix : parmi toutes les solutions réalisables quelles sont les plus risquées, les plus dangereuses, les moins fiables ?

Mais il ne faut pas en conclure que les études de fiabilité ont toujours un caractère probabiliste, du moins pas au sens où elles passeraient par un calcul de probabilités. Le

calcul des probabilités n'est pas la seule méthode de réduction des risques. On est parfois incapable d'évaluer les probabilités de ces risques. Le problème général c'est de se servir de nos facultés prédictives pour réduire les risques.

Les essais sont comme un crible. Ils doivent faire la part du bon et du mauvais. Il s'agit d'exclure le dangereux et l'inefficace et de reconnaître la qualité. On ne peut presque jamais demander un crible parfait, mais on peut souvent trouver d'assez bons cribles qui permettent de réduire les risques.

Dans un contexte international extrêmement concurrentiel, les entreprises doivent maîtriser les différents outils qui leur permettent de rester compétitives et doivent s'engager dans des actions d'amélioration à tous les niveaux. De plus en plus, les entreprises développent ou obtiennent des systèmes sophistiqués sur lesquels ils doivent placer une grande confiance. Dans des circonstances différentes, la focalisation sera mise sur les propriétés différentes de tels services, par exemple, la continuité, l'exécution, la réponse en temps réel, la capacité d'éviter les événements catastrophiques. Dans cette situation, l'une des stratégies importantes exigées aux industriels est la qualité de produits, l'orientation vers la sûreté de fonctionnement. La sûreté de fonctionnement fait partie des enjeux majeurs de ces dernières années et des années à venir. Plus personne aujourd'hui n'accepterait d'utiliser un produit, un système dont tous les aspects liés à son usage ne soient optimisés sous les aspects de la qualité.

On parle de la sûreté de fonctionnement d'un système comme la caractéristique de ce système qui permet à l'utilisateur de placer en lui une confiance justifiée. Lors de la construction d'un système, le constructeur spécifie non seulement les fonctionnalités, mais aussi les objectifs à atteindre en termes de sûreté de fonctionnement. La spécification des objectifs de sûreté de fonctionnement est accompagnée d'une procédure de validation pour vérifier que les objectifs ont été atteints. Des méthodes et des outils de développement spécifiques sont appliqués pour cette démarche.

La règle de l'essai préalable n'est pas sans valeur mais un seul essai n'est en général pas suffisant, pour trois raisons principales :

- Les conditions de l'essai ne sont pas représentatives de toutes les conditions d'utilisation du système. S'il est sensible à l'humidité il peut bien fonctionner le jour de l'essai s'il fait beau mais ne pas fonctionner les jours de pluie, par exemple.
- L'état présent du système n'est représentatif de son état futur ; il peut se transformer et se dégrader.
- Lorsque le système est produit en plusieurs exemplaires, le système essayé peut n'être pas représentatif des autres : dispersion des conditions de fabrication, essai sur un prototype, ...

Faire des études de fiabilité, c'est essayer de faire mieux que la règle de la première fois, c'est donc tenir compte de toutes les utilisations normales du système, de son vieillissement et éventuellement des différences entre les systèmes d'une même série.

Les conditions d'utilisation peuvent être très diverses et les essais de vieillissement sont souvent gourmands en temps. Un des principaux problèmes pour la conception de ces essais, c'est de les accélérer. On voudrait faire des essais qui soient significatifs de l'ensemble des utilisations normales du système, tous ses usages possibles, tout ce qui peut lui arriver, toutes ses façons de réagir, de se transformer et de se dégrader. Mais en pratique il faut se contenter d'un nombre limité d'essais. Il faut concevoir un plan d'essais qui réduise au mieux les risques compte-tenu des moyens disponibles. Comment les résultats de certaines expériences, les utilisations futures du système, peuvent-ils être prédits sur la base des résultats d'autres expériences, les essais? Le rôle des théories prédictives c'est justement cela, prédire des observations futures sur la base d'observations passées. La règle de la première fois est une théorie prédictive simpliste. Outre les moyens matériels et humains et le temps (les délais), les théories prédictives font partie des moyens pour concevoir un plan d'essais. Elles permettent d'augmenter la signification des essais.

Cette thèse va se focaliser sur l'étude de l'estimation de la fiabilité des systèmes par des approches expérimentales et opérationnelles. Une partie de la thèse s'est effectuée dans le cadre d'un projet européen EUREKA/PIDEA/ACLIMAT. Les partenaires de ce projet étaient : Thalès Avionics (Bordeaux), Celistica (Valence en Espagne), Cetecom (Malaga en Espagne), Laboratoire A2X (UMR CNRS 5465, Université Bordeaux 1) et le laboratoire LASQUO (UPRES EA, Université d'Angers). Ainsi, la plupart des exemples présentés au cours de cette thèse est issue d'un cas réel correspondant à une carte électronique embarquée dans un avion.

Le document est structuré en trois chapitres.

Le premier chapitre présente les définitions et les modèles utilisés en fiabilité. Ensuite, les essais accélérés sont étudiés avec l'introduction du modèle Standard de Vie Accélérée (SVA) permettant d'analyser les résultats d'essais. En particulier, nous étudierons deux modèles : paramétrique et semi-paramétrique. Quelques exemples illustreront l'application de ces modèles dans des contextes de faibles tailles d'échantillon et de censure.

Dans le deuxième chapitre, nous abordons l'application de la démarche bayésienne à l'estimation des paramètres des modèles SVA paramétrique et semi-paramétrique. L'objectif est d'obtenir des estimations, avec des intervalles de confiance raisonnables, malgré les faibles tailles d'échantillon et de censure qui sont imposées par le contexte industriel (réduction des coûts et délais). Ainsi, il est possible de capitaliser toute la connaissance disponible sur le produit : les résultats d'essai (vraisemblance) et les informations a priori

sur le niveau de fiabilité du produit. L'efficacité de la démarche passe par la capacité de construire des lois a priori qui synthétisent toute la connaissance disponible. Aussi, nous proposerons une méthodologie permettant d'élaborer les distributions a priori en abordant deux points :

- le choix des distributions a priori qui permet de maximiser le poids des résultats d'essais dans l'estimation. Pour ce faire, nous proposons de rechercher les lois conjuguées bayésiennes par simulation car il n'est pas possible, du fait de la complexité des modèles, de les obtenir analytiquement.
- l'estimation des paramètres de lois par la méthodes des moments.

Ensuite, nous aborderons la problématique de la compatibilité entre les premiers résultats d'essai et la connaissance a priori injectée dans l'inférence bayésienne. Ainsi, s'il y a discordance entre les distributions associées à la vraisemblance et à l'a priori, cela peut conduire à prendre des décisions erronées quant à la fiabilité du produit. Aussi, nous proposerons une méthodologie, permettant de comparer ces deux distributions afin de vérifier la cohérence des résultats d'essai avec la priori et de pondérer celle-ci en fonction du degré de compatibilité. Ce chapitre sera illustré par plusieurs exemples permettant d'indiquer la mise en oeuvre des différentes méthodologies et de vérifier leurs efficacités.

Dans le troisième chapitre, nous aborderons la problématique de l'estimation prévisionnelle de la fiabilité. En effet, lorsque l'on dispose des résultats d'essai, nous cherchons à estimer la fonction de fiabilité dans les conditions normales d'utilisation. Derrière le terme 'conditions normales d'utilisation', nous pouvons introduire plusieurs facteurs. Ainsi, les conditions peuvent fluctuer entre les utilisateurs, les différentes phases de vie du produit (repos, fonctionnement, entretien, ...), les applications (le même matériel équipant une voiture, un avion, un train, ...), ... Aussi, nous présenterons plusieurs techniques permettant d'introduire cette variabilité dans l'estimation de la fonction de fiabilité. Ces techniques utilisent les modèles SVA développés dans le chapitre II. Plusieurs exemples illustreront les différentes approches.

Enfin, des conclusions sur le travail réalisé seront présentées et des perspectives seront proposées.

I

Estimation de la fiabilité

Sommaire

I.1	Sûreté de Fonctionnement	7
I.1.1	Le concept FMDS	7
I.1.2	Grandeurs associées	9
I.2	La fiabilité	11
I.2.1	La fiabilité tout au long du cycle de vie	12
I.2.2	Mesures des performances	14
I.3	L'analyse de données de Fiabilité	19
I.3.1	L'approche non paramétrique	20
I.3.2	L'approche paramétrique	21
I.3.2.1	Le modèle standard de durée de vie	22
I.3.2.1.1	La lois Weibull	23
I.3.2.1.2	La loi exponentielle	23
I.3.2.2	Les estimations de paramètres	24
I.3.2.2.1	Les estimations par la méthode du maximum de vraisemblance (MV)	25
I.3.2.2.2	La méthode bayésienne non informative	27
I.3.2.3	Estimation des caractéristiques de fiabilité	29
I.4	Analyse des modèles de fiabilité et méthodes d'estimation par simulation de Monte Carlo	30
I.4.1	Influence de la taille d'échantillon	30
I.4.1.1	Les modèles non paramétriques	30
I.4.1.2	Les modèles paramétriques	31
I.4.1.3	Comparaison des modèles paramétrique et non paramétrique	36

I.4.2	Influence de la censure	37
I.4.3	Analyse globale	39
I.5	Les essais accélérés	40
I.5.1	Le plan d'essais	41
I.6	Le modèle de vie accélérée (MVA)	44
I.6.1	Principe	44
I.6.2	Lois d'accélération courantes	48
	I.6.2.1 Loi d'Arrhenius	48
	I.6.2.2 Loi de Peck	48
	I.6.2.3 Loi puissance inverse	48
	I.6.2.4 Loi d'Eyring	49
I.6.3	La construction d'un modèle paramétrique	49
I.6.4	La construction d'un modèle semi-paramétrique	52
I.7	Application numérique	55
I.8	Conclusion	59

I.1 Sûreté de Fonctionnement

La **Sûreté de Fonctionnement** (*SdF*) est l'aptitude d'une entité à satisfaire une ou plusieurs fonctions requises dans des conditions données, la propriété qui permet aux utilisateurs du système de placer une confiance justifiée dans le service qu'il leur délivre, la capacité de maintenir le niveau de performance dans le temps.

La sûreté de fonctionnement désigne à la fois un ensemble de moyens et un ensemble de résultats produits par ces moyens [Mortureux, 2001] :

- les méthodes et outils pour caractériser et maîtriser les effets des aléas, des pannes, des erreurs [Villemeur, 1988], [Pagès et Gondran, 1980], [O'Connor, 2002] ;
- les caractéristiques des systèmes pour exprimer la conformité de leurs comportements et leurs actions à long terme [Pagès et Gondran, 1980], [Procaccia et Morilhat, 1996], [Pham, 2003].

Au sens large, la SdF est considérée comme la science des défaillances et des pannes. Différents auteurs définissent la SdF comme :

- la science des défaillances [Villemeur, 1988] ;
- la confiance justifiée dans le service qu'il délivre [Laprie et al., 1995] ;
- le maintien de la qualité dans le temps [Mortureux, 2001] ;
- l'ensemble des propriétés qui décrivent la disponibilité et les facteurs qui conditionnent la fiabilité, la maintenabilité et la logistique de maintenance [ISO 9000].

Plus personne aujourd'hui n'accepterait d'utiliser un produit, un système dont tous les aspects liés à son usage ne soient pas optimisés sous ses 4 aspects de la Sûreté de Fonctionnement [Fournier, 1993] :

1. La **F**iabilité
2. La **D**isponibilité
3. La **M**aintenabilité
4. La **S**écurité

Dans les paragraphes suivants, nous allons définir des quatre attributs FMDS de la SdF et les métriques associées.

I.1.1 Le concept FMDS

Le concept **FMDS** est apparu aux États-Unis dans les années 1960 dans l'industrie du spatial et de l'armement.

C'est un ensemble d'outils et de méthodes qui vont permettre sur toutes les phases de la vie d'un produit de s'assurer qu'il va réaliser la (les) mission(s) pour laquelle (lesquelles)

il a été conçu dans des conditions de Fiabilité, de Maintenabilité, de Disponibilité et de Sécurité définies. Ainsi, la SdF doit être prise en compte tout au long du cycle de développement du produit. Les quatre attributs de la FMDS sont définis de la manière suivante :

La fiabilité

La fiabilité (reliability en anglais) est l'aptitude d'un dispositif à accomplir une fonction requise dans des conditions données pour une période de temps donnée [Chapouille, 1972]. Elle se caractérise par la probabilité $R(t)$ que l'entité E accomplisse ces fonctions, dans les conditions données pendant l'intervalle de temps $[0, t]$, sachant que l'entité n'est pas en panne à l'instant 0.

$$R(t) = P[E \text{ non défaillante sur } [0, t]]$$

La maintenabilité

La maintenabilité (maintainability en anglais) est l'aptitude d'une entité à être maintenue ou rétablie dans un état dans lequel elle peut accomplir une fonction requise, lorsque la maintenance est réalisée dans des conditions données avec des procédures et des moyens prescrits. Elle se caractérise par la probabilité $M(t)$ que l'entité E soit en état, à l'instant t , d'accomplir ces fonctions, sachant que l'entité était en panne à l'instant 0.

$$M(t) = P[E \text{ est réparée sur } [0, t]]$$

La disponibilité

La disponibilité (availability en anglais) est l'aptitude d'une entité à être en état d'accomplir les fonctions requises dans les conditions données et à un instant donné. Elle se caractérise par la probabilité $A(t)$ que l'entité E soit en état, à l'instant t , d'accomplir les fonctions requises dans des conditions données :

$$A(t) = P[E \text{ non défaillante à l'instant } t]$$

La sécurité

La sécurité (safety en anglais) est l'aptitude d'un produit à respecter, pendant toutes les phases de sa vie, un niveau acceptable de risques d'accidents susceptibles d'occasionner une agression du personnel, une dégradation majeure du produit ou de son environnement.

$$S(t) = P[E \text{ évite des événements critiques ou catastrophiques sur } [0, t]]$$

Nous admettrons par la suite que le temps est la variable principale dont dépend la fiabilité. Pour certains appareils, il peut être plus approprié de considérer une autre variable : le nombre de cycles d'ouverture-fermeture pour un relais, le nombre de tours pour un moteur, le nombre de kilomètres pour une voiture, etc.

La fiabilité est une composante essentielle de la sûreté de fonctionnement. Elle participe à la disponibilité d'un équipement. Afin d'envisager une étude de sûreté de fonctionnement exhaustive, il sera nécessaire de réaliser des études complémentaires dans les domaines de la maintenabilité, de la sécurité et des calculs probabilistes de la disponibilité.

I.1.2 Grandeurs associées

Des métriques, basées sur l'estimation de moyennes de temps, peuvent être utilisées pour caractériser la sûreté de fonctionnement d'un système [Villemeur, 1988] :

- **MTTF** : La durée moyenne de fonctionnement d'une entité avant la première défaillance (anglais Mean Time To Failure)

$$MTTF = \frac{\text{temps cumulés de fonctionnement jusqu'à la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre de 1}^{\text{ères}} \text{ défaillances}} \quad (\text{I.1})$$

- **MUT** : Durée moyenne de fonctionnement après réparation (anglais Mean Up Time)

$$MUT = \frac{\text{temps cumulés de fonctionnement après la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre d'intervalles entre 2 défaillances successives}} \quad (\text{I.2})$$

- **MDT** : Durée moyenne d'indisponibilité après défaillance (anglais Mean Down Time)

$$MDT = \frac{\text{temps cumulés d'arrêts}}{\text{nombre total de défaillances après la 1}^{\text{ère}}} \quad (\text{I.3})$$

- **MTBF** : Durée moyenne entre deux défaillances (anglais Mean Time Between Failure)

$$MTBF = \frac{\text{temps cumulés d'exploitation après la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre total de défaillances après la 1}^{\text{ère}}} \quad (\text{I.4})$$

- **MTTR** : Durée moyenne de réparation (anglais Mean Time To Repair)

Ces durées sont représentées dans la figure (I.1.2).

La disponibilité asymptotique est donnée par :

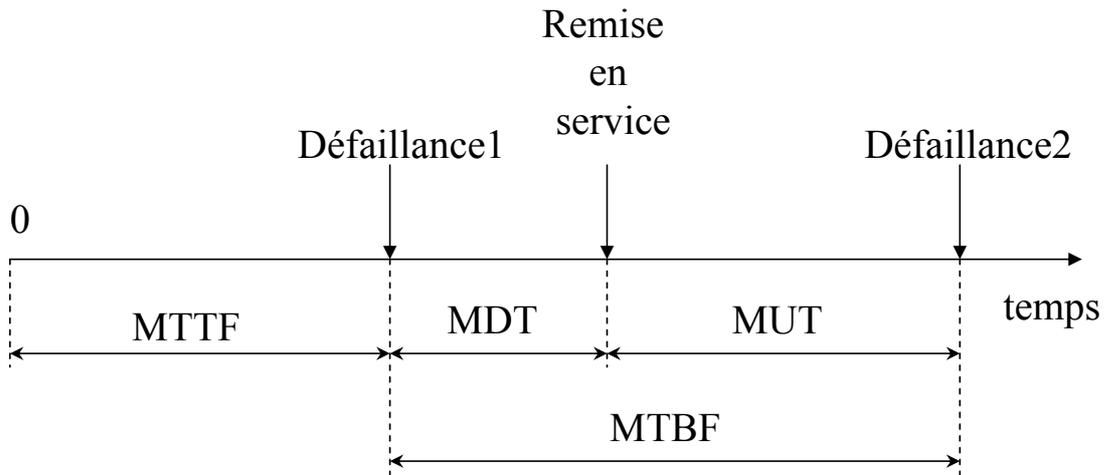


FIG. I.1 – Grandeurs associées à la fiabilité

$$\text{Disponibilité} = \frac{\text{MUT}}{\text{MTBF}} \quad (\text{I.5})$$

Dans la suite de cette thèse, nous allons nous focaliser tout particulièrement sur la fiabilité qui est un élément essentiel de la sûreté de fonctionnement.

I.2 La fiabilité

L'analyse de la fiabilité constitue une phase indispensable dans toute l'étude de la sûreté de fonctionnement. A l'origine, la fiabilité concernait les systèmes à haute technologie et à fort potentiel de risque (nucléaire, aéronautique, spatial, ...). Aujourd'hui, la fiabilité est devenue un paramètre clef de la qualité et de l'aide à la décision dans l'étude de la plupart des composants, produits et processus "grand public" : transport, énergie, bâtiments

De nombreux industriels travaillent à l'évaluation et l'amélioration de la fiabilité de leurs produits au cours de leur cycle de développement, de la conception à la mise en service (conception, fabrication et exploitation) afin de développer leurs connaissances sur le rapport Coût/Fiabilité et maîtriser les sources de défaillance.

La dégradation de la fiabilité d'un produit est directement liée aux mécanismes de dégradation de produit et/ou de ses composants. Parmi les mécanismes de défaillance nous pouvons citer :

- la fatigue
- le fluage
- la fissuration
- l'usure
- la corrosion / l'oxydation
- l'électromigration, etc

Les étapes nécessaires à la réalisation d'une étude complète de fiabilité pour obtenir soit des mesures de performance soit des grandeurs associées au produit sont :

1. Les objectifs de l'étude ;
2. L'analyse fonctionnelle ;
3. La conception (bureau d'études) ;
4. L'identification des risques potentiels ;
5. La modélisation du système ;
6. Les essais de validation sur prototypes, présérie ;
7. La validation par l'analyse de retour d'expérience (REX).

On observe que jusqu'à l'étape 5 (incluse) on peut parler d'une approche prévisionnelle, avant la réalisation du produit final ou de prototypes. C'est une approche très utile apte à fournir des informations pertinentes sur les caractéristiques de la fiabilité du produit en cours du développement. Cette approche permet rapidement et avec de ressources limitées de décider quelles sont les actions à mener pour améliorer l'équipement. Néanmoins, il est

rarement possible d'intégrer les interactions entre les composants et les différents modes de défaillance qui peuvent pourtant dégrader significativement la fiabilité.

Aussi, lors de l'étape suivante, les essais de validation sur des prototypes ou des produits de présérie permettent d'étudier le comportement et les performances en simulant les conditions d'utilisation (stress, sollicitations, environnement, ...) : on parle alors d'une approche expérimentale.

La dernière étape (étape 7) permet de mesurer la fiabilité opérationnelle du produit en traitant les données issues de l'exploitation. Cette étape permet de confirmer les différentes hypothèses prises (mécanismes de défaillance influents, lois de fiabilité, ...) et de capitaliser cette information à travers des recueils de données ou pour servir comme information a priori dans le cadre de démarches Bayésiennes sur de futures versions du produit.

Dans le cadre de la thèse, nous nous sommes intéressés à la fiabilité expérimentale. Dans les paragraphes suivants, nous présentons les différentes techniques d'essais de fiabilité et les statistiques qui peuvent être estimées à partir des résultats d'essais.

I.2.1 La fiabilité tout au long du cycle de vie

La fiabilisation d'un produit au cours du cycle de vie de développement fait appel à de nombreuses phases dans lesquelles on utilise divers méthodes et outils (voir Figure I.2).

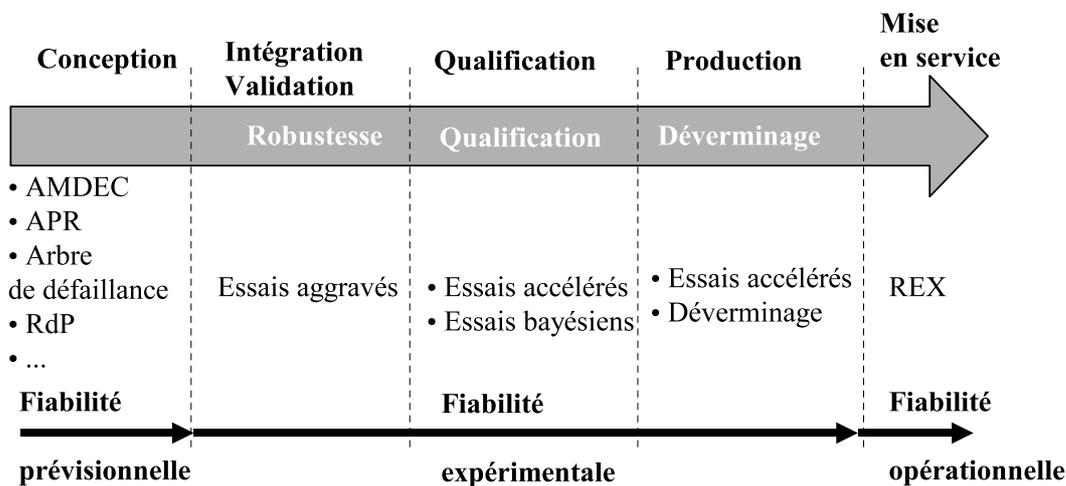


FIG. I.2 – Position des essais dans le cycle de développement d'un produit

L'ensemble des essais de fiabilité (robustesse, estimation et déverminage) contribue largement à la croissance de la fiabilité du produit au cours de son développement et de sa production (voir Figure I.3).

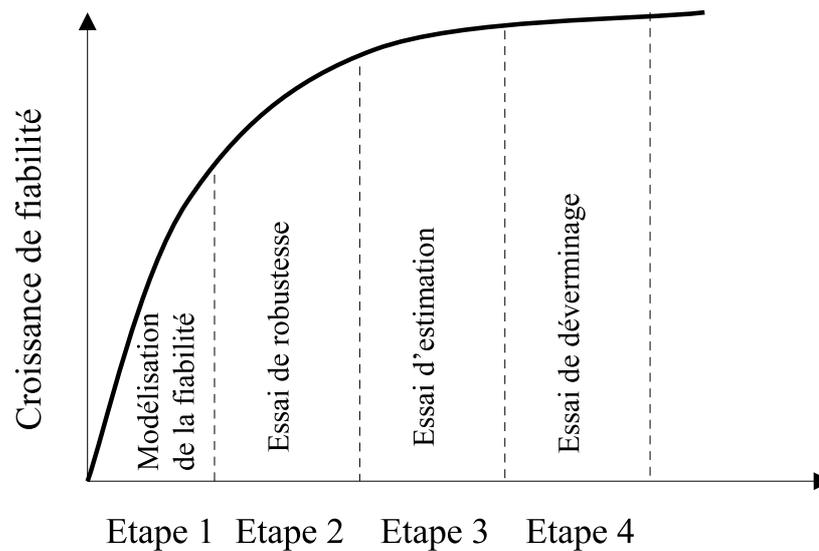


FIG. I.3 – Croissance de la fiabilité au cours du développement d'un produit

Les dernières techniques d'essais développées consistent à accélérer cette croissance de fiabilité par l'utilisation d'**essais sévérés** [Crowe et Feinberg, 2001]. Ainsi, on peut citer :

- essais de robustesse : les essais HALT (Highly **Accelerated** Life Test)
- essais d'estimation : les essais accélérés ALT (**Accelerated** Life Test)
- essais de déverminage : les essais HASS (Highly **Accelerated** Stress Screen)

La cohérence entre les différentes études de fiabilité doit être gardée d'une manière constante pendant les cycles de vie d'un produit comme une partie intégrante de l'ensemble d'outils SdF.

Les estimations expérimentale et opérationnelle de la fiabilité sont basées principalement sur l'analyse et l'interprétation des instants de défaillance. La grande différence entre les deux est les moyens utilisés pour obtenir les temps de défaillance. Par contre, les méthodes et les outils d'analyse sont communs. Lorsque la fiabilité expérimentale utilise des instants de défaillance obtenus à partir d'essai sur des prototypes, la fiabilité opérationnelle utilise les instants de défaillance obtenus à partir de l'exploitation du produit.

Nous allons nous intéresser aux essais d'évaluation qui vont permettre d'évaluer des mesures de performance qui sont généralement des mesures statistiques qu'on va présenter dans le paragraphe suivant.

I.2.2 Mesures des performances

On définit la **défaillance** comme la cessation de l'aptitude d'une entité à accomplir une fonction requise [Villemeur, 1988]. On va considérer T la variable aléatoire qui représente le temps écoulé entre la mise en marche d'un matériel et la première défaillance observée.

Une **variable aléatoire** est une fonction définie sur l'ensemble des résultats possibles d'une expérience aléatoire, telle qu'il soit possible de déterminer la probabilité pour qu'elle prenne une valeur donnée ou qu'elle prenne une valeur dans un intervalle donné [Tiplica, 2007].

La **théorie des probabilités** est l'étude mathématique des phénomènes caractérisés par le hasard et l'incertitude. Plus précisément, les probabilités sont utilisées pour modéliser des situations dans lesquelles une même expérience réalisée à des différents moments mais dans les mêmes conditions, peut produire des résultats différents.

Toutes les mesures des performances suivantes sont déclinées de la théorie plus générale de la probabilité pour une variable aléatoire X , adaptée pour le cas concret $X = T$, et interprétée en association avec leurs sens physiques.

Fonction de fiabilité

La fonction de fiabilité ou fonction de survie $R(t)$, est définie comme l'aptitude d'une entité à accomplir une fonction requise, dans des conditions données, pendant un intervalle de temps donné.

La fiabilité à l'instant t est la probabilité qu'une entité E soit non défaillante sur la durée $[0; t]$, en supposant qu'elle n'est pas défaillante à l'instant $t = 0$.

Par extension, on appelle également fiabilité, la probabilité associée $R(t)$ définie par :

$$R(t) = P(T > t) \tag{I.6}$$

Les aspects théoriques de la fiabilité consisteront donc en une application particulière de la théorie des probabilités aux problèmes de durée de fonctionnement sans incidents qui va utiliser toutes les informations autour du sujet.

La figure I.4 ci-dessous présente une allure de la fonction $R(t)$ en fonction du temps.

Pour compléter l'approche théorique de la notion de fiabilité, il est nécessaire de définir aussi les notions suivantes, qui sont dérivées de la théorie des probabilités.

La fonction de répartition

La fonction $F(t)$ de l'équation I.7 représente la fonction de répartition de la variable aléatoire T . C'est la défiabilité $\overline{R(t)}$ (la probabilité de défaillance du système) ou le complément à 1 de la probabilité de fiabilité $R(t)$ définie par :

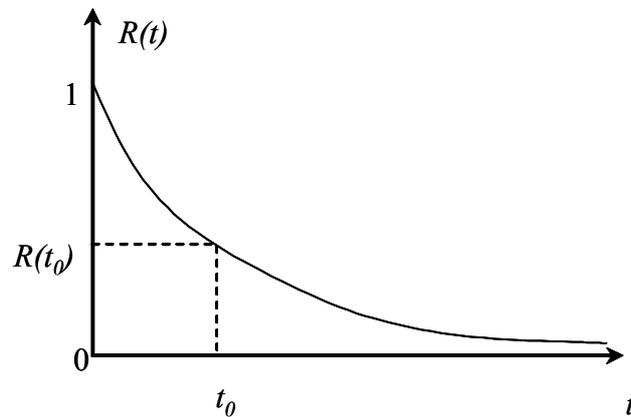


FIG. I.4 – Courbe de survie ou de fiabilité.

$$\begin{aligned} F(t) &= P(T \leq t) \\ F(t) &= \overline{R(t)} = 1 - R(t) \end{aligned} \quad (\text{I.7})$$

La fonction de répartition est la primitive de la densité de probabilité $f(t)$ (voir I.2.2).
La fonction de répartition a les propriétés suivantes :

- $F(t)$ est croissante ;
- $F(t)$ est partout continue à droite ;
- $\lim_{t \rightarrow 0} F(t) = 0$;
- $\lim_{t \rightarrow \infty} F(t) = 1$.

Densité de probabilité

Si dt est un nombre réel positif infiniment petit (pour nous un intervalle de temps), alors la probabilité d'obtenir T inclu dans l'intervalle $[t; t + dt]$ est égale à $f(t)dt$, soit :

$$P(t < T < t + dt) = F(t + dt) - F(t) \quad (\text{I.8})$$

avec :

$$\frac{F(t + dt) - F(t)}{dt} = f(t) \quad (\text{I.9})$$

Finalement, on obtient la relation générale de la fonction de densité :

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt} \quad (\text{I.10})$$

sachant que $f(t)$ est une fonction localement intégrable pour la mesure de Lebesgue .

De même, nous pouvons ajouter les relations suivantes :

$$\begin{aligned}
 F(t) &= \int_0^t f(u)du \\
 R(t) = 1 - F(t) &= 1 - \int_0^t f(u)du = \int_t^{\infty} f(u)du
 \end{aligned}
 \tag{I.11}$$

Le taux de défaillance instantané

Il est aussi possible de définir la notion de taux instantané de défaillance au temps t , notée $\lambda(t)$. La valeur $\lambda(t)dt$ représente la probabilité d'avoir une défaillance dans l'intervalle de temps $[t; t + dt]$, sachant qu'il n'y a pas eu de défaillance dans l'intervalle de temps $[0; t]$.

Ainsi, en appliquant le théorème des probabilités conditionnelles, puis le théorème des probabilités totales, $\lambda(t)$ s'écrit :

$$\begin{aligned}
 \lambda(t)dt &= \frac{\text{Prob(défaillant sur } [t; t + dt] \text{ sans défaillance sur } [0; t])}{\text{Prob(non défaillant sur } [0; t])} \\
 \lambda(t)dt &= \frac{\text{Prob(défaillant sur } [0; t + dt]) - \text{Prob(défaillant sur } [0; t])}{\text{Prob(non défaillant sur } [0; t])}
 \end{aligned}$$

soit finalement

$$\lambda(t) = \frac{f(t)}{R(t)} = -\frac{1}{R(t)} \cdot \frac{dR(t)}{dt}
 \tag{I.12}$$

On peut aussi remarquer que :

$$\begin{aligned}
 \lambda(t) &= -\frac{d \ln [R(t)]}{dt} \\
 \text{d'où : } \ln [R(t)] &= -\int_0^t \lambda(u)du \\
 \text{et : } R(t) &= \exp \left[-\int_0^t \lambda(u)du \right]
 \end{aligned}
 \tag{I.13}$$

Au cours de la vie d'un matériel, de manière empirique, nous pouvons distinguer trois phases sur l'évolution du taux de défaillance $\lambda(t)$ au cours du temps t , à travers une

courbe caractéristique dite en "baignoire" (figure I.5).

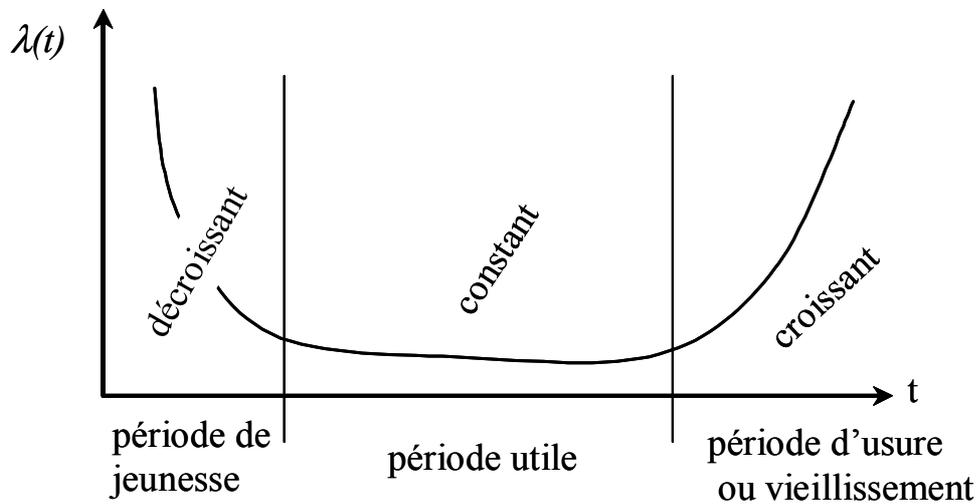


FIG. I.5 – Courbe en baignoire.

Les trois phases sur l'évolution du taux de défaillance sont :

- la première concerne les défaillances précoces dues à des problèmes de conception (mauvais dimensionnement d'un composant, ...) ou de production (dérive d'un processus de fabrication,...);
- la deuxième partie, pouvant être plus ou moins importante selon le type de matériel (plus pour l'électronique et moins pour la mécanique), est caractéristique des défaillances aléatoires;
- la troisième partie correspond aux défaillances dues à des phénomènes d'usure, de vieillissement, etc.

Le percentile

Le percentile de 100*P*-ème d'une distribution $F(t)$ est l'instant τ_P par lequel une proportion P de la population est défaillante. C'est la solution de

$$\tau_P = F^{-1}(P) \quad (\text{I.14})$$

Les moments

On appelle moment d'ordre k de la variable T , lorsqu'il existe, le réel défini par :

$$E [T^k] = \int_{-\infty}^{+\infty} t^k f(t) dt \quad (\text{I.15})$$

$E[X]$ est l'espérance mathématique ou moyenne. La moyenne peut être estimée directement

à partir d'un échantillon de taille n par :

$$E [T] = m = \frac{\sum t_i}{n} \quad (I.16)$$

On appelle variance le réel $V[T]$, lorsqu'il existe, défini par :

$$V[T] = \int_{-\infty}^{+\infty} (t - E[T])^2 f(t) dt \quad (I.17)$$

La variance peut être également estimée par :

$$V[T] = \frac{\sum (t_i - m)^2}{n - 1} \quad (I.18)$$

$\sqrt{V[T]}$ est appelé écart type (noté σ); d'où $V[T] = \sigma^2[T]$

La moyenne et l'écart type permettent de caractériser, respectivement, la tendance et la dispersion.

Mode

Le mode permet également de caractériser la tendance. C'est la valeur du caractère statistique qui apparaît le plus fréquemment. Dans le cas d'une variable continue, on peut entendre parler de classe modale qui serait la classe de plus grand effectif.

I.3 L'analyse de données de Fiabilité

La plupart des cas d'analyse de fiabilité utilisent comme données les instants de défaillance résultant d'essais ou de l'exploitation.

La figure I.6 présente, des instants de défaillance t_i ainsi que des suspensions τ_i pour un échantillon de taille n . La suspension matérialise le fait que l'on ne connaît pas exactement le temps de défaillance d'un équipement. Plusieurs situations peuvent produire des suspensions :

- l'essai est arrêté avant l'apparition de toutes les défaillances (censure à droite),
- la défaillance est constatée après son apparition lors d'une inspection par exemple (censure à gauche),
- la défaillance est apparue entre deux instants d'inspection (temps de censure par intervalle)

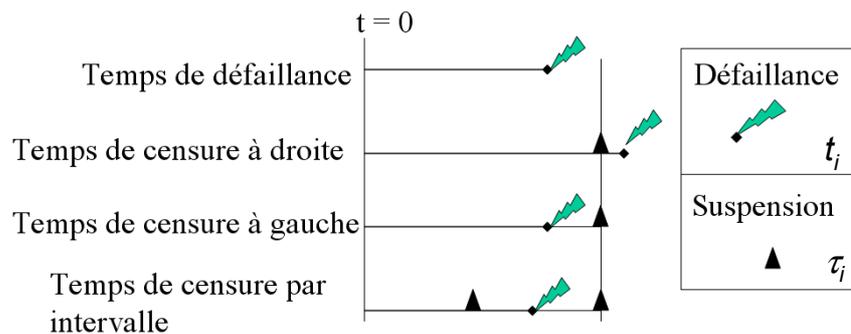


FIG. I.6 – Les types de données

A partir de ces données, des conclusions sur la loi de fiabilité peuvent être tirées. En fonction des informations dont on dispose déjà sur le produit testé, on peut faire des hypothèses plus ou moins restrictives afin d'améliorer les estimations, comme :

- l'existence ou non d'une loi paramétrique de la fonction de fiabilité
- un mode de défaillance unique,
- plusieurs modes de défaillance, etc.

Dans le cadre de la thèse, nous considérons le cas d'un seul mode de défaillance et de données censurées à droite correspondant à la majorité des cas observés lors des essais.

A partir de ces hypothèses, différentes approches sont possibles ; une classification rapide en deux catégories peut être opérée selon la connaissance ou non d'une loi paramétrique pour décrire la courbe de fiabilité :

- L'approche non paramétrique : elle est basée sur aucune hypothèse sur l'existence d'une loi pour approximer la courbe de fiabilité.

- L’approche paramétrique : elle fait l’hypothèse d’une loi continue, dépendante de certains paramètres et qui approxime la courbe réelle de fiabilité.

I.3.1 L’approche non paramétrique

L’approche non paramétrique fournit une solution au problème de l’inférence pour les cas où un modèle paramétrique soit difficile, sinon impossible, à trouver à cause de la complexité mathématique ([Sawa, 1978]). L’avantage de ces modèles est qu’ils n’ont pas la contrainte d’appartenance à une telle classe de distribution.

Dans le cas classique d’un échantillon simple, il existe une méthode universelle d’estimation de toutes les caractéristiques de fiabilité quelque soit le modèle statistique sous-jacent [Nikulin, 2007]. Soit $\mathbf{T} = (T_1, \dots, T_n)^\top$ un échantillon simple avec $F(t) = \mathbf{P} \{T_i \leq t\}$ étant la fonction de répartition de T_i . Ayant une réalisation $t_{réalisé} = (t_1, \dots, t_n)^\top$ de \mathbf{T} , et sachant que les valeurs t_i sont des réalisations de variables aléatoires indépendantes et identiquement distribuées, il est naturel de donner la même masse de probabilité $1/n$ à chaque valeur x_i et construire la fonction de répartition suivante :

$$F_n(t) = \frac{m_n(t)}{n}, t \in \mathfrak{R} \quad (\text{I.19})$$

où :

$$m_n(t) = \sum_{i=1}^n \mathbf{1} \{t_i \leq t\} \quad (\text{I.20})$$

est le compteur du nombre de valeurs t_i , vérifiant la condition $t_i \leq t$:

$$\mathbf{1} \{t_i \leq t\} = \begin{cases} 1, & \text{si } t_i \leq t \\ 0, & \text{sinon} \end{cases}$$

La valeur $F_n(t) = m(t)/n$ de la fonction F_n en n’importe quel point t représente une réalisation de la statistique :

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1} \{t_i \leq t\} \quad (\text{I.21})$$

calculée au point choisi t .

La statistique $\hat{F}_n(t)$ est appelée fonction de répartition empirique. Alors, la valeur estimée de la fiabilité à un instant t_i est :

$$R(t_i) = 1 - \frac{1}{n} \sum_{i=1}^n \mathbf{1} \{t_i \leq t\} \quad (\text{I.22})$$

Dans l'hypothèse d'un échantillon composé par des instants de défaillances et des instants de suspensions, le changement de masse de probabilité prend une valeur dépendante du nombre de suspensions ayant lieu entre deux défaillances successives. Plusieurs méthodes ont été développées :

- la méthode des rangs corrigés de Johnson
- la méthode de Kaplan-Meier
- la méthode de Wayne-Nelson, etc.

Devant le nombre important des méthodes non paramétriques existantes, par des raisons de simplicité de calculs, nous allons présenter plus particulièrement la méthode de Kaplan-Meier que nous allons utiliser par la suite. Ce choix n'impacte pas la méthodologie développée par la suite.

La méthode de Kaplan-Meier permet d'évaluer ponctuellement la fiabilité d'un matériel aux instants de défaillances quand il existe des censures à droite. Soient $\{t_1, t_2, \dots, t_{j-1}, t_j, \dots, t_{i-1}, t_i, \dots, t_k\}$ les instants de défaillances constatés sur un échantillon de taille n . Il y a donc $n - k$ censures à droite.

Ainsi, la fiabilité est obtenue par :

$$R(t_i) = \prod_{j=1}^{j=i} \left(1 - \frac{1}{n_j}\right) \quad (\text{I.23})$$

avec n_j est le nombre de systèmes encore en fonctionnement et non censurés juste avant la j^{eme} défaillance constatée.

La formule de Greenwood I.24 donne une estimation de la variance de $\log(\hat{R}(t))$ [Ansell et Phillips, 1994] :

$$\hat{\text{Var}}\left(\log\left(\hat{R}(t)\right)\right) = \sum_{i:t_i \leq t} \left\{ \frac{1}{n_j(n_j - 1)} \right\} \quad (\text{I.24})$$

La variance de $R(t)$ est donnée par $(R(t))^2 \hat{\text{Var}}(\log(R(t)))$. La figure I.7 présente un exemple d'une fonction de fiabilité définie par la méthode de Kaplan-Meier et l'intervalle de confiance à 90%.

Les approches non paramétriques présentent un inconvénient majeur parce qu'aucune extrapolation est possible au delà du dernier temps de défaillance. Ceci peut devenir critique dans le cas de données fortement censurées.

I.3.2 L'approche paramétrique

L'approche paramétrique, d'après son nom, utilise des lois paramétriques pour approximer la fonction de fiabilité. Pour ce faire, il est nécessaire de choisir la loi adaptée au

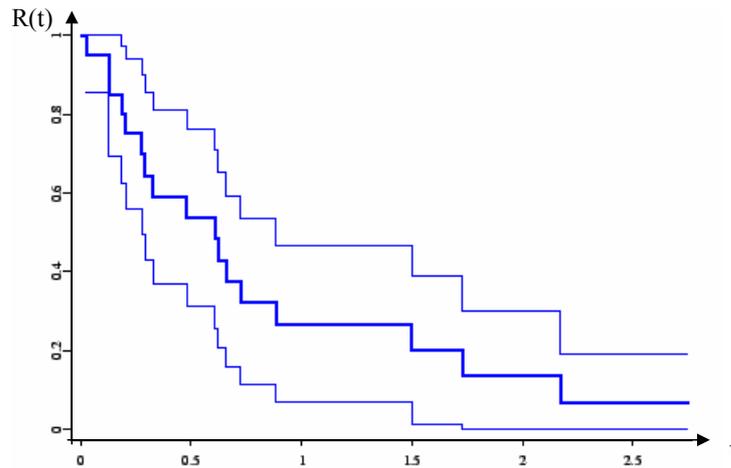


FIG. I.7 – L'estimation de la fiabilité par Kaplan Meier

mécanisme de défaillances ou au produit. Pendant des années, de nombreux chercheurs ont développé différentes lois statistiques et/ou ont associé les lois mathématiques avec le processus physique de défaillance. Dans l'analyse statistique des données, un modèle paramétrique est supposé pour la distribution inconnue ; mais le modèle est employé juste comme une approximation de la distribution inconnue. Il est alors permis de considérer que le degré de confiance d'une analyse paramétrique s'appuie sur la façon dans laquelle le modèle choisi approche la distribution inconnue.

I.3.2.1 Le modèle standard de durée de vie

Une forme générale de fonctions de fiabilité est décrite par le modèle standard de durée de vie. On dit que la variable aléatoire T , temps jusqu'à la première défaillance, suit une loi de fiabilité $R(t)$ si elle représente la probabilité que T soit supérieur à t (équation I.6). La probabilité que la durée de vie soit supérieure à t , $R(t)$ est continue à droite en tout point.

On suppose que la fonction de fiabilité $R(t)$ appartient à une classe de fonctions qui ne dépend que des paramètres d'échelle, de forme et de décalage [Bagdonavicius et al., 2000] :

$$R(t) = R\left(\left(\frac{t - \omega}{\eta}\right)^v\right), (\eta, v > 0) \quad (\text{I.25})$$

avec $\theta = \{\eta, v, \omega\}$ vecteur des paramètres inconnus (η , v et ω respectivement les paramètres d'échelle, de forme et de position (ou de décalage)).

Si $\omega \neq 0$, alors ce terme matérialisera la durée de vie consommée lors d'un déverminage trop agressif ($\omega > 0$) ou d'un délai de mise en service ($\omega < 0$). Dans le cas des essais, on

ne retrouve pas cette situation car les équipements à tester sont produits pour l'essai, le déverminage est contrôlé et le vieillissement est considéré négligeable. Aussi, nous n'allons traiter que les cas où $\omega = 0$.

Les lois classiquement utilisées en fiabilité (Weibull, loglogistique et lognormale) appartiennent à cette famille de modèles.

I.3.2.1.1 La loi Weibull En considérant $R(t) = e^{-t}$ dans l'équation I.25 on obtient la classe de fonction de fiabilité de Weibull avec :

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^v} \quad (\text{I.26})$$

et la fonction de densité de probabilité :

$$f(t) = \frac{v}{\eta} \left(\frac{t}{\eta}\right)^{v-1} e^{-\left(\frac{t}{\eta}\right)^v} \quad (\text{I.27})$$

C'est un modèle pouvant caractériser l'ensemble de la courbe en baignoire tout au long de la vie d'un produit. Ce type de loi est plutôt utilisé pour les composants mécaniques.

Le paramètre de forme v donne des indications sur le mode des défaillances et sur l'évolution du taux de défaillances dans le temps [Lannoy, 1996] :

- $v < 1 \Rightarrow \lambda(t)$ décroît \Rightarrow période de jeunesse
- $v = 1 \Rightarrow \lambda(t)$ constant \Rightarrow indépendance du temps, loi exponentielle (paragraphe I.3.2.1.2)
- $v > 1 \Rightarrow \lambda(t)$ croît \Rightarrow période de vieillissement ou d'usure :
 - $1.5 < v < 2.5 \rightarrow$ phénomène de fatigue
 - $3 < v < 4 \rightarrow$ phénomène d'usure (remarque : $v = 3.5 \Rightarrow$ distribution normale)

I.3.2.1.2 La loi exponentielle C'est une loi Weibull de paramètre $v = 1$:

$$R(t) = e^{-\lambda t}, \quad \text{avec } \lambda = \frac{1}{\eta} \quad (\text{I.28})$$

La loi exponentielle est la plus populaire car elle est la plus simple, mais elle suppose un taux de défaillance constant. Elle ne dépend que d'un seul paramètre : le taux de défaillance λ . Cette loi est associée au processus de Poisson qui génère des événements (défaillances) dont les temps d'occurrence t sont indépendants et distribués identiquement selon une loi exponentielle. La loi exponentielle est souvent utilisée en fiabilité pour mesurer la durée de vie des composants et des circuits électroniques. L'une des spécificités de cette loi est donnée par le fait qu'il n'y a pas d'effet mémoire.

L'équation I.25 permet aussi, pour $R(t) = \frac{1}{1+t}$ d'avoir la classe loglogistique et si $R(t) = 1 - \Phi(\ln t)$, avec Φ la fonction de répartition de la loi normale standard, on a la classe lognormale. Les méthodes d'estimation des paramètres à partir d'un nombre fini d'observations sont basées principalement sur la modélisation stochastique des apparitions des défaillances au cours du temps et sur l'estimation statistique des courbes de fiabilité à partir des résultats d'essai.

I.3.2.2 Les estimations de paramètres

Comparée à l'approche probabiliste, la démarche statistique est fondamentalement une démarche d'inversion, car elle se propose de remonter, même imparfaitement, des 'effets' (observations) aux 'causes' (paramètres du mécanisme générateur). En d'autres termes, ayant observé des variables aléatoires générées suivant un mécanisme dirigé par des paramètres, les méthodes statistiques permettent de conduire, à partir de ces observations, une inférence sur les paramètres, tandis que le modèle probabiliste caractérise le comportement des observations, conditionnellement aux paramètres.

Une fois la classe de loi définie, la démarche statistique permet d'avoir des estimations des paramètres inconnus en considérant les instants de défaillance.

L'estimation d'un paramètre, considéré comme une variable aléatoire, comprend l'estimateur ponctuel et son intervalle de confiance. Plusieurs méthodes possibles sont à la disposition de l'analyste. La suite de la thèse va utiliser les méthodes du maximum de vraisemblance et bayésienne non informative.

Pour chaque méthode, des techniques spécifiques sont prévues pour trouver les intervalles de confiance. Pour la méthode de maximum de vraisemblance, on utilise l'information de Fisher sur l'hypothèse de normalité. Pour la méthode bayésienne non informative nous allons utiliser la construction des fonctions marginales pour caractériser complètement la variable aléatoire.

Les deux méthodes sont basées sur l'utilisation de la fonction de vraisemblance.

On observe les durées successives de fonctionnement de n produits identiques jusqu'à l'instant de censure t_c . Au cours de l'observation, on enregistre les instants de défaillance ($t = (t_1, \dots, t_k)$ avec k le nombre de défaillance observées).

On considère la fonction de fiabilité de paramètres $\theta = (\theta_1, \dots, \theta_r) \in \Omega$. Alors, en utilisant $f(t, \theta)$, on définit la fonction de vraisemblance par :

$$L(t, \theta) = \left(\prod_{j=1}^k f(t_j, \theta) \right) R(t_c, \theta)^{(n-k)} \quad (I.29)$$

I.3.2.2.1 Les estimations par la méthode du maximum de vraisemblance

(MV) : La méthode du maximum de vraisemblance consiste dans un premier temps à évaluer la vraisemblance $L(t, \theta)$ des observations, c'est-à-dire la probabilité d'avoir obtenu un certain nombre de défaillances et de censures à des instants ou des intervalles de temps donnés pour un échantillon, x .

Si $L(t, \theta)$ est dérivable et si le maximum de vraisemblance $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_r)$ existe, alors il satisfait l'équation suivante :

$$\hat{\theta} = \text{Argmax} (L(t, \theta)) \quad (\text{I.30})$$

qui implique :

$$\left. \frac{\partial L(t_1 \dots t_n, \theta_1 \dots \theta_r)}{\partial \theta_j} \right|_{\theta_j = \hat{\theta}_j} = 0 \quad (\text{I.31})$$

Le calcul $R(t|\theta = \hat{\theta})$ permet d'approximer la courbe de fiabilité associée au produit générant des temps de défaillance pour l'estimation ponctuelle $\theta = \hat{\theta}$.

L'information de Fisher est définie comme la matrice des dérivées d'ordre deux de la log vraisemblance calculée pour $\theta = \hat{\theta}$ qui satisfait I.31 :

$$I(\theta) = E \left[\left(\frac{\partial \log L(T, \theta)}{\partial \theta} \right)^2 \Big|_{\hat{\theta}} \right] \quad (\text{I.32})$$

Dans l'hypothèse où la distribution de probabilité dépend de plusieurs paramètres, θ n'est plus un scalaire mais un vecteur $\theta = (\theta_1, \theta_2, \dots)$, la recherche du maximum de vraisemblance ne se résume donc pas à une seule équation mais à un système :

$$E \left[\frac{\partial}{\partial \theta_i} \log f(T; \theta) \right] = 0, \quad \forall i \quad (\text{I.33})$$

Enfin, l'information de Fisher n'est plus définie comme une variance scalaire mais comme une matrice de covariance :

$$I(\theta_i, \theta_j) = E \left[\left(\frac{\partial}{\partial \theta_i} \log f(T; \theta) \right) \left(\frac{\partial}{\partial \theta_j} \log f(T; \theta) \right) \right] \quad (\text{I.34})$$

En effet, le passage de l'espace des observations à l'espace des paramètres est un changement de système de coordonnées. L'inverse de cette matrice permet quant à elle de déterminer les covariances relatives aux estimations conjointes des différents paramètres à partir des observations. Le fait que tous les paramètres soient à estimer simultanément rend l'estimation plus difficile. Ce phénomène est une manifestation de ce qui est parfois

appelé le *fléau de la dimension* [Bellman, 1961].

Soit $\psi(\theta)$ le vecteur d'estimateur $E[\mathbf{T}]$ avec $\mathbf{T} = (T_1, \dots, T_n)^T$.

La borne Cramér-Rao impose que la matrice de covariance pour \mathbf{T} respecte :

$$\text{cov}_\theta(\mathbf{T}) \geq \frac{\partial \psi(\theta)}{\partial \theta} [I(\theta)]^{-1} \left(\frac{\partial \psi(\theta)}{\partial \theta} \right)^T \quad (\text{I.35})$$

avec :

- l'inégalité des matrices $A \geq B$ signifie que la matrice $A - B$ est une matrice semi définie
- $\partial \psi(\theta) / \partial \theta$ est une matrice qui a le ij élément donné par $\partial \psi_i(\theta) / \partial \theta_j$.

Si \mathbf{T} est un estimateur non biaisé de θ (i.e., $\psi(\theta) = \theta$), alors l'inégalité de Cramér-Rao se simplifie [Fisher, 1922] :

$$\text{cov}_\theta(\mathbf{T}) \geq I(\theta)^{-1} \quad (\text{I.36})$$

Alors, pour calculer la variance de $\hat{\theta}_i$, $\hat{V}ar(\hat{\theta}_i)$, la log vraisemblance $L_i(t_i, \theta_i)$ doit être dérivable d'ordre deux. Ensuite, nous construirons la matrice symétrique de Fisher ainsi :

$$\hat{I} = \begin{bmatrix} -\frac{\partial^2 L_i}{\partial \theta_{i1}^2} & -\frac{\partial^2 L_i}{\partial \theta_{i1} \partial \theta_{i2}} & \dots & -\frac{\partial^2 L_i}{\partial \theta_{i1} \partial \theta_{ir}} \\ -\frac{\partial^2 L_i}{\partial \theta_{i2} \partial \theta_{i1}} & -\frac{\partial^2 L_i}{\partial \theta_{i2}^2} & \dots & -\frac{\partial^2 L_i}{\partial \theta_{i2} \partial \theta_{ir}} \\ \dots & \dots & \dots & \dots \\ -\frac{\partial^2 L_i}{\partial \theta_{ir} \partial \theta_{i1}} & -\frac{\partial^2 L_i}{\partial \theta_{ir} \partial \theta_{i2}} & \dots & -\frac{\partial^2 L_i}{\partial \theta_{ir}^2} \end{bmatrix} \quad (\text{I.37})$$

La matrice inverse \hat{I}^{-1} de Fisher représente la matrice estimée de variance – covariance dénommée $\hat{\Sigma}$:

$$\hat{\Sigma} = \begin{bmatrix} Var(\hat{\theta}_{i1}) & Cov(\hat{\theta}_{i1}, \hat{\theta}_{i2}) & \dots & Cov(\hat{\theta}_{i1}, \hat{\theta}_{ir}) \\ Cov(\hat{\theta}_{i2}, \hat{\theta}_{i1}) & Var(\hat{\theta}_{i2}) & \dots & Cov(\hat{\theta}_{i2}, \hat{\theta}_{ir}) \\ \dots & \dots & \dots & \dots \\ Cov(\hat{\theta}_{ir}, \hat{\theta}_{i1}) & Cov(\hat{\theta}_{ir}, \hat{\theta}_{i2}) & \dots & Var(\hat{\theta}_{ir}) \end{bmatrix} = \hat{I}^{-1} \quad (\text{I.38})$$

Le théorème de la limite centrale est un ensemble de résultats sur la convergence faible d'une suite de variables aléatoires en probabilité.

Généralement, les évaluations de MV des paramètres sont asymptotiquement normales, signifiant que, pour des grandes tailles d'échantillon, une distribution de paramètre de la même population serait très près de la distribution normale. Ainsi, si $\hat{\theta}$ est l'estimateur de MV pour θ , dans le cas d'une distribution d'un seul paramètre, estimée à partir

d'une grande taille d'échantillon, on a :

$$z = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} \quad (I.39)$$

avec :

$$P(x \leq z) \rightarrow \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt \quad (I.40)$$

Pour C_1 et C_2 respectant :

$$P(C_1 < \theta < C_2) = \alpha \quad (I.41)$$

avec α un niveau de confiance donné, l'équation I.40 devient :

$$P\left(-w_{1-\frac{\alpha}{2}} < \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} < w_{1-\frac{\alpha}{2}}\right) \cong \alpha \quad (I.42)$$

avec w_α définit par $\alpha = \frac{1}{\sqrt{2\pi}} \int_{w_\alpha}^{\infty} e^{-\frac{t^2}{2}} dt$

Alors, en simplifiant l'équation I.42, pour un certain niveau de confiance α , on a les limites de confiance :

$$\hat{\theta} - w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})} < \theta < \hat{\theta} + w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})} \quad (I.43)$$

Les limites de confiance pour le cas où θ doit être positif sont obtenues en traitant le $\log(\theta)$ comme une variable normal distribuée ; l'équation I.43 devient :

$$\hat{\theta} e^{-\frac{w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})}}{\theta}} < \theta < \hat{\theta} e^{\frac{w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})}}{\theta}} \quad (I.44)$$

I.3.2.2.2 La méthode bayésienne non informative : Cette méthode consiste à appliquer la démarche bayésienne (qui sera très largement présentée dans le Chapitre 2) en considérant un a priori non informatif. Le théorème de Bayes s'écrit dans le cas général :

$$g(\theta|t) = \frac{L(t, \theta) \times g(\theta)}{\int_{\Omega} L(t, \theta) \times g(\theta) d\theta} \quad (I.45)$$

Les fonctions de densité $g(\theta)$ et $g(\theta|t)$ représentent respectivement l'a priori et l'a

posteriori. En faisant l'hypothèse que chaque variable θ_i est indépendante, alors la distribution a priori de θ peut s'écrire comme le produit de fonctions de densité associées à chaque paramètre θ_i :

$$g(\theta) = \prod_{i=1}^r g_i(\theta_i) \quad (I.46)$$

En prenant des lois uniformes non informatives (comme définie dans la figure I.8) les fonctions de densité deviennent indépendantes des paramètres θ_i .

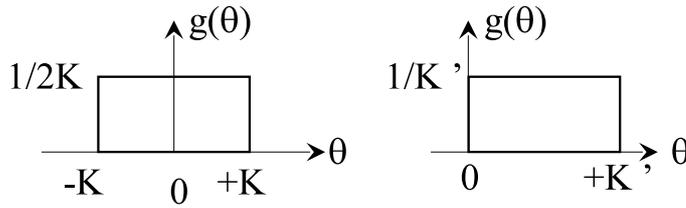


FIG. I.8 – Lois uniformes non informatives

Alors, la fonction de densité a posteriori dépend uniquement de la vraisemblance :

$$g(\theta|t) = \frac{L(t, \theta)}{\int_{\Omega} L(t, \theta) d\theta} \quad (I.47)$$

La fonction de densité $g_i(\theta_i)$ du paramètre θ_i s'écrit :

$$g_i(\theta_i) = \int_{j=1..n, j \neq i} \dots \int g(\theta|t) d\theta_j \quad (I.48)$$

Dans ce contexte, l'estimateur ponctuel est le moment initial d'ordre 1 :

$$\hat{\theta}_i = \int_{\Omega_{\theta_i}} \theta_i g_{\theta_i}(\theta_i) d\theta_i \quad (I.49)$$

La variance est définie par le moment centré d'ordre 2 :

$$V(\hat{\theta}_i) = \int_{\Omega_{\theta_i}} g_{\theta_i}(\theta_i) (\theta_i - \hat{\theta}_i)^2 d\theta_i \quad (I.50)$$

La méthode présentée nécessite de recourir à des méthodes numériques lorsque les expressions analytiques n'existent pas. Les variances sont utilisées pour faire les estimations des intervalles de confiance selon les relations I.43 et I.44. Il est possible également de définir l'intervalle de confiance bilatérale symétrique au niveau de confiance $1 - \alpha$ tel que :

$$prob(\theta_i^{min} \leq \theta_i \leq \theta_i^{max}) = 1 - \alpha \quad (I.51)$$

avec θ_i^{min} et θ_i^{max} solutions de :

$$\int_{D(\theta_i)}^{\theta_i^{min}} g_i(\theta_i) d\theta_i = \alpha/2 \quad (I.52)$$

$$\int_{D(\theta_i)}^{\theta_i^{max}} g_i(\theta_i) d\theta_i = 1 - \alpha/2 \quad (I.53)$$

Dans le paragraphe I.4 est proposée une analyse comparative des deux méthodes présentées.

I.3.2.3 Estimation des caractéristiques de fiabilité

A partir des estimations des paramètres θ de la loi de fiabilité, il est possible d'évaluer des caractéristiques particulières de fiabilité comme par exemple : une probabilité de défaillance pour un temps donné, le MTTF, le L10, ... Ces évaluations doivent s'accompagner d'une estimation d'un intervalle de confiance intégrant la variabilité sur chaque paramètre θ_i . Pour ce faire, nous utiliserons l'algèbre des variables aléatoires.

Soit la caractéristique y de fiabilité à estimer reliée à la fonction de fiabilité, de paramètre θ , par la relation $f(\theta)$ (par exemple, la durée L10 est définie par : $L10 = f(\theta) = R^{-1}(0, 1, \theta)$)

L'estimateur ponctuel de la grandeur y est donné par :

$$\hat{y} \approx f(\hat{\theta}) \quad (I.54)$$

L'intervalle de confiance sur y peut être estimé à partir des relations I.43 et I.44 en utilisant la variance $V(\hat{y})$ définie par :

$$V(\hat{y}) \approx \sum_{i=1}^r \sum_{j=1}^r \left(\frac{\partial f(\theta)}{\partial \theta_i} \right)_{\hat{\theta}} \left(\frac{\partial f(\theta)}{\partial \theta_j} \right)_{\hat{\theta}} COV(\hat{\theta}_i, \hat{\theta}_j) \quad (I.55)$$

Pour des cas complexes, il est possible de faire appel à des méthodes numériques ou de recourir à la technique du Bootstrap.

I.4 Analyse des modèles de fiabilité et méthodes d'estimation par simulation de Monte Carlo

L'objectif de cette simulation est d'étudier l'influence de la taille d'échantillon et de la censure sur les modèles paramétrique et non paramétrique et les méthodes d'estimation. Pour ce faire, nous avons simulé les instants de défaillance par un tirage aléatoire en considérant une loi de Weibull de paramètres, respectivement de forme et d'échelle, $\nu = 2,34$ et $\eta = 114$ mois. Dans un premier temps, nous avons étudié l'influence de la taille d'échantillon puis dans un second temps celle de la censure.

I.4.1 Influence de la taille d'échantillon

Pour étudier l'influence de la taille d'échantillon, nous avons simulé des instants de défaillance pour quatre configurations : $n = 5, 15, 25, 50$. Ensuite, nous avons analysé les instants de défaillance par les approches non paramétrique et paramétrique.

I.4.1.1 Les modèles non paramétriques

La méthode de Kaplan Meier a été appliquée pour traiter les données de simulations. Comme déjà présenté dans la section I.3.1, pour chaque instant de défaillance t_i une valeur de la fiabilité $R(t_i)$ a été estimée. La figure I.9 présente le résultat d'analyse pour les quatre tailles d'échantillon. On peut remarquer que plus la taille d'échantillon est faible plus la fonction de fiabilité est grossière. Néanmoins, aucune hypothèse est faite quant à la nature de la loi. Cette analyse sera complétée avec une comparaison des approches non paramétrique et paramétrique.

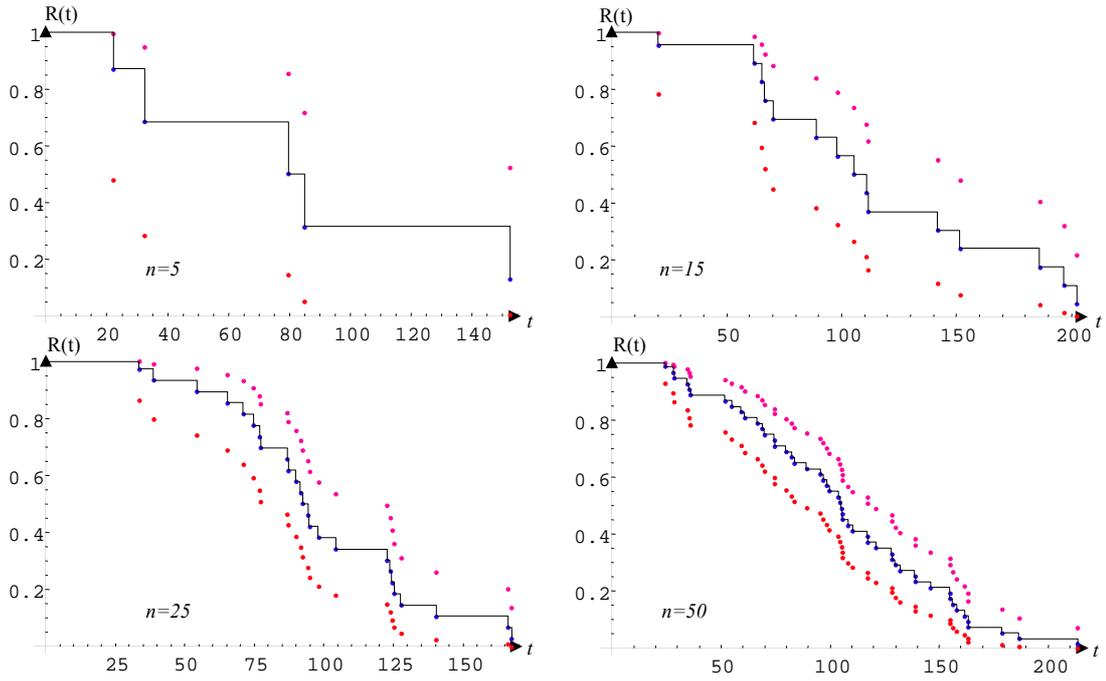


FIG. I.9 – L'estimation de la fiabilité non paramétrique

I.4.1.2 Les modèles paramétriques

Dans le cas de l'approche paramétrique, il est nécessaire de faire une hypothèse sur la nature du modèle qui sera utilisé pour l'analyse. Ce choix se fait selon plusieurs critères comme la nature des composants, le mode de défaillance observé, ... et plusieurs modèles candidats peuvent être testés et le "meilleur" sera retenu. Dans notre cas, nous reprendrons le modèle de Weibull, qui a été utilisé pour faire la simulation, pour analyser les données simulées pour les quatre tailles d'échantillons ($n = 5, 15, 25, 50$).

En considérant un vecteur de n instants de défaillance, t_1, t_2, \dots, t_n , la fonction de vraisemblance (équation I.29) associée au cas Weibull est :

$$L(t|v, \eta) = \prod_{i=1}^n \frac{v}{\eta} \left(\frac{t_i}{\eta} \right)^{v-1} e^{-\left(\frac{t_i}{\eta} \right)^v} \quad (\text{I.56})$$

L'estimation des paramètres inconnus v et η se fera selon les deux méthodes présentées dans I.3.2.2.

L'approche par maximum de vraisemblance

Pour trouver les estimateurs ponctuels par la méthode du maximum de vraisemblance, il suffit de résoudre les deux équations suivantes :

$$\left(\frac{\partial \text{Ln}(L)}{\partial v}\right)_{\hat{v}, \hat{\eta}} = \frac{n}{\hat{v}} - n \ln \hat{\eta} + \sum_{i=1}^{i=n} \ln t_i - \frac{1}{\hat{\eta}^{\hat{v}}} \sum_{i=1}^{i=n} t_i^{\hat{v}} \ln t_i + \frac{1}{\hat{\eta}^{\hat{v}}} (\ln \hat{\eta}) \sum_{i=1}^{i=n} t_i^{\hat{v}} = 0 \quad (\text{I.57})$$

et

$$\left(\frac{\partial \text{Ln}(L)}{\partial \eta}\right)_{\hat{v}, \hat{\eta}} = -\frac{n\hat{v}}{\hat{\eta}} + \frac{\hat{v}}{\hat{\eta}^{\hat{v}+1}} \sum_{i=1}^{i=n} t_i^{\hat{v}} = 0 \quad (\text{I.58})$$

Après simplification, nous obtenons :

$$\frac{n}{\hat{v}} + \sum_{i=1}^{i=n} \ln t_i - n \frac{\sum_{i=1}^{i=n} t_i^{\hat{v}} \ln t_i}{\sum_{i=1}^{i=n} t_i^{\hat{v}}} = 0 \quad (\text{I.59})$$

et :

$$\hat{\eta} = \left(\frac{\sum_{i=1}^{i=n} t_i^{\hat{v}}}{n} \right)^{\frac{1}{\hat{v}}} \quad (\text{I.60})$$

Les intervalles de confiance (définis par I.69 et I.70) sur v et η sont obtenus en considérant les variances estimées par :

$$V(\hat{v}) \cong \frac{0.608}{k} \hat{v}^2 \quad (\text{I.61})$$

et :

$$V(\hat{\eta}) \cong \frac{1.087}{k} \left(\frac{\hat{v}}{\hat{\eta}} \right)^2 \quad (\text{I.62})$$

ou k est le nombre de défaillances observées.

Afin de tenir compte de l'effet aléatoire associé à la simulation de Monte Carlo, nous avons répété 20 fois chaque estimation. Les figures I.10 et I.11 présentent l'estimation ponctuelle et l'écart-type des paramètres de forme v et d'échelle η en fonction de la taille d'échantillon. On peut observer que les estimations sont sensibles à la taille d'échantillon, dans la mesure où la variabilité d'estimation (estimateur ponctuel et écart-type) diminue avec l'augmentation de la taille d'échantillon. Cela était prévisible mais on observe, qu'au delà d'une taille d'échantillon de 35 produits, l'effet sur l'estimateur ponctuel reste constant. Les écart-types, utilisés pour les intervalles de confiance, diminue rapidement avec la taille d'échantillon.

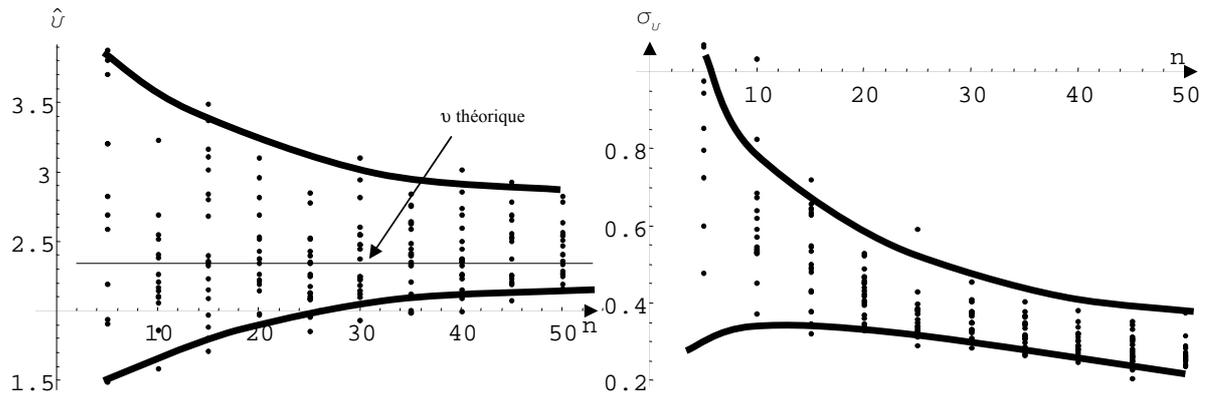


FIG. I.10 – Résultats du test Monte Carlo pour v et σ_v avec l'intervalle de 90% confiance

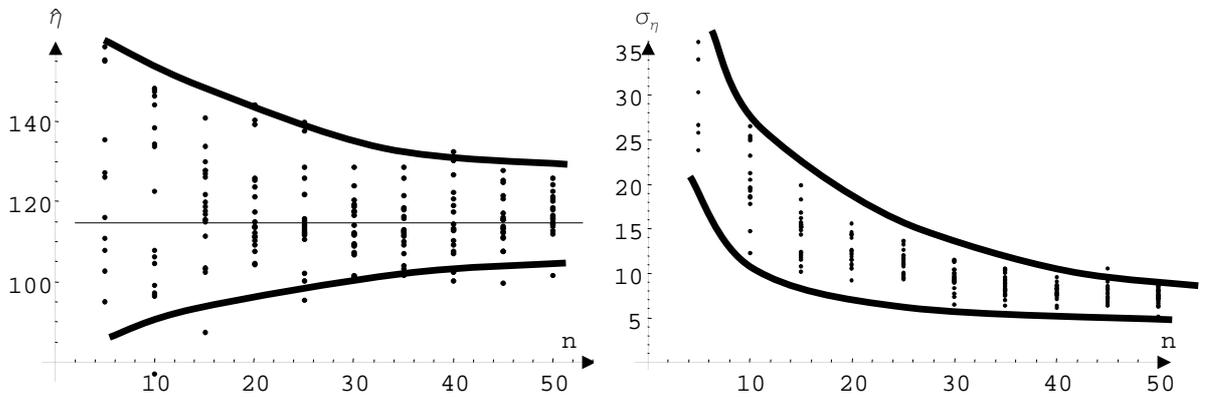


FIG. I.11 – Résultats du test Monte Carlo pour η et σ_η avec l'intervalle de 90% confiance

La méthode bayésienne non informative

A partir de la vraisemblance (équation I.29), nous déterminons les fonctions marginales sur les paramètres v et η :

$$f_\eta(\eta) = \int_0^\infty L(t|v, \eta) dv \quad (\text{I.63})$$

$$f_v(v) = \int_0^\infty L(t|v, \eta) d\eta \quad (\text{I.64})$$

Nous déduisons les estimateurs ponctuels \hat{v} , $\hat{\eta}$ par l'estimation du premier moment :

$$\hat{v} = \bar{v} = \int_0^{\infty} v f_v(v) dv \quad (\text{I.65})$$

$$\hat{\eta} = \bar{\eta} = \int_0^{\infty} \eta f_{\eta}(\eta) d\eta \quad (\text{I.66})$$

Les intervalles de confiance sont donnés par :

$$\hat{v} e^{-\frac{w_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{v})}}{\hat{v}}} < v < \hat{v} e^{\frac{w_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{v})}}{\hat{v}}} \quad (\text{I.67})$$

$$\hat{\eta} e^{-\frac{w_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\eta})}}{\hat{\eta}}} < \eta < \hat{\eta} e^{\frac{w_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{\eta})}}{\hat{\eta}}} \quad (\text{I.68})$$

en considérant les variances :

$$\sigma_v = \sqrt{\int_0^{\infty} f_v(v) (v - \hat{v})^2 dv} \quad (\text{I.69})$$

$$\sigma_{\eta} = \sqrt{\int_0^{\infty} f_{\eta}(\eta) (\eta - \hat{\eta})^2 d\eta} \quad (\text{I.70})$$

L'ensemble des intégrations a été effectué par une méthode numérique (Simpson). Comme pour le cas de la méthode MV, nous avons répété 20 fois chaque estimation afin de considérer l'effet aléatoire de la simulation. Les figures I.12 et I.13 représentent les différentes estimations en fonction de la taille d'échantillon. On observe une forte influence de la taille d'échantillon sur la forme de la fonction marginale $f_{\eta}(\eta)$ ce qui est moins le cas pour $f_v(v)$. Cela se traduit par une plus grande variabilité de l'estimation de l'écart-type du paramètre d'échelle σ_{η} . Sinon, le comportement global des estimations par la méthode bayésienne est identique à celui observé par la méthode MV.

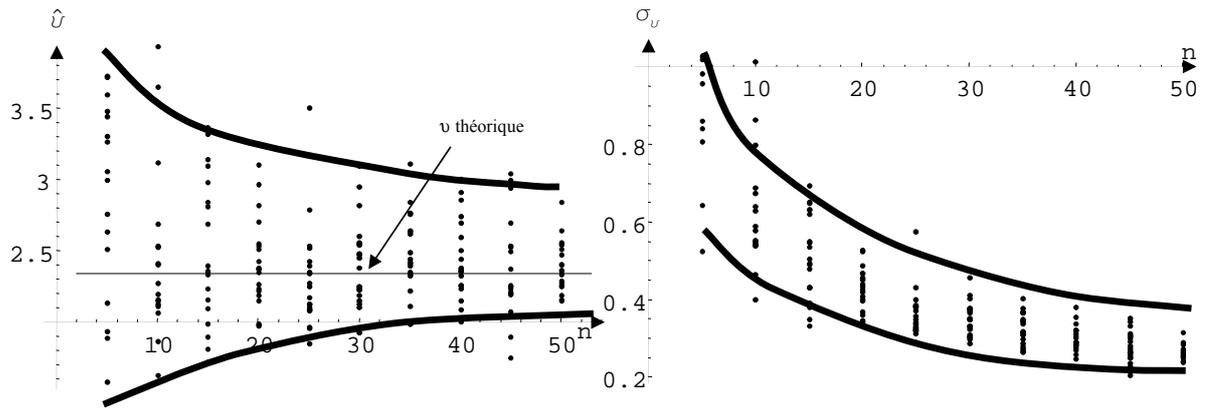


FIG. I.12 – Résultats du test Monte Carlo pour ν et σ_ν avec l'intervalle de 90% confiance

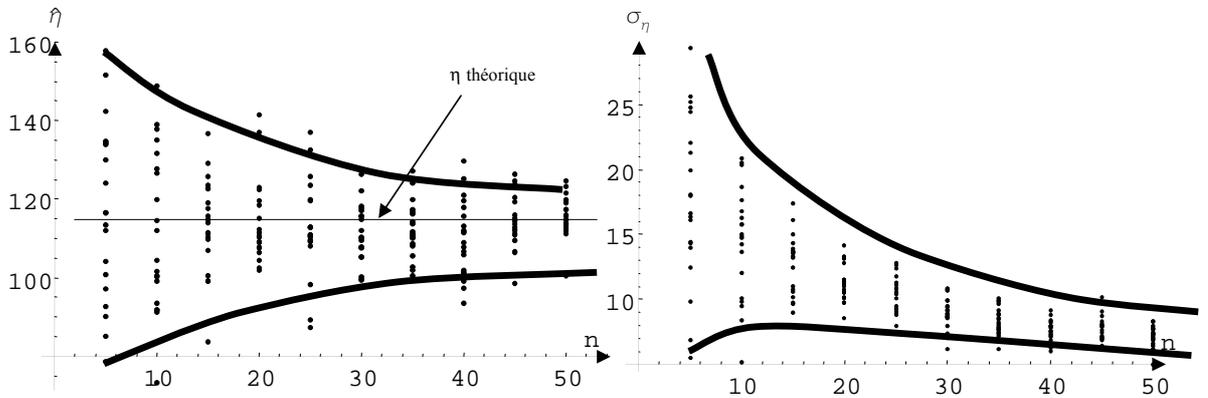


FIG. I.13 – Résultats du test Monte Carlo pour η et σ_η avec l'intervalle de 90% confiance

Comparaison des méthodes d'estimation

Pour chaque itération de la simulation de Monte Carlo, le même jeu de données a été traité par les méthodes d'estimation. Ainsi, les estimateurs ponctuels et les écart-types ont été calculés et les écarts relatifs ont été déterminés par la relation générale suivante :

$$\frac{|\theta_{max} - \theta_{MV}|}{Max(\theta_{max} - \theta_{MV})} \quad (I.71)$$

avec le terme $|\theta_{max} - \theta_{MV}|$ représentant l'écart pour l'itération i et le terme $Max(\theta_{max} - \theta_{MV})$ l'écart max sur l'ensemble des répétitions pour la taille d'échantillon n .

Les figures I.14 et I.15 montrent l'évolution des écarts relatifs des estimations en fonction de la taille d'échantillon.

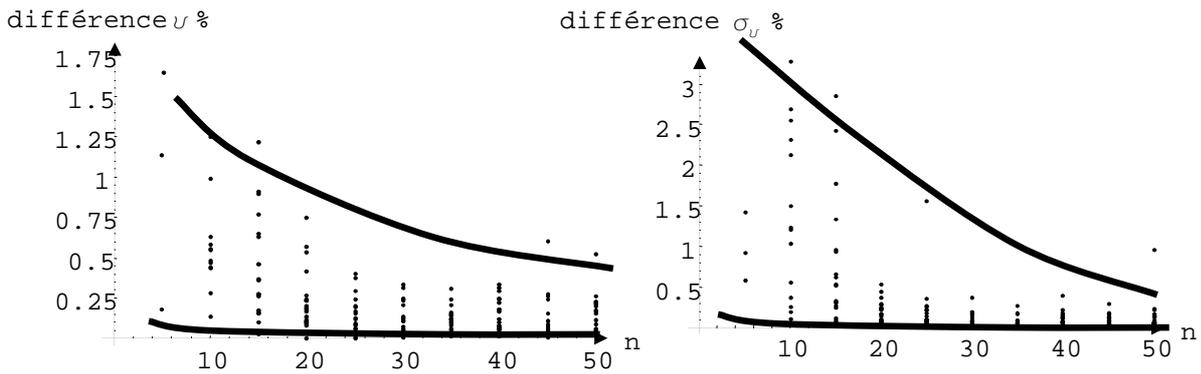


FIG. I.14 – Ecart relatif sur l'estimation de \hat{v} et σ_v

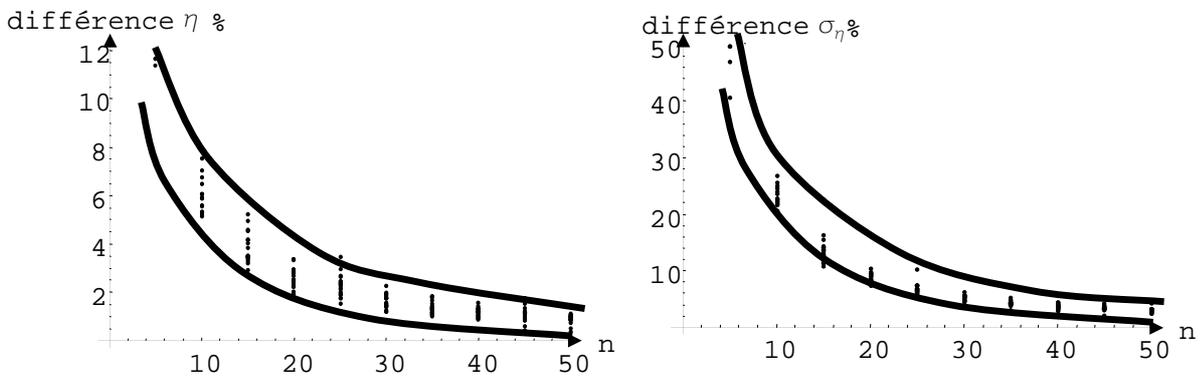


FIG. I.15 – Ecart relatif sur l'estimation de $\hat{\eta}$ et σ_η

Nous constatons des différences très faibles, d'ordre 1.75% entre les deux méthodes en ce qui concerne le paramètre v . Pour son écart-type, les différences restent dans la même ordre, jusqu'à 3%.

En ce qui concerne le paramètre d'échelle, η , les écarts sont plus significatifs pour de faibles tailles d'échantillon. C'est surtout vrai pour l'estimation de l'écart-type σ_η . Cela s'explique par l'effet de la taille d'échantillon sur la forme de la fonction marginale qui aboutit à ces écarts.

I.4.1.3 Comparaison des modèles paramétrique et non paramétrique

Pour les quatre tailles d'échantillon, nous avons comparé les fonctions de fiabilité et les intervalles de confiance (niveau de confiance $1-\alpha = 95\%$) obtenus par les modèles paramétrique ou non. La figure I.16 représente le résultat des quatre configurations.

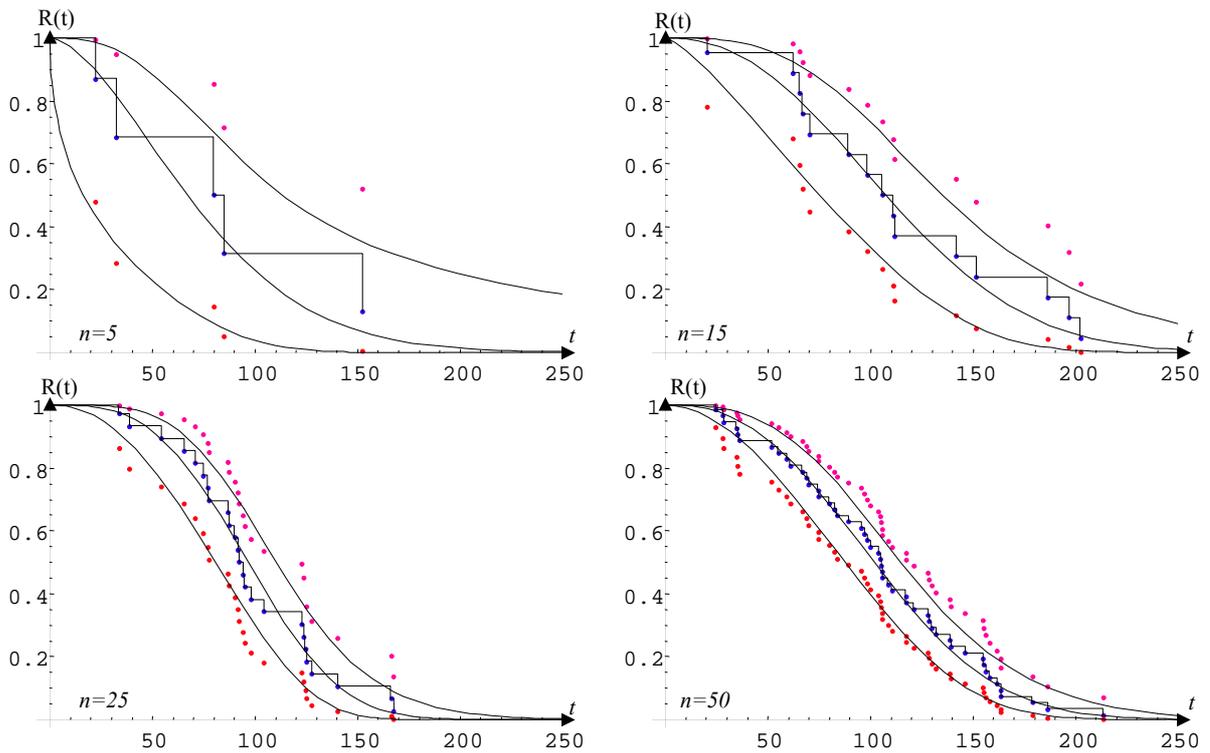


FIG. I.16 – Comparaison des méthodes

On peut noter, comme dans les analyses précédentes, l'influence de la taille d'échantillon sur la construction de la fonction de fiabilité. On peut remarquer que les fonctions de fiabilité et les intervalles de confiance par les deux approches sont comparables. Ainsi, pour de faible taille d'échantillon, le modèle paramétrique permet de faire des estimations plus "lissées" en évitant les sauts visibles sur le modèle non paramétrique. Par contre, dans le cas de taille d'échantillon plus importante, les deux approches sont tout à fait comparables en terme d'utilisation avec un avantage pour le cas non paramétrique puisqu'aucune hypothèse n'est considérée.

I.4.2 Influence de la censure

Afin d'éviter l'influence combinée de la taille d'échantillon et de la censure, nous avons considéré des simulations d'échantillons d'une taille importante de 50. Nous avons traité trois niveaux de censure (aucune censure, censure à 100 mois et censure à 50 mois) en appliquant les modèles paramétrique et non paramétrique. La figure I.17 représente la répétition de 30 simulations où les courbes représentent les estimations ponctuelles des fonctions paramétrique (colonne de gauche) et non paramétrique (colonne de droite) de la fiabilité.

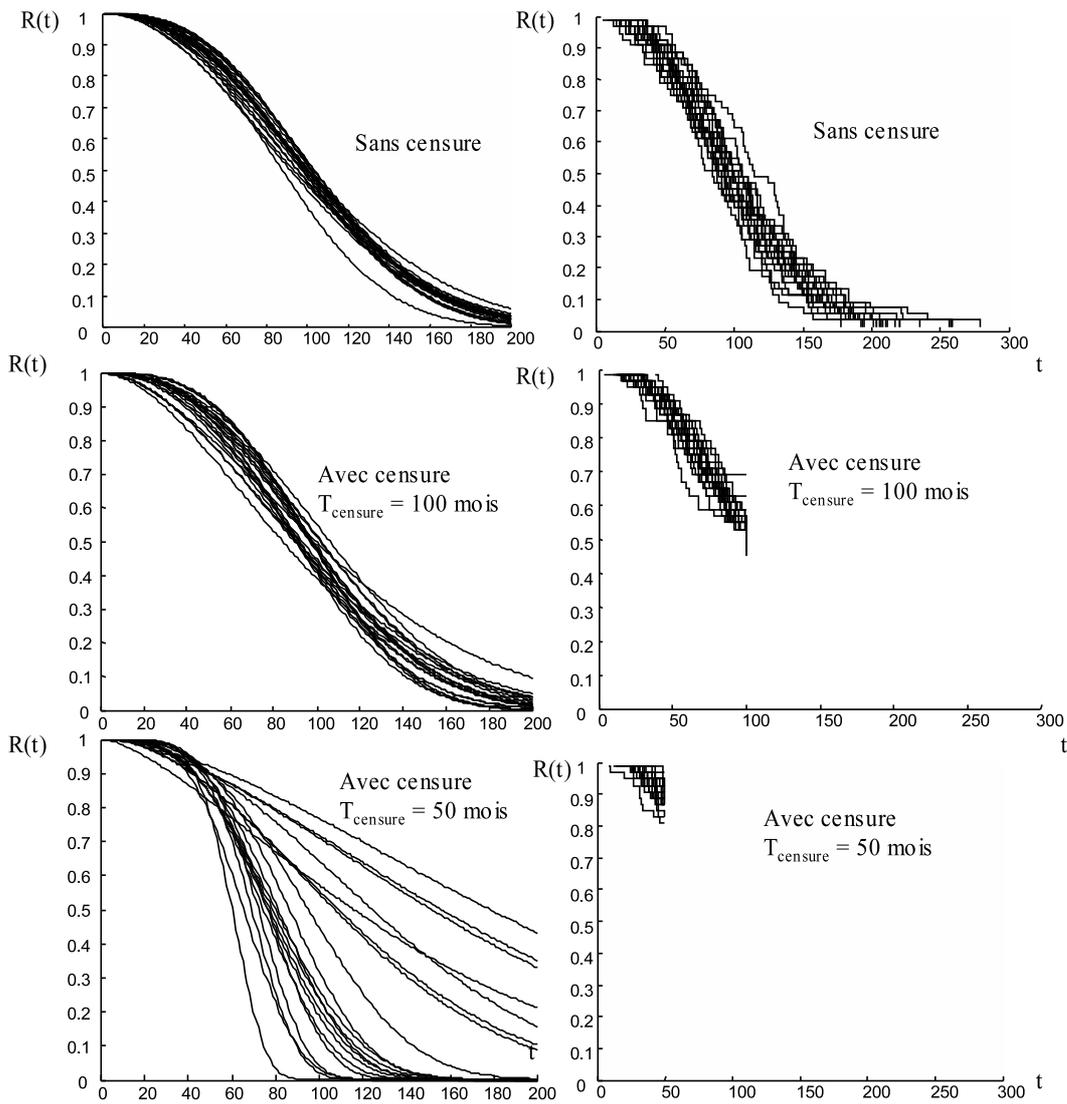


FIG. I.17 – L'estimation de la fiabilité dans un contexte de censure

On peut remarquer que plus le temps de censure diminue et plus la variabilité du modèle paramétrique est grande alors que le modèle non paramétrique reste insensible. Cette variabilité s'explique par le fait que c'est principalement les instants de défaillance qui apportent l'information dans l'estimation. Par contre, le modèle non paramétrique est tronqué au delà du temps de censure. Cela ne permet pas de prédire le comportement du produit au delà de ce temps. Ainsi, l'ingénieur ne peut pas accéder à certaines informations comme le MTTF, la fiabilité pour une période d'exploitation longue, ... ce qui peut être problématique.

I.4.3 Analyse globale

En examinant l'ensemble des simulations, on peut dresser quelques remarques générales :

- la taille d'échantillon impacte les modèles paramétriques et non paramétriques en augmentant les intervalles de confiance associés lorsque la taille diminue,
- les modèles non paramétriques deviennent grossiers avec la faiblesse de la taille d'échantillon mais ils présument aucune hypothèse,
- les estimations ponctuelles des modèles paramétriques sont très sensibles au niveau de censure,
- la troncature des modèles non paramétriques au delà du temps de censure peut être problématique dans certaines applications

En considérant ces remarques, deux voies de développement ont été proposées afin de limiter ces effets. Une première voie consiste à sévérer les conditions d'essai pour précipiter suffisamment de défaillance afin d'améliorer les estimations. Pour permettre d'estimer la fonction de fiabilité en conditions nominales, il est nécessaire d'utiliser un modèle de vie accélérée intégrant la fonction de fiabilité (paramétrique, non paramétrique, ...) et une loi, appelée couramment loi d'accélération, prenant en compte l'effet des variables de sévérisation sur la durée de vie du produit. Une deuxième voie étudiée est celle de l'estimation bayésienne consistant à introduire dans l'inférence statistique toute la connaissance a priori sur la fiabilité du produit. Dans le paragraphe suivant nous détaillons la méthode des essais accélérés.

I.5 Les essais accélérés

Les essais accélérés se composent d'une variété de techniques d'essai pour diminuer la durée de vie des produits ou accélérer les processus de dégradation. Ils sont utilisés pour obtenir plus vite des informations concernant la vie du produit ou concernant le rapport performance/dégradation dans le temps. Les unités testées sont employées plus fréquemment que d'habitude ou sont soumises à des niveaux de sollicitation plus élevés (par exemple l'augmentation de la température, de la tension électrique, de l'humidité, etc.), pour obtenir la loi de fiabilité ou autres caractéristiques de fiabilité (taux de défaillance, temps de défaillance, etc.) des produits (systèmes ou composants) dans des délais plus courts. Pour cela, les niveaux de stress subis par le produit sont augmentés afin d'obtenir des données de vie plus rapidement, qui seront utilisées pour estimer la fiabilité dans les conditions normales de fonctionnement. Les résultats sont employés, par un modèle statistique approprié basé sur la physique de défaillance des composants, pour faire des prévisions de durées de vie du produit soumis aux conditions normales d'utilisation.

Les essais accélérés peuvent s'appliquer à toutes les catégories de matériels en employant différents types de contraintes (mécaniques, électriques, climatiques ...) les plus adaptées vis-à-vis des modes de défaillance attendus sur ces matériels /

- Contraintes mécaniques : torsion, flexion, flexion rotative, chocs mécaniques, vibrations, vibrations acoustiques, traction/compression, etc. La fatigue est le mode de dégradation le plus communément utilisé pour les composants mécaniques qui sont testés sur des pots vibrants,
- Contraintes électriques : cela inclue la tension, l'intensité du courant, la fréquence, choc électrique ... La tension est le stress électrique le plus communément utilisé puisqu'il génère rapidement des défaillances,
- Contraintes climatiques (ou environnementales) : la température et les cycles thermiques sont les stress les plus couramment utilisés. Il est nécessaire d'appliquer des niveaux appropriés pour conserver les modes de défaillance d'origine. D'autres stress peuvent être appliqués comme les ultraviolets, le brouillard salin, la poussière, l'humidité, etc.
- Contraintes spécifiques au matériel.

Ces différentes contraintes peuvent être appliquées combinées ou non aux produits. Dans le cadre général, les stress X peuvent varier au cours du temps et être multidimensionnels [Nelson, 1990] :

$$X = X(\tau), \tau \geq 0 \tag{I.72}$$

I.5.1 Le plan d'essais

La définition d'un plan d'essai accéléré dépend de plusieurs paramètres ([Hoang, 2003], [Nelson, 1990], [Vassilious et Mettas, 2001],...) :

- Les stress d'accélération et les limites opérationnelles : on appelle stress l'ensemble des conditions et facteurs susceptibles d'affecter le bon fonctionnement d'un produit. Les stress peuvent être de toute nature (mécanique, électronique, climatique, ...) et leurs durées de manifestation de tout ordre (constante, échelonnée, progressive, cyclique ou aléatoire). Les nombres et les niveaux des stress appliqués sont choisis en fonction du produit étudié avec la conservation du mécanisme de défaillance d'origine tout en accélérant suffisamment l'essai. Les stress sont parfois désignés par les termes : contraintes et sollicitations. Les limites opérationnelles du produit sont déterminées par un essai aggravé par exemple (donnant les niveaux de stress extrêmes à ne pas dépasser afin d'éviter les fonctionnements dégradés des produits) ou changement de mécanisme de défaillance,
- Les modes et mécanismes de défaillance : dans un essai accéléré, les mécanismes d'endommagement provoqués doivent être représentatifs des conditions normales d'emploi. Chaque mode de défaillance peut être provoqué par un ou plusieurs types de contraintes,
- Les nombres de dispositifs testés à chaque niveau de contrainte, donnant la précision des estimations,
- Un modèle de vie accélérée, qui relie la durée de vie sous les conditions accélérées à celles sous les conditions normales d'utilisation, permet d'analyser les résultats d'essai pour estimer la fonction de fiabilité dans les conditions nominales.

La conception du plan d'essai (quels essais faut-il faire pour montrer que le système est fiable ?) peut intervenir dès le début de la conception du système, dès que les fonctions requises (le cahier des charges) sont connues.

Les stress peuvent être appliqués sur les produits selon différents profils :

1. Le chargement constant (indépendant du temps) : on soumet chaque composant à un niveau de contrainte constant supérieur à la normale (figure I.18),
2. Le chargement variable (en fonction du temps) : la contrainte choisie, évolue en fonction du temps. Elle peut être augmentée d'une manière échelonnée, progressive, cyclique ou aléatoire etc. :
 - La contrainte échelonnée : la contrainte est appliquée de manière échelonnée par des niveaux croissants ou décroissants dans le temps (par paliers) jusqu'à l'apparition d'une défaillance (figure I.19),

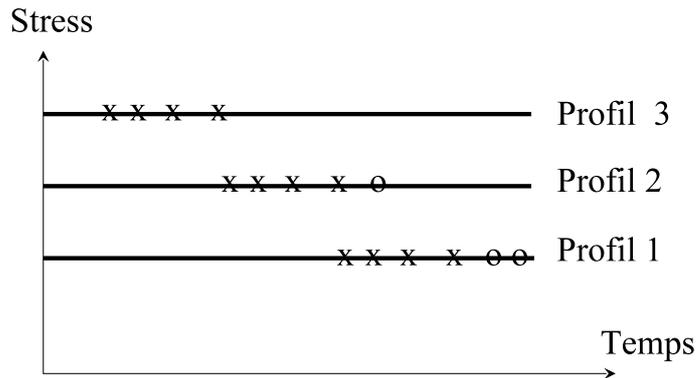


FIG. I.18 – Profil d'essai sous stress constant

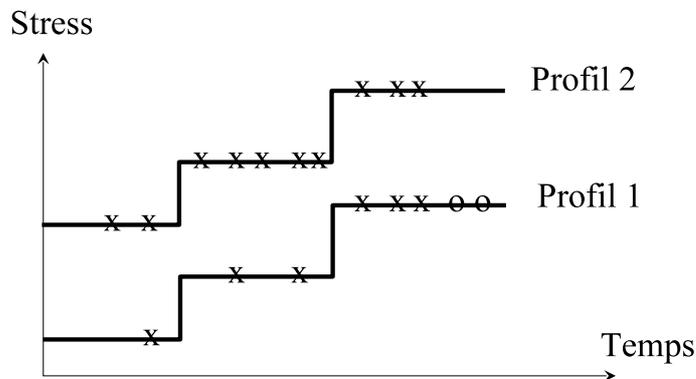


FIG. I.19 – Profil d'essai sous stress échelonné

- La contrainte progressive : la contrainte est augmentée de manière linéaire dans le temps (croissance linéaire),
- La contrainte cyclique : la contrainte est appliquée selon une amplitude et une fréquence données (figure I.20),
- La contrainte aléatoire : la sollicitation suit un profil aléatoire (couramment utilisée en vibration)

Dans le cadre de cette thèse, nous allons utiliser le plan d'essais par chargement constant [Nelson, 1990] avec des données censurées à droite, des niveaux de stress X_1, \dots, X_k constants accélérés : $X_1 < X_2 < \dots < X_k$, avec X_0 le stress usuel, ($X_0 < X_1$). On considère que k groupes d'unités sont observés. On teste le i -ème groupe sous le stress $X_i \in \varepsilon_0$ jusqu'à l'instant T_i . Le stress usuel X_0 n'est pas utilisé.

A chaque niveau de stress X_i on teste n_i produits et on note les temps de défaillance $t_{i1}, \dots, t_{iq_i}, q_i \leq n_i$, produits testés jusqu'à l'instant T_i ; cela veut dire qu'on a des essais censurés à droite. Dans la figure I.18 on a marqué avec **x** les temps de défaillance observés inférieurs à l'instant T_i et avec **o** les temps de défaillance possible supérieurs à l'instant

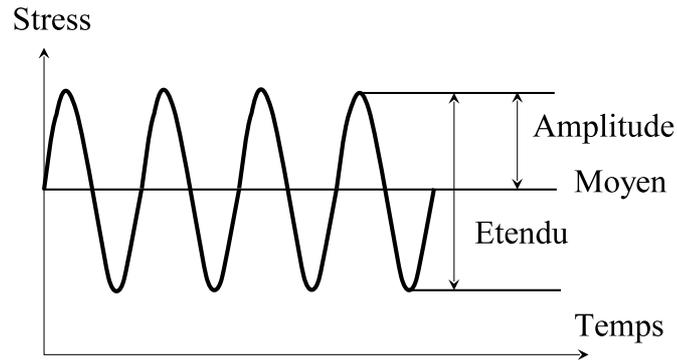


FIG. I.20 – Profil d'essai sous stress cyclique

T_i .

Ce plan de test est simple et les avantages sont les suivants :

- simplicité de solliciter les produits selon des stress constants,
- nombreux modèles disponibles pour traiter les résultats d'essai,
- Plusieurs techniques d'estimation peuvent être utilisées.

Dans le cadre de cette thèse, nous proposons d'utiliser le modèle standard de vie accélérée qui a été très largement étudié par l'équipe de M. Nikulin ([Bagdonavicius et Nikulin, 1995b], [Bagdonavicius et Nikulin, 1995a], [Bagdonavicius et Nikulin, 1997b], [Bagdonavicius et al., 2000], [Bagdonavicius et Nikulin, 2001], [Bagdonavicius et Nikulin, 2001])

I.6 Le modèle de vie accélérée (MVA)

Les modèles statistiques de vie accélérée sont généralement utilisés quand la relation exacte entre les contraintes appliquées et le temps de défaillance du composant est difficile à déterminer selon des principes physico-chimiques. Dans ce cas les composants sont soumis à différents niveaux de contraintes et les instants de défaillance sont utilisés pour ajuster la distribution de vie la plus adéquate. Les instants de défaillance ont le même type de loi à chaque niveau de contrainte, en particulier dans les conditions normales de fonctionnement.

Les modèles statistiques de vie accélérée peuvent s'appliquer à plusieurs domaines comme celui du vivant (sciences médicales), de l'électronique et de la mécanique. Ce qui différencie les diverses applications sont les lois de fiabilité utilisées, les stress employés pour aggraver les essais et la nature des lois d'accélération.

Dans la littérature, on trouve plusieurs définitions théoriques des modèles Statistiques de Vie Accélérée ([Nelson, 1990], [Bagdonavicius et Nikulin, 1995a], [Hoang, 2003]). En général, ils sont constitués de deux composantes principales, la figure I.21 illustre ce genre de modèles :

- un modèle analytique "Durée de vie-stress" appelé aussi loi d'accélération ou modèle d'accélération, traduisant la durée de vie "nominale" du produit soumis à l'essai en fonction des niveaux de contraintes appliquées. Cette durée de vie "nominale" est représentée par une caractéristique de la loi de fiabilité telle que la moyenne, la médiane, l'écart type, un quantile ou un quelconque paramètre de la loi.
- une distribution statistique des durées de vie : Dans un test accéléré, un modèle analytique seul, ne décrit pas le comportement des durées de vie du produit. À chaque niveau de contrainte (stress) le système a une distribution statistique de vie. D'où la combinaison : équation d'accélération et distribution de vie de base.

Devant une telle variété de modèles, un grand intérêt apparaît sur l'unification de ces modèles dans un seul formalisme. C'est le but du modèle standard de vie accélérée.

I.6.1 Principe

Supposons que la durée de vie $T_{X(\bullet)}$ sous n'importe quel stress $X(\bullet)$ est une variable aléatoire non négative absolument continue de fonction de survie :

$$R_{X(\bullet)}(t) = P \{T_{X(\bullet)} > t\}, t \geq 0 \quad (I.73)$$

Considérons un ensemble de stress ε . Formellement, on dit qu'un stress $X_1(\bullet)$ est

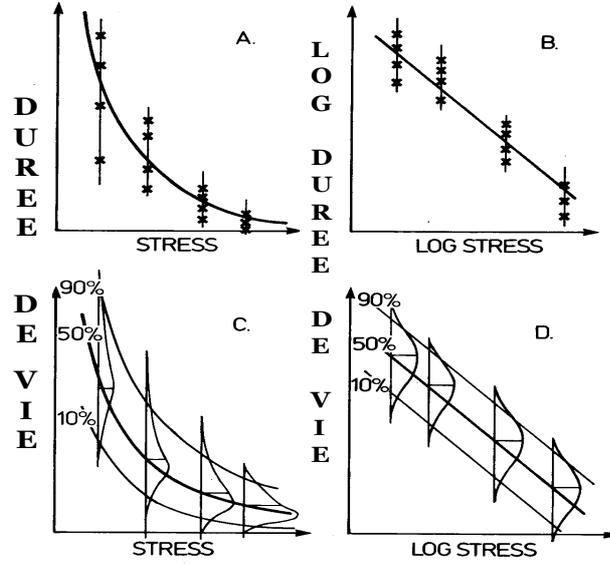


FIG. I.21 – Représentations graphiques d'un Modèle Statistique de vie Accéléré (durée de vie en fonction de la contrainte, tirée de [Nelson, 1990])

supérieur à un stress $X_0(\bullet)$ si $R_{X_0(\bullet)}(t) > R_{X_1(\bullet)}(t)$ pour tout $t \geq 0$. Soit $\varepsilon_0 \subset \varepsilon$ un ensemble de stress constant dans le temps et $X_0 \in \varepsilon_0$. On peut définir la fonction inverse de $R_{X_0}(t)$ comme :

$$R_{X_0}^{-1}(p) = \inf \{t : R_{X_0}(t) \geq p\} \quad (\text{I.74})$$

On note le produit de convolution entre $R_{X_0}^{-1}$ et $R_{X(\bullet)}$ comme $f_{X(\bullet)}(t) = R_{X_0}^{-1} R_{X(\bullet)}(t)$, avec $f_{X(\bullet)}(0) = 0$. Alors, pour tout $X(\bullet) \in \varepsilon$ on a :

$$P \{T_{X_0} \geq f_{X(\bullet)}(t)\} = P \{T_{X(\bullet)} \geq t\} \quad (\text{I.75})$$

Il résulte que pour tout $X(\bullet) \in \varepsilon$ la probabilité de survivre jusqu'à l'instant t sous le stress $X(\bullet)$ est la même que la probabilité de survivre jusqu'à l'instant $f_{X(\bullet)}(t)$ sous le stress X_0 . On appelle la fonction $f_{X(\bullet)}(t)$ la *ressource utilisée sous le stress $X(\bullet)$ jusqu'à l'instant t* ou *fonction de transfert*.

Le *modèle standard de vie accélérée* (SVA) est défini sur ε s'il existe une fonction $r : \varepsilon \rightarrow \mathfrak{R}^+$ telle que pour tout $X(\bullet) \in \varepsilon$:

$$\frac{d}{dt} f_{X(\bullet)}(t) = r[X(t)] \quad (\text{I.76})$$

La fonction r exprime une vitesse de dégradation. Le modèle SVA signifie que la vitesse d'utilisation de la ressource à l'instant t ne dépend que de la valeur du stress appliqué à

l'instant t .

La formule I.76 implique :

$$R_{X(\bullet)}(t) = R_{X_0} \left(\int_0^t r[X(\tau)] d\tau \right) \quad (I.77)$$

Dans le cas où $X(\tau) \equiv X = \text{constant}$, le modèle I.77 implique :

$$R_X(t) = R_{X_0}(r(X)t) \quad (I.78)$$

Donc le stress ne change que l'échelle. Notons que $r(X_0) = 1$.

Le modèle SVA permet d'estimer la fiabilité d'un composant sous des contraintes expérimentalement indisponibles à partir de la fiabilité du même composant sous des conditions d'utilisation. L'analyse statistique du modèle SVA consiste à estimer les paramètres de ce modèle (fiabilité, taux de défaillance, ...). Les techniques d'estimation, dépendent de ce modèle statistique, c'est-à-dire de la connaissance de la fonction de survie R_{X_0} et du choix de la vitesse de dégradation r .

Si la fonction r est complètement inconnue, la fonction R_{X_0} ne peut pas être estimée même si l'on connaît la famille de distributions à laquelle R_{X_0} appartient.

Par exemple, si $R_{X_0}(t) = e^{[-(\frac{t}{\theta})^\alpha]}$ on obtiens $R_{X_i}(t) = e^{[-(\frac{r(X_i)t}{\theta})^\alpha]}$. A partir de temps de défaillance on peut estimer le paramètres $\alpha, \frac{r(X_1)}{\theta}, \dots, \frac{r(X_k)}{\theta}$ mais puisque r est complètement inconnue, $r(X_0)$ et donc $S_{X_0}(t)$ ne peuvent pas être estimés.

Donc la fonction r doit être choisie dans une certaine classe de fonctions. Si le modèle I.78 est vérifié sur un ensemble de stress ε , alors, pour tous X_i, X_j on a

$$R_{X_j}(t) = R_{X_i} \left(\frac{r(X_j)}{r(X_i)} t \right) = R_{X_i}(\rho(X_i, X_j)t) \quad (I.79)$$

D'abord, supposons que le stress $x \in \mathfrak{R}$ soit unidimensionnel. Le taux de changement d'échelle est déterminé par la dérivée :

$$\delta(x) = \lim_{\Delta x \rightarrow 0} \frac{\rho(x, x + \Delta x) - \rho(x, x)}{\Delta x} = [\log r(x)]' \quad (I.80)$$

Donc pour tout $x \in \varepsilon$

$$r(x) = e^{\int_{x_0}^x \delta(v) dv} \quad (I.81)$$

Supposons que $\delta(x)$ soit proportionnelle à une fonction connue $u(x)$ du stress

$$\delta(x) = \beta_1 u(x), \beta_1 > 0 \quad (\text{I.82})$$

Alors

$$r(x) = e^{\beta_0 + \beta_1 z(x)} \quad (\text{I.83})$$

où $z(x)$ est une fonction primitive de $u(x)$, β_0, β_1 sont des paramètres inconnus.

Pour généraliser, si le stress est multidimensionnel : $X = (x_1, \dots, x_m)^T$, on considère des caractéristiques infinitésimales $\delta_i(x)$ données par :

$$\delta_i(X) = \lim_{\Delta x_i \rightarrow 0} \frac{\rho(X, X + \Delta x_i e_i) - \rho(X, X)}{\Delta x_i} = \frac{\partial \log r(X)}{\partial x_i} \quad (\text{I.84})$$

où $e_i = \begin{pmatrix} 0, \dots, 1, \dots, 0 \\ 1 \quad i \quad m \end{pmatrix}^T$. L'unité est dans la i -ème coordonnée. Généralisant le cas unidimensionnel, $\delta_i(X)$ devient :

$$\delta_i(X) = \sum_{j=1}^{k_i} \beta_{ij} u_{ij}(X) \quad (\text{I.85})$$

où les $u_{ij}(X)$ sont des fonctions connues et β_{ij} sont des constantes inconnues. Dans ce cas

$$r(X) = e^{\beta_0 + \sum_{i=1}^m \sum_{j=1}^{k_i} \beta_{ij} z_{ij}(X)} \quad (\text{I.86})$$

avec $z_{ij}(X)$ des fonctions connues.

Alors, en généralisant, les modèles I.77 et I.78 peuvent s'écrire :

$$R_{X(\bullet)}(t) = R_{X_0} \left(\int_0^t e^{\beta^T Z(\tau)} d\tau \right) \quad (\text{I.87})$$

où, pour des stress constants :

$$R_X(t) = R_{X_0} \left(e^{\beta^T Z t} \right) \quad (\text{I.88})$$

où $\beta = (\beta_0, \dots, \beta_m)^T$ un vecteur de paramètres et $Z = (\varphi_0(X), \dots, \varphi_m(X))^T$ avec le fonctions φ_i spécifiées et la première composante z_0 égale à 1.

Dans la littérature ([Nelson, 1990],...) on trouve une grande variété de modèles d'accélération, définis pour chaque type de composants et matériaux. Nous présentons quelques modèles d'accélération les plus utilisés pour des stress à profil constants, des cas particuliers de MVA. Désormais, la durée de vie " nominale " est représentée par la moyenne.

I.6.2 Lois d'accélération courantes

I.6.2.1 Loi d'Arrhenius

Utilisée lorsque le mécanisme d'endommagement est sensible à la température (en particulier : diélectrique, semi-conducteur, batterie, lubrifiant et graisse, plastique et filament de lampe incandescente). La loi d'Arrhenius modélise la durée de vie τ du produit comme fonction de la température T :

$$\tau = Ae^{\left(\frac{+B}{T}\right)} \quad (\text{I.89})$$

avec :

- A et B constantes positives
- T température absolue

Le facteur d'accélération (FA) d'Arrhenius entre la durée de vie τ_1 pour une température T_1 et la durée de vie τ_2 pour une température T_2 est :

$$FA = \frac{\tau_1}{\tau_2} = e^{B\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} \quad (\text{I.90})$$

I.6.2.2 Loi de Peck

Utilisée lorsque le mécanisme d'endommagement est sensible à la température T et à l'humidité H (en particulier : composants électriques, conducteur aluminium et composants mécaniques soumis à la rupture). Le modèle est défini par :

$$\tau = A(H)^{-m} e^{\left(\frac{E_a}{kT}\right)} \quad (\text{I.91})$$

avec

- T : température absolue en °K
- H : humidité en %
- k : $8.6171 \cdot 10^{-5}$ eV/K constante de Boltzmann
- E_a : énergie d'activation en eV
- m paramètre du modèle

I.6.2.3 Loi puissance inverse

Elle est utilisée lorsque le mécanisme d'endommagement est sensible à un stress particulier (par exemple : diélectrique, roulement à billes ou rouleaux, composants optoélectroniques, composants mécaniques soumis à la fatigue et filament de lampe incandescente).

La loi de puissance inverse décrit la cinétique d'une dégradation sous stress constant V , la durée de vie est donnée par l'équation :

$$\tau = \left(\frac{A}{V}\right)^\gamma \quad (\text{I.92})$$

avec :

- V : stress constant
- A et γ : constantes dépendantes de la défaillance et de l'essai

Cas particulier : quand le stress est la température, la loi est appelée la loi de Coffin-Manson.

Le facteur d'accélération entre la durée de vie τ_1 pour un stress V_1 et la durée de vie τ_2 pour un stress V_2 est :

$$FA = \frac{\tau_1}{\tau_2} = \left(\frac{V_2}{V_1}\right)^\gamma \quad (\text{I.93})$$

I.6.2.4 Loi d'Eyring

Utilisée lorsque le mécanisme d'endommagement est sensible à la température et un stress V (en particulier : composants électriques, conducteur aluminium et composants mécaniques soumis à la rupture). Le modèle est défini par :

$$\tau = \left(\frac{A}{T}\right) e^{\left(\frac{B}{kT}\right)} e^{(V(C+\frac{D}{kT}))} \quad (\text{I.94})$$

avec :

- T : température absolue
- V : Stress donné
- k : $8.6171 \cdot 10^{-5}$ eV/K constante de Boltzmann
- A, B, C et D les paramètres du modèle

Dans les paragraphes suivants, nous abordons l'estimation du modèle SVA selon les approches paramétrique et semi-paramétrique.

I.6.3 La construction d'un modèle paramétrique

Les modèles paramétriques sont généralement employés lorsqu'il existe une connaissance suffisante sur les lois de dégradation et les fonctions de fiabilité appropriées au produit à tester. Ainsi, le plan d'essai consiste à solliciter les composants à différents niveaux accélérés de stress. On fait l'hypothèse que la distribution des temps de défaillance est identique aux différents niveaux de stress (y compris dans les conditions initiales) et

que le paramètre de forme de la distribution soit le même pour tous les niveaux de stress. Les temps de défaillance aux différents niveaux de stress sont utilisés pour déterminer la distribution de probabilité la plus appropriée [Shyur et al., 1999].

En reprenant l'équation I.25, on peut noter :

$R(t) = R(e^t)$, $t \in \mathfrak{R}$, $\sigma = \frac{1}{v}$, $v = (v_0, \dots, v_m)$, $\gamma_0 = \ln \eta - \beta_0$ et $\gamma_i = -\beta_i$, $i = (1, \dots, m)$, on peut écrire le modèle I.87 sous la forme :

$$R_X(t) = R\left(\frac{\ln t - \gamma^T Z}{\sigma}\right) \quad (\text{I.95})$$

Notons T_{ij} la durée de vie observée ou censurée de la j -ème unité du i -ème groupe, on a :

$T_{ij} = \ln(T_{ij} \wedge t_i)$, $\delta_{ij} = I\{T_{ij} \leq t_i\}$, $f(u) = -S'$, $\lambda(u) = \frac{f(u)}{S(u)}$ avec \wedge le minimum entre deux valeurs.

Dans ce cas, la fonction de vraisemblance est :

$$L(\gamma, \sigma | T) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda\left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}\right) \right]^{\delta_{ij}} R\left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}\right) \quad (\text{I.96})$$

En dérivant par rapport à γ_i et σ la fonction $\ln L(\gamma, \sigma)$, on obtient :

$$\begin{aligned} U_l(\gamma, \sigma) &= \frac{\partial \ln L(\gamma, \sigma)}{\partial \gamma_l} = \frac{1}{\sigma} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} a_{ij}(\gamma, \sigma), \quad (l = 1, \dots, m), \\ U_{m+1}(\gamma, \sigma) &= \frac{\partial \ln L(\gamma, \sigma)}{\partial \sigma} = \frac{1}{\sigma} \sum_{i=1}^k \sum_{j=1}^{n_i} \{v_{ij}(\gamma, \sigma) a_{ij}(\gamma, \sigma) - \delta_{ij}\}, \end{aligned} \quad (\text{I.97})$$

avec

$$v_{ij}(\gamma, \sigma) = \frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}, \quad a_{ij}(\gamma, \sigma) = \lambda(v_{ij}(\gamma, \sigma)) - \delta_{ij} (\ln \lambda)'(v_{ij}(\gamma, \sigma))$$

Les estimateurs du maximum de vraisemblance $\hat{\sigma}$, $\hat{\gamma}$ peuvent être obtenus en résolvant le système d'équations :

$$U_l(\gamma, \sigma) = 0 \quad (l = 1, \dots, m+1) \quad (\text{I.98})$$

et l'estimateur de la fonction de survie R_{X_0} est

$$\hat{R}_{X_0}(t) = R\left(\frac{\ln t - \hat{\gamma}^T Z^{(0)}}{\hat{\sigma}}\right) \quad (\text{I.99})$$

Pour calculer l'intervalle de confiance $(1-\alpha)$ sur l'estimateur de la fonction de fiabilité, on définit la matrice d'information de Fisher $I^L(\gamma, \sigma)$ avec les éléments suivants :

$$\begin{aligned}
 I_{ls}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} \\
 &= \frac{1}{\sigma^2} \sum_{i=1}^k z_{il} z_{is} \sum_{j=1}^{n_i} c_{ij}(\gamma, \sigma), \quad l, s = 0, \dots, m
 \end{aligned} \tag{I.100}$$

$$\begin{aligned}
 I_{l,m+1}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \gamma_l \partial \sigma} \\
 &= \frac{1}{\sigma} U_l(\gamma, \sigma) + \frac{1}{\sigma^2} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} v_{ij}(\gamma, \sigma) c_{ij}(\gamma, \sigma), \quad l = 0, \dots, m
 \end{aligned} \tag{I.101}$$

$$\begin{aligned}
 I_{m+1,m+1}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \sigma^2} \\
 &= \frac{2}{\sigma} U_{m+1}(\gamma, \sigma) + \frac{1}{\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (v_{ij}^2(\gamma, \sigma) c_{ij}(\gamma, \sigma) + \delta_{ij})
 \end{aligned} \tag{I.102}$$

avec : $c_{ij}(\gamma, \sigma) = \lambda'(v_{ij}(\gamma, \sigma)) - \delta_{ij}(\ln \lambda)''(v_{ij}(\gamma, \sigma))$.

On observe que l'estimateur \hat{R} est une fonction régulière de γ et σ . On va définir :

$$Q_{X_0}(t) = \ln \frac{R_{X_0}(t)}{1 - R_{X_0}(t)} \quad \text{et} \quad \hat{Q}_{X_0}(t) = \ln \frac{\hat{R}_{X_0}(t)}{1 - \hat{R}_{X_0}(t)} \tag{I.103}$$

La fonction $Q_{X_0}(t)$ prend ses valeurs dans \mathfrak{R} donc la vitesse de convergence de \hat{Q}_{X_0} vers la loi limite est plus grande que celle de \hat{R}_{X_0} vers sa loi limite, et on obtient que la loi de

$$\frac{\hat{Q}_{X_0}(t) - Q_{X_0}(t)}{\hat{\sigma}_{Q_0}}$$

est approximée par la loi normale $N(0, 1)$, avec :

$$\begin{aligned}
 (\hat{\sigma}_{Q_0})^2 &= \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) (I^L(\gamma, \sigma)^{-1}) \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) \\
 &= \left(\frac{R'(R^{-1}(\hat{R}_{X_0}(t)))}{\hat{\sigma}^2 \hat{R}_{X_0}(t) (1 - \hat{R}_{X_0}(t))} \right)^2 \times \left[\hat{\sigma}^4 \sum_{l=0}^m \sum_{s=0}^m z_{0l} z_{0s} I_L^{ls}(\hat{\gamma}, \hat{\sigma}) \right. \\
 &\quad \left. - 2\hat{\sigma}^2 \hat{\gamma}^T Z^{(0)} \sum_{l=0}^m z_{0l} I_L^{l,m+1}(\hat{\gamma}, \hat{\sigma}) + (\hat{\gamma}^T Z^{(0)})^2 I_L^{m+1,m+1}(\hat{\gamma}, \hat{\sigma}) \right]
 \end{aligned} \tag{I.104}$$

avec $I_L^{l,s}$ les termes de la matrice $(I^L(\gamma, \sigma))^{-1}$.

L'intervalle de confiance approximatif $(1-\alpha)$ pour $\hat{R}_{X_0}(t)$ est :

$$\left(1 + \frac{1 - \hat{R}_{X_0}(t)}{\hat{R}_{X_0}(t)} \exp \{ \pm \hat{\sigma}_{Q_0} w_{1-\alpha/2} \} \right)^{-1} \quad (\text{I.105})$$

avec w_α la α -quantile de la loi $N(0, 1)$.

Alors, à partir du modèle I.78 on peut créer une multitude de méthodes en choisissant différentes lois d'accélération Arrhenius, Eyring, Peck, Le plus souvent, ces modèles sont appliqués de la façon suivante : on suppose que la fonction $R_{X_0}(t)$ appartient à une certaine classe Γ de répartitions et on observe plusieurs groupes des composants sous des stress accélérés en conformité avec le plan d'essais classique présenté dans I.5.1. On estime alors les paramètres inconnus $\hat{\gamma}$ et $\hat{\sigma}$ de $r(X)$ et la fonction $R_{X_0}(t)$. Un estimateur de la fonction de fiabilité sous le stress "normal" X_0 est obtenu en changeant dans I.78 X par X_0 ainsi que les estimateurs des paramètres inconnus.

Les points faibles de cette méthodologie sont les suivants [Bagdonavicius et Nikulin, 1997a] :

1. impossibilité de vérifier l'hypothèse $H:G \in \Gamma$ dans le cadre de stress "normal" à partir de telles expériences
2. choix limité de la fonction $r(X)$ (log-linéaire en particulier)
3. erreur de prédiction du modèle de régression car le stress nominal X_0 n'est pas utilisé pendant les expériences.

Pour traiter les deuxième et troisième points, des plans ont été proposés, basés sur le modèle non-paramétrique par exemple, où l'on fait quasiment aucune hypothèse sur la classe d'appartenance de $r(X)$ [Tebbi, 2005] et de la fonction de fiabilité. Ces plans ne seront pas traités dans le cadre de cette thèse car ils ne présentent pas d'intérêt dans le contexte abordé. En effet, l'objectif dans le chapitre III est d'utiliser les modèles SVA complets (fonction defaibilité et loi d'accélération connues) pour estimer la fonction de fiabilité d'un produit soumis à un environnement aléatoire.

I.6.4 La construction d'un modèle semi-paramétrique

L'analyse semiparamétrique du modèle SVA dans le cas du plan d'essais classique a été considérée par [Bashu et Ebrahimi, 1982], [Schmoyer, 1986], [Schmoyer, 1991], [Nikulin, 2007].

Nous allons noter $N_i(\tau)$ le nombre de pannes observées du i -ème groupe dans l'intervalle $[0, \tau]$, $Y_i(\tau)$ le nombre d'unités à risque (en état de fonctionnement et non-censurés)

avant le moment τ , $T_{i1} \leq \dots \leq T_{im_i}$ les moments de pannes du i -ème groupe. Les variables aléatoires

$$e^{\beta^T Z_i T_{ij}}, \quad (i = 1, \dots, k; j = 1, \dots, m_i)$$

peuvent être considérées comme des pseudo-pannes observées dans une expérience où $n = \sum_{i=1}^k n_i$ unités avec la fonction de survie R_{x_0} ont été testés et $n_i - m_i$ parmi elles ont été censurées au moment $e^{\beta^T Z_i t_i}$. La fonction de fiabilité sera définie empiriquement par l'estimateur de Kaplan Meier. Alors

$$N^R(\tau, \beta) = \sum_{i=1}^k N_i \left(e^{-\beta^T Z_i \tau} \right)$$

est le nombre de pannes observées dans l'intervalle $[0, \tau]$ et

$$Y^R(\tau, \beta) = \sum_{i=1}^k Y_i \left(e^{-\beta^T Z_i \tau} \right)$$

est le nombre des unités à risque avant l'instant τ .

La fonction de survie R_{X_0} peut être estimée par l'estimateur de Kaplan-Meier. Nous allons noter $\Delta N^R(\tau, \beta) = N^R(\tau, \beta) - N^R(\tau-, \beta)$ et $\Delta N_l(u) = N_l(u) - N_l(u-)$, avec la notation générique $f(u-) = \lim_{\varepsilon \downarrow 0} f(u - \varepsilon)$; alors, pour tous $s \leq \max_i \{e^{\beta^T Z_i t_i}\}$, on a :

$$\tilde{R}_{X_0}(s, \beta) = \prod_{\tau \leq s} \left(1 - \frac{\Delta N^R(\tau, \beta)}{Y^R(\tau, \beta)} \right) = \prod_{\tau \leq s} \left(1 - \frac{\sum_{l=1}^k \Delta N_l \left(e^{-\beta^T Z_l \tau} \right)}{\sum_{l=1}^k Y_l \left(e^{-\beta^T Z_l \tau} \right)} \right) \quad (\text{I.106})$$

On écrit de façon suivante :

$$\tilde{R}_{X_0}(s, \beta) = \prod_{(i,j): T_{ij} \leq \exp\{-\beta^T Z_i\}s} \left(1 - \frac{1}{\sum_{l=1}^k Y_l \left(e^{-\beta^T (Z_i - Z_l) T_{ij}} \right)} \right) \quad (\text{I.107})$$

La fonction de vraisemblance est présentée dans [Bagdonavicius et al., 2000] comme :

$$L(\beta) = \prod_{i=1}^k \prod_{j=1}^{m_i} \left[\tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij-}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right] \tilde{R}_{X_0}^{n_i - m_i} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \quad (\text{I.108})$$

Les auteurs justifient cette formulation par le fait que la fonction de densité $f_{X_0} = -S'_{X_0}$ est inconnue et peut être approchée aux points $e^{\beta^T Z_i T_{ij}}$ par un facteur proportionnel à :

$$f_{X_0} \left(e^{\beta^T Z_i T_{ij}}, \beta \right) \approx \tilde{R}_{X_0} \left(e^{\beta^T Z_i T_{ij-}}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i T_{ij}}, \beta \right) \quad (\text{I.109})$$

Mais, on sait que :

$$\frac{d \left[\tilde{R}(g(u)) \right]}{du} = g'(u) \tilde{R}'(g(u)) \quad (\text{I.110})$$

Alors, on a :

$$\frac{d \left[\tilde{R} \left(e^{\beta^T Z_i t}, \beta \right) \right]}{dt} = e^{\beta^T Z_i} f_{X_0} \left(e^{\beta^T Z_i t}, \beta \right) \quad (\text{I.111})$$

et on va proposer de remplacer l'équation I.108 par :

$$L(\beta) = \prod_{i=1}^k \prod_{j=1}^{m_i} \left[e^{\beta^T Z_i} \left(\tilde{R}_{X_0} \left(e^{\beta^T Z_i T_{ij-}}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i T_{ij}}, \beta \right) \right) \right] \tilde{R}_{X_0}^{n_i - m_i} \left(e^{\beta^T Z_i T_{ij}}, \beta \right) \quad (\text{I.112})$$

Il nous reste simplement à calculer l'estimateur ponctuel :

$$\hat{\beta} = \text{Argmax}_{\beta} L(\beta) \quad (\text{I.113})$$

pour avoir la fonction de survie sous stress normal pour tout $s \leq \max_i \left\{ e^{\hat{\beta}^T Z_i t_i} \right\}$ comme :

$$\hat{R}_{X_0}(s) = \tilde{R}_{X_0} \left(s, \hat{\beta} \right) \quad (\text{I.114})$$

Dans le paragraphe suivant, nous allons appliquer les modèles paramétrique et semi paramétrique sur un exemple simple à un seul stress. Dans le cadre de l'équation I.88, le vecteur de paramètres est unidimensionnel, $\beta = \{\beta_1\}$, en considérant aussi $\beta_0 = 0$. Cela nous va permettre d'afficher en deux dimensions le graphe de transfert de la durée de vie de différents niveaux $X_i, \{i = 1..k\}$ au niveau de stress usuel X_0 . Dans ce contexte, on obtient le $r(X)$ défini dans l'équation I.76 comme :

$$r(X) = e^{\beta_1 Z} \quad (\text{I.115})$$

avec $Z = \varphi_1(X)$ décrivant l'influence du stress.

Le caractère discret de la construction de fonction de fiabilité se retrouve dans la forme

de la fonction de vraisemblance par des discontinuités.

I.7 Application numérique

L'objectif de cet exemple est d'appliquer les modèles paramétrique et semi paramétrique à des données simulées. Pour ce faire, nous avons simulé les résultats d'un plan d'essai complet (sans censure) à deux niveaux de température ($T_1=95^\circ\text{C}$ et $T_2=125^\circ\text{C}$) et pour différentes tailles d'échantillon ($n_1 = n_2 = 5, 15, 25$ et 50). Un modèle Weibull-Arrhénus a été utilisé pour simuler les instants de défaillance :

$$R_{T_i}(t) = e^{-\left(t \times e^{\beta_0} \times e^{\beta_1 \frac{1}{X_i}}\right)^v} \quad (\text{I.116})$$

Avec

- $X_i = 273 + T_i$ la température absolue (Températures d'essai $T_1=95^\circ\text{C}$ et $T_2=125^\circ\text{C}$ et température nominale $T_0 = 27^\circ\text{C}$)
- $\beta_0 = -6.5507$
- $\beta_1 = -\frac{E_a}{k}$ avec $E_a=0.29$ eV et $k=8.6171e-5$ eV/°K
- $v = 2.34$ le paramètre de forme

Pour étudier l'effet aléatoire de la simulation, vingt répétitions sont réalisées pour chaque configuration de plan (température et taille d'échantillon).

Nous commençons par présenter les résultats de l'application du modèle paramétrique. Les figures I.22, I.23 et I.24 montrent les résultats des estimations des trois paramètres inconnus (E_a , $\eta = e^{-\beta_0}$ et v) du modèle Weibull-Arrhénus.

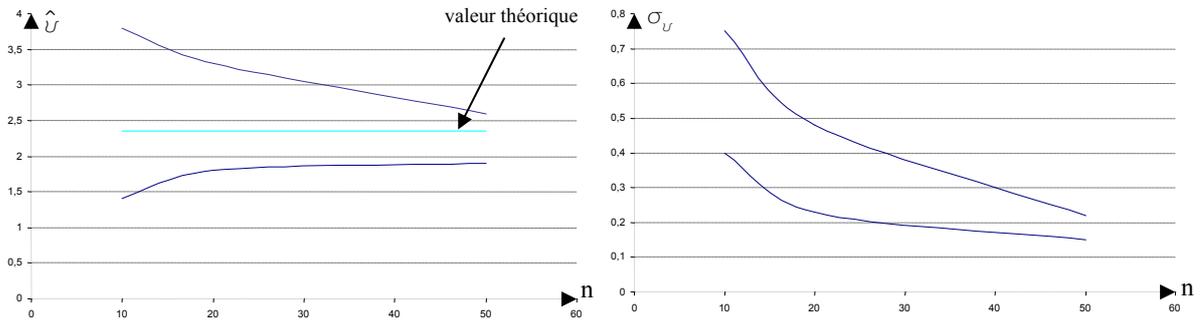


FIG. I.22 – Résultats du test Monte Carlo pour v et σ_v avec l'intervalle de 90% confiance

On peut remarquer que l'estimation du paramètre v est la plus sensible à l'effet de la taille d'échantillon. On montre évidemment qu'avec l'augmentation de la taille d'échantillon les estimateurs ponctuels (par la méthode du maximum de vraisemblance)

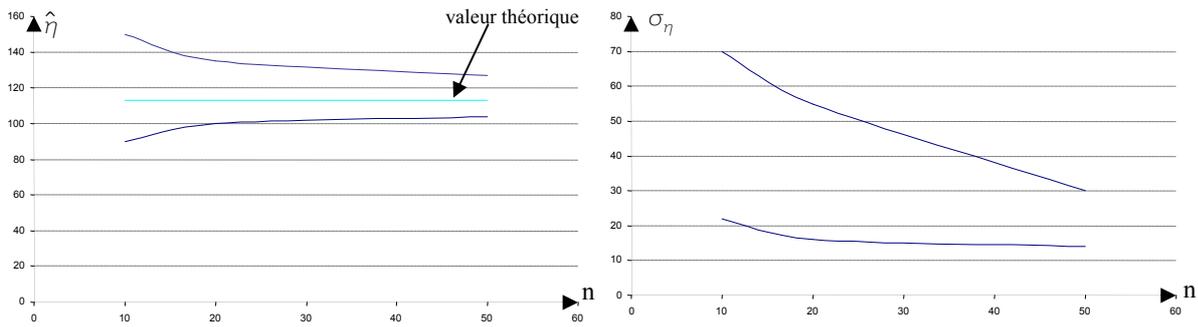


FIG. I.23 – Résultats du test Monte Carlo pour η et σ_η avec l'intervalle de 90% confiance

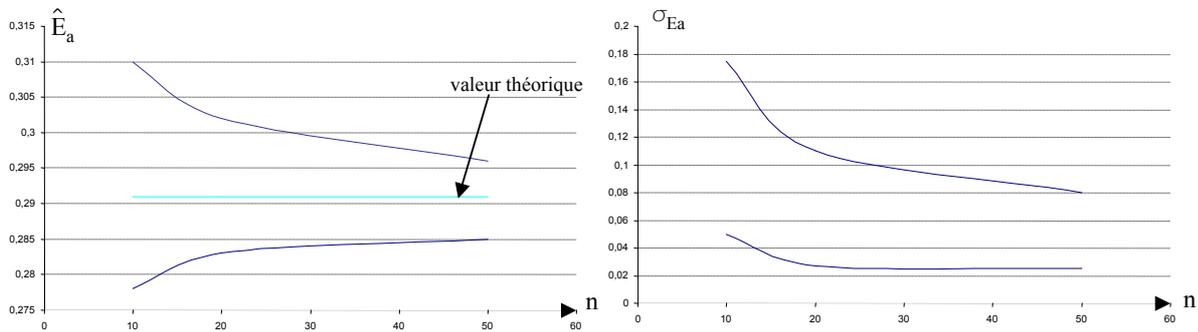


FIG. I.24 – Résultats du test Monte Carlo pour E_a et σ_{E_a} avec l'intervalle de 90% confiance

convergent vers les valeurs théoriques, utilisées pour la simulation, et que les écart-types (estimés à l'aide de la matrice de Fisher) diminuent rapidement.

A partir de ces estimations, il est possible d'estimer les fonctions de fiabilité, pour les différentes tailles d'échantillon,

A partir de ces estimations, on peut caractériser, en utilisant l'équation I.105 l'intervalle de confiance sur la fonction $R(t)$. La figure I.25 présente, pour les quatre tailles d'échantillon $n_1 = n_2 = 5$, $n_1 = n_2 = 15$, $n_1 = n_2 = 25$ et $n_1 = n_2 = 50$ les estimations sur les fonctions $R(t)$ et les intervalles de confiance pour un niveau $\alpha = 0.90\%$. On observe une diminution importante des intervalles avec l'augmentation de la taille d'échantillon.

On peut remarquer, qu'à taille d'échantillon identique, les intervalles de confiance sur la fonction de fiabilité sont plus larges par un plan d'essai accéléré que pour un plan réalisé en conditions nominales. Cela s'explique par la propagation d'incertitudes par le modèle de régression. Par contre, le plan d'essai accéléré permet d'accéder à une estimation de la fonction de fiabilité dans des délais plus réduits.

Maintenant, nous présentons les résultats de l'application du modèle semi-paramétrique. Pour ce faire, nous avons simulé les données selon le même plan et modèle présentés au début du paragraphe et une taille d'échantillon $n_1 = n_2 = 50$. Pour étudier l'influence de l'effet aléatoire sur la simulation, nous avons répété trente fois les simulations. Pour

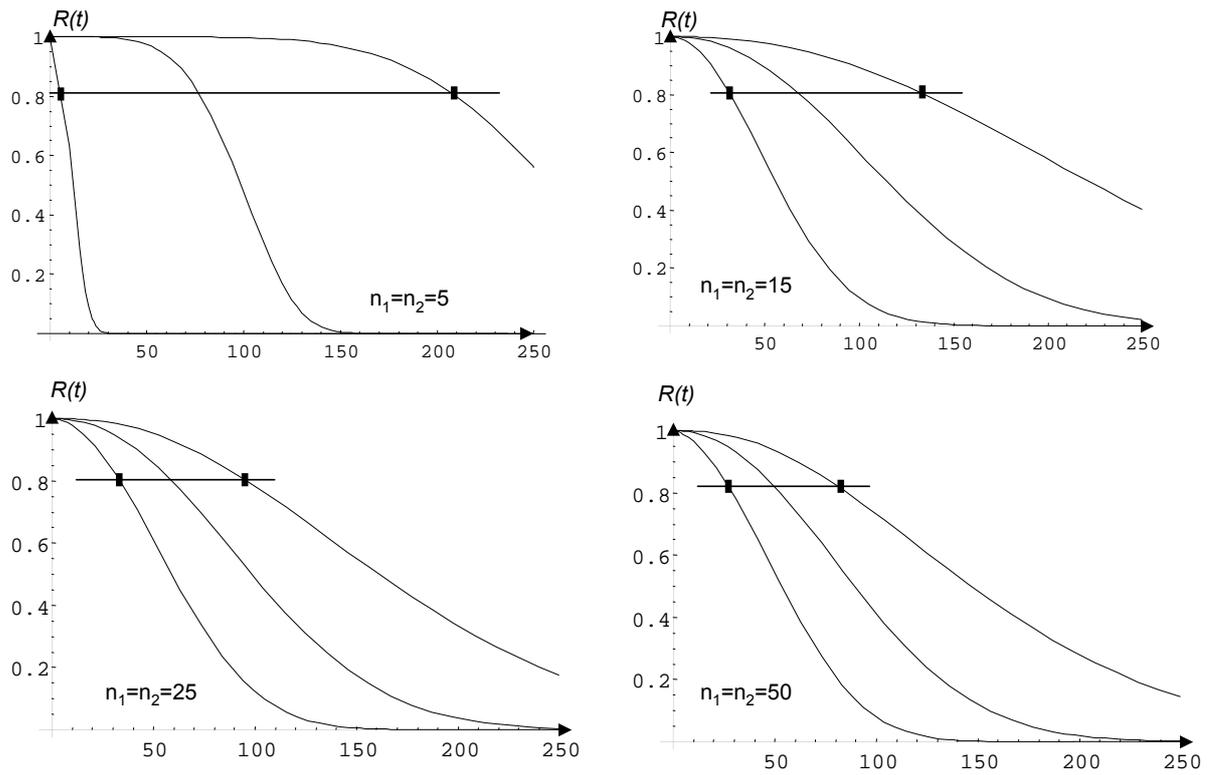


FIG. I.25 – $R(t)$

réaliser l'estimation, nous avons considéré une loi d'Arrhénius $e^{\beta_0 + \beta_1 / (273 + T_i)}$. La figure I.26 présente les estimations des fonctions de fiabilité et I.24 montrent les résultats des estimations des trois paramètres inconnus (E_a , $\eta = e^{-\beta_0}$ et v) du modèle Weibull-Arrhénius.

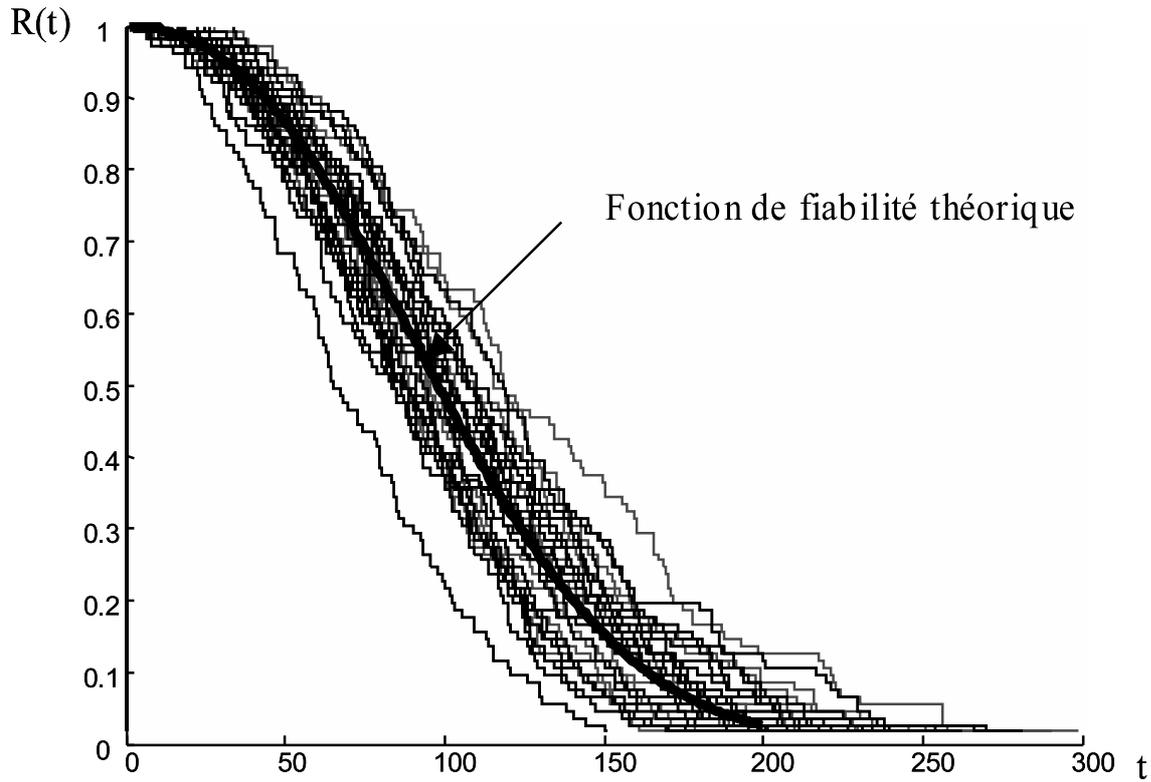


FIG. I.26 – la fonction de fiabilité par le modèle semi paramétrique $R(t)$

On observe la bonne adéquation des fonctions estimées par rapport à celle théorique avec une variabilité limitée. Dans le cas du modèle semi-paramétrique, aucune hypothèse est faite quant à la fonction de fiabilité.

Maintenant, quels sont les critères pour choisir entre les modèles paramétrique et semi-paramétrique ? Cela dépend principalement de la taille d'échantillon et de la censure (effet d'escalier tronqué pour le modèle semi paramétrique) mais également de ce que l'on veut mesurer (la fonction de fiabilité complète, MTTF, B10, ...).

I.8 Conclusion

Ce chapitre a présenté les notions de base de la sûreté du fonctionnement, et plus particulièrement les notions associées à la fiabilité. Ensuite, les méthodes d'analyse de données de fiabilité (instant de défaillance) ont été détaillées.

Puis, les essais accélérés ont été présentés comme alternative pour d'accéder à des estimations, de bonne qualité et rapidement, de la fiabilité lorsque les produits à tester sont très fiables conduisant à des censures importantes s'ils sont testés dans les conditions nominales. Une des plus importantes contraintes de ces essais est de ne pas modifier le mode de défaillance par la sévérisation des conditions d'essai. Le plan étudié est celui par régression qui a pour effet de propager les incertitudes d'estimation obtenues aux niveaux accélérés à celui nominal. Ainsi la figure I.27 illustre le phénomène de propagation d'incertitudes en considérant des couples d'échantillons différents.

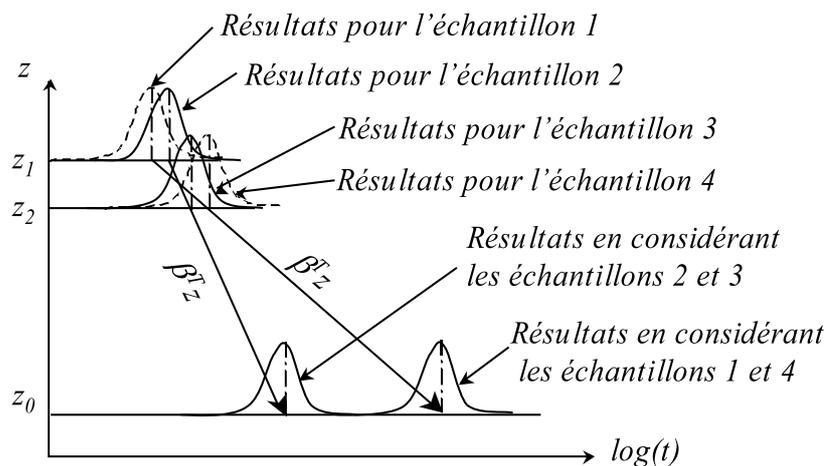


FIG. I.27 – Propagation de l'incertitude

Pour maîtriser les incertitudes, il est possible de rechercher la meilleure configuration de plan d'essai en minimisant l'incertitude d'estimation. De nombreux auteurs traitent la problématique de l'optimisation des plans d'essais accélérés [O'Connor, 2003], [Guerin et al., 2001], [Arizono et al., 2008].

Une autre possibilité pour augmenter la qualité d'estimation sur les paramètres de modèles est d'introduire toute la connaissance disponible sur la fiabilité du système à étudier. En effet, lors du développement d'un produit de nombreuses études sont réalisées afin d'évaluer sa fiabilité (estimation prévisionnelle de la fiabilité par des standards de fiabilité (RdF2000, Mil Hdbk 217, Fides), AMDEC, ...). De plus, assez souvent un nouveau produit est une amélioration d'un ancien se trouvant en exploitation depuis longtemps

et pour lequel nous avons des informations sur sa tenue en terme de fiabilité. D'autre part, il est toujours possible d'interroger un ou plusieurs experts sur le niveau de fiabilité potentiel du produit.

Aussi, nous proposons d'utiliser une démarche bayésienne, pour estimer les modèles standards de vie accélérée, permettant d'intégrer l'information disponible sur la fiabilité du produit. L'objectif est d'obtenir des estimations de bonne qualité (intervalle de confiance "raisonnable") pour des tailles d'échantillon faibles. Le chapitre suivant présente la formalisation de cette démarche.

II

Estimation Bayésienne des modèles SVA

Sommaire

II.1	Introduction	63
II.2	L'inférence Bayésienne dans le cadre de MSVA	66
II.2.1	Application de l'inférence Bayésienne au modèle SVA paramétrique	67
II.2.1.1	Formulation du problème	67
II.2.1.2	Estimation bayésienne des paramètres de modèles	68
II.2.1.2.1	méthode classique	68
II.2.1.2.2	méthode MAP (Maximum of A Priori)	69
II.2.2	L'inférence Bayésienne du modèle SVA semi-paramétrique	73
II.3	L'information a priori	79
II.3.1	Sources d'information	79
II.3.1.1	Recueils de données	80
II.3.1.2	L'avis d'experts	81
II.3.2	Le choix de la loi a priori	82
II.3.2.1	Lois conjuguées	83
II.3.2.2	Détermination de la loi a priori pseudo conjuguée par simulation	84
II.3.3	Construction de la loi a priori	86
II.4	Application numérique	89
II.4.1	A priori cohérent et peu informatif	89
II.4.2	A priori cohérent et très informatif	91
II.4.3	A priori très informatif et incohérent	93

II.5	Compatibilité de l'a priori avec les données	97
II.5.1	Le principe de calcul d'un facteur de compatibilité normalisé	100
II.5.2	L'utilisation du facteur de compatibilité comme facteur de pondération de l'a priori	102
II.5.3	Application au modèle SVA	106
II.6	Exemple numérique	109
II.6.1	Résultats	109
II.6.1.1	A priori peu informatif et cohérent	109
II.6.1.2	A priori très informatif et cohérent	109
II.6.1.3	A priori incohérent et informatif	110
II.6.1.3.1	Les résultats de la méthode Bayésienne classique	112
II.6.1.3.2	Les résultats de la méthode Bayésienne pondérée	112
II.6.2	Synthèse	112
II.7	Conclusion	114

II.1 Introduction

La connaissance est l'état de celui qui connaît ou sait quelque chose. Ce qu'est la connaissance fait encore débat chez les philosophes : il n'y a pas de définition unique sur laquelle ils s'accordent.

On distingue traditionnellement trois types de connaissance :

1. La connaissance *propositionnelle* est le fait de savoir qu'une certaine proposition est vraie, par exemple, savoir que la Terre est ronde,
2. La connaissance *objectuelle*, aussi appelée *acquaintance*, est le fait de connaître une chose, par exemple, connaître Paris,
3. Le *savoir faire* est le fait d'être capable de réussir une action, par exemple, savoir faire des crêpes.

La définition de la *connaissance propositionnelle* est celle qui a le plus attiré l'attention des philosophes. Ils s'accordent généralement sur le fait qu'une connaissance est une croyance qui est vraie, mais aussi qu'elle n'est pas seulement cela. Une autre condition est requise : il faut qu'elle soit justifiée, ou non accidentelle, ou fiable, ou certaine. C'est sur cette condition supplémentaire que les philosophes sont en désaccord.

Le terme *a priori* désigne les connaissances supposées par l'expérience, et qui ne peuvent se réduire à elles. Le concept d'*a priori* au sens kantien désigne ce qui est pensé comme nécessaire et universel et qui, à ce titre, peut être conçu «indépendamment/indépendant de l'expérience». Une connaissance est *a priori* si elle peut être prouvée sans aucune référence à l'expérience.

Comme nous l'avons montré précédemment, les essais d'estimation sont coûteux en temps. Ainsi, il est indispensable d'optimiser les campagnes d'essai en sachant que les niveaux de performances souhaitées à valider sont élevés (probabilité de défaillances et taux de défaillance très faibles) et que la taille des échantillons est telle que l'on a peu de chance d'observer des défaillances.

Dans cette optique, l'approche Bayésienne s'impose puisqu'elle va permettre :

1. d'utiliser le maximum de connaissances disponibles en exploitant les résultats des essais sur des produits similaires et les données issues de l'expertise,
2. d'obtenir des intervalles de confiance plus resserrés par rapport au cas classique,
3. d'intégrer à la connaissance les résultats des séries d'essais au fur et à mesure du développement du produit,
4. d'exploiter le cas du "zéro" défaillance difficile à traiter avec une approche statistique classique qui demanderait une taille d'échantillons très importante.

Il existe de nombreux travaux traitant les techniques bayésiennes. Nous pouvons citer les travaux fondamentaux : [Robert, 1992], [Robert, 2001], [Congdon, 2001], [Ibrahim et al., 2001]... et les travaux concernant les applications à la fiabilité : [Ringler, 1979], [Procaccia et al., 1992], [Procaccia et Morilhat, 1996], [Lannoy et Procaccia, 1994], [Lannoy et Procaccia, 2001], [Sander et Badoux, 1991], [Marts et Walter, 1982], [Ross, 2003].... Le problème difficile à résoudre dans le cadre des techniques bayésiennes reste la modélisation de la connaissance a priori (connaissance initialement disponible). Pourtant, il existe de nombreuses sources de données de fiabilité (des études de sûreté de fonctionnement, étude d'évaluation de risques, d'avis d'expert, du retour d'expérience et des recueils de données).

L'inférence Bayésienne est fondée sur la manipulation d'énoncés probabilistes. Ces énoncés doivent être clairs et concis afin d'éviter toute confusion. Les méthodes bayésiennes se distinguent des méthodes dites standards par l'application systématique de règles formelles de transformation des probabilités.

On nomme inférence Bayésienne la démarche logique permettant de calculer ou réviser la probabilité d'une hypothèse. Cette démarche est régie par l'utilisation de règles strictes de combinaison des probabilités, desquelles dérive le théorème de Bayes. Dans la perspective Bayésienne, une probabilité n'est pas interprétée comme le passage à la limite d'une fréquence, mais plutôt comme la traduction numérique d'un état de connaissance (le degré de confiance accordé à une hypothèse).

La forme continue de théorème Bayes pour la variable aléatoire θ sur le domaine Ω avec t_i , ($i = 1..n$), des résultats d'essai, est :

$$\pi_{apo}(\theta/t_i) = \frac{L(t_1, \dots, t_n/\theta) \pi(\theta)}{\int_{\Omega} L(t_1, \dots, t_n/\theta) \pi(\theta) d\theta} \quad (\text{II.1})$$

avec comme notations $L(t_1, \dots, t_n/\theta)$ la vraisemblance, $\pi(\theta)$ la fonction qui modélise l'information a priori et $\pi_{apo}(\theta/t_i)$ la fonction a posteriori.

Le théorème de Bayes est une conséquence simple des définitions des probabilités conditionnelles. Dans l'approche Bayésienne, le paramètre θ à estimer est une variable aléatoire sur l'espace probabilisable Ω et admet une densité de probabilité.

L'incertitude sur l'occurrence d'un évènement est modélisée par une probabilité subjective, déterminée à partir de connaissances sur des composants similaires ou d'avis d'experts : c'est l'information a priori $\pi(\theta)$.

La fonction de vraisemblance va permettre de valoriser les données issues du retour d'expérience ou d'essais sur le composant étudié : c'est l'information objective $L(t_1, \dots, t_n/\theta)$.

Le théorème de Bayes (équation II.1) va permettre de combiner l'information subjec-

tive avec l'information objective dans une information a posteriori $\pi_{apo}(\theta/t_i)$.

La difficulté d'utiliser les essais accélérés, au cours de phase de conception, est due à la faible taille de l'échantillon à tester. Dans ce contexte, l'approche Bayésienne peut être utilisée pour intégrer dans le processus d'estimation toutes les connaissances disponibles sur l'accélération du modèle de vie (le taux de défaillance, l'énergie d'activation,...). L'application concrète du théorème de Bayes dépend fortement de la capacité à construire une distribution a priori.

Des nombreux articles traitent l'inférence Bayésienne sur les paramètres du modèle de vie accélérée. Dans [Bris, 2000], le cas des essais de démonstration de fiabilité est étudié en considérant l'hypothèse de mortalité exponentielle intégrant un facteur d'accélération. Dans [Madi, 1997], l'estimation de paramètres du modèle de vie accélérée est obtenue en utilisant l'inférence Bayésienne et l'échantillonnage de Gibbs. Dans [Somerville et al., 1997], les auteurs étudient les essais accélérés définis par une conception des expériences en utilisant la distribution a priori Diriclet. Dans [Dorp et Mazzuchi, 2004], l'approche Bayésienne est appliquée aux tests accélérés dans la définition des connaissances a priori directement aux niveaux accélérés. Dans [Dorp et al., 1996], les auteurs étudient l'inférence Bayésienne pour les essais accélérés effectués en contraintes échelonnées.

L'ensemble des travaux concerne des applications et études particulières de l'inférence bayésienne. Dans la suite de ce chapitre, nous proposons d'étudier l'estimation bayésienne des paramètres inconnus des modèles SVA. Nous nous intéresserons tout particulièrement à l'application aux plans par régression en considérant les modèles paramétrique et semi-paramétrique. L'extension aux autres plans et modèles est quasi immédiate car le formalisme utilisé est commun. Une étude sur l'estimation est réalisée avec une comparaison entre les estimateurs classiques et MAP (Maximum of A Priori).

La vraie difficulté, lorsque l'on souhaite mettre en place une démarche bayésienne, est la construction de la loi a priori qui synthétise toute la connaissance disponible sur les paramètres du modèle. Aussi, nous proposons une démarche permettant la construction de celle-ci. Pour finir, nous étudierons la problématique de l'incompatibilité entre la connaissance a priori sur la fiabilité du produit et les premiers résultats d'essais. Aussi, nous proposons un facteur de compatibilité qui indique le degré de cohérence entre l'a priori et la vraisemblance. Cet indicateur est ensuite utilisé, lorsque l'incohérence est révélée, pour pondérer l'a priori afin de diminuer son informativité.

II.2 L'inférence Bayésienne dans le cadre de MSVA

L'intérêt du théorème de Bayes (équation II.1) dans le cadre d'essais est de compenser la faible taille d'échantillon à tester et les censures nombreuses par l'intégration de l'information préalable obtenue par l'expertise ou la prise en compte de résultats d'essais des produits identiques, semblables ou de premières générations. L'information préalable en condition normale d'utilisation du produit sera assimilée à un résultat d'essais hypothétique (a priori) combiné aux valeurs de tests réels en utilisant l'inférence bayésienne.

On peut dire que la distribution a posteriori $\pi_{apo}(\theta/t_i)$ est le produit entre l'information priori $\pi(\theta)$ et l'information obtenue à partir des essais, modulés par la fonction de vraisemblance $L(t_1, \dots, t_n/\theta)$, le tout divisé par un facteur de normalisation afin d'obtenir une fonction densité de probabilité.

L'application du théorème de Bayes à l'estimation de la fiabilité d'un système consiste à appliquer l'inférence Bayésienne représentée dans la figure II.1.

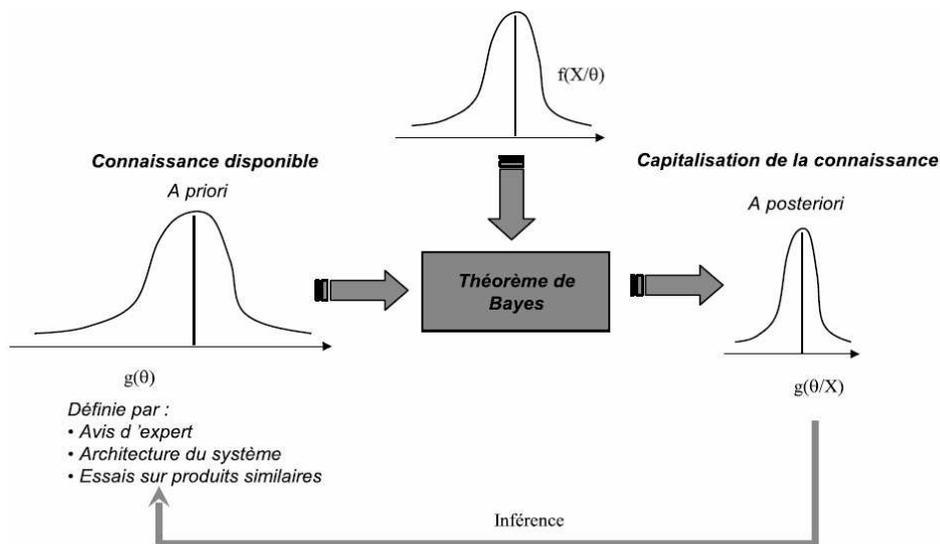


FIG. II.1 – Inférence Bayésienne

La figure II.2 montre un exemple de représentation de ces différentes distributions.

En inférence Bayésienne, le choix crucial est celui de la distribution a priori. Lorsqu'on souhaite dissocier l'apport des données de celui des croyances initiales, on peut prendre une distribution qui ne reflète pas ces croyances, mais qui au contraire formalise un état d'ignorance (distribution neutre ou "non informative"); la distribution finale exprime alors l'apport propre des données : on obtient ainsi, dans la tradition de Laplace à Jeffreys, l'Analyse Bayésienne des Données. Un résultat fondamental est la ré-interprétation laplacienne du seuil observé, pour les modèles d'échantillonnage usuel.

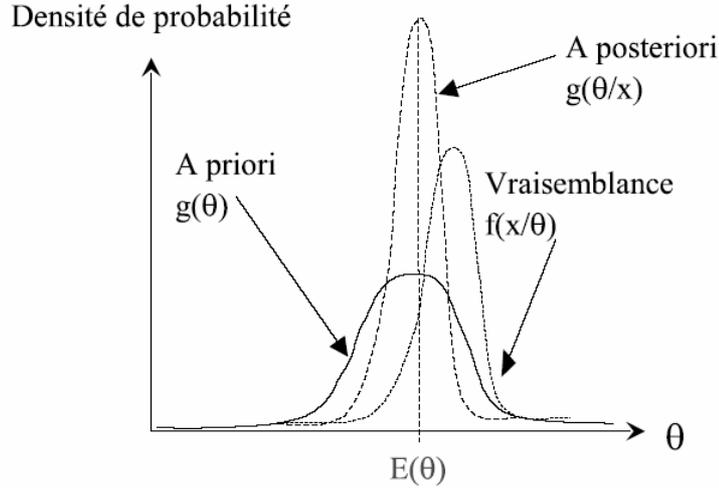


FIG. II.2 – Représentation des distributions a priori, vraisemblance et a posteriori

Nous rappelons que, lorsque le nombre d'observations devient important, la subjectivité de la connaissance a priori perd de son importance au profit des données observées. Par contre, dans le cas d'essais où aucune défaillance n'est observée, l'inférence Bayésienne apporte une solution [Procaccia et al., 1992], [Lannoy et Procaccia, 1994], [Sander et Badoux, 1991], [Usureau et al., 2004].

Dans le cadre de cette thèse, nous proposons d'étudier l'estimation Bayésienne des paramètres des modèles standard de vie accélérée. Cette étude concernera les approches paramétriques présentées dans la section I.6.3 et semi-paramétriques, section I.6.4.

II.2.1 Application de l'inférence Bayésienne au modèle SVA paramétrique

II.2.1.1 Formulation du problème

L'intégration de l'inférence Bayésienne dans le modèle SVA paramétrique présenté dans la section I.6.3 nous permet de compléter le cadre d'analyse de données de fiabilité proposé par [Nikulin, 2007].

La vraisemblance, donnée par l'équation I.96, est :

$$L(\gamma, \sigma | T) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \quad (\text{II.2})$$

Les paramètres inconnus à estimer sont σ et le vecteur γ de m éléments, considérés comme des variables aléatoires. On considère une information a priori sur les paramètres

inconnus modélisée par les fonctions $\pi(\gamma_i)$ et $\pi(\sigma)$. Dans l'hypothèse des variables aléatoires indépendantes, on peut caractériser la totalité de l'information a priori sur les paramètres par :

$$\pi(\gamma, \sigma) = \pi(\sigma) \prod_{i=1}^m \pi(\gamma_i) \quad (\text{II.3})$$

En intégrant l'équation de la fonction de vraisemblance définie par l'équation I.96 dans la forme générale continue du théorème du Bayes (équation II.1), on obtient :

$$\pi_{apo}(\gamma, \sigma | T) = \frac{\left[\prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right] \pi(\gamma, \sigma)}{K_I} \quad (\text{II.4})$$

avec $K_I =$

$$\int_0^\infty \left\{ \int_0^\infty \cdots \int_0^\infty \left[\prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right] \pi(\gamma, \sigma) d\gamma_1 \cdots d\gamma_m \right\} d\sigma,$$

une constante de normalisation qui assure la propriété d'une densité de probabilité pour la fonction $\pi_{apo}(\gamma, \sigma | T)$.

II.2.1.2 Estimation bayésienne des paramètres de modèles

Nous proposons deux approches pour estimer les paramètres inconnus des modèles SVA :

1. la méthode classique consistant à définir les estimateurs ponctuels à partir des premiers moments des fonctions marginales associées aux différents paramètres et les intervalles de confiance par les estimations des variances obtenues par les deuxièmes moments,
2. la méthode MAP (Maximum of A Priori) qui détermine les estimateurs ponctuels à partir des modes des fonctions marginales et les intervalles de confiance par les estimations des variances par la matrice de Fisher.

II.2.1.2.1 méthode classique En faisant l'hypothèse que les variables γ_i et σ sont indépendantes, alors il est possible de définir les fonctions de densité $\pi_i(\theta_i)$ et $\pi(\sigma_i)$ des paramètres γ_i et σ :

$$g_i(\gamma_i) = \int_{-\infty}^0 \int_{j=1..m, j \neq i} \cdots \int \pi_{apo}(\gamma, \sigma | T) d\gamma_j d\sigma \quad (\text{II.5})$$

et

$$g(\sigma) = \int_{i=1..m} \dots \int \pi_{apo}(\gamma, \sigma | T) d\gamma_i \quad (\text{II.6})$$

Dans ce contexte, les estimateurs ponctuels sont les moments initiaux d'ordre 1 :

$$\hat{\gamma}_i = \int_{\Omega_{\gamma_i}} \gamma_i g_i(\gamma_i) d\gamma_i \quad (\text{II.7})$$

et

$$\hat{\sigma} = \int_{\Omega_{\sigma}} \gamma_i g(\sigma) d\sigma \quad (\text{II.8})$$

Les variances sont définies par les moments centrés d'ordre 2 :

$$V(\hat{\gamma}_i) = \int_{\Omega_{\gamma_i}} g_i(\gamma_i) (\gamma_i - \hat{\gamma}_i)^2 d\gamma_i \quad (\text{II.9})$$

et

$$V(\hat{\sigma}) = \int_{\Omega_{\sigma}} g(\sigma) (\sigma - \hat{\sigma})^2 d\sigma \quad (\text{II.10})$$

La méthode nécessite de recourir à des méthodes numériques. Les variances sont utilisées pour faire les estimations des intervalles de confiance selon les relations I.43 et I.44. Il est possible également de définir l'intervalle de confiance bilatérale symétrique au niveau de confiance $1 - \alpha$ tel que :

$$\text{prob}(\gamma_i^{\min} \leq \gamma_i \leq \gamma_i^{\max}) = 1 - \alpha \quad (\text{II.11})$$

avec γ_i^{\min} et γ_i^{\max} solutions de :

$$\int_{\Omega_{\gamma_i}}^{\gamma_i^{\min}} g_i(\gamma_i) d\gamma_i = \alpha/2 \quad (\text{II.12})$$

$$\int_{\Omega_{\gamma_i}}^{\gamma_i^{\max}} g_i(\gamma_i) d\gamma_i = 1 - \alpha/2 \quad (\text{II.13})$$

On procède de la même manière pour le paramètre σ .

II.2.1.2.2 méthode MAP (Maximum of A Priori) Cette méthode considère la fonction de densité a posteriori $\pi_{apo}(\gamma, \sigma | T)$ et les estimateurs ponctuels des paramètres

inconnus (γ, σ) sont estimés de telle sorte qu'ils maximisent :

$$(\hat{\gamma}, \hat{\sigma}) = \text{Arg max}(\pi_{\text{apo}}(\gamma, \sigma | T)) \quad (\text{II.14})$$

Comme dans le cas de la méthode MV classique, ces estimateurs sont obtenus lorsque :

$$\begin{aligned} U_l(\gamma, \sigma) &= \frac{\partial (\ln [\pi_{\text{apo}}(\gamma, \sigma | T)])}{\partial \gamma_l} = \frac{\partial \ln \left[\frac{L(\gamma, \sigma) \pi(\gamma, \sigma)}{K_I} \right]}{\partial \gamma_l} \\ &= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \gamma_l} + \frac{\partial \ln [\pi(\gamma, \sigma)]}{\partial \gamma_l} - \frac{\partial \ln [K_I]}{\partial \gamma_l} \\ &= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \gamma_l} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \gamma_l} \\ &= \frac{1}{\sigma} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} a_{ij}(\gamma, \sigma) + \frac{b-1}{\gamma_l} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \gamma_l}, \quad (l = 1, \dots, m) \\ U_{m+1}(\gamma, \sigma) &= \frac{\partial (\ln [\pi_{\text{apo}}(\gamma, \sigma | T)])}{\partial \sigma} = \frac{\partial \ln \left[\frac{L(\gamma, \sigma) \pi(\gamma, \sigma)}{K_I} \right]}{\partial \sigma} \\ &= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \sigma} + \frac{\partial \ln [\pi(\gamma, \sigma)]}{\partial \sigma} - \frac{\partial \ln [K_I]}{\partial \sigma} \\ &= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \sigma} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \sigma} \\ &= \frac{1}{\sigma} \sum_{i=1}^k \sum_{j=1}^{n_i} \{v_{ij}(\gamma, \sigma) a_{ij}(\gamma, \sigma) - \delta_{ij}\} + \frac{d-1}{\sigma} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \sigma} \end{aligned} \quad (\text{II.15})$$

avec :

$$\begin{aligned} v_{ij}(\gamma, \sigma) &= \frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}, \\ a_{ij}(\gamma, \sigma) &= \lambda(v_{ij}(\gamma, \sigma)) - \delta_{ij} (\ln \lambda)'(v_{ij}(\gamma, \sigma)), \\ \frac{\partial \ln [\pi_{\sigma}(\gamma, \sigma)]}{\partial \sigma} &= \frac{d \left(\ln \left[d \left(\frac{\sigma}{c} \right)^d \right] + \ln [\pi(\gamma, \sigma)] \right)}{d \sigma}, \\ \frac{\partial \ln [\pi_{\gamma_l}(\gamma, \gamma_l)]}{\partial \gamma_l} &= \frac{d \left(\ln \left[b \left(\frac{\gamma_l}{a} \right)^b \right] + \ln [\pi(\gamma, \sigma)] \right)}{d \gamma_l}. \end{aligned}$$

Les estimateurs de maximum d'a posteriori $\hat{\sigma}, \hat{\gamma}$ peuvent être obtenus en résolvant le système d'équations :

$$U_l(\gamma, \sigma) = 0 \quad (l = 1, \dots, m+1) \quad (\text{II.16})$$

et l'estimateur de la fonction de fiabilité R_{X_0} est :

$$\hat{R}_{X_0}(t) = R\left(\frac{\ln t - \hat{\gamma}^T Z^{(0)}}{\hat{\sigma}}\right) \quad (\text{II.17})$$

L'information de Fisher, dans ce cas, respecte le principe de l'équation I.32 et s'applique à la fonction qui décrit l'information apportée sur les paramètres, $\pi_{apo}(\gamma, \sigma | T)$, [Manry et al., 1996] :

$$\begin{aligned} I^{MAP}(\gamma, \sigma) &= E \left[\left(\frac{\partial \log \pi_{apo}(\gamma, \sigma | T)}{\partial \gamma \partial \sigma} \right)^2 \Big|_{\hat{\gamma}, \hat{\sigma}} \right] \\ &= E \left[\left(\frac{\partial \log L(T | \gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \Big|_{\hat{\gamma}, \hat{\sigma}} \right] + E \left[\left(\frac{\partial \log \pi(\gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \Big|_{\hat{\gamma}, \hat{\sigma}} \right] \\ &= I^L(\gamma, \sigma) + I_{\pi}^{AP}(\gamma, \sigma) \end{aligned} \quad (\text{II.18})$$

avec :

$$\begin{aligned} I^L(\gamma, \sigma) &\quad \text{définie dans les équations I.100 – I.102 et} \\ I_{\pi}^{AP}(\gamma, \sigma) &= E \left[\left(\frac{\partial \log \pi(\gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \Big|_{\hat{\gamma}, \hat{\sigma}} \right]. \end{aligned}$$

On observe rapidement que l'information apportée sur les paramètres γ et σ augmente par rapport à l'approche classique avec $I_{\pi}^{AP}(\gamma, \sigma)$. Ceci implique la diminution des variances/covariances des estimateurs ponctuels des paramètres γ et σ , (donnés par l'inégalité Cramer-Rao (équation I.36)).

En reprenant les équations I.100 - I.105, pour calculer l'intervalle de confiance $(1-\alpha)$ sur l'estimateur de la fonction de fiabilité, on obtient la matrice de Fisher $I^{MAP}(\gamma, \sigma)$ avec les termes suivants :

$$\begin{aligned} I_{ls}^{MAP}(\gamma, \sigma) &= I_{ls}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} \right] \\ I_{l,m+1}^{MAP}(\gamma, \sigma) &= I_{l,m+1}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \sigma} \right] \\ I_{m+1,m+1}^{MAP}(\gamma, \sigma) &= I_{m+1,m+1}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma^2} \right] \end{aligned} \quad (\text{II.19})$$

avec :

$$\begin{aligned}\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} &= \begin{cases} 0 & \text{pour } l \neq s \\ -\frac{2(-1+b)}{\gamma_l^2} & \text{pour } l = s, \end{cases} \\ \frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma \partial \gamma_l} &= 0, \\ \frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma^2} &= -\frac{2(-1+d)}{\sigma^2}\end{aligned}$$

Nous avons considéré une indépendance totale entre les variables (équation ??).

De la même manière que pour le cas non-Bayésien, on observe que l'estimateur \hat{R} est une fonction régulière de γ et σ . On va définir :

$$Q_{X_0}(t) = \ln \frac{R_{X_0}(t)}{1 - R_{X_0}(t)} \quad \text{et} \quad \hat{Q}_{X_0}(t) = \ln \frac{\hat{R}_{X_0}(t)}{1 - \hat{R}_{X_0}(t)} \quad (\text{II.20})$$

La fonction $Q_{X_0}(t)$ prend ses valeurs dans \mathfrak{R} donc la vitesse de convergence de \hat{Q}_{X_0} vers la loi limite est plus grande que celle de \hat{R}_{X_0} vers sa loi limite, et on obtient que la loi de

$$\frac{\hat{Q}_{X_0}(t) - Q_{X_0}(t)}{\hat{\sigma}_{Q_0}}$$

est approximée par la loi normale centrée réduite $N(0, 1)$, avec :

$$\begin{aligned}(\hat{\sigma}_{Q_0})^2 &= \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) [I^{MAP}(\gamma, \sigma)]^{-1} \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) \\ &= \left(\frac{R'(R^{-1}(\hat{R}_{X_0}(t)))}{\hat{\sigma}^2 \hat{R}_{X_0}(t) (1 - \hat{R}_{X_0}(t))} \right)^2 \times \left[\hat{\sigma}^4 \sum_{l=0}^m \sum_{s=0}^m z_{0l} z_{0s} I_{MAP}^{ls}(\hat{\gamma}, \hat{\sigma}) \right. \\ &\quad \left. - 2\hat{\sigma}^2 \hat{\gamma}^T Z^{(0)} \sum_{l=0}^m z_{0l} I_{MAP}^{l, m+1}(\hat{\gamma}, \hat{\sigma}) + (\hat{\gamma}^T Z^{(0)})^2 I_{MAP}^{m+1, m+1}(\hat{\gamma}, \hat{\sigma}) \right] \quad (\text{II.21})\end{aligned}$$

avec $I_{MAP}^{l,s}$ les termes de la matrice $[I^{MAP}(\gamma, \sigma)]^{-1}$.

L'intervalle de confiance approximatif $(1-\alpha)$ pour $\hat{R}_{X_0}(t)$ est :

$$\left(1 + \frac{1 - \hat{R}_{X_0}(t)}{\hat{R}_{X_0}(t)} \exp \{ \pm \hat{\sigma}_{Q_0} w_{1-\alpha/2} \} \right)^{-1} \quad (\text{II.22})$$

avec w_α le α -quantile de la loi $N(0, 1)$.

II.2.2 L'inférence Bayésienne du modèle SVA semi-paramétrique

Dans le cas du modèle SVA semi-paramétrique, la démarche bayésienne consiste à capitaliser toute la connaissance disponible sur les paramètres de la loi d'accélération (par exemple l'énergie d'accélération E_a dans le cas de la loi d'Arrhenius).

L'intégration simple de l'inférence Bayésienne dans le modèle SVA semi-paramétrique nous permet de combiner l'analyse discontinue de données de fiabilité proposée par [Nikulin, 2007] avec les fonctions a priori qui définissent la connaissance a priori sur les paramètres.

Pour une question de simplicité et de clarté, nous présenterons uniquement la méthode MAP pour l'estimation des paramètres de modèle. Toutefois, la méthode classique peut être utilisée.

Notons $N_i(\tau)$ le nombre de pannes observées du i -ème groupe dans l'intervalle $[0, \tau]$ et $Y_i(\tau)$ le nombre d'unités à risque (en état de fonctionnement et non-censurés) avant le moment τ , $T_{i1} \leq \dots \leq T_{im_i}$ les moments de pannes du i -ème groupe, les variables aléatoires

$$e^{\beta^T Z_i T_{ij}}, \quad (i = 1, \dots, k; j = 1, \dots, m_i)$$

peuvent être considérées comme des pseudo-pannes *observées* dans une expérience où $n = \sum_{i=1}^k n_i$ unités avec la fonction de fiabilité R_{x_0} ont été testées et $n_i - m_i$ parmi elles ont été censurées au moment $e^{\beta^T Z_i t_i}$.

Alors

$$N^R(\tau, \beta) = \sum_{i=1}^k N_i \left(e^{-\beta^T Z_i \tau} \right)$$

est le nombre de pannes observées dans l'intervalle $[0, \tau]$ et

$$Y^R(\tau, \beta) = \sum_{i=1}^k Y_i \left(e^{-\beta^T Z_i \tau} \right)$$

est le nombre des unités à risque avant l'instant τ .

La fonction de survie R_{X_0} peut être *estimée* par l'estimateur de Kaplan-Meier. Nous allons noter $\Delta N^R(\tau, \beta) = N^R(\tau, \beta) - N^R(\tau-, \beta)$ et $\Delta N_l(u) = N_l(u) - N_l(u-)$, avec la notation générique $f(u-) = \lim_{\varepsilon \downarrow 0} f(u - \varepsilon)$; alors, pour tous $s \leq \max_i \{e^{\beta^T Z_i t_i}\}$, on a :

$$\tilde{R}_{X_0}(s, \beta) = \prod_{\tau \leq s} \left(1 - \frac{\Delta N^R(\tau, \beta)}{Y^R(\tau, \beta)} \right) = \prod_{\tau \leq s} \left(1 - \frac{\sum_{l=1}^k \Delta N_l \left(e^{-\beta^T Z_l \tau} \right)}{\sum_{l=1}^k Y_l \left(e^{-\beta^T Z_l \tau} \right)} \right) \quad (\text{II.23})$$

d'où :

$$\tilde{R}_{X_0}(s, \beta) = \prod_{(i,j): T_{ij} \leq \exp\{-\beta^T Z_i\} s} \left(1 - \frac{1}{\sum_{l=1}^k Y_l (e^{-\beta^T (Z_i - Z_l)} T_{ij})} \right) \quad (\text{II.24})$$

La fonction de vraisemblance s'écrit :

$$L(\beta) = \prod_{i=1}^k \prod_{j=1}^{m_i} \left[e^{\beta^T Z_i} \left(\tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij-}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right) \right] \tilde{R}_{X_0}^{n_i - m_i} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \quad (\text{II.25})$$

En intégrant la fonction de vraisemblance dans la forme générale continue du théorème de Bayes (équation II.1), on obtient :

$$\begin{aligned} \pi_{apo}(\beta | T) &= \frac{1}{K_I} \left\{ \prod_{i=1}^k \prod_{j=1}^{m_i} \left[e^{\beta^T Z_i} \left(\tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij-}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right) \right] \times \right. \\ &\quad \left. \tilde{R}_{X_0}^{n_i - m_i} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right\} \times \pi(\beta) \end{aligned} \quad (\text{II.26})$$

avec :

$$\begin{aligned} K_I &= \int_{\Omega_\beta} \left\{ \prod_{i=1}^k \prod_{j=1}^{m_i} \left[e^{\beta^T Z_i} \left(\tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij-}, \beta \right) - \tilde{R}_{X_0} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right) \right] \times \right. \\ &\quad \left. \tilde{R}_{X_0}^{n_i - m_i} \left(e^{\beta^T Z_i} T_{ij}, \beta \right) \right\} \times \pi(\beta) d\beta \end{aligned}$$

une constante de normalisation.

Comme dans le cas paramétrique, nous ne disposons pas d'expression analytique. De plus, la fonction de fiabilité R_{X_0} est discontinue et engendre une fluctuation aléatoire dans la fonction de vraisemblance (équation II.25). La figure II.3 illustre ce phénomène pour un exemple simple à un seul paramètre. Pour éviter la propagation de ce phénomène, nous avons proposé d'approximer la vraisemblance par une méthode par noyaux (Noyaux de Laplace). Ainsi, sur l'exemple précédent, la figure II.6 présente le résultat de l'approximation.

L'application de la méthode par noyaux aux données permet d'obtenir une approximation de la fonction de vraisemblance. Cette méthode consiste à :

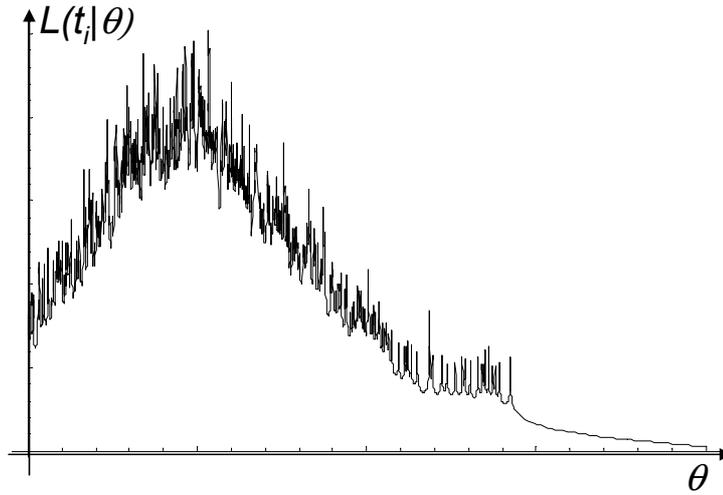


FIG. II.3 – La vraisemblance semiparamétrique

1. chaque paramètre β_i , définir le domaine d'observation $\{\beta_{i_{min}}, \beta_{i_{max}}\}$,
2. discrétiser ces intervalles en w classes
3. estimer la valeur moyenne de $L(t_i|\beta)$ sur chacune des classes
4. définir les centres de classe.

Ainsi, on obtient un histogramme qui est ensuite normalisé. Pour permettre d'appliquer la méthode par noyaux, il est nécessaire de disposer d'un échantillon de β_i caractéristique de la distribution matérialisée par l'histogramme. Pour ce faire, nous simulons n valeurs aléatoires du paramètres β_i respectant les fréquences de l'histogramme.

Pour le cas d'une variable, la figure II.4 illustre l'approximation par cette méthode tandis que la figure II.5 illustre l'approximation pour un cas à deux variables ($n = 2$).

Pour une réalisation particulière $\beta = (\beta_1, \dots, \beta_n)^T$ d'un échantillon, on donne une masse de probabilité $1/n$ pour chaque valeur β_i . On va définir comme noyau la fonction $K(\beta)$, une densité de probabilité unimodale avec un pic en zéro, telle que :

$$\begin{aligned}
 \int_{-\infty}^{\infty} \beta K(\beta) d\beta &= 0 \\
 \int_{-\infty}^{\infty} \beta^2 K(\beta) d\beta &< 0 \\
 k = \int_{-\infty}^{\infty} K^2(\beta) d\beta &< 0
 \end{aligned}
 \tag{II.27}$$

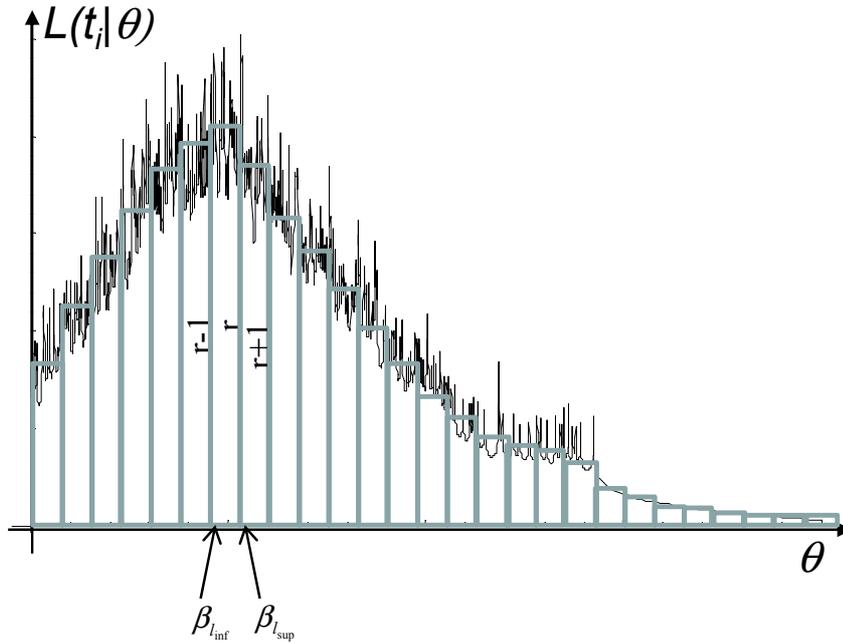


FIG. II.4 – L’approximation de la vraisemblance pour le cas uni variable

Pour toute $h_n > 0$, la fonction :

$$L_{in}^*(\beta) = \frac{1}{h_n} K \left(\frac{\beta - \beta_i}{h_n} \right) \quad (\text{II.28})$$

est une densité unimodale avec un pic en β_i . On étale la masse de probabilité $1/n$ autour de β_i d’après la loi donnée par $L_{in}^*(\beta)$. Donc l’intervalle $] -\infty, \beta]$ obtient la masse :

$$\frac{1}{n} \int_{-\infty}^{\beta} \frac{1}{h_n} K \left(\frac{u - \beta_i}{h_n} \right) du = \frac{1}{n} G \left(\frac{\beta - \beta_i}{h_n} \right) \quad (\text{II.29})$$

où $G(\beta) = \int_{-\infty}^{\beta} K(u) du$. Donc toute la masse de probabilité 1 est étalée d’après la loi de la fonction de répartition et de la densité suivante :

$$L_n^*(\beta) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K \left(\frac{\beta - \beta_i}{h_n} \right) \quad (\text{II.30})$$

Le noyau de Laplace est :

$$K(\beta) = \frac{1}{2} e^{-\beta} \quad (\text{II.31})$$

La figure II.6 présente l’application de la méthode des noyaux [Nikulin, 2007] à la vraisemblance dans le cas de l’inférence Bayésienne semi-paramétrique à un seul paramètre.

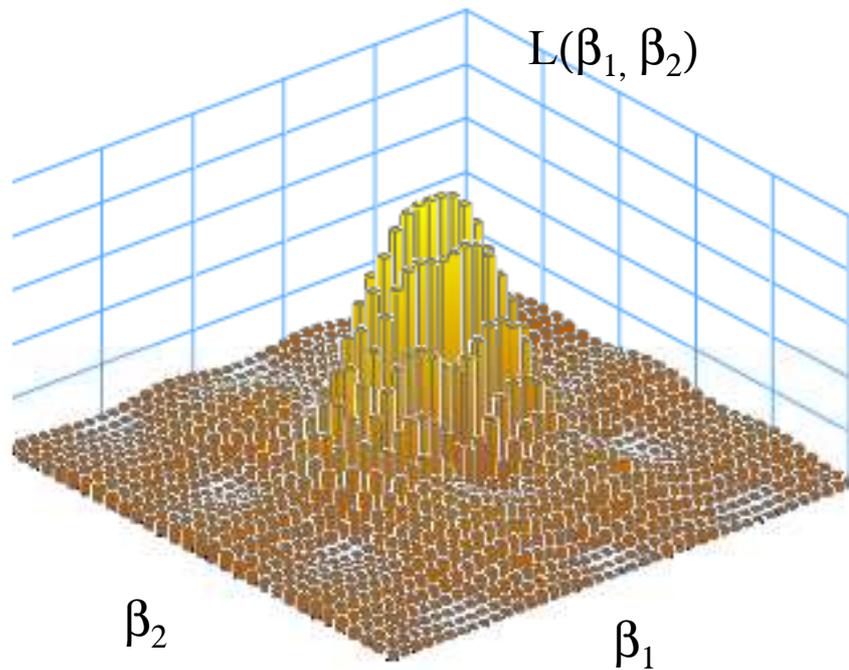


FIG. II.5 – L'approximation de la vraisemblance pour le cas bivariable

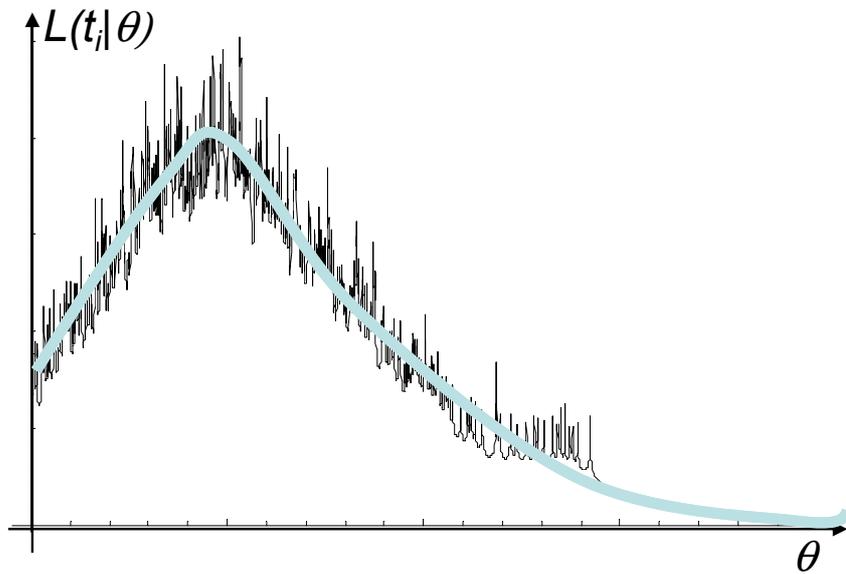


FIG. II.6 – L'approximation par la méthode des noyaux

Une fois la fonction de vraisemblance approximée, les techniques d'estimation des paramètres présentées dans le paragraphe I.3.2.2 permettent de caractériser l'ensemble des variables β .

Dans le cas où la vraisemblance (définie par II.25) est multidimensionnelle, on applique le même principe. On définit l'hyperespace de variation des paramètres $\beta = \{\beta_0, \beta_1, \dots, \beta_q\}$.

Cet hyperespace est discrétisé en $s \times q$ éléments (correspondant à des pas de discrétisation $\Delta\beta_i$) pour obtenir la matrice :

$$\begin{bmatrix} \beta_0^{min}, \dots, \beta_q^{min} \\ \vdots, \dots, \vdots \\ \beta_0^i, \dots, \beta_q^i \\ \vdots, \dots, \vdots \\ \beta_0^{max}, \dots, \beta_q^{max} \end{bmatrix} \quad (\text{II.32})$$

Pour chaque domaine de discrétisation, la moyenne de la fonction de vraisemblance est estimée $\bar{L}(\beta)$. On obtient finalement un hyper-histogramme (voir l'exemple de la figure II.5). Ensuite, nous déterminons une approximation de l'histogramme associé à la fonction marginale du paramètre β_i . Pour ce faire, nous estimons pour chaque centre d'intervalle $\tilde{\beta}_i^j$ ($j = 1, \dots, s$) l'approximation de la fréquence par :

$$f(\tilde{\beta}_i^j) \approx \sum \dots \sum_{j=1, \dots, q \text{ et } j \neq i} \left(\bar{L}(\beta) \prod_{k=1, \dots, q \text{ et } k \neq i} \pi(\beta_k) \times \prod_{k=1, \dots, q \text{ et } k \neq i} \Delta\beta_k \right) \quad (\text{II.33})$$

Après, cet histogramme est normalisé et nous déterminons la fonction de densité par la méthode des noyaux en appliquant la procédure présentée précédemment.

Toute cette première partie a concerné l'application de la démarche bayésienne aux modèles SVA paramétrique et semi-paramétrique. L'intérêt et l'efficacité dépendent principalement dans la capacité de construire la loi a priori. Dans le paragraphe suivant, nous abordons cette problématique.

II.3 L'information a priori

L'introduction d'une loi a priori $\pi(\theta)$ sur les paramètres divise toujours les statisticiens. Le passage de la notion d'un paramètre inconnu à celle d'un paramètre aléatoire représente un saut que beaucoup se refusent à effectuer [Robert, 1992]. Si, dans certains modèles, le paramètre à estimer résulte de l'action de multiples facteurs en amont et peut donc effectivement être perçu comme partiellement aléatoire (on pourrait ainsi citer la physique quantique et sa représentation probabiliste du monde), il est d'autres modèles où la perception du paramètre comme produit d'une expérience aléatoire semble impossible. Ainsi, les adversaires de l'approche Bayésienne avancent l'exemple de l'estimation de quantités physiques, telle la vitesse de la lumière, c . On peut rétorquer à cette objection que la véritable valeur de c n'est jamais connue, du fait de la limitation des appareils de mesure, et qu'il est donc légitime de considérer c comme une variable uniforme dans $[c_0 - \epsilon, c_0 + \epsilon]$, ϵ étant la précision maximale des instruments de mesure et c_0 la mesure retenue. Nous soulignons que l'importance de la loi a priori n'est pas tant dans le fait que le paramètre puisse ou ne puisse pas être conceptuellement envisagé comme la réalisation d'une variable aléatoire, que dans le fait que $\pi(\bullet)$ représente un moyen efficace (le seul ?) de résumer l'information a priori disponible sur les paramètres, ainsi que l'incertitude sur la valeur de cette information, et donc de permettre une appréciation quantitative de ces paramètres. C'est également l'unique manière cohérente de raisonner conditionnellement (au sens probabiliste) aux observations [Robert, 2001].

L'utilisation d'une loi $\pi(\theta)$ implique trois niveaux de discussion :

- La source : quelle est la source d'information
- La forme : quelle est le choix de la loi
- Les paramètres : quelles sont les valeurs des paramètres à choisir

II.3.1 Sources d'information

Il existe plusieurs sources d'informations où sont synthétisées toute la connaissance disponible sur les paramètres influençant la fiabilité des composants. Nous distinguerons, dans le cadre de cette thèse, deux sources distinctes :

- recueils de données issus d'analyse de REX,
- avis d'expert.

II.3.1.1 Recueils de données

Le recueil des données est un ensemble de données validées et/ou élaborées après un long processus d'expertise et de traitement, relatif à un domaine de connaissance et organisé pour être offert aux consultations d'utilisateurs.

En électronique, un domaine dans lequel le calcul de la fiabilité est pratiqué depuis de nombreuses années, les bases de données de fiabilité sont disponibles et nombreuses. En mécanique l'utilisation des recueils de données est plus récente. Il y a une certaine difficulté à constituer ces recueils, compte tenu de la complexité des composants. Néanmoins, ils sont de plus en plus utilisés.

En pratique, on utilise souvent des bases de données connues ou, mieux encore, quand cela est possible, les données issues de retours d'expériences auprès des fabricants des composants. Pour les composants mécaniques ces données sont difficiles à obtenir parce qu'elles ne sont pas communiquées ou publiées et, dans la plupart des cas, elles sont exploitées confidentiellement.

Ces recueils sont mis à jour régulièrement pour tenir compte des évolutions technologiques. En électronique, la durée de vie des données fournies par les recueils est relativement courte (de 3 à 6 ans) ; elle est plus longue en mécanique (de 30 à 40 ans). En électronique, les recueils des données sont souvent exhaustifs sachant que les composants suivent une loi exponentielle. En mécanique, l'hypothèse du taux de défaillance constant est rarement justifiée.

Dans la majorité des recueils, les informations disponibles sont la dénomination du composant, la moyenne des temps de bon fonctionnement (MTTF), le taux de défaillance moyen ou calculé dans l'hypothèse selon laquelle il est constant, l'intervalle de confiance associé et un coefficient de pondération du taux de défaillance dépendant de l'environnement. Le tableau II.1 est un extrait d'un tableau du recueil NPRD-3 [NPRD-3, 1985] avec les données suivantes sur les composants : le taux de défaillance (caractérisé par les valeurs minimale, moyenne et maximale), le nombre des défaillances et les temps cumulés d'exécution. Ainsi dans le même tableau, nous observons que le nombre de défaillances est directement lié à l'intervalle de confiance, un faible nombre de défaillances implique un grand intervalle de confiance et vice-versa [Mihalache, 2007].

Un autre exemple traite le lien entre les modes de défaillances et les valeurs d'énergie d'activation. La figure II.7 présente un extrait d'un recueil dans le domaine de l'électronique.

Dans le domaine du logiciel il y a peu de recueils de données. Les erreurs sont introduites lors de la conception ou bien lors de la programmation. Microsoft a estimé qu'un programmeur introduit environ 6 défaillances sur 1000 lignes de code [Lyonnet, 2006],

Composant	λ_{moy} $10^{-6}h^{-1}$	Limite inf. 20%	Limite sup 20% 20%	Nbr. de défaillances	Temps cumulé h^{-6}
soupape	214.286	109.372	396.948	3	0.014
Capteur	672.367	651.597	693.921	747	1.111
Actionneur	1111.11	1053.916	1171.927	270	0.243
Led	0.267	0.264	0.289	363	1312.883

TAB. II.1 – Extrait du recueil NPRD-3

[MIL-HDBK-338B, 1998].

Les données collectées permettent d'établir des modèles de fiabilité prévisionnelle pour les composants mécaniques, électroniques et logiciel.

II.3.1.2 L'avis d'experts

En sûreté de fonctionnement comme dans d'autres domaines, le jugement d'expert peut être considéré comme l'une des réponses possibles à un problème technique, un appoint à un retour d'expérience déficient ou un support d'aide à la décision. Il est en effet utilisé pour compléter ou suppléer les données objectives éventuellement disponibles sur le problème considéré lorsque ces données sont peu nombreuses, dispersées voire inexistantes ou lorsqu'elles sont douteuses ou inapplicables. C'est bien souvent la seule source d'information disponible pour effectuer un état de l'art sur un sujet difficile, complexe, nouveau ou peu compris [Poncet, 2005].

Dans le domaine technique, l'expertise peut être considérée comme une source de données, souvent qualitative, quelquefois quantitative. Mais comme toute donnée, les données d'expertise sont une représentation imparfaite et incertaine de ce qu'elles sont censées représenter. La qualité de ces données dépend essentiellement :

- de l'information disponible et de l'expert et, en particulier, de sa formation : son éducation, sa culture, sa profession, sa formation théorique et pratique sur le sujet, son expérience, son environnement, sa personnalité, et d'une façon générale de la qualité intrinsèque de l'expert, mais aussi de sa motivation et de son sérieux, de sa façon de recueillir, appréhender, transmettre, corriger et agréger ses connaissances,
- de la complexité relative du problème et de l'état général des connaissances sur ce problème, et de la manière dont cette complexité est modélisée dans l'esprit de l'expert,
- de la façon dont va s'instaurer le dialogue entre l'expert et l'analyste chargé de recueillir puis d'interpréter son expertise.

D'après Meyer et Booker [MEY93], est considéré comme un expert dans un domaine, toute personne ayant une connaissance particulière avec le niveau de détail nécessaire à

Mécanismes défaillance	Causes défaillance	Mode de défaillance	Etapas de fabrication	Types défaillance	E_a eV	Modes accélér.	Types de défaut
1) Contamination ionique	Présence d'ions alcalins dans oxyde de grille	Dérive de la tension de seuil	Photoresist. Passivation	Jeunesse Vieillesse	1 1,4	E, T	D
2) Charges de surface	Présence de charges en surfaces des oxydes de grille	Dérive de la tension de seuil	Diffusion Trait. à chaud Passivation	Vieillesse	0,5 1	E, T	D
3) Inversion de charges	Charges induites	Dérive de la tension de seuil	Passivation	Vieillesse	1	E, T	D
4) Accumulation de charges en surface	Impuretés ioniques dans la résine	Dérive de la tension de seuil	Encapsulation		1 1,35	E, T	D
5) Perte de charges	Perte de charges entraînant une dérive de la tension de seuil	Mauvais fonctionnement mémoire		Jeunesse Aléatoires Vieillesse	0,6 1,4	T	D
6) Polarisation	Présence de molécules polarisables dans oxyde de grille	Dérive de la tension de seuil	Dopage phosphore Passivation	Vieillesse	1	E, T	D
7) Défauts d'oxyde	Cassures, trous fêlures dans oxyde	Circuits ouverts	Croissance oxyde	Jeunesse Aléatoires	0,3 0,5	E, T	C
8) Porteurs chauds	Ions piégés dans l'oxyde	Dérive de la tension de seuil	Inadéquation conception procédés	Vieillesse	-1 -1,4	$T < 0^\circ\text{C}$ V	D
9) Claquage diélectrique	ESD, surtension	Circuits ouverts Fuites	Photoresist Etching	Jeunesse Aléatoires	0,2 1	E, T	C
10) Interaction (Al-SiO ₂)	Destruction de l'oxyde	Court-circuits	Croissance oxyde	Vieillesse	1,77 2,6	E, T	C

FIG. II.7 – Des valeurs pour l'énergie d'activation par rapport aux modes de défaillances

l'étude de ce domaine.

L'inférence Bayésienne peut se porter sur l'avis d'un ou des plusieurs experts.

II.3.2 Le choix de la loi a priori

Le choix d'une distribution a priori est une démarche complexe et peut laisser place à une appréciation subjective. Ce choix dépend principalement de la connaissance préalable, de la qualité des informations disponibles sur le paramètre étudié.

La distribution a priori doit intégrer, d'une part, toutes les connaissances disponibles sur le paramètre étudié et, d'autre part, toutes les incertitudes liées à ce paramètre. En fonction de la connaissance disponible, nous pouvons définir les principales distributions susceptibles de caractériser la distribution a priori. Si la connaissance initiale est faible, l'a priori sera une distribution simple, peu informative. Si tout paramètre cherché a la même probabilité d'être choisi, l'a priori suit une distribution uniforme. Si la connaissance initiale est solide, l'a priori sera une distribution plus sophistiquée.

Concernant la modélisation effective de la loi $\pi(\theta)$ a priori à partir de la connaissance a priori, de nombreuses méthodologies ont été développées pour modéliser la connaissance

a priori à partir du type d'informations disponibles.

Plusieurs principes peuvent être appliqués :

- le maximum d'entropie : la loi a priori $\pi(\theta)$ maximisant l'entropie correspond en effet à la loi minimisant l'information a priori apporté sur θ ,
- les méthodes Bayésiennes empiriques : utilisent les observations et la loi marginale pour estimer les paramètres de la loi a priori,
- les modèles hiérarchiques : modélisent le manque d'information sur les paramètres de la loi a priori par une nouvelle distribution sur ces paramètres, etc.

Même dans les situations où l'information a priori est disponible, il est rare de pouvoir proposer une détermination exacte de la loi a priori. Il importe ainsi de pouvoir mesurer les conséquences de cette imprécision sur l'inférence a posteriori, afin de contrôler que l'apport arbitraire dans la loi a priori n'est pas prédominant. Cette étude est dite analyse de sensibilité ou de robustesse. Nous proposons d'utiliser les pseudo-lois conjuguées. En effet, pour étendre le concept d'indifférence au choix de la meilleure connaissance a priori en fonction des données observées, Clarotti démontre dans [Clarotti, 1998] et [Procaccia, 1998] l'intérêt particulier des lois conjuguées.

II.3.2.1 Lois conjuguées

Pour des raisons de calcul purement théorique, la distribution a priori dépend de la famille mathématique de la fonction de vraisemblance ; cette dernière condition n'est pas obligatoire pour une résolution numérique.

En fait, le problème consiste à définir quelle densité a priori sera susceptible de laisser le pouvoir le plus cohérent entre les données et la connaissance a priori. Il est donc nécessaire de trouver une série de densités de probabilités a priori qui permette de déterminer avec bon sens quel pouvoir de décision sera laissé aux données.

Il est ainsi démontré que les familles de lois dites conjuguées ne se contentent pas uniquement de simplifier les calculs mais surtout que ce type de lois maximise le pouvoir des données observées.

Le tableau II.2 donne les principales lois conjuguées utilisées - [Procaccia et al., 1992], [Suhner, 1994]

La recherche de ces lois peut être problématique car nous ne disposons pas des expressions analytiques pour définir la loi a priori conjuguée d'une vraisemblance définie par un modèle SVA. Néanmoins, nous proposons de trouver une loi a priori par simulation vérifiant les propriétés de conjugaison.

$f(x \theta)$	$\pi(\theta)$	$\pi_{apo}(\theta x)$
Normale $N(\theta, \sigma^2)$	Normale $N(\mu, \tau^2)$	$N(\rho(\sigma^2\mu + r^2x), \rho\sigma^2r^2)$ avec $\rho^{-1} = \sigma^2 + r^2$
Poisson $P(\theta)$	Gamma $G(\alpha, \beta)$	$G(\alpha + x, \beta + 1)$
Gamma $G(v, \theta)$	Gamma $G(\alpha, \beta)$	$G(\alpha + v, \beta + x)$
Binomiale $B(n, \theta)$	Bêta $Be(\alpha, \beta)$	$Be(\alpha + x, \beta + n - x)$
Négative Binomiale $Neg(m, \theta)$	Bêta $Be(\alpha, \beta)$	$Be(\alpha + m, \beta + x)$
Multinomiale $M_k(\theta_1, \dots, \theta_k)$	Dirichlet $D(\alpha_1, \dots, \alpha_k)$	$D(\alpha_1 + x_1, \dots, \alpha_k + x_k)$
Bêta $Be(\alpha, \theta)$	Exponentielle $Exp(\lambda)$	$exp(\lambda - \log(1 - x))$
Normale $N(\mu, 1/\theta)$	Gamma $G(\alpha, \beta)$	$G(\alpha + 0.5, \beta + (\mu - x)^2/2)$

TAB. II.2 – Lois conjuguées naturelles des familles exponentielles classiques

II.3.2.2 Détermination de la loi a priori pseudo conjuguée par simulation

Pour obtenir la forme de la distribution a priori qui soit la conjuguée bayésienne d'une vraisemblance donnée, il est nécessaire de procéder en deux étapes :

1. la première consiste à déterminer la distribution a posteriori en considérant un a priori non informatif :

$$\Pi_{apos1}(\theta|t_i) = \frac{L(t_1, \dots, t_n|\theta) \cdot \Pi_0(\theta)}{\int_{\Omega} L(t_1, \dots, t_n|\theta) \cdot \Pi_0(\theta) \cdot d\theta}$$

2. dans un deuxième temps, nous construisons la distribution a priori $\Pi_1(\theta)$ selon la forme définie par $\Pi_{apos1}(\theta|t_i)$ et ensuite nous déterminons la distribution a posteriori :

$$\Pi_{apos2}(\theta|t_i) = \frac{L(t_1, \dots, t_n|\theta) \cdot \Pi_1(\theta)}{\int_{\Omega} L(t_1, \dots, t_n|\theta) \cdot \Pi_1(\theta) \cdot d\theta}$$

Nous vérifions que le type de distribution de $\Pi_{apos2}(\theta|t_i)$ est identique à $\Pi_1(\theta)$. Si cela est le cas, ce type respecte les conditions de conjugaison bayésienne.

Dans le cas des modèles SVA, il est difficile de traiter ces deux étapes analytiquement. Aussi, nous proposons de déterminer approximativement les lois conjuguées bayésiennes par simulation.

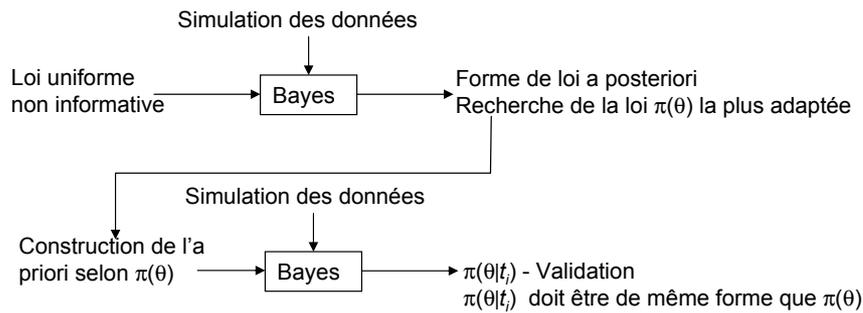


FIG. II.8 – Recherche d'une pseudo loi conjuguée

La procédure des simulations est la suivante :

1. on fixe le plan d'essai (niveau d'accélération, taille des échantillons à tester sous chaque niveau de stress) permettant de construire la vraisemblance,
2. à l'aide des informations a priori sur les paramètres, on cherche les valeurs les plus probables associées aux paramètres [Poncet et al., 2004a],
3. on construit la simulation de Monte-Carlo en répétant (n fois) les étapes suivantes :
 - on simule des instants de défaillance aléatoires sur les hypothèses estimées au point 2,
 - on calcule les estimations des paramètres par MV.
4. on construit les histogrammes des estimateurs MV,
5. on fait des hypothèses sur les formes des lois les plus proches des histogrammes et on vérifie la validité de ces hypothèses (test du Chi-Deux, test de Kolmogorov Smirnov, ...).
6. on construit les lois a priori selon les formes obtenues dans l'étape 5 à partir des informations de l'étape 2,
7. on réitère les étapes 3, 4 et 5.
8. on vérifie que les formes des distributions obtenues dans l'étape 7 sont identiques à celles de l'étape 5.

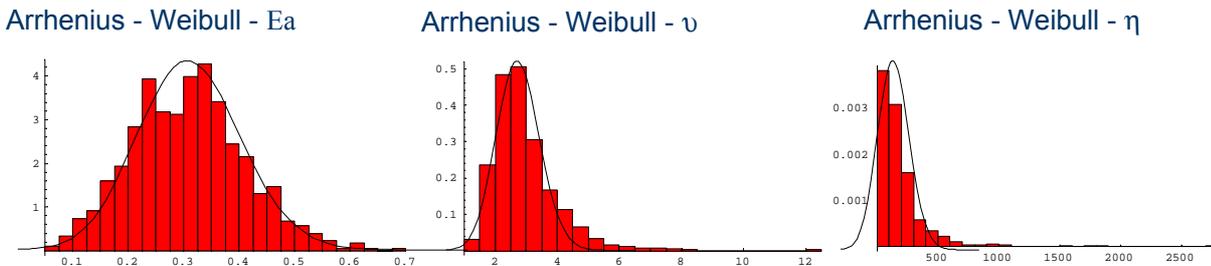
La construction d'a priori serait faite en respectant la forme de lois validée au point 8.

La figure II.8 synthétise la procédure.

Une fois la modélisation a priori réalisée et les décisions prises, concernant le type de loi, il reste un point à traiter. Il faut en effet vérifier que la loi a priori déterminée est la plus pertinente et, notamment, qu'elle n'influence pas la vraisemblance de manière à dégrader le résultat. Cette partie constitue un élément peu étudié à ce jour. La figure II.9 montre que la forme des distributions a posteriori des paramètres d'un modèle SVA

paramétrique (Weibull-Arrhénus) de paramètres η (échelle), ν (forme) et E_a (énergie d'activation) ne changent pas de forme notablement avec la taille d'échantillon ($n = 9$ et $n = 60$). Il reste à déterminer les paramètres des lois a priori à partir des informations disponibles.

taille d'échantillon $n = 9$



taille d'échantillon $n = 60$

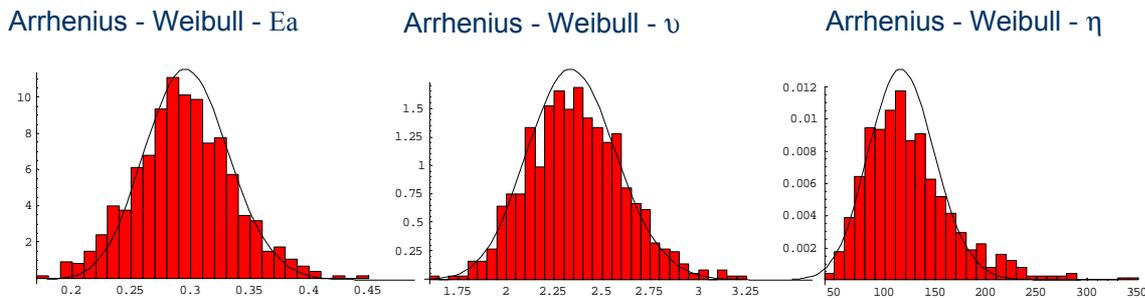


FIG. II.9 – Application pour un cas Arrhenius - Weibull

II.3.3 Construction de la loi a priori

Une fois que l'on a choisi les types de loi à utiliser pour modéliser la distribution a priori, il reste ensuite à déterminer les paramètres de lois à partir des différentes informations disponibles. Rappelons que l'on fait l'hypothèse que tous les paramètres de modèles sont indépendants qui implique que la distribution a priori est le produit des distributions a priori individuelles associées aux différents paramètres ($\Pi(\theta) = \prod_{i=1}^q \Pi_i(\theta_i)$). En fonction du nombre de paramètres de la loi à modéliser, il faudra disposer d'un ou de plusieurs types d'informations.

En fonction des informations fournies, plusieurs méthodes sont disponibles pour estimer les paramètres de lois [Suhner, 1994], [Finetti, 1974] et [Goldstein, 1999] :

- la méthode séquentielle,
- la méthode des échantillons équilibrés,
- la méthode des moments pondérés,

- la méthode des doubles moments,
- la méthode des similitudes comme l'ADEPT ou autre
- la méthode du maximum d'entropie,
- les techniques de transformation de groupe,
- la méthode des moments (moyenne et écart - type),
- la méthode des fractiles.

La plupart des sources d'informations donne celles-ci sous forme d'intervalle. Aussi, nous proposons dans la suite de la thèse d'utiliser la méthode des moments. Le principe consiste à déterminer la moyenne et la variance de loi uniforme informative bornée par les limites d'intervalle. Ainsi, lorsque l'on dispose d'une information a priori, sur le paramètre θ_i , par un intervalle $\theta_i \in [\theta_i^{min}, \theta_i^{max}]$, il est possible d'estimer la moyenne et la variance par :

$$\bar{\theta}_i = \frac{\theta_i^{min} + \theta_i^{max}}{2} \quad \text{et} \quad V(\theta_i) = \frac{(\theta_i^{max} - \theta_i^{min})^2}{12} \quad (\text{II.34})$$

Ensuite, nous choisissons un type de loi a priori et estimons ses paramètres permettant de respecter la moyenne et la variance obtenues ci-dessus.

Pour illustrer cette méthode, considérons le cas d'un modèle SVA paramétrique (Weibull-Arrhénius) pour lequel nous disposons des informations a priori (tableau II.3).

paramètre	limite inférieure	limite supérieure
v	0	4.68
η	0	230
E_a	0eV	0.58eV

TAB. II.3 – Connaissances a priori

La figure II.9 montre le résultat de la recherche des pseudo lois conjuguées bayésiennes par simulation. Ainsi, nous avons trouvé que des lois Gamma (de paramètres α et β) étaient les plus appropriées pour définir les distributions a priori. La moyenne et la variance pour une loi Gamma, en fonction des paramètres, sont définies par :

$$\bar{\theta}_i = \frac{\alpha}{i}\beta \quad \text{et} \quad V(\theta_i) = \frac{\alpha}{\beta^2} \quad (\text{II.35})$$

Aussi, nous pouvons déterminer les paramètres de la loi Gamma pour θ_i par :

$$\alpha = \frac{\bar{\theta}_i^2}{V(\theta_i)} \quad \text{et} \quad \beta = \frac{\bar{\theta}_i}{V(\theta_i)} \quad (\text{II.36})$$

Pour les paramètres du modèle SVA du tableau II.3, nous obtenons les paramètres des

lois Gamma suivants :

paramètre	α	β
v	3	1.28
η	3	0.026
E_a	3	10.34

TAB. II.4 – Paramètres des lois Gamma a priori

II.4 Application numérique

L'exemple numérique reprend en partie les paramètres de simulation de l'exemple de la section I.4. Ainsi, nous avons simulé les résultats d'un plan d'essai complet (sans censure) à deux niveaux de température ($T_1=95^\circ\text{C}$ et $T_2=125^\circ\text{C}$) et pour différentes tailles d'échantillon. Un modèle Weibull-Arrhénius a été utilisé pour simuler les instants de défaillance :

$$R_{T_i}(t) = e^{-\left(t \times e^{\beta_0} \times e^{\beta_1 \frac{1}{X_i}}\right)^v} \quad (\text{II.37})$$

Avec

- $X_i = 273 + T_i$ la température absolue (Températures d'essai $T_1=95^\circ\text{C}$ et $T_2=125^\circ\text{C}$ et température nominale $T_0 = 27^\circ\text{C}$)
- $\beta_0 = -6.5507$
- $\beta_1 = -\frac{E_a}{k}$ avec $E_a=0.29$ eV et $k=8.6171e-5$ eV/ $^\circ\text{K}$
- $v = 2.34$ le paramètre de forme

Pour le cas Arrhenius-Weibull, aucune forme analytique ne permet de trouver la forme des lois conjuguées. Aussi, nous utiliserons la méthodologie présentée précédemment pour rechercher les pseudo lois conjuguées bayésiennes et déterminer les paramètres. Pour étudier l'effet aléatoire de la simulation, dix répétitions sont réalisées pour chaque configuration de plan (température et taille d'échantillon).

Trois cas de connaissances a priori sont présentés :

1. un cas d'un a priori cohérent (c'est-à-dire centré sur la valeur théorique) et peu informatif (variance importante),
2. un cas d'un a priori très informatif et cohérent,
3. un cas d'un a priori très informatif et incohérent (décentré par rapport à la valeur théorique).

II.4.1 A priori cohérent et peu informatif

Pour le premier cas, des distributions a priori peu informatives sur les paramètres $\theta = \{v, \eta, E_a\}$ ont été considérées.

Le tableau II.5 donne l'opinion d'un expert sur les paramètres.

La figure II.9 montre le résultat de la recherche des pseudo lois conjuguées bayésiennes par simulation. Ainsi, nous avons trouvé que des lois Gamma (de paramètres α et β)

paramètre	limite inférieure	limite supérieure
ν	0	4.68
η	0	230
E_a	0eV	0.58eV

TAB. II.5 – Prior intervalles

étaient les plus appropriées pour définir les distributions a priori. Le tableau II.6 donne les paramètres des lois Gamma a priori par la méthode des moments.

paramètre	α	β
ν	3	1.28
η	3	0.026
E_a	3	10.34

TAB. II.6 – Paramètres des lois Gamma a priori

Les résultats de simulations sont présentés dans les figures II.10, II.11 et II.12. Les courbes (a) et (b) représentent, respectivement, les estimations classiques et bayésiennes.

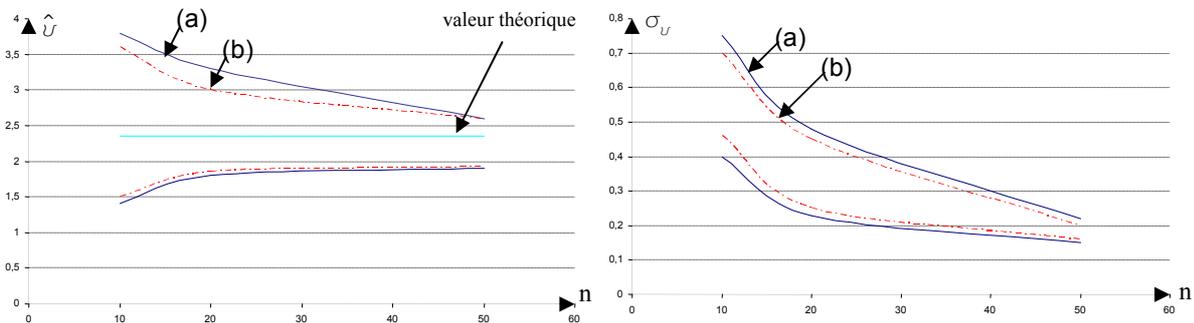


FIG. II.10 – Résultats du test Monte Carlo pour $\hat{\nu}$ et σ_ν avec l'intervalle de confiance 90%

Nous constatons par rapport aux résultats de la section I.7 :

- les estimateurs ponctuels (MAP) restent proches de la valeur théorique ce qui se justifie par le fait que les trois distributions Gamma a priori ont leur moyenne égale au valeur théorique,
- la taille d'échantillon, entre $n_1 = n_2 = 10$ et $n_1 = n_2 = 20$, a une influence importante sur les estimations des paramètres,
- l'estimation de l'écart-type du η montre une sensibilité à l'apport d'une faible information a priori.

On constate que l'intérêt de la démarche Bayésienne diminue avec l'information objective (taille d'échantillon).

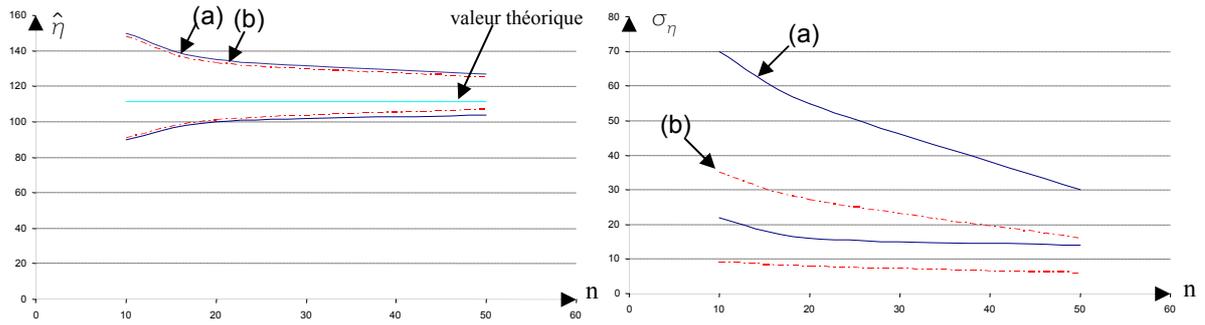


FIG. II.11 – Résultats du test Monte Carlo pour $\hat{\eta}$ et σ_{η} avec l'intervalle de confiance 90%

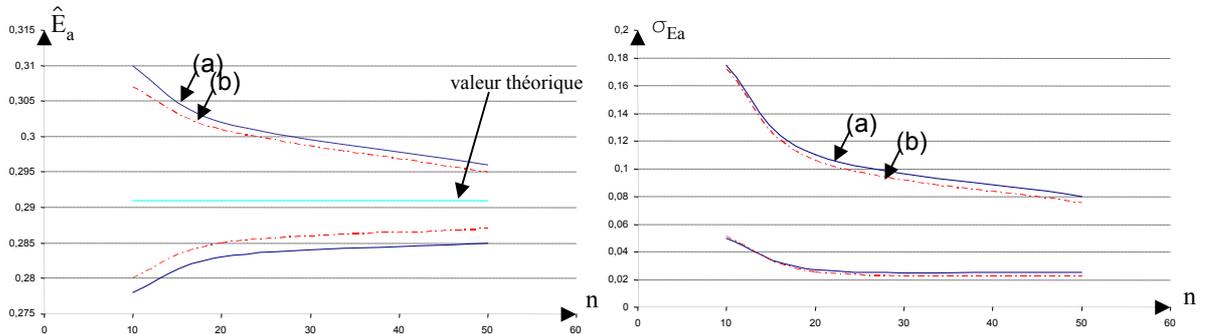


FIG. II.12 – Résultats de la simulation de Monte Carlo pour \hat{E}_a et σ_{E_a} avec l'intervalle de confiance 90%

A partir de ces estimations, il est possible d'estimer la fonction de fiabilité $R(t)$, en utilisant l'équation II.22, et son intervalle de confiance.

La figure II.13 présente, pour quatre situations $n_1 = n_2 = 5$, $n_1 = n_2 = 15$, $n_1 = n_2 = 25$ et $n_1 = n_2 = 50$, les estimations sur les fonctions $R(t)$ dans les conditions nominales et les intervalles de confiance pour le niveau de confiance $\alpha = 0.95\%$.

Nous constatons que les intervalles de confiance sont sensibles à la taille d'échantillon.

II.4.2 A priori cohérent et très informatif

Pour le deuxième cas, des distributions a priori très informatives et cohérentes sur les paramètres ont été considérées.

Par une distribution a priori très informative, on comprend une distribution caractérisée par sa variance très faible. Par une distribution a priori très cohérente on comprend une distribution avec sa moyenne très proche de la valeur théorique.

Le tableau II.7 donne l'opinion d'expert sur les paramètres.

De ces informations sont déduits les paramètres des lois Gamma a priori selon la méthode des moments présentée dans le paragraphe II.3.3.

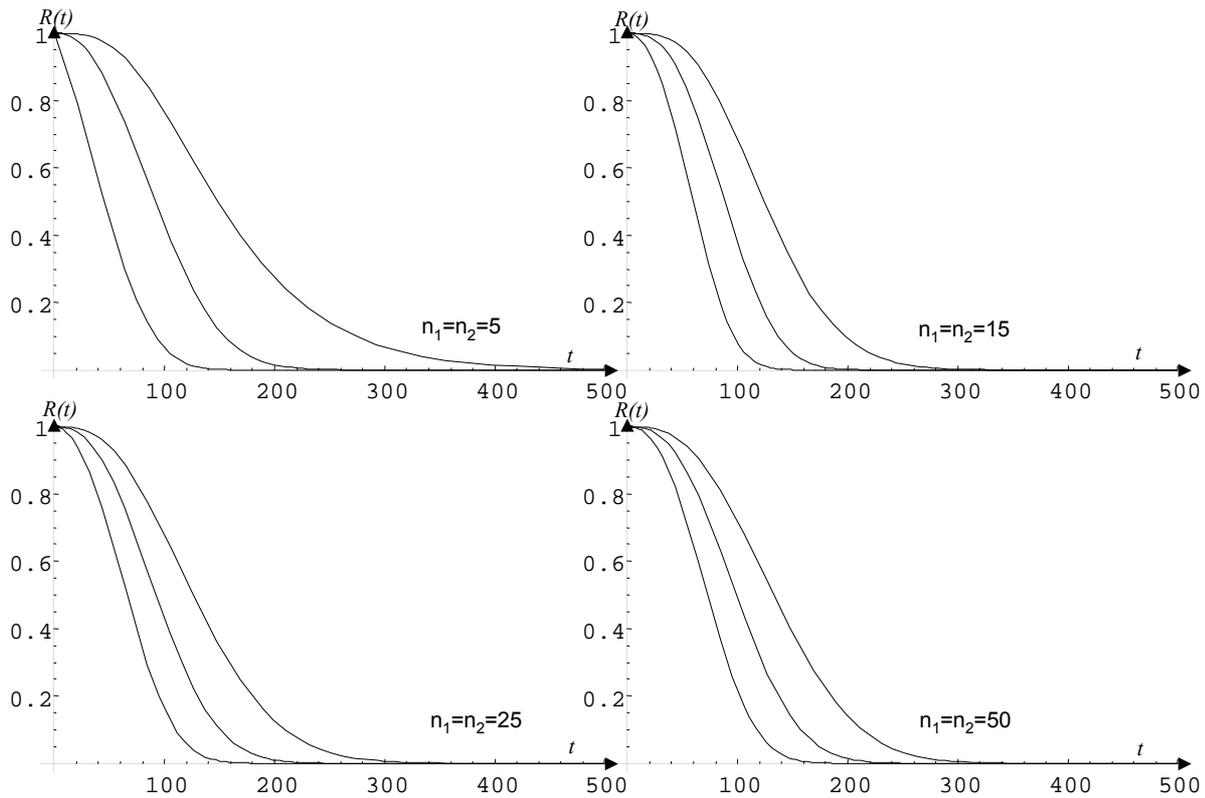


FIG. II.13 – $R(t)$

paramètre	limite inférieure	limite supérieure
ν	2.0475	2.6325
η	100	120
E_a	0.25375eV	0.32625eV

TAB. II.7 – Prior intervalles

Les courbes présentées dans les figures II.14, II.15 et II.16 présentent :

- (a) l'estimation classique
- (b) l'estimation bayésienne, présentée dans le paragraphe précédent, avec un a priori cohérent et peu informatif,
- (c) l'estimation bayésienne, présentée dans ce paragraphe, avec un a priori cohérent et très informatif,

Nous constatons, par rapport aux résultats de la section I.7, que pour le cas bayésien avec l'a priori cohérent et très informatif (courbes c) :

- les estimateurs ponctuels bayésiens sont très proches des valeurs théoriques,
- les écart-types des estimateurs sont peu influencés par la taille d'échantillon (cela s'explique par le fait que l'information apportée par les résultats d'essai est faible par rapport à celle contenue dans les a priori),

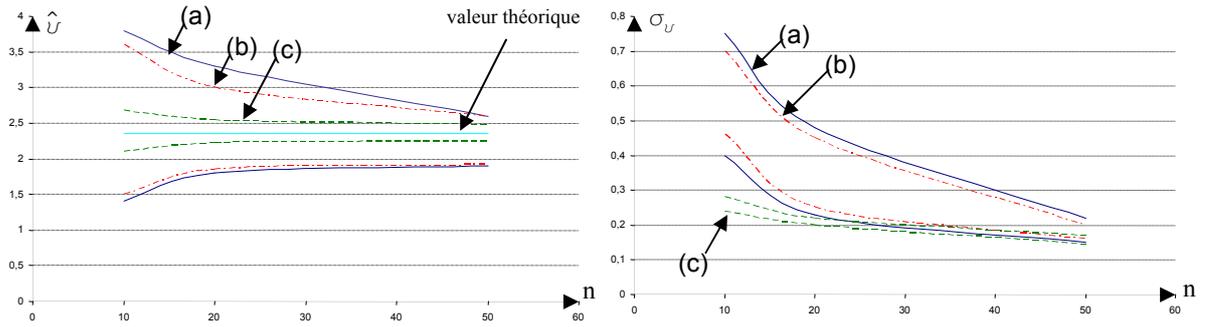


FIG. II.14 – Résultats de la simulation de Monte Carlo pour \hat{v} et σ_v avec l'intervalle de confiance 90%

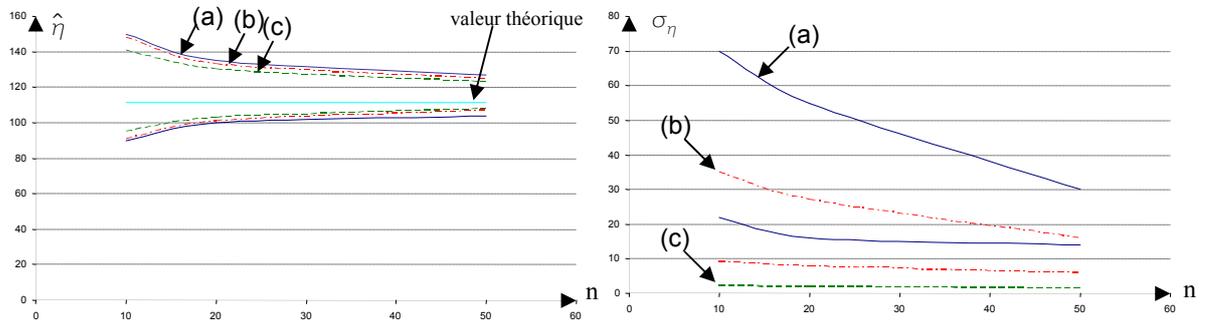


FIG. II.15 – Résultats de la simulation de Monte Carlo pour $\hat{\eta}$ et σ_η avec l'intervalle de confiance 90%

- les intervalles du paramètres η pour les trois cas, sont relativement proches montrant une certaine insensibilité de ce paramètre à l'apport d'information.

Suite aux estimations de paramètres on peut caractériser, en utilisant l'équation II.22 l'intervalle de confiance sur la fonction $R(t)$.

La figure II.17 présente, pour quatre situations $n_1 = n_2 = 5$, $n_1 = n_2 = 15$, $n_1 = n_2 = 25$ et $n_1 = n_2 = 50$, les estimations sur les fonctions $R(t)$ et les intervalles de confiance pour le niveau de confiance $\alpha = 0.95\%$.

II.4.3 A priori très informatif et incohérent

Pour le troisième cas, une très mauvaise information a priori sur les paramètres a été considérée. Le tableau II.8 donne l'opinion de l'expert sur les paramètres.

De ces informations sont déduits les paramètres des lois Gamma a priori selon la méthode des moments présentée dans le paragraphe II.3.3.

Les courbes présentées dans les figures II.14, II.15 et II.16 présentent :

- (a) l'estimation classique
- (b) l'estimation bayésienne, présentée dans un paragraphe précédent, avec un a

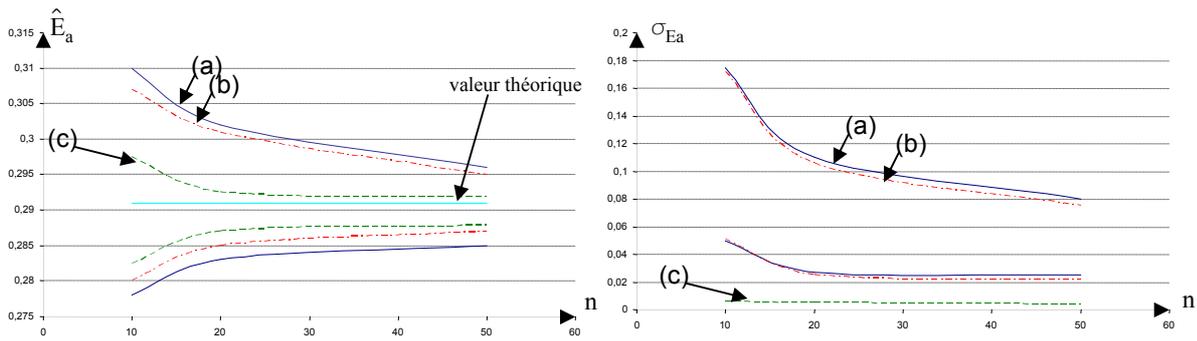


FIG. II.16 – Résultats de la simulation de Monte Carlo pour \hat{E}_a et $\hat{\sigma}_{E_a}$ avec l'intervalle de confiance 90%

paramètre	limite inférieure	limite supérieure
ν	2.5	2.6
η	80	90
E_a	0.33	0.35

TAB. II.8 – Prior intervalles

priori cohérent et peu informatif,

- (c) l'estimation bayésienne, présentée dans le paragraphe, avec un a priori cohérent et très informatif,
- (d) l'estimation bayésienne, présentée dans ce paragraphe, avec un a priori incohérent et très informatif.

La forte information a priori s'impose face aux résultats d'essai. Ainsi, à cause de l'écart existant entre les intervalles a priori et les valeurs théoriques, tous les résultats sont décalés par rapport aux vraies valeurs :

- les estimateurs ponctuels pour ν et pour E_a sont sur-évalués,
- l'estimateur ponctuel pour η est sous évalué

Puisque les a priori sont très informatifs, les écart-types sont faibles et peu sensibles aux tailles d'échantillon. Il en découle des intervalles de confiance serrés mais décentrés.

Comme dans ce cas, la connaissance objective donnée par la vraisemblance peut être dégradée par le choix d'un mauvais a priori. L'analyse Bayésienne peut amener à prendre une décision a posteriori erronée. Dans le paragraphe suivant, nous proposons d'étudier le problème de la compatibilité de l'a priori et les premiers résultats d'essais.

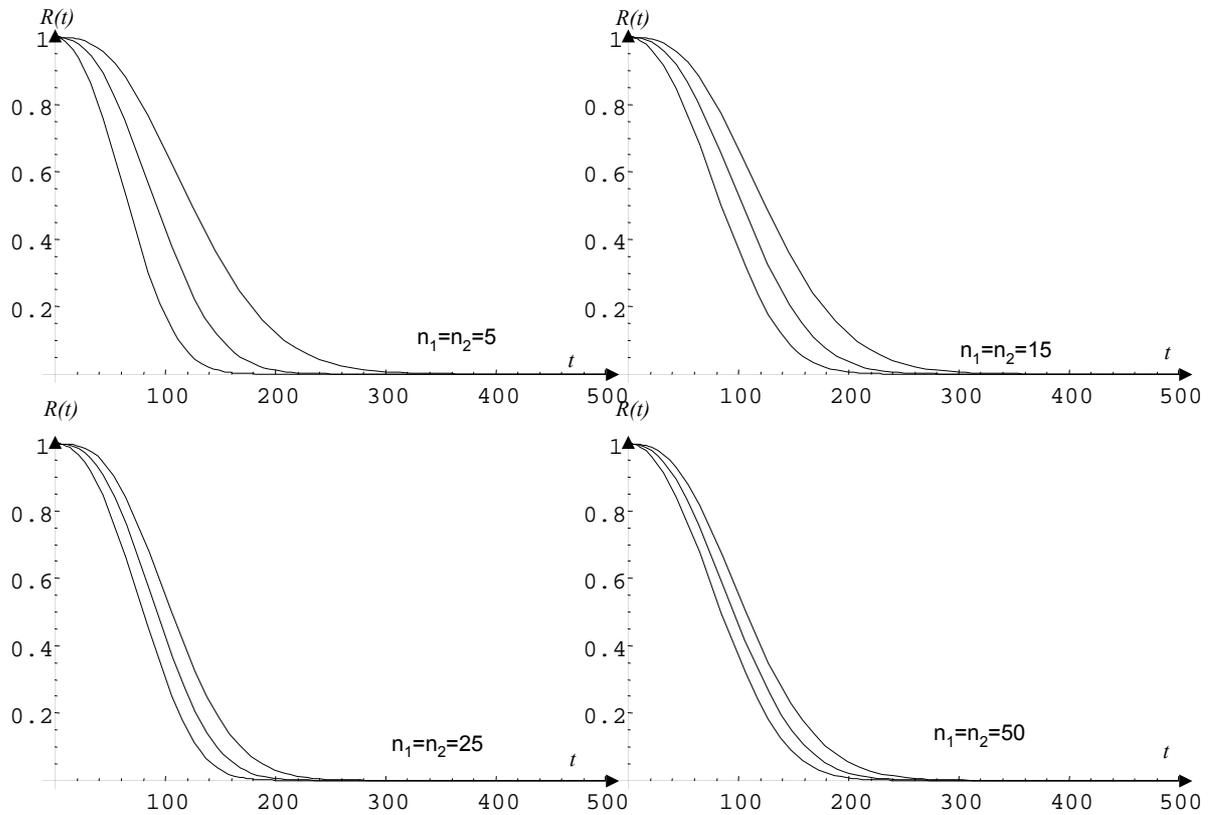


FIG. II.17 – $R(t)$

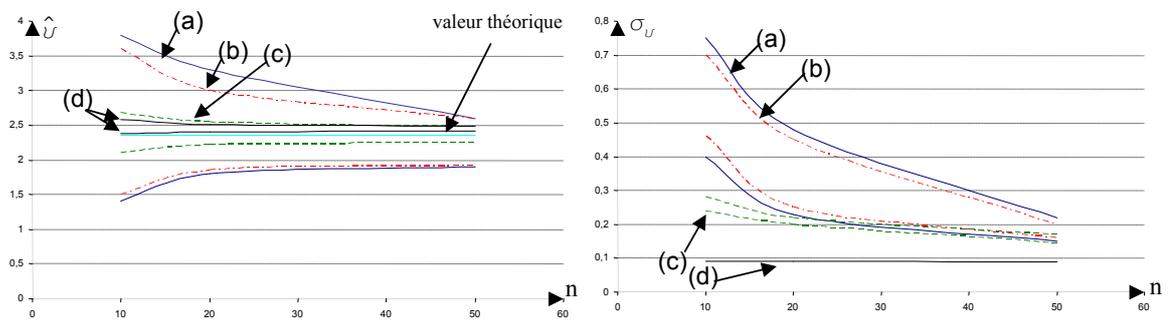


FIG. II.18 – Résultats de la simulation de Monte Carlo pour \hat{v} et σ_v avec l'intervalle de confiance 90%

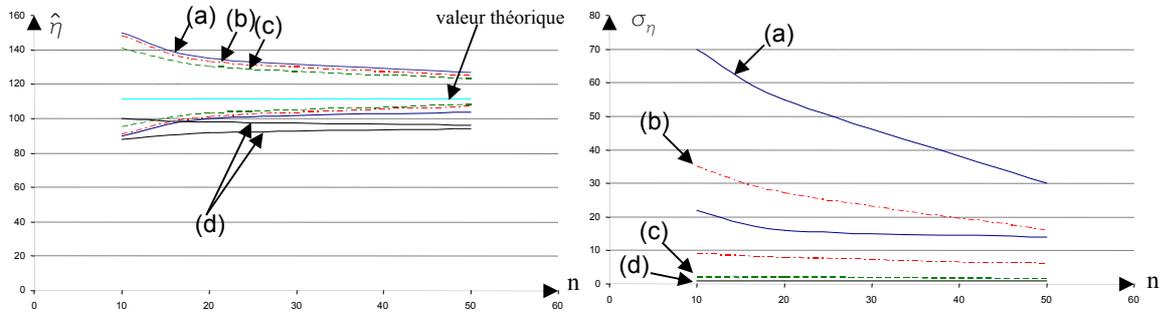


FIG. II.19 – Résultats de la simulation de Monte Carlo pour $\hat{\eta}$ et σ_η avec l'intervalle de confiance 90%

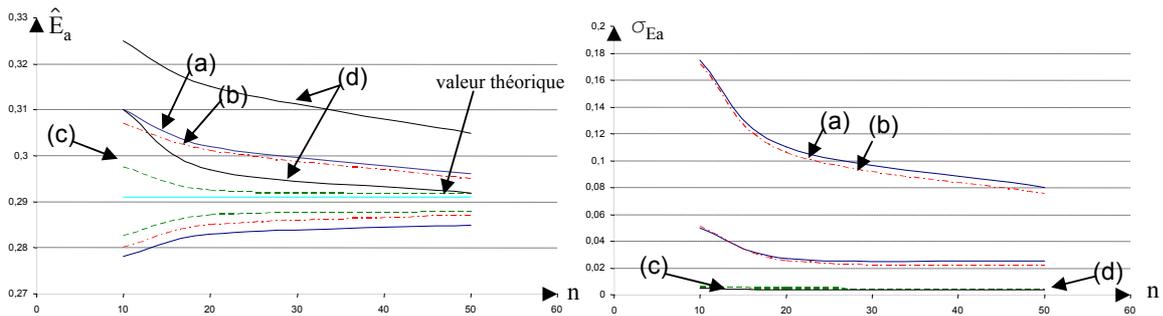


FIG. II.20 – Résultats de la simulation de Monte Carlo pour \hat{E}_a et σ_{E_a} avec l'intervalle de confiance 90%

II.5 Compatibilité de l'a priori avec les données

Parmi les articles sur l'inférence Bayésienne subjective pour estimer le paramètre θ , un grand nombre d'entre eux insiste pour le comportement contrariant du résultat a posteriori quand l'information préalable est mal conditionnée. Ainsi nous considérons une information fautive fournie par les avis experts (ou toute autre source d'information subjective) et une distribution a priori $\pi(\theta)$ dans la limite de la région de confiance de l'approche fréquentielle sur le paramètre θ apportée par des données observées.

Ils faut noter que, en cas de données fortement censurées et de petites tailles d'échantillons, quand l'approche Bayésienne informative est très recommandée [Robert, 2001], l'hétérogénéité de données peut avoir les mêmes effets qu'un mauvais a priori avec des données correctes. Un a priori peut être très bien choisi mais l'a posteriori peut impliquer de mauvaises décisions. Par exemple, imaginer un système industriel dont les données de défaillance ont été notées dans le passé. Après quelques amendements le système fonctionne actuellement. Un expert donne le temps probable de défaillance à l'avenir qui peut être décalé, en raison de l'évolution technique. Dans ce cas, certaines données antérieures peuvent être considérées comme une pollution de la connaissance objective. Dans la plupart des cas, les articles sont portés sur un renfort de la robustesse a posteriori. Cependant, une première difficulté est justement de considérer que la connaissance a priori n'est pas contradictoire.

De plus, une connaissance a priori et une vraie connaissance peuvent avoir des rôles symétriques en conflit ; particulièrement dans un contexte industriel, les réponses antérieures de calibrage répondent bien au tremblement des données observées. Curieusement, ce sujet ne semble pas avoir été beaucoup étudié bien qu'il constitue, de notre point de vue, un préliminaire important à l'inférence subjective Bayésienne. Il n'existe pas d'outils simples de statistique pour définir l'accord entre une distribution antérieure et une connaissance objective apportées par des données observées.

Il faut noter dans l'observation stricte du paradigme bayésien que la distribution a priori et la vraisemblance constituent un modèle entier de décision et il semble à première vue non approprié de déterminer les degrés d'anomalie entre eux, d'un point de vue théorique. C'est pourquoi beaucoup d'articles sur l'inférence Bayésienne sont portés sur l'étude de l'accord des données avec le choix du modèle bayésien et la détection de la surprise [Bayarri et Berger, 1998]. À notre connaissance, il y a peu d'articles traitant la détection du conflit entre l'a priori et la vraisemblance. Le principe est de mettre en place des indicateurs habituellement construits pour permettre de modifier l'antérieur ou la probabilité en détectant et en enlevant des données mal positionnées ou des données

trop influentes. La question est d'obtenir des a priori robustes permettant un impact a posteriori négligeable quand un conflit est détecté.

[Han, 1997] a développé des méthodes pour construire la distribution antérieure hiérarchique. Récemment, quelques résultats ont été obtenus sur la méthode Bayésienne hiérarchique pour traiter des données de la vie. Mais tous ces résultats obtenus au moyen de méthodes bayésiennes hiérarchiques comportent des intégrations compliquées.

[Dunson et Taylor, 2005] suppose que les données se composent d'un échantillon aléatoire d'une fonction de distribution F_Y , qui est inconnue, et que l'intérêt s'appuie sur les inférences de paramètres données comme un vecteur des quantiles de F_Y . Quand la fonction de probabilité n'est pas entièrement indiquée, une densité postérieure ne peut pas être calculée et l'inférence Bayésienne est difficile. Les auteurs prennent en considération une approche qui se fonde sur une vraisemblance de substitution caractérisée par un vecteur des quantiles.

[Cross et al., 2007] propose une technique Bayésienne pour l'évaluation simultanée de la taille initiale équivalente *flaw size* (EIFS) et les distributions de taux de propagation de la fissure dans les modèles de prévision de la vie basés par mécanique de rupture des données d'inspection. La méthode d'inférence mesure séparément les différentes sources de variabilité pour obtenir une distribution indépendante EIFS sans contraintes pour la connaissance a priori des incertitudes sur d'autres paramètres. La capacité prévoit simultanément et exactement le bruit du taux de propagation de la fissure de distribution et d'EIFS d'une petite taille d'échantillon.

Le théorème de Bayes (equation II.1) témoigne d'une certaine audace, puisqu'il met sur un pied d'égalité causes et effets, tous deux pouvant être probabilisés. La différence entre observations et paramètres s'estompe, les manipulations conditionnelles permettant d'interchanger leurs positions respectives.

Quand le paradigme bayésien doit être mis en œuvre pour estimer les paramètres d'un modèle de décision, la modélisation a priori de ces paramètres doit être calibrée très soigneusement. Si la vraisemblance des données observées et la connaissance a priori se révèlent en trop grand désaccord sur la zone de confiance qu'elles induisent pour chacun des paramètres, l'inférence Bayésienne peut conduire à des mauvaises décisions. La détection d'un tel conflit revêt un aspect important - et curieusement sous-estimé - en statistique appliquée à des problèmes concertés.

Dans l'ouvrage [Tribus, 1972], qui constitue une somme des techniques bayésiennes dont la grande émergence se fait à cette époque, il est montré que le χ^2 constitue un exemple de passage à la limite du psi-test (test de plausibilité) bayésien lorsque le nombre de valeurs en présence devient important - ce qui est la condition de travail des statistiques

classiques, mais pas nécessairement des bayésiennes. Le raccord entre les deux disciplines, qui sont asymptotiquement convergentes, est ainsi complet.

Le psi-test bayésien (qui est utilisé pour déterminer la plausibilité d'une distribution par rapport à des observations) est asymptotiquement convergent avec le χ^2 des statistiques classiques dans la mesure où le nombre d'observations devient grand. Le choix apparemment arbitraire d'une distance euclidienne dans le χ^2 est ainsi parfaitement justifié a posteriori par le raisonnement bayésien.

[Idée, 2001] propose d'employer un test de Fisher entre les mesures antérieures et empiriques de l'incertitude. Aussi, dans le cadre de [Usureau, 2001], on trouve une méthode de vérification des lois a priori modélisées en étudiant la compatibilité avec la vraisemblance ([Usureau et al., 2004], [Guerin et al., 2003], [Guerin et al., 2004]). Ces auteurs ont défini un indicateur évaluant la pertinence de la loi a priori et ont mis en place une méthode permettant une meilleure utilisation de celle-ci par rapport à la distribution de vraisemblance. Le domaine d'existence de cet indicateur est $[0, 1]$. Plus la valeur est proche de 1, plus les deux distributions sont cohérentes. Les auteurs vont plus loin en pondérant l'information a priori avec ce facteur pour diminuer son influence sur l'a posteriori.

L'un des plus récents travaux sur ce sujet a été dirigé dans [Bousquet, 2006]. Dans le cadre de ce rapport, l'auteur propose la validation de l'a priori en employant un critère de cohérence qui mesure la comparabilité entre la modélisation a priori π et l'information apportée par les données. Le critère consiste à calculer le ratio :

$$\text{DAC}^J(\pi | y_n) = \frac{KL(\pi^J(\cdot | y_n) \| \pi)}{KL(\pi^J(\cdot | y_n) \| \pi^J)} \quad (\text{II.38})$$

avec $KL(\pi_1 \| \pi_2) = \int_{\Theta} \pi_1(\theta) \log \frac{\pi_1(\theta)}{\pi_2(\theta)} d\theta$ la divergence Kullback-Leibler entre les distributions π_1 et π_2 et π^J l'a priori non informative. DAC est considéré comme critère de concordance entre les données $. | y_n$ et l'a priori π . Une valeur $\text{DAC}^J(\pi | y_n) \leq 1$ caractérise une concordance entre les données $. | y_n$ et l'a priori π . Une valeur $\text{DAC}^J(\pi | y_n) > 1$ signale un conflit. La méthodologie n'est pas applicable aussi facilement pour des cas plus complexes que les cas idéaux à cause des ajustements nécessaires à faire sur π^J suite à sa impropreté ($\int_{\Theta} \pi(\theta) d\theta \rightarrow \infty$).

Une fois que l'une de ces modélisations a été vérifiée et qu'il est prouvé que l'a priori n'est pas compatible avec l'information apportée par les données, trois solutions peuvent être envisagées :

- revenir à l'utilisation des méthodes classiques, ce qui implique soit la perte d'intérêt pour les méthodes Bayésiennes, soit la disparition d'une partie d'information qui se traduit dans des pertes économiques,

- recommencer la modélisation de la connaissance a priori, ce qui peut se révéler long et coûteux en terme d’analyse de données ou d’interrogation des experts,
- pondérer la connaissance a priori pour la rendre compatible avec les résultats réellement obtenus (vraisemblance) et donc ne retenir qu’une partie des informations a priori. En effet, même mal modélisée, la connaissance a priori est en partie basée sur des données correctes qui peuvent être exploitées dans le cadre de l’inférence Bayésienne.

II.5.1 Le principe de calcul d’un facteur de compatibilité normalisé

Dans le cadre Bayésien, la probabilité retrouve sa sémantique naturelle : évaluer l’incertitude allant du connu (les données statistiques) vers l’inconnu (les hypothèses d’intérêt). En combinant les données avec une distribution initiale qui synthétise les informations autres que les données, la règle de Bayes conduit à une distribution finale qui permet de calculer les probabilités des hypothèses.

L’inférence Bayésienne classique dans le cadre de l’analyse de la fiabilité expérimentale, par la formulation mathématique donnée par l’équation II.1, en faisant le produit direct entre la fonction a priori et la vraisemblance, donne le même poids aux deux distributions : la vraisemblance et la connaissance a priori.

Pour les cas où l’une des deux serait trop informative, on observe que l’autre ne compte plus.

Dans le cadre de la démarche présentée dans [Usureau, 2001], il est montré une approche qui permet de pondérer l’information a priori en fonction du degré de compatibilité entre l’a priori et la vraisemblance. Nous proposons d’étudier l’application de cette approche à notre problématique.

Cette méthodologie permet de déterminer un facteur K de pondération normalisé compris entre 0 et 1 qui permet :

- de déterminer le degré de compatibilité de la modélisation de la connaissance a priori avec les données réelles,
- de pondérer la distribution a priori afin de la rendre compatible avec les données observées et pour ne pas biaiser les résultats de l’analyse.

Un tel facteur de pondération doit pouvoir être utilisé dans le cas général, exploitable et efficace dans tous les cas extrêmes possibles.

La méthodologie compare l’information a priori qu’on a pour chaque paramètre θ_i avec l’information contenue par la vraisemblance définie par les fonctions marginales (voir

section I.3.2.2.2).

Ce facteur doit mesurer la similitude entre la distribution a priori initiale modélisée à partir des avis d'experts, des données et/ou simulations sur des produits similaires et la fonction de vraisemblance représentant, elle, les données réelles observées.

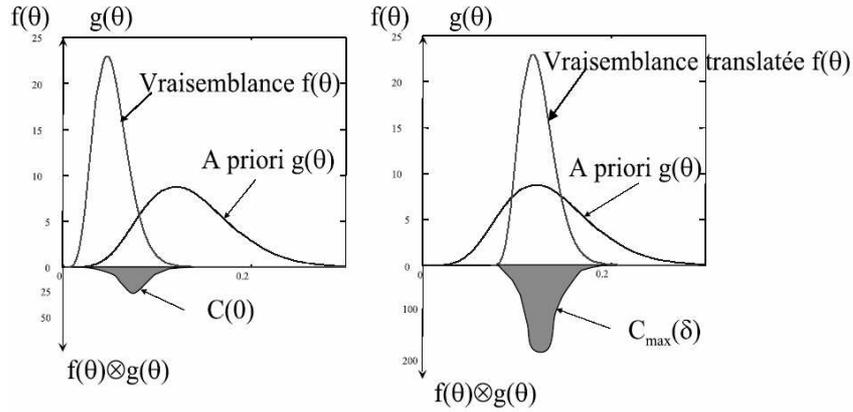


FIG. II.21 – Principe de détermination du coefficient de pondération normalisé

Cette similitude est proportionnelle au produit de convolution (voir Figure II.21) entre l'a priori $g_0(\theta)$ et la vraisemblance $f(\theta)$. Il est ainsi possible de définir le réel $C(\delta)$ par le produit de convolution $g_0(\theta) \otimes f(\theta)$, soit :

$$C(\delta) = \int_{D(\theta)} g_0(\theta) f(\theta - \delta) d\theta \quad (\text{II.39})$$

Certes, on peut procéder d'une autre manière en écrivant $C(\delta) = \int_{D(\theta)} g_0(\theta - \delta) f(\theta) d\theta$, cela ne change en rien le calcul.

La valeur de $C(\delta)$ évolue bien avec la similitude entre la distribution a priori $g_0(\theta)$ et la distribution de vraisemblance $f(p)$ puisque plus les deux fonctions se recoupent et plus $C(\delta)$ est élevée. Dans le cas contraire, plus les deux fonctions ne sont pas similaires et plus $C(\delta)$ est faible.

Il reste maintenant à trouver une solution pour réaliser la pondération de la distribution a priori. Pour ce faire, il suffit de définir un valeur cohérente avec $C(\delta)$ qui évolue uniquement entre 0 et 1 . L'expression normalisée de $C(\delta)$ suffit à répondre à ce problème, notons la K :

$$K = \frac{C(0)}{C_{\max}(\delta)} \quad (\text{II.40})$$

avec

- $C(0)$ obtenue pour $\delta = 0$
- $C_{\max}(\delta)$ représentant la valeur maximum de $C(\delta)$, c'est-à-dire pour simplifier le maximum de l'aire qui se trouve sous la courbe produit des deux fonctions $g_0(\theta)$ et $f(\theta)$ quand on fait glisser les deux courbes l'une sur l'autre (voir figure II.21).

Finalement, K n'est autre qu'un facteur de compatibilité normalisé entre la distribution a priori et la distribution de vraisemblance.

II.5.2 L'utilisation du facteur de compatibilité comme facteur de pondération de l'a priori

Nous disposons maintenant d'un facteur K de pondération qui permet de statuer sur la bonne cohérence entre la modélisation a priori et les essais réels symbolisés par la vraisemblance. Bien entendu, le facteur K est déjà utile dans le sens où il va permettre de statuer sur la pertinence de l'a priori.

Il est possible d'aller plus loin dans la démarche. En effet, quel que soit l'état de la distribution a priori, le facteur K peut être utilisé pour pondérer cette connaissance afin d'améliorer la compatibilité avec la vraisemblance.

Plus K sera proche de 1 et plus la connaissance a priori originale sera intégrée, permettant ainsi d'exploiter au mieux l'un des intérêts des méthodes bayésiennes, à savoir de réaliser des gains considérables sur les plans d'essais de Sécurité/Fiabilité.

Plus K sera proche de 0, moins la connaissance a priori initiale sera intégrée dans l'inférence Bayésienne, évitant ainsi de prendre des décisions pouvant être lourdes de conséquences sur le plan de la Sécurité ou sur le plan économique, ceci sur la base de résultats faux.

Le but ainsi serait de pondérer une distribution de telle façon que la distribution pondérée soit $1/K$ fois moins informative.

L'information de Fisher est une notion qui quantifie l'information relative à un paramètre contenu dans une distribution. Sa définition est fortement liée à la variance I.32. On interprète la variance comme l'espérance des carrés des écarts à l'espérance. Elle permet de caractériser la dispersion des valeurs par rapport à la moyenne. Ainsi, une distribution avec une même espérance et une variance plus grande apparaîtra comme plus étalée.

Le résultat est qu'on ne modifie pas le mode de la distribution a priori, par contre la variance devient plus importante affaiblissant cette distribution. Le facteur de pondération K évoluant avec le niveau de pondération entre a priori et vraisemblance, la modélisation a priori initiale sera d'autant plus intégrée dans l'inférence Bayésienne qu'elle est compa-

tible avec la vraisemblance. Il en sera de même en cas d'incompatibilité entre l'a priori et la vraisemblance car le facteur K sera proche de 0. Le poids de la vraisemblance dans l'inférence Bayésienne sera donc prépondérant. Aussi, la modélisation a priori incompatible au départ sera très peu exploitée.

Ainsi, le premier moment de la nouvelle distribution serait identique avec l'ancien mais le deuxième moment serait $1/K$ fois plus grand.

Pour le calcul pratique nous distinguons deux cas possibles :

- Le cas où l'on peut trouver la solution en résolvant directement les égalités entre les équations décrivant les moments d'ordre 1 et 2.

Par exemple, considérons le cas d'une distribution a priori $f(x)$ de type gamma de paramètres α et λ , de densité de probabilité $f(x) = x^{\alpha-1} \frac{e^{-\frac{x}{\lambda}}}{\Gamma(\alpha)\lambda^\alpha}$, avec l'espérance $\alpha\lambda$ et la variance $\alpha\lambda^2$. La pondération de la distribution a priori par le facteur K donne une nouvelle distribution de paramètres tels que :

$$\begin{cases} \alpha\lambda = \alpha'\lambda' \\ \alpha\lambda^2 = K\alpha'\lambda'^2 \end{cases} \quad (\text{II.41})$$

Qui donnent $\alpha' = K\alpha$ et $\lambda' = \frac{\lambda}{K}$.

- Le cas où, pour résoudre les égalités précédentes, il faut appliquer une méthode numérique d'optimisation.

Par exemple, considérons le cas d'une distribution a priori $f(x)$ de type Weibull de paramètres β et η , de densité de probabilité $f(x) = \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^\beta}$, avec l'espérance $\eta\Gamma\left(1 + \frac{1}{\beta}\right)$ et la variance $\eta^2\Gamma\left(1 + \frac{1}{\beta}\right) - \left(\eta\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2$. La pondération de la distribution a priori par le facteur K donne une nouvelle distribution de paramètres vérifiant :

$$\begin{cases} \eta\Gamma\left(1 + \frac{1}{\beta}\right) = \eta'\Gamma\left(1 + \frac{1}{\beta'}\right) \\ \eta^2\Gamma\left(1 + \frac{1}{\beta}\right) - \left(\eta\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2 = K\eta'^2\Gamma\left(1 + \frac{1}{\beta'}\right) - \left(\eta'\Gamma\left(1 + \frac{1}{\beta'}\right)\right)^2 \end{cases} \quad (\text{II.42})$$

Les paramètres β' et η' sont obtenus en résolvant numériquement ces équations.

A titre d'illustration, les figures suivantes présentent des exemples d'inférence Bayésienne utilisant une distribution a priori pondérée dans la construction de celle correspondant à l'a posteriori en considérant les systèmes à mortalité exponentielle (à gauche) et binômiale (à droite).

Distributions a priori très pessimiste ou optimiste => K très faible

Dans les figures ci-dessous, que la distribution a priori soit très optimiste (figure II.22) ou bien très pessimiste (figure II.23), le facteur de pondération K est très faible.

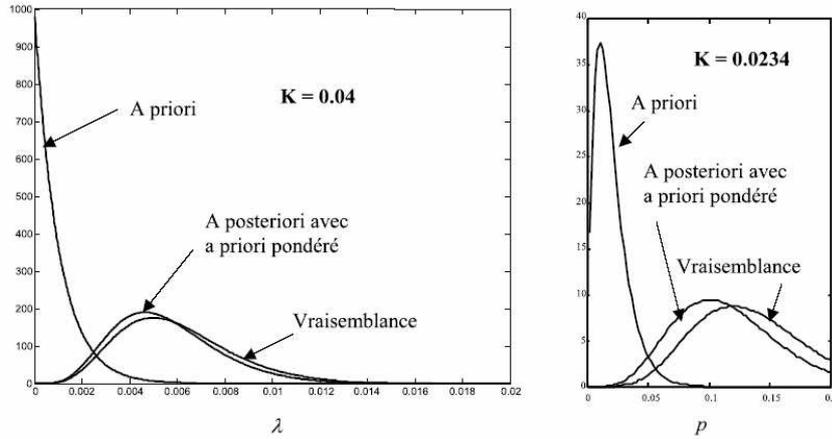


FIG. II.22 – Distribution a priori décalée et trop optimiste

La pondération, par ce facteur, de la distribution a priori permet de donner un maximum de poids à la vraisemblance.

La distribution a posteriori est donc modélisée principalement en exploitant les résultats de la vraisemblance puis, à une moindre échelle, par les données cohérentes de l'a priori. Les résultats sont donc plus proches de la réalité et l'inférence Bayésienne a mieux été optimisée.

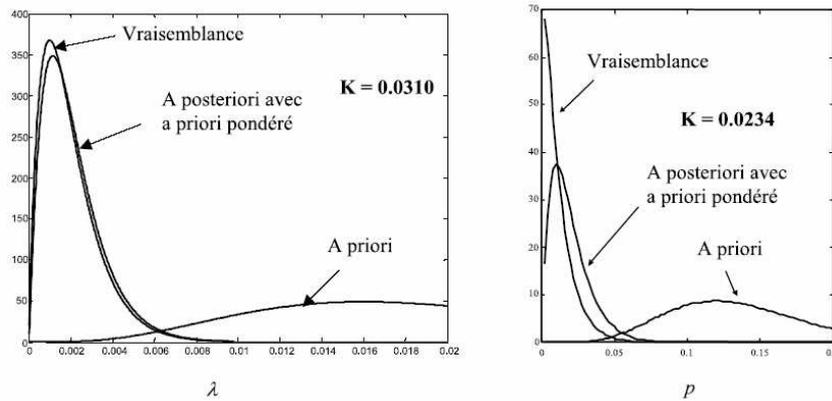


FIG. II.23 – Distribution a priori décalée et trop pessimiste

Distribution a priori non cohérente => K minimum

Dans le cas de la figure II.24, la distribution a priori n'est absolument pas cohérente et beaucoup trop optimiste. La pondération par K , qui cette fois-ci est nul, permet de ne

pas prendre en compte l'erreur liée à l'a priori. La distribution a posteriori est identique à la distribution de vraisemblance, ce cas correspond au cas des essais classiques.

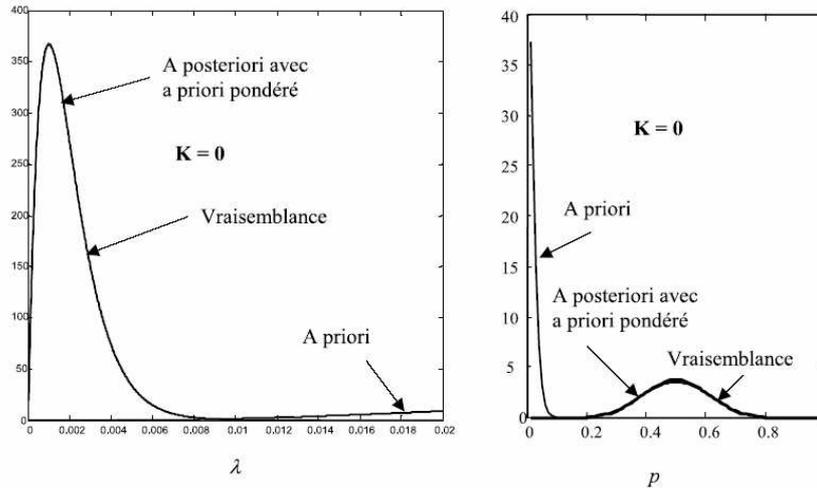


FIG. II.24 – Distribution a priori incohérente

Distribution a priori cohérente => K maximum

Dans le cas de la figure II.25, la pondération entre a priori et la vraisemblance est maximale. La pondération par le facteur K égal à 1 n'a aucun effet sur l'a posteriori.

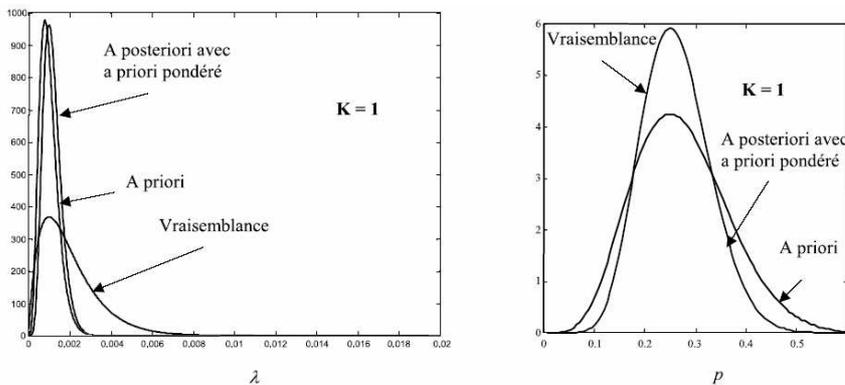


FIG. II.25 – Distribution a priori cohérente

Distribution a priori en partie cohérente => 0 < K < 1

Le cinquième cas (figure II.26) est une illustration de toutes les éventualités comprises lorsque K n'est ni minimum, ni maximum. Dans ce cas, seule la partie cohérente de l'a priori est prise en compte dans l'inférence Bayésienne, laissant plus ou moins de poids à la vraisemblance en fonction de la similitude des deux fonctions. Ceci permet d'inhiber

les erreurs liées à la première modélisation a priori et de fournir des résultats fiables à l'analyste et au décideur.

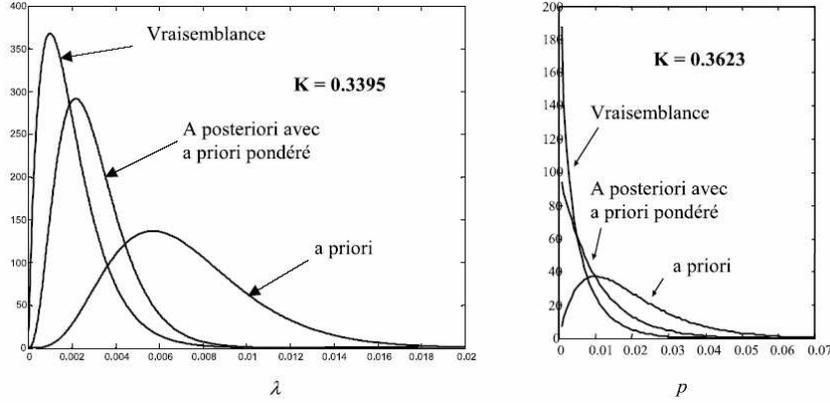


FIG. II.26 – Distribution a priori partiellement cohérente

II.5.3 Application au modèle SVA

La recherche de la compatibilité entre la connaissance a priori et la vraisemblance consiste à déterminer les facteurs K pour chacun des paramètres du modèle SVA considéré (paramétrique ou semi-paramétrique). Pour ce faire, nous déterminons les fonctions marginales ($g_\theta(\theta)$) associées aux paramètres θ ($f_{\theta_i}(\theta_i)$ pour le paramètre θ_i) du modèle par :

$$f_{\theta_i}(\theta_i) = \frac{\int_{j=1 \dots q, j \neq i} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1 \dots q, j \neq i} d\theta_j}{\int_{j=1 \dots q} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1 \dots q} d\theta_j} \quad (\text{II.43})$$

Ensuite, nous déterminons le facteur K_{θ_i} selon la méthode présentée dans le paragraphe II.5.1 en considérant les distributions marginales $f_{\theta_i}(\theta_i)$ et a priori $\Pi_{\theta_i}(\theta_i)$. Une fois les facteurs K_{θ_i} estimés, nous déterminons les distributions a priori pondérées d'après la méthode détaillée dans le paragraphe précédent. Cette étape n'est pas systématique, car l'incompatibilité peut s'expliquer par d'autres causes qu'un problème d'a priori. Ainsi, nous obtenons la nouvelle formulation de l'inférence bayésienne :

$$\pi_{apo}(\theta/t_i) = \frac{L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i})}{\int_{\Omega} L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i}) d\theta_i} \quad (\text{II.44})$$

avec comme notations $L(t_1, \dots, t_n/\theta)$ la vraisemblance (définie par les équations II.2, pour le cas paramétrique, ou II.25, pour le cas semi-paramétrique), $\pi(\theta_i, K_{\theta_i})$ la distribution a priori pondérée pour le paramètre θ_i et $\pi_{apo}(\theta/t_i)$ la fonction a posteriori.

Pour illustrer cette technique de pondération, considérons l'exemple présenté dans le paragraphe II.6.1.3 en étudiant uniquement le paramètre v . La figure II.27 présente les fonctions de densité de l'a priori (courbe 1) et de la vraisemblance (courbe 2) et l'a posteriori (courbe 3). On observe que l'a posteriori est décalé par rapport à la valeur théorique et quasiment superposé à l'a priori. Cela s'explique par le choix d'un a priori incohérent et très informatif.

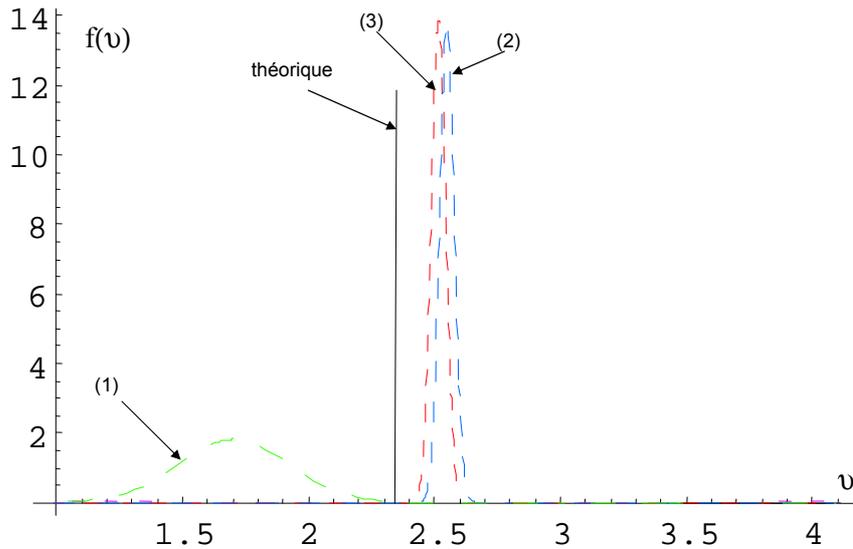


FIG. II.27 – Estimation classique pour v

L'application du critère DAC ([Bousquet, 2006]) donne une valeur de 1.4 ce qui invalide le choix d'a priori pour le jeu de données. Le calcul de facteur de pondération K donne une valeur de 0.135, valeur qui est suffisamment petite pour invalider aussi le choix d'a priori. En plus, en allant plus loin, la pondération de l'a priori avec le facteur $K = 0.135$ donne la courbe (4) de la figure II.28. L'application de démarche Bayésienne sur l'a priori ainsi pondérée donne l'estimation a posteriori (courbe 5), estimation moins informative mais proche de la valeur théorique.

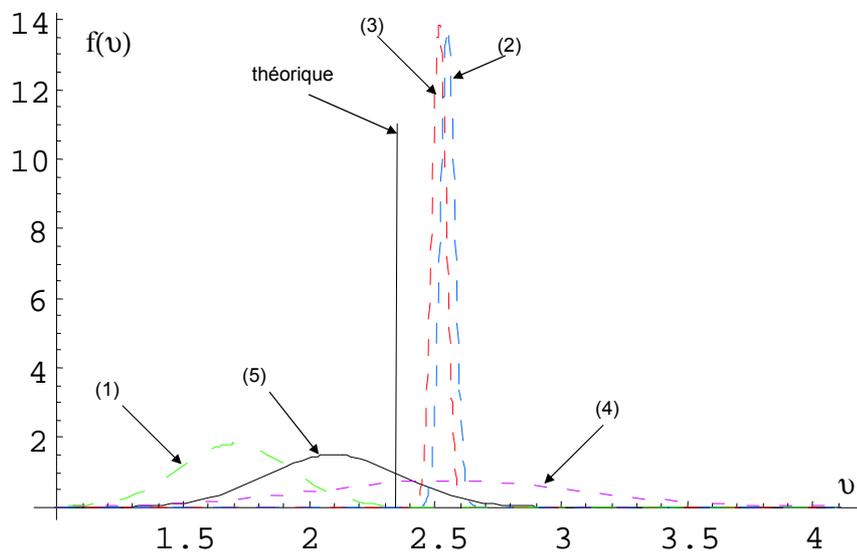


FIG. II.28 – Pondération pour v

II.6 Exemple numérique

L'objectif de cet exemple est d'appliquer la méthodologie de la pondération de l'a priori. Pour ce faire, nous reprenons les mêmes paramètres de simulation et de connaissance a priori de l'exemple du paragraphe II.6.

II.6.1 Résultats

Pour étudier l'influence de la simulation, nous avons répété 10 fois les simulations. Nous reprenons les trois cas de connaissances a priori :

1. un cas d'un a priori cohérent (c'est-à-dire centré sur la valeur théorique) et peu informatif (variance importante),
2. un cas d'un a priori très informatif et cohérent,
3. un cas d'un a priori très informatif et incohérent (décentré par rapport à la valeur théorique).

II.6.1.1 A priori peu informatif et cohérent

Pour le premier cas, une assez faible information a priori sur les paramètres a été considérée. Pour chaque paramètre, un intervalle assez large a été considéré comme définissant les limites des fonctions uniformes a priori caractérisant le groupe de paramètres $\theta = v, \eta, E_a$.

Alors, nous avons les limites suivantes :

- $v \in (0, 4.68)$
- $\eta \in (0, 230)$ mois
- $E_a \in (0, 0.58)$ eV

Au cours de la simulation, les valeurs du facteur K calculées, selon la méthode présentée dans le paragraphe II.5.1, sont comprises dans l'intervalle $[0.88, 0.99]$. Aussi, les estimations sont comparables à celles du paragraphe II.6.1.1 car l'a priori n'est quasiment pas pondéré.

II.6.1.2 A priori très informatif et cohérent

Pour le deuxième cas, un a priori très informatif et cohérent a été considéré. Pour chaque paramètre, un intervalle de $\pm 1/8$ de la valeur théorique et centré sur cette dernière a été considéré.

Alors, nous avons les intervalles suivants :

- $v \in (2.0475, 2.6325)$
- $\eta \in (100, 120)$ mois
- $E_a \in (0.25375, 0.32625)$ eV

Au cours de simulations, les valeurs du facteur K ont été calculées et elles étaient comprises dans l'intervalle $[0.88, 0.99]$. Comme dans le cas précédent, les estimations sont conformes à celles obtenues dans le paragraphe II.6.1.2 puisque l'a priori n'était quasiment pas pondéré.

II.6.1.3 A priori incohérent et informatif

Pour le troisième cas, une mauvaise information a priori sur les paramètres a été considérée. C'est-à-dire que l'on a choisi des intervalles serrés et décentrés des valeurs théoriques.

Nous avons considéré les limites suivantes :

- $v \in (2.5, 2.6)$
- $\eta \in (80, 90)$ mois
- $E_a \in (0.33, 0.35)$ eV

Les résultats sur l'estimation des paramètres sont présentés dans les figures II.29, II.30 et II.31.

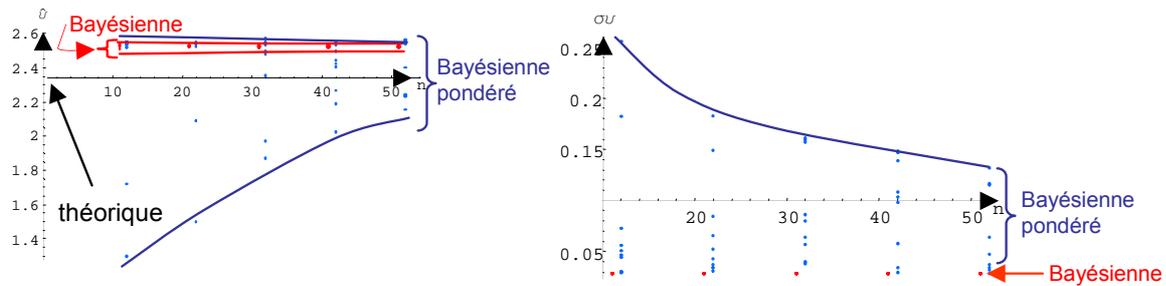


FIG. II.29 – Résultats de la simulation de Monte Carlo pour v et σ_v avec l'intervalle de confiance 90%

Suite aux estimations des paramètres, on peut caractériser, en utilisant l'équation II.22 l'intervalle de confiance sur la fonction $R(t)$.

La figure II.32 présente, pour quatre situations $n_1 = n_2 = 5$, $n_1 = n_2 = 15$, $n_1 = n_2 = 25$ et $n_1 = n_2 = 50$ les estimations sur les fonctions $R(t)$ et les intervalles de confiance pour un niveau de confiance $\alpha = 0.95\%$.

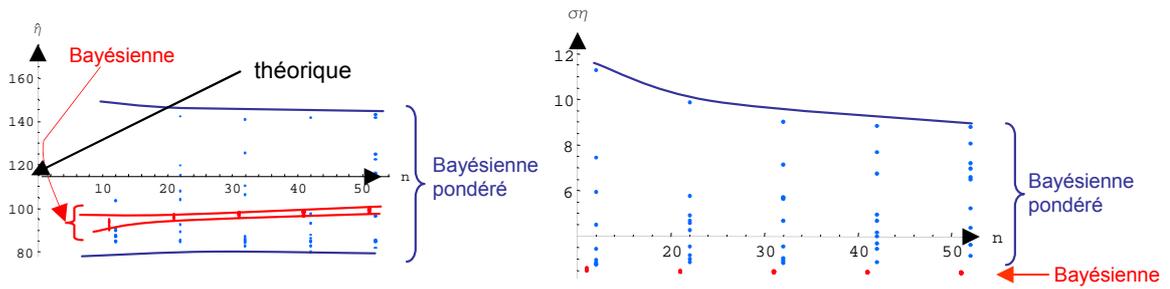


FIG. II.30 – Résultats de la simulation de Monte Carlo pour η et σ_η avec l'intervalle de confiance 90%

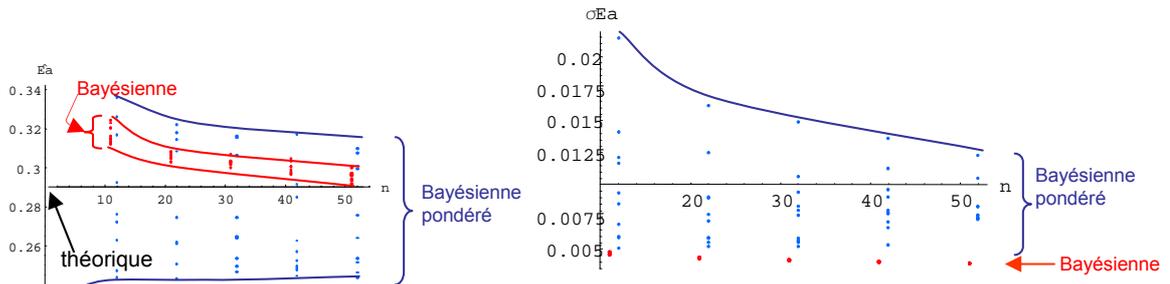


FIG. II.31 – Résultats de la simulation de Monte Carlo pour E_a et σ_{E_a} avec l'intervalle de confiance 90%

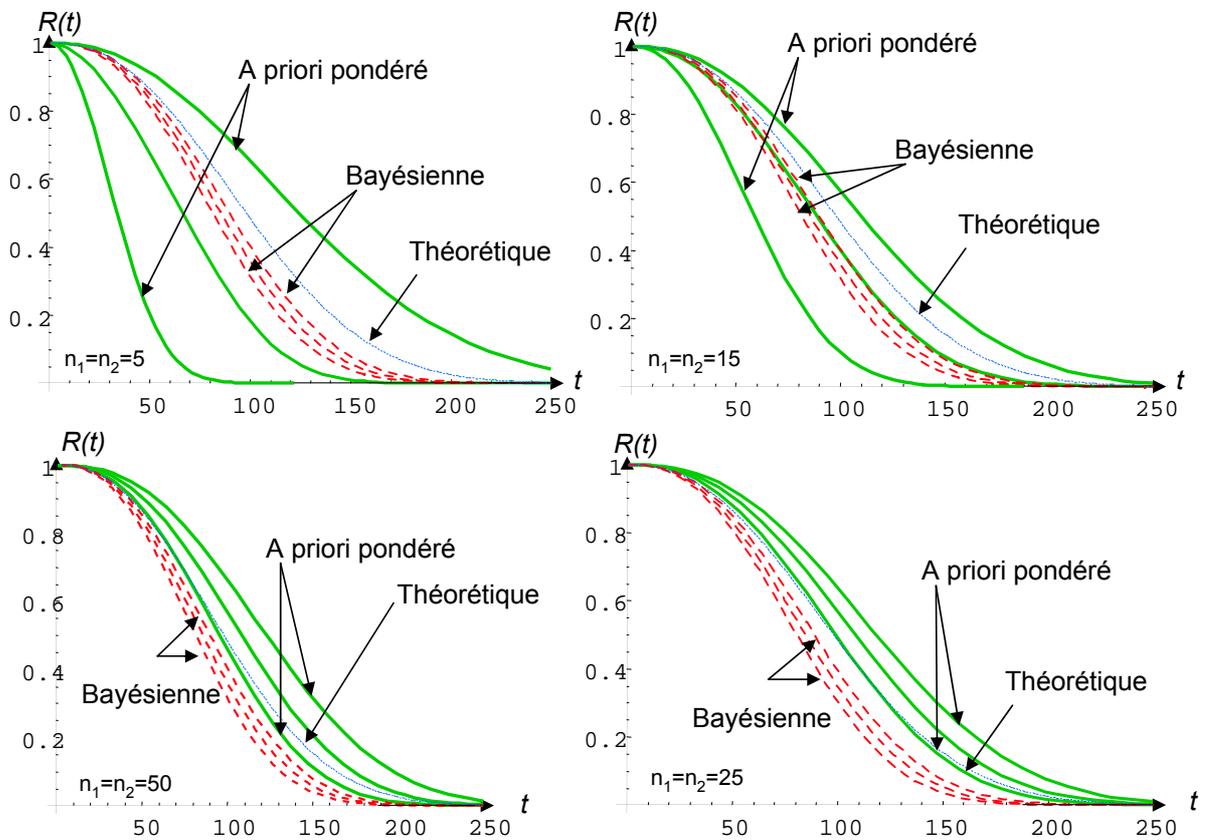


FIG. II.32 – $R(t)$

II.6.1.3.1 Les résultats de la méthode Bayésienne classique La forte information contenue dans l'a priori l'emporte face aux résultats de test. Ainsi, les estimations ne tiennent presque pas du tout compte des résultats d'essai. Ainsi, à cause de l'écart existant entre les intervalles a priori et la valeur théorique, tous les résultats sont décalés par rapport aux vraies valeurs :

- les estimateurs ponctuels de v et de E_a sont surévalués
- l'estimateur ponctuel de η est sous-évalué

Du fait des informations a priori très informatives par rapport aux résultats d'essai, les intervalles de confiance obtenus sont très serrés et décalés des valeurs théoriques.

II.6.1.3.2 Les résultats de la méthode Bayésienne pondérée La forte information a priori n'est pas validée par la méthodologie. Ainsi, des valeurs très faibles des facteurs K sont trouvées et une diminution importante de l'informativité de l'a priori s'impose. C'est pour cela que, en appliquant la pondération dans ce cas, les résultats sont très sensibles aux données d'essai.

En plus, par rapport à l'inférence Bayésienne classique, les intervalles de confiance à 90% contiennent les valeurs théoriques.

II.6.2 Synthèse

En résumé, et sur la base des exemples présentés ci-dessus, la pondération par le facteur de pondération K présente un intérêt évident, ceci quelle que soit la situation à analyser. En effet, dans le cas extrême d'a priori non cohérent, il permet de ne pas utiliser la loi a priori. Ce cas revient à utiliser les méthodes classiques en évitant à l'analyste de s'engager inutilement sur des plans d'essais trop importants et sans aucune justification réelle. A l'opposé, même si l'a priori est parfaitement cohérent, la pondération n'a aucun impact, ce qui revient à réaliser une inférence Bayésienne normale, assurant ainsi d'exploiter au mieux l'intérêt de ces méthodes. Enfin, entre ces deux cas, la pondération par K apporte le fait de n'exploiter que la partie utile de la modélisation a priori, ce qui permet de ne pas biaiser les résultats.

Plus K est proche de 1, plus la distribution a priori doit être prise en considération. Cela se manifestera par la quasi conservation du nombre de défaillances équivalent k_0 et du nombre d'essais équivalent N_0 (dans le cas de systèmes à mortalité binômiale) ou du temps cumulé d'essai équivalent T_0 (dans le cas de systèmes à mortalité exponentielle) obtenus à partir des données du retour d'expériences sur des produits similaires, des résultats d'études de conception, des avis d'experts, ...

En revanche, si le facteur K est inférieur à 1, on peut décider soit de rejeter l'a priori pour une valeur de K inférieure à une valeur limite, soit pour "affaiblir" l'a priori. On admet alors des informations moins "riches" pour l'a priori qui se manifeste par un nombre de défaillances réduit, une taille d'échantillon (dans le cas de système à mortalité binômiale), un temps cumulé d'essai (dans le cas de système à mortalité exponentielle) et sans pour autant modifier le maximum de la distribution a priori.

Seul l'écart type de la distribution est affecté d'une manière d'autant plus importante que la taille d'échantillon ou le temps cumulé diminue.

A l'inverse, si $K = 1$, il n'existe aucune raison de rejeter l'a priori et l'essai réel à réaliser sera réduit.

Au fur et à mesure que l'on va enrichir les essais réels, on va "renforcer" la vraisemblance et l'on peut s'attendre à ce que K diminue.

Quand la taille de l'échantillon d'essais réels ou le temps cumulé d'essais réels tend vers l'infini, on retrouve le résultat des essais classiques car $K \rightarrow 0$ (on ne tient plus compte de l'a priori).

II.7 Conclusion

Ce chapitre a présenté l'introduction de la démarche bayésienne dans l'estimation des paramètres des modèles SVA. Elle permet de réduire les intervalles de confiance en apportant une connaissance a priori. La démarche a été appliquée aux modèles paramétrique et semi-paramétrique en étudiant les estimations classiques et MAP (Maximum of A Priori).

Ensuite, nous avons traité un aspect important dans la démarche bayésienne en proposant une méthodologie de construction de la distribution a priori :

Pour finir, nous avons abordé un aspect peu étudié qui est la compatibilité entre la priori (fiabilité présumée du système et son incertitude) et la vraisemblance (données réelles de fiabilité). Ainsi, nous avons proposé un indicateur de compatibilité, noté K et compris entre 0 et 1, permettant par la suite de pondérer la priori. Les nombreux exemples traités montrent l'efficacité de la méthode proposée. Néanmoins, une étude complémentaire est nécessaire afin de tester cette méthode de pondération lorsque l'incompatibilité est uniquement sur une partie des paramètres. Il reste à étendre la démarche bayésienne aux autres modèles SVA (non paramétrique, Cox, ...) et plan (stress échelonné, ...).

Le modèle SVA permet ensuite d'estimer la fiabilité en conditions nominales. Cette étape est importante puisqu'elle permet de vérifier la tenue des objectifs contractuels. Pour ce faire, il est nécessaire de connaître le profil de mission du produit. Dans de nombreux cas, il est difficile d'obtenir cette information car le produit peut être utilisé dans différents environnements, par divers utilisateurs, ... Nous proposons dans le chapitre suivant d'intégrer cette variabilité dans l'estimation de la fiabilité.

III

Estimation de la fiabilité en environnement aléatoire

Sommaire

III.1 Introduction	116
III.2 Définition des profils de mission	120
III.3 La variabilité d'usage	122
III.3.1 Estimation directe de la fonction de fiabilité	122
III.3.2 Exemple d'illustration	123
III.3.3 Estimation de la fiabilité par la méthode contrainte-résistance .	125
III.4 Variabilité temporelle des sollicitations	127
III.4.1 introduction	127
III.4.2 Stress défini par une variable aléatoire indépendante du temps .	128
III.4.2.1 Principe de l'analyse	128
III.4.2.2 Caractérisation de la variable aléatoire	130
III.4.2.3 Histogramme ou temps de maintien dans une classe d'amplitudes	134
III.4.2.4 Comptage du nombre de dépassements d'un niveau donné	134
III.4.3 Estimation de la fiabilité	136
III.4.4 Séries temporelles	136
III.4.5 Discrétisation du profil de stress	139
III.4.5.1 Principe de la discrétisation	139
III.4.5.2 Modèle d'endommagement cumulé (MEC)	139
III.5 Exemple d'application	142
III.6 Conclusion	147

III.1 Introduction

Les modèles de vie décrits dans les deux chapitres précédents permettent une estimation de la fiabilité à n'importe quel niveau du stress. Suite aux tests en laboratoire, en intégrant toutes les connaissances a priori, une estimation de la fiabilité peut être obtenue. Mais, il y a souvent des différences importantes entre les estimations obtenues en laboratoire et le comportement réel du produit. Pour bien approximer le comportement réel de la vie du produit, on est obligé de tenir compte de la variabilité d'utilisation du système.

La prise en compte des conditions réelles d'usage devient un enjeu technologique et économique. Mais elle entraîne un profond changement d'attitude puisque les causes sont envisagées d'une manière probable à partir de suppositions d'utilisation par un segment significatif d'une population. Et, naturellement, le calcul des effets sera présenté en termes de probabilité de contrainte et de fiabilité [Lannoy, 2004]. Cette approche est possible car deux maillons de la modélisation sont bien maîtrisés actuellement. Tout d'abord, les conséquences des chargements variés appliqués sur les composants peuvent être analysées et calculées en termes de surcharges dynamiques [Savard, 2004] [Bedard, 2000]. L'ingénieur concepteur peut alors mettre en place des méthodes de calcul conciliant au mieux les connaissances actuelles et les objectifs qu'il doit atteindre. Ainsi, en amont de l'approche, il lui faut connaître les chargements à partir des conditions imposées par les utilisateurs. Et, en aval de l'approche, il lui faut calculer les conséquences de ces chargements.

Dans le but d'estimer une probabilité de défaillance du produit au cours du temps, une analyse de l'environnement de ce produit soumis à une variabilité des conditions d'utilisation et de fonctionnement est nécessaire. Par exemple, des produits identiques utilisés par différents utilisateurs peuvent être sollicités différemment. Alors, il est indispensable de tenir compte de cette variabilité dans l'estimation de la fiabilité.

Un autre aspect est la diversité d'action du temps agissant sur les composants d'un système. En effet, certains mécanismes de défaillance sont sensibles au temps calendaire, au nombre de cycles, au kilométrage, ... De plus, certains équipements peuvent être garantis selon plusieurs modalités temporelles (par exemple pour l'automobile, les garanties peuvent être définies pour 5 ans ou 100 000 km).

Dans la suite du travail, nous aborderons la variabilité à travers 3 notions :

1. Les phases de vie correspondent à des modalités de fonctionnement différentes du produit. On peut citer les différentes phases de vie d'un avion :
 - fonctionnement (décollage, vol, atterrissage, fonctionnement agressif ou non , etc.)

- repos
 - maintenance
 - stockage, etc.
2. Les missions sont associées à des usages différents du produit qui impliquent des répartitions spécifiques des phases de vie. En reprenant l'exemple de l'aéronautique, on peut citer les missions suivantes pour un avion :
- long courrier,
 - moyen courrier,
 - low cost,
 - etc.
3. L'environnement définit les conditions dans lesquelles le produit fonctionne. Les éléments d'un environnement peuvent être :
- climatique (température, cycles de température, humidité, etc.)
 - mécanique (vibrations, chocs, etc.)
 - chimique (pollution, etc.),
 - électrique, (tension, intensité, décharge, etc.) etc.

Aujourd'hui, on vient modéliser toute la variabilité par des profils de mission qui décrivent pour une mission donnée (par exemple pour un moyen courrier), les différentes phases de vie (fonctionnement, repos, maintenance, ... de l'avion) et les niveaux de stress par phase. La figure III.1 présente un profil de mission pour une carte électronique montée en soute dans un moyen-long courrier.

Profil d'emploi de type aéronautique civil moyen / long courrier, calculateur en soute														
Contrainte		Thermique et Humidité			Cyclage thermique				Mécanique	Chimique			Induit	
Intitulé de la phase	Temps calendaire (heures)	On/Off	Température ambiante (°C)	Taux d'humidité (%)	ΔT (°C)	Nombre de cycles (/an)	Durée du cycle (heures)	Température maximale au cours du cyclage (°C)	Vibrations aléatoires (Gms)	Pollution saline	Pollution industrielle	zone d'application	Niveau de protection	Π application
Sol-operating - 1	797	On	47	30	33	330	2	47	-	Faible	Urbaine	Non habitable	Non hermétique	4.8
Sol-operating - 2	1193	On	55	30	15	647	1,5	55	-	Faible	Urbaine	Non habitable	Non hermétique	2.0
Sol-roulage	84	On	47	5	-	-	-	-	5	Faible	Négligeable	Non habitable	Non hermétique	1.2
Vol-operating	4083	On	40	5	-	-	-	-	0,6	Faible	Négligeable	Non habitable	Non hermétique	1.1
Sol-dormant	2603	Off	14	70	10	108	24	19	-	Faible	Urbaine	Non habitable	Non hermétique	3.3

FIG. III.1 – Profil d'emploi de type aéronautique civil moyen / long courrier [FIDES, 2004]

Néanmoins, ces profils sont souvent définis de manière conservatrice (surévaluation des niveaux de stress) pour minimiser les risques. Pourtant, dans la réalité les niveaux de stress peuvent fluctuer dans des zones plus faibles (ou plus hautes). A titre d'exemple, la figure III.2 présente l'évolution temporelle de la température. Quand cette information est disponible, nous proposons de l'introduire dans l'estimation de la fiabilité.

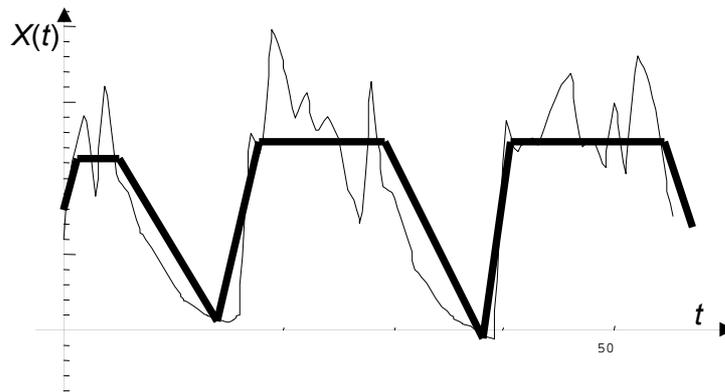


FIG. III.2 – Estimation d'un profil de mission

Nous proposons, dans le cadre de cette thèse, de traiter la question de la prise en compte de la variabilité lors de l'estimation de la fiabilité. Pour ce faire, nous allons considérer plusieurs cas, correspondant à différentes situations rencontrées dans l'industrie, qui intègrent plus ou moins la variabilité. Ainsi, nous proposons trois approches d'estimation de la fonction de fiabilité :

1. la première prend en compte uniquement un profil de mission déterministe,
2. la deuxième permet de prendre en compte la variabilité de l'usage (diversité d'utilisateurs). A l'aide d'une approche Contrainte-Résistance, nous pouvons estimer la fiabilité d'un équipement pour un temps donné. Ainsi, cette approche permet de déterminer la proportion défaillante d'une population durant, par exemple, une période de garantie,
3. la dernière approche consiste à intégrer l'évolution temporelle des sollicitations (température, pression, humidité, ...) dans l'estimation de la fonction de fiabilité. Cette approche nécessite de suivre en temps réel (ou périodiquement) le fonctionnement et les conditions environnementales des produits. Ceci est possible à l'aide d'une instrumentation spécifique ou dans le cadre de programmes "Health monitoring" qui permettent de prévoir des maintenances en fonction du niveau réel de dégradation du produit. Cette dégradation est directement liée aux niveaux de sollicitation du produit. Dans le cadre de cette thèse, nous proposons d'utiliser trois méthodes :
 - estimation en intégrant les stress définis par des distributions indépendantes du temps,
 - estimation de la fonction de fiabilité intégrant les séries temporelles des stress,
 - estimation en considérant un profil discrétisé.

L'ensemble de ces approches nécessite de disposer de tous les modèles SVA associés aux

différents stress listés dans les environnements. Dans le paragraphe suivant, nous présentons comment sont déterminés les profils de mission et quelles seront les informations disponibles.

III.2 Définition des profils de mission

Le profil de mission décrit de quelle manière un produit sera utilisé et dans quelles conditions. Aussi, la construction d'un profil de mission est déterminante dans le dimensionnement du produit et dans l'évaluation prévisionnelle de sa fiabilité. En effet, sa tenue en exploitation dépendra en grande partie de la couverture du profil de mission des conditions opérationnelles. Cette construction se fait par la capitalisation de toute l'information disponible sur l'utilisation de produits équivalents, de versions antérieures, ..., ou issue d'avis d'experts. Généralement, le profil est composé de deux volets (voir figure III.2) :

- un premier qui décrit les différentes phases de vie du produit (fonctionnement, repos, entretien, ...)
- un deuxième qui associe à chaque phase de vie les niveaux de sollicitation (température, humidité, mécanique, ...),

Dans la réalité, chaque produit sera employé par un utilisateur particulier, qui aura son propre usage, dans un environnement spécifique en fonction, par exemple, du lieu d'exploitation. Aussi, un profil de mission doit tenir compte de ces variabilités et il est nécessaire d'introduire celles-ci dans l'estimation de la fiabilité. Ainsi, la figure III.3 présente une démarche possible permettant de collecter et de caractériser la variabilité d'usage d'une population d'utilisateurs.

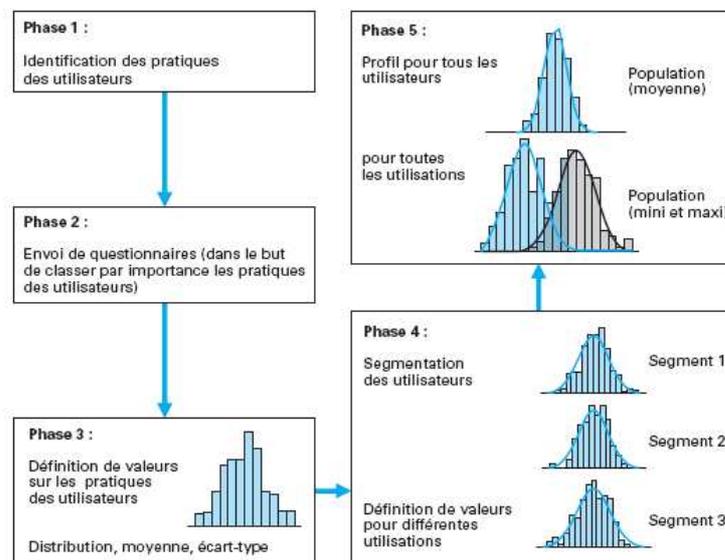


FIG. III.3 – Prise en compte de conditions d'usage [et R. KOUTA, 2007]

Dans le cas où l'on souhaite combiner plusieurs types de sollicitations, il est possible de récupérer un histogramme qui nous donne le temps passé dans chaque combinaison de

sollicitations. A titre d'exemple, la figure III.4 présente le nombre d'heures associé à différentes combinaisons de températures et de vibration pour un équipement aéronautique

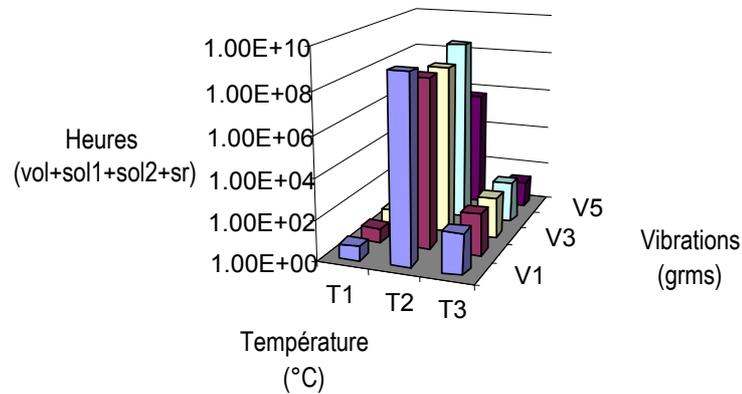


FIG. III.4 – Nombre des heures par température et vibration

Il est également possible d'enregistrer en temps réel les niveaux de sollicitation subis par le produit (III.5).

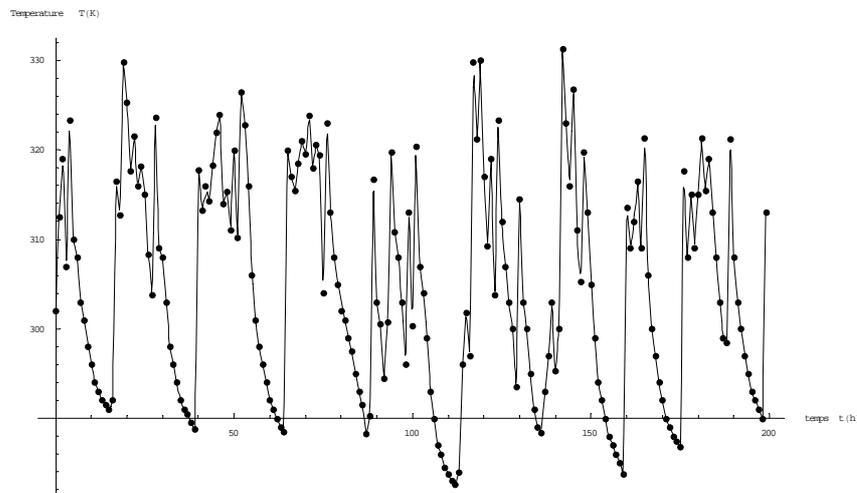


FIG. III.5 – Exemple de variation de la température pour une voiture [Smith, 2002]

Finalement, toutes ces informations sont synthétisées dans le profil de mission. Généralement pour simplifier les estimations, les variabilités sont résumées par un profil majorant et déterministe (voir le profil défini figure III.2). Nous proposons dans le cadre de cette thèse de traiter la variabilité à deux niveaux :

1. le premier concerne la prise en compte de la variabilité d'usage (plusieurs phases de vie, différents utilisateurs),

2. le deuxième permet d'intégrer la variabilité des sollicitations au cours du temps.

Ensuite, il est possible d'imbriquer ces deux niveaux.

III.3 La variabilité d'usage

Dans ce paragraphe, nous considérons la variabilité associée à plusieurs phases de vie et à différents utilisateurs. Nous faisons l'hypothèse que la variabilité concerne principalement une répartition différente des durées de phase et du temps d'utilisation.

III.3.1 Estimation directe de la fonction de fiabilité

Dans cette section, nous allons nous intéresser à introduire toute la variabilité dans la construction de la fonction de fiabilité.

Notons le stress S_{kji} de type i ($i = 1, \dots, m_{jl}$), pour la phase de vie j ($j = 1, \dots, n_l$) et pour l'utilisateur k ($k = 1, \dots, l$). Le taux de passage par la phase de vie j pour l'utilisateur k est donnée par p_{kj} . La durée de la phase de vie j pour l'utilisateur k pour la période τ est donnée par $T_{kj} = \tau \cdot p_{kj}$

Le fait d'examiner la répartition des durées des phases de vie sur une période donnée, permettra d'intégrer la variabilité d'utilisation. En effet, si l'on considère le kilométrage d'une population de voitures au bout d'un an de roulage, on observe une variabilité importante. En considérant que le niveau de stress S_{kji} de type i pour la phase de vie j est constant sur la période T_{kj} , alors on peut écrire la fiabilité par :

$$R_{ki}(\tau) = e^{-\int_0^{\tau \cdot p_{kj}} \lambda(u, S_{kji}) du} = e^{-\sum_{j=1}^{n_l} \int_{\tau \cdot p_{k,j-1}}^{\tau \cdot p_{k,j}} \lambda(u, S_{kji}) du} = \prod_{j=1}^{n_l} R_{S_{ij}}(\tau \cdot p_{kj}) \quad (\text{III.1})$$

avec

- τ le temps exprimé selon l'unité de la période choisie (par exemple en année)
- $\lambda(u, S_{kji})$ le taux de défaillance sous le stress S_{kji} et au temps u
- $T_{k,0} = 0$
- R_{kji} la fiabilité associée au stress i , à la phase de vie j et à l'utilisateur k (toutes phases de vie confondues).
- R_{ki} la fiabilité associée au stress i et à l'utilisateur k (toute phase de vie confondue).

Pour finir, si on considère que le stress i engendre un mécanisme de défaillance indépendant des autres, alors il est possible d'écrire la fonction de fiabilité globale pour

l'utilisateur k à partir des modèles SVA :

$$R_k(\tau) = \prod_{j=1}^{n_l} \prod_{i=1}^{m_{jl}} R_{S_{ij}}(\tau \cdot p_{kj}) \quad (III.2)$$

Ensuite, il est possible de tracer la fonction de fiabilité pour chaque utilisateur k (voir figure III.6).

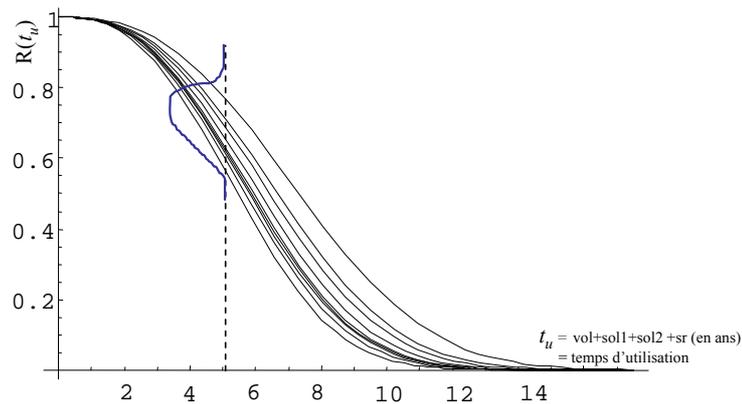


FIG. III.6 – Courbes de fiabilité pour plusieurs clients

III.3.2 Exemple d'illustration

Nous proposons un exemple sur un calculateur de soute embarquée dans un moyen/long courrier. Le profil est donné, pour un seul utilisateur, par le tableau suivant :

Profil d'emploi de type aéronautique civil moyen / long courrier, calculateur en soute														
Contrainte		Thermique et Humidité			Cyclage thermique				Mécanique	Chimique				Induit
Intitulé de la phase	Temps calendaire (heures)	On/Off	Température ambiante (°C)	Taux d'humidité (%)	ΔT (°C)	Nombre de cycles (/an)	Durée du cycle (heures)	Température maximale au cours du cyclage (°C)	Vibrations aléatoires (Gms)	Pollution saline	Pollution industrielle	zone d'application	Niveau de protection	II application
Sol-operating - 1	797	On	47	30	33	330	2	47	-	Faible	Urbaine	Non habitable	Non hermétique	4.8
Sol-operating - 2	1193	On	55	30	15	647	1,5	55	-	Faible	Urbaine	Non habitable	Non hermétique	2.0
Sol-roulage	84	On	47	5	-	-	-	-	5	Faible	Négligeable	Non habitable	Non hermétique	1.2
Vol-operating	4083	On	40	5	-	-	-	-	0,6	Faible	Négligeable	Non habitable	Non hermétique	1.1
Sol-dormant	2603	Off	14	70	10	108	24	19	-	Faible	Urbaine	Non habitable	Non hermétique	3.3

FIG. III.7 – Profil d'emploi de type aéronautique civil moyen / long courrier [FIDES, 2004]

Nous avons ainsi les phases suivantes dont les taux p_j sont obtenues par la proportion du temps calendaire associé à chaque phase pour un an. Ainsi :

- P_1 : sol-operating 1, $p_1 = 9.10\%$
- P_2 : sol-operating 2, $p_2 = 13.62\%$
- P_3 : sol-roulage, $p_3 = 0.96\%$

stress	loi acceleration	paramètres du modèle
H, T	Peck : $\eta(H, T) = CH^{-n}e^{\frac{B}{T}}$	$n = 1.24, B = 1638.5, C = 0.411, v = 2.67$
ΔT	Coffin-Manson : $\eta(\Delta T) = (\frac{a}{\Delta T})^n$	$a = 3.87E + 4, n = 1.24, v = 2.33$
$Grms$	Inverse Power : $\eta(V) = (\frac{a}{V})^n$	$a = 0.08, n = 2.19, v = 1.46$

TAB. III.1 – Lois de fiabilité associées et estimations de paramètres

- P_4 : vol-operating, $p_4 = 46.38\%$
- P_5 : sol-dormant, $p_5 = 29.71\%$

Nous avons associé à chaque phase P_i les 3 stress suivants ($m_j = 3$) :

- H, T : température, humidité
- ΔT : température cyclique
- $Grms$: vibrations

Les lois de fiabilité associées à chaque stress sont des lois de Weibull de paramètres v (forme) et η (échelle) définis par des lois d'accélération présentées dans le tableau III.2.

Ainsi, l'évaluation de la fiabilité globale du système s'exprime :

$$R_{\text{sys}}(t) = \left\{ \prod_{j=1}^5 [R_{P_j}^{H,T}(p_j t)] \right\} \left\{ \prod_{j=1}^5 [R_{P_j}^{\Delta T}(p_j t)] \right\} \left\{ \prod_{j=1}^5 [R_{P_j}^{Grms}(p_j t)] \right\} \quad (\text{III.3})$$

Nous pouvons déterminer la fonction de fiabilité globale en fonction du temps calendaire (figure III.8)

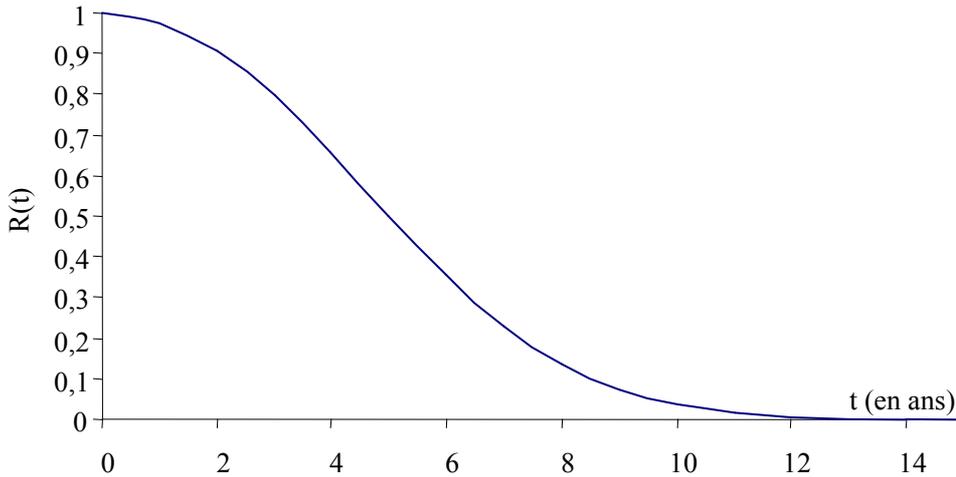


FIG. III.8 – $R(t)$ conforme au Tab. III.2

Cette première approche considère un utilisateur moyen employant l'équipement dans des conditions données. L'approche prend en compte que très peu la variabilité (usage,

phase de vie, environnement, etc.) et elle conduit assez souvent à une sous estimation de la fiabilité car on se place naturellement dans le pire cas.

III.3.3 Estimation de la fiabilité par la méthode contrainte-résistance

Dans cette méthode, on considère indépendamment la contrainte exercée par les différents utilisateurs et la résistance du composant à ces contraintes définie par la loi de fiabilité ([Marcovici et Ligeron, 1974], [Villemeur, 1988]). C'est-à-dire que la variabilité est matérialisée par un taux d'utilisation différent. On peut citer l'exemple de l'automobile où la variabilité d'usage s'exprime par une distribution du kilométrage pour une période donnée (par exemple une loi Log-Normale). Dans ce cas, on suppose que chaque véhicule subira proportionnellement le même profil. Ainsi, il est possible de construire la fonction de fiabilité selon l'équation III.2 où la variable temporelle τ est le kilométrage. Dans ce cas, il est possible d'utiliser la méthode contrainte-résistance. Elle est employée couramment pour le dimensionnement de pièces mécaniques. Récemment, cette méthode est apparue dans le dimensionnement des essais de validation, le principe restant le même [Perroud, 2000], [Poncet et al., 2004a], [Poncet et al., 2004b]. Dans ce cas, la contrainte n'est plus mécanique mais une sollicitation appliquée au composant comme un nombre de cycles, un nombre d'heures.

Le calcul de fiabilité consiste à déterminer la probabilité que la résistance à la contrainte de la pièce étudiée soit supérieure aux contraintes effectivement appliquées pendant toute la mission considérée.

On a donc :

$$R = Prob(C < \rho) \quad (III.4)$$

La probabilité d'avoir une résistance à la contrainte ρ supérieure à C_1 est :

$$P[\rho > C] = \int_C^\infty f_2(\rho) d\rho \quad (III.5)$$

La fiabilité est donc :

$$R = \int_{-\infty}^\infty f_1(C) \left[\int_{-\infty}^\infty f_2(\rho) d\rho \right] dC \quad (III.6)$$

En considérant au départ une résistance à la contrainte ρ_2 , on peut montrer que :

$$R = P[\rho > C] = \int_{-\infty}^\infty f_2(\rho) \left[\int_{-\infty}^{\rho_2} f_1(C) dC \right] d\rho_2 \quad (III.7)$$

$$R = 1 - Pf = 1 - \int_0^{\infty} \left(\int_0^c f_R(u) \cdot du \right) \cdot f_C(c) \cdot dc = 1 - \int_0^{\infty} F_R(c) \cdot f_C(c) \cdot dc \quad (\text{III.8})$$

Ainsi, la fiabilité, tenant compte de la résistance définie par l'équation III.2, s'écrit :

$$R(t) = 1 - \int_0^{\infty} \left(\int_0^c \prod_{i=1}^{j=1} \prod_{m_{ji}} f_{S_{ij}}(u) \cdot du \right) \cdot f_C(c, \tau) \cdot dc \quad (\text{III.9})$$

avec $f_{S_{ij}}(u)$ la fonction de densité du modèle SVA pour le stress S_{ij} et $f_C(c, t)$ la contrainte d'utilisation pour le temps t .

La figure III.9 présente un exemple de résultats de l'application de la méthode contrainte-résistance. Dans cette figure, nous avons présenté deux distributions *contraintes* pour un et deux ans d'utilisation (distributions à gauche) et la résistance (la fonction de fiabilité définie par l'équation III.2)

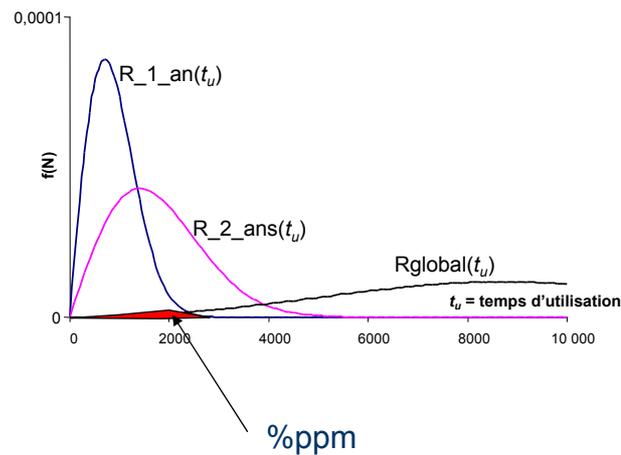


FIG. III.9 – Estimation de la proportion de défectueux par la méthode Résistance/contrainte

Ainsi, une probabilité de défaillance peut être estimée après un an (ou deux ans) de fonctionnement.

III.4 Variabilité temporelle des sollicitations

III.4.1 introduction

La notion de *stress* donne lieu à un historique de sollicitations (appelé aussi trajectoire). Cet événement est un état de charge de service, produisant, au sein de chaque composant considéré, une variation de sollicitations.

Le comptage de l'historique de sollicitations devient une méthode de plus en plus attractive suite aux développements techniques permettant l'intégration de fonctions du type *health monitoring* dans des cartes électroniques. Il s'agit en fait de fonctions qui permettent de suivre l'évolution de l'état de la carte et qui mesure en temps réel l'évolution des sollicitations, permettant ainsi de pouvoir tracer l'historique de ces sollicitations.

Un événement stress (une situation d'usage donnée), avec des paramètres d'exploitation bien déterminés (l'endroit où les observations s'effectuent), correspond en général à une trajectoire de chargements et de sollicitations. Par exemple, la figure III.10 présente un signal temporel mesuré et digitalisé.

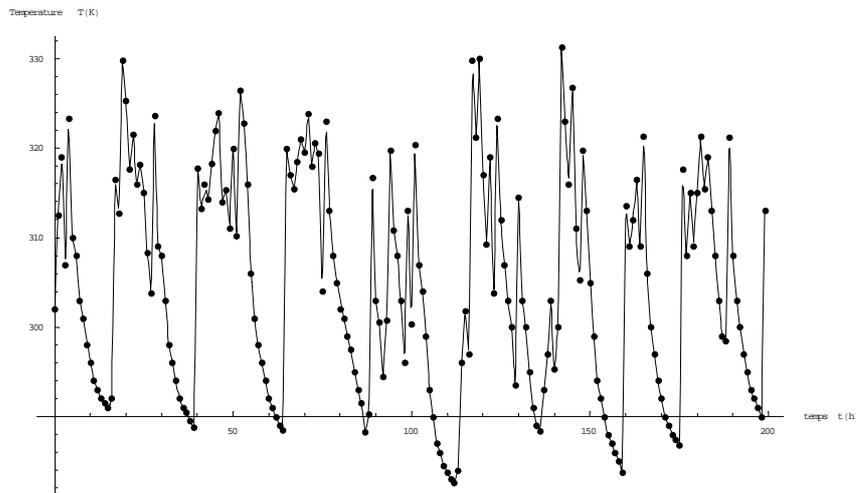


FIG. III.10 – Exemple de variation de la température pour une voiture [Smith, 2002]

Ce dernier est défini sur une durée relativement courte par rapport à la durée totale de vie du produit. En considérant que l'échantillon est suffisamment représentatif pour décrire la vie globale du produit, on peut modéliser le comportement de l'évolution de stress au cours du temps par des séries temporelles. Ainsi, nous proposons dans le cadre de cette thèse, d'intégrer cette variabilité dans l'estimation de la fiabilité en considérant l'évolution de la sollicitation par trois approches (voir figure III.11) :

1. Variable aléatoire indépendante du temps

2. Série temporelle
3. Discrétisation du profil

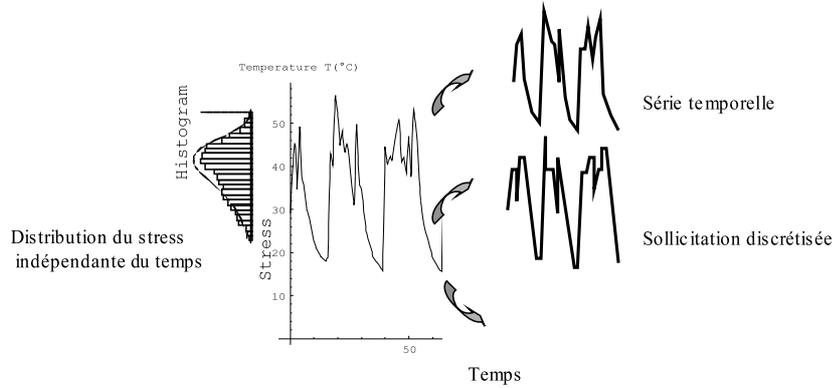


FIG. III.11 – Trois approximations du profil en température

III.4.2 Stress défini par une variable aléatoire indépendante du temps

III.4.2.1 Principe de l'analyse

À partir de la trajectoire du stress ou cours du temps, le problème consiste à obtenir les informations nécessaires pour disposer d'un histogramme, ou d'une loi de distribution de sollicitations que l'on appelle spectre de charges ou de sollicitations [Grubisic, 1994], et qui n'est en réalité qu'une représentation approchée de l'ensemble des charges ou sollicitations appliquées. On remarque aussi que l'obtention de spectres de charges réduit l'information, en ce sens que l'on perd l'aspect chronologique des cycles de variations des charges. Le calcul ultérieur de l'endommagement ne pourra donc pas tenir compte d'une éventuelle interaction entre les cycles successifs de variations de contraintes dues à ces charges. On peut, toutefois, admettre que beaucoup de phénomènes sont largement aléatoires et qu'il est illusoire, au stade de la prévision du comportement des systèmes et composants mécaniques, de prétendre avoir une quelconque connaissance de l'ordre précis d'apparition, par exemple, des valeurs des étendues de variations de contraintes. On met ainsi l'accent sur l'étude de la distribution statistique de l'étendue de la variation de contrainte. Et pour certaines applications, la contrainte moyenne de chaque cycle est parfois retenue. On supposera, dans la suite du mémoire, que la moyenne globale des sollicitations est nulle sur la durée de la trajectoire.

Bien que la prise en compte de l'évolution complète du stress au cours de la vie du produit est l'approche la plus adaptée, c'est loin d'être possible dans la pratique. Même

avoir une description complète de l'évolution du stress au cours du temps d'un échantillon n'est pas suffisant car il y a très peu de situations dans lesquelles une population de produits rencontre le même environnement de fonctionnement. Aussi, l'analyse de la représentativité des valeurs mesurées est un point important à discuter.

Hormis quelques cas particuliers de processus (trajectoire périodique sinusoïdale, processus stationnaire gaussien à bande étroite, c'est-à-dire avec peu de fréquences d'excitation caractéristiques), il est en général difficile d'associer une étendue de variation de contrainte à un cycle (figure III.12).

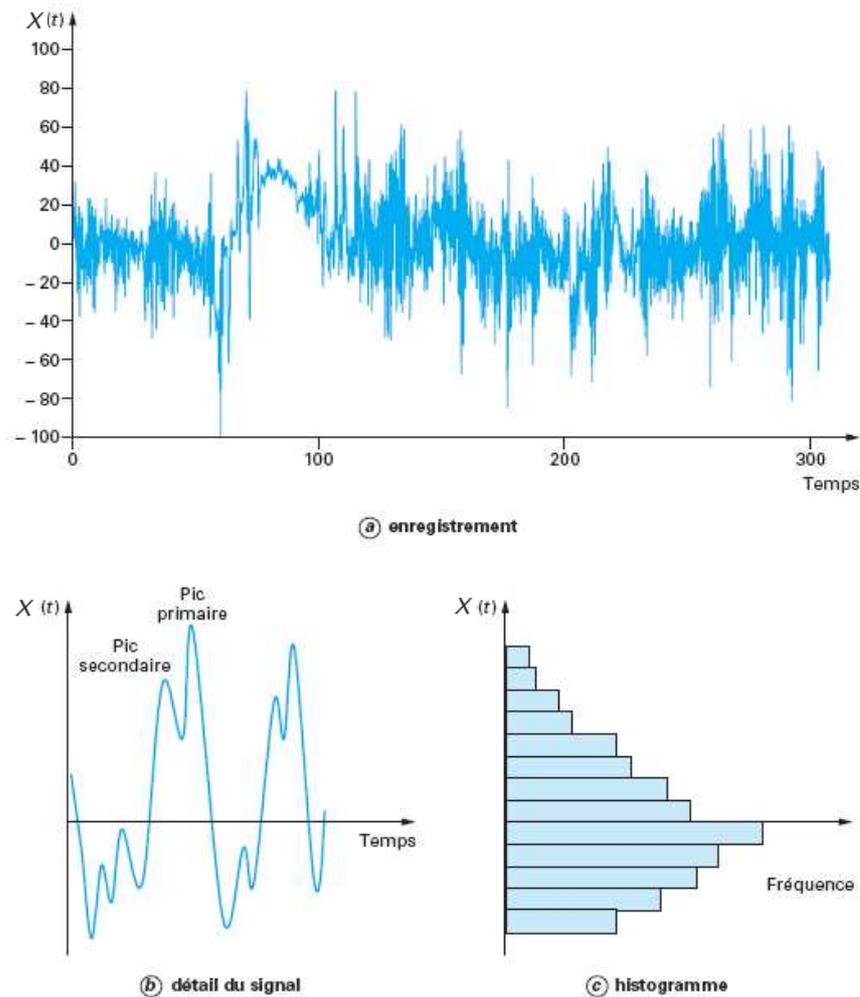


FIG. III.12 – Visualisation d'une sollicitation $X(t)$ en conditions réelles d'usage [et R. KOUTA, 2007]

Dans le cas d'une trajectoire de sollicitation très irrégulière, comme celle de la figure III.12, les pics secondaires posent problèmes. Et toute définition a priori de la façon de compter des étendues de variation de contrainte risque de conduire à des écarts de prévision par rapport à la réalité, si elle n'est pas confortée par une vérification expérimentale.

Les lois d'endommagement reposent sur des modèles plus ou moins simples, et la seule façon de dire si une méthode d'identification de cycles dommageables est meilleure qu'une autre est de corrélérer les résultats du modèle étudié avec ceux de l'expérience dans le cas où c'est possible de la réaliser (échelle de temps et/ou coût compatibles...). En réalité, les méthodes existantes donnent des résultats assez dispersés en comparaison des résultats publiés. Pour ces différentes raisons, l'extraction d'informations d'une sollicitation aléatoire doit s'effectuer avec précaution. Les différents types d'informations à extraire peuvent se présenter selon les trois formes suivantes :

- Analyse globale où l'ensemble des amplitudes de la sollicitation est étudié sans distinction sur la forme géométrique de la trajectoire (amplitude extrême, à pente positive ou négative, courbure vers le haut ou vers le bas). Cette analyse s'effectue grâce à l'histogramme de la sollicitation ou au suivi d'amplitudes particulières dans la sollicitation étudiée.
- Analyse locale par l'intermédiaire de l'étude des valeurs extrêmes selon la forme géométrique de la trajectoire. Dans ce cas, les valeurs extrêmes sont séparées en quatre familles statistiques : les pics positifs, les creux positifs, les pics négatifs et les creux négatifs. Les amplitudes qui ne présentent pas un changement de direction dans le signal digitalisé ne sont pas étudiées.
- Analyse des étendues de la sollicitation et/ou des cycles : quand la sollicitation aléatoire étudiée est une contrainte, il est judicieux de raisonner en étendue ou en cycle de contrainte. Cette définition est en adéquation avec ce qui est fait lors d'essais de vibration sous sollicitations sinusoïdales où la durée de vie est comptabilisée en nombres de cycles. Dans le cas d'une sollicitation sinusoïdale, une étendue ne concerne que la moitié d'un cycle. Dans le cas d'une sollicitation aléatoire $X(t)$, la définition d'un cycle est moins facile.

Dans le paragraphe suivant, nous présentons les conditions à vérifier et les méthodes pour construire la distribution du stress.

III.4.2.2 Caractérisation de la variable aléatoire

Les processus stochastiques (ou aléatoires) permettent de modéliser des systèmes dont le comportement n'est que partiellement prévisible. La théorie est fondée sur le calcul des probabilités et des statistiques.

Le calcul des probabilités classique concerne des épreuves où chaque résultat possible (ou réalisation) est un nombre, ce qui conduit à la notion de variable aléatoire. Un processus stochastique ou processus aléatoire (voir Calcul stochastique) ou fonction aléatoire (voir Probabilité) représente une évolution, généralement dans le temps, d'une variable

aléatoire.

La notion de processus continu correspond à un type de processus stochastique utilisé dans la description des signaux physiques, fonctions généralement régulières du temps.

Un processus stochastique est souvent appelé processus aléatoire (voir Calcul stochastique) ou, plus rarement, fonction aléatoire (voir Probabilité). Cette dernière expression aide à comprendre intuitivement de quoi il s'agit.

De même qu'une variable aléatoire est en fait un ensemble de nombres d'où l'on peut extraire un nombre particulier, une fonction aléatoire est un ensemble de fonctions ordinaires (tout au moins dans le cadre restreint considéré ici) d'où l'on peut extraire une fonction particulière.

Cette fonction particulière s'appelle réalisation du processus. Elle représente l'aspect temporel du phénomène que l'on peut caractériser par des moyennes temporelles, moyennes des valeurs de la fonction élevées aux différentes puissances.

Le processus possède aussi un aspect probabiliste si on considère l'ensemble des valeurs à un instant donné comme une variable aléatoire. Cette dernière est décrite par une densité de probabilité ou - ce qui revient au même sous certaines conditions de régularité - par la suite des moments statistiques que l'on nomme moyennes d'ensemble.

Les moyennes d'ensemble dépendent, en principe, de l'instant auquel elles se rattachent. Si en fait elles n'en dépendent pas, les propriétés statistiques restent constantes au cours du temps et l'on dit que le processus est stationnaire.

Si, de plus, les moyennes temporelles sont identiques aux moyennes d'ensemble, on dit que le processus est ergodique. Dans ces conditions, le phénomène est entièrement décrit par une seule réalisation. C'est donc cette notion d'ergodicité qui permet de répondre à la question posée initialement : comment interpréter en termes probabilistes un phénomène temporel même s'il est régi par des lois déterministes ?

L'interprétation de stress comme une variable aléatoire permet d'associer au stress certaines mesures et propriétés typiques pour la statistique.

Afin d'appliquer cette méthode, quelques hypothèses sur les valeurs d'entrée de stress doivent être faites pour que l'on puisse traiter en tant que processus par l'analyse des propriétés statistiques du processus entier.

Un processus est un ensemble $X(t)$, de fonctions ordinaires $x(t)$, chacune d'elles étant une réalisation du processus. On peut caractériser ce processus en lui associant à chaque instant t_0 , une densité de probabilité $p_X(x, t_0)$.

A une réalisation donnée, on peut associer les moyennes temporelles :

$$\overline{x(t)^n} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t)^n dt \quad (\text{III.10})$$

A la densité de probabilité $p_X(x, t_0)$, on peut associer les moments appelés moyennes d'ensemble :

$$E[X(t_0)^n] = \int_{-\infty}^{\infty} x^n p_X(y, t_0) dx \quad (\text{III.11})$$

Si ces moyennes d'ensemble, et par conséquent la densité de probabilité, ne dépendent pas de l'instant t_0 , on parle de processus stationnaire. Si, de plus, les moyennes temporelles leur sont égales, il s'agit d'un processus ergodique.

En fait, il ne s'agit là que des propriétés au premier ordre, les propriétés aux ordres supérieurs faisant intervenir la densité de probabilité jointe à des instants différents. Elles impliquent également des moyennes temporelles et des moyennes d'ensemble. Parmi ces dernières, la plus importante est l'autocovariance statistique $E[X(t_0)X(t_0 + \tau)]$. Si le processus est stationnaire et ergodique au second ordre, elle est identique à l'autocovariance temporelle, elle-même équivalente à la densité spectrale. Alors, un processus stationnaire, par définition, est un processus aléatoire où toutes ses propriétés statistiques ne changent pas au cours du temps [Cooper et McGillem, 1988].

Le passage de l'analyse statistique d'une sollicitation aléatoire à une modélisation probabiliste avec une analyse mathématique est toujours délicat. En effet, l'objet de la modélisation des chargements aléatoires est de déterminer une probabilité théorique d'apparition d'événements, à partir de l'observation d'une série de valeurs d'une variable aléatoire. Si la statistique descriptive cherche à porter un jugement sur la variable aléatoire par rapport à des valeurs présélectionnées, la statistique mathématique cherche à porter un jugement sur la probabilité d'apparition de différentes valeurs. L'histogramme (figure III.14 a) devient un « estimateur de densité » et le diagramme des cumuls est l'« estimateur de probabilité » (fonction de répartition).

En pratique, la loi de X est continue, la densité $f_X(x)$ est obtenue à partir de la série observée x_1, \dots, x_N . Appelons f_N la fonction qui associe, à x réel, le nombre $f_N(x)$ égal à la hauteur du rectangle qui est relatif à la classe contenant x (figure III.14 a). L'objectif de la modélisation consiste à affirmer que, pour tout x , les deux valeurs $f_N(x)$ et $f_x(x)$ sont « proches ». Des tests d'ajustement sont utilisés avec les précautions nécessaires à l'interprétation d'un test statistique (tests du χ^2 , de Kolmogorov, etc.)

L'intégration de propriétés statistique de stress dérive de la méthode plus générale de traitement de propriétés statistique d'un processus. La distribution obtenue au-dessus pour le stress est introduite dans la fiabilité de la fonction de produit soit en intégrant

la fonction de fiabilité sur le domaine de définition de stress, en tenant compte de la distribution précédemment calculée de stress ([Mettas, 2005] et [Gaudoin, 2002]), ou en calculant la fonction de densité multivariée [Al-Mutairi, 1996].

Dans le cas d'une sollicitation aléatoire, un nouveau facteur de forme est défini ; il est appelé facteur d'irrégularité I . Ce facteur est égal au rapport du nombre de passages à zéro N_0 du signal centré et du nombre d'extrémums N_e contenus dans la sollicitation étudiée : $I = N_0/N_e$, I compris entre 0 et 1 (figure III.13).

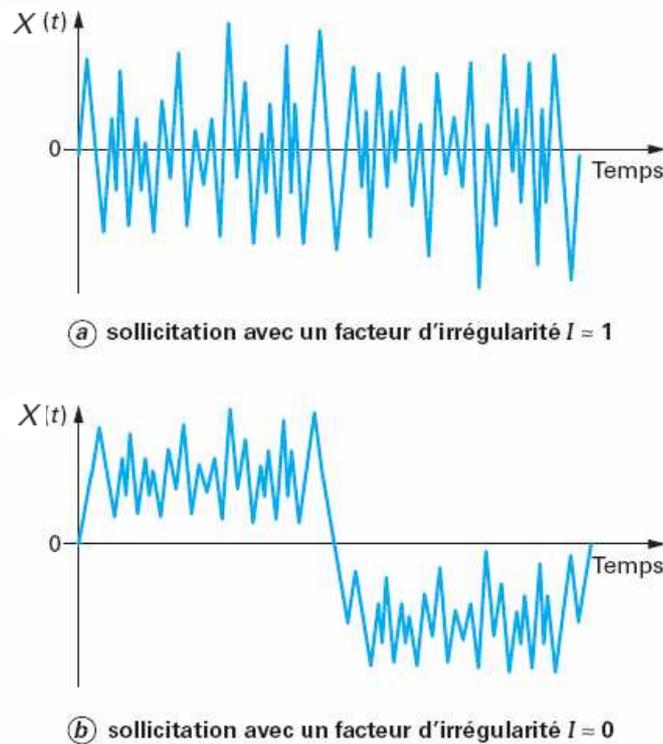


FIG. III.13 – Caractérisation des sollicitations en fonction du facteur d'irrégularité I [et R. KOUTA, 2007]

Plus I s'approche de 1, plus la régularité de la sollicitation est importante. Ainsi, presque chaque extrémum est systématiquement suivi par un passage à zéro (figure III.13 a). Dans cette situation, la mesure étudiée est qualifiée de « sollicitation à bande étroite ». Lorsque I s'approche de 0, la régularité de la sollicitation devient de plus en plus faible. Ainsi, entre deux passages par le niveau zéro, le signal passe par de nombreux extrémums relatifs de même signe (figure III.13 b). Dans cette situation, la mesure étudiée est qualifiée de « sollicitation à bande large ». Le facteur d'irrégularité donne une indication importante sur l'évolution temporelle d'une sollicitation, ce que les autres facteurs de forme ignorent totalement. Dans les paragraphes suivants, nous présentons deux méthodes de construction des histogrammes de stress à partir desquels il est possible de définir les fonction de

densité.

III.4.2.3 Histogramme ou temps de maintien dans une classe d'amplitudes

Cette méthode considère le signal digital enregistré comme un échantillon statistique en ignorant l'aspect temporel. L'échantillon est regroupé en classes d'amplitudes (figure III.14 a traits pointillés horizontaux). Dans ce cas, aucune distinction n'est faite entre les valeurs extrêmes et les autres. L'avantage de cette méthode réside dans la possibilité immédiate de faire une modélisation statistique et de proposer un modèle de densité de probabilité (partie droite de la figure III.14 a). Étant donné qu'entre deux points successifs il y a un pas de temps prédéfini par la méthode de mesure, le nombre de points comptabilisés dans une classe une fois multiplié par le pas de temps donne le temps global de maintien de la sollicitation étudiée dans cette classe d'amplitudes. Cette méthode de comptage doit être uniquement réservée aux sollicitations homogènes (ou dont la source est considérée homogène), c'est-à-dire des cas sans changement important de nature de chargement. En effet, cette méthode est très dépendante de la vitesse (dérivée première sur la courbe considérée comme une trajectoire avec une signature dynamique spécifique) et de l'accélération (dérivée seconde) de la sollicitation étudiée. Dans le cas où la sollicitation présente plusieurs types d'information (liés aux freinages, virages, chargements variés...), le signal perd son homogénéité et le comptage par classe sera altéré par ces différentes utilisations dont la signature dynamique n'est pas la même. La figure III.14 a présenté une sollicitation digitalisée où 28 points et 9 classes d'amplitudes sont définis [Kouta et Play, 2006]. Le comptage par classe (à partir de la classe du bas) conduit successivement à 2, 7, 3, 5, 1, 2, 5, 1 et 2 amplitudes par classe.

III.4.2.4 Comptage du nombre de dépassements d'un niveau donné

Cette méthode, comme la précédente, demande que soient définies au préalable des classes d'amplitudes (figure III.14 b). Le comptage, pour un niveau donné, est déclenché chaque fois que le signal franchit un niveau avec une pente positive (d'où le nom de dépassement de niveau). Un comptage du nombre de dépassements d'un niveau donné n'est pertinent que si une attitude de sélection des petites oscillations est définie. Ces petites oscillations peuvent apporter des effectifs (nombre) de charge sans intérêt du point de vue du calcul de l'endommagement et du calcul de la durée de vie. Ainsi, pour le comptage du nombre de dépassements de niveau, un incrément de signal noté Δ est défini. Il correspond souvent, dans le cas d'un composant mécanique, à un intervalle de contrainte au-dessous d'une limite de fatigue définie sur une courbe de Wöhler. Cet incrément Δ est

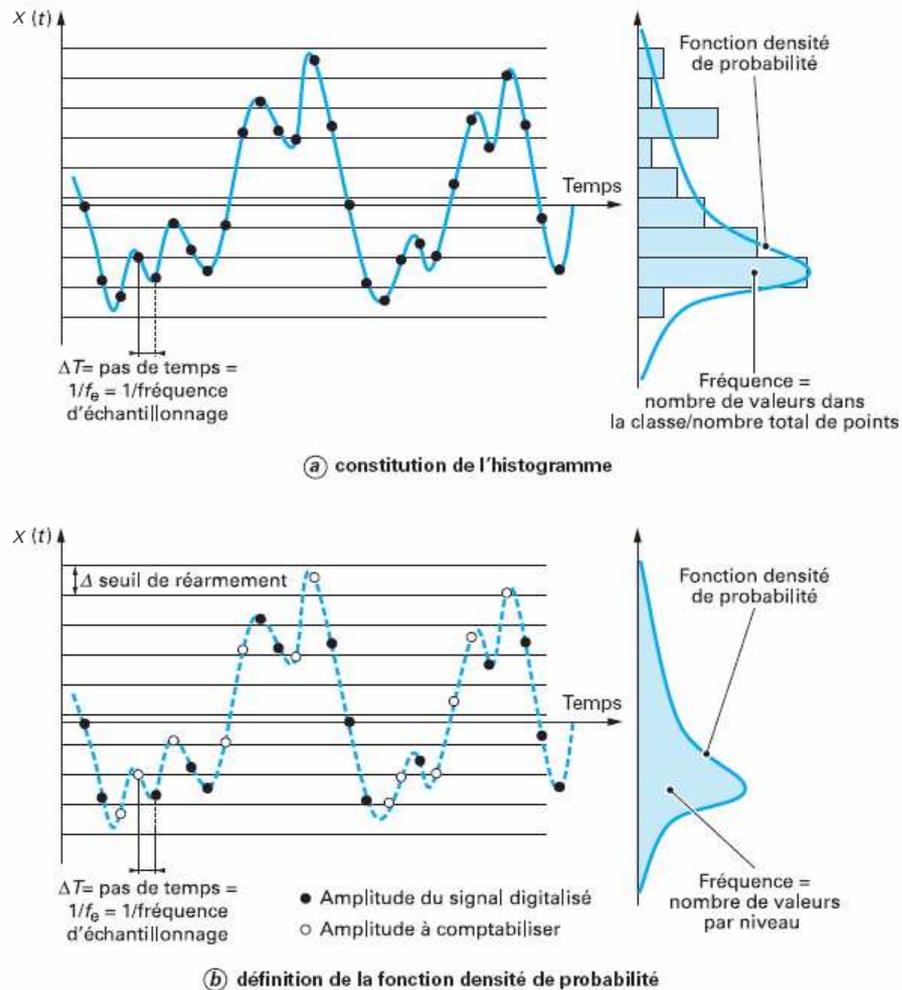


FIG. III.14 – Statistique d'une sollicitation [et R. KOUTA, 2007]

considéré comme un seuil de réarmement dans le processus de comptage. Historiquement, plusieurs méthodes de comptage ont été proposées. La méthode la plus intéressante compte un niveau seulement si la sollicitation a déjà traversé au moins une fois le seuil Δ , cela sans distinction de nature de pente. Il faut rappeler aussi que le comptage s'effectue avec un signal digitalisé, ainsi une règle de proximité doit être mise en place afin de compter une amplitude très proche des niveaux déterminés. Cette méthode de comptage permet - comme la précédente - de construire un modèle de densité de probabilité. L'application de cette méthode se focalise sur les niveaux définis par les bornes des classes. Le comptage par niveau (avec la sollicitation de la figure III.14 b, à partir du niveau du bas) conduit successivement à 0, 0, 2, 1, 0, 1, 1 et 2 dépassements par niveau.

III.4.3 Estimation de la fiabilité

Une fois ayant la modélisation probabiliste, l'influence des propriétés statistiques de chaque stress X_l sur les caractéristiques de produit peut être obtenue de deux façons :

1. à partir de la règle des probabilités conditionnelles moyennes applique sur la fonction de densité de probabilité [Pitman, 1993] :

$$f(t) = \int_{DX_l} f(t|X_l) f_{X_l}(X_l) dX_l \quad (\text{III.12})$$

2. par intégration directe dans la fonction de fiabilité du produit :

$$R(t) = \int_{DX_l} R(t|X_l) f_{X_l}(X_l) dX_l \quad (\text{III.13})$$

et, à partir de I.88, I.25 et III.13, nous obtenons l'estimation de la fonction de fiabilité $R(t)$ d'un produit subit au stress multidimensionnelle X par :

$$R(t) = - \int_{DX_l} R_0 \left(\left(\frac{e^{\beta^T Z t}}{\theta} \right)^v \right) \frac{b(X_l/a)^{-1+b}}{a} S'_0 \left(\left(\frac{X_l}{a} \right)^b \right) dX_l \quad (\text{III.14})$$

Les analyses statistiques ont l'avantage de conduire à des modèles probabilistes [Demoulin, 1990] qui offrent des possibilités de modélisation de la dispersion naturelle des sollicitations étudiées et de leurs conséquences (fissuration, fatigue, endommagement, durée de vie, etc.). L'inconvénient de ces analyses statistiques réside cependant dans le fait qu'elles ignorent l'historique des événements.

Pour permettre la prise en compte de l'historique, nous proposons d'utiliser les séries temporelles.

III.4.4 Séries temporelles

Une série temporelle est une série de nombres (ici des valeurs du stress) indicés par le temps, chacun d'entre eux étant une réalisation d'une variable aléatoire. Ceci peut s'exprimer aussi en disant que la série temporelle est une réalisation d'une famille de variables aléatoires indicées par le temps. Cette famille de variables aléatoires s'appelle un processus aléatoire.

Les séries temporelles et les mécanismes générateurs que sont les processus, ont une structure. En d'autres termes, les variables aléatoires composant le processus ne sont pas forcément indépendantes les unes des autres mais établissent certaines relations qui donnent à l'ensemble une sorte de structure.

Il existe deux catégories de modèles pour rendre compte d'une série temporelle [Delignières, 2000]. La première considère que les données sont une fonction du temps ($X = f(t)$). Cette catégorie de modèle peut être ajustée par la méthode des moindres carrés, ou d'autres méthodes itératives. L'analyse des modèles par transformée de Fourier est une version sophistiquée de ce type de modèle. Une seconde catégorie de modèles cherche à déterminer chaque valeur de la série en fonction des valeurs qui la précède ($X_t = f(X_{t-1}, X_{t-2}, \dots)$). C'est le cas des modèles ARIMA (Auto - Regressive - Integrated - Moving Average).

A noter que le choix de l'un ou l'autre type de modèle est surtout théorique : est-il raisonnable de penser que dans un phénomène donné, les points sont fondamentalement fonction des points précédents et de leurs erreurs, plutôt qu'un signal, périodique ou non, entaché de bruit. On peut noter cependant que souvent, on a recours à l'analyse de variance pour traiter les séries temporelles. Or une des hypothèses majeures de l'ANOVA est que les résidus des différentes mesures ne sont pas auto corrélés. Ce n'est évidemment pas le cas si la performance à l'essai t est liée à la performance réalisée à l'essai $t-1$. Les processus autorégressifs supposent que chaque point peut être prédit par la somme pondérée d'un ensemble de points précédents, plus un terme aléatoire d'erreur. Le processus d'intégration suppose que chaque point présente une différence constante avec le point précédent. Les processus de moyenne mobile supposent que chaque point est fonction des erreurs associées aux points précédents, plus sa propre erreur. Un modèle ARIMA, noté ARIMA (p, d, q), dépend des variables suivantes :

- p est le nombre de termes auto-régressifs
- d est le nombre de différences
- q est le nombre de moyennes mobiles.

L'examen d'une série temporelle (une réalisation d'un processus) permet, en général, de lui reconnaître trois types de composantes : une tendance, une composante saisonnière et une variation aléatoire. D'autres caractéristiques évolutives, comme les chocs, peuvent être également observées mais sont moins intimement liées à la structure de la série. Il est alors utile de séparer ces composantes, et ceci pour deux raisons. La première est de répondre à des questions de bon sens comme celle de la croissance ou la décroissance générale du phénomène observé. L'extraction de la tendance et l'analyse de celle-ci répondront à cette question. Il est intéressant aussi de mettre en évidence la présence éventuelle d'une variation périodique grâce à l'analyse de la composante saisonnière. La seconde de ces raisons est de débarrasser le phénomène de sa tendance et de ses variations périodiques pour observer plus aisément le phénomène aléatoire.

Ce modèle utilise souvent des valeurs retardées (d'où le terme Auto Regressive) et de

chocs aléatoires qui sont en général de moyenne nulle, de variance constante et non auto corrélés (bruit blanc); quand la variable qui représente ces chocs est retardée, on parle de moyenne mobile.

L'estimation des modèles ARIMA suppose que l'on travaille sur une série stationnaire. Ceci signifie que la moyenne de la série est constante dans le temps, ainsi que la variance. La meilleure méthode pour éliminer toute tendance est de différencier, c'est-à-dire de remplacer la série originale par la série des différences adjacentes. Une série temporelle qui a besoin d'être différenciée pour atteindre la stationnarité est considérée comme une version intégrée d'une série stationnaire (d'où le terme Integrated).

La correction d'une non stationnarité en termes de variance peut être réalisée par des transformation de type logarithmique (si la variance croît avec le temps) ou à l'inverse exponentielle. Ces transformations doivent être réalisées avant la différenciation.

Le modèle ARMA est un cas particulier. En effet, le modèle ARMA ne permet de traiter que les séries dites stationnaires (des moments du premier ordre qui sont invariants au cours du temps). Les modèles ARIMA permettent de traiter les séries non stationnaires après avoir déterminé le niveau d'intégration (le nombre de fois qu'il faut différencier la série avant de la rendre stationnaire).

Bien que possédant d'excellentes qualités prévisionnelles, le modèle ARIMA ou ARMA souffre d'une lacune majeure : il est incapable de traiter simultanément plus d'une variable (série). Pour contourner ce problème, il faut pouvoir généraliser le modèle ARIMA dans le cas à plusieurs variables. C'est ce qu'a fait en partie Sims en proposant en 1980 le modèle Vector Auto Regressive (VAR) qui permet de traiter concomitamment plusieurs variables. Mais, contrairement au modèle structurel à plusieurs variables, dans les modèles VAR, toutes les variables sont endogènes. Cette manière de modéliser en faisant abstraction d'une théorie économique a donné naissance à ce que l'on a appelé l'Econométrie sans théorie.

La notation $AR(p)$ désigne le modèle autoregressive d'ordre p :

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t. \quad (\text{III.15})$$

avec φ_1, \dots et φ_p les paramètres du modèle, c une constante et ε_t le terme d'erreur. La constante n'est souvent pas prise en compte pour des raisons de simplicité. Le terme d'erreur ε_t suit une loi normale et représente la partie aléatoire due à la variation du stress.

L'influence des propriétés statistiques de chaque stress X_t sur les caractéristiques de produit peut être obtenue de deux façons :

1. à partir de la règle des probabilités conditionnelles moyennes appliquée sur la fonction de densité de probabilité [Pitman, 1993] :

$$f(t) = \int_{D\varepsilon_t} f(t|X_t) f_{\varepsilon_t}(\varepsilon_t) d\varepsilon_t \quad (\text{III.16})$$

2. par intégration directe dans la fonction de fiabilité du produit :

$$R(t) = \int_{D\varepsilon_t} R(t|X_t) f_{\varepsilon_t}(\varepsilon_t) d\varepsilon_t \quad (\text{III.17})$$

Dans le cas où il n'est pas possible d'obtenir une distribution ou une série temporelle, nous proposons de discrétiser le signal.

III.4.5 Discrétisation du profil de stress

III.4.5.1 Principe de la discrétisation

Cette méthode consiste à discrétiser le profil de stress afin d'obtenir un profil échelonné comme représenté sur la figure .

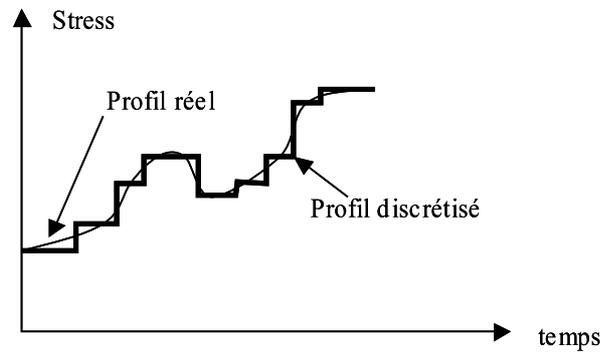


FIG. III.15 – Discrétisation du profil réel

A partir de ce profil, il est possible d'estimer la fiabilité en utilisant le modèle d'endommagement cumulé.

III.4.5.2 Modèle d'endommagement cumulé (MEC)

Le modèle d'endommagement cumulé est basé sur l'expression générale des modèles SVA :

$$R_{X(\bullet)}(t) = R_0 \left(\left(\frac{\int_0^t e^{\beta T Z(\tau)} d\tau}{\theta} \right)^v \right), (\theta, v > 0) \quad (\text{III.18})$$

Pour le cas où $R_0(u) = e^{-u}$ on obtient la classe de fonctions de Weibull, alors III.18 devient :

$$R_{X(\bullet)}(t) = e^{-\left(\left(\frac{\int_0^t e^{\beta T Z(\tau)} d\tau}{\theta}\right)^v\right)} \quad (\text{III.19})$$

et, par identification, les paramètres de forme et d'échelle s'expriment :

$$\begin{cases} v = \beta \\ \eta(t|X(t)) = \frac{e^{\beta T Z(t)}}{\theta} \end{cases} \quad (\text{III.20})$$

Un des modèles très utilisés, dans le cadre des essais accélérés, est le modèle d'endommagement cumulé car il peut intégrer dans ses équations la variation de stress au cours de temps. Par une étude, qui combine l'approche statistique et celle basée sur la notion de ressources utilisées, Finkelstein [Finkelstein, 2008] montre que le modèle MEC est justifié seulement pour le cas particulier où $r[X(\tau)]t$ (équation I.76) a une forme linéaire.

Pour présenter le modèle MEC, nous considérons un profil de stress échelonné, représenté par la figure III.16, avec le stress $S_1 < S_2 < S_3$ et avec $2t_1 = t_2$ et $3t_1 = t_3$. On observe un changement de distribution en fonction du temps ($R(t)$ représentée en pointillé sur la figure III.16). Pour le modèle d'endommagement cumulé, le paramètre influencé par la variation de stress est le paramètre d'échelle, $\eta(t)$ de la distribution Weibull.

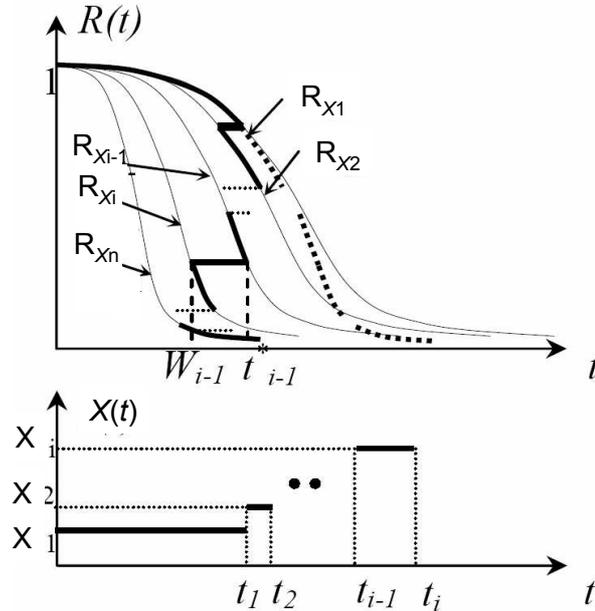


FIG. III.16 – Step-stress profile and the corresponding life distributions

Ainsi, le modèle d'endommagement cumulé consiste à intégrer la variation de stress

dans l'estimation du paramètre d'échelle $\eta(t)$ de la fonction Weibull :

$$R_i(t) = e^{-\left(\frac{t}{\eta_i}\right)^\beta} \quad (\text{III.21})$$

Pour $t_{i-1} \leq t \leq t_i$, la fonction de fiabilité devient :

$$R(t) = e^{-\left(\frac{t-t_{i-1}+W_{i-1}}{\eta_i}\right)} \quad (\text{III.22})$$

Le transfert d'un niveau de stress aux autres est supposé instantané. Alors, en faisant l'égalité entre les percentiles pour deux niveaux consécutifs de stress, on obtient les équations suivantes :

$$\left\{ \begin{array}{l} \frac{\eta_{i-1}}{\eta_i} W_{i-1} = (t_{i-1} - t_{i-2}) W_{i-2} = \Delta t_{i-1} - W_{i-2} \\ \frac{\eta_{i-2}}{\eta_{i-1}} W_{i-2} = (t_{i-2} - t_{i-3}) W_{i-3} = \Delta t_{i-2} - W_{i-3} \\ \vdots \\ \frac{\eta_2}{\eta_3} W_2 = (t_2 - t_{i-1}) W_1 = \Delta t_2 - W_1 \\ \frac{\eta_1}{\eta_2} W_1 = (t_1 - t_0) = \Delta t_1 \end{array} \right. \quad (\text{III.23})$$

Après la multiplication de chaque équation par $\prod_{k=2}^j \frac{\eta_{k-1}}{\eta_k}$ avec $j = 2..(i-1)$, on obtient la relation :

$$\frac{W_{i-1}}{\eta_i} = \sum_{k=1}^{i-1} \frac{\Delta t_k}{\eta_k} \quad (\text{III.24})$$

En remplaçant III.24 dans III.22, la forme générale de $R(t)$ devient :

$$R_{[i-1,i]}(t) = e^{-\left\{ \frac{t-t_{i-1}}{\eta_i} + \sum_{k=1}^{i-1} \frac{\Delta t_k}{\eta_k} \right\}^\beta} \quad (\text{III.25})$$

En prenant $\Delta t_i \rightarrow 0$, on obtient :

$$\left\{ \begin{array}{l} \frac{t-t_{i-1}}{\eta_i} \rightarrow 0 \\ \sum_{k=1}^{i-1} \frac{\Delta t_k}{\eta_k} \rightarrow \int_0^t \frac{1}{\eta(t)} dt \end{array} \right. \quad (\text{III.26})$$

d'où :

$$R(t|X) = e^{-\left(\frac{t}{\int_0^t \frac{d\tau}{\eta(\tau|X(\tau))}}\right)^\beta} \quad (\text{III.27})$$

L'équation III.27 décrit le modèle d'endommagement cumulé pour une loi de Weibull en considérant le paramètre de forme β constant avec la variation de stress.

III.5 Exemple d'application

Nous reprenons l'exemple du calculateur de soute embarquée dans un moyen/long courrier. Pour une question de clarté, nous proposons de reprendre le profil de mission présenté dans la figure III.17 auquel nous ajoutons un profil aléatoire en température (profil défini par la figure III.18) pour les phases de vie *vol-operating* et *sol-operating*¹.

Profil d'emploi de type aéronautique civil moyen / long courrier, calculateur en soute														
Contrainte		Thermique et Humidité			Cyclage thermique				Mécanique	Chimique			Induit	
Intitulé de la phase	Temps calendaire (heures)	On/Off	Température ambiante (°C)	Taux d'humidité (%)	ΔT (°C)	Nombre de cycles (/an)	Durée du cycle (heures)	Température maximale au cours du cyclage (°C)	Vibrations aléatoires (Gms)	Pollution saline	Pollution industrielle	zone d'application	Niveau de protection	II application
Sol-operating - 1	797	On	47	30	33	330	2	47	-	Faible	Urbaine	Non habitable	Non hermétique	4.8
Sol-operating - 2	1193	On	55	30	15	647	1,5	55	-	Faible	Urbaine	Non habitable	Non hermétique	2.0
Sol-roulage	84	On	47	5	-	-	-	-	5	Faible	Négligeable	Non habitable	Non hermétique	1.2
Vol-operating	4083	On	40	5	-	-	-	-	0,6	Faible	Négligeable	Non habitable	Non hermétique	1.1
Sol-dormant	2603	Off	14	70	10	108	24	19	-	Faible	Urbaine	Non habitable	Non hermétique	3.3

FIG. III.17 – Profil d'emploi de type aéronautique civil moyen / long courrier [FIDES, 2004]

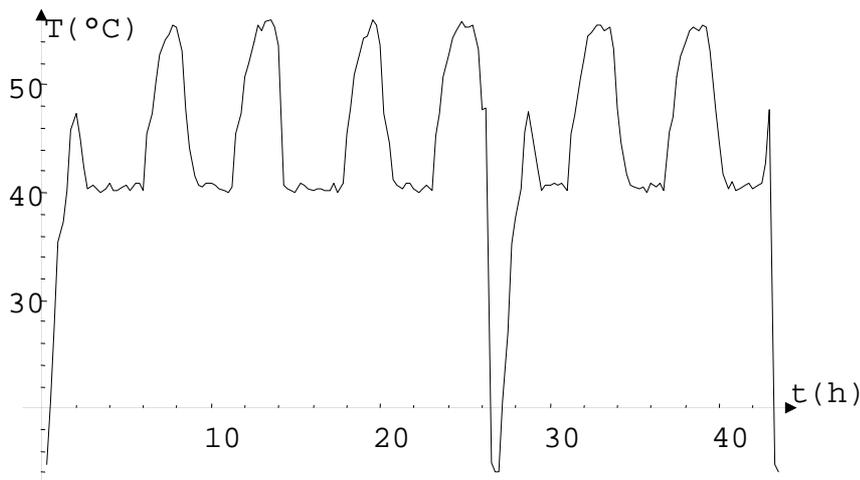


FIG. III.18 – Profil de mission réel

Nous avons ainsi les taux de passage par les phases p_j qui sont obtenus par la proportion du temps calendaire associé à chaque phase pour un an :

- P_1 : sol-operating 1, $p_1 = 13\%$
- P_2 : sol-operating 2, $p_2 = 20\%$
- P_3 : sol-roulage, $p_3 = 2\%$
- P_4 : vol-operating, $p_4 = 65\%$

Ainsi, nous allons pouvoir étudier les trois approches d'estimation de la fonction de fiabilité :

stress	loi acceleration	paramètres du modèle
H, T	Peck : $\eta(H, T) = CH^{-n}e^{\frac{B}{T}}$	$n = 1.24, B = 1638.5, C = 0.411, v = 2.67$
ΔT	Coffin-Manson : $\eta(\Delta T) = \left(\frac{a}{\Delta T}\right)^n$	$a = 3.87E + 4, n = 1.24, v = 2.33$
$Grms$	Inverse Power : $\eta(V) = \left(\frac{a}{V}\right)^n$	$a = 0.08, n = 2.19, v = 1.46$

TAB. III.2 – Lois de fiabilité associées et valeurs des paramètres

1. Distribution de la température indépendante du temps
2. Profil de température approximé par une série temporelle
3. Profil de température discrétisé

Pour estimer les différentes fonctions de fiabilité pour les différentes phases de vie (sauf celle *sol-dormant*), nous utilisons les modèles de vie accélérée correspondant aux trois stress suivants :

- H, T : température, humidité
- ΔT : température cyclique
- $Grms$: vibrations

Les modèles de vie accélérée associés à chaque stress sont des lois de Weibull de paramètres v (forme) et η (échelle) définis par des lois d'accélération présentées dans le tableau III.2.

Distribution de la température indépendante du temps

Nous considérons que la température, lors des phases *vol-operating* et *sol-operating1*, est décrite par deux distributions normales :

- *vol-operating* : $N(\mu, \sigma) = N(40^\circ\text{C}, 1^\circ\text{C})$,
- *sol-operating1* : $N(\mu, \sigma) = N(58^\circ\text{C}, 6^\circ\text{C})$,

Alors, la loi de fiabilité associée à la température aléatoire et à la phase de vie P_j s'écrit :

$$R_{P_j}^T(t) = \int_{\Omega_T} e^{-\left(\frac{t}{CH_j^{-n}e^{\frac{B}{T}}}\right)^v} \times \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{T-\mu}{\sigma}\right)^2} dT$$

Les valeurs de C, n, B et v sont définies dans le tableau III.2 et les valeurs de H sont prises dans la figure III.17 selon la phase de vie considérée. Ensuite, nous estimons la fonction de fiabilité globale (figure III.19) :

$$R_{\text{sys}}(t) = \left\{ \prod_{j=1}^4 [R_{P_j}^{H,T}(p_j t)] \right\} \left\{ \prod_{j=1}^5 [R_{P_j}^{\Delta T}(p_j t)] \right\} \left\{ \prod_{j=1}^4 [R_{P_j}^{Grms}(p_j t)] \right\} \quad (\text{III.28})$$

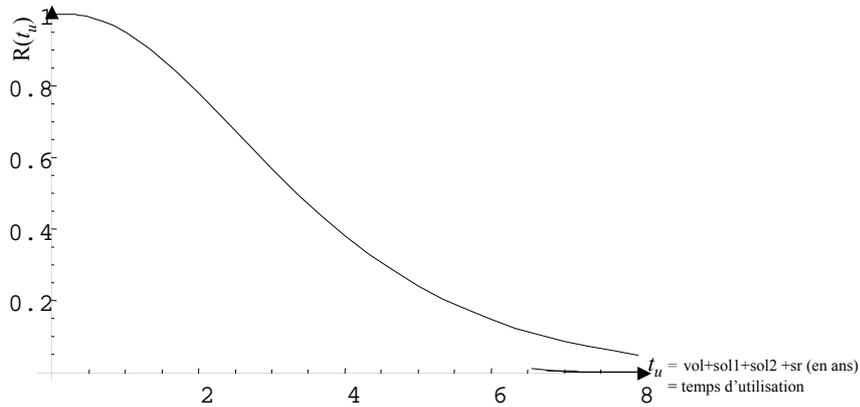


FIG. III.19 – Estimation de la fiabilité

Profil de température approximé par une série temporelle

Nous considérons la température définie par la figure III.18 et nous avons déterminé les paramètres du modèle ARIMA(2, 1, 0) :

- const. : 0,22682
- p(1) : 0,23066
- p(2) : 0,21855

L'analyse des résidus a montré qu'ils étaient négligeables. En appliquant les équations III.17 (avec la fonction de densité $f_{\epsilon_t}(\epsilon_t)$ considérée négligeable) et III.28, la courbe de fiabilité globale prend la forme présentée dans la figure III.20.

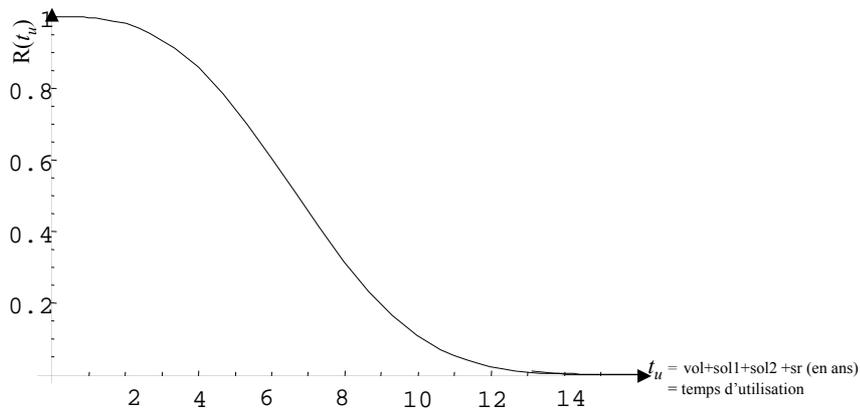


FIG. III.20 – Estimation de la fiabilité

Profil de température discrétisé

Nous considérons la température définie par la figure III.18 et nous avons discrétisé le profil par un pas de 10 minutes. En appliquant la méthode MEC, nous construisons la fonction de fiabilité $R^{T,H}$ en répétant plusieurs fois le profil de température afin de couvrir tout le domaine d'exploitation (figure III.21).

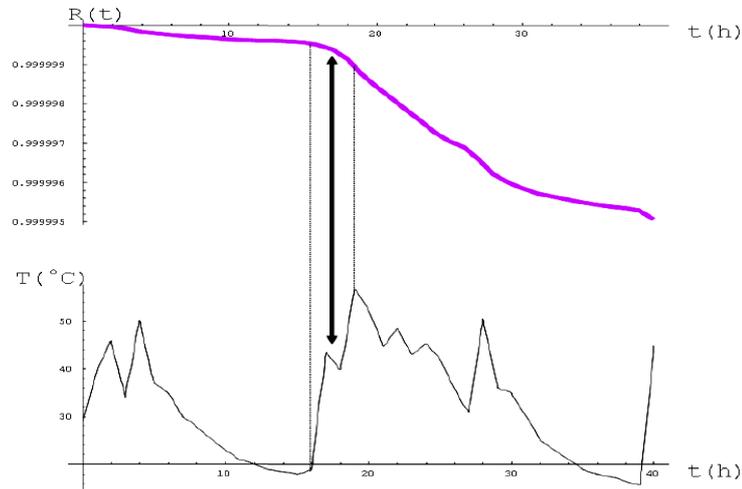


FIG. III.21 – La variation de $R(t)$ avec le stress (temperature)

Ensuite, il est possible d'estimer la fonction de fiabilité globale par l'équation III.28 (figure III.22).

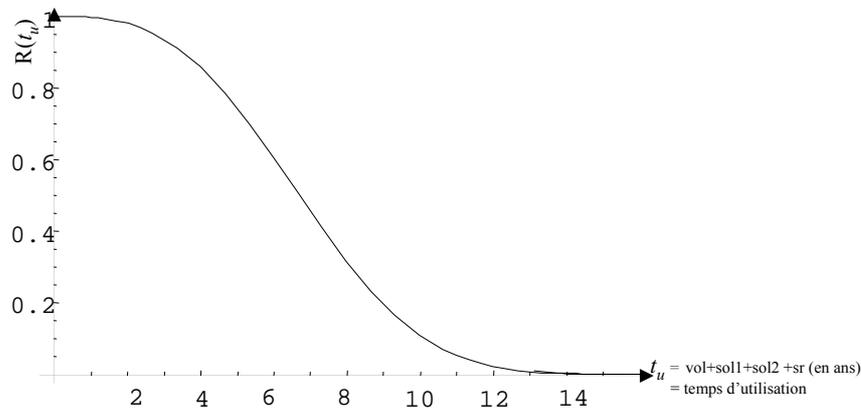


FIG. III.22 – La fiabilité en appliquant le modèle d'endommagement cumulative

Finalement, la figure III.23 représente les fonctions de fiabilité globale obtenues par les trois approches :

- courbe *a* : série temporelle,
- courbe *b* : discrétisation du profil,
- courbe *c* : distribution du stress indépendante du temps.

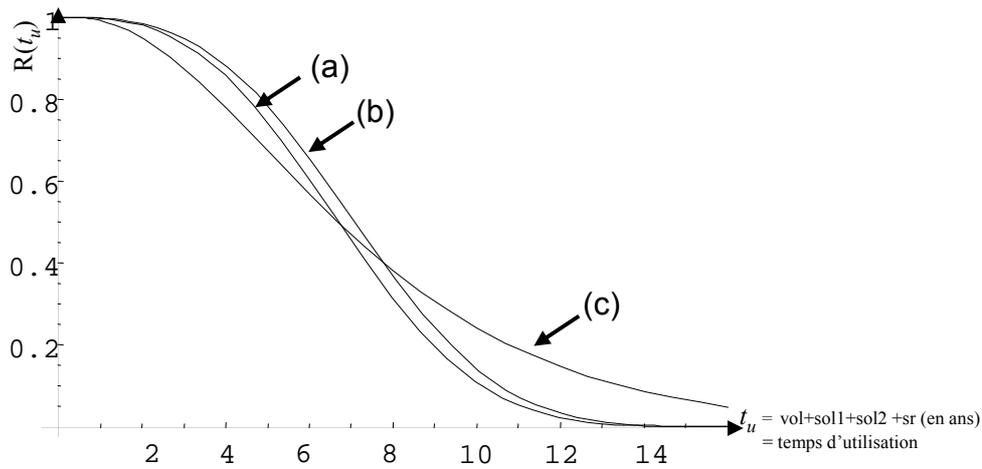


FIG. III.23 – Résultats

On observe que les courbes (a) et (b) sont proches. Cela s'explique par le fait que les deux méthodes utilisent tout le profil de température avec quasiment aucune perte d'information (pour la série temporelle). Au contraire, avec la méthode où la température est définie par des distributions à deux niveaux, nous perdons de l'information qui se traduit par une courbe de fiabilité décalée.

III.6 Conclusion

Ce chapitre a présenté une méthodologie permettant de considérer les variabilités d'usage, de conditions d'utilisation, d'utilisateurs, ... dans l'estimation de la fiabilité. L'objectif de ce travail était d'obtenir des estimations en phase de développement proches de celles observées en condition opérationnelle qui intègrent naturellement ces variabilités.

Pour ce faire, nous avons rappelé comment était défini un profil de mission et les éléments le constituant : phases de vie, mission et environnement. La variabilité peut intervenir sur ces trois éléments impactant directement la fiabilité du produit. Aussi, nous avons proposé d'introduire cette variabilité dans l'estimation de la fiabilité en considérant deux cas :

- Variabilité d'usage qui intègre différentes phases de vie et divers utilisateurs (l'environnement est déterministe),
- Variabilité d'environnement qui considère des stress aléatoires.

Pour tenir compte de la variabilité d'usage, nous avons présenté une méthode directe et une autre basée sur l'approche Résistance-Contrainte. L'effet aléatoire des stress sur le produit peut se modéliser de trois manières différentes :

- Par des distributions des stress indépendantes du temps,
- Par des séries temporelles,
- Par la discrétisation des profils.

Pour chaque cas, nous avons proposé une méthode d'estimation de la fiabilité. Nous avons observé que les méthodes étaient comparables. Pour aller plus loin dans l'analyse, il faudrait pouvoir comparer les résultats avec ceux obtenus à partir de données opérationnelles. Une limite de cette méthodologie concerne la difficulté à récupérer des profils de mission complets et détaillés. Une extension de ce travail concerne l'application des modèles proposés à la problématique du " Health-Monitoring ". En effet, depuis quelques années, des capteurs de mesure de paramètres environnementaux sont intégrés sur les cartes afin d'estimer en temps réel la durée de vie consommée, la durée de vie résiduelle et le risque de défaillance et ainsi optimiser les politiques de maintenance.

IV

Conclusions et perspectives

Compte tenu des marchés de plus en plus concurrentiels, l'estimation de la fiabilité des systèmes est devenue un enjeu majeur des entreprises et ce dès les premières phases de développement. En effet, plus tôt les caractéristiques d'un produit ou d'un système sont appréhendées dans son cycle de vie et moins les risques (financiers ou liés à la sécurité des installations), dus à la non réalisation des performances attendues sont élevés. Dans un contexte exigeant de systèmes de plus en plus fiables et sûrs, et de durées de garanties croissantes, il est impératif de vérifier le plus rapidement possible que les performances des systèmes soient conformes au cahier des charges. La démarche idéale pour identifier la fiabilité d'un produit ou d'un système avant même sa fabrication en série, est de procéder à des séries d'essais sur des prototypes, quand ils existent. Cette démarche nécessite un investissement trop important en temps et en nombre de prototypes, car ces derniers étant de plus en plus fiables, l'observation de défaillances est de moins en moins probable. Ainsi, pour réduire ces durées, on peut réaliser des essais accélérés. Pour cela, le produit est soumis à des sollicitations d'utilisation ou d'environnement amplifiées, afin d'accélérer les mécanismes d'endommagement (les mécanismes de défaillances provoqués doivent être représentatifs des conditions normales d'emploi) et de réduire la durée nécessaire pour l'estimation de certaines caractéristiques comportementales du produit dans les conditions normales d'emploi. Toutefois, la réalisation d'essai nécessite de disposer de suffisamment de produits afin d'obtenir des estimations satisfaisantes. Malheureusement, dans la plupart des cas, nous ne disposons pas assez de produits. C'est dans ce contexte que s'est inscrite cette thèse afin de proposer une méthodologie permettant d'estimer les paramètres de modèles de vie accélérée en intégrant toute la connaissance disponible sur la fiabilité du produit étudié. Pour ce faire, nous nous sommes appuyés sur la démarche bayésienne.

Dans une première étape, nous avons introduit l'analyse des données non accélérées en faisant une comparaison des méthodes paramétrique et non paramétrique. Nous avons

montré les limites des méthodes lorsque l'on considère des petits échantillons et de forte censure. Nous avons étudié l'estimation par la méthode du maximum de vraisemblance et par celle bayésienne non informative.

Pour réduire les délais de cycle de vie, et en particulier la qualification des produits, nous avons étudié l'analyse des résultats d'essais accélérés pour diminuer le temps de test. Pour ce faire, nous avons utilisé le modèle de vie standard de vie accélérée en considérant les cas paramétrique et semi-paramétrique.

Ensuite, pour obtenir des estimations correctes avec des tailles d'échantillon faible, nous avons introduit l'estimation Bayésienne des modèles SVA. Nous avons étudié deux techniques d'estimation : maximum d'a posteriori et la méthode classique. L'efficacité de la démarche Bayésienne dépend principalement dans le choix des lois a priori. Pour cela, nous avons proposé une méthodologie permettant de définir des pseudo lois conjuguées par une technique de simulation puisqu'il n'est pas possible de les déterminer analytiquement. Pour finir, nous avons traité le problème de la compatibilité entre la vraisemblance et l'a priori. Dans le cas où nous observons une incompatibilité partielle ou totale, nous avons proposé une méthodologie de pondération.

Le troisième chapitre aborde la problématique de l'estimation de la fiabilité lorsque les conditions d'utilisation sont variables. Nous avons présenté une méthodologie permettant de considérer les variabilités d'usage, de conditions d'utilisation, d'utilisateurs, ... dans l'estimation de la fiabilité. L'objectif de ce travail était d'obtenir des estimations en phase de développement proches de celles observées en opérationnelle qui intègrent naturellement ces variabilités.

En terme de perspectives, nous envisageons plusieurs voies de développement :

- réaliser une étude détaillée de la méthode proposée de pondération. En effet, quel est le comportement de l'estimation bayésienne pondérée lorsque l'incohérence concerne uniquement un seul paramètre ?
- étendre la démarche Bayésienne à d'autres modèles (Generalized Proportional Hazard Model, Changing Shape and Scale Model, ...) et plans (Essais sous stress échelonnés, ...),
- optimiser les plans d'essai afin de réduire le coût d'essai (nombre de produits à tester et temps d'essai) tout en maximisant la qualité d'estimation.
- appliquer la démarche bayésienne aux essais accélérés basés sur les processus de dégradation et non plus uniquement sur les processus d'apparition des défaillances. En effet, nous pouvons observer des défaillances par dérive où une performance du système se dégrade au cours du temps jusqu'à atteindre un niveau inacceptable.
- proposer des plans permettant de traiter plusieurs modes de défaillance,

-
- utiliser la méthode ϕ_2 pour estimer la fiabilité en environnement aléatoire.

Bibliographie

- [Al-Mutairi, 1996] Al-Mutairi (1996). Multivariate failure models generated by random environments : A unified approach. *Journal of Statistical Planning and Inference* 55.
- [Ansell et Phillips, 1994] Ansell, I. J. et Phillips, M. (1994). *Practical methods for reliability data analysis*. Oxford Science Publications.
- [Arizono et al., 2008] Arizono, I., Kawamura, Y., et Takemoto, Y. (2008). Reliability tests for weibull distribution with variational shape parameter based on sudden death lifetime data. *Eur J Oper Res*, 189(2) :570–574.
- [Bagdonavicius et al., 2000] Bagdonavicius, V., Gerville-Réache, L., Nikoulina, V., et Nikoulin, M. (2000). Expériences accélérées : Analyse statistique du modèle standard de vie accélérée. *Statistique Appliquée*, XLVII(3) :5–38.
- [Bagdonavicius et Nikulin, 1995a] Bagdonavicius, V. et Nikulin, M. (1995a). On accelerated testing of systems. *European Journal of Diagnosis and Safety in Automation*, vol. 5 :307–316.
- [Bagdonavicius et Nikulin, 1995b] Bagdonavicius, V. et Nikulin, M. (1995b). *Semi-parametrics Models in Accelerated Life Testing*. Queen’s University.
- [Bagdonavicius et Nikulin, 1997a] Bagdonavicius, V. et Nikulin, M. (1997a). On application of step stresses in accelerated life testing. *C. R. Acad. Sci. Ser. I Math.*, 325(5) :523–526.
- [Bagdonavicius et Nikulin, 1997b] Bagdonavicius, V. et Nikulin, M. (1997b). Transfer functional and semi-parametric regression. *Biometrika*, vol. 84 :365 – 378.
- [Bagdonavicius et Nikulin, 2001] Bagdonavicius, V. et Nikulin, M. (2001). *Mathematical models in the theory of accelerated experiments*. World Scientific Publishing Co.
- [Bashu et Ebrahimi, 1982] Bashu, A. P. et Ebrahimi, N. (1982). Nonparametric accelerating life testing. *IEEE Trans. on Reliability*, 31 :432–435.
- [Bayarri et Berger, 1998] Bayarri, M. et Berger, J. (1998). Quantifying surprise in the data and model verification (with discussion). *Bayesian Statistics*, 6 :53–82.

- [Bedard, 2000] Bedard, S. (2000). Comportement des structures de signalisation aérienne en aluminium soumises à des sollicitations cycliques . *Mémoire de maîtrise ès sciences appliquées*,.
- [Bellman, 1961] Bellman, R. (1961). *Adaptive Control Processes*.
- [Bousquet, 2006] Bousquet, N. (2006). Subjective bayesian statistics : agreement between prior and data. Technical report, Institut National De Recherche En Informatique Et En Automatique.
- [Bris, 2000] Bris, R. (2000). Bayes approach in rdt using accelerated and long-term life data. *Reliability Engineering and System Safety*.
- [Chapouille, 1972] Chapouille, P. (1972). *La Fiabilité*.
- [Clarotti, 1998] Clarotti, C. (1998). Les techniques fréquentielle et bayésienne au service le l'ingénieur de sûreté de fonctionnement. Technical Report 08/96, ISDF.
- [Congdon, 2001] Congdon, P. (2001). *Bayesian statistical modelling*. Ed Wiley.
- [Cooper et McGillem, 1988] Cooper, G. R. et McGillem, C. D. (1988). *Probabilistic Methods of Signal and System Analysis*. Oxford University Press US.
- [Cross et al., 2007] Cross, R., Makeev, A., et Armanios, E. (2007). Simultaneous uncertainty quantification of fracture mechanics based life prediction model parameters. *Int J Fatigue*, 29(8) :1510–1515.
- [Crowe et Feinberg, 2001] Crowe, D. et Feinberg, A. (2001). *Design for reliability*. Ed CRC Press.
- [Delignières, 2000] Delignières, D. (2000). Séries temporelles - modèles arima. *Séminaire EA "Sport - Performance - Santé"*.
- [Demoulin, 1990] Demoulin, B. (1990). *Base documentaire « Mesures. Généralités »*, chapitre Processus aléatoires.
- [Dorp et Mazzuchi, 2004] Dorp, J. R. et Mazzuchi, A. T. (2004). A general bayes exponential inference model for accelerated life testing. *Journal of statistical planning and inference*.
- [Dorp et al., 1996] Dorp, J. R. V., Mazzuchi, T., Fornell, G., et Pollock, L. (1996). A bayes approach to step-stress accelerated life testing. *IEEE transaction on reliability*, vol 45(3).
- [Dunson et Taylor, 2005] Dunson, D. et Taylor, J. (2005). Approximate bayesian inference for quantites. *J. Nonparametric Stat.*, 17(3) :385–400.
- [et R. KOUTA, 2007] et R. KOUTA, D. P. (2007). *Analyse de chargements aléatoires - BM5030*. Techniques de l'Ingénieurs.

-
- [FIDES, 2004] FIDES (2004). *FIDES Guide 2004 : Reliability Methodology for Electronic Defense System*.
- [Finetti, 1974] Finetti, B. D. (1974). *Theory of probability : a critical introductory treatment*. Wiley.
- [Finkelstein, 2008] Finkelstein, M. (2008). Virtual age of nonrepairable objects. *Reliability Engineering & System Safety*, In Press, Accepted Manuscript :-.
- [Fisher, 1922] Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Trans. of the Royal Society of London*, A 222 :309–368.
- [Fournier, 1993] Fournier, J. (1993). *Fiabilité du logiciel*. Hermes, Paris.
- [Gaudoin, 2002] Gaudoin, O. (2002). Modèles stochastiques et méthodes statistiques pour la fiabilité des systèmes. *Habilitation à Diriger des Recherches*, Institut National Polytechnique de Grenoble.
- [Goldstein, 1999] Goldstein, M. (1999). Bayes linear analysis. *Encyclopaedia of statistical sciences*, 3 :29–34.
- [Grubisic, 1994] Grubisic, V. (1994). Determination of load spectra for design and testing. *International Journal of Vehicle Design*, 15.
- [Guerin et al., 2003] Guerin, F., Dumon, B., et Usureau, E. (2003). New methodology of weighting of prior distribution used bayesian approach applied to reliability estimation. In *Probability Method Conference*, Los Angeles, USA.
- [Guerin et al., 2004] Guerin, F., Dumon, B., et Usureau, E. (2004). *Compatibility between prior and likelihood distributions and weighting approach applied to reliability*. Progresses In Quality and Dependability. Editors, Bacivarov, Dumon, Wiele. A paraître.
- [Guerin et al., 2001] Guerin, F., Tebbi, O., et Dumon, B. (2001). Accelerated testing based on mechanical damage model. In *IEEE International Conference on Reliability and Maintainability*, Philadelphia, Pennsylvania, USA.
- [Han, 1997] Han, M. (1997). The structure of hierarchical prior distribution and its applications. *Chinese Operations Research and Management Science*, 6 :31–40.
- [Hoang, 2003] Hoang, P. (2003). *Handbook of reliability engineering*. ed Springer.
- [Ibrahim et al., 2001] Ibrahim, J., Chen, M., et Sinha, D. (2001). *Bayesian survival analysis*. Ed Springer.
- [Idée, 2001] Idée, E. & Lannoy, A. . M. T. (2001). Estimation of a lifetime law for equipment on the basis of a highly right multicensored sample and expert assesment.

- [Kouta et Play, 2006] Kouta, R. et Play, D. (2006). DurÉE de vie d'un systÈme mÉca-
nique : Analyse de chargements aléatoires. *Techniques de l'Ingénieur*.
- [Lannoy, 1996] Lannoy, A. (1996). *Analyse quantitative et utilité du retour d'expérience
pour la maintenance des matériels et la sécurité*. Eyrolles.
- [Lannoy, 2004] Lannoy, A. (2004). Introduction à la fiabilité des structures. *Base docu-
mentaire «Sécurité et gestion des risques»*.
- [Lannoy et Procaccia, 1994] Lannoy, A. et Procaccia, H. (1994). *Méthodes avancées
d'analyse des bases de données du retour d'expérience industriel*. Ed Eyrolles.
- [Lannoy et Procaccia, 2001] Lannoy, A. et Procaccia, H. (2001). *L'utilisation du jugement
d'expert en sûreté de fonctionnement*. Ed. TEC&DOC.
- [Laprie et al., 1995] Laprie, J. C., Arlat, J., Blanquart, J. P., Costes, A., Crouzet, Y.,
Deswarte, Y., Fabre, J. C., Guillermain, H., Kaâniche, M., Kanoun, K., Mazet, C.,
Powell, D., Rabéjac, C., et Thévenod, P. (1995). Guide de la sûreté de fonctionnement.
Cepadues, Toulouse.
- [Lyonnet, 2006] Lyonnet, P. (2006). *Ingénierie de la fiabilité*.
- [Madi, 1997] Madi, T. M. (1997). Bayesian inference for partially accelerated life tests
using gibbs sampling. *Microelectronics Reliability*, 37(8).
- [Manry et al., 1996] Manry, M., Hsieh, C.-H., Dawson, M., Fung, A., et S.J.Apollo (1996).
Cramer rao maximum a-posteriori bounds on neural network training error for non-
gaussian signals and parameters. *Internat. Journ. of Intelligent Control and Syst.*,
1 :381–391.
- [Marcovici et Ligeron, 1974] Marcovici, C. et Ligeron, J.-C. (1974). *Utilisation des tech-
niques de fiabilité en mécanique*.
- [Marts et Walter, 1982] Marts, H. et Walter, R. (1982). *Bayesian reliability analysis*. Ed
John Willey& sons.
- [Mettas, 2005] Mettas, A. (2005). Reliability predictions based on customers usage stress
profile. *Proceedings Annual Reliability & Maintainability*.
- [Mihalache, 2007] Mihalache, A. G. (2007). *MODELISATION ET EVALUATION DE
LA FIABILITE DES SYSTEMES MECATRONIQUES : APPLICATION SUR SYS-
TEME EMBARQUE*. PhD thesis, Ecole Doctorale d'Angers - Institut des Sciences et
Techniques de l'Ingénieur d'Angers.
- [MIL-HDBK-338B, 1998] MIL-HDBK-338B (1998). *Military handbook electronic reliabi-
lity design handbook*.

-
- [Mortureux, 2001] Mortureux, Y. (2001). La sûreté de fonctionnement : méthodes pour maîtriser les risques. *Techniques de l'Ingénieur*, AG4670 :1–17.
- [Nelson, 1990] Nelson, W. (1990). *Accelerated Testing : Statistical Models, Test Plans and Data Analyses*. John Wiley & Sons.
- [Nikulin, 2007] Nikulin, M & Gerville-Réache, L. . C. V. (2007). *Statistique des essais accélérés*. Hermes Science.
- [O'Connor, 2003] O'Connor, P. (2003). Testing for reliability. *Quality and Reliability Engineering international*, vol 19 :73–84.
- [O'Connor, 2002] O'Connor, P. D. T. (2002). *Practical Reliability Engineering Fourth Edition*. Wiley.
- [Pagès et Gondran, 1980] Pagès, A. et Gondran, M. (1980). *Fiabilité des systèmes*. Eyrolles, Paris.
- [Perroud, 2000] Perroud, G. (2000). Accéder au juste nécessaire par une expérimentation. *extrait de Phoebus no 13, Les essais accélérés*.
- [Pham, 2003] Pham, H. (2003). *Reliability Engineering*. Springer-Verlang, New York, USA.
- [Pitman, 1993] Pitman, J. (1993). *Probability*. Springer Texts in Statistics.
- [Poncet, 2005] Poncet, A. (2005). Fiabilité expérimentale : dimensionnement optimale d'essais sous contraintes. Master's thesis, UNIVERSITE JEAN MONNET.
- [Poncet et al., 2004a] Poncet, A., Dumas, C., Schimmerling, P., Sarriau, O., et Miazga, C. (2004a). Dimensionnement de chapeaux de paliers de moteurs turbo diesel par une approche fiabiliste. *Congrès SF2M, Paris*.
- [Poncet et al., 2004b] Poncet, A., Schimmerlin, P., Lyonnet, P., et andC. Jouzeau, C. T. (2004b). Le dimensionnement des essais de validation par la méthode contrainte - résistance. *extrait du congrès "LambdaMu" 14, Bourges*.
- [Procaccia, 1998] Procaccia, H. (1998). Fondements et applications des approches fréquentielle bayésiennes dans le domaine de la sûreté de fonctionnement. *Edition interne ISDF*.
- [Procaccia et Morilhat, 1996] Procaccia, H. et Morilhat, P. (1996). *Fiabilité des structures des installations industrielles*. Eyrolles, Paris.
- [Procaccia et al., 1992] Procaccia, H., Piepszownik, L., et Clarotti, C. (1992). *Fiabilité des équipements et théorie de la décision statistique fréquentielle et bayésienne*. Ed Eyrolles.

- [Ringler, 1979] Ringler, J. (1979). Utilisation des techniques bayésienne dans le cas de système complexe. *Revue de Statistique Appliquée*, XXVII(2) :-.
- [Robert, 1992] Robert, C. (1992). *L'analyse statistique bayésienne*. Edition ECONOMICA.
- [Robert, 2001] Robert, C. (2001). *The Bayesian Choice. From Decision-Theoretic Motivations to Computational Implementation (second édition)*. Springer-Verlag : New York.
- [Ross, 2003] Ross, T. D. (2003). Accurate confidence intervals for binomial proportion and poisson rate estimation. *Computers in biology and medicine*, vol 33 :509–531.
- [Sander et Badoux, 1991] Sander, P. et Badoux, R. (1991). *Bayesian Methods in Reliability*. Kluwer Academic Publisher.
- [Savard, 2004] Savard, M. (2004). Étude de la sensibilité d'un pont routier aux effets dynamiques induits par la circulation routière. *11^{ème} colloque sur la progression de la recherche québécoise des ouvrages d'art, université Laval, Québec*.
- [Sawa, 1978] Sawa, T. (1978). Information criteria for discriminating among alternative regression models. *Econometrica*, 46 :1273–1292.
- [Schmoyer, 1986] Schmoyer, R. (1986). An exact distribution-free analysis for accelerated life testing at several levels of a single stress. *Technometrics*, 28 :165–175.
- [Schmoyer, 1991] Schmoyer, R. (1991). Nonparametric analysis for two-level single-stress accelerated life tests. *Technometrics*, 33 :175–186.
- [Shyur et al., 1999] Shyur, H.-J., Elsayed, E., et Luxhøj, J. (1999). A general model for accelerated life testing with time-dependent covariates. *Nav Res Logist*, 46(2-3) :303–321.
- [Smith, 2002] Smith, T. (2002). Life consumption monitoring in electronic circuits. *CEPA 2 Reliability Workshop*.
- [Somerville et al., 1997] Somerville, F. I., Dietrich, L. D., et Mazzuchi, A. T. (1997). Bayesian reliability analysis using the dirichlet prior distribution with emphasis on accelerated life testing run in random order. *Nonlinear analysis theory. Methods and applications*, 30.
- [Suhner, 1994] Suhner, M. (1994). *Utilisation de l'analyse bayésienne pour optimiser la démarche de Fiabilité*. Thèse, Université de Nancy.
- [Tebbi, 2005] Tebbi, O. (2005). *ESTIMATION DES LOIS DE FIABILITE EN MECANIQUE PAR LES ESSAIS ACCELERES*. PhD thesis, Ecole Doctorale d'Angers - Institut des Sciences et Techniques de l'Ingénieur d'Angers.

-
- [Tiplica, 2007] Tiplica, T. (2007). Cours de statistique et probabilités.
- [Tribus, 1972] Tribus, M. (1972). *Décisions rationnelles dans l'incertain*.
- [Usureau, 2001] Usureau, E. (2001). Application des méthodes bayésiennes pour l'optimisation des coûts de développement des produits nouveaux. Master's thesis, Institut des Sciences et Techniques de l'Ingenieur d'Angers.
- [Usureau et al., 2004] Usureau, E., Guerin, F., et Dumon, B. (2004). *Démarche bayésienne et applications à la Sûreté de Fonctionnement*. éditions HERMES Lavoisier. Rédaction des chapitres Modélisation de l'a priori et compatibilité de l'a priori et de la vraisemblance.
- [Vassilious et Mettas, 2001] Vassilious, A. et Mettas, A. (2001). Understanding accelerated life-testing analysis. In *IEEE Proceedings Annual Reliability and Maintainability Symposium*, USA.
- [Villemeur, 1988] Villemeur, A. (1988). *Sûreté de fonctionnement des systèmes industriels : fiabilité, facteurs humains, informatisation*. Eyrolles.