

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 ÉTAT DE L'ART	5
1.1 Phénomène de cavitation	5
1.1.1 Présentation générale	5
1.1.2 Cas particulier des turbines Francis	8
1.1.3 Érosion par cavitation	11
1.2 Modélisation par trajectoire de dégradation	14
1.3 Processus stochastiques	16
1.3.1 Définitions	16
1.3.2 Processus de Poisson	18
1.3.3 Processus de Wiener et de diffusion	19
1.3.4 Processus inverse gaussien	24
1.3.5 Processus gamma	25
1.3.5.1 Processus gamma avec covariables comme modèle de dégradation	32
1.3.5.2 Estimation des paramètres	32
1.3.5.3 Simulation d'un processus gamma	36
1.3.5.4 Calcul du temps d'atteinte et durée de la de vie résiduelle	37
1.4 Sélection de modèles	39
1.4.1 Critères de sélection	41
1.4.2 Tests d'adéquation	42
1.4.3 Méthode graphique	43
1.5 Traitement des incertitudes	45
1.5.1 Intervalles de confiance asymptotiques à partir des propriétés de convergence de l'EMV	46
1.5.2 Méthode delta	48
1.5.3 Méthodes de ré-échantillonnage	49
1.5.3.1 Méthode paramétrique	49
1.5.3.2 Méthodes non paramétriques	50
1.6 Conclusion	52
CHAPITRE 2 ÉTUDE DE LA MÉTHODE D'ESTIMATION PAR MAXIMUM DE VRAISEMBLANCE	55
2.1 Introduction	55
2.2 Performance de la méthode d'EMV	55
2.2.1 Analyse des résultats	62
2.3 Estimation des paramètres d'un processus gamma en deux étapes	65
2.4 Conclusion	74

CHAPITRE 3	MODÉLISATION DES DONNÉES DE LABORATOIRE	77
3.1	Introduction	77
3.2	Dispositifs de tests de cavitation érosive	78
3.3	Matériaux étudiés	80
3.4	Comparaison des matériaux	83
3.4.1	Analyse statistique préliminaire	83
3.4.2	Estimation des paramètres - Comparaison des modèles	84
3.4.3	Analyse des résultats	86
3.5	Influence de l'incertitude des paramètres sur le temps d'atteinte	89
3.5.1	Construction des intervalles de confiance	89
3.5.1.1	Méthode asymptotique	90
3.5.1.2	Méthodes de ré-échantillonnage	91
3.5.2	Propagation des incertitudes - Méthode delta	93
3.5.3	Indicateurs statistiques de performance	94
3.5.4	Analyse des résultats	95
3.6	Conclusion	101
CHAPITRE 4	MODÉLISATION DE LA DÉGRADATION DES ROUES DE TURBINE À PARTIR DES DONNÉES DE MONITORING	103
4.1	Acquisition et pré-traitement des données	103
4.1.1	Dispositif de monitoring en continu	103
4.1.2	Distinction entre cavitation érosive et non érosive	106
4.1.3	Regroupement par grappes (<i>clustering</i>)	106
4.1.4	Calcul de la perte de masse	111
4.1.5	Correction des données	113
4.2	Analyse préliminaire	114
4.3	Modélisation par processus gamma - Année 2011	119
4.3.1	Estimation des paramètres du processus	119
4.3.2	Calcul des intervalles de confiance sur les paramètres	120
4.3.3	Vérification - Simulation de trajectoires	122
4.3.4	Correction de la variance des processus	123
4.4	Vérification et validation avec les données disponibles	126
4.4.1	Regroupement des MO inconnus	128
4.4.2	Prédiction de la dégradation à partir de la modélisation de l'année 2011	133
4.5	Conclusions	140
CONCLUSION ET RECOMMANDATIONS	141
BIBLIOGRAPHIE	147

LISTE DES TABLEAUX

		Page
Tableau 1.1	Types de processus.....	17
Tableau 1.2	Plan d'expérience.	29
Tableau 2.1	Plan d'expérience.	56
Tableau 2.2	Indicateurs statistiques utilisés pour évaluer les performance de la méthode d'estimation MV.	57
Tableau 2.3	Résultats de simulation et d'estimation.	58
Tableau 2.4	Résultats de simulation et d'estimation (suite).	59
Tableau 2.5	Résultats de simulation et d'estimation (suite).	60
Tableau 2.6	Résultats de simulation et d'estimation (suite).	61
Tableau 2.7	Plan d'expérience.	67
Tableau 2.8	Indicateurs statistiques de performance.....	69
Tableau 2.9	Test de sensibilité de la méthode de simulation <i>GSS</i>	72
Tableau 2.10	Comparaison de deux méthodes d'estimation.	73
Tableau 3.1	Temps d'incubation $t_{inc,j}$ des essais de cavitation réalisés sur deux matériaux usuels.	83
Tableau 3.2	Paramètres estimés d'un PGNH.....	87
Tableau 3.3	Simulation du temps d'atteinte.....	89
Tableau 3.4	Indicateurs statistiques utilisés pour caractériser les performances des méthodes <i>bootstrap</i>	95
Tableau 3.5	Estimations de θ par la méthode MV et 3 techniques de <i>bootstrap</i> différentes - Approximation des intervalles de confiance asymptotiques (à 95%) de θ	96
Tableau 3.6	Indicateurs statistiques pour les différentes méthodes <i>bootstrap</i>	97

Tableau 3.7	Estimation des quantiles du temps d'atteinte : $P_{T_p}^{2,5\%}$, $P_{T_p}^{50\%}$, $P_{T_p}^{97,5\%}$ et de leur incertitude associée pour les 3 méthodes <i>bootstrap</i> et MV + MD.	98
Tableau 4.1	Extrait des données d'opération pour l'année 2011.	110
Tableau 4.2	Dénombrement des instants mesurés pour l'année 2011 par grappe.	114
Tableau 4.3	Analyse des modes opératoires et des incréments de dégradation associés pour l'année 2011.	119
Tableau 4.4	Estimation des paramètres du processus pour chaque grappe.	119
Tableau 4.5	$IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2011.	121
Tableau 4.6	Analyse des modes opératoires et des incréments de dégradation associés pour 2012.	129
Tableau 4.7	Analyse des modes opératoires et des incréments de dégradation associés pour l'année 2013.	129
Tableau 4.8	Estimation des paramètres du processus pour chaque grappe - Données de l'année 2012.	132
Tableau 4.9	$IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2012.	132
Tableau 4.10	Estimation des paramètres du processus pour chaque grappe - Données de l'année 2013.	133
Tableau 4.11	$IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2013.	133
Tableau 4.12	Estimation des paramètres du processus pour chaque grappe - Concaténation des années 2011 et 2012.	136
Tableau 4.13	$IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données des années 2011+2012.	137
Tableau 4.14	Estimation des paramètres du processus pour chaque grappe - Concaténation des années 2011, 2012 et 2013.	137
Tableau 4.15	$IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données des années 2011+2012+2013.	137

LISTE DES FIGURES

	Page
Figure 1.1	Diagramme d'état de l'eau..... 6
Figure 1.2	Zones de dépression et de surpression lors d'un écoulement sur un profil hydrodynamique. 6
Figure 1.3	Exemples de cavitation se formant sur un hydrofoil : a) deux bulles de cavitation sur l'extrados de la pale, b) nuage de bulles de cavitation en mouvement sur une pale. Tirée et adaptée de d'Agostino & Salvetti (2008). 8
Figure 1.4	Représentation schématique d'une installation hydroélectrique typique avec une turbine Francis. 9
Figure 1.5	Représentation schématique d'une turbine Francis. Tirée et adaptée de Avellan (2004). 10
Figure 1.6	Exemple de colline de rendement d'un modèle réduit de turbine Francis. Tirée et adaptée de Guo, Wang, Sun & Luo (2017)..... 11
Figure 1.7	Bulles hémisphériques implosant à la surface d'une paroi. Tirée et adaptée de Benjamin & Ellis (1966). 12
Figure 1.8	Photographies montrant le développement de micro-jets lors de l'implosion d'une bulle. Tirée et adaptée de (Tomita & Shima, 1990). 13
Figure 1.9	Exemples d'érosion observée sur les pales de turbine Francis. 13
Figure 1.10	Exemple de trajectoires simulées pour un processus de Poisson..... 20
Figure 1.11	Exemple de trajectoires simulées pour un processus de Wiener avec dérive. 22
Figure 1.12	Exemple de trajectoires simulées pour un processus d'Ornstein-Uhlenbeck. 23
Figure 1.13	Exemple de trajectoires simulées pour un processus inverse gaussien..... 25
Figure 1.14	Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène..... 29
Figure 1.15	Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène..... 30

Figure 1.16	Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène.	31
Figure 1.17	Exemple de notation des données.	34
Figure 1.18	Temps d'atteinte T_ρ d'une trajectoire de dégradation au seuil ρ	37
Figure 1.19	Classification simple de processus stochastiques couramment rencontrés en fiabilité.	40
Figure 1.20	Exemple de diagramme quantile-quantile pour la comparaison de deux processus stochastiques.	45
Figure 2.1	Comparaison du biais relatif et de l'erreur relative d'estimation pour un nombre de trajectoires m et pour des paramètres de PG : (1; 1; 1).	62
Figure 2.2	Comparaison du biais relatif et de l'erreur relative d'estimation pour un nombre de trajectoires m et pour des paramètres de PG : (1; 1; 2).	63
Figure 2.3	Comparaison du biais relatif et de l'erreur relative d'estimation pour différentes tailles d'échantillon N et valeurs de paramètres de PG.	63
Figure 2.4	Enveloppe des simulations d'un processus gamma.	70
Figure 2.5	Enveloppe des simulations d'un processus gamma (suite).	71
Figure 2.6	Distributions des estimations du processus gamma par la méthode en deux étapes.	74
Figure 3.1	Représentation schématique du dispositif de test par jet cavitant (CavitJet).	79
Figure 3.2	Représentation schématique du dispositif de test par cavitation vibratoire. Tirée et adaptée de ASTM (2016).	80
Figure 3.3	Étapes caractéristiques de l'évolution de la courbe d'érosion en fonction du temps. Tirée et adaptée de ASTM (2016).	81
Figure 3.4	Trajectoires de dégradation pour le matériau E309L.	83
Figure 3.5	Matériau E309L - Simulations d'un processus gamma : a) homogène et b) non homogène - 10000 trajectoires simulées.	85
Figure 3.6	Matériau E309L - Densités de probabilité du niveau de dégradation aux instants : 360, 600 et 960 min pour des trajectoires simulées d'un processus gamma : a) homogène et b) non homogène.	85

Figure 3.7	Matériau E309L - Densités de probabilité du temps d'atteinte pour $\rho = 100$ mg de trajectoires simulées d'un processus gamma : a) homogène et b) non homogène.	86
Figure 3.8	Matériau AISI 415 - a) Simulations d'un processus gamma non homogène - b) Densités de probabilité du niveau de dégradation aux instants 360, 600 et 960 min - c) Densité de probabilité du temps d'atteinte au seuil $\rho = 100$ mg.	88
Figure 3.9	Comparaison des distributions du temps d'atteinte pour les seuils $\rho = (20, 50, 100, 150)$ mg et les deux matériaux.	88
Figure 3.10	$\hat{\theta}_{EB}$, $\hat{P}_{\hat{\theta}_{EB}^*}^{2,5\%}$ et $\hat{P}_{\hat{\theta}_{EB}^*}^{97,5\%}$ - Analyse de convergence sur le nombre d'itérations <i>bootstrap</i> B à réaliser pour la méthode EB.	93
Figure 3.11	Comparaison des IC asymptotiques et <i>bootstrap</i>	95
Figure 3.12	Propagation des incertitudes sur la distribution du temps d'atteinte pour $\rho = 100$ mg.	99
Figure 3.13	Distribution du temps d'atteinte T_ρ (avec $\rho = 100$ mg) pour les 3 méthodes <i>bootstrap</i> et MV + MD.	100
Figure 4.1	Processus mixte d'acquisition et pré-traitement des données de monitoring.	104
Figure 4.2	Zone de manifestation de la cavitation et emplacements des capteurs (gauche) et possibles chemins empruntés par les vibrations (droite). Figure tirée et adaptée de Escaler, Egusquiza, Farhat, Avellan & Coussirat (2006).	105
Figure 4.3	Exemples de cavitation rencontrés sur une turbine de type Francis : a) entrée extradados, b) torche hélicoïdale. Figures tirées et adaptées de Franc, Avellan, Belahadji, Billard, Briançon-Marjollet, Fréchou, Fruman, Karimi, Kueny & Michel (1995) et Escaler <i>et al.</i> (2006)	107
Figure 4.4	Schématisation de la chambre d'équilibre de la centrale et de ses 8 turbines.	108
Figure 4.5	Dendrogramme des modes opératoires. Tirée et adaptée de Bodson-Clermont (2016).	112
Figure 4.6	Trajectoire de dégradation du groupe T/A étudié pour l'année 2011 - Perte de masse cumulée en fonction du temps.	116

Figure 4.7	Trajectoires de dégradation par grappe pour l'année 2011 - Perte de masse cumulée en fonction du temps.	117
Figure 4.8	Comparaison des intensités de cavitation par grappe.	118
Figure 4.9	Comparaison des trajectoires simulées à partir des différents processus gamma.	122
Figure 4.10	Algorithme de simulation de trajectoires indexé sur les données réelles.	124
Figure 4.11	Enveloppe des trajectoires simulées - comparaison à la trajectoire de référence pour l'année 2011.	125
Figure 4.12	Enveloppe des trajectoires simulées pour une variance multipliée par : a) 10 et b) 100 - Trajectoire de référence année 2011.	127
Figure 4.13	Enveloppe des trajectoires simulées pour une variance multipliée par 1000 - Trajectoire de référence année 2011.	128
Figure 4.14	Trajectoire de dégradation du groupe T/A étudié pour l'année 2012 - Perte de masse cumulée en fonction du temps.	130
Figure 4.15	Trajectoire de dégradation du groupe T/A étudié pour l'année 2013 - Perte de masse cumulée en fonction du temps.	131
Figure 4.16	Trajectoires de dégradation pour les années 2011 à 2013 - Perte de masse cumulée en fonction du temps.	134
Figure 4.17	Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir de l'année 2011.	135
Figure 4.18	Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir des années 2011 et 2012.	138
Figure 4.19	Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir des années 2011 à 2013.	139

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AIC	Critère d'Information d'Akaike (<i>Akaike Information Criterion</i>)
BIC	Critère d'Information Bayésien (<i>Bayesian Information Criterion</i>)
BP	<i>Bootstrap</i> paramétrique
CDF	Fonction de densité cumulée (<i>Cumulative Distribution Function</i>)
EB	Efron <i>Bootstrap</i>
EMV	Estimation du Maximum de Vraisemblance
EQM	Erreur Quadratique Moyenne
GSS	<i>Gamma Sequential Sampling</i>
IC	Intervalle de Confiance
i.i.d.	indépendant identiquement distribué
i.n-i.d.	indépendant non-identiquement distribué
MBB	<i>Moving Block Bootstrap</i>
MD	Méthode Delta
MO	Mode Opérateur
MM	Méthode des Moments
MV	Maximum de Vraisemblance
PDF	Fonction de densité de probabilité (<i>Probability Density Function</i>)
PG	Processus Gamma
PGH	Processus Gamma Homogène
PGNH	Processus Gamma Non Homogène
Q-Q	Quantile-Quantile
RUL	Durée de vie résiduelle (<i>Remaining Useful Lifetime</i>)
T/A	Turbine Alternateur
TCL	Théorème Central Limite

LISTE DES SYMBOLES

\mathbb{R}	l'ensemble des nombre réels
\mathbb{R}^+	l'ensemble des nombre réels non-négatifs
\mathbb{R}_*^+	l'ensemble des nombre réels positifs
\mathbb{N}	l'ensemble des nombre entiers naturels
$X = (X_t)_{t \geq 0}$	un processus stochastique
\mathbb{F}	fonction de répartition
ρ	seuil de dégradation
T_ρ	premier temps d'atteinte du niveau de dégradation ρ
u	paramètre d'échelle d'un processus gamma
$v(t)$	fonction de forme d'un processus gamma
c	paramètre de la fonction de forme d'un processus gamma
b	paramètre de la fonction de forme loi puissance
$Ga(x u, v)$	densité de probabilité de la distribution gamma de paramètres u et v
$\Gamma()$	fonction gamma incomplète
Φ	la CDF d'une loi normale centrée réduite
$\psi()$	fonction di-gamma
\mathbf{D}_{obs}	jeu de données de dégradation
$x_{i,j}$	niveau de dégradation observé à l'instant $t_{i,j}$ de la trajectoire j
n_j	nombre d'observation de la trajectoire j
m	nombre de trajectoires de \mathbf{D}_{obs}
N	nombre total d'observations
$t_{i,j}$	instant d'inspection i de la trajectoire j
$\delta_{i,j}$	incrément de dégradation entre les instants $t_{i,j}$ et $t_{i+1,j}$

$\mathbb{E}(\cdot)$	espérance de \cdot
$\mathbb{V}(\cdot)$	variance de \cdot
\mathcal{L}	fonction de vraisemblance
ℓ	fonction log-vraisemblance
θ	vecteur des paramètres d'une distribution gamma
$\hat{\cdot}$	valeur estimée de \cdot
B	Nombre d'échantillons <i>bootstrap</i>
p^q	Quantile d'ordre q de \cdot

INTRODUCTION

Le domaine du pronostic connaît une forte croissance et un intérêt de plus en plus fort depuis déjà plusieurs années (Venkatasubramanian, 2005). L'attention portée à ce domaine s'est d'autant plus accrue avec l'arrivée des objets connectés (*internet of things (IoT)*) dans le milieu industriel. Encore plus récemment, avec le déploiement de la technologie de télécommunication 5G, le nombre d'applications de suivi en temps réel devrait faire bondir les opportunités en pronostic.

Historiquement, le pronostic était déjà très répandu dans les domaines de l'industrie tels que : l'aéronautique, la médecine, l'électronique ou encore l'énergie ; des secteurs d'activité régis à la fois par de fortes contraintes réglementaires de sûreté, et des intérêts économiques importants. Pour rappel, le pronostic (en industrie) consiste à prédire l'état de dégradation futur d'un composant ou d'un système et donc d'en déduire sa durée de vie résiduelle (le délai avant qu'il ne soit défaillant) (Lebold & Thurston, 2001 ; Muller, Suhner & Iung, 2008). Le besoin d'une telle démarche est justifié par la mise en œuvre de politiques de maintenance ou de processus décisionnels qui sont basés sur la prise en compte de l'état actuel et passé d'un composant, de l'environnement dans lequel il évolue, et éventuellement des conditions futures d'opération de celui-ci. Par l'adoption de cette approche, les industriels montrent leur volonté de se défaire le plus possible des approches de maintenance classiques comme la systématique (planifiée).

C'est d'ailleurs dans cette optique que s'inscrit le présent projet de recherche, initié par deux grands acteurs de l'énergie : Hydro-Québec (QC, Canada) et Electricité De France (France). En particulier, ces derniers s'intéressent à la prévision de dégradation des roues de turbines hydroélectriques. En effet, ces composants font partie d'installations stratégiques, en particulier au Québec dont l'électricité est à 99% (60% à l'échelle du Canada) d'origine hydroélectrique (contre environ 11% pour la France) (ACH, 2015). Avec des puissances installées respectives de 78 000 MW (ACH, 2016) et 25 000 MW (EDF, 2019), le Canada et la France possèdent un parc de machines hydrauliques conséquent dont la maintenance représente un enjeu économique fort.

En effet, la relative flexibilité d'opération des centrales hydroélectriques, offerte notamment par les réservoirs d'eau de grandes capacités (encore plus vrai pour le Canada), leur permet de lisser la production d'électricité non constante des autres installations. Une pression de plus en plus forte est exercée sur le parc de machines hydroélectriques, due notamment à l'augmentation des parts d'énergies renouvelables de type solaire et éolien et à un marché de l'énergie de plus en plus concurrentiel (en France, la production d'électricité d'appoint est assurée par les centrales hydrauliques notamment, en raison du fonctionnement en base des installations nucléaires d'où provient 70% de l'énergie totale produite). Ceci résulte en des sollicitations plus fréquentes des groupes turbine/alternateur (nombre de démarrages et arrêts en augmentation) et en une utilisation en dehors des plages nominales de fonctionnement menant à une dégradation prématurée des équipements, notamment causée par la cavitation érosive.

Dans ce contexte, les projets THAUR (Turbines Hydrauliques Adaptées à leur Utilisation Réelle) et AMPH (Asset Management pour l'Hydraulique) ont vu le jour au sein de l'Institut de Recherche d'Hydro Québec (IREQ) et EDF R&D respectivement, afin de pallier les problèmes de maintenance et de limiter les indisponibilités des machines. Parmi les nombreux objectifs de ces deux projets, on peut citer les suivants :

- Réduire la durée des arrêts de groupe turbine/alternateur,
- Réduire le coût lié aux opérations de réfection des machines,
- Réduire les temps de réfection et ainsi les pertes d'exploitation.

Pour parvenir à ces objectifs, il est indispensable de considérer quels sont les paramètres influençant la dégradation de la turbine, à savoir par exemple : les conditions d'exploitation, les conditions dommageables, les matériaux employés, la nature électrochimique de l'eau. En ayant les connaissances nécessaires sur ces paramètres, il est alors possible d'adapter l'utilisation de la turbine à ces conditions, dans le cas d'une installation existante, ou bien d'adapter la turbine à son utilisation réelle, dans le cas de remplacement d'équipements.

Le présent projet de recherche s'inscrit directement dans la feuille de route des deux programmes de recherche introduits précédemment, en proposant un outil d'évaluation de l'état de santé des roues de turbines hydroélectriques.

CHAPITRE 1

ÉTAT DE L'ART

Dans ce chapitre, nous introduisons en premier lieu le phénomène de dégradation par cavitation érosive. Par la suite, seront présentés différents modèles probabilistes de dégradation fréquemment utilisés dans le domaine de la fiabilité des matériels industriels. Enfin, un volet traitant des incertitudes liées aux estimations des paramètres des modèles de dégradation sera introduit.

1.1 Phénomène de cavitation

1.1.1 Présentation générale

Comme évoqué dans l'introduction, les turbines hydroélectriques étudiées font l'objet d'une dégradation par cavitation érosive. En premier lieu, il convient d'introduire quelques notions relatives à ce phénomène thermodynamique et mécanique.

Franc *et al.* (1995) définit la cavitation comme "la rupture du milieu continu de liquide sous l'effet de contraintes excessives". Plus simplement, on peut définir ce phénomène physique par l'apparition de bulles de vapeur dans un environnement liquide due à des variations de pression. Comme le montre le diagramme d'état de l'eau en Figure 1.1, un changement d'état de liquide à gazeux peut apparaître lorsque la température augmente ou lorsque la pression diminue, en passant en dessous de la pression de vapeur saturante (limite modélisée entre la région bleue et verte de la Figure 1.1). C'est ce dernier mécanisme qui opère pour la cavitation, appelée aussi ébullition à froid.

Lorsqu'on considère maintenant le profil de pale d'une machine hydraulique (voir Figure 1.2), l'écoulement du fluide sur celui-ci génère des zones de surpression et de dépression (zones en rouge, les flèches indiquent la direction des champs de pression). C'est dans ces zones de dépression, situées notamment sur l'extrados des pales de turbine, que la cavitation peut se manifester sous certaines conditions.

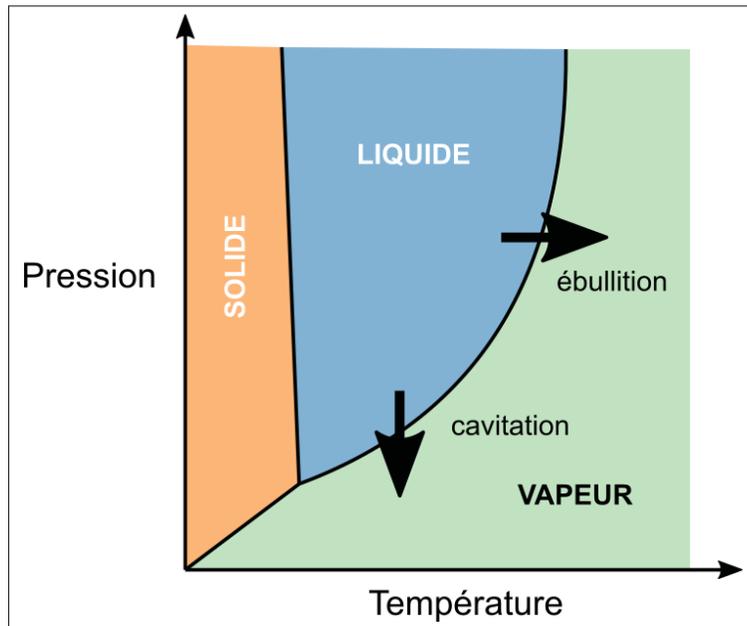


Figure 1.1 Diagramme d'état de l'eau.

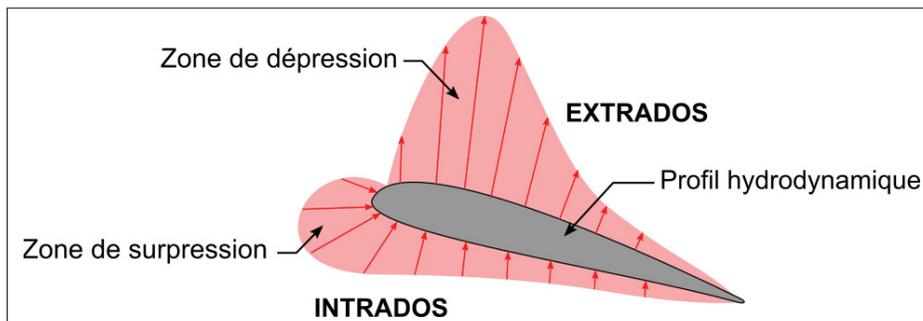


Figure 1.2 Zones de dépression et de surpression lors d'un écoulement sur un profil hydrodynamique.

Rayleigh fut le premier en 1917 à décrire les conséquences de l'apparition de la cavitation sur les hélices de bateau. À cette occasion, on commence alors à parler de cavitation érosive pour caractériser la dégradation résultante. Il donna d'ailleurs son nom à l'équation de Rayleigh-Plesset qui définit la dynamique d'une bulle sphérique plongée dans un corps infini et incompressible (Brennen, 2014). Cette équation différentielle, fréquemment utilisée dans l'étude du phénomène

de cavitation, est présentée dans l'expression (1.1) :

$$R(t) \frac{d^2 R(t)}{dt^2} + \frac{3}{2} \left(\frac{dR(t)}{dt} \right)^2 + \frac{4\nu_L}{R(t)} \frac{dR(t)}{dt} + \frac{2\gamma}{\rho_L R(t)} + \frac{\Delta P(t)}{\rho_L} = 0 \quad (1.1)$$

avec $R(t)$ le rayon variable de la bulle en fonction du temps, ρ_L la densité du liquide environnant (supposé constant), ν_L la viscosité cinétique du liquide environnant (supposé constant), γ la tension de surface à l'interface bulle-liquide, $\Delta P(t) = P_\infty(t) - P_B(t)$, avec $P_B(t)$ la pression à l'intérieur de la bulle et $P_\infty(t)$ la pression du fluide à une distance infinie de la bulle. Cette expression permet notamment de décrire l'évolution d'une bulle : initialement dans un état de nucléation, elle est transportée par un fluide dont la pression varie selon son écoulement le long d'une pale. La bulle va alors croître dans les zones de basse pression, devenant macroscopique, pour finalement imploser dans une région où la pression devient plus importante (d'Agostino & Salvetti, 2008). Cette évolution est illustrée par les photographies présentées dans la Figure 1.3.

Pour caractériser le potentiel de cavitation d'un écoulement, on utilise le nombre adimensionnel σ , appelé paramètre de cavitation. Celui-ci est donné dans l'équation (1.2) (d'Agostino & Salvetti, 2008) :

$$\sigma = \frac{p_{ref} - p_v}{\frac{1}{2}\rho v^2}, \quad (1.2)$$

avec p_{ref} la pression locale du liquide en un point donné, p_v la pression de vapeur du fluide, ρ la masse volumique du fluide et v la vitesse du fluide. Pour les applications à des profils hydrodynamiques (de type pale, par exemple), on prendra p_{ref} en un point éloigné de ce profil. Intuitivement, plus la différence entre p_{ref} et p_v est faible, plus le liquide se rapproche de la limite de changement d'état de liquide à gazeux. Ce point limite est atteint pour σ_i , le paramètre de cavitation naissante. Pour s'éloigner de ce point critique, et avoir ainsi un écoulement sans cavitation, p_{ref} doit donc être élevé et par conséquent σ aussi. Selon le même raisonnement, on se rapproche du point critique σ_i lorsque les pressions restent inchangées mais que la vitesse v de l'écoulement augmente.

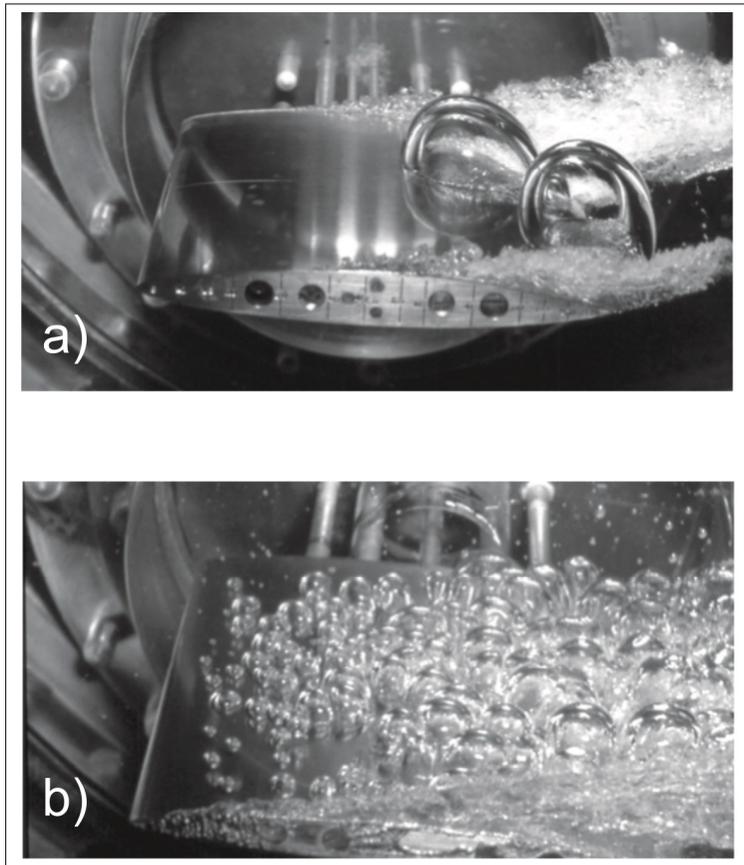


Figure 1.3 Exemples de cavitation se formant sur un hydrofoil : a) deux bulles de cavitation sur l'extrados de la pale, b) nuage de bulles de cavitation en mouvement sur une pale. Tirée et adaptée de d'Agostino & Salvetti (2008).

1.1.2 Cas particulier des turbines Francis

Dans le cas des installations hydroélectriques, une des préoccupations lors de la conception est de s'assurer du fonctionnement des turbines sans l'apparition de cavitation. Pour cela, les ingénieurs peuvent notamment s'appuyer sur des règles de conception simples issues de la mécanique des fluides. Par analogie avec le paragraphe précédent, il est aussi possible de déterminer le caractère cavitant d'un écoulement au travers d'une machine hydraulique (comme une turbine Francis par exemple). Afin d'introduire les futures notations, les Figures 1.4 et 1.5 font respectivement apparaître la géométrie générique d'un barrage hydroélectrique et la géométrie d'une turbine de

type Francis. La hauteur de chute brute, notée H_g , représente la différence entre les niveaux amont et aval des réservoirs d'un barrage. La hauteur nette H est une représentation de la hauteur de chute équivalente une fois les différentes pertes retranchées (par friction par exemple).

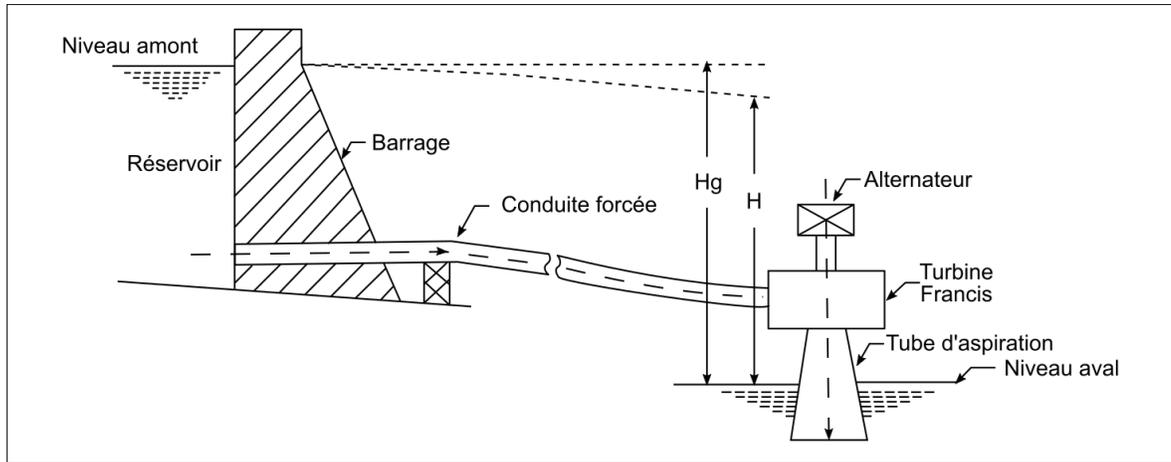


Figure 1.4 Représentation schématique d'une installation hydroélectrique typique avec une turbine Francis.

Enfin, la hauteur nette d'aspiration h_s représente la hauteur entre la bêche spirale et la sortie du tube d'aspiration. On définit alors le nombre de Thoma σ_i (adimensionnel également), représentant les propriétés de cavitation du fluide, avec l'expression suivante :

$$\sigma_i = \frac{p_{atm} - \rho g h_s - p_i}{\rho g H}, \quad (1.3)$$

avec p_{atm} la pression atmosphérique, ρ la densité de l'eau, g l'accélération de la pesanteur et p_i la pression de vapeur saturante de l'eau. Le nombre de Thoma de la centrale σ_p représente quant à lui les propriétés de cavitation de la centrale avec une pression de vapeur équivalente p_v et s'exprime de la façon suivante :

$$\sigma_p = \frac{p_{atm} - \rho g h_s - p_v}{\rho g H}. \quad (1.4)$$

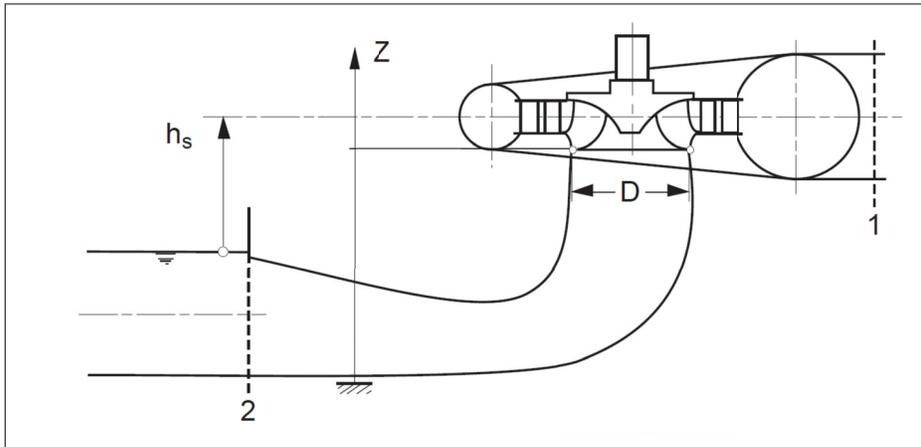


Figure 1.5 Représentation schématique d'une turbine Francis.
Tirée et adaptée de Avellan (2004).

Pour assurer le fonctionnement de la turbine sans cavitation, les conditions d'opération doivent vérifier $\sigma_p > \sigma_i$ (Celebioglu, Altintas, Aradag & Tascioglu, 2017). Chaque installation hydroélectrique étant différente, les fabricants de turbines doivent prouver numériquement (par des simulations en mécanique des fluides), puis à l'aide d'essais sur modèles réduits, que leur nouveau prototype sera exempt de cavitation dans les conditions de fonctionnement nominales. En particulier, les turbines sont généralement conçues pour fonctionner dans des conditions qui maximisent leur rendement. Pour cela, les fabricants s'appuient sur les collines de rendement générées à partir des essais de modèles réduits. En pratique, il s'agit du meilleur compromis entre puissance générée et volume d'eau turbiné. La Figure 1.6 montre un exemple de colline de rendement. Ces mêmes graphiques sont aussi souvent utilisés pour mettre en évidence les conditions de fonctionnement limites auxquelles la cavitation peut apparaître.

D'un point de vue industriel, le phénomène de cavitation est dans une grande majorité des cas considéré comme un phénomène nuisible (Brennen, 2014 ; Gindroz, 1991 ; Lafleur, 2012 ; Plastre, Devals, Gauthier, Nennemann, Trépanier & Guibault, 2019). En effet, le mécanisme d'implosion des bulles de cavitation entraîne des contraintes mécaniques à la surface des matériaux qui résulte en une érosion de celle-ci. Ce mécanisme de dégradation sera décrit plus amplement dans la section suivante. On peut néanmoins souligner que des applications industrielles tirent

bénéfice du phénomène de cavitation (Chahine, Kapahi, Choi & Hsiao, 2016; Ohl, Arora, Dijkink, Janve & Lohse, 2006; Yusof, Babgi, Alghamdi, Aksu, Madhavan & Ashokkumar, 2016). Entre autres, on retrouve souvent des méthodes de nettoyage par ultrason qui reposent sur l'explosion de bulles à la surface des matériaux traités. Cette technologie permet d'atteindre des zones isolées, pour lesquelles une action mécanique serait impossible. En médecine, cette méthode est aussi employée pour réaliser des clichés d'imagerie ou dans le cadre de traitements localisés non-invasifs (Burgess, Apostolakis & Konofagou, 2018; Escudero, Goudot, Vion, Tanter & Pernot, 2018).

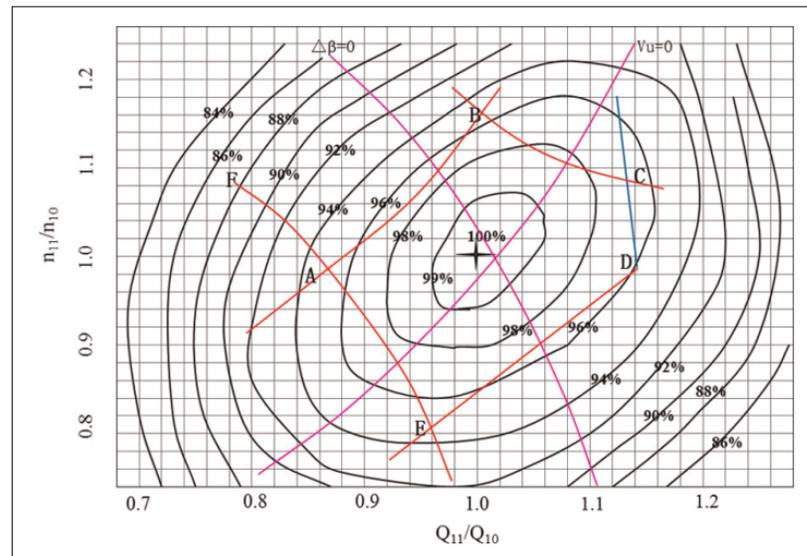


Figure 1.6 Exemple de colline de rendement d'un modèle réduit de turbine Francis. Tirée et adaptée de Guo *et al.* (2017).

1.1.3 Érosion par cavitation

Les dommages engendrés par la cavitation sont largement constatés dans de nombreux domaines de l'ingénierie (d'Agostino & Salvetti, 2008; Jian, Petkovšek, Houlin, Širok & Dular, 2015; Kumar & Saini, 2010). Les experts en cavitation s'accordent à dire que deux mécanismes de dégradation entrent en jeu : le premier consiste en la formation d'un micro-jet violent lors

de l'implosion de la bulle de vapeur. Ce micro-jet, avec une vitesse évaluée jusqu'à 100 m/s (Franc, Riondet, Karimi & Chahine, 2012), entraîne très localement à la surface du matériau une contrainte élevée résultant en la formation de trous (cette contrainte excédant la contrainte à la rupture du matériau). Le deuxième repose sur l'onde de choc de large amplitude générée lors de l'implosion de la bulle (d'Agostino & Salvetti, 2008). Ces deux mécanismes sont illustrés respectivement avec les Figures 1.7 et 1.8. Ils ont pour conséquence de solliciter mécaniquement le matériau, de façon répétée et localisée, entraînant des ruptures par fatigue. À la longue, ces micro-ruptures résultent en une érosion prononcée du matériau dans les zones exposées à la cavitation. La Figure 1.9 montre une pale de turbine Francis ayant été exposée à de la cavitation. La sévérité de l'érosion peut même aller jusqu'à la perforation de la pale, comme montré dans cette même figure (noter que la pale a subi au préalable une préparation de la surface avant réparation).

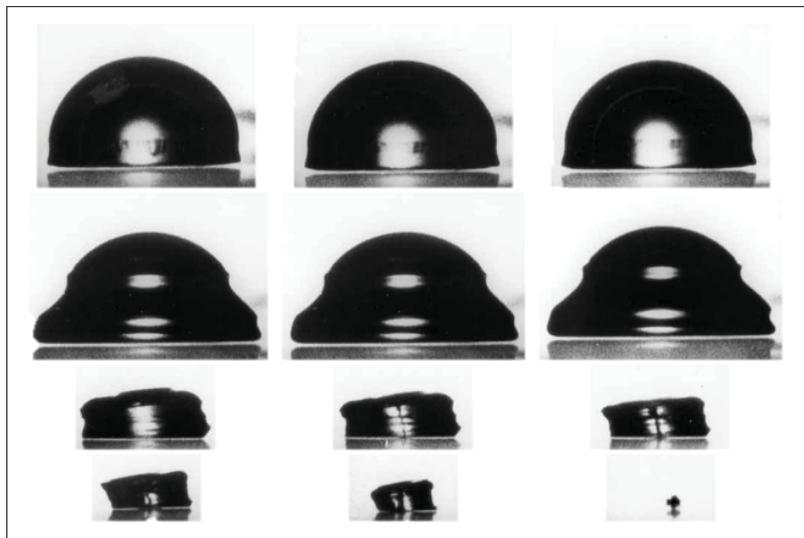


Figure 1.7 Bulles hémisphériques implosant à la surface d'une paroi. Tirée et adaptée de Benjamin & Ellis (1966).

Bien qu'intensément étudiée (Crum, 1995 ; Jian *et al.*, 2015 ; Plastre *et al.*, 2019), l'érosion par cavitation dans les turbines à réaction, comme les Francis, semble inévitable selon Kumar & Saini (2010). Un constat similaire est dressé par Peters, Lantermann & el Moctar (2018) à propos des hélices de bateaux, plus largement répandues. En effet, bien que conçues pour fonctionner sans

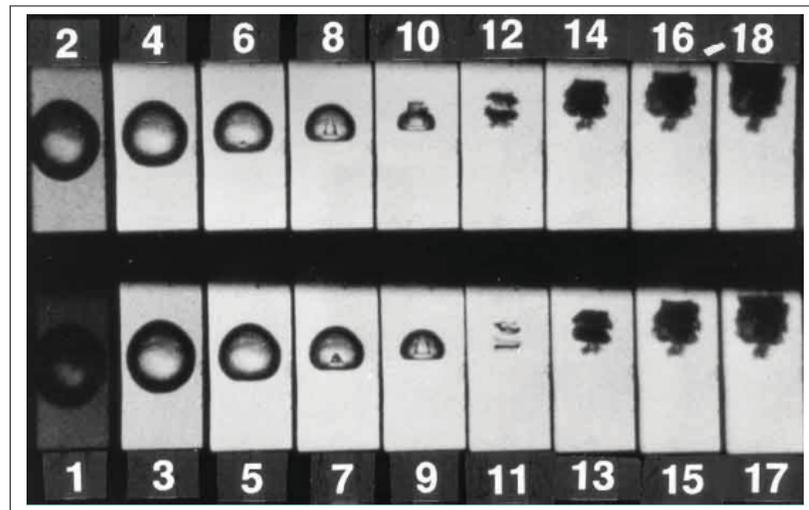


Figure 1.8 Photographies montrant le développement de micro-jets lors de l'implosion d'une bulle. Tirée et adaptée de (Tomita & Shima, 1990).

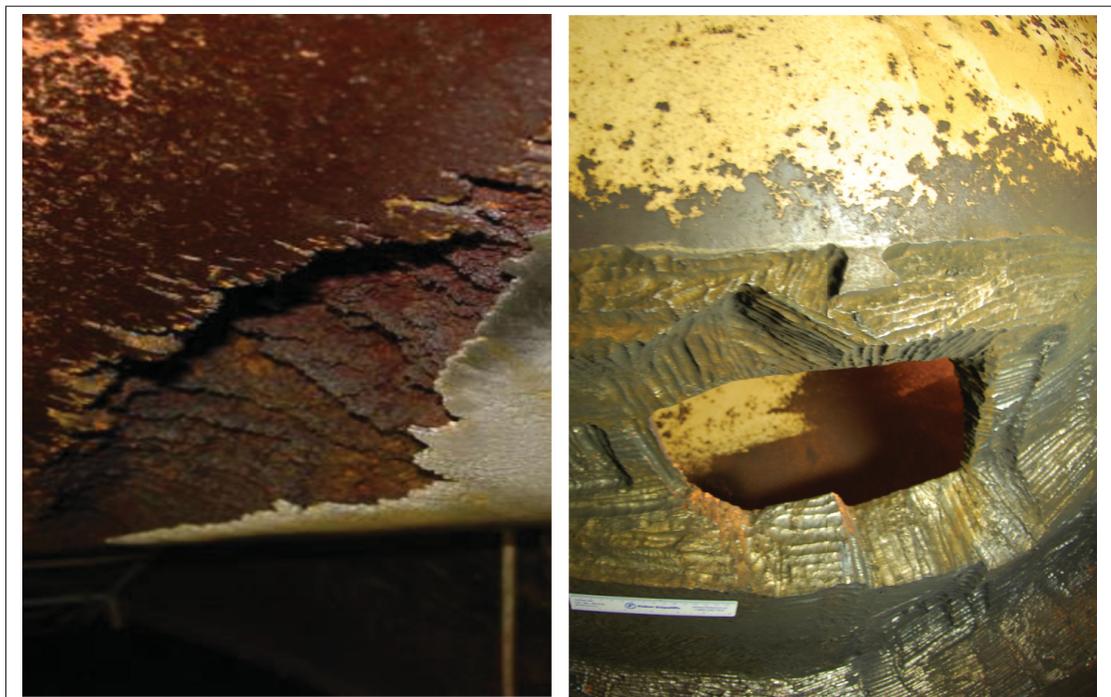


Figure 1.9 Exemples d'érosion observée sur les pales de turbine Francis.

faire apparaître de cavitation, les turbines sont de plus en plus exploitées en dehors des régions d'opération nominales (certaines raisons sont évoquées dans le Chapitre 4). De ce fait, même si certaines études s'intéressent à changer la géométrie des roues de turbines afin de les adapter à leur utilisation réelle (Celebioglu *et al.*, 2017 ; Yao, Xiao, Wang & Yang, 2015), la plupart cherchent à prédire les zones de manifestation de la cavitation et éventuellement leur intensité (Krumenacker, Fortes-Patella & Archer, 2014). Cependant, le phénomène, complexe, est à l'interface entre des écoulements instables et le comportement du matériau exposé. Cette combinaison rend difficiles les études en mécanique des fluides même pour des profils hydrodynamiques simples. Par ailleurs, les études récentes sur le sujet montrent encore des difficultés à prédire quantitativement l'intensité de la cavitation érosive (Biluš, Hočevár, Dular & Lešnik, 2020 ; Gohil & Saini, 2015), alors que qualitativement les résultats semblent prometteurs.

Ainsi, malgré la relative maturité des modèles en mécanique des fluides et des capacités de puissance de calcul accrues, la prédiction de l'érosion par cavitation à l'aide de modèles physiques n'est pas aboutie. Pour ces raisons, nous proposons d'étudier les modèles probabilistes basés sur les données de dégradation mesurées.

1.2 Modélisation par trajectoire de dégradation

Une des approches classiques pour modéliser un processus de dégradation (lorsque des données sont disponibles) repose sur l'hypothèse que le phénomène physique sous-jacent est monotone et déterministe, mais présente une variabilité d'un composant observé à un autre (Kahle, Mercier & Paroissin, 2016). Si on prend comme notation X_{ij} pour désigner le niveau de dégradation d'un composant i à l'instant t_{ij} , $i \in \{1, \dots, n\}$, avec n le nombre de composants et $j \in \{1, \dots, m_i\}$, avec m_i le nombre de mesures pour le composant i , alors on peut l'écrire de la façon suivante (Coble & Hines, 2011 ; Xu, Hong & Jin, 2016) :

$$X_{ij} = \mathcal{D}(t_{ij}, \beta_{1i}, \dots, \beta_{li}) + \epsilon_{ij}, \quad (1.5)$$

avec \mathcal{D} une fonction déterministe, ϵ_{ij} un résidu suivant une loi de Gauss $\mathcal{N}(0, \sigma_\epsilon^2)$ et β_{ki} , $k \in (1, \dots, l)$, des paramètres aléatoires. Ces derniers intègrent à la fois les effets fixes (comme les caractéristiques du design) sur la dégradation et les effets aléatoires propres à chaque composant i (tels que les variations dans la qualité de fabrication d'un matériel à l'autre). De ce fait, les paramètres β_{ki} peuvent différer d'un composant à l'autre ou être identiques (pour certains) pour tous les composants. Meeker & Escobar (2014) appellent ces modèles, les modèles à trajectoire de dégradation (*Degradation Path Models* ou *general path models* en anglais). Ces derniers sont semblables aux populaires modèles de régression, tels que présentés par McCulloch & Neuhaus (2005), au sens où la dégradation est décrite par une fonction issue de la physique sous-jacente et qu'il est possible d'intégrer de l'aléa au travers des termes ϵ_{ij} . Une fois que les paramètres β_{ki} sont déterminés (pour un composant donné), la fonction \mathcal{D} , tout comme le temps de défaillance associé (premier instant auquel la dégradation atteint un seuil ρ donné) deviennent alors déterministes (Ye & Xie, 2015).

Ce type de modèle est particulièrement adapté lorsqu'une description déterministe de la dégradation est disponible, basée par exemple sur des modèles physiques ou mécaniques. Ye & Xie (2015) liste les avantages des modèles par trajectoire et réalise une comparaison générale avec les processus stochastiques (qui seront présentés dans la suite de ce chapitre). On retient par exemple que les *degradation path models* présentent, en plus de leur simplicité, la possibilité d'intégrer facilement des effets aléatoires ou des covariables. Si on prend un modèle à trajectoire de dégradation dans sa forme la plus simple, c'est à dire une régression, la "vraie" trajectoire de dégradation peut être décrite par une fonction dépendant uniquement du temps. Ainsi, quelconque déviation de la dégradation observée par rapport à la "vraie" trajectoire de dégradation est imputée à des erreurs de mesure. Cette hypothèse de modélisation (simplificatrice) considère uniquement une partie de l'aléatoire alors qu'il peut y avoir plusieurs sources distinctes d'incertitude, parmi lesquelles : la variabilité intrinsèque à chaque composant (comme l'hétérogénéité à l'intérieur d'un matériau), la variabilité due à l'environnement (comme les fluctuations de conditions d'opération), et enfin, les erreurs de mesures, qui peuvent être insignifiantes lorsque les appareils utilisés présentent une grande précision. Ainsi, les modèles

de dégradation à trajectoire attribuent l'aléa aux erreurs de mesure à l'aide des ϵ_{ij} et ne peuvent prendre en compte les différentes sources de cet aléa. Ces modèles ne sont donc à privilégier que lorsque la variabilité due aux phénomènes environnementaux est faible ou lorsque ceux-ci peuvent être modélisés à l'aide de covariables par exemple (Ye & Xie, 2015). Si ces dernières conditions ne sont pas vérifiées, l'estimation de la durée de vie du composant peut alors être erronée. Dans ce cas, il convient donc de considérer d'autres modèles de dégradation.

Dans notre contexte, les modèles physiques de cavitation érosive ne sont pas assez matures pour être intégrés à ces modèles de trajectoire (au travers de la fonction \mathcal{D} de l'équation (1.5)). De plus, le phénomène de dégradation étudié présente une variabilité naturelle importante. Les modèles à trajectoire seront donc écartés au profit des processus stochastiques. Ils ont la particularité de modéliser l'ensemble d'une trajectoire (de dégradation par exemple) et de prendre en compte l'aspect temporel de l'aléa, tandis que les modèles précédemment introduits s'avèrent moins riches et complets pour traiter ces aspects. Par conséquent, les processus stochastiques sont particulièrement adaptés pour décrire l'évolution temporelle d'une variable aléatoire représentant la dégradation d'un matériel et sont présentés par la suite.

1.3 Processus stochastiques

Quelques processus stochastiques classiques, que l'on retrouve fréquemment dans le contexte de la fiabilité des matériels industriels, sont présentés dans cette section. Afin d'assurer une notation cohérente tout au long de la thèse, mais aussi de présenter les notions essentielles à la compréhension de cette section, quelques définitions sont d'abord introduites.

1.3.1 Définitions

Un processus stochastique temporel se définit comme une famille de variables aléatoires indexées par un paramètre continu ou discret, représentant le temps, dans un espace d'état D_X (Cinlar, 2013). On note $X_{1,n} = \{x_{t_1} \dots x_{t_n}\}$, $n \in \mathbb{N}$, le vecteur contenant n réalisations d'un phénomène aléatoire représenté par un processus $(X_t)_{t \geq 0}$ aux instants t_1, \dots, t_n . Dans le cas des processus à

temps continu (voir paragraphe suivant), le processus stochastique sera noté $X(t)$ à la place de X_t .

Espaces de temps et d'état

Parmi les processus stochastiques, on peut en distinguer 4 types, classés en fonction de leurs espaces de temps et d'état, dont des exemples sont donnés à titre indicatif dans le Tableau 1.1.

Tableau 1.1 Types de processus.

Type	Exemple
Temps discret, espace d'état discret	La réussite d'un golfeur à atteindre un nouveau trou en un coup. Les chaînes de Markov en font partie.
Temps discret, espace d'état continu	L'évolution d'une moyenne mobile pour chaque nouvelle observation.
Temps continu, espace d'état discret	Le nombre de visiteurs d'un musée à chaque instant.
Temps continu, espace d'état continu	L'évolution de la température à chaque instant.

Dans le contexte particulier de la dégradation, on se place généralement dans la famille de processus à espaces d'état et de temps continus bien que les observations soient réalisées en temps discrets.

Processus de Markov

Un processus est dit de Markov si, connaissant $X(t)$, la réalisation de $X(\tau)$, avec $\tau > t$, est indépendante des valeurs antérieures $X(u)$, $u < t$ (sans mémoire) (Barlow & Proschan, 1996).

Dans un espace d'état discret, cette propriété se traduit par :

$$\mathbb{P}(X(\tau) = x | X(u), 0 \leq u \leq t) = \mathbb{P}(X(\tau) = x | X(t)) \quad (1.6)$$

avec $\tau > t$ et $x \in D_X$.

Dans le cas d'un espace de temps continu, la probabilité d'obtenir une valeur donnée x , $\mathbb{P}(X(\tau) = x)$ est généralement nulle. Les probabilités d'apparition de l'équation précédente sont alors remplacées par des noyaux de transition de y à x , notés $p(x|y)$ et la propriété markovienne devient :

$$p(x_n|x_{n-1}, x_{n-2}, \dots, x_1, x_0) = p(x_n|x_{n-1}) \quad (1.7)$$

Cette dernière expression traduit que la réalisation du processus à un instant t_i dépend uniquement de la réalisation à l'instant précédent t_{i-1} et que les réalisations antérieures n'ont pas d'influence sur celle-ci.

Processus de Lévy

Un processus stochastique $\{X(t)\}_{t \geq 0}$ à temps continu est appelé processus de Lévy s'il vérifie les propriétés suivantes (Applebaum, 2009) :

- $X(0) = 0$,
- $\forall 0 \leq t_1 \leq \dots \leq t_n < \infty$, les incréments $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$, sont indépendants,
- $\forall s < t$, $X(t) - X(s)$ est égal en loi à $X(t - s)$: les incréments sont stationnaires,
- $t \mapsto X(t)$ est continue à droite, limite à gauche.

On rappelle qu'un processus stochastique possède des incréments stationnaires si la distribution de probabilité des incréments $X(t+h) - X(t)$ dépend uniquement de h , la longueur de l'incrément de temps, $\forall t, h \geq 0$.

1.3.2 Processus de Poisson

Le processus de Poisson, qui doit son nom au mathématicien français, est un processus de Lévy (dans sa forme homogène) à sauts (Ross, Kelly, Sullivan, Perry, Mercer, Davis, Washburn,

Sager, Boyce & Bristow, 1996). Noté $(N_t)_{t \geq 0}$, ce processus de comptage présente les propriétés suivantes :

- $\forall 0 \leq t_1 \leq \dots \leq t_n < \infty$, les incréments $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$, sont indépendants,
- $N_\tau - N_t \sim \mathcal{P}(\Lambda(\tau) - \Lambda(t))$, $\forall 0 \leq t \leq \tau$: les incréments suivent une distribution de Poisson de taux cumulé $\Lambda(t)$,
- le processus est stationnaire si $\Lambda(t) = \lambda t$, $\forall t \geq 0$, $\lambda > 0$.

C'est néanmoins sous la forme de processus de Poisson composé (*compound Poisson*) qu'il est le plus souvent utilisé pour modéliser la dégradation (Lawless & Crowder, 2010; Singpurwalla & Wilson, 1998). Celui-ci peut être décrit par l'expression suivante :

$$C(t) = \sum_{i=1}^{N_t} D_i \quad (1.8)$$

avec N_t un processus de Poisson homogène de taux cumulé Λ et D_i des variables aléatoires et indépendantes, aussi appelées tailles de saut. Caballé, Castro, Pérez & Lanza-Gutiérrez (2015) présentent une étude théorique sur la défaillance d'un matériel sujet à deux mécanismes de dégradation : une première dégradation progressive et une seconde par chocs. Un processus de Poisson composé (non-homogène) est utilisé pour modéliser les instants d'initiation de la dégradation progressive. Les incréments suivent dans cette étude une loi gamma (voir section 1.3.5).

La Figure 1.10 présente une simulation de trajectoires d'un processus de Poisson homogène pour différentes valeurs du paramètre λ .

1.3.3 Processus de Wiener et de diffusion

Le mouvement brownien standard, aussi connu sous le nom de processus Wiener sans dérive, doit son nom au botaniste Brown, le premier à décrire le mouvement de diffusion aléatoire

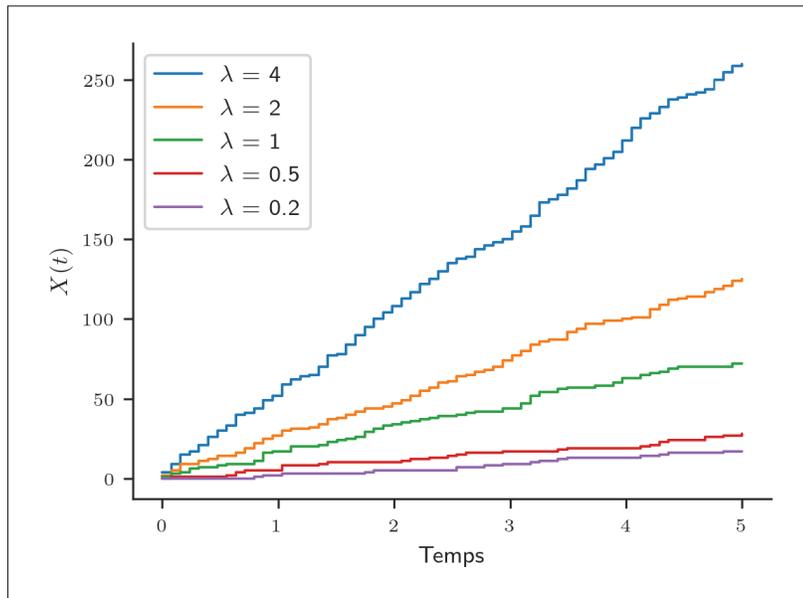


Figure 1.10 Exemple de trajectoires simulées pour un processus de Poisson.

et incessant de particules plongées dans un liquide (Karlin, 2014). C'est par la suite que le mathématicien Wiener proposa un modèle abouti permettant de décrire ce phénomène physique.

Le processus de Wiener standard, noté $W(t)$, est un processus stochastique markovien et demeure parmi les plus utilisés dans la famille de Lévy. Il présente en particulier les propriétés suivantes :

- $W(0) = 0$,
- $W(\tau) - W(t) \sim \mathcal{N}(0, \tau - t), \forall \tau > t \geq 0$.

Le mouvement brownien (processus de Wiener) est un cas particulier de processus de diffusion (Si, Wang, Hu, Zhou & Pecht, 2012). Un processus de Markov est un processus de diffusion s'il est solution de l'équation différentielle stochastique suivante :

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \quad (1.9)$$

avec $\mu(t, X(t))$ le terme de tendance moyenne (ou *drift*) et $\sigma(t, X(t))$ le terme de diffusion (ou de volatilité) qui correspond à l'écart-type instantané (Poznyak, 2009).

C'est cependant dans sa forme générale, appelée processus de Wiener avec dérive (linéaire), qu'on le retrouve le plus souvent pour modéliser la dégradation de composants (Jin & Matthews, 2014; Si, Wang, Hu, Chen & Zhou, 2013; Wang, Jiang, Guo & Cheng, 2014). Cette forme généralisée présente alors les propriétés suivantes :

- $W(0) = 0$,
- $W(\tau) - W(t) \sim \mathcal{N}(\eta(\tau - t), \sigma^2(\tau - t))$, $\mathbb{E}(W(t)) = \eta t$ et $\mathbb{V}(W(t)) = \sigma^2 t$, $\forall \tau > t \geq 0$ et $\sigma > 0$.

Si on prend $\eta > 0$, le processus de Wiener avec dérive (linéaire) impose alors une tendance moyenne croissante de la dégradation au cours du temps, contrairement à sa version standard. Le processus de Wiener avec dérive η (constante) et coefficient de diffusion σ se note alors :

$$X(t) = \eta t + \sigma W(t), \quad (1.10)$$

avec $X(t)$ le niveau de dégradation à l'instant t et $W(t)$ le processus de Wiener standard. La Figure 1.11 présente une simulation de trajectoires d'un processus de Wiener avec dérive pour différentes valeurs des paramètres μ et σ .

En particulier, on peut souligner qu'en raison de ses incréments qui suivent une loi normale, le processus de Wiener peut présenter entre deux instants successifs des incréments positifs ou négatifs. Dans un contexte de dégradation, ce type de modèle peut s'avérer efficace pour décrire le comportement d'un composant présentant une capacité d'auto-réparation ou encore étant exposé à des niveaux de charge fluctuants (résultant en la fermeture d'une fissure par exemple (Zhang, Si, Hu & Lei, 2018)). Cette modélisation est aussi adaptée en présence d'incertitudes de mesures ou de signal bruité, pouvant entraîner des incréments de dégradation observés négatifs entre deux instants. On notera néanmoins qu'il est possible de modéliser le phénomène de dégradation et le bruit de mesure des observations séparément (Le Son, Fouladirad & Barros, 2016).

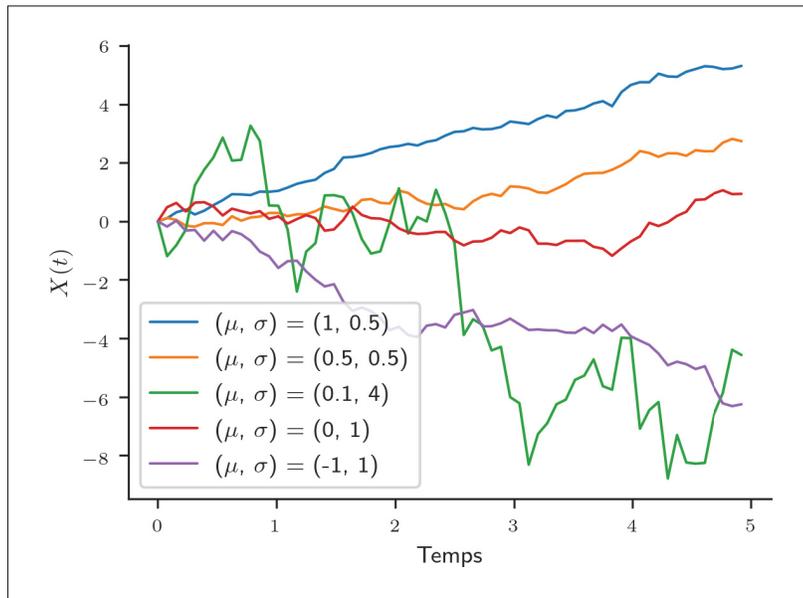


Figure 1.11 Exemple de trajectoires simulées pour un processus de Wiener avec dérive.

Une limite des processus stationnaires ¹ (dont l'accroissement sur un intervalle de temps ne dépend que de la durée entre les deux instants d'observation) réside dans le fait que la vitesse moyenne de dégradation est constante au cours du temps. Cette hypothèse s'avère souvent contraignante dans le cas de dynamiques de dégradation qui évoluent au cours du temps. Pour pallier cet inconvénient, il est possible d'introduire au processus de Wiener une dérive non-linéaire, qui reflète par exemple une accélération (ou un ralentissement) de la dégradation dans le temps (Huang, Xu, Wang & Sun, 2015). Dans ce cas, le processus n'est alors plus considéré comme stationnaire dans la mesure où la loi de probabilité de ses incréments de dégradation ne dépend plus uniquement de la largeur de l'intervalle τ , mais plutôt des instants formant cet incrément de temps. Li, Wang, Zhang, Liu & Fu (2018) estiment la durée de vie résiduelle des LED à partir de l'intensité lumineuse de celles-ci à différents instants. Le comportement de cet indicateur n'est pas monotone, mais présente une tendance moyenne décroissante, d'où l'emploi d'un processus de Wiener avec dérive non-linéaire pour le modéliser.

¹ Cette limite est valable pour tous les processus stochastiques stationnaires présentés dans ce chapitre.

On retrouve aussi fréquemment dans la littérature le processus d'Ornstein-Uhlenbeck comme processus de diffusion. Il peut convenir à la modélisation de composants dont la dégradation évolue et se stabilise autour d'un point d'équilibre (Nguyen, Fouladirad & Grall, 2018). Ce processus est décrit par l'équation différentielle stochastique suivante :

$$dX(t) = \theta(\mu - X(t))dt + \sigma dW(t), \quad (1.11)$$

où $\theta > 0$, $\sigma > 0$ et $\mu \in \mathbb{R}$ sont des paramètres déterministes et $W(t)$ est un processus de Wiener. Deng, Barros & Grall (2015) justifient l'emploi de ce type de processus par la possibilité de construire un modèle dont la tendance moyenne et la variance peuvent être choisies librement. Un cas d'étude est présenté pour modéliser la longueur de fissures d'un composant de centrale électrique dont les données comportent un grand nombre de trajectoires avec peu d'observations.

La Figure 1.12 présente une simulation de trajectoires d'un processus d'Ornstein-Uhlenbeck pour différentes valeurs de paramètres.

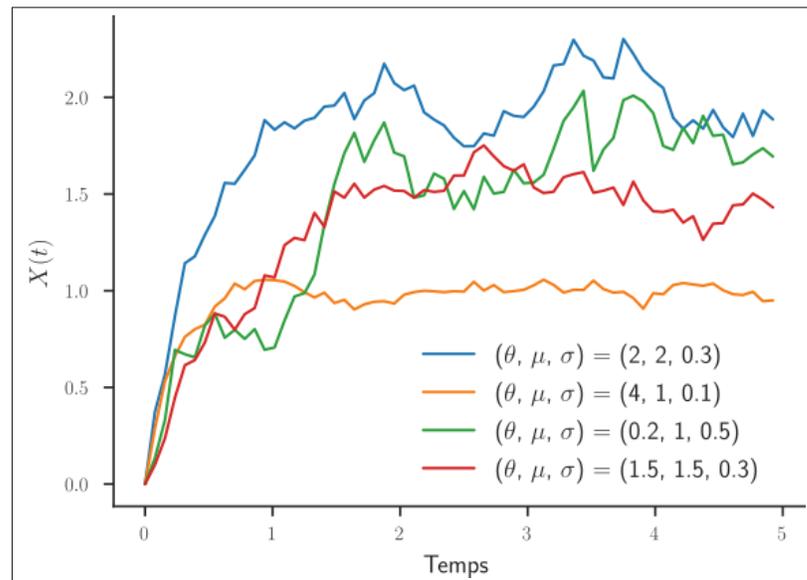


Figure 1.12 Exemple de trajectoires simulées pour un processus d'Ornstein-Uhlenbeck.

1.3.4 Processus inverse gaussien

Wasan (1968) fut le premier à introduire le processus inverse gaussien dans la littérature. Ce processus de Lévy (dans sa forme homogène) présente des incréments strictement positifs et permet donc la modélisation de dégradation cumulative et monotone. Pour ces raisons, il est d'ailleurs souvent comparé au processus gamma (voir section suivante). De plus, ces deux processus sont considérés comme des limites d'un processus de Poisson composé (Ye & Chen, 2014).

En particulier, le processus stochastique inverse gaussien $(X(t))_{t \geq 0}$ satisfait les propriétés suivantes (Ye, Chen, Tang & Xie, 2014a) :

- $X(0) = 0$,
- $X(\tau) - X(t) \sim \mathcal{IG}(\Lambda(\tau) - \Lambda(t), \eta[\Lambda(\tau) - \Lambda(t)]^2), \forall \tau > t \geq 0$,
- $X(t)$ a des incréments indépendants et distincts,

avec $\Lambda(t)$ une fonction monotone croissante et $\mathcal{IG}(\mu, \lambda)$, $\mu > 0, \lambda > 0$, la distribution inverse gaussienne dont la fonction de densité (PDF) est donnée par :

$$f(x|\mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right], x > 0. \quad (1.12)$$

On notera que μ et λ désignent respectivement la moyenne et le paramètre de forme de la distribution inverse gaussienne.

Bien que largement étudié (Chhikara, 1988 ; Wang & Xu, 2010 ; Ye & Chen, 2014 ; Ye *et al.*, 2014a), le processus inverse gaussien a reçu un accueil mitigé pour la modélisation de dégradation (Ye & Chen, 2014). Ce constat est en partie attribué au fait que ce processus ne répond pas clairement aux attentes des ingénieurs fiabilistes sur les questions d'explicabilité physique du modèle (Ye & Xie, 2015). Néanmoins, Wang & Xu (2010) utilisent ce type de processus pour modéliser la dégradation d'un laser sujet à une augmentation de courant au fil du temps, entraînant la défaillance du composant lorsque le courant dépasse un seuil donné.

La Figure 1.13 présente une simulation de trajectoires issues d'un processus inverse gaussien pour différentes valeurs de paramètres.

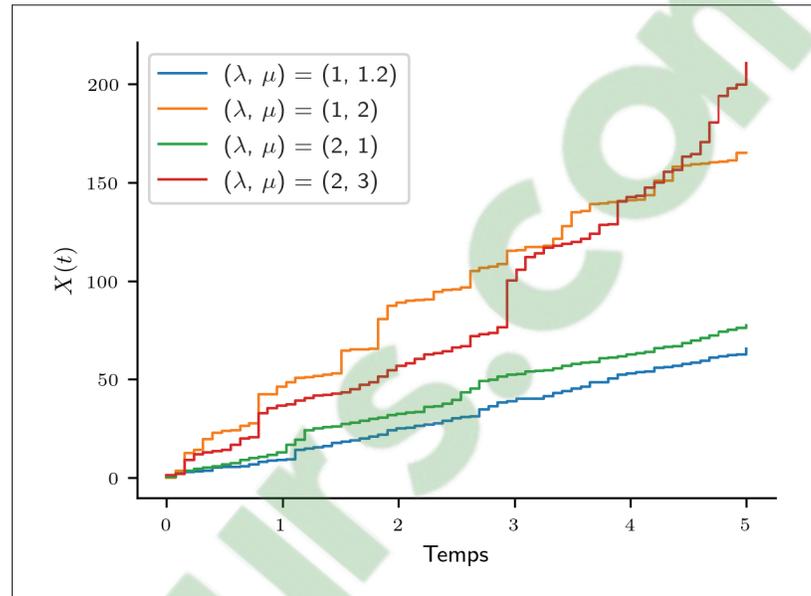


Figure 1.13 Exemple de trajectoires simulées pour un processus inverse gaussien.

1.3.5 Processus gamma

Le processus gamma (PG) est un processus croissant présentant un nombre infini de sauts sur un intervalle de temps fini. Appartenant à la famille des processus de Lévy (dans sa forme homogène), il est particulièrement adapté pour la modélisation des phénomènes dont la dégradation est strictement croissante et s'accumule sous forme de petits incréments (Kahle *et al.*, 2016). Ce processus est ainsi souvent considéré avec le processus inverse gaussien, comme les seuls capables de modéliser une dégradation strictement croissante (Wang & Xu, 2010; Ye & Chen, 2014; Zhang *et al.*, 2018). Le processus gamma est néanmoins souvent privilégié dans le contexte de la fiabilité de matériels industriels en raison de son rattachement au sens physique. Celui-ci sera privilégié par la suite en raison des nombreux résultats liés aux structures utilisant ce processus.

Abdel-Hameed (1975) fut le premier à utiliser le processus gamma pour modéliser une dégradation se manifestant de façon aléatoire dans le temps. Van Noortwijk (2009) et Alaswad & Xiang (2017), au travers de leur revue de littérature respective, soulignent la capacité du processus gamma à modéliser les dommages graduels et strictement cumulatifs, tels que la corrosion, la fatigue, les glissements, l'érosion, *etc.* Van Noortwijk (2009) décrit le processus gamma comme un processus stochastique avec des incréments indépendants, non-négatifs, ayant une distribution gamma dont le paramètre d'échelle est constant.

Afin de définir les bases mathématiques du processus gamma, introduisons la variable aléatoire positive X de loi gamma. Cette dernière possède une distribution gamma, notée $Ga(u, v)$, avec $u > 0, v > 0$ le paramètre de taux et le paramètre de forme respectivement, si sa PDF s'écrit de la façon suivante :

$$f_X(x|u, v) = Ga(x|u, v) = \frac{u^v}{\Gamma(v)} x^{v-1} e^{-ux} I_{\mathbb{R}_+^*}(x), \quad (1.13)$$

avec $I_A(x) = 1, \forall x \in A$ et $I_A(x) = 0, \forall x \notin A$, et $\Gamma(a)$ est la fonction gamma incomplète $\forall a > 0$ (Arfken, 1985). Sa fonction de répartition (CDF) est donnée par l'expression suivante :

$$\mathbb{F}_X(x|u, v) = \frac{\gamma(v, ux)}{\Gamma(v)}, \quad (1.14)$$

avec $\gamma(v, ux)$ la fonction gamma incomplète inférieure (Chaudhry & Zubair, 1994). On trouve souvent dans la littérature une autre paramétrisation classique de la distribution gamma : (α, β) avec $\beta = 1/u$ et $\alpha = v$. Dans ce cas, β est appelé paramètre d'échelle. Par abus de langage, ces deux paramétrisations sont souvent confondues.

Le processus gamma $X(t)$ avec pour paramètre de taux $u > 0$ et pour fonction de forme $v(t) > 0$ possède les propriétés suivantes :

- $X(0) = 0,$
- $X(\tau) - X(t) \sim Ga(u, v(\tau) - v(t)), \forall \tau > t > 0,$

- $\forall 0 \leq t_1 \leq \dots \leq t_n < \infty$, les incréments $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$, sont indépendants.

v doit être une fonction strictement croissante, continue à droite, à valeurs réelles dans \mathbb{R}_+ avec $v(0) = 0$. Dès lors que la fonction de forme v dépend du temps, le processus n'est plus considéré comme stationnaire (ou homogène) en raison des incréments qui ne sont plus identiquement distribués.

L'espérance et la variance de $X(t)$ avec pour paramètres u et $v(t)$ sont données par :

$$\mathbb{E}(X(t)) = \frac{v(t)}{u}, \quad \mathbb{V}(X(t)) = \frac{v(t)}{u^2} \quad (1.15)$$

La fonction de forme $v(t)$ reflète la dynamique moyenne (à une constante près) de la dégradation en fonction du temps. Un des choix classiques pour cette fonction est $v(t) = ct^b$ avec $c > 0, b > 0$ (Kahle *et al.*, 2016). En effet, les études empiriques ou les connaissances techniques montrent que l'espérance d'une dégradation en fonction du temps est souvent proportionnelle à une fonction puissance (Van Noortwijk, 2009). D'un point de vue pratique, les paramètres c et b influencent respectivement l'intensité d'un incrément de dégradation et l'évolution de la vitesse moyenne de la dégradation. Lorsque $b = 1$, le processus gamma est homogène (PGH, ou stationnaire); pour $b \neq 1$, le processus gamma est non-homogène (PGNH, ou non-stationnaire). Dans leur étude, Park & Padgett (2005) identifient une valeur appropriée pour b basée sur le critère d'information d'Akaike (AIC, voir section 1.4.1), qui est un indicateur statistique classique pour la sélection de modèles. Plusieurs études modélisant une dégradation ont été réalisées lorsque b est *a priori* connu : celle du béton due à la corrosion de sa matrice de renforcement ($b = 1$) (Cinlar, Bazant & Osman, 1977) ou pour des cas de fluages ($b = 1/8$) (Ellingwood & Mori, 1993). Un autre choix classique pour la fonction de forme repose sur la fonction exponentielle : $v(t) = e^{-ct} - 1$, avec $c > 0$.

Un des avantages de l'utilisation du processus gamma dans la modélisation de la dégradation, repose sur la possibilité de déduire mathématiquement un certain nombre d'indicateurs d'intérêt

(cette remarque est aussi valable dans une certaine mesure pour les processus de Wiener et inverse gaussien). Parmi ces indicateurs, un particulièrement intéressant est la durée de vie résiduelle à une certaine date (souvent rencontrée comme *remaining useful lifetime (RUL)* dans la littérature) connaissant le niveau de dégradation du composant étudié (cf. section 1.3.5.4).

On notera que différentes variantes du processus gamma sont présentes, notamment en fiabilité (Van Noortwijk, 2009). Parmi celles-ci, on retrouve le processus gamma étendu (*extended gamma process* (Çinlar, 1980)), qui permet de s'affranchir du ratio constant variance - espérance ($\frac{\mathbb{V}(X(t))}{\mathbb{E}(X(t))} = u$) au cours du temps. Al Masry, Mercier & Verdier (2017) présentent une comparaison du processus gamma étendu avec le processus gamma standard (présenté plus haut) dans un contexte d'optimisation de coût de maintenance. Un processus gamma transformé (*transformed gamma process*) est utilisé par Giorgio, Guida & Pulcini (2016) pour modéliser la dégradation de LED au cours du temps. Cette variante du processus gamma permet notamment de modéliser facilement une dépendance des incréments de dégradation sur des intervalles de temps disjoints.

La Figure 1.14 présente une simulation de trajectoires d'un processus gamma pour différentes valeurs de paramètres dont la fonction de forme est du type $v(t) = ct^b$. En particulier, la trajectoire verte représente une réalisation d'un processus gamma non homogène tandis que les autres sont issues de processus gamma homogènes.

Afin de mieux appréhender l'influence des paramètres d'un processus gamma sur les incréments de dégradation, plusieurs faisceaux de trajectoires ont été simulés à partir de différents jeux de paramètres. La méthode de simulation employée repose sur la technique de *Gamma Sequential Sampling* (voir la section 1.3.5.3) et les paramètres sont résumés dans le Tableau 1.2.

Les trajectoires simulées ainsi que la distribution empirique des incréments de dégradation associés $\delta_{i,j}$, $1 \leq i \leq n_j$, $1 \leq j \leq m$, pour des processus gamma homogènes et gamma non homogènes, sont représentées respectivement par les Figures 1.15 et 1.16. Afin de présenter de façon lisible les différentes figures, ces dernières utilisent des échelles sur l'axe des abscisses différentes. La Figure 1.15 présente des trajectoires de dégradation cumulée issues d'un processus gamma homogène. On observe que les distributions empiriques de la variable aléatoire modélisant

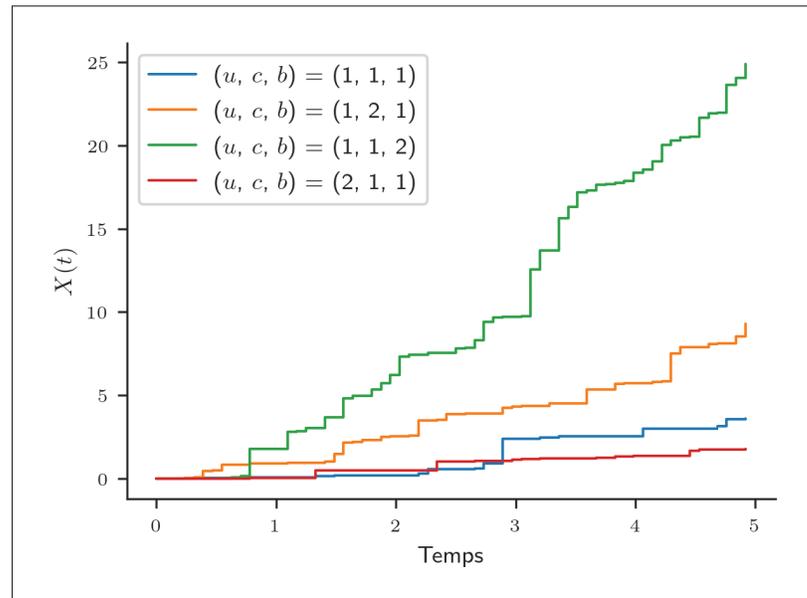


Figure 1.14 Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène.

Tableau 1.2 Plan d'expérience.

Paramètres	Valeurs
Paramètres PG $(u; c; b)$	$(1; 1; 1)$
	$(1; 2; 1)$
	$(2; 1; 1)$
	$(1; 1; 2)$
Longueur de trajectoire n_j	150
Nombre de trajectoires m	50

les incréments de dégradation sont semblables aux distributions classiques d'une loi gamma. Lorsqu'on s'intéresse maintenant à la Figure 1.16, la fonction de forme $v(t) = ct^b$ prend dans ce cas pour paramètre $b = 2$ (cas non homogène). On notera que pour cette figure, il a été choisi de représenter de deux façons différentes la distribution des incréments de dégradation issus d'un même faisceau de trajectoires (pour des paramètres de processus gamma donnés). L'histogramme (a), représente la distribution empirique des incréments de dégradation $\delta_{i,j}$

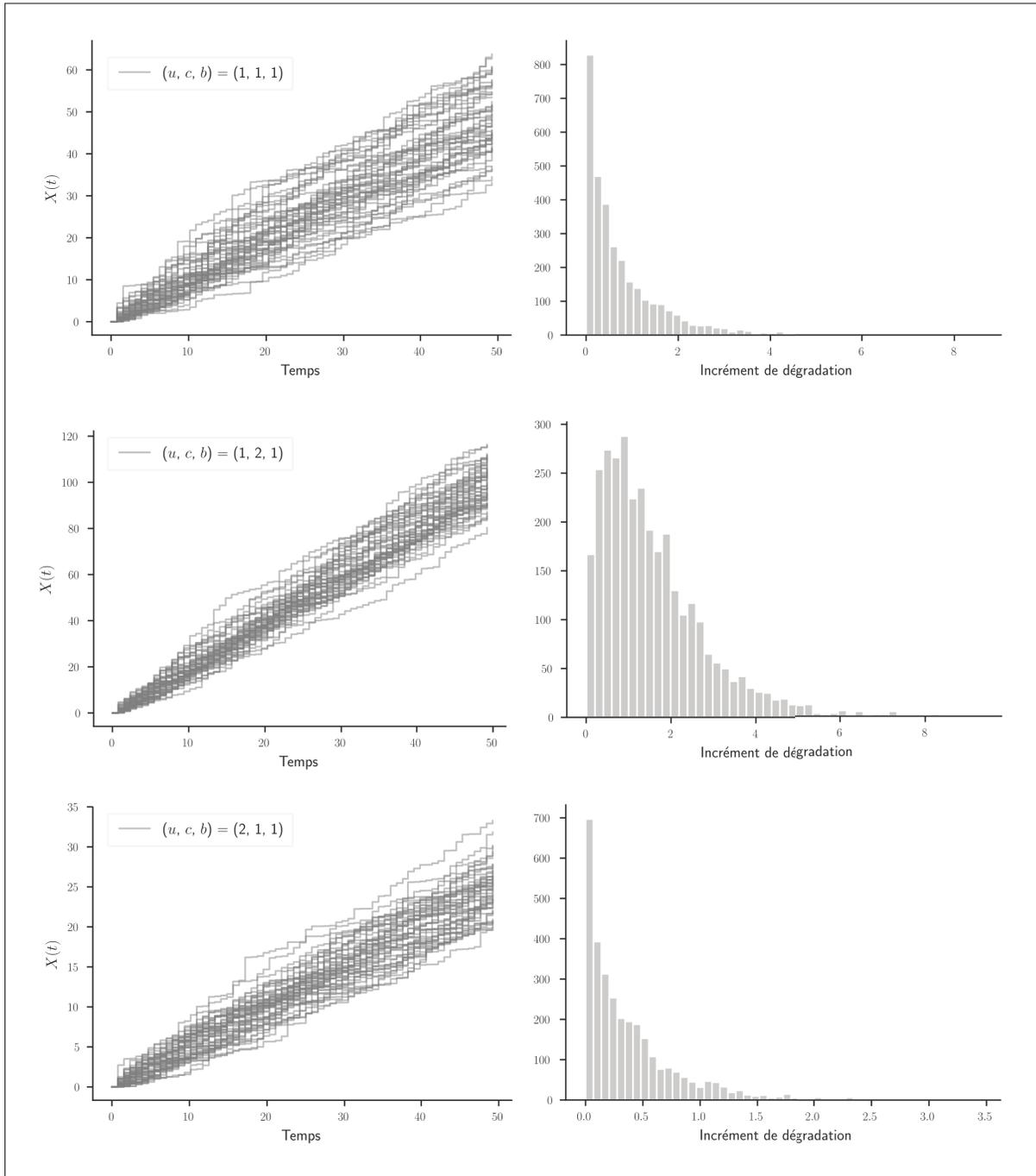


Figure 1.15 Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène.

de l'ensemble des trajectoires m et pour la totalité du temps de simulation T (qui peut se traduire par $1 \leq i \leq n_j$). En revanche, les histogrammes **(b)** représentent les distributions

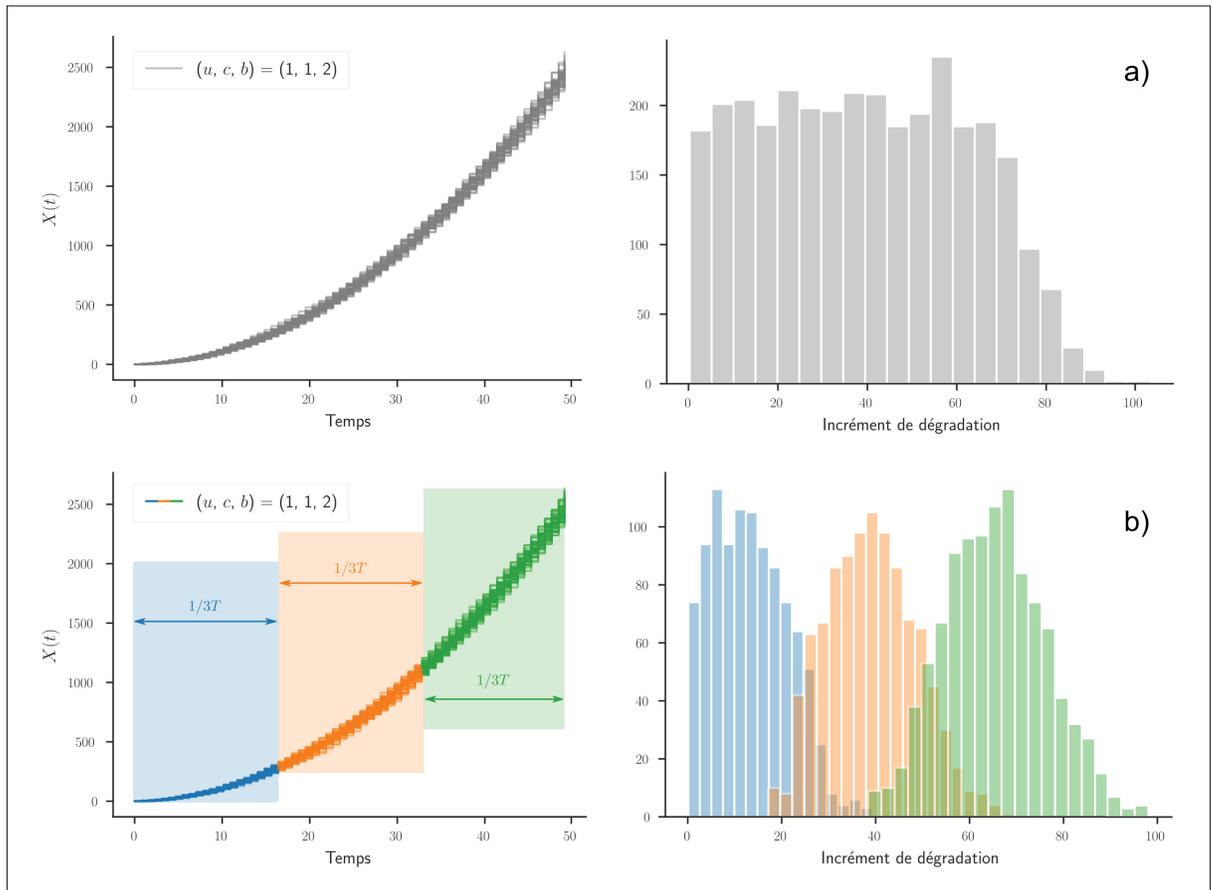


Figure 1.16 Exemple de trajectoires simulées pour un processus gamma homogène et non-homogène.

empiriques des incréments de dégradation sur les intervalles de temps $[0; 1/3T]$ ($1 \leq i \leq 1/3n_j$), $[1/3T; 2/3T]$ ($1/3n_j \leq i \leq 2/3n_j$) et $[2/3T; T]$ ($2/3n_j \leq i \leq n_j$) respectivement. Il est alors possible d'observer la non stationnarité du processus sur les incréments de dégradation : plus les incréments $\delta_{i,j}$ sont pris à des instants t_i importants, plus la tendance centrale des histogrammes augmente (c.à.d. que les incréments de dégradation augmentent en moyenne en fonction du temps). On soulignera aussi la forte influence du paramètre b comparativement aux autres paramètres.

1.3.5.1 Processus gamma avec covariables comme modèle de dégradation

Comme c'est le cas pour les autres processus stochastiques présentés dans les sections précédentes, le processus gamma peut intégrer des covariables. Ces dernières désignent généralement des variables continues dont les mesures sont prises en même temps que la variable principale reflétant le niveau de dégradation. Pan & Balakrishnan (2011) ont recours à un processus gamma intégrant des covariables (ici un modèle bivarié) pour modéliser la dégradation et estimer la durée de vie d'un composant sujet à l'évolution de 2 fissures dépendantes l'une de l'autre. Ils utilisent par ailleurs l'approximation de Birnbaum-Saunders, présentée en section 1.3.5.4, pour estimer la durée de vie de ce composant. D'autres exemples de processus stochastiques intégrant des covariables sont présentés dans la littérature : Duan & Wang (2018) utilisent un processus global (appelé processus de dispersion exponentielle), incluant les processus de Wiener, inverse gaussien et gamma comme des cas particuliers. Les covariables introduites dans la modélisation reflètent l'influence de l'environnement extérieur sur l'évolution de la dégradation de lasers au cours du temps. D'autres exemples de modélisations impliquant un processus gamma avec des covariables sont présentés dans (Deloux, Fouladirad & Bérenguer, 2016 ; Lawless & Crowder, 2004 ; Park & Padgett, 2005 ; Zhang *et al.*, 2018).

1.3.5.2 Estimation des paramètres

Afin de modéliser l'évolution de la dégradation d'un composant par un processus gamma à partir de données de dégradation mesurées, le recours à des méthodes statistiques d'estimation des paramètres est nécessaire. Cinlar *et al.* (1977) furent les premiers à proposer les deux méthodes d'estimation les plus rencontrées dans la littérature, à savoir l'estimation par maximum de vraisemblance (MV) et la méthode des moments (MM) (Van Noortwijk, 2009). Cette dernière est particulièrement adaptée lorsque la valeur de b , pour une fonction de forme de type loi puissance par exemple, est connue à l'avance (comme la valeur issue d'un modèle physique). Ainsi, le processus gamma non homogène peut être ramené à un processus gamma homogène à l'aide d'une transformation (Van Noortwijk, 2009) lorsque les incréments de temps sont tous identiques. Ceci a pour avantage de permettre une estimation des paramètres plus rapide et

facile, en raison du caractère explicite des expressions. En revanche, lorsque la vitesse de la dégradation n'est pas connue à l'avance, le recours à la méthode MV est préféré. C'est d'ailleurs le cas pour les données de dégradation étudiées dans les chapitres suivants dont la valeur de b est inconnue dans la fonction de forme $v(t) = ct^b$. Pour cette raison, nous avons décidé de présenter en détail cette méthode d'estimation.

Supposons qu'on observe m trajectoires indépendantes d'un même processus stochastique. Soit n_j le nombre d'observations successives de la trajectoire de dégradation j , $1 \leq j \leq m$. Alors, la $j^{\text{ième}}$ trajectoire est observée $n_j + 1$ fois aux instants $t_{0,j} < t_{1,j} < \dots < t_{n_j,j}$. La notation suivante est adoptée pour tout $1 \leq j \leq m$ (voir Figure 1.17) :

- $t_{0,j} = t_0 = 0$;
- $x_{0,j} = 0$, ce qui signifie que pour t_0 il n'y a pas dégradation ;
- $\delta_{i,j} = x_{i+1,j} - x_{i,j}$: l'incrément de dégradation entre deux instants successifs $t_{i,j}$ et $t_{i+1,j}$;
- $\mathbf{D}_{obs} = \cup_{j=1}^m \cup_{i=0}^{n_j} \{(t_{i,j}, x_{i,j})\} = \cup_{j=1}^m \cup_{i=0}^{n_j-1} \{(t_{i,j}, t_{i+1,j}, \delta_{i,j})\}$: le jeu de données de dégradation observé organisé en m trajectoires indépendantes de n_j observations, $1 \leq j \leq m$.

La méthode MV consiste à inférer les valeurs des paramètres, notés $\theta = (u, c, b)$, d'un processus à partir d'un jeu de données \mathbf{D}_{obs} , en maximisant la fonction de vraisemblance \mathcal{L} . La fonction de vraisemblance \mathcal{L} d'un échantillon issu d'un processus gamma est donnée dans l'équation (1.16) (voir Van Noortwijk (2009)). Pour des raisons pratiques, on choisit plutôt de maximiser la fonction log-vraisemblance ℓ exprimée dans l'équation (1.17).

$$\begin{aligned} \mathcal{L}(\theta|\mathbf{D}_{obs}) &= \mathcal{L}(u, c, b|\mathbf{D}_{obs}) = \prod_{j=1}^m \prod_{i=1}^{n_j} Ga(\delta_{i,j}|u, v(t_{i,j}) - v(t_{i-1,j})) \\ &= \prod_{j=1}^m \prod_{i=1}^{n_j} \frac{u^{c[t_{i,j}^b - t_{i-1,j}^b]}}{\Gamma(c[t_{i,j}^b - t_{i-1,j}^b])} \delta_{i,j}^{c[t_{i,j}^b - t_{i-1,j}^b] - 1} e^{-u\delta_{i,j}} \end{aligned} \quad (1.16)$$

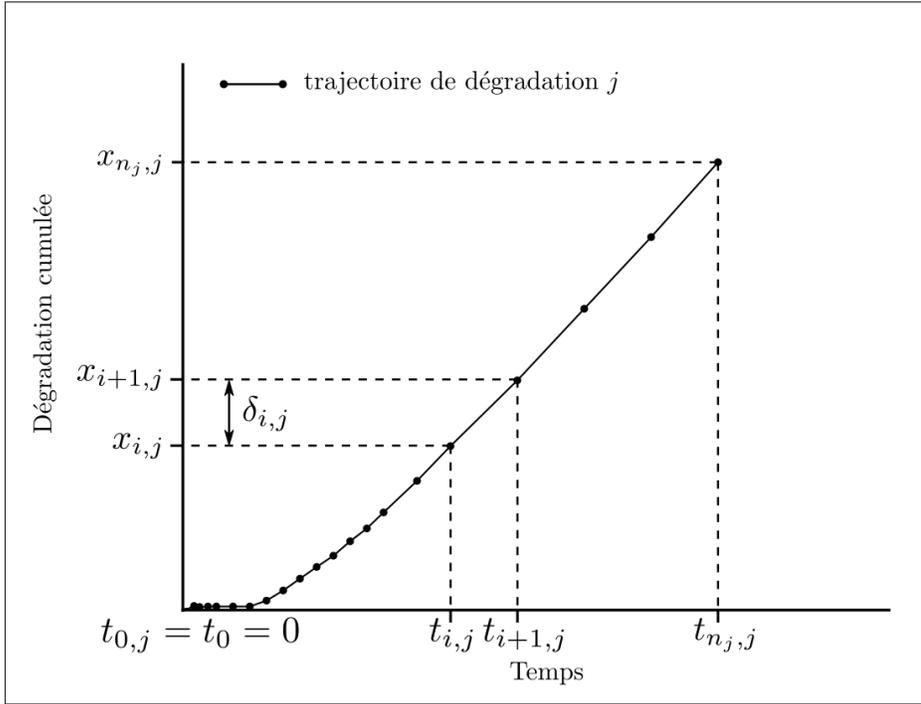


Figure 1.17 Exemple de notation des données.

$$\begin{aligned}
 \ell(\boldsymbol{\theta}|\mathbf{D}_{obs}) &= \log(\mathcal{L}(u, c, b|\mathbf{D}_{obs})) \\
 &= \sum_{j=1}^m \sum_{i=1}^{n_j} c[t_{i,j}^b - t_{i-1,j}^b] \log(u) - \log\left[\Gamma\left(c[t_{i,j}^b - t_{i-1,j}^b]\right)\right] \\
 &\quad + \left(c[t_{i,j}^b - t_{i-1,j}^b] - 1\right) \log(\delta_{i,j}) - u\delta_{i,j}
 \end{aligned} \tag{1.17}$$

En dérivant partiellement ℓ par rapport aux trois paramètres, l'estimateur du maximum de vraisemblance (EMV) $\hat{\boldsymbol{\theta}}_{MV} = (\hat{u}_{MV}, \hat{c}_{MV}, \hat{b}_{MV})$ de $\boldsymbol{\theta} = (u, c, b)$ consiste en la solution du système à trois équations non-linéaires suivant :

$$\hat{u}_{MV} = \hat{c}_{MV} \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} \hat{t}_{i,j}^{\hat{b}_{MV}} - \hat{t}_{i-1,j}^{\hat{b}_{MV}}}{\sum_{j=1}^m \sum_{i=1}^{n_j} \delta_{i,j}} = \hat{c}_{MV} \frac{\sum_{j=1}^m \hat{t}_{n_j,j}^{\hat{b}_{MV}}}{\sum_{j=1}^m x_{n_j,j}}, \tag{1.18}$$

$$\sum_{j=1}^m \sum_{i=1}^{n_j} \left[t_{i,j}^{\hat{b}_{MV}} - t_{i-1,j}^{\hat{b}_{MV}} \right] \left\{ \psi \left(\hat{c}_{MV} \left[t_{i,j}^{\hat{b}_{MV}} - t_{i-1,j}^{\hat{b}_{MV}} \right] \right) - \log(\delta_{i,j}) \right\} = \sum_{j=1}^m (t_{n_j,j}^{\hat{b}_{MV}}) \log(\hat{u}_{MV}), \quad (1.19)$$

$$\begin{aligned} & \sum_{j=1}^m \sum_{i=1}^{n_j} \hat{c}_{MV} \left[t_{i,j}^{\hat{b}_{MV}} \log(t_{i,j}) - t_{i-1,j}^{\hat{b}_{MV}} \log(t_{i-1,j}) \right] \left[\psi \left(\hat{c}_{MV} \left[t_{i,j}^{\hat{b}_{MV}} - t_{i-1,j}^{\hat{b}_{MV}} \right] \right) - \log(\delta_{i,j}) \right] \\ &= \sum_{j=1}^m \left(\hat{c}_{MV} t_{n_j,j}^{\hat{b}_{MV}} \log(t_{n_j,j}) \right) \log(\hat{u}_{MV}), \end{aligned} \quad (1.20)$$

avec $\psi(a)$, la fonction di-gamma, qui est la dérivée de la fonction log-gamma :

$$\psi(a) = \frac{\Gamma'(a)}{\Gamma(a)} = \frac{d}{da} \log \Gamma(a). \quad (1.21)$$

Ainsi, \hat{c}_{MV} and \hat{b}_{MV} peuvent soit être évalués numériquement en résolvant les équations (1.19) et (1.20) (en remplaçant \hat{u}_{MV} par son expression donnée dans (1.18) qui est fonction de \hat{c}_{MV} et \hat{b}_{MV}) ou en évaluant d'abord l'EMV de c et b puis en calculant \hat{u}_{MV} avec l'équation (1.18). Chacune des méthodes est numérique et implique des outils d'optimisation spécifiques pour résoudre les équations (1.18), (1.19) et (1.20). Nous avons dans un premier temps décidé d'utiliser la méthode de détermination de racines `hybr` qui repose sur les routines de MINPACK `hybrd` et `hybrj` (une version modifiée de la méthode de Powell, voir Scipy (2019) pour plus de détails).

On notera que d'autres méthodes d'estimation des paramètres ne reposant pas sur une approche fréquentiste (comme c'est le cas pour la méthode MV) sont présentes dans la littérature. On compte parmi elles la méthode bayésienne, permettant de tenir compte des incertitudes statistiques (Bousquet, Fouladirad, Grall & Paroissin, 2015 ; Dufresne, Gerber & Shiu, 1991 ; Newby & Barker, 2006 ; Nicolai, Budai, Dekker & Vreijling, 2004). Dans le cadre bayésien, on fait appel à un expert en raison de l'absence de données de dégradation suffisantes pour inférer les paramètres de la fonction de forme du processus.

1.3.5.3 Simulation d'un processus gamma

La simulation de trajectoires issues d'un processus gamma peut s'avérer utile dans différentes situations : pour évaluer les performances d'un estimateur à l'aide de simulations de *Monte Carlo* (voir section 1.5) ou pour juger de la robustesse d'un modèle dont les paramètres ont été estimés à partir de trajectoires de dégradation (voir section 1.4), pour ne citer qu'elles.

Plusieurs méthodes de simulation sont proposées pour approximer une trajectoire d'un processus gamma. On parle en effet d'approximation du fait que la simulation est réalisée à des intervalles de temps discrets et dénombrables alors qu'un processus gamma est défini par une infinité de sauts sur l'intervalle $[0; T]$. Parmi les méthodes de simulation disponibles, on peut citer : *Random Walk Approximation* (RWA) (Rosiński, 2014), *Gamma Bridge Sampling* (GBS) (Kahle *et al.*, 2016) et *Gamma Sequential Sampling* (GSS).

En raison de ses bonnes performances, la méthode GSS (Avramidis, L'ecuyer & Tremblay, 2003) pour générer une trajectoire d'un processus gamma non homogène de paramètres $(u, v(t))$ est présentée. Le pas de temps h entre deux incréments de dégradation est donné par : $h = 2^{-k}T$, où $[0, T]$ représente l'intervalle de temps de la simulation.

Cet algorithme a été employé pour la simulation des trajectoires présentées en Figure 1.14.

Algorithme 1.1 Algorithme de simulation par la méthode *Gamma Sequential Sampling* (GSS).

1	Entrée : Paramètres de simulation d'un processus gamma $\theta = (u, v(t))$
2	Sortie : Trajectoire simulée D_{sim}
3	Dégradation initiale : $G(0) = 0$
4	Pas de temps entre deux instants : $h = 2^{-k}T$
5	for $i = 1$ to k do
6	Generate $Q \sim Ga(u, v(ih) - v((i-1)h))$
7	$G(ih) = G((i-1)h) + Q$
8	Next i
9	end
10	$D_{sim} \leftarrow G$

1.3.5.4 Calcul du temps d'atteinte et durée de la de vie résiduelle

Temps d'atteinte

Un composant peut être considéré comme défaillant dès lors que son niveau de dégradation dépasse un seuil donné ρ . Ce dernier ne caractérise pas nécessairement l'inopérabilité du composant mais plutôt un niveau de dégradation à partir duquel il est considéré que celui-ci ne répond pas correctement à des critères de fonctionnement (ou de performance) donnés. Soit T_ρ l'instant auquel la défaillance se manifeste, aussi appelé premier temps d'atteinte du niveau de dégradation ρ :

$$T_\rho = \inf\{t \geq 0, X(t) \geq \rho\}. \quad (1.22)$$

Une représentation de la notion du temps d'atteinte est donnée dans la Figure 1.18.

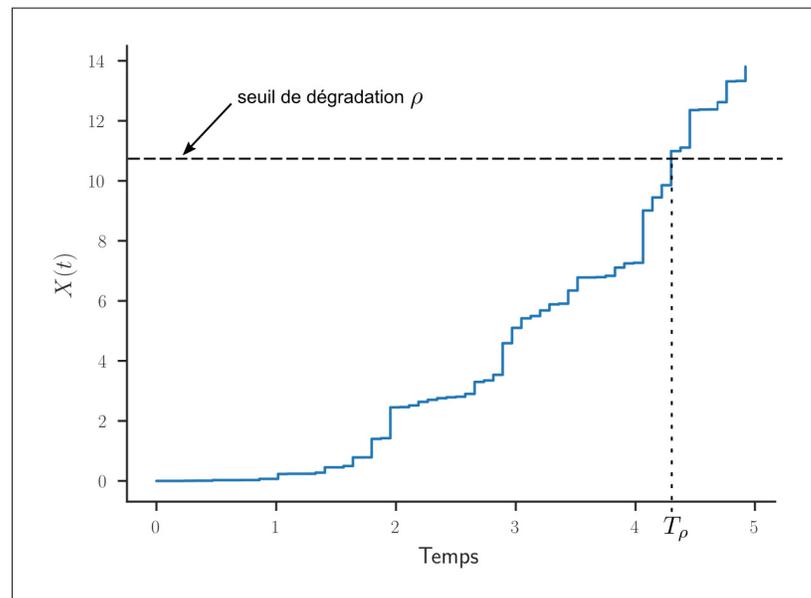


Figure 1.18 Temps d'atteinte T_ρ d'une trajectoire de dégradation au seuil ρ .

Les incréments de dégradation suivant des distributions de type gamma (voir équation (1.13)), la fonction de répartition de T_ρ est donnée dans l'expression suivante (Van Noortwijk, 2009) :

$$\begin{aligned}\mathbb{F}_{T_\rho}(t|u, v(t)) &= \mathbb{P}(T_\rho \leq t) = \mathbb{P}(X(t) \geq \rho) \\ &= \int_{x=\rho}^{\infty} f_{X(t)}(x|u, v(t)) dx = \frac{\gamma(v(t), \rho)}{\Gamma(v(t))}\end{aligned}\quad (1.23)$$

De plus, si on considère par la suite une fonction de forme du type : $v(t) = ct^b$, la CDF est dérivable et la fonction de densité de probabilité s'écrit alors :

$$\forall t \geq 0, \quad f_{T_\rho}(t|u, v(t)) = \frac{v'(t)}{\Gamma(v(t))} \int_{u\rho}^{\infty} \{\log(z) - \psi(v(t))\} z^{v(t)-1} e^{-z} dz. \quad (1.24)$$

On notera que d'autres expressions de la PDF du temps d'atteinte T_ρ sont données par Paroissin & Salami (2014). Cependant, en raison du comportement numérique de la fonction gamma, évaluer les équations (1.23) et (1.24) peut s'avérer en pratique compliqué même si ces expressions donnent la distribution exacte du premier temps d'atteinte du seuil ρ (Park & Padgett, 2005). Pour s'épargner cette contrainte d'ordre numérique, Park & Padgett (2005) présentent la distribution de Birnbaum-Saunders, qui permet d'approcher la distribution précédente. Cette approximation de la CDF du temps d'atteinte, notée \mathbb{F}_{BS} , est plus simple à évaluer que l'expression exacte de \mathbb{F}_{T_ρ} . Paroissin & Salami (2014) montrent d'ailleurs dans son étude théorique que cette approximation est aussi valable dans le cas des processus gamma non stationnaires (pour $v(t)$ de type loi puissance par exemple). L'équation (1.23) peut alors être approchée avec l'expression suivante :

$$\begin{aligned}\mathbb{F}_{T_\rho}(t|u, v(t)) &= \mathbb{P}(T_\rho \leq t) \simeq \mathbb{F}_{BS}(t|u, v(t)) \\ \text{avec } \mathbb{F}_{BS}(t|u, v(t)) &= -\Phi\left(\frac{u\rho - v(t)}{\sqrt{v(t)}}\right) = \Phi\left[\sqrt{u\rho}\left(\sqrt{\frac{v(t)}{u\rho}} - \sqrt{\frac{u\rho}{v(t)}}\right)\right],\end{aligned}\quad (1.25)$$

avec Φ la fonction de répartition de la distribution gaussienne standard.

Durée de vie résiduelle

La durée de vie résiduelle (*remaining useful lifetime (RUL)*) est définie à l'instant t comme le temps restant à un matériel observé à un instant t avant que son niveau de dégradation ne dépasse un seuil ρ donné. Dans le cas d'un processus gamma, celle-ci peut être définie à partir du premier temps d'atteinte (introduit plus haut) et de l'instant d'inspection t . Si on note $\mathbf{D}_{obs,t}$ l'ensemble des observations jusqu'à l'instant t , la *RUL* s'écrit alors (Le Son *et al.*, 2016) :

$$RUL(t) = \inf\{h : X(t+h) \geq \rho | X(t) < \rho, \mathbf{D}_{obs,t}\} \quad (1.26)$$

La distribution de la durée de vie résiduelle a pour expression :

$$\mathbb{F}_{RUL(t)}(h) = \mathbb{P}(RUL(t) < \rho, \mathbf{D}_{obs,t}) = \mathbb{P}(X(t+h) > \rho | X(t) < \rho, \mathbf{D}_{obs,t}) \quad (1.27)$$

Si la valeur exacte de $X(t)$ est contenue dans $\mathbf{D}_{obs,t}$, alors :

$$\begin{aligned} \mathbb{F}_{RUL(t)}(h) &= \mathbb{P}(X(t+h) > \rho | X(t) < \rho, X(t) = x) = \mathbb{P}(X(t+h) > \rho - x) \\ &= \frac{\gamma(v(t+h) - v(t), (\rho - x)u)}{\Gamma(v(t+h) - v(t))} \end{aligned} \quad (1.28)$$

1.4 Sélection de modèles

Le choix d'un processus stochastique plutôt qu'un autre pour modéliser la dynamique d'une dégradation est souvent basé sur des hypothèses formulées par l'analyste. Celles-ci peuvent s'appuyer sur le phénomène physique entraînant la dégradation ou bien sur une analyse préliminaire des données de cette dégradation. Kang, Gong & Chen (2020) proposent au travers de leur étude une classification simple des processus stochastiques couramment rencontrés en fiabilité. Celle-ci est représentée par la Figure 1.19.

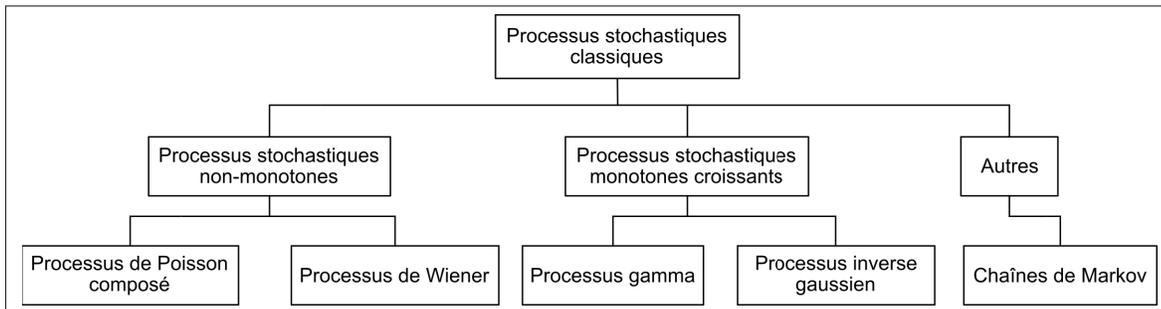


Figure 1.19 Classification simple de processus stochastiques couramment rencontrés en fiabilité.

On remarque que cette hiérarchisation ² repose essentiellement sur le caractère monotone ou non des processus. Cette caractéristique est en effet primordiale dans le choix d'un processus stochastique. Même si on retrouve souvent en fiabilité des phénomènes de dégradation strictement croissants (comme c'est le cas pour la cavitation érosive), l'image de cette dégradation, issue de mesures, peut ne pas présenter la monotonie attendue. C'est le cas en présence de mesures bruitées ou encore d'erreurs de mesure (liées à la méthode employée ou au composant lui-même). Ainsi, Wang *et al.* (2014) proposent de modéliser la perte de capacité d'un condensateur par un processus de Wiener avec dérive (ici les mesures sont sujettes à des incertitudes).

Nguyen *et al.* (2018) abordent le problème du choix de processus stochastiques au travers d'une étude comparant différents critères classiques et pronostics. Parmi les critères classiques, on compte les tests d'adéquation et les critères de sélection. Les premiers quantifient la différence entre les observations issues des données de dégradation et les valeurs issues du modèle choisi, tandis que les deuxièmes intègrent la notion de complexité dans leur évaluation du modèle. Les critères pronostics quantifient quant à eux l'écart entre les temps de défaillance observés et les distributions de durée de vie estimées par les modèles considérés. Leur étude est menée sur deux familles de processus stochastiques communément rencontrés en dégradation : les processus gamma et les processus browniens / d'Ornstein-Uhlenbeck. Ils étudient respectivement deux modèles appartenant à la première famille de processus et quatre pour la seconde. À l'aide de simulations pour chacun des processus, ils montrent que les indicateurs statistiques

² On notera que pour des cas particuliers, le processus de Poisson composé peut être monotone croissant.

classiques présentent de meilleures performances que les critères de pronostics pour choisir à quelle famille de processus appartiennent les données. Néanmoins, dans de nombreux cas, ces mêmes indicateurs statistiques ne parviennent pas à sélectionner (à l'intérieur d'une famille de processus) le modèle dont sont issues les données de dégradation.

1.4.1 Critères de sélection

Les tests d'adéquation, par leur construction, peuvent tendre à favoriser les modèles présentant une complexité relative plus élevée (Myung, 2000). Pour intégrer cette notion de complexité, des critères alternatifs pour la sélection de modèles existent, comme le critère d'information d'Akaike (*AIC*) ou encore le critère d'information bayésien (*BIC*). Ils sont tous les deux basés sur la théorie de l'information (Akaike, 1974). En particulier l'*AIC* sélectionne le modèle dont la distribution s'approche le plus de la "vraie" distribution des données (selon la divergence de Kullback–Leibler) et présente le moins de paramètres. Ainsi, si on considère un jeu de données \mathbf{D}_{obs} , l'*AIC* d'un modèle M ayant k paramètres, est donné par :

$$AIC = 2 \ln \mathcal{L}(\mathbf{D}_{obs}|M) + 2k, \quad (1.29)$$

avec $\mathcal{L}(\mathbf{D}_{obs}|M)$ le maximum de la fonction vraisemblance du modèle. Le modèle présentant la plus faible valeur d'*AIC* sera alors privilégié.

Le deuxième critère, le *BIC*, prend en compte, en plus du nombre de paramètres du modèle, la taille n du jeu de données. Celui-ci est donné par :

$$BIC = 2 \ln L(\mathbf{D}_{obs}|M) + k \ln n. \quad (1.30)$$

De même que pour l'*AIC*, le modèle présentant la plus faible valeur de *BIC* est sélectionné. Ainsi, *BIC* tient compte à la fois du nombre k de paramètres du modèle et de la taille n de l'échantillon. La fonction log-vraisemblance étant croissante selon k et n , celle-ci est "pénalisée" par le critère de *BIC* lorsque la taille n de l'échantillon de données ou le nombre de paramètre augmente.

1.4.2 Tests d'adéquation

De nombreux tests statistiques sont à la disposition de l'analyste pour vérifier si un jeu de données suit une distribution connue (par exemple gaussienne) ou encore si deux jeux de données suivent une même quelconque distribution (Mecklin & Mundfrom, 2004). Parmi eux, on peut citer les classiques tests de Kolmogorov-Smirnov (KS) (Justel, Peña & Zamar, 1997), d'Anderson-Darling (souvent considéré comme une alternative au test KS) (Engmann & Cousineau, 2011) ou encore celui de Cramer-von-Mises (Darling, 1957). En particulier les tests précédemment évoqués s'intéressent à quantifier l'écart entre les données étudiées et le modèle proposé. Ces tests d'adéquation sont valides sous l'hypothèse que les données soient indépendantes et identiquement distribuées (i.i.d.). Si les données sont non-identiquement distribuées (i.n-i.d.), ces tests ne peuvent plus être appliqués.

Très populaires pour les distributions gaussiennes, les tests d'adéquation le sont moins pour des applications liées aux processus stochastiques. En particulier, choisir un modèle adéquat de processus revient à proposer une distribution adaptée aux incréments de dégradation. Kahle *et al.* (2016) présentent deux tests statistiques et un outil graphique permettant d'écarter ou non l'hypothèse que les incréments de dégradation issus d'un jeu de données suivent une distribution gamma. Pour ces trois outils, l'hypothèse que les incréments de dégradation sont i.i.d. doit être vérifiée. Par conséquent, ils ne peuvent être employés qu'en présence d'un processus gamma (supposé) homogène.

Le premier test statistique, introduit par Henze, Meintanis & Ebner (2012), est basé sur la transformation de Laplace des incréments de dégradation. Cette transformation, notée \mathcal{L} , est motivée par le fait que la distribution gamma, $Ga(u, v)$, $u, v \in \mathbb{R}_+^*$, est l'unique solution de l'équation différentielle suivante :

$$(1 + vt)\mathcal{L}'(t) + uv\mathcal{L}(t) = 0 \quad (1.31)$$

avec pour condition initiale : $\mathcal{L}(0) = 1$. Soit (\hat{u}_n, \hat{v}_n) un estimateur de (u, v) , on pose alors $Y_i = \hat{v}_n X_i$, $i \in \{1, \dots, n\}$ et $\hat{\mathcal{L}}_n$ est la transformée de Laplace des Y_1, \dots, Y_n . Les X_i désignent

un échantillon de n variables aléatoires i.i.d. À partir de l'équation (1.31), la statistique suivante est proposée :

$$T_n = \int_0^\infty Z_n^2(t)w(t)dt, \quad (1.32)$$

avec w une fonction de pondération continue positive telle que : $\int_0^\infty t^4w(t)dt < \infty$ et où :

$$Z_n(t) = \sqrt{n} [(1+t)\hat{\mathcal{L}}'_n(t) + \hat{u}_n\hat{\mathcal{L}}_n(t)]. \quad (1.33)$$

Henze *et al.* (2012) proposent de prendre $w(t) = \exp(-ct)$ ou $w(t) = \exp(-ct^2)$ avec $c > 0$. L'hypothèse H_0 , selon laquelle les variables X_i suivent une distribution $Ga(u, \nu)$, est rejetée pour des grandes valeurs de T_n .

Un deuxième test statistique est basé sur le ratio de deux estimateurs différents de la variance. Les estimateurs sont obtenus à partir des équations utilisées dans la méthode des moments (Van Noortwijk, 2009). Cette méthode, introduite par Villaseñor & González-Estrada (2015), est semblable au très populaire outil développé par Shapiro & Wilk (1965) pour tester la normalité d'un échantillon de réalisation d'une variable aléatoire.

1.4.3 Méthode graphique

Cette méthode, proposée par Ye, Xie, Tang & Chen (2014b), permet de comparer la qualité de modélisation par différents processus stochastiques stationnaires à partir d'une représentation graphique. Si on considère une réalisation d'un processus stochastique $(X_i)_{i \in \mathbb{N}}$, ses incréments de dégradation entre les instants t_{i-1} et t_i sont notés δ_i , $i \in 1, \dots, n$. On suppose que ces derniers suivent une distribution de paramètre θ_i pouvant être estimé par une des méthodes présentées en section 1.3.5.2 et noté $\hat{\theta}_i$. Ainsi, si les données sont issues du modèle dont la CDF s'écrit : $\mathbb{F}_{X_i}(\cdot|\theta_i)$, alors $U_i := \mathbb{F}_{X_i}(\delta_i|\theta_i)$ est uniformément distribué sur $[0; 1]$. Donc, en remplaçant θ_i par une estimation $\hat{\theta}_i$ et si le modèle utilisé est adéquat, la distribution des variables aléatoires $\hat{U}_i := \mathbb{F}_{X_i}(\delta_i|\hat{\theta}_i)$ devrait être proche d'une distribution uniforme (Kahle *et al.*, 2016).

Un diagramme quantile-quantile (*Q-Q plot*) basé sur $\hat{U}_1, \dots, \hat{U}_n$ peut alors être utilisé pour comparer différents processus stochastiques.

Afin d'illustrer la méthode, des trajectoires d'un processus gamma homogène de paramètres $\theta_{Ga} = (u; c) = (1/100; 1)$ ont été tirées ($n = 100$) par simulation (les incréments de temps sont tous pris identiques). Ensuite, à partir de ce jeu de données, une estimation des paramètres pour une distribution normale, pour une distribution inverse gaussienne et pour une distribution gamma a été réalisée. Ces paramètres estimés sont notés respectivement $\hat{\theta}_N = (\hat{\mu}; \hat{\sigma}) = (0,011; 0,012)$, $\hat{\theta}_{IG} = (\hat{\mu}; \hat{\lambda}) = (0,012; 0,086)$ et $\hat{\theta}_{Ga} = (\hat{u}; \hat{c}) = (0,85; 0,011)$. Les deux diagrammes quantile-quantile associés ont été représentés dans la Figure 1.20 à partir des CDF $\mathbb{F}_{X_i}(\delta_i; \hat{\theta}_{Wi})$, $\mathbb{F}_{X_i}(\delta_i; \hat{\theta}_{IG})$ et $\mathbb{F}_{X_i}(\delta_i; \hat{\theta}_{Ga})$. En noir est représentée la CDF d'une distribution uniforme sur $[0; 1]$. On peut constater à partir de ce graphique que, logiquement, les paramètres estimés d'une distribution gamma induit une CDF plus proche de la courbe noire que ceux issus d'une distribution normale (pour rappel, le jeu de données utilisé est issu d'une distribution gamma). On constate également, que pour le paramètre θ_{Ga} choisi, la distribution des tirages peut être également approximée par une distribution inverse gaussienne de paramètre $\hat{\theta}_{IG}$. En effet, les diagrammes quantile-quantile associés sont très proches l'un de l'autre.

Comme expliqué précédemment, lorsque les incréments de dégradation sont supposés appartenir à une réalisation d'un processus stochastique non stationnaire, les tests d'adéquation classiques ne peuvent pas être employés. En effet, chaque incrément suit une distribution \mathcal{D} différente $X(\tau) - X(t) \sim \mathcal{D}(\theta(\tau) - \theta(t))$ et l'hypothèse H_0 ne peut alors être vérifiée. Pour pallier ce problème, des tests d'adéquations sont en cours de développement pour les variables aléatoires i.n.i.d. ; ces derniers s'appuient notamment sur la théorie de la profondeur (*depth test*) (Zhang, Xiang & Shen, 2012). En définitif, peu d'outils sont à la disposition de l'analyste lorsque celui-ci doit vérifier l'adéquation d'un processus non stationnaire à un jeu de données ou bien pour comparer deux processus entre eux.

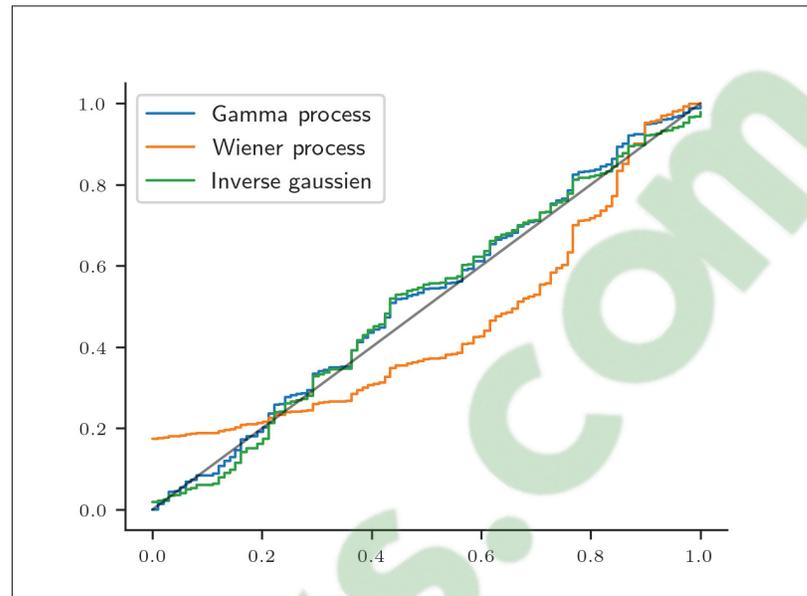


Figure 1.20 Exemple de diagramme quantile-quantile pour la comparaison de deux processus stochastiques.

1.5 Traitement des incertitudes

Dans la perspective d'obtenir une évaluation de la durée de vie résiduelle robuste pour un composant donné, il est nécessaire en amont d'évaluer correctement les incertitudes d'estimation des paramètres du modèle de dégradation. En effet, même si les processus stochastiques, comme les processus gamma, peuvent se montrer mathématiquement adaptés à la modélisation de la dégradation, l'estimation des paramètres par maximum de vraisemblance (ou d'autres méthodes) se base uniquement sur l'information disponible contenue dans les données, qui peut s'avérer limitée en pratique (coût d'acquisition des mesures, matériels très fiables, ...). Dans les prochains chapitres, nous nous intéresserons à modéliser des données de dégradation issues d'essais en laboratoire dans un premier temps, pour ensuite traiter des mesures collectées en centrale. Les premières présentent peu d'observations par trajectoire, mais quelques répliques sont disponibles, tandis que dans la deuxième situation, beaucoup d'observations sont disponibles mais sur une seule trajectoire. Si on considère deux trajectoires de dégradation différentes issues d'un même composant, l'estimation des paramètres d'un processus pour chacune des trajectoires produira des valeurs différentes. Ceci s'explique par le fait que la méthode d'estimation repose

sur les données de dégradation et que ces dernières présentent une certaine variabilité en raison de différents facteurs.

D'un point de vue industriel, ces variations dans l'estimation des paramètres du modèle de dégradation impliquent des conséquences en cascade : typiquement, la durée de vie résiduelle du matériel estimée risque d'être erronée entraînant un arrêt anticipé ou tardif du matériel en question et dans un même temps des pertes d'exploitation et/ou des coûts de maintenance plus importants. Pour prévenir les conséquences d'une information limitée, il est parfois nécessaire de collecter de nouvelles données pour réduire ces incertitudes et obtenir *in fine* un modèle de dégradation plus "précis".

Concrètement, l'analyste va calculer des intervalles de confiance pour les paramètres, qui reflètent la qualité de l'information recueillie. Ces intervalles contiennent *a priori* la vraie valeur du paramètre θ (inconnue). Afin d'évaluer la robustesse de la méthode d'estimation, plusieurs outils mathématiques applicables aux processus stochastiques sont présentés dans les sections suivantes.

1.5.1 Intervalles de confiance asymptotiques à partir des propriétés de convergence de l'EMV

Si on considère que $N = \sum_{j=1}^m n_j$, le nombre total d'observations (voir section 1.3.5.2), devient grand, alors la fonction log-vraisemblance ℓ présente des propriétés intéressantes. Tout d'abord, $\ell(\theta|\mathbf{D}_{obs})$ présente une convexité de plus en plus importante au voisinage de $\hat{\theta}_{MV}$ (Wald, 1949). Ensuite, $\ell(\theta|\mathbf{D}_{obs})$ se montre plus symétrique autour de $\hat{\theta}_{MV}$. Ce comportement de symétrie montre simplement que plus la taille de l'échantillon des observations \mathbf{D}_{obs} croît, plus nous sommes confiants que la loi de $\hat{\theta}_{MV}$ s'approche d'une loi gaussienne. Ainsi, la courbure de la fonction de vraisemblance (*i.e.* la dérivée seconde de $\ell(\theta|\mathbf{D}_{obs})$ par rapport à θ) s'avère être une caractéristique importante. C'est sur cette propriété que repose la construction d'intervalles de confiance asymptotiques autour de la valeur inconnue θ . Les précédentes affirmations s'appuient sur les propriétés asymptotiques des dérivées premières et secondes de $\ell(\theta|\mathbf{D}_{obs})$ données par le théorème central limite (TCL), dans le cas d'incrémentés indépendants et identiquement

distribués (i.i.d.) (Birolini, 2007 ; Sweeting, 1980). Hoadley (1971) donne les conditions pour lesquelles la convergence asymptotique du maximum de vraisemblance peut être démontrée, dans le cas particulier d'incrémentés indépendants mais non identiquement distribués (i. non-i.d.), ce qui est le cas des données exploitées dans les Chapitres 3 et 4.

Pour alléger les notations, $\ell(\boldsymbol{\theta}|\mathbf{D}_{obs})$ sera simplement noté par la suite $\ell(\boldsymbol{\theta})$. Dans le contexte du maximum de vraisemblance, la normalité asymptotique, telle que définie par Rao (1973), s'écrit :

$$(\hat{\boldsymbol{\theta}}_{MV} - \boldsymbol{\theta}) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}\left(0, I^{-1}(\boldsymbol{\theta})\right), \quad (1.34)$$

avec $I(\boldsymbol{\theta})$ la matrice de l'information de Fisher espérée. Cette dernière est donnée par l'expression :

$$I(\boldsymbol{\theta}) = -\mathbb{E} \left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(\boldsymbol{\theta}) | \boldsymbol{\theta} \right] \quad (1.35)$$

Aussi, \xrightarrow{d} désigne la convergence en loi lorsque N tend vers l'infini. Étant donné qu'il est impossible de calculer l'information de Fisher espérée lorsque son expression analytique est absente, on utilise plutôt l'information de Fisher observée $J(\boldsymbol{\theta})$, présentée par Efron & Hinkley (1978) et donnée dans l'expression suivante :

$$J(\boldsymbol{\theta}) = -\nabla^2 \ell(\boldsymbol{\theta}) \quad (1.36)$$

Ainsi, il est possible de calculer des intervalles de confiance sur $\boldsymbol{\theta}$ à partir de la normalité asymptotique du maximum de vraisemblance. En considérant chaque composant θ_i du vecteur des paramètres de $\boldsymbol{\theta}$, l'intervalle de confiance au niveau $1 - \alpha$, $0 \leq \alpha \leq 1$ associé se calcule avec :

$$\left[\hat{\boldsymbol{\theta}}_{MV,i} - s \sqrt{\left(J(\hat{\boldsymbol{\theta}})^{-1}\right)_{ii}} ; \hat{\boldsymbol{\theta}}_{MV,i} + s \sqrt{\left(J(\hat{\boldsymbol{\theta}})^{-1}\right)_{ii}} \right], \quad (1.37)$$

avec $s = \Phi\left(1 - \frac{\alpha}{2}\right)$ et Φ la CDF d'une loi normale centrée réduite.

1.5.2 Méthode delta

La méthode delta (MD), aussi appelée méthode de propagation des erreurs, permet d'obtenir la variance de la transformée d'une ou plusieurs variables aléatoires (Casella & Berger, 2002). Ainsi, par exemple, connaissant la loi asymptotique de $\hat{\theta}$, il est possible de dériver la loi asymptotique d'autres indicateurs découlant de θ .

Oehlert (1992) introduit la MD avec les notations suivantes : on note $(W_n)_{n \in \mathbb{N}}$ la série de vecteurs aléatoires à valeurs dans \mathbb{R}^p . On suppose qu'il existe $\mu_W \in \mathbb{R}^p$ et Σ une matrice $p \times p$ définie positive tels que :

$$\sqrt{n}(W_n - \mu_W) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma) \quad (1.38)$$

Soit r fonctions réelles g_1, \dots, g_r dont la dérivée partielle première est continue en μ_W , et supposons qu'au moins une de ces dérivées est non nulle. Pour $i \in \{1, \dots, r\}$ et pour $n \in \mathbb{N}^*$ quelconque, on pose $Z_{i,n} = g_i(W_n)$, $Z_n = (Z_{1,n}, \dots, Z_{r,n})^T$ et $\mu_Z = (g_1(\mu_W), \dots, g_r(\mu_W))^T$. Alors la série $(Z_n)_{n \in \mathbb{N}^*}$ est aussi asymptotiquement normale :

$$\sqrt{n}(Z_n - \mu_Z) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, K\Sigma K^T), \quad (1.39)$$

avec K la matrice $r \times p$ ayant pour éléments $k_{i,j} = \frac{\partial g_i}{\partial w_j}$ pour $i \in \{1, \dots, r\}$ et $j \in \{1, \dots, p\}$. La matrice K est aussi appelée matrice jacobienne des fonctions g_i . Ce résultat permet donc d'évaluer, sous certaines conditions, la distribution asymptotique (lorsque $n \rightarrow \infty$) de $g_i(W_n)$ à partir de celle de $(W_n)_{n \in \mathbb{N}}$.

Fouladirad, Paroissin & Grall (2018) illustrent la méthode delta par une étude de sensibilité qui compare l'influence de différentes approches de maintenance sur l'estimation de la durée de vie d'un composant.

D'un point de vue pratique, on s'intéressera dans le Chapitre 3 à "propager les erreurs" d'estimation des paramètres du modèle de dégradation afin d'obtenir l'incertitude d'estimation sur la distribution du temps d'atteinte pour un certain seuil de dégradation associée. On appliquera

la MD aux incertitudes d'estimation d'un processus gamma. Pour cela, il conviendra de choisir minutieusement une fonction g correspondant à la transformation d'intérêt.

1.5.3 Méthodes de ré-échantillonnage

Le concept de ré-échantillonnage, qui se retrouve souvent sous son nom anglais *bootstrap*, fut initialement présenté par (Efron et al., 1979) dans le cadre de variables indépendantes, puis l'application de la méthode fut élargie ensuite par de nouveaux auteurs (Härdle, Horowitz & Kreiss, 2003). L'idée générale du *bootstrap* consiste à reformer de nouveaux échantillons à partir d'un échantillon initial pour réaliser des inférences statistiques. L'objectif de cette méthode est de remplacer des hypothèses probabilistes (condition asymptotique par exemple), dont les conditions d'applicabilité ne sont pas toujours vérifiées. Ces méthodes de ré-échantillonnage permettent ainsi d'approcher la distribution d'un estimateur lorsqu'on ne connaît pas la loi de l'échantillon étudié par exemple. Il est aussi possible de dériver la variabilité d'autres indicateurs (en particulier celle de la *RUL* ou de la fonction de répartition du temps d'atteinte) à partir de la distribution des paramètres estimés sur les nouveaux échantillons. En particulier, diverses méthodes *bootstrap* ont été introduites pour les modèles de processus stochastiques ou séries temporelles (Berkowitz & Kilian, 2000 ; Bose & Politis, 1993 ; Bühlmann, 2002 ; Carey, 2005 ; Härdle *et al.*, 2003 ; Hongyi Li & Maddala, 1996). D'un point de vue pratique, il est assez facile numériquement d'utiliser la méthode *bootstrap*, ce qui explique son développement depuis les années 1970, le grand nombre de réplifications de tirages nécessitant une certaine puissance de calcul.

1.5.3.1 Méthode paramétrique

Dans le cas du *bootstrap* paramétrique (BP), on fait l'hypothèse que les données initiales sont issues d'un modèle paramétrique donné, de paramètre θ . Les nouveaux échantillons sont alors simulés à partir du modèle de paramètre θ .

En pratique, la méthode employée dans les prochains chapitres repose sur les étapes suivantes. Étant donné les observations \mathbf{D}_{obs} , de taille totale $N = \sum_{j=1}^m n_j$ (voir section 1.3.5.2), on estime d'abord $\hat{\theta}$ (Genest, Rémillard et al., 2008 ; Stute, Manteiga & Quindimil, 1993). Ensuite, $\hat{\theta}$ est utilisé pour ré-échantillonner $B \in \mathbb{N}$ fois m trajectoires de dégradation de longueur n_j , $1 \leq j \leq m$, à partir du modèle considéré (un processus stochastique dans notre cas). Ainsi, les B échantillons *bootstrap* sont utilisés afin d'obtenir B estimations de $\hat{\theta} : \hat{\theta}_{BP,w}^*$, $1 \leq w \leq B$. À partir de ces B valeurs de $\hat{\theta}_{BP}^*$, on obtient une estimation de l'incertitude sur $\hat{\theta}$ permettant donc d'évaluer le comportement de l'estimateur *bootstrap* $\hat{\theta}$ par rapport à θ .

La méthode de *bootstrap* paramétrique est généralement privilégiée en présence de peu de données (moins de 10 observations) ou lorsque l'analyste a une idée du modèle dont sont issues ces données. En effet, pour les méthodes non paramétriques (voir la section suivante), le ré-échantillonnage est réalisé à partir du jeu de données initial de taille N . Ainsi si N est petit, la structure fine des données peut ne pas être correctement capturée dans cet échantillon. La simulation de *Monte Carlo*, pour la méthode paramétrique, permettrait, *a priori*, de pallier cet inconvénient.

1.5.3.2 Méthodes non paramétriques

À la différence du *bootstrap* paramétrique, les méthodes non paramétriques ne font, quant à elles, pas l'hypothèse que les données sont issues d'un modèle paramétrique donné. Parmi la multitude de méthodes de ré-échantillonnage non paramétrique disponibles (voir Kreiss & Lahiri (2012)), deux d'entre elles sont présentées par la suite en raison de leur relative popularité.

Moving Block Bootstrap

Lorsque les données sont organisées sous forme de séries temporelles, ré-échantillonner une seule observation (à partir des données de dégradation, dans le contexte présent), comme suggéré par Efron & Tibshirani (1994), peut produire des résultats décevants (Kreiss & Lahiri, 2012). En effet, en présence d'une dépendance temporelle (ou entre composants étudiés), l'ordre des observations

n'est pas aléatoire. La méthode de *Moving Block Bootstrap* (MBB) peut donc s'avérer efficace pour ré-échantillonner des données présentant une dépendance temporelle, comme c'est le cas pour les données de dégradation cumulative. Celle-ci repose sur un ré-échantillonnage qui sélectionne un bloc d'observations ordonnées dans le temps. De cette façon, une partie de la dépendance temporelle est préservée (Kreiss & Lahiri, 2012).

Ainsi, considérons la $j^{\text{ième}}$ trajectoire de dégradation de longueur n_j comme une série temporelle, et supposons que les incréments de dégradation $\delta_{i,j}$ sont observés entre les instants $t_{i,j}$ et $t_{i+1,j}$, $0 \leq i \leq (n_j - 1)$. On note $\mathbf{Y}_j = \{Y_{0,j}, \dots, Y_{n_j-1,j}\}$, avec $Y_{i,j} = \{t_{i,j}, t_{i+1,j}, \delta_{i,j}\}$, $1 \leq j \leq m$. Soit l un entier avec $1 \leq l \leq n_j$. Les blocs superposés $B_{0,j}, \dots, B_{p,j}$, $p = n_j - l$ de longueur l , appartenant à \mathbf{Y}_j sont définis de la façon suivante :

$$\begin{aligned}
 B_{0,j} &= (Y_{0,j}, Y_{1,j}, \dots, Y_{l-1,j}), \\
 B_{1,j} &= (Y_{1,j}, \dots, Y_{l-1,j}, Y_{l,j}), \\
 &\dots \qquad \qquad \qquad \dots \\
 B_{p,j} &= (Y_{n_j-l,j}, \dots, Y_{n_j-1,j})
 \end{aligned} \tag{1.40}$$

Pour des questions de simplification, on considère que n_j est un multiple de l . Soit $b_j = n_j/l$. Pour générer un échantillon MBB, b_j blocs sont tirés aléatoirement avec remise à partir de l'échantillon initial $\{B_{0,j}, \dots, B_{p,j}\}$. Le jeu de données ré-échantillonnées $\mathbf{Y}_j^* = \{Y_{0,j}^*, \dots, Y_{n_j-1,j}^*\}$ contient $b_j \times l = n_j$ éléments puisque chaque bloc tiré possède l éléments et que b_j blocs sont tirés. Idéalement, pour des données temporellement dépendantes, l devrait satisfaire la condition : $l = Cn_j^{1/k}$, avec $k = 3$ ou 4 , où $C \in \mathbb{R}$ est une constante (Kreiss & Lahiri, 2012). Dans le cas général où l n'est pas un multiple de n_j , $b_j = b_0$ blocs doivent être tirés avec $b_0 = \min\{k \geq 1 : kl \geq n_j\}$ et seulement les n_j premières valeurs ré-échantillonnées sont conservées pour former le nouvel ensemble \mathbf{Y}_j^* .

Efron Bootstrap

La méthode *Efron bootstrap* (EB) initialement présentée par Efron (1992) peut être assimilée à un cas spécial de la méthode MBB. En effet, si on prend $l = 1$, celle-ci est semblable à la méthode de *bootstrap* non paramétrique décrite par Efron & Tibshirani (1994) pour des données i.i.d. (voir application de la méthode dans Croteau, Cleroux & Leger (2005)).

1.6 Conclusion

Ce chapitre propose une recherche bibliographique portant sur un phénomène important de dégradation rencontré sur les turbines hydroélectriques, la cavitation érosive, mais aussi sur les processus stochastiques couramment utilisés pour la modélisation de la dégradation à partir de données. Aussi, une partie du chapitre est dédiée à la prise en compte des incertitudes et à leur intégration aux modèles précédemment considérés. Cette étude peut être résumée à l'aide des points suivants :

- Bien que largement étudiée, la modélisation de la cavitation par modèle physique ou à l'aide d'outils numériques (simulation en mécanique des fluides) montre une maturité limitée pour des applications à des matériels complexes, telle qu'une roue de turbine Francis.
- Les processus stochastiques permettent à la fois d'intégrer la variabilité du phénomène étudié, mais aussi de réactualiser le modèle à partir de nouvelles observations.
- En particulier, les processus gamma et inverse gaussien présentent de bonnes propriétés pour la modélisation de phénomènes de dégradation cumulatifs monotones et croissants. Néanmoins, le processus gamma sera privilégié par la suite en raison de son explicabilité physique et des nombreux résultats mathématiques disponibles. En effet, la bonne compréhension et l'appropriation du modèle de dégradation retenu sont des caractéristiques importantes dans le contexte de cette thèse industrielle.

La suite de cette thèse sera développée selon les étapes suivantes :

- Chapitre II : Étude des performances de la méthode d'estimation par maximum de vraisemblance. Dans cette partie théorique, nous nous intéressons au comportement de l'estimateur

des paramètres d'un processus gamma à partir de données simulées. Une variante de la méthode d'estimation des paramètres par maximum de vraisemblance sera aussi proposée afin de pallier certains problèmes numériques.

- Chapitre III : Modélisation de la cavitation érosive observée en laboratoire. À partir des données de dégradation issues d'essais expérimentaux en laboratoire, une comparaison de deux matériaux différents est réalisée selon un critère de durée de vie. De plus, les incertitudes d'estimation des paramètres du modèle seront propagées à l'aide de différentes méthodes présentées en section 1.5.
- Chapitre IV : Cas d'étude réalisé à partir de données de monitoring en continu. Ce chapitre sera dédié à la modélisation de la cavitation érosive observée sur une turbine réelle. Les contraintes liées à l'acquisition du niveau de dégradation ainsi que les problèmes d'estimation du modèle seront présentés dans cette partie.

CHAPITRE 2

ÉTUDE DE LA MÉTHODE D'ESTIMATION PAR MAXIMUM DE VRAISEMBLANCE

2.1 Introduction

L'objectif de ce chapitre est d'étudier les performances de la méthode d'estimation par maximum de vraisemblance (EMV) présentée dans le chapitre précédent dans le cas d'un processus gamma. Pour cela, cette méthode sera appliquée à l'estimation de paramètres d'un processus gamma non homogène. Dans un deuxième temps, une méthode alternative à celle précédemment introduite sera proposée afin de pallier les problèmes numériques rencontrés avec la première.

2.2 Performance de la méthode d'EMV

Comme introduit dans le chapitre précédent, l'estimation des paramètres d'un modèle donné influence *in fine* les stratégies de maintenance qui pourront être mises en œuvre à partir des résultats de pronostic. Il convient donc d'évaluer, dans un premier temps, les performances de la méthode d'estimation par MV qui a été retenue afin de quantifier, entre autres, l'influence de la taille des données sur la "précision" des estimations. En effet, il est rappelé que nous disposons principalement de deux sources et typologies différentes de données : la première issue de tests de cavitation menés en laboratoire (peu d'observations et quelques trajectoires répliquées) et la deuxième provient d'un système de surveillance en centrale (beaucoup d'observations mais pas de réplication). D'un point de vue industriel, l'estimation du temps d'atteinte du niveau de dégradation pour un seuil ρ donné est essentielle. Pour cette raison, nous nous sommes intéressés à évaluer les performances de la méthode MV sur ce critère, à partir de trajectoires simulées dont les "vrais" paramètres sont connus (a minima les paramètres de simulation dont sont issues les trajectoires générées sont connus). Le plan d'expérience retenu est présenté dans le Tableau 2.1.

Des valeurs classiques de paramètres d'un processus gamma (u, c, b) ont été choisies pour simuler des trajectoires de dégradation présentant plus ou moins de variabilité (au travers du

Tableau 2.1 Plan d'expérience.

Paramètres	Valeurs
Paramètres PG ($u; c; b$)	(1; 1; 1)
	(0,5; 1; 1)
	(1; 0,5; 1)
	(1; 1; 2)
	(2; 1; 1)
	(1; 1; 0,5)
Longueur de trajectoire n_j	1000
	100
	10
Nombre de trajectoires m	1000
	100
	10
	1

ratio $v(t)/u^2$) et différentes vitesses moyennes de dégradation (pente et convexité/concavité de la tendance moyenne de la trajectoire au travers du paramètre b). À partir de ces trajectoires de dégradation, les performances de la méthode MV sont étudiées en termes d'estimation des paramètres d'un processus gamma et d'espérance sur la distribution du temps d'atteinte T_ρ . Noter que dans le cas $n_j = 1000$, le nombre maximal de trajectoires m est limité à 100 en raison de contraintes d'ordre numérique.

Les indicateurs statistiques utilisés pour évaluer les performances de la méthode MV sont présentés dans le Tableau 2.2. L'espérance empirique du temps d'atteinte T_ρ notée \hat{E} , est un estimateur de la moyenne de T_ρ . La variance de $\hat{E}(T_\rho)$, notée $\hat{V}(\hat{E}(T_\rho))$, est obtenue directement à partir de la méthode delta (présentée en section 1.5.2). Les résultats de simulation et d'estimation sont présentés dans les Tableaux 2.3 à 2.6.

Tableau 2.2 Indicateurs statistiques utilisés pour évaluer les performances de la méthode d'estimation MV.

	Expression mathématique
Biais	$\hat{\mathbb{E}}(T_\rho) - \mathbb{E}(T_\rho)$
Biais relatif	$\frac{\hat{\mathbb{E}}(T_\rho) - \mathbb{E}(T_\rho)}{\mathbb{E}(T_\rho)}$
Erreur relative	$\frac{\sqrt{\hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))}}{\hat{\mathbb{E}}(T_\rho)}$
Erreur quadratique moyenne relative	$\frac{\sqrt{\text{biais}^2 + \hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))}}{\mathbb{E}(T_\rho)}$

Les Figures 2.1 et 2.2 montrent l'influence du nombre de trajectoires m pour un même nombre total d'observations N sur le biais et l'erreur relatifs de $\mathbb{E}(T_\rho)$. Les différentes couleurs de points représentent la taille N du jeu de données tandis que la taille du point représente le nombre de trajectoires m utilisées pour estimer les paramètres ; plus le point est gros, plus m devient important. Afin d'améliorer la compréhension, ces figures ne montrent qu'un jeu de paramètres d'un processus gamma à la fois. Néanmoins, des résultats similaires sont observés pour les autres jeux de paramètres étudiés. On observe dans la Figure 2.1 que le biais et l'erreur relatifs tendent tous deux vers zéro lorsque le nombre de trajectoires m est supérieur à 100 ou lorsque la taille des données N est supérieure à 1000.

La Figure 2.3 montre pour deux nombres de trajectoires m différents et une même longueur de trajectoire n_j le comportement du biais et de l'erreur relatifs pour différents jeux de paramètres d'un PG. Dans cette figure, les cas $[m = 10 ; n_j = 100]$ et $[m = 100 ; n_j = 100]$ sont les seuls considérés. En raison de l'hétérogénéité des résultats pour le cas $(u = 1 ; v = 1 ; b = 0,5)$, ces derniers ne sont pas présentés sur la Figure 2.3.

Tableau 2.3 Résultats de simulation et d'estimation.

Param sim. (u, c, b)	PG traj. n_j	Nb de traj. m	Total N	\hat{u}	\hat{c}	\hat{b}	$\hat{\mathbb{E}}(T_\rho)$	$\hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))$	Biais rel.	Erreur rel.	EQM rel.
$(1; 1; 1)$	1000	100	10^5	1,0030	1,0127	0,9986	99,692	0,422	-0,31%	0,65%	0,007
		10	10^4	0,9994	0,9657	1,0069	100,241	4,18	0,24%	2,04%	0,021
		1	10^3	1,0100	1,2820	0,9682	90,926	38,6	-9,07%	6,83%	0,110
	100	1000	10^5	0,9978	0,9992	0,9999	99,879	0,0613	-0,12%	0,25%	0,003
		100	10^4	1,0044	1,0031	1,0000	100,146	0,608	0,15%	0,78%	0,008
		10	10^3	1,0007	1,1169	0,9830	96,856	5,87	-3,14%	2,50%	0,040
	$\mathbb{E}(T_\rho) = 100$	1	10^2	0,9914	1,1014	0,9872	95,419	53,9	-4,58%	7,69%	0,087
		1000	10^4	1,0000	0,9826	1,0026	100,533	0,0721	0,53%	0,27%	0,006
		100	10^3	0,9689	0,9821	0,9978	99,679	0,741	-0,32%	0,86%	0,009
	Moyenne	1000	10	10^2	0,7911	0,8041	0,9985	99,044	9,03	-0,96%	3,03%
1			10	1,6730	1,4473	1,0283	101,446	41,2	1,45%	6,33%	0,066
1000			10^4	1,0400	1,0634	0,9974	98,533	14,06	-1,47%	2,84%	0,035
100		100	10^5	0,4970	0,9667	1,0040	50,604	0,164	1,21%	0,80%	0,015
		10	10^4	0,5002	1,0382	0,9929	49,529	1,634	-0,94%	2,58%	0,027
		1	10^3	0,4529	0,7929	1,0254	51,679	16,96	3,36%	7,97%	0,089
$(0,5; 1; 1)$		1000	10^5	0,4960	0,9899	1,0004	50,032	0,024	0,06%	0,31%	0,003
		100	10^4	0,5047	1,0016	1,0008	50,233	0,243	0,47%	0,98%	0,011
		10	10^3	0,5167	0,9864	1,0048	51,415	2,469	2,83%	3,06%	0,042
$\mathbb{E}(T_\rho) = 50$		1	10^2	0,5746	1,0749	1,0156	50,294	18,07	0,59%	8,45%	0,085
	1000	10^4	0,5008	1,0027	0,9999	49,959	0,029	-0,08%	0,34%	0,003	
	100	10^3	0,5083	1,0133	0,9998	50,200	0,290	0,40%	1,07%	0,011	
Moyenne	10	10^2	0,4513	0,8628	1,0060	51,090	3,361	2,18%	3,59%	0,043	
	1	10	0,8067	2,4757	0,9369	41,207	13,55	-17,59%	8,93%	0,191	
	1000	10^4	0,5281	1,1096	0,9988	49,658	5,163	-0,68%	3,46%	0,047	

Tableau 2.4 Résultats de simulation et d'estimation (suite).

Param sim. PG (u, c, b)	Longueur traj. n_j	Nb de traj. m	Total N	\hat{u}	\hat{c}	\hat{b}	$\hat{\mathbb{E}}(T_\rho)$	$\hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))$	Biais rel.	Erreur rel.	EQM rel.
(1; 0,5; 1)	1000	100	10^5	0,9998	0,5143	0,9954	199,242	0,164	-0,38%	0,64%	0,007
		10	10^4	1,0403	0,5814	0,9795	199,469	16,59	-0,27%	2,04%	0,021
		1	10^3	1,0132	0,6644	0,9578	190,272	170,2	-4,86%	6,86%	0,081
$\mathbb{E}(T_\rho) = 200$	100	1000	10^5	1,0044	0,5013	1,0004	199,885	0,268	-0,06%	0,28%	0,003
		100	10^4	1,0227	0,4995	1,0032	201,272	2,651	0,64%	0,90%	0,010
		10	10^3	1,1241	0,4913	1,0183	207,503	25,04	3,75%	2,41%	0,045
Moyenne	10	1	10^2	1,1544	0,6520	0,9824	194,201	217,6	-2,90%	7,59%	0,079
		1000	10^4	1,0032	0,5000	1,0005	200,173	0,314	0,09%	0,28%	0,003
		100	10^3	0,9930	0,5208	0,9933	197,534	3,150	-1,23%	0,90%	0,015
(1; 1; 2)	1000	100	10^5	1,0068	1,0439	1,9937	9,891	0,000	-1,09%	0,15%	0,011
		10	10^4	0,9850	1,0137	1,9950	9,914	0,002	-0,86%	0,47%	0,010
		1	10^3	1,0158	1,0254	1,9988	9,967	0,021	-0,33%	1,46%	0,015
$\mathbb{E}(T_\rho) = 10$	100	1000	10^5	0,9095	1,1267	1,9633	9,361	0,000	-6,39%	0,05%	0,064
		100	10^4	0,9048	1,1300	1,9619	9,337	0,000	-6,63%	0,16%	0,066
		10	10^3	0,9197	1,1667	1,9590	9,294	0,002	-7,06%	0,52%	0,071
Moyenne	10	1	10^2	0,8746	1,0223	1,9753	9,510	0,025	-4,90%	1,67%	0,051
		1000	10^4	0,0600	0,3149	1,7194	5,553	0,000	-44,47%	0,24%	0,445
		100	10^3	0,0603	0,3189	1,7178	5,535	0,002	-44,65%	0,77%	0,447
Moyenne	1	10	10^2	0,0626	0,3315	1,7178	5,532	0,018	-44,68%	2,39%	0,447
		1	10	0,0480	0,2549	1,7171	5,526	0,232	-44,74%	8,72%	0,450
		0,6225	0,7954	1,8836	8,1291	0,028	-18,71%	1,51%	0,189		

Tableau 2.5 Résultats de simulation et d'estimation (suite).

Param sim. PG (u, c, b)	Longueur traj. n_j	Nb de traj. m	Total N	\hat{u}	\hat{c}	\hat{b}	$\hat{\mathbb{E}}(T_\rho)$	$\hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))$	Biais rel.	Erreur rel.	EQM rel.
(2; 1; 1)	1000	100	10^5	2,0023	1,0143	0,9980	199,539	1,025	-0,23%	0,51%	0,006
		10	10^4	1,9727	0,9203	1,0133	199,795	10,10	-0,10%	1,59%	0,016
		1	10^3	2,0229	0,8472	1,0233	210,842	108,3	5,42%	4,93%	0,075
	100	1000	10^5	1,9924	0,9903	1,0010	200,097	0,140	0,05%	0,19%	0,002
		100	10^4	1,9946	0,9827	1,0027	200,075	1,40	0,04%	0,59%	0,006
		10	10^3	2,0455	0,9499	1,0095	204,699	14,05	2,35%	1,83%	0,030
	$\mathbb{E}(T_\rho) = 200$	1	10^2	1,8063	0,7114	1,0366	208,834	156,0	4,42%	5,98%	0,076
		1000	10^4	2,0089	1,0085	0,9994	199,856	0,157	-0,07%	0,20%	0,002
		100	10^3	2,0495	1,0033	1,0024	201,696	1,565	0,85%	0,62%	0,011
	Moyenne	10	10	10^2	2,0618	0,9427	1,0154	201,508	15,32	0,75%	1,94%
1			10	2,1741	1,2948	0,9754	191,094	142,7	-4,45%	6,25%	0,074
1000			10^4	2,0119	0,9696	1,0070	201,6394	40,97	0,82%	2,24%	0,029
(1; 2; 1)	1000	100	10^5	0,9937	1,9817	1,0002	50,114	0,101	0,23%	0,63%	0,007
		10	10^4	0,9831	2,1101	0,9899	48,449	0,972	-3,10%	2,03%	0,037
		1	10^3	1,0664	1,9562	1,0116	52,065	9,93	4,13%	6,05%	0,075
	100	1000	10^5	1,0022	1,9970	1,0004	50,099	0,012	0,20%	0,22%	0,003
		100	10^4	0,9960	1,9291	1,0046	50,710	0,127	1,42%	0,70%	0,016
		10	10^3	1,0636	2,0378	1,0063	50,910	1,22	1,82%	2,17%	0,029
	$\mathbb{E}(T_\rho) = 50$	1	10^2	0,9912	1,6075	1,0272	55,284	14,73	10,57%	6,94%	0,131
		1000	10^4	0,9880	1,9679	1,0003	50,139	0,015	0,28%	0,24%	0,004
		100	10^3	1,1366	2,2871	0,9991	49,861	0,127	-0,28%	0,72%	0,008
	Moyenne	10	10	10^2	1,1288	2,3171	0,9973	49,236	1,257	-1,53%	2,28%
1			10	0,7826	2,0000	0,9626	45,115	16,54	-9,77%	9,02%	0,127
1000			10^4	1,0120	2,0174	1,0000	50,1801	4,093	0,36%	2,82%	0,042

Tableau 2.6 Résultats de simulation et d'estimation (suite).

Param sim. PG (u, c, b)	Longueur traj. n_j	Nb de traj. m	Total N	\hat{u}	\hat{c}	\hat{b}	$\hat{\mathbb{E}}(T_\rho)$	$\hat{\mathbb{V}}(\hat{\mathbb{E}}(T_\rho))$	Biais rel.	Erreur rel.	EQM rel.
		1000	10^5	0,9984	0,4879	0,5893	8344	12273	-16,56%	1,33%	0,166
(1;1;0,5)	100	100	10^4	1,0249	0,4278	0,6075	8260	115032	-17,39%	4,11%	0,177
		10	10^3	1,0576	0,5358	0,5737	10021	1960735	0,21%	13,97%	0,140
$\mathbb{E}(T_\rho) = 10^3$		1	10^2	0,8291	0,4222	0,6098	5762	4982072	-42,37%	38,73%	0,479
	10	10	10^2	1,0330	0,3123	0,6440	8168	2090302	-18,31%	17,80%	0,233
		1	10	1,8640	0,8704	0,5902	8891	15359164	-11,08%	44,07%	0,407
Moyenne				1,1345	0,5094	0,6024	8241	4086596	-17,6%	20,0%	0,239
Moyenne globale				1,0545	1,0368	1,1000	770,6	335904	-4,34%	4,06%	0,08

On soulignera que les résultats présentés dans les Tableaux 2.3 à 2.6 pour les cas où $n_j = 1$ et les faibles valeurs de m sont plus ou moins reproductibles et doivent être considérés avec un recul suffisant. En effet, comme la méthode sous-jacente employée pour la simulation des trajectoires (méthode *GSS* présentée en section 1.3.5.3) réalise des tirages d'une distribution gamma, pour des combinaisons de petites valeurs de n_j et m , N peut s'avérer trop faible pour donner une représentation correcte de la réalité : les conditions du théorème central limite ne sont pas vérifiées. Dans certains cas, cela peut mener à des résultats contre-intuitifs comme un biais relatif plus faible pour $n_j = 10$ et $m = 1$ que pour $n_j = 100$ et $m = 10$ (voir le cas pour des paramètres de PG $(1; 1; 1)$ dans le Tableau 2.3 par exemple).

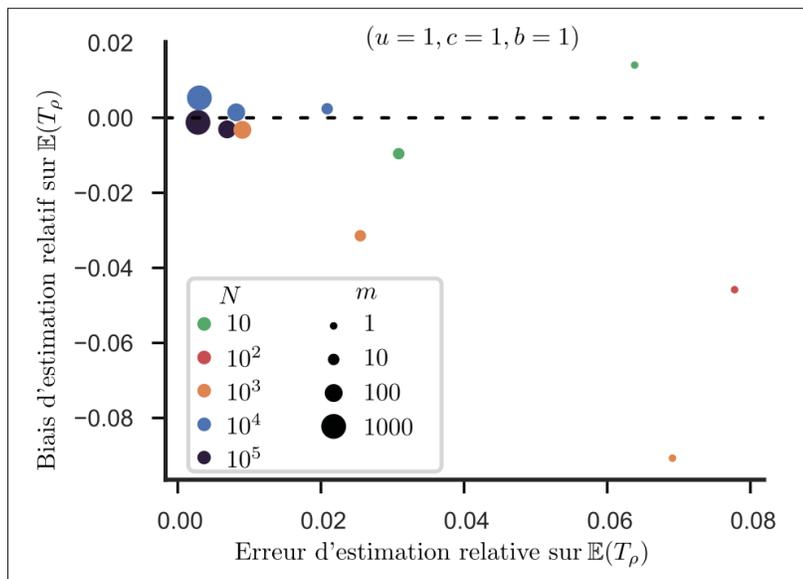


Figure 2.1 Comparaison du biais relatif et de l'erreur relative d'estimation pour un nombre de trajectoires m et pour des paramètres de PG : $(1; 1; 1)$.

2.2.1 Analyse des résultats

Tout d'abord, on notera que des difficultés d'ordre numérique ont été rencontrées pour les cas avec $b = 0,5$ (et plus généralement lorsque $b \in [0,5; 0,75]$ comme illustré dans le Tableau 2.6 avec certaines combinaisons absentes pour la longueur de trajectoire n_j et le nombre de trajectoires m). Nous avons en effet décidé de ne pas présenter les résultats (incohérents) issus d'une procédure

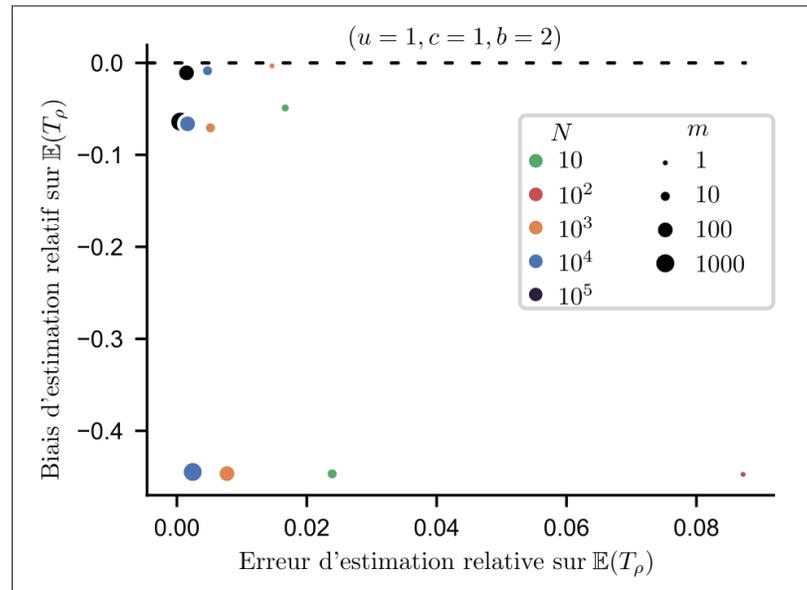


Figure 2.2 Comparaison du biais relatif et de l'erreur relative d'estimation pour un nombre de trajectoires m et pour des paramètres de PG : $(1; 1; 2)$.

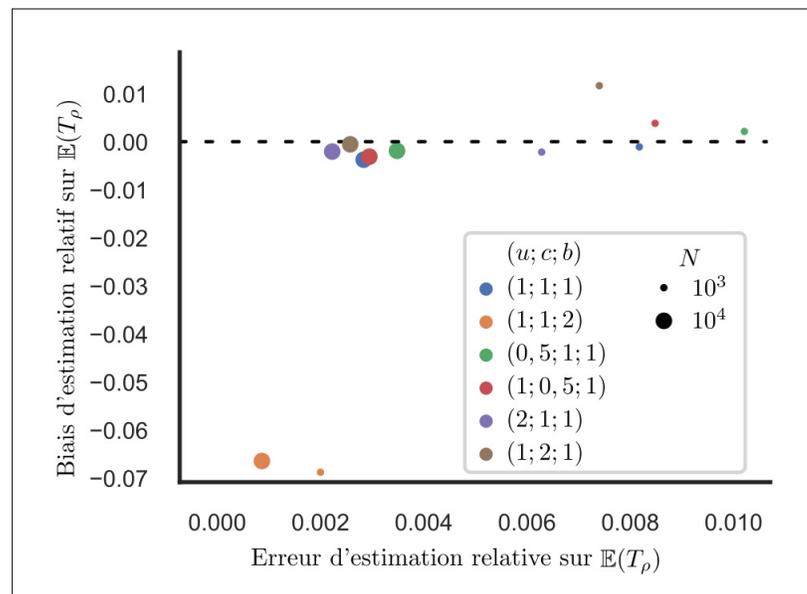


Figure 2.3 Comparaison du biais relatif et de l'erreur relative d'estimation pour différentes tailles d'échantillon N et valeurs de paramètres de PG.

d'optimisation ayant échoué. En regardant en détail la forme de la fonction log-vraisemblance $\ell(\theta)$, nous avons remarqué que cette dernière présentait une région presque "plate" pour les faibles valeurs de paramètre b , donnant des solutions locales hasardeuses. Par ailleurs, pour les autres résultats, issus d'une optimisation ayant abouti avec succès, on observe un biais relatif, à la fois sur les paramètres et l'espérance du temps d'atteinte, bien plus important que pour les autres combinaisons de paramètres dont $b \neq 0,5$. Ces résultats montrent que la méthode de MV est inefficace pour des valeurs de b autour de 0,5.

À partir des résultats des Tableaux 2.3 à 2.5 et de la Figure 2.1, on observe que pour un nombre total d'observations N , le biais et l'erreur relatifs diminuent lorsque le nombre de trajectoires m augmente (avec un facteur deux ou trois dans plusieurs cas). Cette tendance est observée pour presque tous les jeux de paramètres dont $b = 1$, correspondant au cas particulier de processus gamma homogène (PGH). Lorsque $b = 1$, les incréments de dégradation sont alors identiquement distribués, ce qui peut expliquer les bonnes performances de l'EMV. Au contraire, lorsque $b \neq 1$ (PGNH), on observe des résultats opposés (voir Figure 2.2). L'estimation du paramètre b semble être plus difficile lorsque les conditions suivantes sont réunies : la valeur du paramètre b est "éloignée" de un (1) et les longueurs de trajectoires n_j sont "courtes" (par exemple pour $10 \leq n_j \leq 100$, voir Tableau 2.4). Pour ces valeurs de paramètres, une variance plus importante est alors constatée, en raison du manque d'information engendré par le faible nombre d'observations n_j . De même, un ordre de grandeur semblable est observé pour l'EQM relatif entre les cas PGH ($b = 1$) et PGNH ($b \neq 1$). Pour $b = 1$, la méthode MV montre de meilleures performances lorsque le nombre de trajectoires augmente tout en conservant un nombre d'observations N identique. Du point de vue de l'ingénieur, afin de construire un modèle robuste, celui-ci devrait privilégier, lorsque possible, plusieurs répétitions d'une trajectoire de dégradation plutôt qu'un plus grand nombre d'observations n_j par trajectoires pour un même nombre total d'observations N .

À partir de la Figure 2.3, on observe que l'erreur et le biais relatifs (pour un jeu de paramètres de PG donné) semblent suivre la même tendance lorsque la taille des données N augmente. En effet, les points sont regroupés selon le même motif pour $N = 10^4$ et $N = 10^3$. Aussi, on

constate que certains jeux de paramètres montrent des biais et erreurs relatifs plus faibles que d'autres, peu importe le nombre total d'observations N . Ces résultats montrent que la méthode MV présente de meilleures performances pour certains jeux de paramètres de PG.

Bien que les résultats pour les jeux de paramètres ($u = 1; c = 1; b = 0,5$) aient été écartés (en raison des difficultés d'optimisation de la méthode MV rencontrées), la moyenne du biais relatif de $\hat{E}(T_\rho)$ pour l'ensemble du plan d'expérience est de l'ordre de -3% , ce qui montre que la méthode MV semble sous-estimer l'espérance du temps d'atteinte. Ces résultats, d'un point de vue fiabilité, peuvent être considérés comme conservatifs : le composant est considéré défaillant avant qu'il ne le soit réellement.

On notera également que la moyenne globale du biais relatif est de l'ordre de $4\% \neq 0$, ce qui montre que la propriété asymptotique de l'EMV n'est pas vérifiée ; néanmoins, cette tendance théorique est observée lorsque N augmente.

2.3 Estimation des paramètres d'un processus gamma en deux étapes

En raison des difficultés numériques rencontrées lors de l'estimation des paramètres d'un processus gamma non homogène (comme c'est d'ailleurs le cas dans la section précédente), une méthode alternative à celle décrite dans la section 1.3.5.2 est étudiée. En effet, la méthode initiale repose sur la détermination des racines d'un système d'équations à dérivées partielles. Celles-ci font apparaître trois (3) inconnues correspondant aux paramètres des fonctions de forme et d'échelle d'un processus gamma (u, c, b). L'un des paramètres pouvant être exprimé en fonction des deux autres (voir équation 1.18), le système se ramène à deux degrés de liberté seulement. Il est alors possible d'estimer numériquement les paramètres restants à l'aide d'une méthode d'optimisation. Dans notre étude, celle retenue repose sur une version modifiée de Powell incluse dans la librairie `Scipy` de Python (voir Fletcher & Powell (1963); Scipy (2019)).

Lors de nos expérimentations numériques, nous avons constaté, en fonction de la valeur "vraie" des paramètres, que le présent algorithme d'optimisation pouvait s'avérer très sensible aux valeurs d'initialisation (valeurs supposées proches des valeurs "vraies" des paramètres qui sont

renseignées pour le démarrage de l'optimisation). Ainsi, il était parfois difficile d'obtenir une optimisation qui aboutit et donc des valeurs cohérentes des paramètres estimés. À cause de ces inconvénients, nous avons décidé d'étudier une méthode alternative permettant de pallier ces problèmes numériques. Cette nouvelle méthode consiste toujours à estimer les paramètres d'un processus gamma non homogène pour une trajectoire ou un ensemble de trajectoires donné, mais cette fois en deux temps.

La première étape repose sur l'hypothèse que la trajectoire de dégradation suit globalement une fonction puissance de la forme $y : x \mapsto a \times x^b$. On cherche alors à estimer la valeur de b . Pour cela, une régression linéaire est réalisée à partir de la transformation logarithmique de la précédente fonction, s'écrivant : $\log(y) = \log(a) + b \times \log(x)$. Il est alors possible d'estimer la puissance b par la méthode des moindres carrés, notée \hat{b}_{mc} (on utilise dans notre cas la méthode `optimize least squares` de la librairie `Scipy`). Une fois b estimé, la deuxième étape consiste à estimer c à l'aide des équations (2.1) et (2.2) issues de la vraisemblance, analogues à celles présentées en section 1. Le système se ramène alors à une équation à une seule inconnue rendant l'optimisation plus aisée (car l'équation ne présente plus qu'un degré de liberté). La résolution de l'équation (2.2) se fait alors à l'aide de la précédente optimisation et ensuite l'expression (2.1) permet de déduire l'estimation du paramètre u .

$$\hat{u}_{MV} = \hat{c}_{MV} \frac{\sum_{i=1}^n t_i^{\hat{b}_{mc}} - t_{i-1}^{\hat{b}_{mc}}}{\sum_{i=1}^n \delta_i} \quad (2.1)$$

$$\sum_{i=1}^n \left[t_i^{\hat{b}_{mc}} - t_{i-1}^{\hat{b}_{mc}} \right] \left\{ \psi \left(\hat{c}_{MV} \left[t_i^{\hat{b}_{mc}} - t_{i-1}^{\hat{b}_{mc}} \right] \right) - \log(\delta_i) \right\} = \sum_{j=1}^m (t_{n,j}^{\hat{b}_{mc}}) \log(\hat{u}_{MV}), \quad (2.2)$$

N.B. 1 : Nous utilisons jusqu'alors cette approche, qui consiste à estimer la valeur du paramètre b par la méthode des moindres carrés, pour obtenir une valeur initiale du paramètre dans l'étape d'optimisation (qui en cas de succès, estime les paramètres (triplet) maximisant la fonction de vraisemblance). Lors du traitement des données issues de centrale (*cf.* Chapitre 4), nous avons été confrontés à des échecs d'optimisation malgré l'utilisation de valeurs initiales considérées

comme adéquates (c.à.d. : des valeurs (très) proches des valeurs estimées à la suite d'une optimisation réussie). Ce problème numérique (une très forte sensibilité aux valeurs initiales) a alors été levé en utilisant la méthode d'estimation en deux temps.

N.B. 2 : la méthode `optimize least squares` permet également de déterminer l'inverse de la matrice hessienne, exprimée aux valeurs de l'optimisation. Cette information permet alors de déterminer un intervalle de confiance sur le paramètre b .

Afin d'étudier les performances de la méthode d'estimation présentée plus haut, le plan d'expérience présenté dans le Tableau 2.7 a été réalisé.

Tableau 2.7 Plan d'expérience.

Paramètres	Valeurs	Commentaires
Paramètres PG ($u; c; b$)	(16; 8; 1)	Variance moyenne
	(16; 8; 1,2)	
	(16; 8; 2)	
	(100;10;1)	Variance faible
	(100;10;1,2)	
	(100;10;2)	
	(4;8;1)	Variance importante
	(4;8;1,2)	
	(4;8;2)	
Longueur de trajectoire n	256	
Nombre de trajectoires k	1000	
Temps final de simulation T	40	

La longueur des trajectoires et leur nombre ont été choisis pour se rapprocher des conditions asymptotiques (théorie de la loi des grands nombres). Les valeurs des paramètres du processus gamma simulé ont quant à elles été prises pour présenter 3 niveaux différents de variance (de faible à important) et 3 valeurs de b différentes (trajectoire homogène/linéaire : 1, légèrement convexe : 1,2 et très convexe : 2). Le choix de la valeur de $b = 1,2$ pour ce plan d'expérience a

été motivé suite aux observations réalisées sur les trajectoires de dégradation issues des données de laboratoire et de centrale (cf. Chapitres 3 et 4). En effet, les trajectoires observées présentent une légère convexité dont la modélisation par une fonction puissance montre une valeur de b proche de 1,2.

Afin de valider cette méthode d'estimation en deux étapes, des trajectoires suivant un processus gamma ont été simulées selon les paramètres du Tableau 2.7. Pour cela, la méthode du *Gamma Sequential Sampling* (voir section 1.3.5.3) a été utilisée. Les Figures 2.4 et 2.5 montrent l'enveloppe des trajectoires simulées pour les paramètres de processus suivants : $(u; c; b) = (16; 8; 1)$, $(16; 8; 1,2)$ et $(16; 8; 2)$. La courbe en rouge représente le tracé d'une fonction puissance dont l'espérance est identique à l'espérance théorique du processus gamma simulé. Cette courbe constitue un repère pour analyser l'enveloppe des trajectoires simulées. Les simulations sont réalisées sur l'intervalle de temps fixe $[0, T]$ dont les pas de temps sont homogènes entre eux. L'espérance $\mathbb{E}(X(T))$ et la variance $\mathbb{V}(X(T))$ théoriques de la dégradation $X(t)$ à l'instant T sont données dans le Tableau 2.9.

À partir du Tableau 2.9 et des Figures 2.4 et 2.5, on peut observer l'influence du paramètre b sur la taille de l'enveloppe des simulations : plus b est grand, plus la variance est importante (voir expression 1.15) et donc l'enveloppe des simulations est plus large (noter que l'échelle sur l'axe des abscisses est différente d'une Figure à l'autre). Aussi, on peut souligner que l'enveloppe ne présente pas de dissymétrie flagrante par rapport à la courbe rouge sur les figures. Cette observation permet entre autres de montrer les qualités de la méthode *GSS* pour la simulation de trajectoires de processus gamma.

Afin d'étudier les performances du couplage des méthodes *GSS* et d'estimation des paramètres en deux temps (indiquée *MC*), quelques indicateurs statistiques ont été calculés. Aussi, à des fins de comparaison, cette dernière méthode est confrontée à la méthode d'estimation reposant exclusivement sur la fonction de vraisemblance (indiquée *MV* et présentée en section 1.3.5.2). On note par la suite $\hat{\mathbb{E}}(X(T))$ et $\hat{\mathbb{V}}(X(T))$ respectivement l'espérance empirique des niveaux de dégradation $x_i(T)$ des trajectoires i , $1 \leq i \leq k$ à l'instant T et la variance empirique de ces

mêmes niveaux de dégradation. Le Tableau 2.8 présente l'expression mathématique de $\hat{\mathbb{E}}(X(T))$ et du biais relatif. On note aussi, \hat{u} , \hat{c} et \hat{b} la moyenne empirique des \hat{u}_i , \hat{c}_i et \hat{b}_i , $1 \leq i \leq k$, les estimation de $(u; c; b)$ à partir de la i -ème trajectoire respectivement. Le calcul du biais relatif est détaillé dans le Tableau 2.8 pour l'espérance mais l'expression reste valable pour les différentes variables de l'étude. Ces indicateurs ont été évalués pour les deux méthodes d'estimation comparées.

N.B. : on notera que les indicateurs sur l'espérance, la variance et leur biais relatif associé, caractérisent uniquement les performances de la méthode de simulation *GSS*, contrairement aux autres indicateurs du Tableau 2.10 qui reflètent les performances du couplage simulation/estimation.

Tableau 2.8 Indicateurs statistiques de performance.

	Expression mathématique
Espérance empirique	$\hat{\mathbb{E}}(X(T)) = \frac{\sum_{i=1}^k x_i(T)}{k}$
Biais relatif	$\frac{\hat{\mathbb{E}}(X(T)) - \mathbb{E}(X(T))}{\mathbb{E}(X(T))}$

Les résultats des calculs statistiques sont présentés dans le Tableau 2.10 pour les différents jeux de paramètres étudiés. Tout d'abord, on peut noter les bonnes performances de la méthode de simulation *GSS* en ce qui concerne l'espérance finale : en moyenne, les trajectoires simulées présentent un biais relatif sur l'espérance finale (pour $t = T$) de moins de 0,5%. Par contre, en ce qui concerne sa capacité à reproduire la variabilité d'un processus gamma, ses performances sont moins bonnes avec un biais relatif de la variance qui peut avoisiner dans certains cas les 4%.

Pour ce qui est des méthodes d'estimation étudiées, on observe qu'en ce qui concerne le paramètre b , celles-ci ont tendance à sous-estimer systématiquement la valeur de cette puissance lorsqu'elle est égale à 2 (et au delà, même si les résultats ne sont pas présentés dans cette étude). On peut

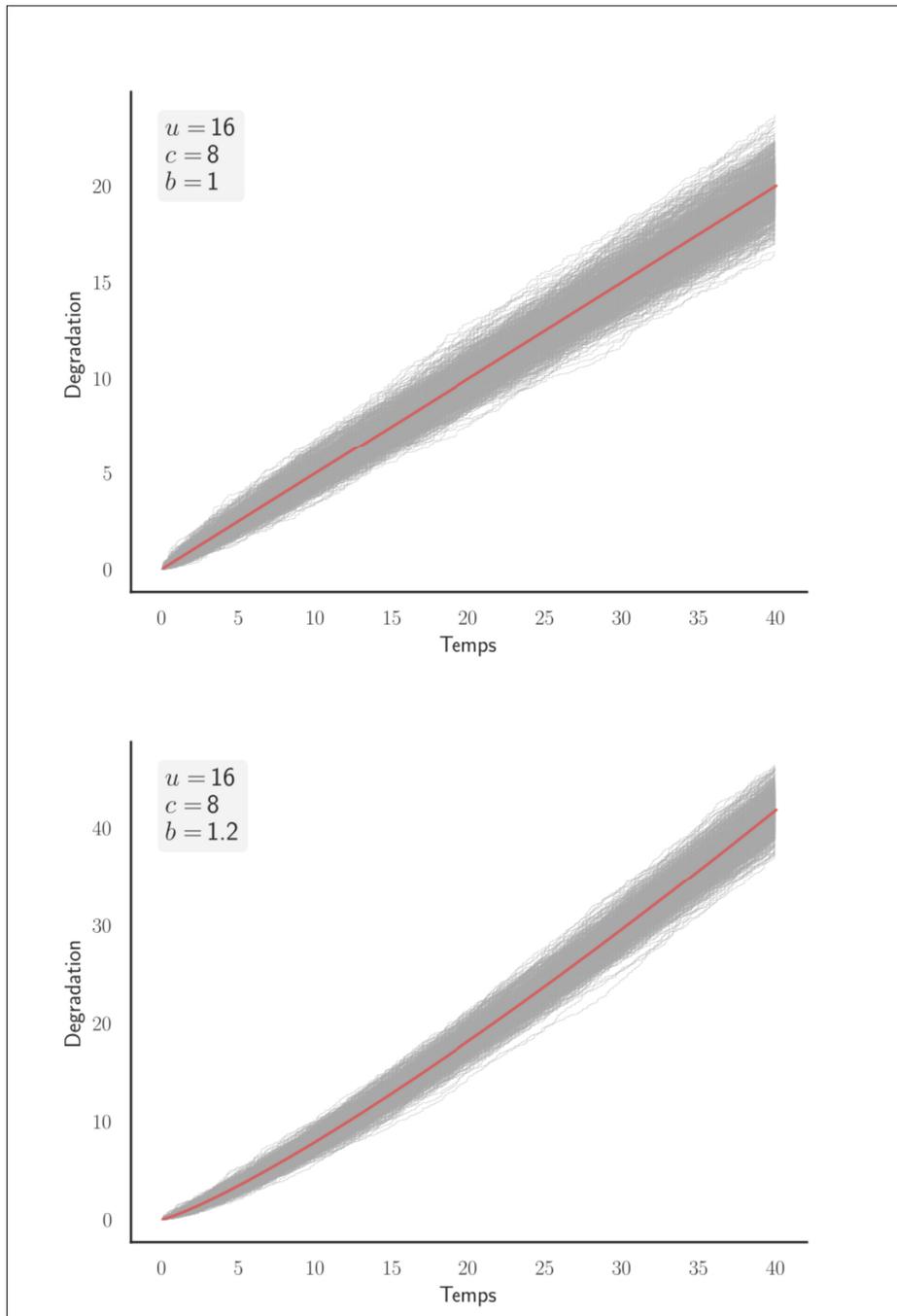


Figure 2.4 Enveloppe des simulations d'un processus gamma.

par ailleurs noter que le biais relatif associé à l'espérance empirique de $\mathbb{E}(X(T))$ ne présente pas de tendance particulière pouvant expliquer cette observation (par exemple : un biais relatif toujours négatif, donnant globalement des trajectoires dont la réalisation à l'instant final T est en

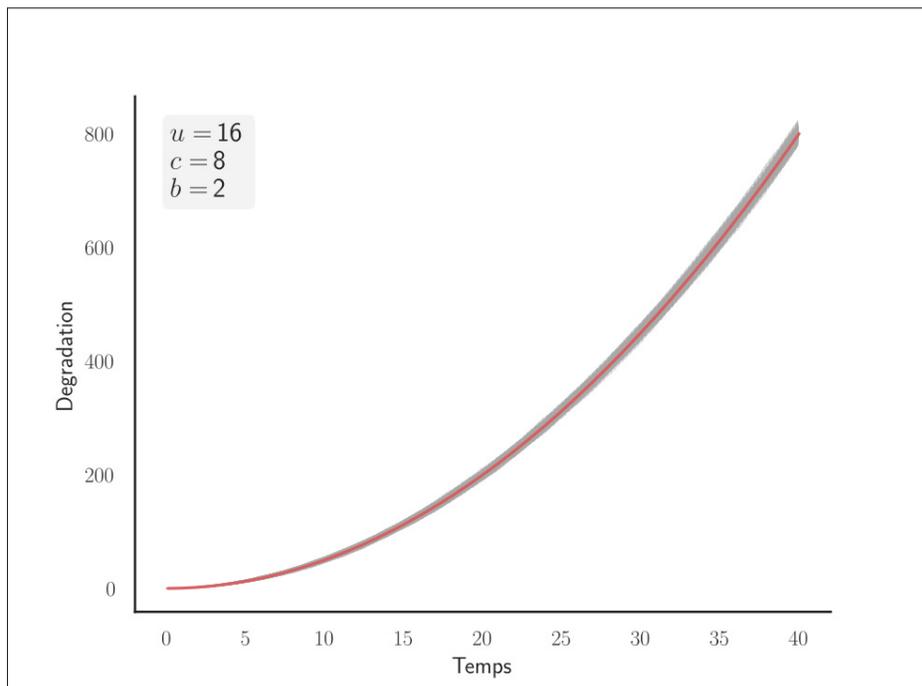


Figure 2.5 Enveloppe des simulations d'un processus gamma (suite).

dessous de la courbe rouge théorique). Néanmoins, malgré cette sous-estimation du paramètre b , le biais relatif associé à ce paramètre reste homogène aux autres valeurs pour les cas où $b \neq 2$. La méthode MV présente un biais relatif sur \hat{b} très faible pour les trajectoires avec pour paramètre $b = 1$ et inférieur à la méthode MC . Globalement, on n'observe pas de biais relatif sur \hat{b} supérieur à 1,2% pour la méthode MC et à 0,9% pour la méthode MV , ce qui sous-tend le caractère robuste de ces deux méthodes d'estimation et encourage leur utilisation. Lorsqu'on s'intéresse à l'estimation du paramètre c , ces méthodes montrent plus de difficultés avec un biais relatif sur \hat{c} pouvant avoisiner 10% pour $b = 2$. On notera d'ailleurs que dans ce dernier cas, les deux méthodes semblent plus sensibles aux erreurs d'estimation sur ce paramètre quant il s'agit d'estimer ensuite c que lorsque $b \neq 2$. Ces observations peuvent être aussi réalisées pour la plupart à l'aide de la distribution des estimations pour chaque paramètre représentée dans la Figure 2.6. Pour chaque distribution, est représentée en rouge la valeur du paramètre de la simulation et en trait pointillé la moyenne des estimations du même paramètre. D'une manière générale, on ne peut, à partir de cette étude, faire ressortir une méthode plus performante qu'une

autre sur le jeu de paramètres testé. Néanmoins, pour les jeux de paramètres avec $b = 1, 2$, illustrant la convexité observée des trajectoires réelles, on peut souligner des biais relatifs sur les trois paramètres systématiquement plus faibles pour la méthode *MC*.

N.B. : en raison du caractère déductif de \hat{u} à partir des deux autres paramètres, on ne commentera pas les résultats afférent à ce paramètre.

Comme précisé plus haut dans la présente section, les trajectoires de dégradation réelles, issues des données de laboratoire ou de centrale, présentent pour certaines une convexité pouvant être modélisée par une fonction puissance dont le paramètre b est proche de 1,2. Dans d'autres cas, la trajectoire semble plutôt présenter un comportement linéaire. Pour ces raisons, les conclusions tirées sur les performances de la méthode d'estimation en deux temps portent surtout sur les cas d'étude où $b = 1$ ou 1,2. Comme souligné, l'estimation du paramètre b dans ces deux cas présente un biais relatif inférieur à 1%. Les biais relatifs sur \hat{c} et \hat{u} sont respectivement inférieurs à 3% et 1,1%. Au regard de ces résultats, et des hypothèses de modélisation qui seront prises dans la suite du document, la méthode d'estimation des paramètres en deux temps sera employée dans le Chapitre 4.

Tableau 2.9 Test de sensibilité de la méthode de simulation *GSS*.

Param. sim. ($u; c; b$)	$\mathbb{E}(X(T))$	$\mathbb{V}(X(T))$	$\hat{\mathbb{E}}(X(T))$	$\hat{\mathbb{V}}(X(T))$	Biais relatif	
					$\hat{\mathbb{E}}$	$\hat{\mathbb{V}}$
(16; 8; 1)	20	1,25	19,93	1,23	-0,36%	-1,67%
(16; 8; 1,2)	41,83	2,61	41,84	2,70	0,03%	3,31%
(16; 8; 2)	800	50	799,84	50,07	-0,02%	0,14%
(100; 10; 1)	4	0,04	3,98	0,04	-0,48%	-1,25%
(100; 10; 1,2)	8,37	0,08	8,35	0,09	-0,21%	3,48%
(100; 10; 2)	160	1,60	160,05	1,63	0,03%	1,62%
(4; 8; 1)	80	20	79,63	20,41	-0,46%	2,04%
(4; 8; 1,2)	167,30	41,83	166,96	40,18	-0,20%	-3,94%
(4; 8; 2)	3200	800	3199,97	774,54	0,00%	-3,18%

Tableau 2.10 Comparaison de deux méthodes d'estimation.

Param. sim. ($u; c; b$)	Méthode en deux étapes					Méthode MV						
	\hat{u}_{MC}	\hat{c}_{MC}	\hat{b}_{MC}	\hat{u}_{MC}	Biais relatif \hat{c}_{MC}	\hat{b}_{MC}	\hat{u}_{MV}	\hat{c}_{MV}	\hat{b}_{MV}	Biais relatif \hat{c}_{MV}	\hat{b}_{MV}	
(16; 8; 1)	16,25	8,18	1,01	1,59%	2,25%	0,63%	16,29	8,21	1,00	1,80%	2,58%	0,09%
(16; 8; 1,2)	16,18	8,44	1,19	1,10%	5,50%	-0,44%	16,21	8,49	1,19	1,29%	6,17%	-0,75%
(16; 8; 2)	16,24	8,85	1,98	1,51%	10,62%	-1,03%	16,38	8,74	1,98	2,35%	9,30%	-0,77%
(100; 10; 1)	101,60	10,25	1,00	1,60%	2,48%	0,32%	101,86	10,24	1,00	1,86%	2,43%	0,07%
(100; 10; 1,2)	101,74	10,60	1,19	1,74%	6,04%	-0,55%	101,91	10,69	1,19	1,91%	6,88%	-0,85%
(100; 10; 2)	101,45	11,09	1,98	1,45%	10,92%	-1,07%	102,37	10,93	1,98	2,37%	9,28%	-0,76%
(4; 8; 1)	4,04	8,15	1,00	1,05%	1,90%	0,49%	4,05	8,16	1,00	1,27%	2,04%	0,06%
(4; 8; 1,2)	4,06	8,45	1,19	1,46%	5,67%	-0,43%	4,07	8,51	1,19	1,65%	6,32%	-0,76%
(4; 8; 2)	4,04	8,87	1,98	0,94%	10,82%	-1,12%	4,07	8,74	1,98	1,83%	9,22%	-0,83%

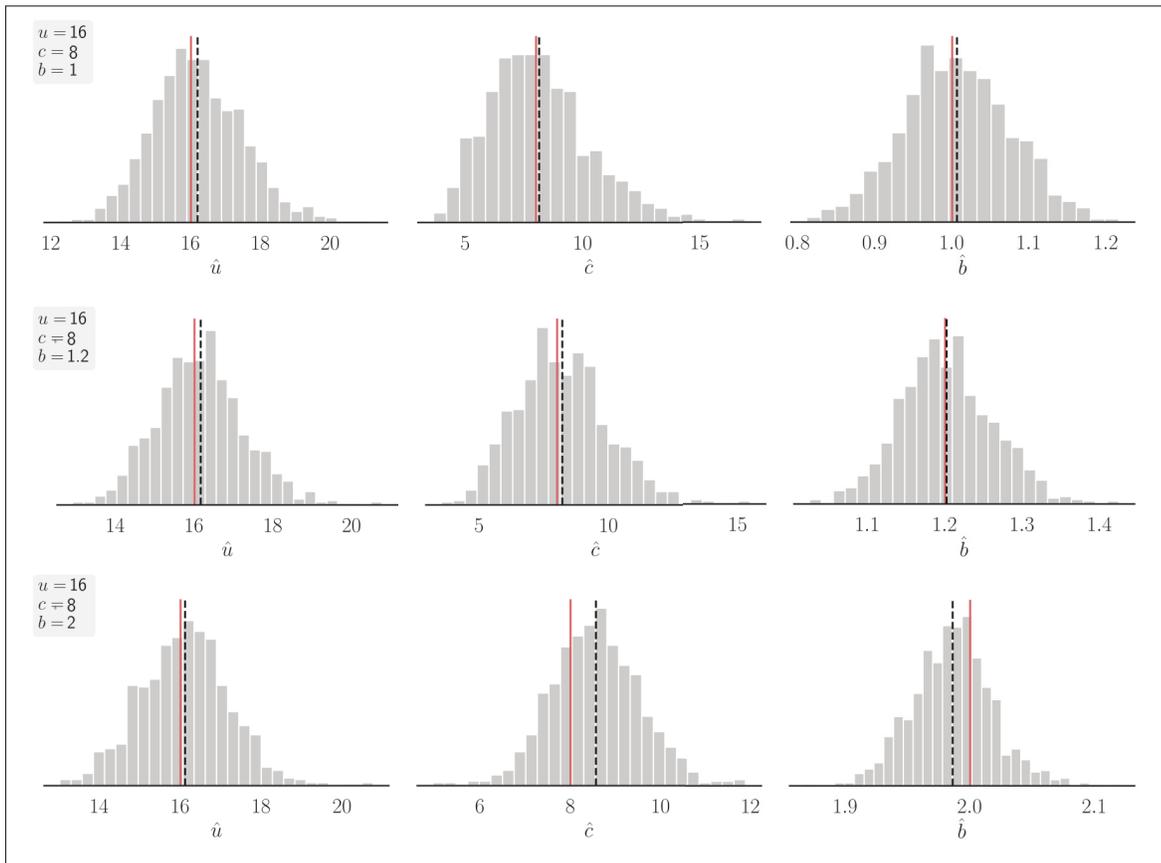


Figure 2.6 Distributions des estimations du processus gamma par la méthode en deux étapes.

2.4 Conclusion

Ce chapitre présente une étude de performance de la méthode d'estimation par maximum de vraisemblance des paramètres d'un processus gamma, ainsi qu'une variante de cette méthode permettant d'éliminer certains problèmes d'optimisation.

Cette étude de la méthode MV propose une estimation à la fois des paramètres d'un PG et de l'espérance du temps d'atteinte associé. Il est montré, que pour le plan d'expérience choisi, les paramètres u et c ont un impact plus faible sur la qualité d'estimation du MV que le nombre de trajectoires m ou encore que le paramètre $b \neq 1$ (processus gamma non homogène). Pour les cas où b est dans l'intervalle $[0; 0,75]$, des méthodes alternatives d'estimation, comme la méthode

des moments ou une approche bayésienne devraient être explorées. Le RMSE relatif pourrait être utilisé à titre de comparaison des méthodes.

Une étude similaire pourrait être réalisée pour des cas où l'analyste connaît à l'avance la valeur du paramètre b (à partir d'un modèle physique par exemple ou grâce à la méthode d'estimation en deux temps présentée plus haut). Ce type de cas permettrait de réduire le nombre de degrés de liberté (passant de 3 paramètres à 2 seulement), mais aussi d'avoir recours à différentes méthodes d'estimation et de les comparer. Aussi, afin d'estimer les intervalles de confiance sur les paramètres ainsi que sur l'espérance du temps d'atteinte, dans les cas de "petits" jeux de données (pour $N \leq 10^3$ par exemple), des méthodes de ré-échantillonnage devraient être considérées (voir 1.5.3) en remplacement de la méthode delta. En effet, pour rappel, ces techniques permettent en plus d'évaluer d'éventuels biais sur les estimations.

La variante de la méthode d'estimation par maximum de vraisemblance présentée dans ce chapitre sera employée pour les données de centrale présentées dans le Chapitre 4. En plus de résoudre certains problèmes numériques, il a été constaté que l'estimation des paramètres par cette méthode réduisait drastiquement les temps de calcul. Ce constat s'explique notamment par la simplification des équations mises en jeu (réduction du nombre de degrés de liberté). Pour les jeux de données importants, comme c'est le cas pour les données en centrale, cette rapidité d'estimation représente un certain avantage.

CHAPITRE 3

MODÉLISATION DES DONNÉES DE LABORATOIRE

3.1 Introduction

Dans ce chapitre, nous nous intéressons aux données de dégradation obtenues à partir de tests de cavitation érosive réalisés en laboratoire. Les données présentées seront par la suite modélisées à l'aide d'un processus gamma. Cette première étape, dans la modélisation de données de dégradation réelles, peut être considérée comme une preuve de concept avant d'aborder dans un deuxième temps les données issues du dispositif de suivi en continu de la cavitation (installé sur des roues de turbines opérées industriellement). Un des principaux avantages des données de laboratoire réside dans leurs conditions d'obtention : en général, les essais sont conduits dans un environnement contrôlé (au regard de l'environnement dans lequel évolue une machine industrielle) avec des instruments faisant preuve d'une précision (résolution) adaptée. Ceci permet de limiter le nombre de paramètres (covariables) ayant une influence sur la dégradation mais aussi les incertitudes de mesure par exemple.

D'un point de vue industriel, les essais de cavitation en laboratoire ont pour but d'étudier la résistance d'un matériau (ou d'un revêtement) exposé à de la cavitation. Pour cela, des échantillons du matériau étudié sont fabriqués dans des conditions identiques (ou proches) à celles de la roue de turbine. Ils sont ensuite exposés à de la cavitation et l'érosion résultante est quantifiée à l'aide de la mesure, à certains instants, de la perte de masse de l'échantillon.

Dans un premier temps, deux dispositifs de laboratoire permettant d'étudier l'érosion par cavitation seront présentés puis une analyse de durée de vie est réalisée pour deux matériaux employés dans la fabrication de roues de turbines. À cette occasion, une comparaison entre un processus gamma homogène et non homogène sera réalisée afin de montrer l'influence de la modélisation de la vitesse moyenne de dégradation. La suite du chapitre se concentrera sur les données d'un des matériaux présentés en proposant une étude de l'influence des estimations de paramètres sur la durée de vie d'un composant.

3.2 Dispositifs de tests de cavitation érosive

Plusieurs typologies de tests sont disponibles en laboratoire pour étudier le phénomène de cavitation érosive. En particulier, le partenaire industriel Hydro Québec a recours à deux dispositifs exploitant des technologies différentes : le premier repose sur l'exposition de l'échantillon étudié à un jet d'eau cavitant. Pour cela, un écoulement d'eau est porté à haute pression puis projeté à travers une buse dont la géométrie est étudiée pour entraîner un jet cavitant. Ce dispositif, appelé CavitJet (technologie développée en interne), génère une cavitation intense provoquant l'érosion rapide de l'échantillon étudié. Représenté sur la Figure 3.1, il est utilisé uniquement à des fins de comparaison de matériaux (ou de revêtements), ce dispositif ne permet pas d'étudier finement l'évolution de la cavitation (ou son mécanisme physique). De plus, les tests reposant sur ce dispositif ne permettent de réaliser qu'un nombre réduit de mesures de perte de masse au cours du temps. En raison du faible nombre de données disponibles, ces dernières n'ont pas été exploitées dans le cadre de cette thèse. Le deuxième dispositif, présenté en détails par la suite, repose sur la génération d'un nuage de cavitation à proximité de l'échantillon testé. Celui-ci est provoqué par la vibration d'un cône plongé dans une eau à température constante.

Dispositif de cavitation vibratoire

Le cas d'étude présenté dans la section suivante repose sur les données de cavitation érosive collectées à l'issue d'essais réalisés en laboratoire à l'aide du dispositif de cavitation vibratoire. Cette méthode expérimentale, décrite précisément par la norme G32-06, a été initialement élaborée dans le but de "comparer la résistance à la cavitation érosive de différents matériaux et/ou d'étudier en détails la nature et l'évolution de la dégradation pour un matériau donné" (ASTM, 2016). De ce fait, même si les données de laboratoire ne peuvent pas être transposées directement aux machines hydrauliques réelles (en raison de la technique utilisée pour générer la cavitation), le mécanisme de dégradation est considéré comme similaire à celui observé sur les roues de turbines Francis sujettes à la cavitation érosive. Dans cette étude, les essais de cavitation érosive sont réalisés à l'aide d'un dispositif à ultrasons tel que proposé dans la norme ASTM G32. Le dispositif expérimental est représenté dans la Figure 3.2. Les échantillons du matériau

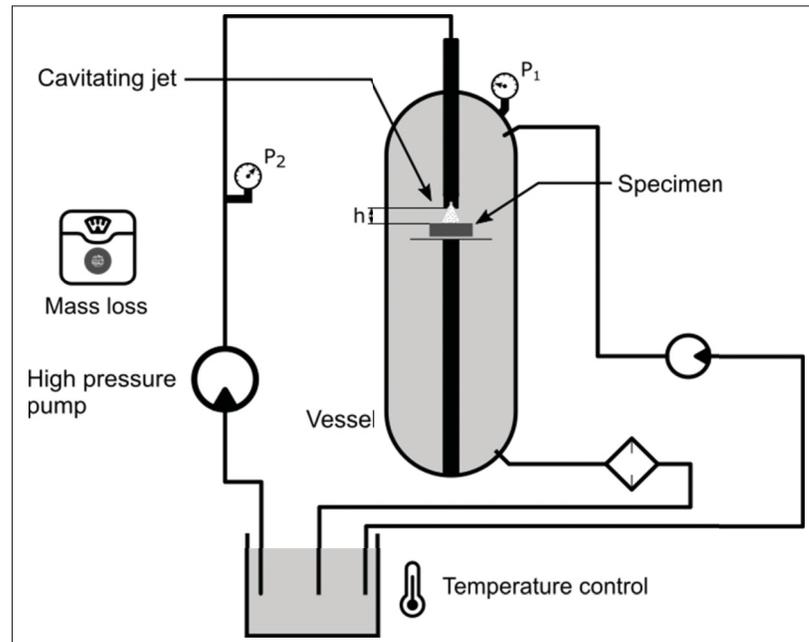


Figure 3.1 Représentation schématique du dispositif de test par jet cavitant (CavitJet).

testé sont maintenus de façon statique sous un cône vibrant à une distance de $0,5 \pm 0,02$ mm de celui-ci (méthode indirecte). La méthode directe consiste quant à elle à fixer l'échantillon testé directement à l'extrémité du cône vibrant. La vibration générée par le cône lors du test a les caractéristiques suivantes : une fréquence de 20 kHz et une amplitude de déplacement du cône, crête à crête, de 50 μm .

Selon la méthodologie décrite dans la norme ASTM, l'essai doit suivre les étapes suivantes :

- L'échantillon est plongé dans un récipient rempli avec le liquide de test (eau distillée ou autre, si un environnement particulier doit être reproduit), maintenu à une température donnée tout au long de l'essai.
- L'échantillon est pesé méticuleusement avant que l'essai ne débute, puis à intervalles périodiques (après séchage) de façon à obtenir un historique de la perte de masse en fonction du temps.

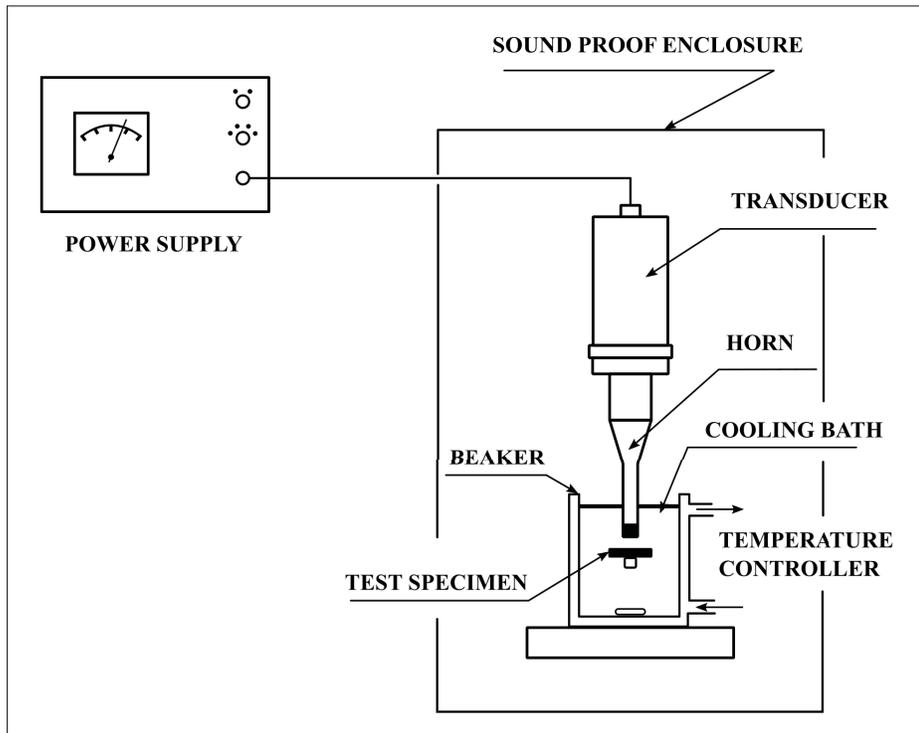


Figure 3.2 Représentation schématique du dispositif de test par cavitation vibratoire. Tirée et adaptée de ASTM (2016).

Les définitions suivantes, initialement proposées par la norme ASTM, seront adoptées dans la suite du chapitre :

- **Temps nominal d'incubation (A)** : obtenu à partir de l'intersection entre l'axe représentant le temps d'exposition et la pente maximale de la courbe d'érosion cumulée ; bien que ce temps ne puisse être considéré comme une mesure en tant que telle, il permet de placer sur cette courbe la ligne du taux d'érosion maximale.
- **Taux d'érosion maximale (B)** : défini comme le maximum du taux d'érosion instantanée lors du test. Celui-ci est ensuite suivi par une diminution du taux d'érosion.

3.3 Matériaux étudiés

Dans cette étude, les matériaux comparés sont deux aciers inoxydables : le AISI 415 et le E309L. Ces deux matériaux ont été choisis en raison de leur utilisation courante dans la fabrication

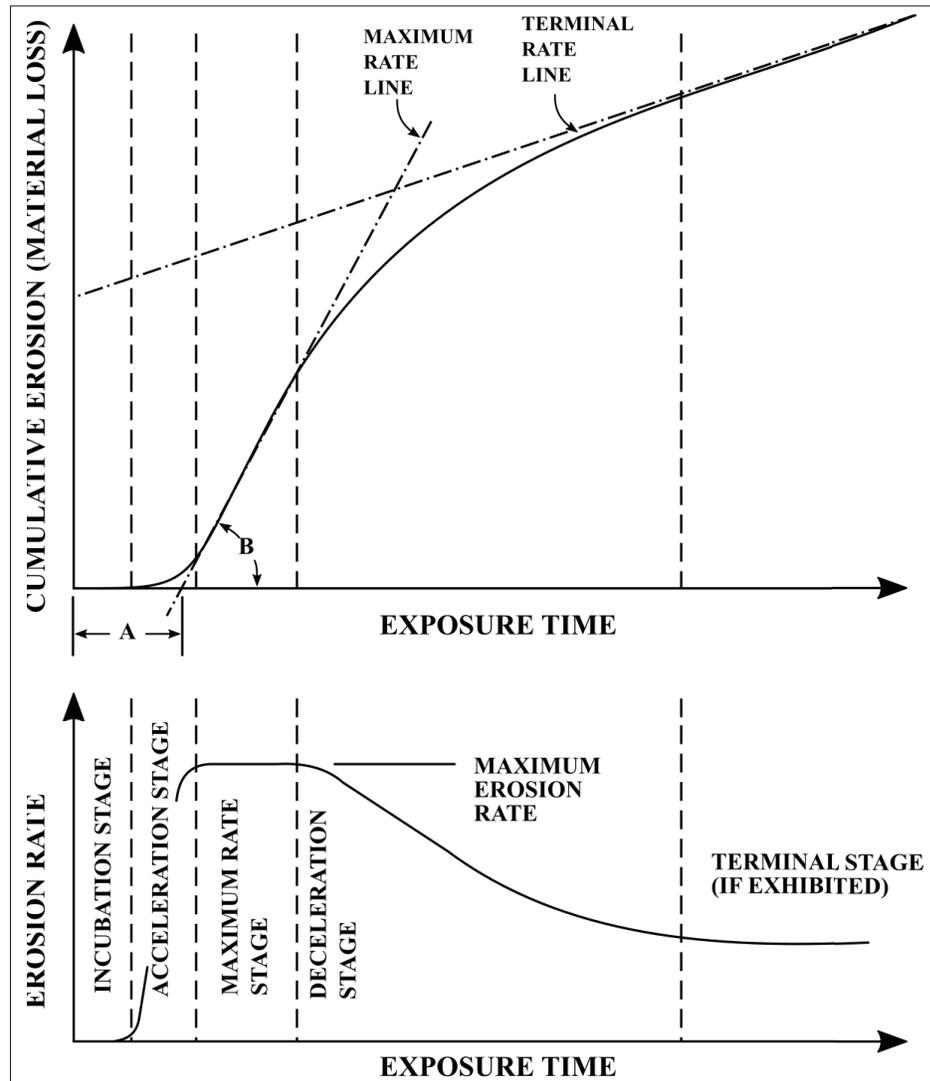


Figure 3.3 Étapes caractéristiques de l'évolution de la courbe d'érosion en fonction du temps. Tirée et adaptée de ASTM (2016).

de turbines de type Francis. Les échantillons extraits de ces deux matériaux, des cylindres de un (1) pouce de diamètre, sont issus de deux aubes de turbine fabriquées en fonderie (procédé de coulage). Le E309L est classiquement utilisé lors de la réparation des roues subissant des dommages par fatigue ou par cavitation érosive tandis que le AISI 415 est équivalent au CA6NM, le matériau employé pour le coulage des roues. Des essais par cavitation vibratoire ont été réalisés pour les deux matériaux selon la méthodologie présentée dans la section précédente. Afin de capturer les différentes sources de variabilité présentes lors d'un test, deux (2) répliques de

l'essai ont été réalisées pour le matériau E309L et une (1) pour le AISI 415. Ainsi, on dénombre un total de 49 observations pour le E309L et 41 pour le AISI 415. On notera que les jeux de données étudiés présentent un nombre relativement faible d'observations pouvant entraîner des incertitudes importantes sur les paramètres estimés.

Pour des considérations industrielles, nous nous sommes concentrés sur les données appartenant à la région d'accélération de la cavitation (*acceleration stage*) présente dans la Figure 3.3. En effet, la région d'incubation (*incubation stage*) n'est pas observée expérimentalement pour les données issues des systèmes de monitoring de machines réelles. Même si les mécanismes physiques mis en jeu sont supposés communs aux deux sources de données, l'intensité de cavitation régnant à l'intérieur de la turbine semble être bien plus importante que celle générée à l'aide du dispositif de cavitation vibratoire en laboratoire. Ces différences d'intensité de cavitation peuvent ainsi expliquer l'absence de région d'incubation dans la dynamique de dégradation des matériels industriels. Néanmoins, il aurait été possible de modéliser les deux régions précédentes à l'aide de méthodes de détection de changement de régime par exemple (Lorden & Eisenberger, 1973). Cette option n'a pas été retenue ici, et dans cette perspective, uniquement la région d'accélération a été modélisée par la suite. Ainsi, le temps d'incubation, noté $t_{inc,j}$, pour chaque essai j , a été au préalable estimé graphiquement à l'aide de la pente de la courbe d'érosion cumulée (tel que détaillé dans la norme ASTM du test de cavitation vibratoire). Les résultats sont présentés dans le Tableau 3.1. Les trajectoires de dégradation, pour le matériau E309L, sont représentées dans la Figure 3.4. Les régions d'incubation et d'accélération sont distinguées pour une meilleure compréhension.

Pour des questions de simplicité et afin que les données manipulées soient conformes aux notations introduites dans la section 1.3.5.2, les données de dégradation associées aux régions d'accélération de la cavitation ont subi une transformation (changement de repère) telle que : $t_{0,j} = t_{inc,j}$ et $x_{0,j}(t_{0,j}) = 0$. Ce changement de repère translate les coordonnées de l'origine avec : $t_{new,j} = t_{inc,j}$ sur l'axe temporel et $x_{new,j} = x(t_{new,j})$ sur l'axe de la dégradation (*cf.* notations introduites en section 1.3.5.2).

Tableau 3.1 Temps d'incubation $t_{inc,j}$ des essais de cavitation réalisés sur deux matériaux usuels.

Essai	Matériau	
	AISI 415	E309L
1	90 min	150 min
2	90 min	180 min
3	-	180 min

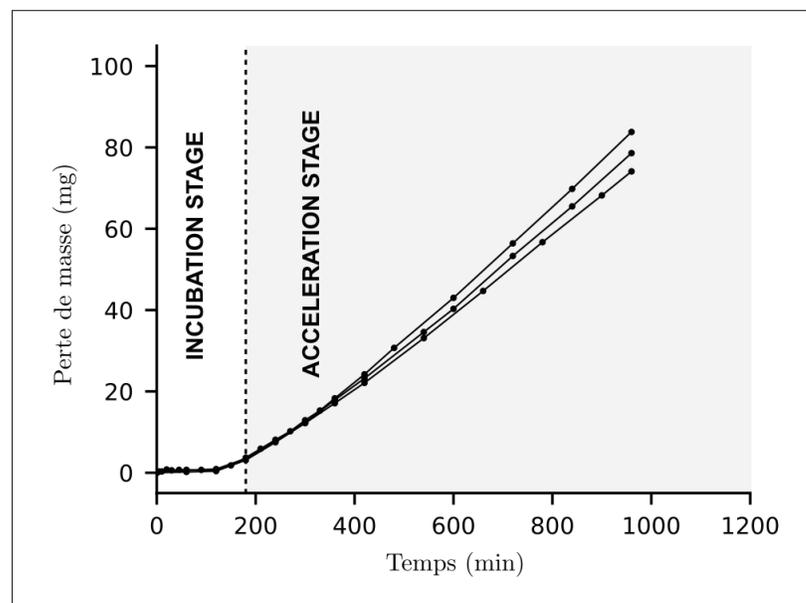


Figure 3.4 Trajectoires de dégradation pour le matériau E309L.

3.4 Comparaison des matériaux

3.4.1 Analyse statistique préliminaire

Comme présenté dans la section 1.3.5, le modèle de loi puissance pour la fonction de forme d'un processus gamma a été privilégié pour la suite de la modélisation. Ce choix a été motivé suite à

l'analyse des résultats de quatre (4) modèles de régressions différents (basée sur le score *AIC*, voir section 1.4.1) appliqués aux données de dégradation transformées. Ainsi, les régressions de type : logarithme, linéaire, polynomiale (d'ordre 2 et 3) et puissance ont été évaluées à l'aide du logiciel R. C'est à l'issue de ces régressions que le modèle de type loi puissance $v(t) = ct^b$ a été retenu pour modéliser la vitesse moyenne de la dégradation au travers de la fonction $v(t)$. La valeur de b approximée par la régression précédente a été par la suite utilisée en tant que valeur initiale de l'optimisation dans la méthode de MV. On notera que dans cette partie, l'estimation du paramètre b a été réalisée à l'aide d'une optimisation de la fonction log-vraisemblance et non pas par la méthode des moindres carrés.

3.4.2 Estimation des paramètres - Comparaison des modèles

Afin d'apporter des éléments de justification en faveur d'un processus gamma non homogène, mais aussi de comparer ce dernier avec son cas particulier homogène, deux fonctions de forme ont été étudiées. Pour la première, le cas particulier de la fonction de forme $v(t) = ct^b$ avec $b = 1$ résulte en une fonction linéaire. Son expression suppose ainsi que la vitesse moyenne de la dégradation est constante au cours du temps. Ensuite, la même fonction, sous sa forme non homogène, $v(t) = ct^b$, $b \neq 1$ est considérée. La méthode MV a été employée pour estimer les paramètres u , c (et b) d'un processus gamma avec les deux fonctions de forme $v(t)$.

Les deux modèles présentés plus haut sont comparés à l'aide des données du matériau E309L, ce dernier ayant le plus de trajectoires observées (avec 2 répliques du test). Par la suite, le modèle offrant les meilleures performances sera privilégié pour la modélisation de la dégradation. Au préalable, une étude de convergence portant sur la moyenne et la variance des estimateurs des paramètres estimés $\hat{\theta}$ à partir de trajectoires simulées, est réalisée afin de déterminer le nombre B minimum de trajectoires permettant d'atteindre un comportement asymptotique acceptable de l'estimateur. Cette étude montre que celle-ci est atteinte pour $B > 8000$ trajectoires. Dans la Figure 3.5, 10000 trajectoires sont simulées sur l'intervalle de temps discret $[0; 2000]$ min. Les lignes en noir représentent les trajectoires de dégradation observées. Un seuil de dégradation est arbitrairement fixé avec $\rho = 100$ mg afin d'illustrer la notion de temps d'atteinte dans ce

cas d'application. Les trajectoires de dégradation sont simulées sur un intervalle de temps assurant le dépassement du seuil ρ par les trajectoires. Les densités de probabilité du niveau de dégradation aux instants 360, 600 et 960 min sont calculées à l'aide des trajectoires simulées et sont présentées dans la Figure 3.6. Les points noirs représentent la dégradation expérimentale aux niveaux précédents pour les trajectoires observées. La Figure 3.7 présente les densités de probabilités du temps d'atteinte pour le seuil $\rho = 100$ mg.

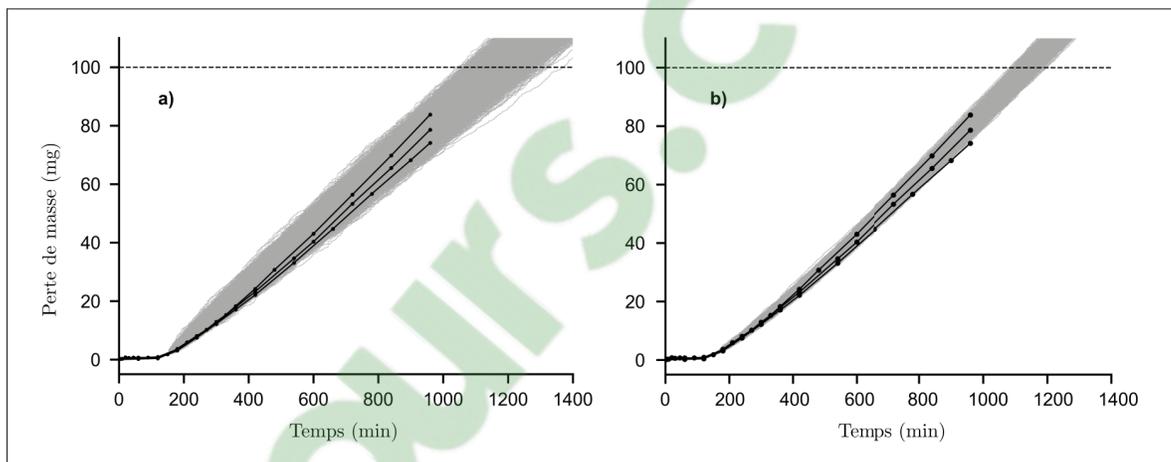


Figure 3.5 Matériau E309L - Simulations d'un processus gamma : a) homogène et b) non homogène - 10000 trajectoires simulées.

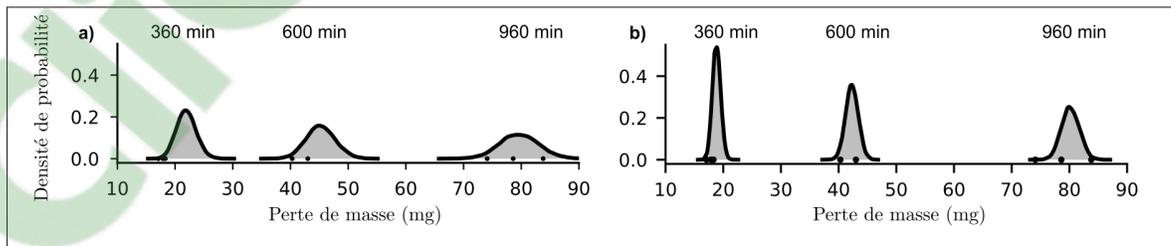


Figure 3.6 Matériau E309L - Densités de probabilité du niveau de dégradation aux instants : 360, 600 et 960 min pour des trajectoires simulées d'un processus gamma : a) homogène et b) non homogène.

Les résultats présentés dans les Figures 3.5 et 3.7 montrent que la dégradation est mieux modélisée en utilisant un processus gamma non homogène (PGNH) (fonction de forme $v(t) = ct^b$, $b \neq 1$). On observe que les trajectoires simulées par un PGNH présentent une meilleure adéquation aux

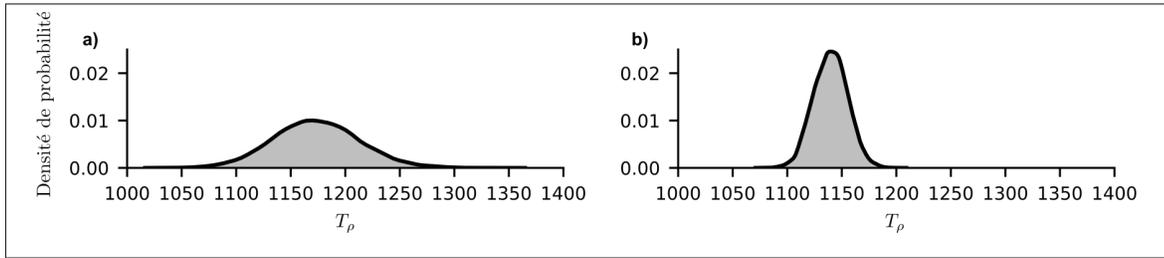


Figure 3.7 Matériau E309L - Densités de probabilité du temps d'atteinte pour $\rho = 100$ mg de trajectoires simulées d'un processus gamma : a) homogène et b) non homogène.

trajectoires observées que pour les simulations d'un PGH. L'EMV estime la valeur du paramètre b à 1,13 (pour les trajectoires des deux matériaux) ce qui est cohérent avec les résultats obtenus à partir des régressions utilisant un modèle de loi puissance. Les estimations des paramètres de forme et d'échelle sont donnés dans le Tableau 3.2 pour un PGNH. On observe sur la Figure 3.7 que les densités de probabilité des trajectoires simulées pour un PGNH sont plus centrées autour des niveaux de dégradation observés et présentent une étendue plus faible que pour celles d'un PGH. Ces résultats montrent ainsi que la dynamique de la dégradation est mieux modélisée avec un PGNH. Pour la suite du chapitre, la dégradation des matériaux sera alors modélisée par ce type de modèle.

3.4.3 Analyse des résultats

On s'intéresse ici à la comparaison des deux matériaux présentés en section 3.3 vis-à-vis de leur résistance à la cavitation. Pour cela, les paramètres d'un PGNH ont été estimés pour les deux jeux de données puis des simulations de trajectoires utilisant ces paramètres ont été réalisées, cette fois, pour le matériau AISI 415. Les résultats sont présentés à l'aide de la Figure 3.8 et des Tableaux 3.2 et 3.3 à des fins de comparaison. Les intervalles de confiance (IC) asymptotiques sont calculés et présentés dans le Tableau 3.2 à l'aide de la matrice de Fisher (*cf.* section 1.5). Pour rappel, si on note $\hat{\theta} = (\hat{u}, \hat{c}, \hat{b})$, l'erreur relative de l'estimation est exprimée par :

$$\text{Erreur relative} = \frac{\sqrt{\mathbb{V}(\hat{\theta}_i)}}{\hat{\theta}_i}, \quad (3.1)$$

avec $\mathbb{V}(\hat{\theta}_i)$ qui représente la variance de l'estimateur du paramètre $\hat{\theta}_i$, $1 \leq i \leq 3$.

On peut souligner que les jeux de données utilisés lors de l'estimation des paramètres d'un PGNH pour les deux matériaux sont de même taille totale N . Intuitivement, on peut alors s'attendre à une "précision" semblable entre les deux estimations des paramètres d'un PGNH. Cependant, l'erreur relative sur les estimations, présentée dans le Tableau 3.2, est plus faible pour le E309L dont l'échantillon de données se compose de 3 trajectoires de dégradations au lieu de 2 pour le AISI 415. À partir de ces résultats, on observe donc pour une même taille totale de l'échantillon, l'estimation des paramètres semble présenter de meilleures performances lorsque le nombre de trajectoires est plus important. Cette observation est cohérente avec les résultats tirés de l'étude de l'EMV présentée dans la section 2.2

Tableau 3.2 Paramètres estimés d'un PGNH.

Matériau		u	c	b
E309L	Estimation	32,5	1,272	1,135
	IC (à 95%)	[16,3 ; 48,7]	[0,61 ; 1,94]	[1,11 ; 1,16]
	Erreur relative	0,254	0,268	0,012
AISI 415	Estimation	10,3	0,494	1,131
	IC (à 95%)	[5,18 ; 15,4]	[0,22 ; 0,77]	[1,09 ; 1,17]
	Erreur relative	0,254	0,284	0,018

Le Tableau 3.3 regroupe les résultats du temps d'atteinte à différents seuils ρ pour les deux matériaux étudiés. Ici, les distributions du temps d'atteinte ont été estimées à partir des simulations des trajectoires d'un PGNH et non à partir de l'approximation de Birnbaum-Saunders. On note que pour un seuil de dégradation donné ρ , $p_{T_\rho}^\alpha$ représente le quantile de la distribution du temps d'atteinte T_ρ au niveau α . À partir du Tableau 3.2, il est possible de déterminer si les IC des deux matériaux présentent une superposition des intervalles. C'est le cas par exemple pour l'IC sur c du E309L qui couvre environ 12 % de celui du AISI 415, tandis que les IC sur u sont disjoints et que les IC sur b sont pratiquement identiques. Cela signifie qu'on ne peut pas

rejeter l'hypothèse que le paramètre c soit identique pour les deux matériaux. Cependant, si on considère à présent la distribution du temps d'atteinte à un seuil donné (représentée dans la Figure 3.9 et le Tableau 3.3), on observe pour chaque valeur de ρ que les distributions sont disjointes. Ainsi, il est possible d'affirmer à partir de ces résultats que les deux matériaux n'ont pas le même comportement vis-à-vis de la résistance à la cavitation. Par ailleurs, le matériau E309L montre la meilleure résistance à la cavitation.

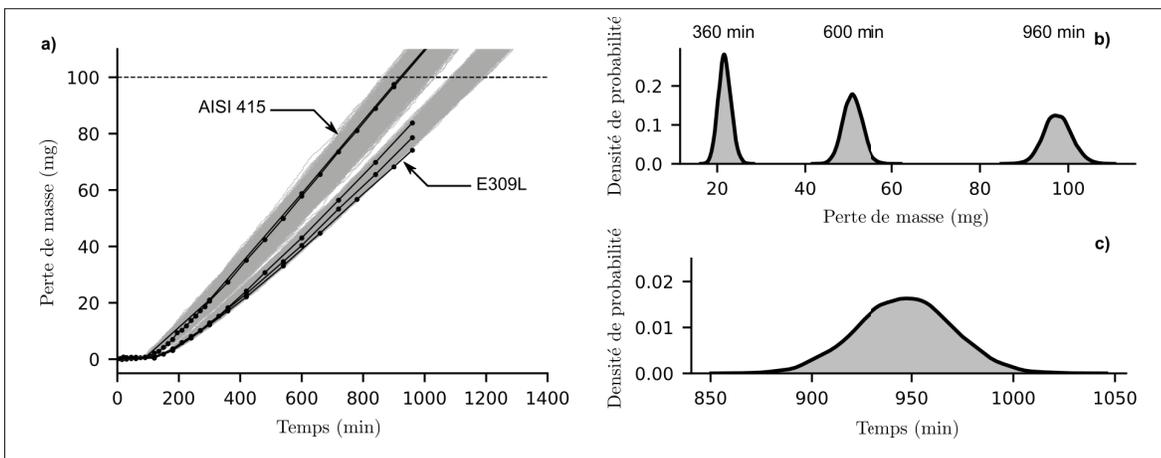


Figure 3.8 Matériau AISI 415 - a) Simulations d'un processus gamma non homogène - b) Densités de probabilité du niveau de dégradation aux instants 360, 600 et 960 min - c) Densité de probabilité du temps d'atteinte au seuil $\rho = 100$ mg.

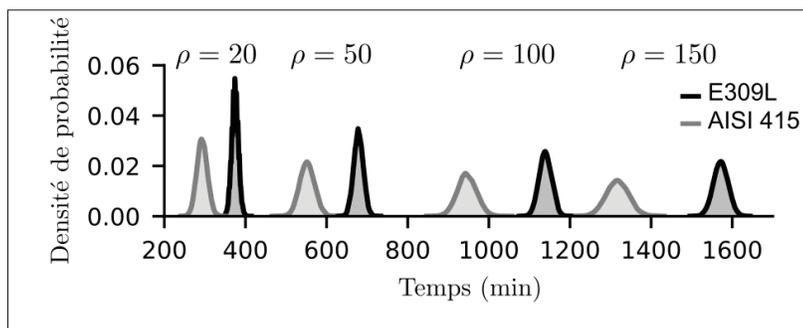


Figure 3.9 Comparaison des distributions du temps d'atteinte pour les seuils $\rho = (20, 50, 100, 150)$ mg et les deux matériaux.

Tableau 3.3 Simulation du temps d'atteinte.

Matériau		$\rho = 20$ mg	$\rho = 50$ mg	$\rho = 100$ mg	$\rho = 150$ mg
E309L	Moyenne (T_ρ)	374	679	1140	1572
	$[p_{T_\rho}^{2,5\%}; p_{T_\rho}^{97,5\%}]$	[360; 392]	[656; 699]	[1111; 1173]	[1534; 1610]
	Erreur relative	0,022	0,017	0,013	0,012
AISI 415	Moyenne (T_ρ)	293	552	946	1317
	$[p_{T_\rho}^{2,5\%}; p_{T_\rho}^{97,5\%}]$	[269; 318]	[515; 586]	[903; 989]	[1269; 1373]
	Erreur relative	0,043	0,033	0,025	0,021

3.5 Influence de l'incertitude des paramètres sur le temps d'atteinte

Dans cette section, nous nous intéressons à l'influence de l'incertitude des paramètres d'un PG sur le temps d'atteinte. Contrairement à la section précédente qui estime la distribution du temps d'atteinte à partir d'un faisceau de trajectoires simulées dépassant un certain seuil ρ , les méthodes statistiques introduites dans la section 1.5 sont mises en œuvre. La variabilité de la distribution du temps d'atteinte dans la section précédente reposait uniquement sur la variabilité de la distribution gamma induite par les valeurs des paramètres de simulation utilisées. Ici, les incertitudes sur les paramètres d'un processus gamma sont propagées à l'aide des méthodes *bootstrap* et *delta* présentées dans la section 1.5.

3.5.1 Construction des intervalles de confiance

Les différentes méthodes de construction d'intervalles de confiance (IC) introduites en section 1.5 sont appliquées à des données de dégradation en laboratoire. En particulier, on s'intéressera aux données de dégradation relatives au matériau E309L présentées plus haut, qui présentent l'avantage d'avoir le plus grand nombre de trajectoires m .

3.5.1.1 Méthode asymptotique

Pour rappel, à partir de la normalité asymptotique du maximum de vraisemblance, il est possible de construire des IC sur θ . Ainsi, pour chaque composant θ_i du vecteur des paramètres θ , l'intervalle de confiance au niveau $1 - \alpha$ s'obtient avec :

$$\left[\hat{\theta}_{MV,i} - s\sqrt{\left(J(\hat{\theta})^{-1}\right)_{ii}}; \hat{\theta}_{MV,i} + s\sqrt{\left(J(\hat{\theta})^{-1}\right)_{ii}} \right], \quad (3.2)$$

avec $s = \Phi\left(1 - \frac{\alpha}{2}\right)$ et Φ la CDF d'une loi normale centrée réduite et $1 \leq i \leq 3$.

Il est aussi rappelé que la matrice de Fisher observée $J(\theta)$ est donnée par :

$$J(\theta) = -\nabla^2 \ell(\theta) \quad (3.3)$$

Pour la fonction log-vraisemblance $\ell(\theta)$, le vecteur gradient de la dérivée partielle première du processus gamma peut s'écrire de la façon suivante :

$$\nabla \ell(\theta) = \begin{pmatrix} \frac{\partial \ell(\theta)}{\partial u} \\ \frac{\partial \ell(\theta)}{\partial c} \\ \frac{\partial \ell(\theta)}{\partial b} \end{pmatrix} \quad (3.4)$$

La matrice Hessienne avec la dérivée partielle seconde de cette même fonction $\ell(\theta)$ s'écrit quant à elle :

$$\nabla^2 \ell(\theta) = \begin{pmatrix} \frac{\partial^2 \ell(\theta)}{\partial u^2} & \frac{\partial^2 \ell(\theta)}{\partial u \partial c} & \frac{\partial^2 \ell(\theta)}{\partial u \partial b} \\ \frac{\partial^2 \ell(\theta)}{\partial c \partial u} & \frac{\partial^2 \ell(\theta)}{\partial c^2} & \frac{\partial^2 \ell(\theta)}{\partial c \partial b} \\ \frac{\partial^2 \ell(\theta)}{\partial b \partial u} & \frac{\partial^2 \ell(\theta)}{\partial b \partial c} & \frac{\partial^2 \ell(\theta)}{\partial b^2} \end{pmatrix} \quad (3.5)$$

3.5.1.2 Méthodes de ré-échantillonnage

Les trois méthodes de ré-échantillonnage présentées dans la section 1.5.3 sont appliquées aux données de dégradation du matériau E309L. Afin d'assurer la bonne compréhension de la mise en œuvre de ces méthodes, les algorithmes utilisés sont présentés pour chaque méthode. On prendra pour la suite comme échantillon \mathbf{D}_{obs} , les données de dégradation des trajectoires observées du matériau E309L dans la région "accélération" uniquement (et une fois la transformation décrite dans la section 3.3 réalisée).

Bootstrap paramétrique

La procédure de ré-échantillonnage paramétrique (BP) (présentée en section 1.5.3) peut être résumée par l'algorithme suivant :

1. À partir des observations de dégradation \mathbf{D}_{obs} , estimer les paramètres d'un processus gamma $\hat{\theta}_{MV}$.
2. Simuler m trajectoires de dégradation de longueur n_j , $1 \leq j \leq m$, suivant un processus gamma avec pour paramètre $\hat{\theta}_{MV}$ sur l'intervalle de temps discret $[0, t_{n_j}]$.
3. Estimer $\hat{\theta}_{BP}^*$ à partir du nouvel échantillon obtenu à l'étape 2.
4. Répéter les étapes 2. et 3. B fois afin d'obtenir B estimations $\hat{\theta}_{BP,w}^*$, $1 \leq w \leq B$, des B jeux de données simulées.
5. À partir des B estimations réalisées $\hat{\theta}_{BP,w}^*$, $1 \leq w \leq B$, calculer la moyenne empirique $\overline{\hat{\theta}_{BP}^*} = \frac{1}{B} \sum_{w=1}^B \hat{\theta}_{BP,w}^*$ et les quantiles empiriques d'ordre q , $0 < q < 1$, $\hat{P}_{\hat{\theta}_{BP}^*}^q$. Ces quantiles donnent une image de la dispersion de $\hat{\theta}_{BP}^*$ comparé à $\hat{\theta}_{MV}$ et permettent d'évaluer statistiquement l'impact de l'incertitude d'estimation de θ sur le temps d'atteinte.

Par la suite on considère que $\hat{\theta}_{BP}$ est l'estimateur *bootstrap* naturel de θ , avec $\hat{\theta}_{BP} = \overline{\hat{\theta}_{BP}^*} = \frac{1}{B} \sum_{w=1}^B \hat{\theta}_{BP,w}^*$.

Moving Block Bootstrap

La procédure de ré-échantillonnage MBB (présentée en section 1.5.3) peut être résumée par l'algorithme suivant :

1. Définir un bloc de longueur l .
2. Tirer avec remise un bloc de l observations consécutives à partir de \mathbf{Y}_j .
3. Répéter b_j fois l'étape 2. jusqu'à soit que l'échantillon *bootstrap* \mathbf{Y}_j^* possède $b_j \times l \geq n_j$ éléments.
4. Conserver uniquement les n_j premières valeurs ré-échantillonnées.
5. Répéter les étapes 2-4 pour chaque trajectoire de dégradation j , $1 \leq j \leq m$, afin d'obtenir un jeu de m trajectoires $\mathbf{Z}^* = \{\mathbf{Y}_1^*, \dots, \mathbf{Y}_m^*\}$
6. Répéter l'étape 5. B fois pour générer B échantillons *bootstrap* \mathbf{Z}_w^* , $1 \leq w \leq B$.
7. Calculer $\hat{\theta}_{MBB,w}^*$ à l'aide de la méthode MV pour chaque échantillon *bootstrap* \mathbf{Z}_w^* .

De même que pour la méthode BP, à partir des B valeurs de $\hat{\theta}_{MBB,w}^*$, $1 \leq w \leq B$, on pose $\hat{\theta}_{MBB} = \overline{\hat{\theta}_{MBB}^*} = \frac{1}{B} \sum_{w=1}^B \hat{\theta}_{MBB,w}^*$, l'estimateur MBB naturel de θ , et les quantiles empiriques d'ordre q , $0 < q < 1$, $\hat{P}_{\hat{\theta}_{MBB}^*}^q$. Dans ce cas d'application, on prend une longueur de bloc $l = 3$ afin de réaliser un compromis entre la dépendance temporelle de la structure et la longueur relativement "courte" n_j des trajectoires.

Méthode bootstrap d'Efron

Comme évoqué dans la section 1.5.3, la méthode Efron *bootstrap* (EB) peut être considérée comme un cas particulier de la méthode MBB en prenant une longueur de bloc $l = 1$. Ainsi, de même que pour les méthodes de ré-échantillonnage précédentes, à partir des B valeurs de $\hat{\theta}_{EB,w}^*$, $1 \leq w \leq B$, on peut introduire la moyenne empirique $\hat{\theta}_{EB} = \overline{\hat{\theta}_{EB}^*} = \frac{1}{B} \sum_{w=1}^B \hat{\theta}_{EB,w}^*$ qui est considérée comme une estimation EB de θ , et les quantiles empiriques d'ordre q , $0 < q < 1$, $\hat{P}_{\hat{\theta}_{EB}^*}^q$.

Afin de déterminer un nombre adéquat B d'échantillons *bootstrap* à tirer (ou générer), une analyse de convergence a été réalisée. Pour cela, le précédent échantillon de données \mathbf{D}_{obs} est utilisé et les trois méthodes de *bootstrap* sont appliquées pour différentes valeurs de B , $10 \leq B \leq 10^4$. La Figure 3.10 présente $\hat{\theta}_{EB}$ et les quantiles, $\hat{P}_{\hat{\theta}_{EB}^*}^{2,5\%}$ et $\hat{P}_{\hat{\theta}_{EB}^*}^{97,5\%}$, représentés avec les barres d'erreur définies plus haut. On observe que la convergence asymptotique est atteinte autour de $B = 10^3$ tirages. Pour la suite, on prendra $B = 10^4$ tirages pour les trois méthodes *bootstrap* afin d'assurer une convergence correcte.

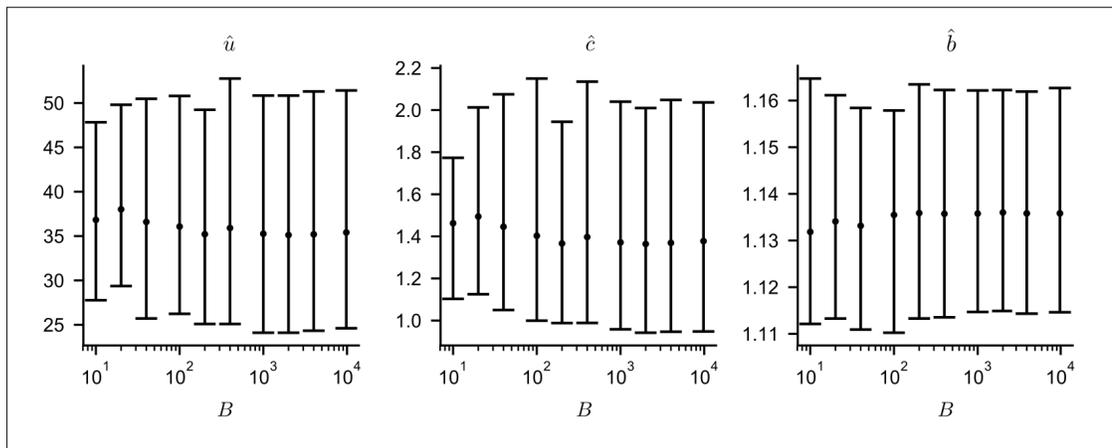


Figure 3.10 $\hat{\theta}_{EB}$, $\hat{P}_{\hat{\theta}_{EB}^*}^{2,5\%}$ et $\hat{P}_{\hat{\theta}_{EB}^*}^{97,5\%}$ - Analyse de convergence sur le nombre d'itérations *bootstrap* B à réaliser pour la méthode EB.

3.5.2 Propagation des incertitudes - Méthode delta

Dans cette section, nous nous intéressons à évaluer comment les incertitudes d'estimation sur θ peuvent impacter la distribution du temps d'atteinte T_ρ , et en particulier les quantiles d'ordre q , $0 < q < 1$, données par : $P_{T_\rho}^q = \mathbb{F}_{T_\rho}^{-1}(q|\theta) \approx \mathbb{F}_{BS}^{-1}(q|\theta)$. Si on considère que l'hypothèse de convergence introduite dans l'expression (1.34) est vérifiée, alors la méthode delta (MD) peut être appliquée. Ainsi, lorsque $N = \sum_{j=1}^m n_j$ est suffisamment grand, $g(\hat{\theta}_{MV})$ suit la distribution normale : $\mathcal{N}\left(g(\theta), J_g(\theta)J^{-1}(\hat{\theta}_{MV})J_g(\theta)^T\right)$, avec $J_g(\theta)$ la jacobienne de g , c'est-à-dire la matrice 1×3 des dérivées partielles de g selon les composants du vecteur de θ et J , l'expression définie dans l'équation (1.36).

Pour cela, si on considère l'approximation de Birnbaum-Saunders donnée dans l'expression (1.25), on doit alors inverser la CDF de cette dernière expression et prendre comme fonction

$$g(\boldsymbol{\theta}) = g_q(\boldsymbol{\theta}) = \mathbb{F}_{BS}^{-1}(q|\boldsymbol{\theta})$$

dans l'équation suivante :

$$\begin{aligned} g_q(\boldsymbol{\theta}) &= \left(\Phi[h(P_{T_\rho}^q | \boldsymbol{\theta})] \right)^{-1} = h^{-1} \left(\Phi^{-1}(q) | \boldsymbol{\theta} \right) \\ &= 2^{-1/b} \left(\frac{2c\rho u + 2c \operatorname{erf}_c^{-1}(2q)^2 - \sqrt{2}c \operatorname{erf}_c^{-1}(2q) \sqrt{4\rho u + 2 \operatorname{erf}_c^{-1}(2q)^2}}{c^2} \right)^{1/b}, \end{aligned} \quad (3.6)$$

avec $h(t|\boldsymbol{\theta}) = \sqrt{u\rho} \left(\sqrt{\frac{v(t)}{u\rho}} - \sqrt{\frac{u\rho}{v(t)}} \right)$ l'approximation de Birnbaum-Saunders $\mathbb{F}_{BS}(t|\boldsymbol{\theta}) = \Phi(h(t|\boldsymbol{\theta}))$ dans l'équation (1.25) et $\operatorname{erf}_c^{-1}$ la fonction inverse de erf_c , la fonction d'erreur complémentaire donnée par l'expression suivante :

$$\operatorname{erf}_c(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty \exp(-t^2) dt. \quad (3.7)$$

3.5.3 Indicateurs statistiques de performance

Dans le Tableau 3.4, nous introduisons quelques indicateurs statistiques classiques afin d'évaluer les performances des méthodes *bootstrap* et de les comparer avec la combinaison de la méthode de maximum de vraisemblance et la méthode delta (MV + MD). Si on considère par exemple l'estimation des paramètres du modèle, le biais mesure la qualité d'approximation de la vraie valeur (inconnue) du paramètre par $\hat{\boldsymbol{\theta}}$. De même, la variance et l'étendue des $\hat{\boldsymbol{\theta}}_w^*$ donnent une indication sur la précision de l'estimation de $\boldsymbol{\theta}$ par $\hat{\boldsymbol{\theta}}$. Enfin, l'erreur quadratique moyenne (EQM) combine ces deux indicateurs afin de pouvoir comparer les méthodes entre elles. $\overline{\hat{\boldsymbol{\theta}}^*}$ désigne la moyenne de $\hat{\boldsymbol{\theta}}_w^*$, $1 \leq w \leq B$, comme introduit dans chacun des algorithmes *bootstrap*.

Tableau 3.4 Indicateurs statistiques utilisés pour caractériser les performances des méthodes *bootstrap*.

	Expression mathématique
Variance	$\frac{1}{B} \sum_{w=1}^B (\hat{\theta}_w^* - \overline{\hat{\theta}^*})^2$
Erreur relative	$\frac{\sqrt{\text{variance}}}{\overline{\hat{\theta}^*}}$
Étendue (à 95%)	$\hat{P}_{\hat{\theta}^*}^{97,5\%} - \hat{P}_{\hat{\theta}^*}^{2,5\%}$
Erreur quadratique moyenne	biais ² + variance

3.5.4 Analyse des résultats

Estimation des incertitudes du modèle

Les trois méthodes *bootstrap* présentées plus haut ont été appliquées aux données \mathbf{D}_{obs} . Les IC asymptotiques sont calculés selon la méthode présentée dans la section 3.5.1.1 et permettent d'obtenir les approximations des intervalles de confiance symétriques à 95% sur θ .

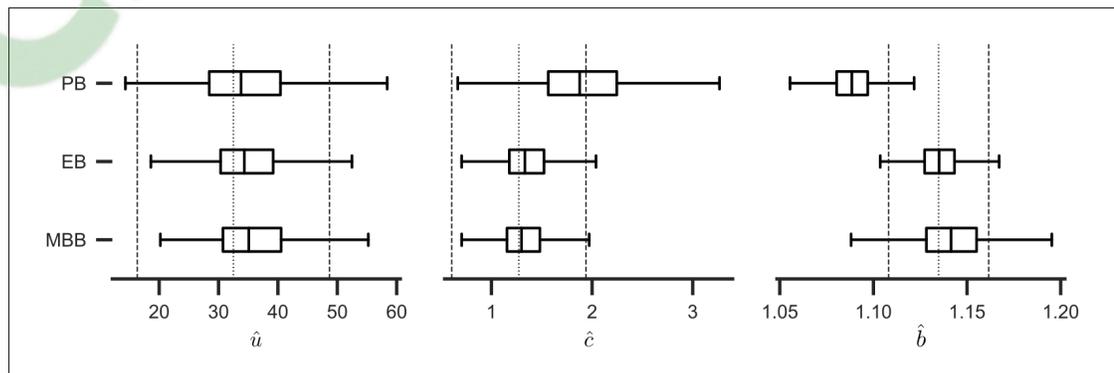


Figure 3.11 Comparaison des IC asymptotiques et *bootstrap*.

Tableau 3.5 Estimations de θ par la méthode MV et 3 techniques de *bootstrap* différentes
- Approximation des intervalles de confiance asymptotiques (à 95%) de θ .

	\hat{u}	\hat{c}	\hat{b}
MV	32,501	1,2722	1,1348
Approximation IC asympt. (à 95%)	[16,32 ; 48,69]	[1,101 ; 1,939]	[1,108 ; 1,162]
BP	35,154	1,9516	1,0888
EB	34,971	1,3572	1,1361
MBB	36,420	1,3549	1,1421

Le Tableau 3.6 présente une synthèse des indicateurs statistiques pour les résultats des méthodes *bootstrap* et MV. On observe dans le cadre de cette étude que les méthodes présentant les meilleures performances sont EB et MBB, avec une erreur relative autour de 25% pour les paramètres u et c , ce qui peut être considéré comme plutôt élevé. Cependant, la méthode EB semble produire de meilleures estimations de θ que MBB en termes d'incertitudes d'estimation relatives. Pour une meilleure illustration et comparaison, la Figure 3.11 présente les boîtes à moustaches de $\hat{\theta}_w^*$, $1 \leq w \leq B$, pour les trois méthodes *bootstrap*. Les moustaches inférieures et supérieures représentent $\hat{P}_{\hat{\theta}^*}^{2,5\%}$ et $\hat{P}_{\hat{\theta}^*}^{97,5\%}$ respectivement ; les boîtes représentent quant à elles, les premiers $\hat{P}_{\hat{\theta}^*}^{25\%}$ et troisièmes $\hat{P}_{\hat{\theta}^*}^{75\%}$ quantiles ainsi que la médiane $\hat{P}_{\hat{\theta}^*}^{50\%}$. Les lignes en pointillés à gauche et à droite représentent respectivement les bornes inférieures et supérieures de l'intervalle de confiance asymptotique à 95% du Tableau 3.5, obtenus à partir de la normalité asymptotique du MV. La ligne en pointillés au centre représente $\hat{\theta}_{MV} = (\hat{u}_{MV}, \hat{c}_{MV}, \hat{b}_{MV})$. Dans le Tableau 3.6 et la Figure 3.11, on observe le biais le plus important (positif, pouvant entraîner une possible sur-estimation ; de même que pour le biais relatif) pour la méthode BP (pour les paramètres c et b , autour de 10% et 50% respectivement) ainsi que la plus grande dispersion sur u et c , tandis que le biais est de l'ordre de 10% pour les méthodes EB et MBB. Ces dernières montrent d'ailleurs des résultats similaires aussi pour la dispersion en dehors de l'estimation de b par la méthode EB qui présente un biais proche de zéro. Une des principales hypothèses de l'approximation des IC asymptotiques obtenus à partir de l'EMV (*cf.* section 1.5.1) est soulignée

Tableau 3.6 Indicateurs statistiques pour les différentes méthodes *bootstrap*.

		MV	Techniques <i>bootstrap</i>		
			BP	EB	MBB
Biais	\hat{u}		2,653	2,470	3,919
	\hat{c}	0 (supposé)	0,6794	0,0850	0,0827
	\hat{b}		-0,0461	0,0013	0,0073
Biais relatif	\hat{u}		0,0816	0,0760	0,1206
	\hat{c}	0 (supposé)	0,5340	0,0668	0,0650
	\hat{b}		-0,0406	0,0011	0,0064
Variance	\hat{u}	68,20	87,56	48,59	58,11
	\hat{c}	0,1158	0,2926	0,0787	0,0865
	\hat{b}	0,0002	0,0002	0,0002	0,0004
Erreur relative	\hat{u}	0,2541	0,2879	0,2145	0,2346
	\hat{c}	0,2675	0,4252	0,2206	0,2312
	\hat{b}	0,0120	0,0109	0,0108	0,0182
Étendue à 95%	\hat{u}	32,64	36,77	25,84	29,31
	\hat{c}	0,838	2,099	1,040	1,133
	\hat{b}	0,0540	0,0486	0,0485	0,0814
EQM	\hat{u}	68,195	94,69	56,74	73,89
	\hat{c}	0,1158	0,7544	0,08961	0,09553
	\hat{b}	0,0002	0,0023	0,0002	0,0003

dans le Tableau 3.6 : $\hat{\theta}_{MV}$ est supposé centré sur la vraie valeur (inconnue) du paramètre θ et ainsi ne présente aucun biais (asymptotiquement, ce qui peut s'avérer une hypothèse erronée pour un faible échantillon de données, comme présenté dans l'analyse de convergence de la Figure 3.10). Cependant, en considérant que l'EMV ne présente pas de biais, l'EQM la plus faible est obtenue pour la méthode EB sur les trois paramètres u , b et c . On observe dans la Figure 3.11 que les boîtes à moustaches sont asymétriques alors que l'approximation des

IC asymptotiques est centrée autour de $\hat{\theta}_{MV}$. Ce résultat n'est pas vérifié sur les distributions *bootstrap* $\hat{\theta}^*$. Les Tableaux 3.5 et 3.6 ainsi que la Figure 3.11 présentent une erreur relative faible sur les estimations du paramètre b (environ 1%) pour les méthodes non paramétriques et montrent que la valeur de la puissance $b = 1$ n'est pas incluse dans l'IC à 95%, justifiant donc l'utilisation d'un PGNH pour modéliser les données d'érosion par cavitation.

Tableau 3.7 Estimation des quantiles du temps d'atteinte : $P_{T_\rho}^{2,5\%}$, $P_{T_\rho}^{50\%}$, $P_{T_\rho}^{97,5\%}$ et de leur incertitude associée pour les 3 méthodes *bootstrap* et MV + MD.

Quantiles de T_ρ		MV+MD	BP	EB	MBB
$P_{T_\rho}^{2,5\%}$	Médiane	976	950	976	976
	Étendue à 95%	31	40	40	54
	Erreur relative	0.032	0.042	0.041	0.055
$P_{T_\rho}^{50\%}$	Médiane	1006	979	1005	1006
	Étendue à 95%	31	40	40	54
	Erreur relative	0,031	0,041	0,040	0,054
$P_{T_\rho}^{97,5\%}$	Médiane	1037	1011	1036	1036
	Étendue à 95%	34	44	41	56
	Erreur relative	0,033	0,044	0,040	0,054

Influence des incertitudes du modèle sur le temps d'atteinte

Les Figures 3.12 et 3.13 illustrent le concept de propagation des incertitudes. En effet, l'influence de l'incertitude d'estimation de θ sur les distributions du temps d'atteinte \mathbb{F}_{T_ρ} peut être obtenue en ayant recours à l'approximation de Birnbaum-Saunders (seuil $\rho = 100 \text{ mg}$) à partir de la méthode delta et des résultats des méthodes *bootstrap*. Afin d'étudier l'impact des incertitudes sur la distribution complète de T_ρ , chaque quantile d'ordre $q = \{2, 5\%, 5\%, \dots, 97, 5\%\}$: $\mathbb{F}_{BS}^{-1}(q|\hat{\theta}_{ML})$ est représenté dans la Figure 3.12, ainsi que sa dispersion en utilisant la méthode delta conjointement avec $\mathbb{F}_{BS}^{-1}(q|\hat{\theta}_w^*)$, avec $1 \leq w \leq B$, pour les trois méthodes *bootstrap*.

Les distributions alors obtenues ainsi que leur boîte à moustaches sont présentées dans la Figure 3.12 (les mêmes règles de construction ont été employées que pour la Figure 3.11). Pour des questions de comparaison, les précédents résultats sont concaténés dans la Figure 3.13 qui représente uniquement la valeur centrale de la boîte à moustaches et respectivement les premiers et troisièmes quantiles. La ligne en noir représente la médiane de la distribution des quantiles tandis que l'aire en gris représente $\hat{P}_{T_p}^{75\%} - \hat{P}_{T_p}^{25\%}$ des précédentes distributions. On observe clairement dans la Figure 3.12 que la dispersion semble rester constante pour les différents ordres de quantiles pour toutes les méthodes *bootstrap*. On remarquera que les boîtes sont légèrement asymétriques (de façon plus significative pour la méthode MBB) : la plupart des valeurs sont concentrées sur la moitié inférieure de la distribution des ordres de quantiles. Cette tendance est aussi observable sur la Figure 3.11 pour les trois paramètres du PGNH. MV+MD, EB et MBB donnent des résultats très proches pour les quantiles d'ordre inférieur ou supérieur mais aussi pour les quantiles d'ordre médian. Pour les méthodes MV+MD et EB, on constate également des résultats semblables en termes de dispersion des estimations des quantiles. Uniquement MBB produit une dispersion plus importante sur les distribution des quantiles du temps d'atteinte. Si on regarde maintenant l'erreur relative de $P_{T_p}^q$ dans le Tableau 3.3, on peut souligner que T_p semble être mieux estimé que θ . Ce résultat peut s'expliquer par un effet de compensation des incertitudes d'estimation des paramètres u , c et b sur la distribution du temps d'atteinte.

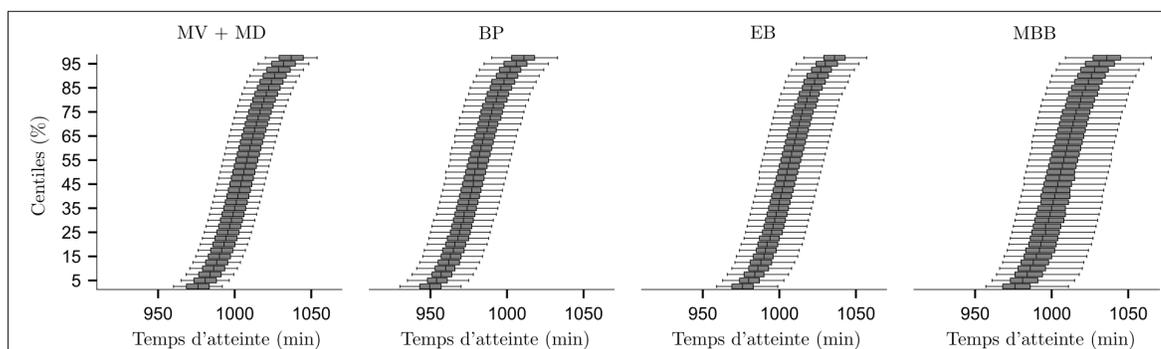


Figure 3.12 Propagation des incertitudes sur la distribution du temps d'atteinte pour $\rho = 100 \text{ mg}$.

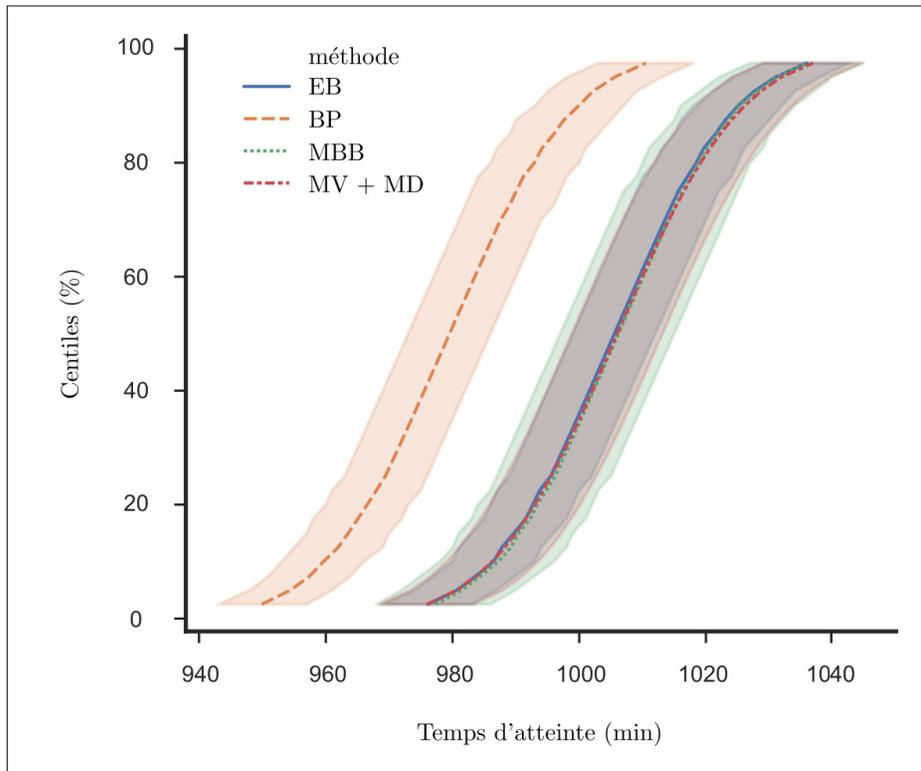


Figure 3.13 Distribution du temps d'atteinte T_ρ (avec $\rho = 100 \text{ mg}$) pour les 3 méthodes *bootstrap* et MV + MD.

Au regard des résultats du paragraphe consacré aux estimations d'incertitudes, l'utilisation d'un PGNH s'avère justifié par le fait que l'IC sur b ne contient pas la valeur 1 (c'est-à-dire, équivalent au cas homogène). Les méthodes de *bootstrap* non paramétriques semblent donner des résultats proches de ceux de la méthode MV, tandis que le *bootstrap* paramétrique présente le biais et la dispersion les plus grands sur les paramètres estimés. Même si ce résultat peut sembler contre intuitif, on remarque que BP ne parvient pas à produire des simulations de trajectoires satisfaisantes lorsque leur longueur n_j est petite (inférieure à 10^3 observations, comme présenté dans l'étude de performance sur l'EMV du Chapitre 2). Cette tendance est particulièrement vraie pour les PGNH, dont les performances de l'EMV sont moins bonnes que pour un PGH (avec de petites trajectoires aussi).

MBB et EB fournissent de bonnes estimations de la dynamique de la dégradation au travers du paramètre b , avec un biais relatif et une dispersion inférieure à 1% pour les deux méthodes. Ce

même indicateur atteint environ 10% pour les paramètres u et c , montrant que les méthodes *bootstrap* non paramétriques éprouvent plus de difficultés pour estimer ces paramètres. Parmi les méthodes *bootstrap*, EB semble produire les plus proches estimations par rapport à $\hat{\theta}_{MV}$. Ce résultat peut paraître surprenant sachant que la méthode MBB est supposée mieux capturer la dépendance temporelle des séries. Une explication peut venir du fait que l'échantillon de données est relativement limité (avec des trajectoires courtes). Ainsi, les trajectoires ré-échantillonnées ne fournissent pas beaucoup plus d'information comparées aux initiales. Cette remarque est aussi applicable aux autres méthodes *bootstrap* dans le cas présent.

À partir du Tableau 3.6 et de la Figure 3.13, on observe que l'incertitude sur le paramètre b produit la plus forte influence sur la distribution du temps d'atteinte. Plus la dispersion sur b est importante, plus la dispersion estimée des quantiles (à un ordre donné) de T_ρ est importante. Une exploration plus poussée pourrait être réalisée pour quantifier l'influence de chaque paramètre sur le temps d'atteinte à travers une analyse de sensibilité. La combinaison de MV et MD produit des résultats proches de EB pour un temps de calcul significativement plus faible : environ 0,0029 seconde ¹ pour MV et MD et environ 2035 secondes pour les méthodes *bootstrap* (pour générer $B = 10^4$ échantillons et estimer les paramètres associés). Néanmoins, les développements mathématiques demandent plus d'effort pour la MD et l'incertitude de la distribution du temps d'atteinte. De plus, les méthodes *bootstrap* donnent la distribution complète des estimateurs des paramètres alors que les IC asymptotiques ne fournissent qu'un intervalle centré sur $\hat{\theta}_{MV}$. Ceci permet, pour les méthodes *bootstrap*, d'estimer les possibles biais entre $\hat{\theta}_{MV}$ et θ . Pour ces raisons et parce que MV repose sur plusieurs hypothèses (qui ne sont pas nécessairement vérifiées dans notre cas), EB montre de meilleures performances et semble être plus robuste que la combinaison MV et MD.

3.6 Conclusion

À partir des données de dégradation produites par un dispositif de laboratoire, nous avons estimé les paramètres d'un processus gamma (homogène et non homogène) ainsi que leurs

¹ Réalisé sur un processeur Quad Core Intel i7 2.8GHz avec 16GB de RAM

incertitudes associées. Ce travail a servi à deux intérêts distincts : le premier, la comparaison de la résistance à la cavitation de deux matériaux concurrents. On a montré qu'à partir de la simulation (avec les paramètres estimés de l'échantillon de données de dégradation) d'un faisceau de trajectoires qu'il était possible de prouver que les deux matériaux présentaient des tenues à la cavitation différentes. Pour cela, la distribution empirique du temps d'atteinte (à un seuil donné) a été obtenue grâce aux trajectoires simulées. Le deuxième intérêt repose sur l'étude de l'influence des incertitudes des paramètres sur le temps d'atteinte. Pour cela, la méthode de maximum de vraisemblance (MV) couplée à la méthode delta (MD) ont été utilisées, aux côtés de trois méthodes de ré-échantillonnage. Ces dernières, coûteuses en temps de calcul, s'avèrent être faciles à implémenter et fournissent des informations supplémentaires aux méthodes de maximum de vraisemblance et delta. De plus, leur utilisation s'avère intéressante lorsqu'on est en présence d'un jeu de données limité.

CHAPITRE 4

MODÉLISATION DE LA DÉGRADATION DES ROUES DE TURBINE À PARTIR DES DONNÉES DE MONITORING

Dans ce chapitre, il est question de modéliser la dégradation de roues de turbines de type Francis pour des unités hydrauliques dont la puissance nominale développée est typiquement supérieure à 100 MW. Cette dégradation résulte de l'effet de la cavitation érosive à la surface des aubes de la roue de turbine. Comme mentionné précédemment, en raison des fluctuations de pression de l'eau au passage de cette dernière dans la turbine, des bulles de vapeur peuvent être créées et ainsi arracher de la matière en impactant la surface de la roue, générant à l'occasion des vibrations dans la machine. Au fil du temps, certaines surfaces de la roue peuvent être creusées sous l'action mécanique de la cavitation érosive, entraînant une perte d'efficacité du groupe Turbine/Alternateur (T/A).

4.1 Acquisition et pré-traitement des données

4.1.1 Dispositif de monitoring en continu

Dans le cadre de cette étude, les données de dégradation considérées sont issues du couplage de systèmes de suivi de cavitation et de conditions d'opération des groupes T/A. Afin d'opérer ces derniers dans les centrales hydroélectriques et de disposer d'un historique des conditions d'opération, celles-ci sont mesurées et échantillonnées à intervalle régulier par un système dédié présent dès la construction des installations par le fabricant des groupes T/A. Ce système permet à l'exploitant de disposer d'une image en direct et d'enregistrer périodiquement l'état de certaines quantités d'intérêt telles que : la hauteur de chute, la puissance développée, l'ouverture des vannes de guidage, la combinaison des groupes T/A fonctionnant en simultanément dans une même centrale, etc. Ces variables d'opération sont stockées dans une base de données comme illustré par l'onglet "Données Op." de la Figure 4.1.

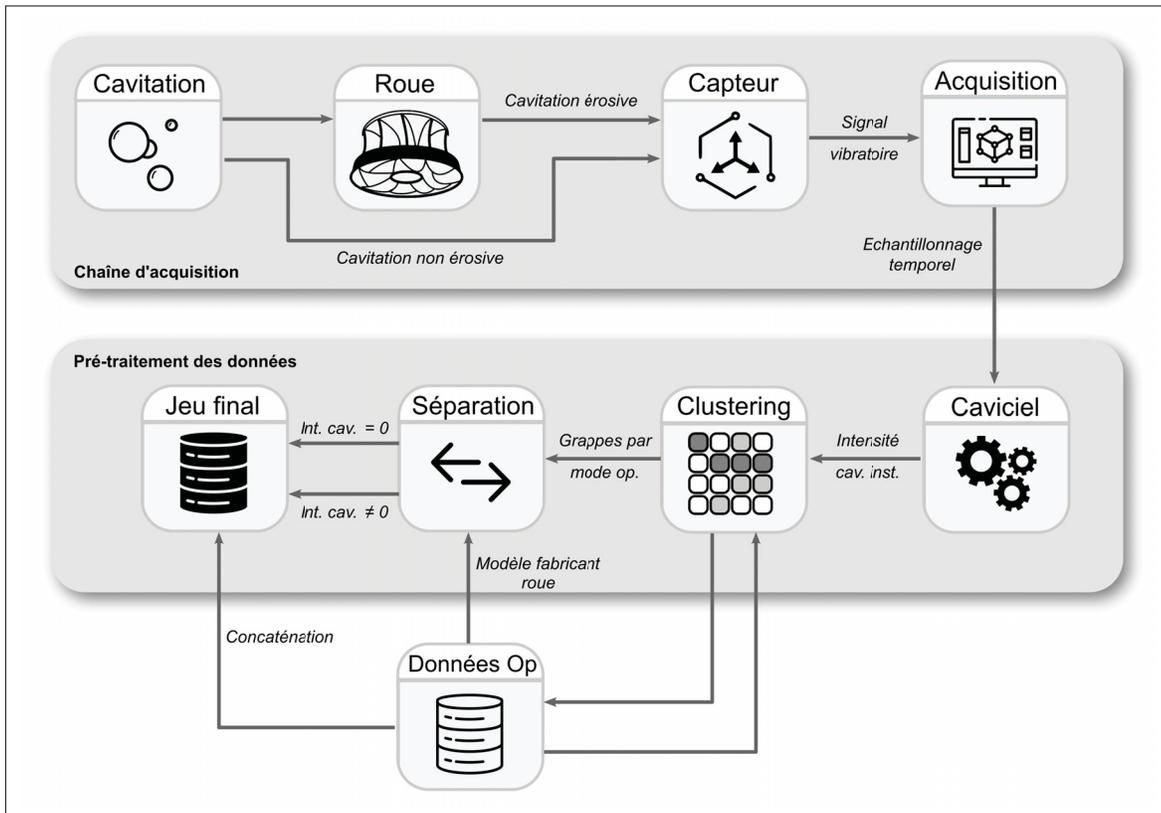


Figure 4.1 Processus mixte d'acquisition et pré-traitement des données de monitoring.

Le deuxième dispositif d'acquisition de données consiste en un système de suivi de l'intensité de cavitation perçue par les groupes T/A (mesure exprimée en kg/10 000h de fonctionnement). Ce dernier repose sur la mesure du signal vibratoire généré par la cavitation à l'intérieur d'une roue de turbine. Contrairement au système de suivi des conditions d'opération, ce dispositif est installé et maintenu par l'exploitant hydroélectrique sur certains de ses groupes T/A. L'objectif de ce système est d'estimer l'intensité de la cavitation à l'intérieur de la turbine par le biais d'une mesure indirecte. Pour rappel, le meilleur indicateur de l'état de dégradation dû à la cavitation érosive demeure la quantification de la masse perdue par le composant dégradé. Dans le cas des roues de turbines hydroélectriques en fonctionnement, la mesure de la masse perdue au cours du temps est rendue impossible par des contraintes techniques. Pour pallier cette absence de mesure directe, l'intensité de cavitation au sein de la roue de turbine est estimée, puis l'incrément de dégradation associé à celle-ci est calculé. Dans un premier temps, pour

estimer l'intensité de la cavitation, une mesure du signal vibratoire au niveau du palier guide inférieur de la turbine est réalisée avec un accéléromètre (cf. zone entourée sur la Figure 4.2). Cette mesure est réalisée à intervalles réguliers (pas forcément homogènes) et il est considéré que la période d'échantillonnage n'affecte pas la représentativité des conditions d'opération. De même, la durée de l'échantillonnage est considérée suffisamment longue pour bien caractériser la distribution statistique des impacts de cavitation au point de fonctionnement de la mesure. Pour de plus amples renseignements sur l'architecture du dispositif d'acquisition, le lecteur est invité à se référer à Bourdon (2000).

Ainsi mesuré, le signal vibratoire est ensuite filtré au travers d'un passe-bas puis la valeur quadratique moyenne (RMS) ainsi que la modulation du signal sont analysées sur la période d'acquisition. L'intensité de la cavitation instantanée (en [kg/10 000 h]) durant la période de mesure est obtenue après un calcul à l'aide de l'outil *Caviciel* (développé par Hydro Québec, Lafleur (2012)) à travers une fonction de transfert (cf Figure 4.1).

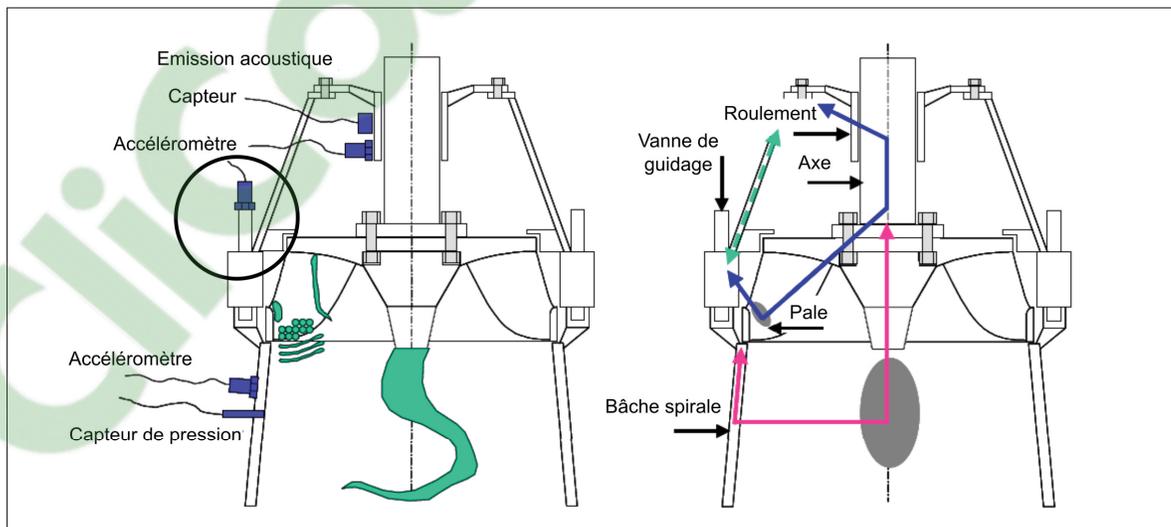


Figure 4.2 Zone de manifestation de la cavitation et emplacements des capteurs (gauche) et possibles chemins empruntés par les vibrations (droite). Figure tirée et adaptée de Escaler *et al.* (2006).

4.1.2 Distinction entre cavitation érosive et non érosive

Dans le cas des turbines de type Francis, la cavitation peut apparaître dans différentes zones, comme illustré sur les Figures 4.2 et 4.3 (Escaler *et al.*, 2006; Franc *et al.*, 1995; Gindroz, 1991). Certaines zones de cavitation, suffisamment éloignées de toutes surfaces, n'entraînent pas de dégradation de celles-ci. C'est le cas par exemple pour la cavitation en torche hélicoïdale présentée en Figure 4.3, qui génère des vibrations au travers de la roue de turbine sans la dégrader. Il convient alors de distinguer lorsqu'une excitation vibratoire est résultante de la cavitation érosive ou non. Pour cela, les exploitants s'appuient sur les modèles développés en mécanique des fluides par les fabricants de turbines, destinés initialement à la conception du prototype. Ceux-ci permettent de déterminer, en fonction des conditions d'opération, si de la cavitation va apparaître et quelle zone de la turbine est susceptible d'être affectée. Suivant les zones concernées, la cavitation est alors considérée comme érosive ou non. Ces modèles numériques font aussi l'objet d'un processus de validation en laboratoire à l'aide de modèles réduits pour corroborer les résultats obtenus. *In fine*, il est alors possible de distinguer si l'excitation vibratoire mesurée est l'image de la cavitation érosive ou de la cavitation non érosive.

Dans le cadre de cette étude, d'après le constructeur de la turbine, seule la cavitation se formant à l'entrée de l'extrados des pales (*cf.* Figure 4.3 a)) est considérée comme érosive. Par conséquent, si les conditions d'opération de la turbine impliquent l'apparition de cavitation en dehors de la zone précédemment citée, on considère, par simplification, que l'intensité de cavitation est nulle car non érosive (car n'entraînant *a priori* pas de dégradation). Cette étape est représentée par l'onglet "séparation" dans le processus de pré-traitement de la Figure 4.1.

4.1.3 Regroupement par grappes (*clustering*)

Le groupe T/A étudié dans ce chapitre opère dans une centrale hydroélectrique composée d'un total de huit (8) groupes T/A semblables entre eux. En fonction des besoins en énergie et de la planification de la production en électricité, il est possible de faire fonctionner de zéro (0) à huit (8) groupes T/A simultanément. Parmi les bonnes pratiques à respecter pour opérer une centrale

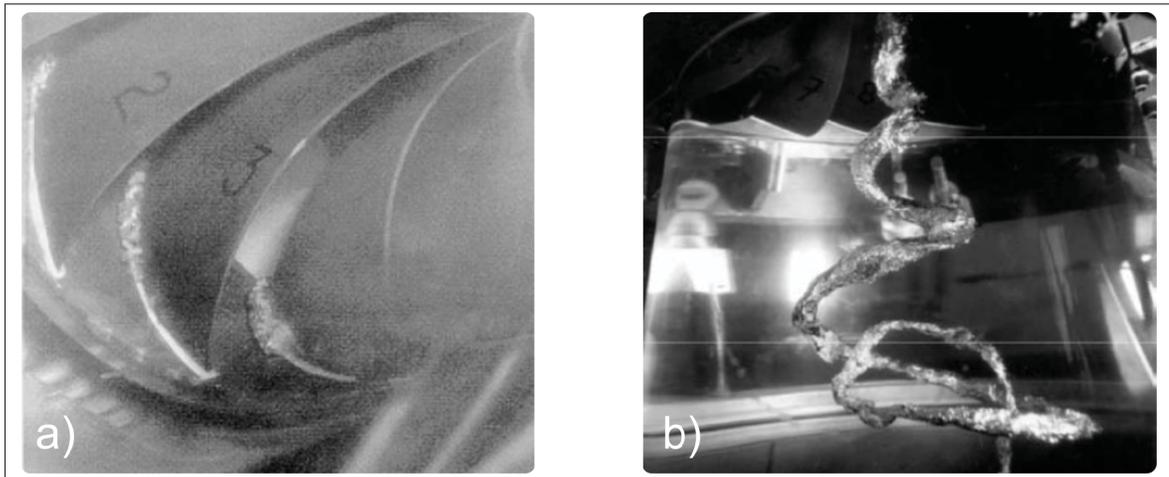


Figure 4.3 Exemples de cavitation rencontrés sur une turbine de type Francis : a) entrée extradoss, b) torche hélicoïdale. Figures tirées et adaptées de Franc *et al.* (1995) et Escaler *et al.* (2006)

hydroélectrique, l'ordonnancement des groupes T/A en fait partie. Elle consiste à démarrer et arrêter les groupes T/A dans une séquence prédéfinie afin d'optimiser l'énergie produite. En effet, bien que similaires en termes de conception, les groupes T/A présentent des variations de rendement entre eux à cause notamment de l'aménagement hydroélectrique de la centrale. Dans la séquence d'ordonnancement, sont donc privilégiés, par ordre décroissant, les groupes T/A présentant le meilleur rendement. Bien qu'idéale, cette séquence n'est pas toujours respectée par l'exploitant de la centrale pour des raisons variées liées à l'opération (groupes indisponibles, contraintes de planification de production). Les différentes combinaisons de groupes T/A fonctionnant en simultané sont appelées modes opératoires (MO). Par ailleurs, Bodson-Clermont (2016) a démontré statistiquement l'influence du mode opératoire sur l'intensité de cavitation à laquelle une turbine peut être exposée. Ainsi, pour un même groupe T/A fonctionnant sous des conditions d'opération stables et homogènes, l'intensité de cavitation est fortement influencée par le fonctionnement simultané d'autres groupes T/A partageant la même chambre d'équilibre, comme illustré dans la Figure 4.4. En effet, la hauteur de chute perçue par la turbine est liée directement au niveau de la chambre d'équilibre qui se voit modifiée lorsque plusieurs groupes T/A sont opérés simultanément. Ainsi, les MO constituent un critère essentiel pour caractériser le niveau d'intensité de cavitation.

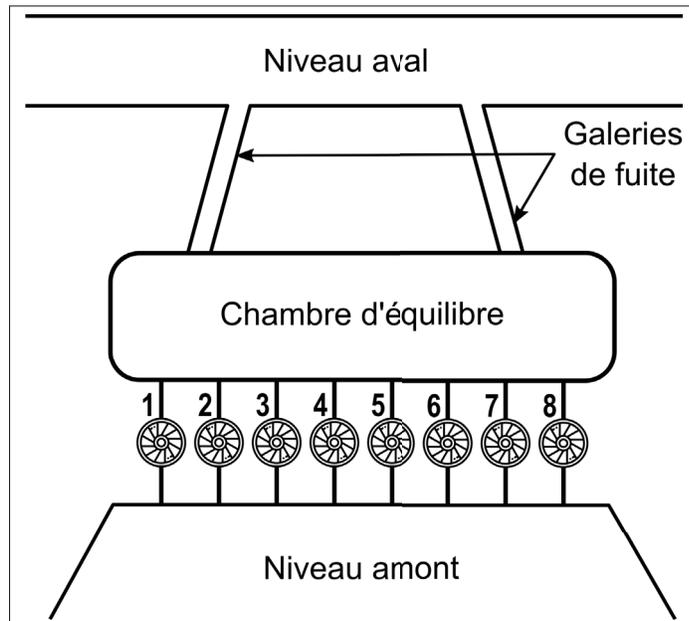


Figure 4.4 Schématisation de la chambre d'équilibre de la centrale et de ses 8 turbines.

Pour l'année 2011, on dénombre dans l'historique des conditions d'opération plus de 70 MO différents. En raison de ce nombre élevé, et par simplification, il a été décidé en 2012 de regrouper par "grappe" les MO présentant des conditions d'opération semblables, à l'aide d'une méthode statistique de *clustering*. L'intérêt de les regrouper ainsi réside dans la possibilité de décrire le comportement d'un ensemble de MO avec un nombre réduit de sous-modèles de dégradation (en faisant l'hypothèse que le nombre de grappes choisi résulte en une disparité acceptable des conditions d'opération à l'intérieur d'une même grappe). Bodson-Clermont (2016) détaille la méthodologie employée pour effectuer le *clustering* en grappes. En résumé, pour chaque MO, la moyenne, l'écart-type et la médiane sont calculés sur les 10 variables d'opération suivantes :

- **courant** (A), intensité délivrée par l'alternateur ;
- **débit** (m^3/s), volume d'eau traversant la turbine par unité de temps (mesure indirecte) ;
- **niveau amont** (m), hauteur de l'eau au niveau du réservoir ;
- **niveau aval** (m), hauteur de l'eau au niveau de la chambre d'équilibre ;

- **ouverture de vannage (%)**, pourcentage d'ouverture des vannes de guidage, ajustant le débit traversant la turbine ;
- **puissance active (MW)**, composante de la puissance dont la tension est en phase avec l'intensité ;
- **puissance réactive (MVAR)**, composante de la puissance dont la tension est en quadrature de phase avec l'intensité ;
- **puissance effective de stabilité (MW)**, puissance maximale développée par la turbine dans son régime transitoire ;
- **tension (kV)**, différence de potentiel en sortie de l'alternateur ;
- **hauteur de chute (m)**, différence de hauteur entre niveau amont et aval.

N.B. : le regroupement par grappe réalisé dans l'étude de (Bodson-Clermont, 2016) ne prend pas en compte la variable "intensité de cavitation".

Un extrait des données est présenté dans le Tableau 4.1. À ces 30 indices statistiques (3 indicateurs calculés sur 10 variables), est ajoutée la fréquence F_{MO} du MO correspondant. Cette fréquence est donnée par l'expression (4.1).

$$F_{MO} = \frac{T_{MO}}{T_{Tot}}, \quad (4.1)$$

avec T_{MO} le temps cumulé passé par le groupe T/A dans le MO opératoire et T_{Tot} , le temps cumulé total d'opération du groupe sur une année. Au total, 31 valeurs statistiques par MO sont utilisées pour réaliser le regroupement hiérarchique (*clustering*) avec comme fonction de lien le critère de Ward (*cf.* Ward Jr (1963)). Contrairement aux fonctions de lien classiques reposant sur un calcul de distance euclidienne normée (par exemple simple ou complète) comme critère de regroupement, Ward cherche à minimiser la variance intra-grappe (aussi appelée inertie). Une fois ces regroupements réalisés (on peut chercher à obtenir jusqu'à une seule grappe), il est possible de visualiser les MO formant des grappes à l'aide d'un dendrogramme, comme présenté dans la Figure 4.5.

Tableau 4.1 Extrait des données d'opération pour l'année 2011.

Date	Courant	Débit	Niv. amont	Niv. aval	Ouv.	P. réac.	P. act.	P. eff. de stabilité	Tension	H. chute
'2011-01-01 00 :10 :28'	13,11	236,36	173,91	33,01	72,56	63,99	302,40	366,23	13,35	140,9
'2011-01-01 00 :22 :16'	13,21	236,18	173,91	32,91	71,81	64,26	302,40	366,13	13,35	141,0
'2011-01-01 00 :34 :04'	13,12	236,62	173,91	32,91	71,81	63,45	302,94	366,14	13,35	141,0
'2011-01-01 00 :45 :51'	13,13	236,74	173,91	32,98	71,81	63,45	302,94	366,13	13,35	140,9
'2011-01-01 00 :57 :38'	13,25	237,25	173,91	33,02	72,56	65,88	303,48	366,11	13,35	140,9
'2011-01-01 01 :09 :26'	13,13	236,27	173,91	32,96	71,81	65,88	302,40	366,04	13,35	141,0
'2011-01-01 01 :21 :14'	12,90	231,64	173,91	32,98	71,12	67,23	296,46	366,14	13,35	140,9

Dans le cadre de cette étude, il a été retenu de regrouper les MO en 5 grappes distinctes. Ce choix est motivé à la fois par l'homogénéité des MO au sein d'une grappe (par exemple la grappe 1 contient des MO opérant peu de groupes T/A) et aussi par l'absence de grappe isolée (contenant peu de MO). À l'issue de ce regroupement, la base contenant les données d'opération a été modifiée pour faire apparaître le numéro de grappe associé à chaque MO.

4.1.4 Calcul de la perte de masse

Un bon indicateur de la dégradation par cavitation d'une roue de turbine est la perte de masse de celle-ci. De ce fait, après avoir mesuré à intervalles de temps réguliers l'intensité de cavitation (pour rappel, exprimée en [kg/10 000 h]) régnant dans la turbine, on cherche à se ramener à une perte de masse entre deux mesures successives. Pour cela, l'expression suivante est utilisée :

$$\delta deg_i = \delta t_i \times cav_inst_i, \quad (4.2)$$

avec δdeg_i l'incrément de perte de masse associé à l'incrément de temps $\delta t_i = (t_i - t_{i-1})$ et cav_inst_i la cavitation instantanée mesurée à l'instant i , $0 \leq i \leq n$ avec n le nombre total d'observation pour une trajectoire (voir Tableau 4.2).

NB : les incréments de dégradation dans le cas de la cavitation sont strictement positifs ou nuls. En effet, l'estimation d'un incrément de dégradation repose sur l'intégration dans le temps d'une fonction strictement positive. Cette méthode de calcul est cohérente avec le mécanisme physique de la cavitation érosive, qui consiste en une succession de petits arrachements de matière à la surface des pales de la turbine.

L'estimation de la masse totale perdue, M_p , sur la période $[0, \tau]$, est obtenue avec l'expression suivante :

$$M_p = \sum_{i=0}^n \delta deg_i \quad (4.3)$$

avec $t_0 = 0$ et $t_n = \tau$.

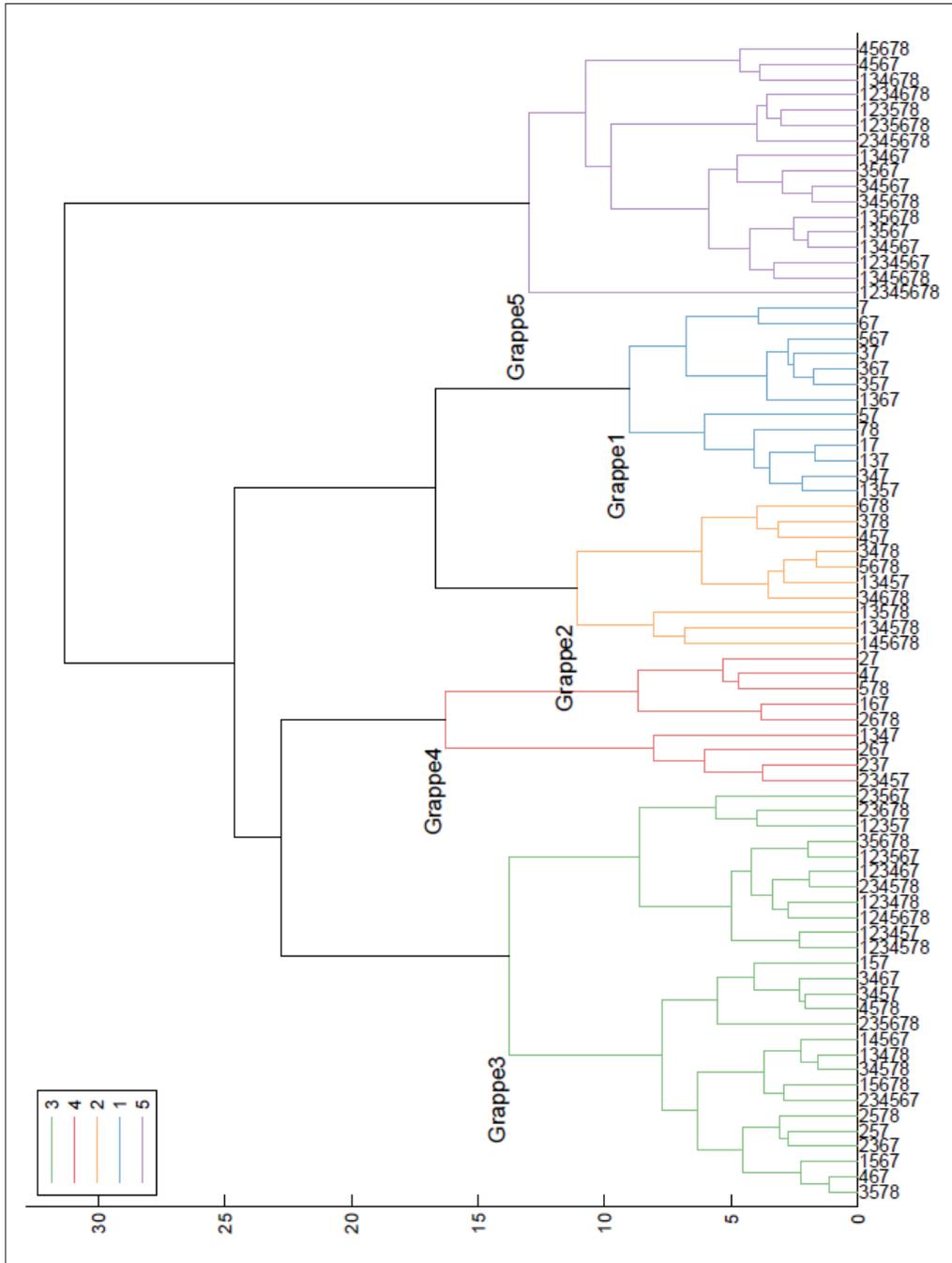


Figure 4.5 Dendrogramme des modes opératoires. Tirée et adaptée de Bodson-Clermont (2016).

4.1.5 Correction des données

Bien que dans ce chapitre un seul groupe T/A fasse l'objet d'une étude, 3 autres groupes de la même centrale sont équipés du dispositif de suivi de cavitation. Afin de réduire la complexité du système, un capteur de vibration est installé sur chaque groupe T/A, mais une seule chaîne d'acquisition est chargée de réaliser successivement les mesures de l'intensité de cavitation. Ces mesures sont prises à intervalles réguliers et cycliques : une fois la mesure sur un groupe réalisée, le système mesure le suivant quelques minutes après et ainsi de suite. Néanmoins, si le MO au moment de la mesure ne contient pas le groupe instrumenté (groupe à l'arrêt), le système passe au groupe suivant dans la séquence de prise de mesure. Ainsi, en fonction du MO, l'intervalle de temps δt_i entre deux mesures successives peut varier, allant de 5 min à 35 min environ. De plus, aucune mention dans la base de données n'indique l'arrêt ou le redémarrage du groupe lorsque la situation se produit. Pour ces raisons, il est nécessaire de déterminer dans la base de données, les instants t_i correspondant à un arrêt ou à un démarrage du groupe T/A et ainsi de corriger les incréments de dégradation associés. En l'absence de cette information, l'estimation de la masse perdue peut alors être surévaluée. Cette éventuelle sur-estimation aurait pour conséquence d'atteindre plus rapidement le seuil à partir duquel on considère que la masse perdue nécessite une intervention de maintenance et donc d'être plus conservatif.

Pour cela, tous les incréments de temps δt_i de la base de données sont calculés. On considère qu'un incrément de temps est "incohérent" dès lors qu'on vérifie : $\delta t_i \geq 1$ heure. Pour les incréments de temps vérifiant la précédente inégalité, on fixe alors l'incrément de dégradation associé $\delta deg_i = 0$, le groupe étant supposé à l'arrêt.

Pour l'ensemble de l'année 2011 (du 01/01 au 31/12), on dénombre 34553 instants mesurés. Ceux-ci sont répartis par grappe selon la colonne "Mesures non-corrigées" du Tableau 4.2. Lorsqu'on ne considère que les mesures dont la cavitation associée est non nulle et après correction, les données sont alors réparties selon la colonne "Mesures corrigées".

Tableau 4.2 Dénombrement des instants mesurés pour l'année 2011 par grappe.

Grappe	Mesures non-corrigées	Mesures corrigées
1	7175	6329
2	1641	1393
3	1916	1722
4	206	204
5	23615	4553
Total	34553	14201

On observe en particulier que seulement 1/5 des instants mesurés de la grappe 5 présente des conditions de cavitation érosive. Sur l'ensemble des instants mesurés, seulement environ 40% constituent des données de dégradation à proprement parlé.

4.2 Analyse préliminaire

Une analyse descriptive des données d'un groupe T/A donné pour l'année 2011 permet de faire ressortir simplement et rapidement quelques tendances intéressantes. La Figure 4.6 présente avec la courbe en blanc l'évolution de la perte de masse cumulée du groupe (estimation à partir du système de monitoring) en fonction du temps pour une année d'opération. Les franges de couleur superposées représentent les différentes grappes associées aux MO rencontrés. On peut ainsi constater que le groupe T/A étudié fonctionne majoritairement dans les MO de la grappe 5 (en violet).

Afin de visualiser l'évolution de la dégradation au cours du temps pour chaque grappe, les incréments de dégradation ont été concaténés par grappe en formant des trajectoires. Cette concaténation a été réalisée à partir des données corrigées (après application du modèle du fabricant et en retirant les incréments de temps incohérents - colonne "Mesures corrigées" du Tableau 4.2). Pour cette raison, on remarquera que les instants cumulés (en jours) de l'ensemble des grappes ne totalisent pas 365 jours. Les Figures 4.7 comparent les trajectoires de dégradation

pour les différentes grappes (la Figure du bas est un agrandissement du cadre en pointillés). Les droites en pointillés servent de repère pour identifier une éventuelle convexité des trajectoires de dégradation. On peut ainsi observer que les trajectoires des grappes 1 et 5 semblent ne pas adopter un comportement linéaire, alors que cette tendance est moins évidente pour les autres trajectoires/grappes. Aussi la trajectoire de la grappe 2 est semblable en termes de pente initiale à celle de la grappe 1. Quant à la grappe 4 (en rouge), sa contribution à la perte de masse cumulée est bien moins significative que pour les incréments de dégradation des autres grappes (environ 2% de la masse totale perdue sur une année).

Afin de pouvoir comparer les incréments de dégradation pour chaque grappe, ceux-ci ont été homogénéisés (normalisés) en les divisant par leur incrément de temps associé. Cette opération permet de comparer les données de dégradation des différentes grappes en se ramenant à une intensité de cavitation (cette fois-ci exprimée en kg/jour). Les distributions des intensités de cavitation ainsi obtenues sont tracées sur les Figures 4.8. Par souci de lisibilité, certaines distributions sont représentées à plusieurs reprises pour comparaison. Ces distributions d'intensité de cavitation viennent entre autres confirmer certaines observations faites à partir de la Figure 4.7 : les incréments de dégradation par unité de temps pour les grappes 1 et 2 partagent une partie de leur distribution ; aussi la distribution pour la grappe 3 est en moyenne à droite de celle de la grappe 2 ce qui explique la dynamique plus lente de la dégradation pour cette dernière. On observe également une forte disparité pour la grappe 1 des intensités de cavitation comparativement aux autres grappes. Les distributions associées aux grappes 2 et 3 montrent aussi une similarité quand on considère leur tendance moyenne. On notera que pour les grappes présentant graphiquement des distributions semblables d'intensité de cavitation, des tests de Kolmogorov-Smirnov ont été réalisés. Les résultats de ces tests montrent que chaque grappe de MO semble donc disposer d'une distribution d'intensité de cavitation propre.

N. B. : On soulignera que la pertinence de ces tests peut être questionnée au regard de la taille de l'échantillon testé (très grand) et pour les raisons évoquées en section 1.4.

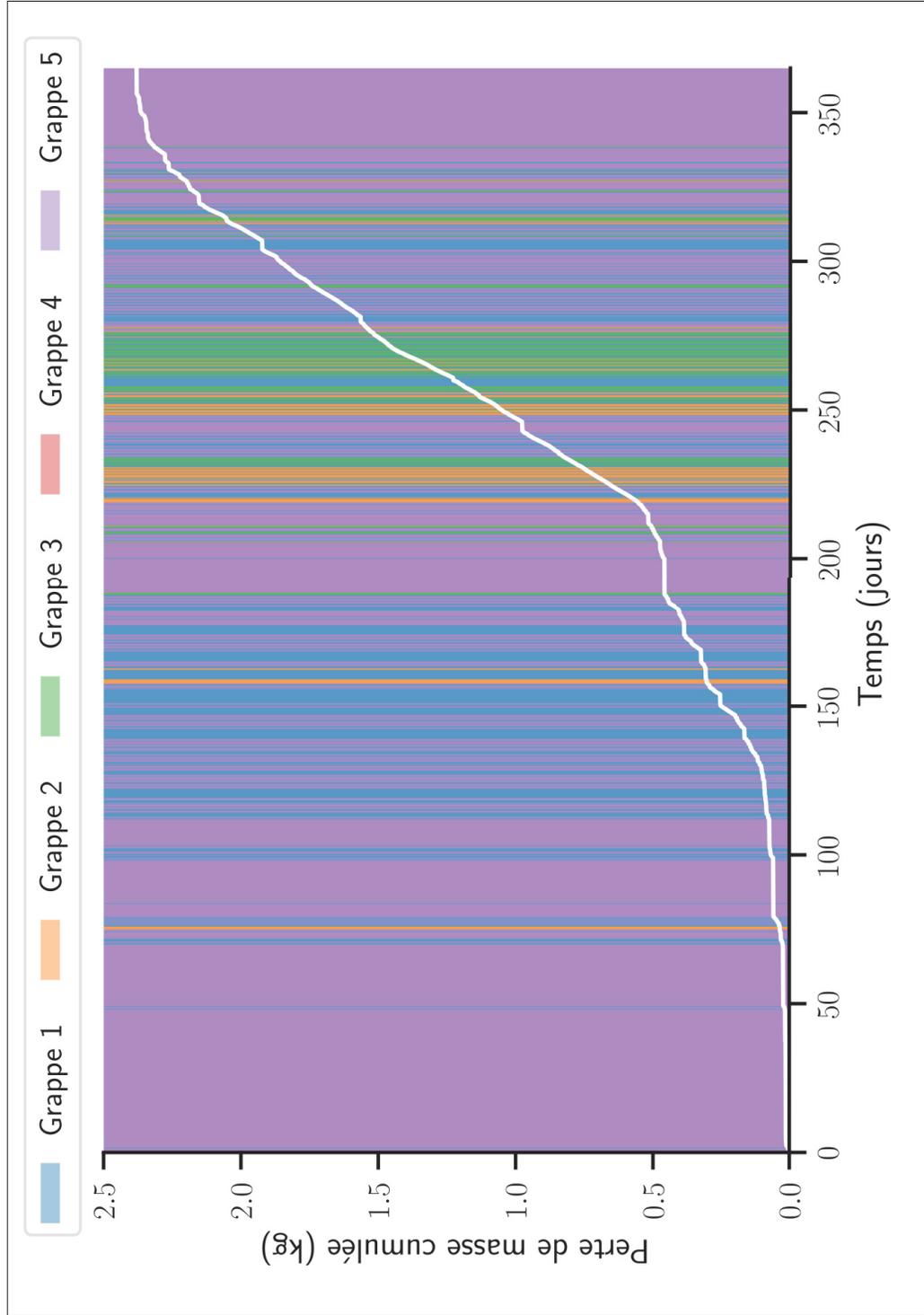


Figure 4.6 Trajectoire de dégradation du groupe T/A étudié pour l'année 2011 - Perte de masse cumulée en fonction du temps.

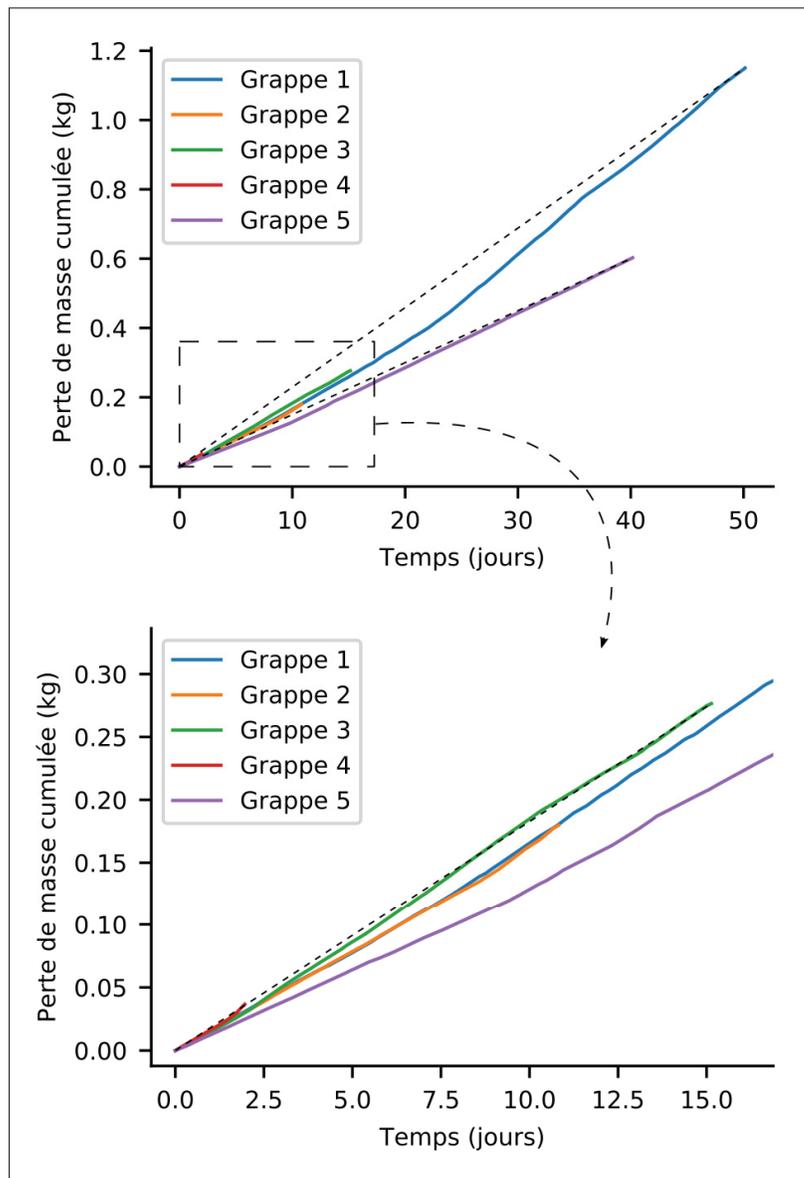


Figure 4.7 Trajectoires de dégradation par grappe pour l'année 2011 - Perte de masse cumulée en fonction du temps.

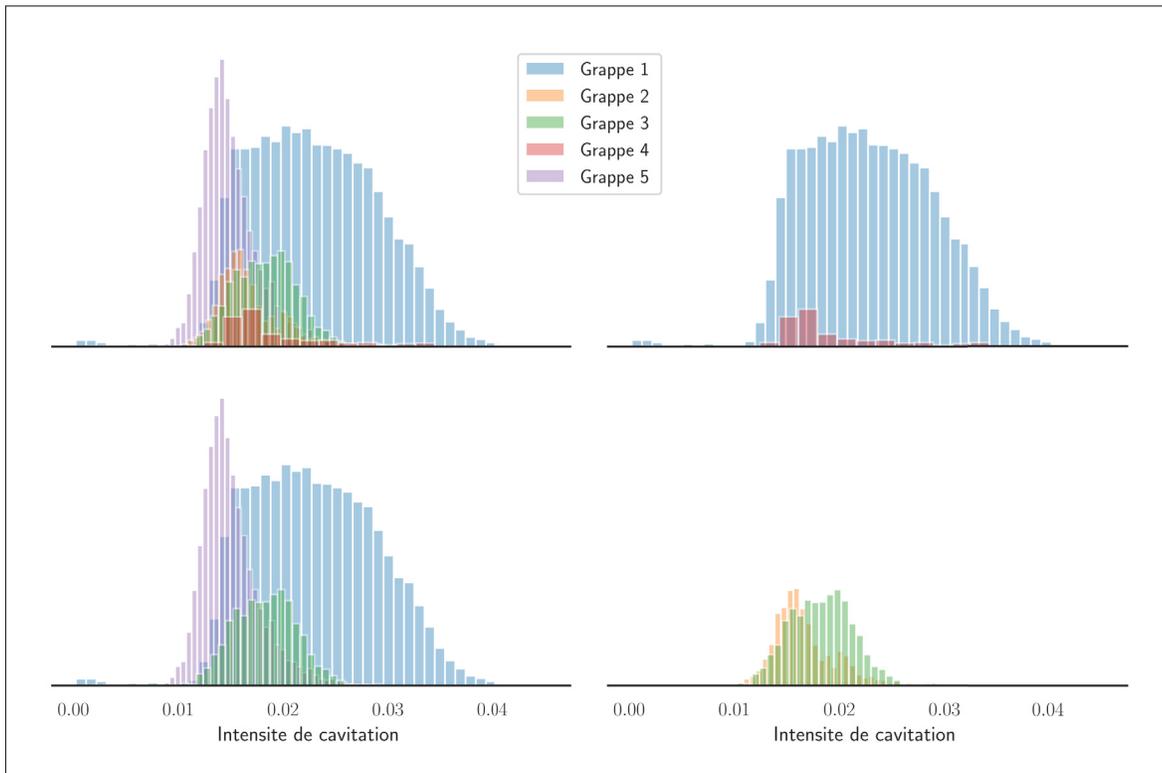


Figure 4.8 Comparaison des intensités de cavitation par grappe.

Afin de pouvoir mettre en évidence certaines tendances dans les données, pour chaque grappe ont été calculés : le temps passé dans la grappe et la moyenne des incréments de dégradation par unité de temps de la grappe. On notera que le temps précédemment introduit est calculé à partir des mesures corrigées du Tableau 4.2. La moyenne empirique des incréments normalisés $\delta x_{i,c} / \delta t_{i,c}$, $1 \leq i \leq n_c$, $1 \leq c \leq 5$, notée $\overline{\delta x_{norm,c}}$ est exprimée en [g/h] et s'obtient avec l'expression suivante :

$$\overline{\delta x_{norm,c}} = \frac{\sum_{i=1}^{n_c} \delta x_{i,c}}{n_c \sum_{i=1}^{n_c} \delta t_{i,c}}, \quad (4.4)$$

avec $\delta t_{i,c}$ l'incrément de temps associé à l'incrément de dégradation $\delta x_{i,c}$ et n_c le nombre totale de mesures corrigées appartenant à la grappe c , $1 \leq c \leq 5$. Les résultats pour l'année 2011 sont présentés dans le Tableau 4.3.

Tableau 4.3 Analyse des modes opératoires et des incréments de dégradation associés pour l'année 2011.

Grappe	1	2	3	4	5
Temps passé dans la grappe (jours)	50,13	10,82	15,13	1,96	40,16
$\overline{\delta x}_{norm}$ (g/h)	0,96	0,69	0,76	0,78	0,63

À partir du tableau précédent, on peut constater que le temps passé dans les différentes grappes est assez inégal en plus de présenter des disparités en termes d'incrément de dégradation moyen (normalisé).

4.3 Modélisation par processus gamma - Année 2011

4.3.1 Estimation des paramètres du processus

Pour chaque grappe de l'année 2011, les paramètres d'un processus gamma non homogène ont été estimés à l'aide de la méthode en deux temps (*cf.* section 2.3). Les valeurs obtenues sont résumées dans le Tableau 4.4. À titre indicatif, les valeurs d'espérance et de variance (à $t = 100$ jours) associées sont aussi données.

Tableau 4.4 Estimation des paramètres du processus pour chaque grappe.

Grappe	1	2	3	4	5
\hat{u}	76952	245971	223953	94593	230421
\hat{c}	848	2911	2411	1456	2257
\hat{b}	1,12	1,05	1,08	1,03	1,07
Espérance à $t = 100$	1,890	1,507	1,564	1,784	1,326
Variance à $t = 100$	$2,46 \cdot 10^{-5}$	$6,13 \cdot 10^{-6}$	$6,98 \cdot 10^{-6}$	$1,89 \cdot 10^{-5}$	$5,75 \cdot 10^{-6}$

Quelques observations rapides à partir du Tableau 4.4 peuvent être réalisées. Tout d'abord, comme pressenti dans la section 4.2, la grappe 1 présente la convexité la plus importante, mais aussi l'espérance (à $t = 100$ jours) la plus élevée. Si on considère la tendance moyenne des

distributions présentées sur la Figure 4.8, on peut vérifier qu'elles sont rangées dans le même ordre que les espérances du Tableau 4.4 : par exemple, la grappe 5 possède la tendance moyenne la plus faible en accord avec la valeur de son espérance. Ces observations permettent *a priori* de vérifier l'absence d'erreur manifeste dans l'estimation des paramètres des différents processus gamma. Aussi, on notera, pour l'ensemble des grappes, une valeur de variance très faible. Ce résultat peut sembler contre intuitif au regard des données modélisées, issues de systèmes d'exploitation réels qu'on peut supposer entachées d'une variabilité importante. Enfin, on vérifie bien que les grappes 2 et 3 montrent des modèles proches l'un de l'autre.

4.3.2 Calcul des intervalles de confiance sur les paramètres

Il est possible à partir de la matrice de Fisher de déterminer des intervalles de confiance (IC) sur les paramètres des processus gamma estimés pour chaque grappe. La méthode de calcul complète est détaillée dans la section 1.5. Pour rappel, cette matrice est obtenue à partir de la matrice hessienne associée à la fonction log-vraisemblance d'un processus gamma. Pour l'estimation du paramètre b par la méthode des moindres carrés, il est aussi possible d'obtenir la matrice hessienne et de construire les intervalles de confiance associés à ce paramètre. L'équation (1.37) est utilisée pour déterminer les IC présentés dans le Tableau 4.5. Afin de permettre une analyse plus aisée, les IC symétriques sont exprimés en pourcentage du paramètre estimé $\hat{\theta}_i$ (valeur centrale). En effet, pour ce type d'analyse, la largeur relative de l'intervalle est plus importante que la valeur intrinsèque des bornes de cet intervalle. On notera que les IC associés au paramètre b , issus de l'estimation par la méthode MV, sont notés ϵ en raison de leur largeur d'intervalle très faible.

À partir du Tableau 4.5, il est possible de vérifier les hypothèses formulées dans la section 4.2. En effet, pour chaque grappe, il a été choisi de modéliser la dégradation associée par un processus gamma non homogène. Cette hypothèse signifie que le paramètre $b \neq 1$. En considérant l'IC associé au paramètre b pour chaque grappe, on vérifie alors si la valeur $b = 1$ est bien exclue de l'IC. Le cas échéant, on ne peut pas rejeter l'hypothèse que la dégradation de la grappe en question peut être modélisée par un processus gamma homogène. C'est le cas par exemple pour

Tableau 4.5 $IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2011.

Grappe	1	2	3	4	5
$IC_{MV}(u) = \hat{u}$	$\pm 4,03\%$	$\pm 8,40\%$	$\pm 6,71\%$	$\pm 25,59\%$	$\pm 4,13\%$
$IC_{MV}(c) = \hat{c}$	$\pm 3,41\%$	$\pm 7,39\%$	$\pm 7,71\%$	$\pm 20,81\%$	$\pm 4,25\%$
$IC_{MV}(b) = \hat{b}$	$\pm \epsilon \%$				
$IC_{MC}(b) = \hat{b}$	$\pm 2,59\%$	$\pm 6,00\%$	$\pm 4,91\%$	$\pm 16,5\%$	$\pm 3,36\%$

les grappes 2 et 4 du Tableau 4.4, si on prend l'intervalle de confiance construit à partir de la méthode des moindres carrés. Les Tableaux 4.4 et 4.5 permettent également de déduire que les grappes 2 et 3 partagent une partie de leur IC pour les 3 paramètres : les données de ces grappes pourraient éventuellement être regroupées sous une seule grappe.

Aussi, il est possible de justifier l'emploi d'un processus gamma homogène ou non homogène en comparant les trajectoires de dégradation simulées à celles des données de dégradation réelles (voir l'algorithme de simulation présenté en détail dans la section suivante). La Figure 4.9 présente ce travail de comparaison : en bleu est représentée la trajectoire de dégradation réelle pour l'année 2011 ; en vert, une trajectoire d'une simulation par des processus gamma homogènes et en orange par des processus gamma non homogènes (pour chaque grappe, un processus gamma de paramètres estimés donnés est simulé). Bien que les deux modélisations estiment correctement le niveau de dégradation final, les trajectoires de celles-ci s'éloignent de la trajectoire de référence entre les jours 150 et 250. On observe néanmoins que l'écart entre trajectoire simulée et trajectoire de référence est moindre pour la modélisation par un processus gamma non homogène.

N.B. : La figure précédente présente uniquement une trajectoire pour chaque modélisation par souci de lisibilité. On souligne ici, qu'un faisceau de trajectoires donnerait une meilleure idée du comportement de chaque modélisation. Néanmoins, la variance associée aux jeux de paramètres estimés (et utilisés pour la simulation des trajectoires) étant très faible, l'enveloppe des trajectoires associée à chaque modélisation est très étroite. Ainsi, le comportement de

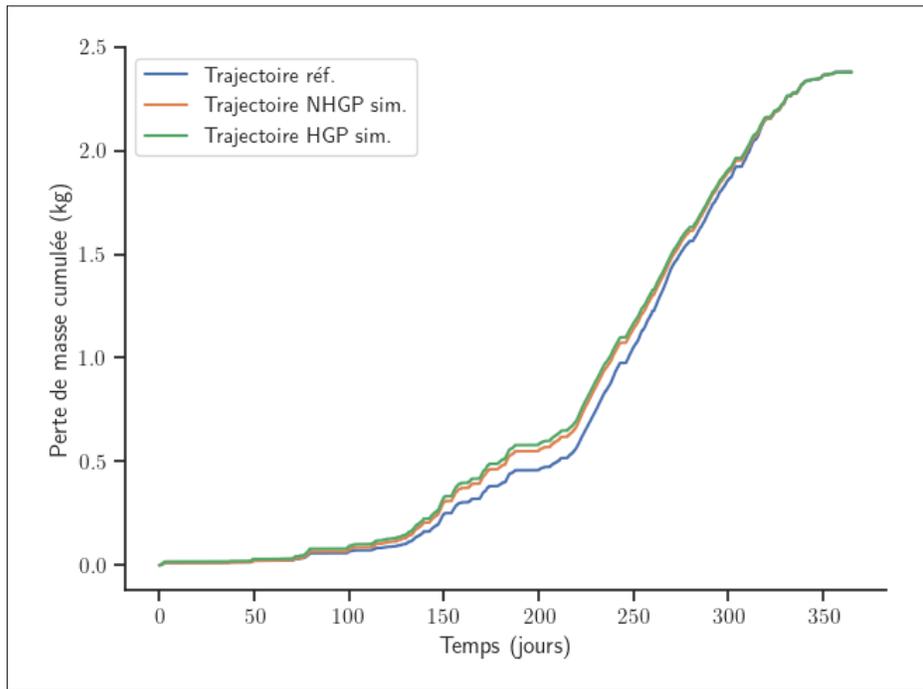


Figure 4.9 Comparaison des trajectoires simulées à partir des différents processus gamma.

ces deux trajectoires simulées reflète la tendance qui serait observée avec des enveloppes de trajectoires.

4.3.3 Vérification - Simulation de trajectoires

En raison de la non homogénéité des incréments de temps, il est impossible de réaliser un test statistique d'adéquation sur les incréments de dégradation associés pour vérifier s'ils sont ou non distribués suivant une loi gamma ; ce qui permettrait de justifier ou au contraire d'écarter un choix particulier de modélisation (voir section 1.4). Pour pallier ce problème, il est choisi de vérifier la cohérence des estimations des paramètres par simulation de faisceaux de trajectoires. Cette simulation est réalisée en prenant comme paramètres du processus gamma les paramètres précédemment estimés pour les données de l'année 2011. La méthode de simulation employée est décrite dans le Chapitre 1 et repose sur l'algorithme *GSS*. Dans le cas présent, il est nécessaire de prendre pour la simulation les mêmes incréments de temps que ceux utilisés pour estimer les

paramètres. L'algorithme de simulation d'une trajectoire de dégradation indexé sur les données réelles est donné en Figure 4.10. On peut noter en particulier que l'incrément de dégradation simulé inc_deg_i entre deux instants t_i et t_{i-1} est généré en tenant compte du numéro de grappe c entre ces deux instants (et des paramètres estimés, notés $\hat{\theta}_{c,2011}$, $1 \leq c \leq 5$, du processus gamma associé à cette grappe).

L'algorithme précédemment présenté a été utilisé pour simuler un faisceau de 100 trajectoires avec pour paramètre $\hat{\theta}_{c,2011}$, associé à la grappe c , $1 \leq c \leq 5$. La Figure 4.11 présente l'enveloppe des trajectoires de dégradation simulées (en gris) ainsi que la trajectoire de celle observée pour l'année 2011 (en noir). La Figure du bas détaille l'agrandissement du cadre en pointillés de la Figure du haut. On remarque qu'avec une variance faible impliquée par les valeurs des paramètres $\hat{\theta}_{c,2011}$, $1 \leq c \leq 5$ (voir Tableau 4.4), l'enveloppe des trajectoires simulées est très étroite. La vue de détail montre bien la variabilité des trajectoires simulées (bien que très faible). Cette enveloppe n'inclut pas la trajectoire de référence sur l'intervalle de temps 100 à 300 jours, traduisant les limites de la présente modélisation.

4.3.4 Correction de la variance des processus

La modélisation de la dégradation par un processus gamma a été en partie motivée par sa capacité à "capturer" la variabilité du phénomène de dégradation étudié. Néanmoins, comme présenté dans la section précédente, les valeurs estimées des paramètres pour chaque grappe impliquent une variance très faible et donc une variabilité des trajectoires simulées qui suit cette tendance. Ce résultat est intuitivement étonnant quand on considère que les données modélisées sont extraites d'un système de suivi de cavitation installé sur un groupe T/A exploité de façon industrielle. Le phénomène de dégradation étant sensible aux conditions d'opération, on pourrait s'attendre à plus de variabilité dans les données modélisées. Une explication de ce résultat pourrait être imputé au système de suivi de cavitation en temps réel et à la méthode d'échantillonnage de son signal. En effet, une mesure est réalisée à un instant t pendant une durée δt . Le signal mesuré est ensuite filtré, moyenné puis passé à travers une fonction de transfert pour aboutir à une intensité

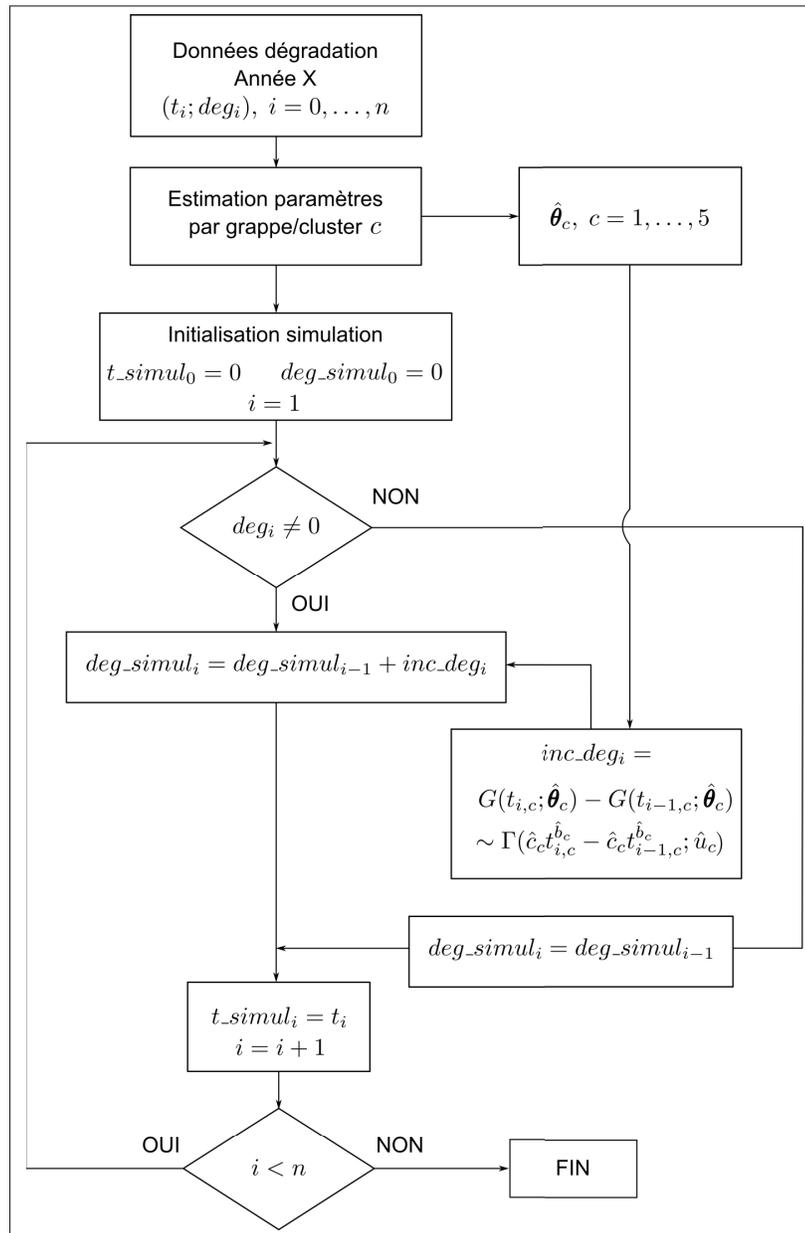


Figure 4.10 Algorithme de simulation de trajectoires indexé sur les données réelles.

de cavitation. Ces différentes étapes peuvent impliquer un "lissage" des intensités de cavitation en bout de chaîne du système et donc une plus faible variabilité dans les valeurs observées.

Afin d'introduire artificiellement plus de variabilité dans les trajectoires simulées, et donc obtenir une enveloppe plus large, il est proposé de modifier les valeurs de paramètres estimés $\hat{\theta}_{c,2011}$,

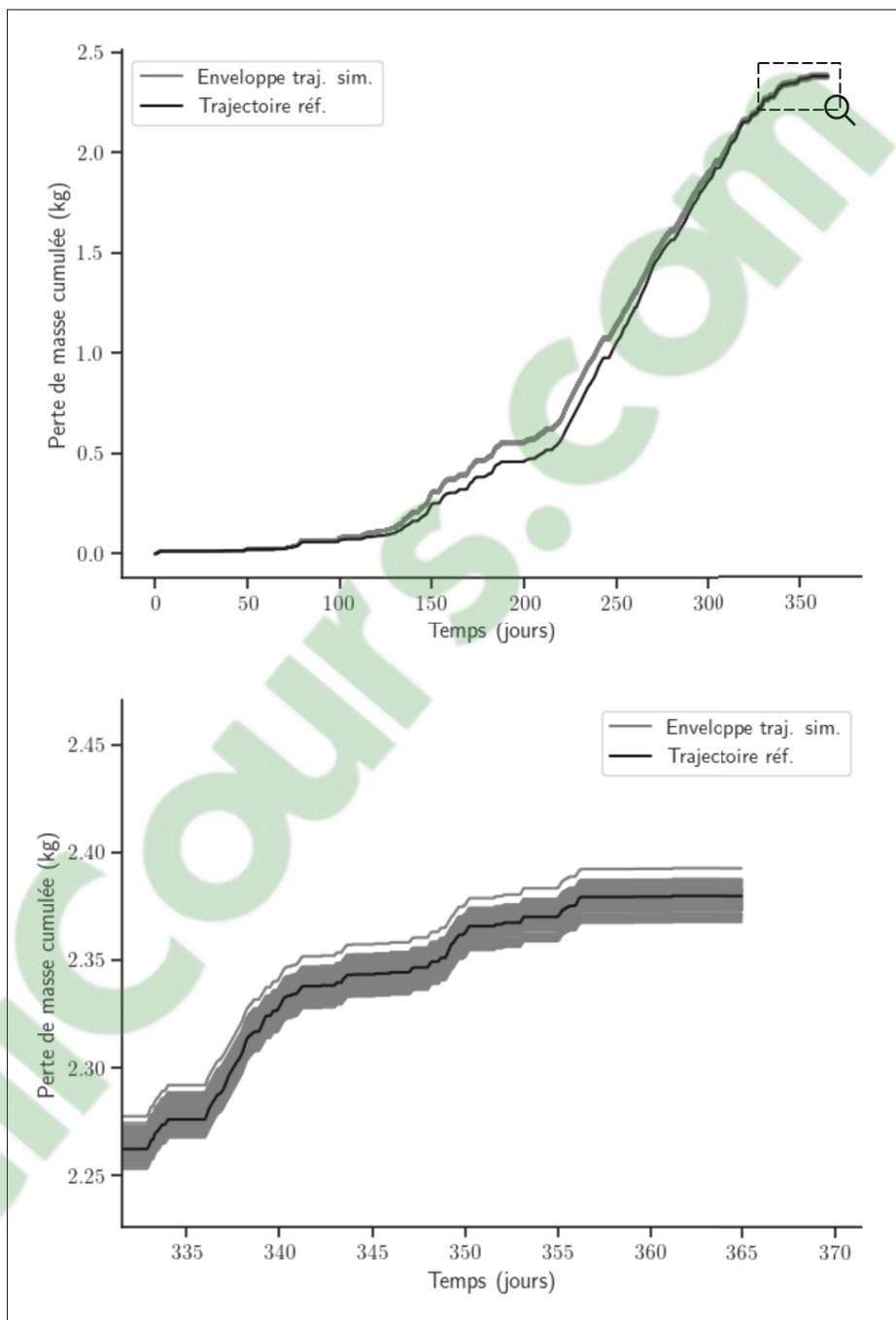


Figure 4.11 Enveloppe des trajectoires simulées - comparaison à la trajectoire de référence pour l'année 2011.

$1 \leq c \leq 5$, de telle façon que l'espérance soit conservée mais la variance modifiée par un facteur

k . Pour cela, les expressions de l'espérance et de la variance sont rappelées :

$$\mathbb{E}(X(t)) = \frac{v(t)}{u} = \frac{ct^b}{u}, \quad \mathbb{V}(X(t)) = \frac{v(t)}{u^2} = \frac{ct^b}{u^2} \quad (4.5)$$

Les expressions (4.5) montrent que le ratio c/u doit être conservé pour garantir une espérance constante. Ainsi, si les paramètres u et c sont multipliés par un facteur k , on obtient bien :

$$\mathbb{E}_{mod}(X(t)) = \frac{ckt^b}{uk} = \frac{ct^b}{u}, \quad \mathbb{V}_{mod}(X(t)) = \frac{ckt^b}{(uk)^2} = \frac{1}{k} \frac{ct^b}{u^2} = \frac{1}{k} \mathbb{V}(X(t)) \quad (4.6)$$

La variance modifiée, notée $\mathbb{V}_{mod}(X(t))$, est multipliée par $1/k$ lorsqu'on multiplie les paramètres u et c par un facteur k . L'espérance, quant à elle, est bien conservée.

Les Figures 4.12 et 4.13 montrent l'enveloppe de trajectoires simulées pour des valeurs de variance multipliées respectivement par 10, 100 et 1000, soit des paramètres u et c multipliés par $1/10$, $1/100$ et $1/1000$. Pour les trois figures, 100 trajectoires avec les nouveaux paramètres ont été générées. Comme anticipé, plus la variance est importante, plus l'enveloppe des trajectoires simulées est large. On remarque que la trajectoire de référence est incluse dans l'enveloppe des trajectoires simulées pour une valeur de $k = 1/1000$. Néanmoins, l'enveloppe présente toujours un biais positif dans l'intervalle [100; 300] jours.

4.4 Vérification et validation avec les données disponibles

L'objectif de cette section est de présenter une méthode graphique, reposant sur la simulation de faisceaux de trajectoires, permettant la validation du modèle par rapport aux données de dégradation disponibles. Pour cela, on s'intéresse dans un premier temps à vérifier l'adéquation du modèle proposé aux données ayant servi à estimer les paramètres du modèle. Ensuite, ce même modèle est utilisé pour la prédiction de la dégradation future.

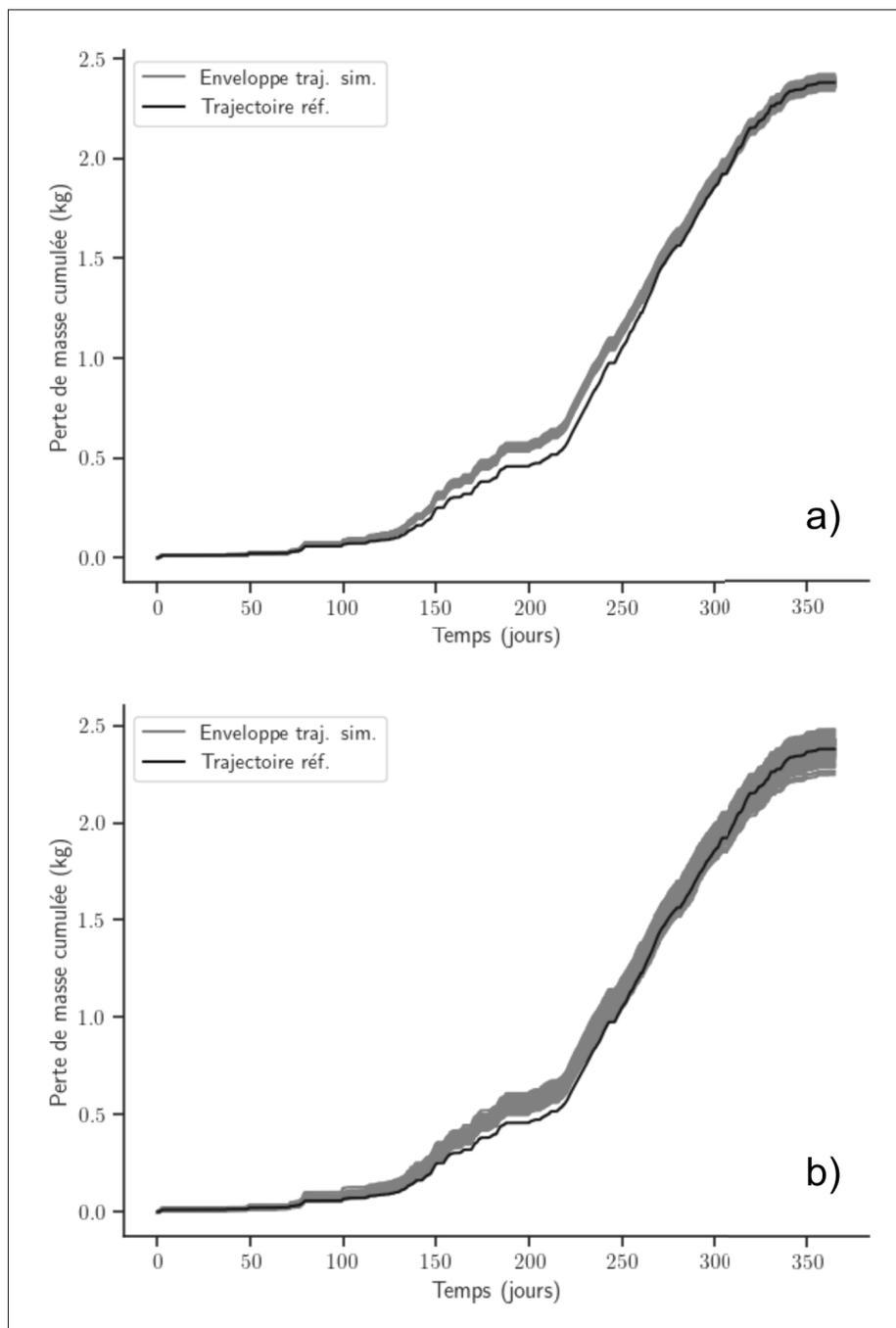


Figure 4.12 Enveloppe des trajectoires simulées pour une variance multipliée par : a) 10 et b) 100 - Trajectoire de référence année 2011.

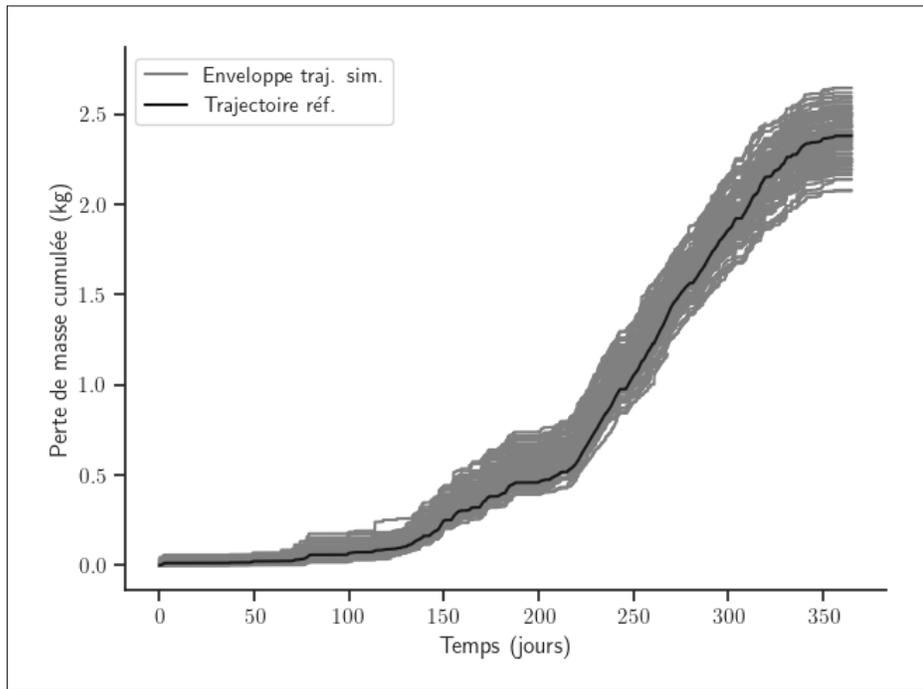


Figure 4.13 Enveloppe des trajectoires simulées pour une variance multipliée par 1000 - Trajectoire de référence année 2011.

4.4.1 Regroupement des MO inconnus

Le regroupement des MO par grappe ayant été réalisé à partir des données de l'année 2011, des MO non répertoriés peuvent apparaître pour les autres années étudiées. Ceci s'explique par la multitude de combinaisons de groupes T/A pouvant fonctionner simultanément : l'ensemble des MO n'a donc pas été observé durant l'année 2011. Afin d'associer ces nouveaux MO aux grappes existantes, une méthodologie simplificatrice a été employée. Pour cela, la moyenne des intensités de cavitation non nulles, notée $\overline{cav_inst}_c$, a été calculée pour chaque grappe c , $1 \leq c \leq 5$ sur l'année 2011. De même, pour chaque nouveau MO apparaissant dans les données des années disponibles, la moyenne des intensités de cavitation non nulles a été calculée, notée $\overline{cav_inst}_{new_MO}$. L'écart, en valeur absolue, entre $\overline{cav_inst}_{new_MO}$ et $\overline{cav_inst}_c$ est calculé selon l'équation (4.7) pour chaque grappe c de l'année 2011.

$$ecart_moy_{new_MO,c} = |\overline{cav_inst}_c - \overline{cav_inst}_{new_MO}| \quad (4.7)$$

Le nouveau mode opératoire est alors rattaché à la grappe c qui minimise $ecart_moy_{new_MO,c}$. Cette méthode de regroupement *a posteriori* présente l'avantage de s'affranchir de la lourdeur numérique des méthodes de regroupement dynamiques qui consistent à refaire le regroupement à chaque nouvelle apparition d'un cas non répertorié. De plus, dans le présent contexte industriel, il est impossible de balayer (ou d'observer), sur une année donnée, l'ensemble des combinaisons de groupes T/A fonctionnant simultanément et formant un MO.

Les Figures 4.14 et 4.15 montrent la répartition des MO au cours du temps ainsi que la dégradation cumulée pour les années 2012 et 2013 respectivement. On remarque aisément à partir de ces figures que l'année 2012 présente macroscopiquement une répartition semblable des modes opératoires au cours de l'année à celle de 2011. En revanche, pour l'année 2013, le fonctionnement du groupe T/A dans les modes associés à la grappe 3 est nettement plus marqué que pour les deux précédentes. On peut donc en déduire que le groupe a été opéré différemment lors de l'année 2013. Cette observation est confirmée par les résultats d'analyse des MO présentés dans les Tableaux 4.6 et 4.7.

Tableau 4.6 Analyse des modes opératoires et des incréments de dégradation associés pour 2012.

Grappe	1	2	3	4	5
Temps passé dans la grappe (jours)	44,95	5,27	4,75	0,55	60,73
$\overline{\delta x_{norm}}$ (g/h)	0,89	0,82	0,68	0,64	0,71

Tableau 4.7 Analyse des modes opératoires et des incréments de dégradation associés pour l'année 2013.

Grappe	1	2	3	4	5
Temps passé dans la grappe (jours)	18,73	45,95	58,84	6,07	43,54
$\overline{\delta x_{norm}}$ (g/h)	0,90	0,62	0,54	0,75	0,54

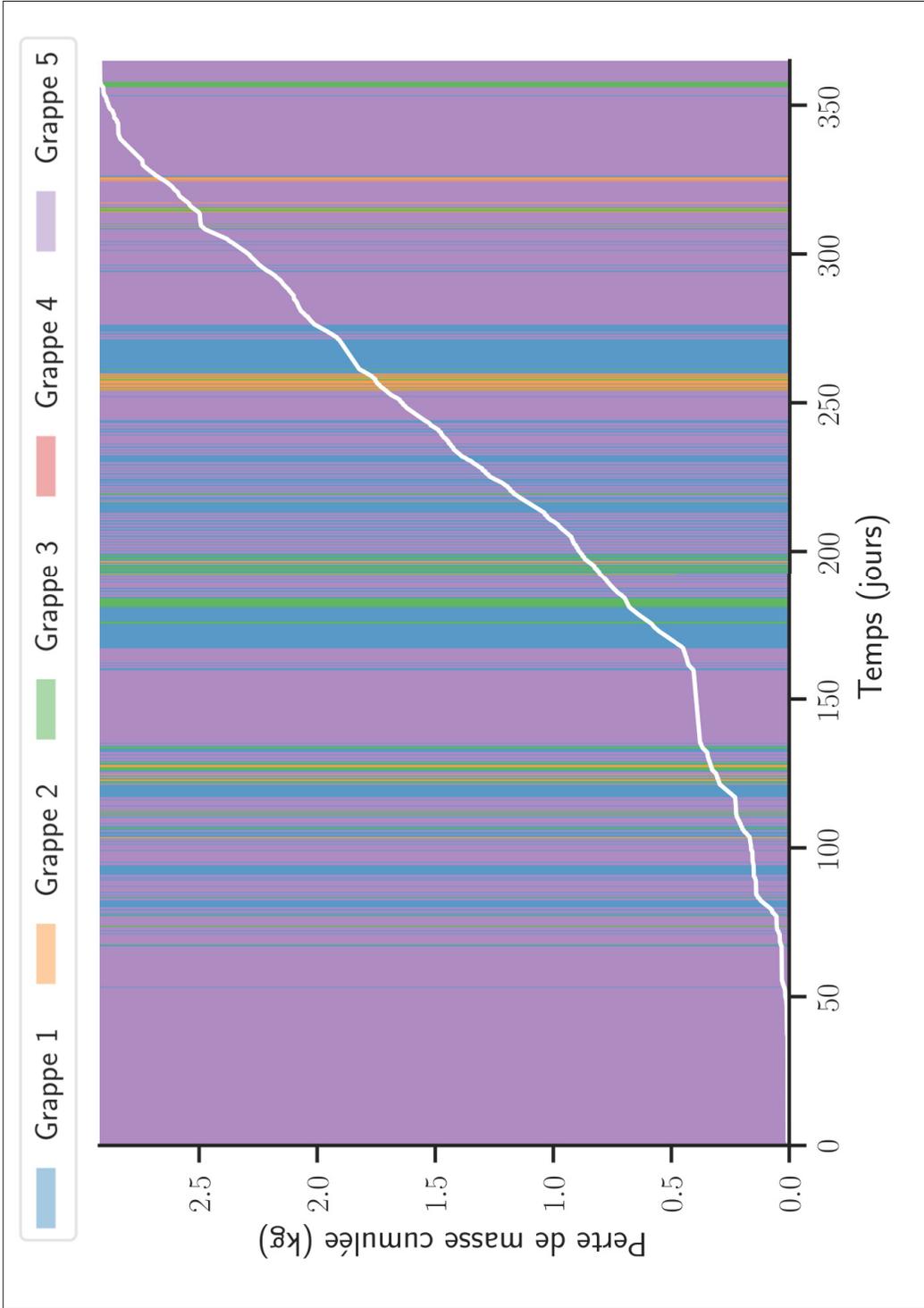


Figure 4.14 Trajectoire de dégradation du groupe T/A étudié pour l'année 2012 - Perte de masse cumulée en fonction du temps.

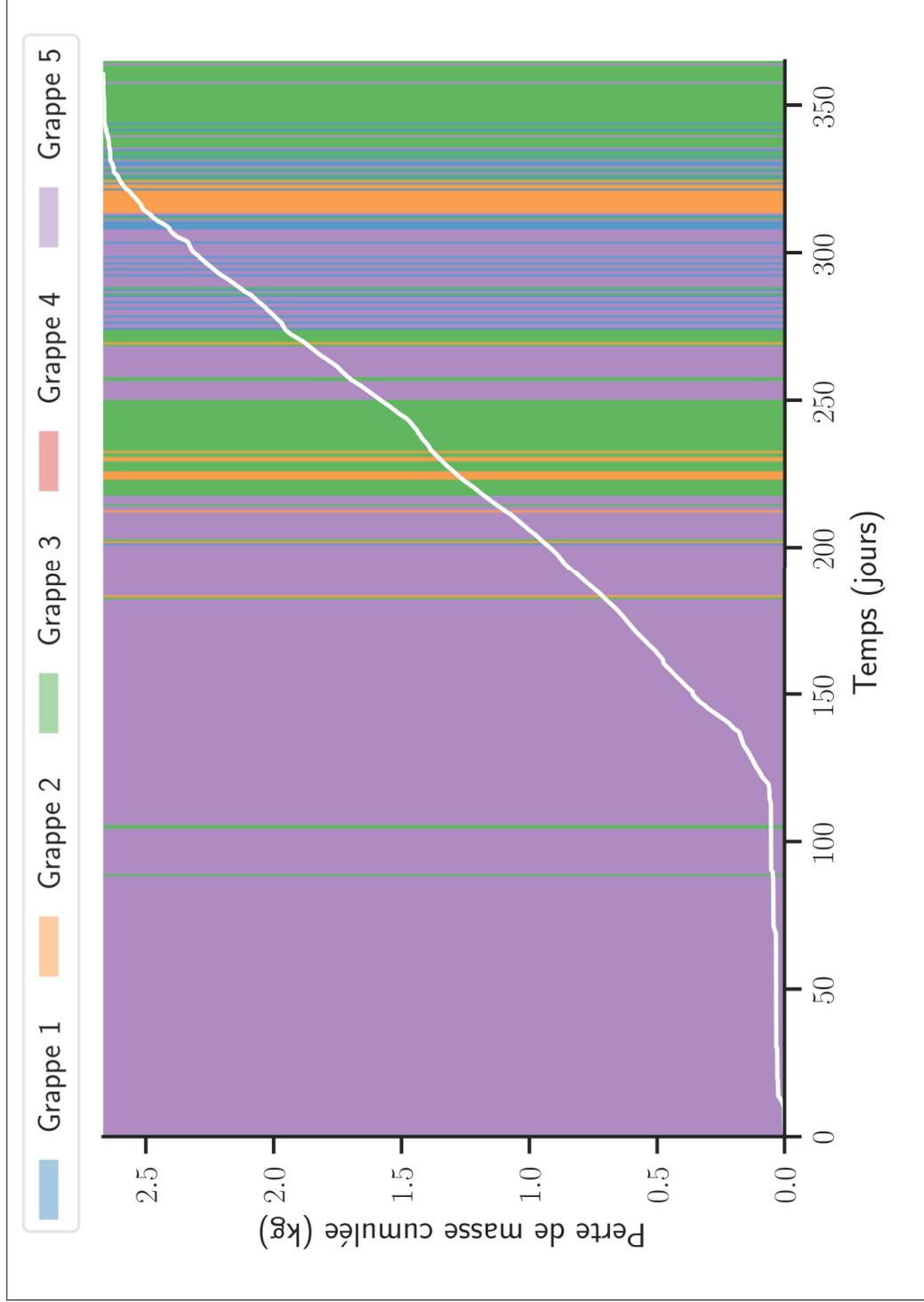


Figure 4.15 Trajectoire de dégradation du groupe T/A étudié pour l'année 2013 - Perte de masse cumulée en fonction du temps.

Afin de compléter l'étude de chacune des années 2012 et 2013, une modélisation des trajectoires par un processus gamma non homogène est réalisée à l'aide de leurs données respectives et en suivant le même processus que pour l'année 2011. Les résultats présentés avec les IC associés dans les Tableaux 4.8 à 4.11, permettent de comparer la dynamique de la dégradation pour chaque année.

Tableau 4.8 Estimation des paramètres du processus pour chaque grappe - Données de l'année 2012.

Grappe	1	2	3	4	5
\hat{u}	78346	181673	360452	142083	221499
\hat{c}	1305	2279	4303	1547	2715
\hat{b}	1,04	1,07	1,05	1,06	1,05
Espérance à $t = 100$	2,003	1,718	1,491	1,429	1,554
Variance à $t = 100$	$2,56.10^{-5}$	$9,45.10^{-6}$	$4,14.10^{-6}$	$1,01.10^{-5}$	$7,02.10^{-6}$

Tableau 4.9 $IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2012.

Grappe	1	2	3	4	5
$IC_{MV}(u) = \hat{u}$	$\pm 4,46\%$	$\pm 13,62\%$	$\pm 15,32\%$	$\pm 46,86\%$	$\pm 3,39\%$
$IC_{MV}(c) = \hat{c}$	$\pm 3,75\%$	$\pm 11,42\%$	$\pm 13,62\%$	$\pm 38,87\%$	$\pm 3,34\%$
$IC_{MV}(b) = \hat{b}$	$\pm \epsilon \%$				
$IC_{MC}(b) = \hat{b}$	$\pm 1,06\%$	$\pm 1,19\%$	$\pm 1,14\%$	$\pm 2,48\%$	$\pm 1,02\%$

On soulignera que les modélisations proposées ici pour les données de 2012 et 2013 respectivement, considèrent une dégradation nulle à l'instant associé t_0 (premier instant observé). Néanmoins, étant donné que le matériel ne subit pas d'opération de maintenance entre les différentes années, le niveau de dégradation au début de l'année $a + 1$ est dans la réalité égal à celui à la fin de l'année a . Il en est de même pour les instants observés. Pour des considération de pronostic, dans la section suivante, les données des années étudiées seront regroupées afin de proposer un modèle permettant d'estimer un niveau de dégradation futur.

Tableau 4.10 Estimation des paramètres du processus pour chaque grappe - Données de l'année 2013.

Grappe	1	2	3	4	5
\hat{u}	147069	362373	227524	241083	367539
\hat{c}	2229	6517	3411	2901	6449
\hat{b}	1,06	0,97	0,98	1,06	0,95
Espérance à $t = 100$	1,96	1,56	1,35	1,60	1,41
Variance à $t = 100$	$1,34.10^{-5}$	$4,32.10^{-6}$	$5,93.10^{-6}$	$6,66.10^{-6}$	$3,84.10^{-6}$

Tableau 4.11 $IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données de l'année 2013.

Grappe	1	2	3	4	5
$IC_{MV}(u) = \hat{u}$	$\pm 3,05\%$	$\pm 8,96\%$	$\pm 6,11\%$	$\pm 19,59\%$	$\pm 3,44\%$
$IC_{MV}(c) = \hat{c}$	$\pm 2,56\%$	$\pm 7,52\%$	$\pm 5,58\%$	$\pm 16,42\%$	$\pm 2,90\%$
$IC_{MV}(b) = \hat{b}$	$\pm \epsilon \%$				
$IC_{MC}(b) = \hat{b}$	$\pm 1,63\%$	$\pm 2,69\%$	$\pm 2,86\%$	$\pm 5,64\%$	$\pm 1,54\%$

4.4.2 Prédiction de la dégradation à partir de la modélisation de l'année 2011

L'objectif de cette section est de montrer dans quelle mesure la modélisation réalisée à partir des données de l'année 2011 est capable de prédire la dégradation des années suivantes. Pour cela, les trajectoires des années 2012 et 2013 sont exploitées à des fins de comparaison. Cette étude repose sur l'hypothèse qu'à partir de la connaissance du MO et de sa grappe associée, il est possible de simuler un faisceau de trajectoires avec pour paramètres les valeurs estimées pour l'année 2011. La Figure 4.16 montre les trajectoires réelles de la dégradation pour les années 2011 à 2013.

N.B. : On notera que la figure précédente a pour but principal de comparer graphiquement le comportement de la dégradation entre les différentes années. Il est par ailleurs admis chez les experts qu'un phénomène de saisonnalité est observé dans le comportement de

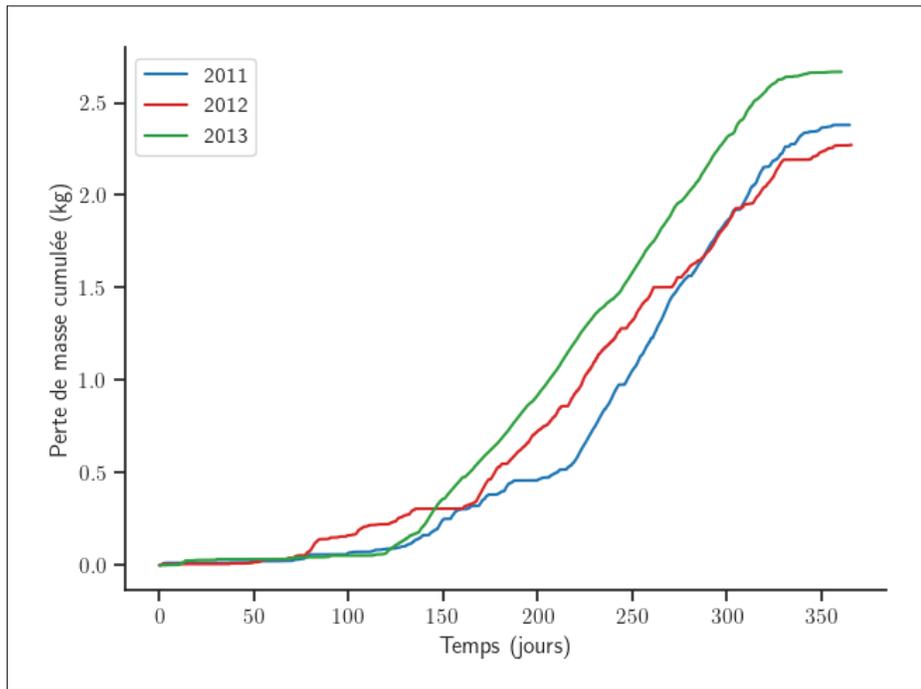


Figure 4.16 Trajectoires de dégradation pour les années 2011 à 2013 - Perte de masse cumulée en fonction du temps.

la cavitation apparaissant en centrale. Cette saisonnalité de la dégradation a d'ailleurs été montrée statistiquement dans l'étude de Bodson-Clermont (2016). La Figure 4.16 montre macroscopiquement deux régimes principaux : un premier avec une vitesse de dégradation lente et un deuxième avec une vitesse de dégradation plus importante, correspondant respectivement aux saisons hivernales et estivales (le découpage en 4 saisons est moins marqué au Québec, en particulier pour les régions au nord, où sont installés les équipements étudiés). Néanmoins, on retiendra que cette représentation ne considère pas la continuité des trajectoires de dégradation d'une année à l'autre (comme c'est cette fois le cas dans la Figure 4.17). Ainsi, la trajectoire d'une année suivante ne débute pas à l'origine du repère mais, bien à la fin de la trajectoire de l'année précédente en raison, notamment de l'absence de remise en état initial du matériel (par une éventuelle opération de maintenance).

Pour la Figure 4.17, 50 trajectoires ayant pour paramètres $\hat{\theta}_{c,2011}$, $1 \leq c \leq 5$, ceux estimés pour l'année 2011 (avec une variance modifiée, multipliée par 1000, voir section 4.3.4), ont été

simulées selon l'algorithme présenté en section 4.3.3. Les données des années 2011 à 2013 ont été considérées dans l'algorithme uniquement pour la boucle de simulation. On soulignera que, pour l'estimation des paramètres $\theta_{c,2011}$, $1 \leq c \leq 5$, les données des années 2012 et 2013 n'ont pas été utilisées.

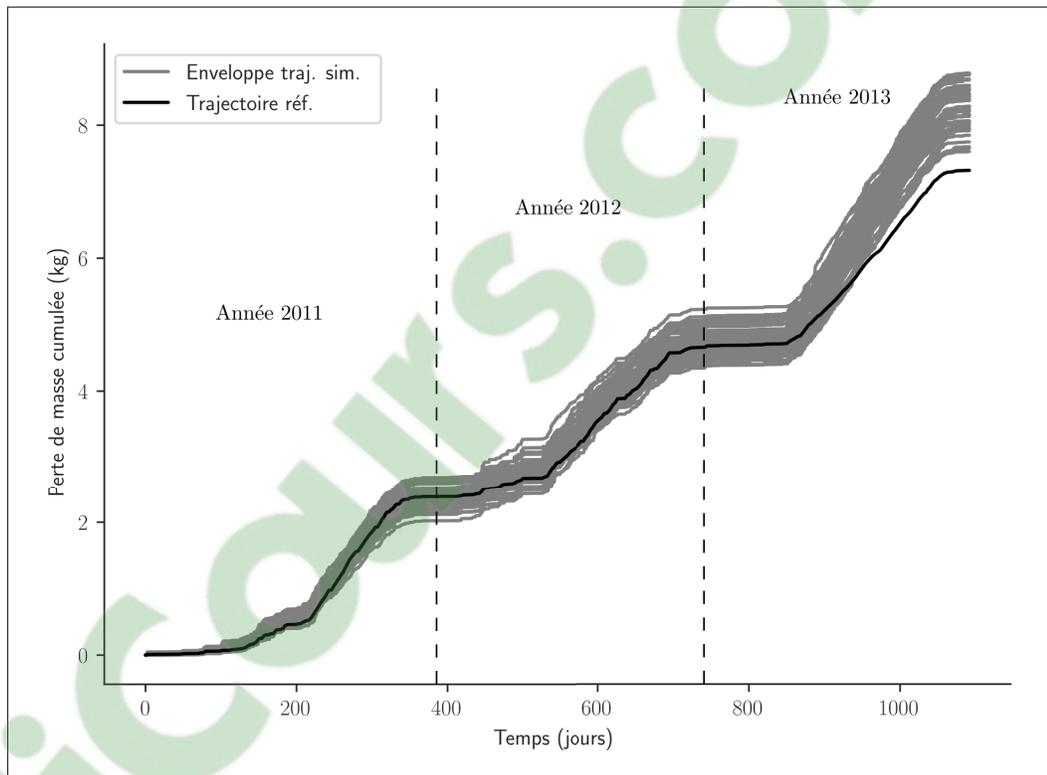


Figure 4.17 Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir de l'année 2011.

On observe à partir de la Figure 4.17 que l'enveloppe des trajectoires simulées est cohérente avec la trajectoire de référence de l'année 2012 : cette dernière est proche de la tendance centrale de l'enveloppe et globalement son comportement est semblable à celui de l'année 2011. Pour rappel, il a été observé dans la section précédente que ces années présentent un comportement semblable en terme de répartition (macroscopique) des MO sur une année. En revanche pour l'année 2013, on peut noter que la trajectoire de référence présente un comportement bien plus lisse, en plus d'une perte de masse sensiblement supérieure, que pour les années précédentes. De plus, les régimes saisonniers semblent plus marqués pour cette année que pour les précédentes.

L'enveloppe des trajectoires simulées présente une tendance à surestimer la dégradation sur une grande partie de l'année 2013. Sur la deuxième moitié de l'année on constate même que la trajectoire de référence est en dehors de l'enveloppe des simulations. Ces observations montrent l'incapacité du modèle estimé, avec les données de l'année 2011, à prédire convenablement la dégradation de l'année 2013.

Pour pallier ce défaut de modélisation, les données des années 2012 et 2013 sont par la suite exploitées pour l'estimation des paramètres θ_c , $1 \leq c \leq 5$. L'hypothèse sous-jacente repose sur la disponibilité d'information caractérisant la dégradation dans les données des années suivantes. Dans un premier temps, les données des années 2011 et 2012 sont concaténées pour former un nouvel échantillon \mathbf{D}_{obs} et estimer les paramètres θ_c à partir de ce dernier échantillon. La concaténation consiste à regrouper les données de dégradation des années sélectionnées de façon à former une seule trajectoire continue en temps et niveau de dégradation. La même procédure est employée en considérant dans un deuxième temps les données de dégradation des années 2011 à 2013 à des fins de comparaison. Les résultats d'estimation sont présentés dans les Tableaux 4.12 et 4.14. Les simulations de trajectoires associées sont quant à elles représentées dans les Figures 4.18 et 4.19 respectivement.

Tableau 4.12 Estimation des paramètres du processus pour chaque grappe - Concaténation des années 2011 et 2012.

Grappe	1	2	3	4	5
\hat{u}	41201	121752	160554	26835	107142
\hat{c}	4633	6238	8617	1110	5803
\hat{b}	0,76	0,84	0,84	0,86	0,83
Espérance à $t = 100$	3,775	2,500	2,547	2,200	2,489
Variance à $t = 100$	$9,16 \cdot 10^{-5}$	$2,05 \cdot 10^{-5}$	$1,58 \cdot 10^{-5}$	$8,19 \cdot 10^{-5}$	$2,32 \cdot 10^{-5}$

On constate au travers des Tableaux 4.12 et 4.14 que, contrairement aux estimations réalisées à partir des données de l'année 2011 uniquement, la fonction de forme $v(t) = ct^b$ présente un caractère concave (avec des valeurs de b statistiquement inférieures à 1, déduit à partir des IC

Tableau 4.13 $IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données des années 2011+2012.

Grappe	1	2	3	4	5
$IC_{MV}(u) = \hat{u}$	$\pm 3,05\%$	$\pm 8,96\%$	$\pm 6,11\%$	$\pm 19,59\%$	$\pm 3,44\%$
$IC_{MV}(c) = \hat{c}$	$\pm 2,56\%$	$\pm 7,52\%$	$\pm 5,58\%$	$\pm 16,42\%$	$\pm 2,90\%$
$IC_{MV}(b) = \hat{b}$	$\pm \epsilon \%$				
$IC_{MC}(b) = \hat{b}$	$\pm 1,63\%$	$\pm 2,69\%$	$\pm 2,86\%$	$\pm 5,64\%$	$\pm 1,54\%$

Tableau 4.14 Estimation des paramètres du processus pour chaque grappe - Concaténation des années 2011, 2012 et 2013.

Grappe	1	2	3	4	5
\hat{u}	41766	240527	189635	58361	134430
\hat{c}	5134	12124	10502	2852	7442
\hat{b}	0,76	0,85	0,82	0,86	0,83
Espérance à $t = 100$	3,984	2,481	2,437	2,616	2,477
Variance à $t = 100$	$9,54 \cdot 10^{-5}$	$1,03 \cdot 10^{-5}$	$1,28 \cdot 10^{-5}$	$4,48 \cdot 10^{-5}$	$1,84 \cdot 10^{-5}$

associés des Tableaux 4.13 et 4.15), alors qu'on observait $b > 1$ pour les données de 2011). Néanmoins, on peut remarquer que l'espérance moyenne de la dégradation pour $t = 100$ jours (de chaque grappe) est supérieure à celle issue des estimations pour les données de 2011 uniquement,

Tableau 4.15 $IC(\theta_i)$ exprimé en pourcentage du paramètre estimé $\hat{\theta}_i$ - Niveau de confiance fixé à 95% - Données des années 2011+2012+2013.

Grappe	1	2	3	4	5
$IC_{MV}(u) = \hat{u}$	$\pm 2,86\%$	$\pm 3,15\%$	$\pm 2,97\%$	$\pm 10,94\%$	$\pm 2,18\%$
$IC_{MV}(c) = \hat{c}$	$\pm 2,41\%$	$\pm 3,16\%$	$\pm 3,55\%$	$\pm 9,08\%$	$\pm 2,11\%$
$IC_{MV}(b) = \hat{b}$	$\pm \epsilon \%$				
$IC_{MC}(b) = \hat{b}$	$\pm 1,46\%$	$\pm 1,50\%$	$\pm 1,45\%$	$\pm 3,04\%$	$\pm 1,30\%$

illustrant le phénomène de compensation des paramètres. On rappelle, que ces trajectoires concaténées ne peuvent pas être considérées comme des répliques de la trajectoire initialement étudiée (celle de 2011). Ceci implique, en raison de la modélisation par des processus gamma non homogènes, que les instants associés à un incrément de dégradation doivent être connus et ne peuvent être remplacés par leur incrément de temps associé (contrairement à une modélisation par un processus gamma homogène). Enfin, il est possible d'observer à partir des Tableaux 4.13 et 4.15 une diminution systématique de la largeur des IC avec l'augmentation de la longueur de la trajectoire étudiée.

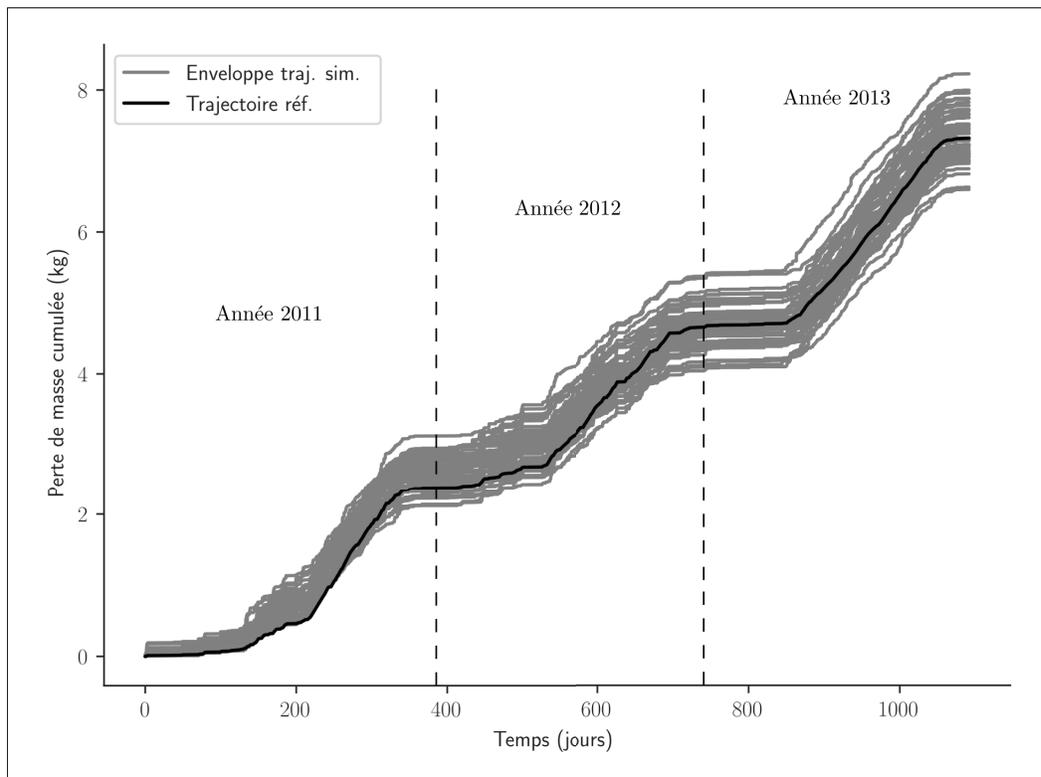


Figure 4.18 Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir des années 2011 et 2012.

On constate sur les Figures 4.18 et 4.19 que l'enveloppe des trajectoires simulées (en gris) englobe l'ensemble de la trajectoire de référence (en noir) contrairement à la modélisation s'appuyant uniquement sur les données de 2011. On observe donc une amélioration de la modélisation de la dégradation lorsque les données des années 2011 et 2012 sont prises en

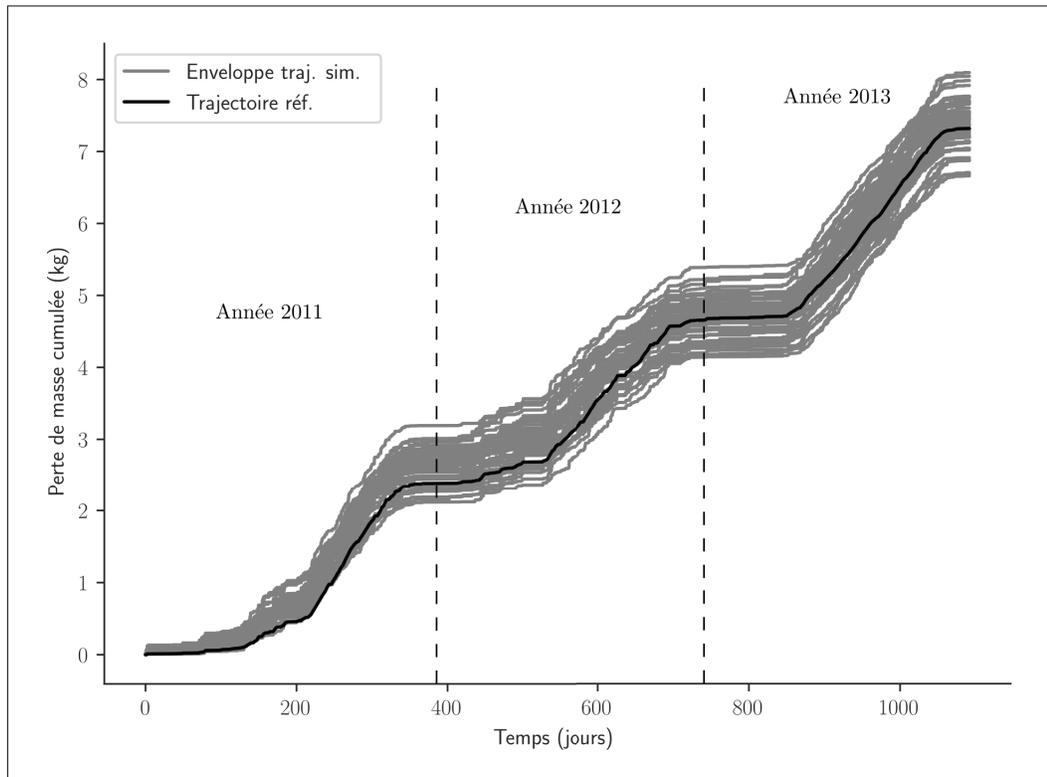


Figure 4.19 Enveloppe des trajectoires simulées avec variance modifiée - Modélisation à partir des années 2011 à 2013.

compte. On notera également que l'enveloppe des simulations est plus large que celle présente dans la Figure 4.17. Ce constat est cohérent étant donné que la variance associée aux nouveaux paramètres estimés est plus élevée. Par ailleurs, la trajectoire observée pour l'année 2011 est contenue dans la partie basse de l'enveloppe des simulations ce qui était déjà le cas dans la Figure 4.17 : cette nouvelle modélisation ne pénalise pas particulièrement l'estimation de la dégradation associée à l'année 2011 (en plus d'avoir un effet plutôt conservatif). Si on s'attarde à présent sur la qualité de prédiction du niveau de dégradation pour l'année 2013, les deux modélisations semblent présenter des résultats similaires ; que ce soit en termes de valeurs des paramètres, ou en observant les faisceaux des trajectoires simulées. En effet, logiquement, la modélisation incluant la trajectoire de l'année 2013 est capable de correctement prédire le niveau de dégradation pour cette année. La Figure 4.19 constitue un élément de vérification du modèle. En revanche, il est possible de montrer à partir de la Figure 4.18 que la modélisation

reposant sur les données des années 2011 et 2012 permet de prédire correctement le niveau de dégradation de 2013 : la trajectoire de cette dernière année est centrée par rapport à l'enveloppe simulée (en considérant une variance modifiée). Une prédiction du niveau de dégradation pour l'année 2014 (et des suivantes) pourrait permettre de confirmer que les données de dégradation de deux années consécutives (ou non consécutives) sont suffisantes pour estimer correctement les paramètres du modèle. Ce même travail devrait être réalisé sur les données issues d'autres groupes T/A.

4.5 Conclusions

Ce chapitre propose l'étude des données de cavitation érosive obtenues par mesure indirecte en centrale. Celle-ci s'est appuyée sur des données de monitoring de 3 années successives. Il a été montré qu'une modélisation reposant sur l'exploitation des deux premières années permettait de prédire de façon satisfaisante le niveau de dégradation de l'année suivante. Cette modélisation a la particularité de proposer un processus gamma non homogène pour chaque grappe différente regroupant les différents modes opératoires observés. Ces derniers sont en effet considérés comme principaux responsables de l'apparition de conditions cavitantes dans une centrale hydroélectrique. Une prédiction du niveau de dégradation peut alors être réalisée à partir des conditions d'opération et en particulier du mode opératoire dans lequel est exploitée une turbine donnée. On notera que les outils mathématiques classiques d'estimation du temps d'atteinte ne peuvent être utilisés pour ce type de modélisation proposant un modèle par grappe. L'estimation du temps d'atteinte peut alors être réalisée à partir de simulations en supposant au préalable les conditions d'opération (par exemple combien de temps un groupe T/A est opéré dans un MO donné) à l'aide de méthodes statistiques ou de modèles de prévision plus complexes. Enfin, l'exploitation des données de dégradation d'un autre (ou plusieurs) groupe T/A permettrait d'apporter un éclairage supplémentaire à ces travaux.

CONCLUSION ET RECOMMANDATIONS

Conclusions générales

Les présents travaux ont porté sur la problématique de l'utilisation d'outils probabilistes adaptés à la prévision de la durée de vie de matériels industriels. En particulier, nous nous sommes intéressés à modéliser une dégradation importante rencontrée lors du fonctionnement de turbines hydroélectriques : la cavitation érosive. Pour cela, les processus stochastiques ont été privilégiés, en particulier pour leur capacité à modéliser l'ensemble des trajectoires de dégradation ainsi qu'à décrire l'évolution temporelle de celle-ci. La motivation de ce projet étant l'estimation de la durée de vie résiduelle de roues de turbines hydroélectriques, nous nous sommes focalisés principalement sur le processus gamma, particulièrement adapté pour décrire la dynamique d'une dégradation monotone croissante. En effet, l'érosion par cavitation présente la particularité d'être cumulative. Bien qu'appliqués à une dégradation spécifique, la méthodologie ainsi que les outils employés conservent un aspect générique et peuvent donc être utilisés pour d'autres types de dégradation croissante et monotone.

La simulation de trajectoires issues d'un processus gamma a permis de mettre en avant les qualités et limites de la méthode d'estimation par maximum de vraisemblance à l'aide d'une étude de ses performances. Le plan d'expérience adopté, bien que restreint, montre qu'il est préférable de privilégier, pour un même nombre total d'observations de niveau de dégradation, plusieurs trajectoires observées. Ce résultat permet à l'analyste d'orienter une future campagne d'expérimentation. De plus, il a été montré que la méthode d'estimation étudiée pouvait présenter des difficultés pour certaines valeurs de paramètres d'un processus gamma.

Nous avons appliqué notre modèle de dégradation à deux sources de données différentes : une première issue d'essais de cavitation érosive en laboratoire et une deuxième à partir d'un dispositif de suivi de l'intensité de cavitation installé sur un groupe T/A en centrale. La première

étude a tout d'abord montré qu'il était possible de comparer la tenue de deux matériaux couramment utilisés dans la fabrication des roues de turbines en termes de résistance à la cavitation (Chatenet, Gagnon, Remy, Fouladirad & Tahan, 2020). Cette comparaison s'appuie en particulier sur le temps d'atteinte à un seuil donné de la dégradation. Dans un deuxième temps, l'influence des paramètres estimés (donc l'incertitude inhérente à ces estimations) sur la distribution du temps d'atteinte a été étudiée. Une application de la méthode delta avec celle de maximum de vraisemblance a permis de propager les incertitudes au temps d'atteinte de la dégradation.

Enfin, une modélisation des données de cavitation d'un groupe T/A installé en centrale a été réalisée. Une estimation des paramètres à partir de l'année a a été utilisée pour modéliser les années $a + 1$ et $a + 2$. En particulier, la prise en compte de l'influence des conditions d'opération sur la dégradation s'est appuyée sur les résultats d'un regroupement par grappes réalisé en amont du présent projet. Cette dernière partie présente un cas d'application dans le contexte industriel de l'étude. À partir des conditions d'exploitation des matériels industriels, il est possible d'estimer la perte de masse cumulée sur un intervalle de temps écoulé ou bien de prédire une perte de masse à une date future. À partir d'une analyse statistique ou à l'aide de modèles de planification plus complexes, il est possible de réaliser une estimation du temps qui sera passé dans chaque mode opératoire. Ces estimations de l'état de dégradation du matériel permettront à l'exploitant d'ajuster ses opérations de maintenance afin de maximiser la disponibilité de ses équipements mais aussi de réduire les pertes d'exploitation et les coûts liés aux réparations.

Recommandations et perspectives

Pour améliorer et compléter la méthodologie proposée, certaines perspectives pour de futurs travaux de recherche sont suggérées :

Perspectives scientifiques

- Les processus inverses gaussiens pourraient être considérés en remplacement des processus gamma. Ceux-ci montrent en effet des caractéristiques semblables aux processus gamma (croissants et monotones) et font l'objet depuis quelques années d'un intérêt croissant dans le domaine de la fiabilité (Hao, Yang & Berenguer, 2019 ; Peng, Zhu & Shen, 2019 ; Zhang & Revie, 2016).
- Bien que les experts en cavitation bénéficient d'une expérience importante sur le sujet, il nous a semblé difficile d'intégrer cette source d'information au travers d'une approche bayésienne. En effet, contrairement à d'autres matériels industriels, le nombre de défaillances observé pour ce type de matériel est faible en plus d'être très espacé dans le temps (certains groupes T/A sont en opération depuis près de 50 ans). Il serait donc difficile pour un expert de déterminer, par exemple, un intervalle de durée du vie d'un matériel donné (sachant que cette durée de vie est de plus, fortement affectée par les conditions d'exploitations), qui pourrait par la suite servir d'*a priori* dans un modèle bayésien. Néanmoins, cette approche a fait ses preuves dans de nombreuses applications dont les données disponibles sont limitées et pour la modélisation des incertitudes du modèle. Une nouvelle exploration de cette méthode pourrait être envisagée dans le futur, en particulier pour les données de laboratoire.
- Dans la suite "idéologique" du regroupement par grappes (*clustering*) des modes opératoires, les futurs développements pourraient s'intéresser à intégrer les conditions d'opération (d'intérêt) sous forme de covariables dans les modèles de dégradation. Ceci présenterait notamment l'avantage de pouvoir se passer du regroupement par grappe, qui présente des

limites lorsqu'un cas n'a pas été classé par le passé par exemple. De nombreux travaux sont disponibles dans la littérature et pourraient servir comme point de départ.

- L'étude sur les performances de la méthode de maximum de vraisemblance devrait adopter un plan d'expérience plus conséquent (présentant plus de cas) pour permettre de tirer des conclusions plus génériques. Celui présenté dans cette thèse permet néanmoins de faire ressortir un certain nombre de tendances.
- Les méthodes de ré-échantillonnages présentées se limitent à trois techniques. Néanmoins, de nombreuses alternatives plus ou moins populaires existent (telle que le *Markov Bootstrap*) et pourraient être explorées, notamment pour les applications utilisant les données de laboratoire.

Perspectives industrielles

- La classification des différents modes d'opération présentés dans le Chapitre 4 est le résultat d'une étude statistique récente menée il y a quelques années par Hydro-Québec (Bodson-Clermont, 2016). Ce dernier ne prend pas en compte comme critère l'intensité de cavitation fournie par le système de suivi en temps réel *Caviciel*. Pourtant, si on considère que l'intensité de cavitation est liée aux combinaisons de groupe T/A fonctionnant en simultané, la classification des modes opératoires devrait être impactée par la prise en compte de ce critère. Une nouvelle classification des modes opératoires dont les critères pris en compte se limitent aux conditions d'opération pertinentes, ainsi qu'à l'intensité de cavitation, devrait être réalisée afin d'affiner le modèle de durée de vie.
- Après une éventuelle période de transmission des connaissances et des outils informatiques, le matériel scientifique développé dans le présent projet pourrait être intégré à un outil d'aide à la prise de décision. Cette contribution s'intégrerait alors dans un macro-modèle des modes de défaillances observés pour les groupes T/A.

- Une étude basée sur la théorie de la valeur de l'information permettrait d'évaluer la meilleure stratégie à adopter quant à l'instrumentation des matériels industriels (*cf.* Straub (2014); Straub, Chatzi, Bismut, Courage, Döhler, Faber, Köhler, Lombaert, Omenzetter, Pozzi et al. (2017)). Parmi les scénarios possibles, il serait envisageable d'installer le dispositif de mesure de cavitation à des fins de calibration du modèle uniquement, et non de façon permanente comme c'est le cas actuellement.
- La dégradation relevée par le système de monitoring en continu installé en centrale repose sur une fonction de transfert traduisant un signal vibratoire en perte de masse. Ainsi, la mesure de la dégradation étant indirecte, celle-ci peut ne pas correctement refléter le niveau de dégradation réel. Afin de produire un modèle de dégradation robuste, deux voies d'amélioration se dessinent : soit porter un effort particulier pour rendre la fonction de transfert plus fidèle (meilleure image de l'intensité de cavitation relevée par rapport à la réalité), soit ajuster le modèle de dégradation dès qu'une inspection permet d'évaluer le niveau réel de la perte de masse. On notera d'ailleurs que ces deux améliorations peuvent être conduites en parallèle et ne sont pas exclusives.
- En raison de la durée d'échantillonnage associée à la mesure de l'intensité de cavitation, un moyennage (lissage) de la valeur réelle peut apparaître. Ceci a pour conséquence, entre autres, de capturer de façon imparfaite la variabilité propre au phénomène de cavitation. Une expérimentation portant sur divers temps d'échantillonnage ou une nouvelle analyse du signal vibratoire permettrait de valider ou non la variabilité ajoutée numériquement au modèle dans le Chapitre 4. Jusqu'alors, cette modification de la variance repose uniquement sur un avis d'expert.
- Comme évoqué dans le Chapitre 4, on considère qu'une roue de turbine est défaillante (pour la dégradation par cavitation) dès lors que la perte de masse cumulée dépasse un certain seuil. Néanmoins, le dépassement de ce seuil n'entraîne pas forcément l'arrêt du groupe T/A, ce dernier pouvant continuer à fonctionner. La détermination de ce seuil est à considérer

attentivement pour la suite des travaux. En effet, une étude technico-économique du coût de maintenance en fonction du seuil de dégradation fixé devrait être réalisée avant l'intégration du modèle à un outil d'aide à la décision.

Publications

Les travaux présentés dans cette thèse ont fait l'objet des publications suivantes :

Congrès internationaux

- Chatenet, Q., Gagnon, M., Tôm-Thât, L., Fouladirad, M., Remy, E., & Tahan, A. S. (2019, March). Stochastic modelling of cavitation erosion in Francis runner. *In IOP Conference Series : Earth and Environmental Science* (Vol. 240, No. 6, p. 062017).
- Chatenet, Q., Fouladirad, M., Remy, E., Gagnon, M., & Tahan, A. S. (2019, January). On the performance of the Maximum Likelihood Estimation method for Gamma Process. *Conference : Proceedings of the 29th European Safety and Reliability Conference (ESREL)*

Revue scientifique

- Chatenet, Q., Gagnon, M., Tôm-Thât, L., Remy, E., Fouladirad, M., & Tahan, A. (2020). Stochastic modeling of cavitation erosion in francis runner. *International Journal of Fluid Machinery and Systems*, 13(1), 21-29.
- Chatenet, Q., Remy, E., Gagnon, M., Fouladirad, M., & Tahan, A. Modeling cavitation erosion using non-homogeneous gamma process, *Reliability Engineering & System Safety* (en deuxième relecture).

RÉFÉRENCES

- Abdel-Hameed, M. (1975). A gamma wear process. *IEEE transactions on Reliability*, 24(2), 152–153.
- ACH. (2016). *L'hydroélectricité en 5 points*. Canadian-Hydropower-Association.
- ACH, P. E. a. A. (2015). *L'hydroélectricité dans l'économie canadienne : investissements et emplois dans la principale source d'électricité au Canada*. Canada. Repéré à <https://canadahydro.ca/wp-content/uploads/2015/09/ACH-L-hydro%C3%A9lectricit%C3%A9-dans-l-%C3%A9conomie-canadienne-May-6-2015.pdf>.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716–723.
- Al Masry, Z., Mercier, S. & Verdier, G. (2017). A condition-based dynamic maintenance policy for an extended gamma process. *10th International Conference on Mathematical Methods in Reliability*.
- Alaswad, S. & Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157, 54–63.
- Applebaum, D. (2009). *Lévy processes and stochastic calculus*. Cambridge university press.
- Arfken, G. (1985). The incomplete gamma function and related functions. *Mathematical Methods for Physicists*, 565–572.
- ASTM. (2016). *Standard Test Method for Cavitation Erosion Using Vibratory Apparatus*. West Conshohocken, PA. Repéré à www.astm.org.
- Avellan, F. (2004). Introduction to cavitation in hydraulic machinery. *The 6th International Conference on Hydraulic Machinery and Hydrodynamics*.
- Avramidis, A., L'ecuyer, P. & Tremblay, P.-A. (2003, Dec). Efficient simulation of gamma and variance-gamma processes. *Proceedings of the 2003 Winter Simulation Conference*, 1, 319-326.
- Barlow, R. E. & Proschan, F. (1996). *Mathematical theory of reliability*. Siam.
- Benjamin, T. B. & Ellis, A. T. (1966). The collapse of cavitation bubbles and the pressures thereby produced against solid boundaries. *Philosophical Transactions for the Royal Society of London. Series A, Mathematical and Physical Sciences*, 221–240.

- Berkowitz, J. & Kilian, L. (2000). Recent developments in bootstrapping time series. *Econometric Reviews*, 19(1), 1–48.
- Biluš, I., Hočevár, M., Dular, M. & Lešnik, L. (2020). Numerical Prediction of Various Cavitation Erosion Mechanisms. *Journal of Fluids Engineering*, 142(4).
- Birolini, A. (2007). *Reliability engineering*. Springer.
- Bodson-Clermont, P.-M. (2016). Modélisation statistique de l'érosion de cavitation d'une turbine hydraulique selon les paramètres d'opération.
- Bose, A. & Politis, D. N. (1993). *A review of the bootstrap for dependent samples*. Purdue University. Department of Statistics.
- Bourdon, P. (2000). *Détection vibratoire de l'érosion de cavitation des turbines Francis*. (Thèse de doctorat, EPFL).
- Bousquet, N., Fouladirad, M., Grall, A. & Paroissin, C. (2015). Bayesian gamma processes for optimizing condition-based maintenance under uncertainty. *Applied Stochastic Models in Business and Industry*, 31(3), 360–379.
- Brennen, C. E. (2014). *Cavitation and bubble dynamics*. Cambridge University Press.
- Bühlmann, P. (2002). Bootstraps for time series. *Statistical science*, 52–72.
- Burgess, M., Apostolakis, I. & Konofagou, E. (2018). Power cavitation-guided blood-brain barrier opening with focused ultrasound and microbubbles. *Physics in Medicine & Biology*, 63(6), 065009.
- Caballé, N. C., Castro, I., Pérez, C. & Lanza-Gutiérrez, J. M. (2015). A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. *Reliability Engineering & System Safety*, 134, 98–109.
- Carey, V. J. (2005). Resampling Methods for Dependent Data. SN Lahiri. *Journal of the American Statistical Association*, 100, 712–713.
- Casella, G. & Berger, R. L. (2002). *Statistical inference*. Duxbury Pacific Grove, CA.
- Celebioglu, K., Altintas, B., Aradag, S. & Tascioglu, Y. (2017). Numerical research of cavitation on Francis turbine runners. *International Journal of Hydrogen Energy*, 42(28), 17771–17781.

- Chahine, G. L., Kapahi, A., Choi, J.-K. & Hsiao, C.-T. (2016). Modeling of surface cleaning by cavitation bubble dynamics and collapse. *Ultrasonics sonochemistry*, 29, 528–549.
- Chatenet, Q., Gagnon, M., Remy, E., Fouladirad, M. & Tahan, A. (2020). Stochastic modeling of cavitation erosion in francis runner. *International Journal of Fluid Machinery and Systems*, 13(1), 21–29.
- Chaudhry, M. A. & Zubair, S. M. (1994). Generalized incomplete gamma functions with applications. *Journal of Computational and Applied Mathematics*, 55(1), 99–123.
- Chhikara, R. (1988). *The Inverse Gaussian Distribution : Theory : Methodology, and Applications*. CRC Press.
- Çınlar, E. (1980). On a generalization of gamma processes. *Journal of Applied Probability*, 17(2), 467–480.
- Cinlar, E. (2013). *Introduction to stochastic processes*. Courier Corporation.
- Cinlar, E., Bazant, Z. P. & Osman, E. (1977). Stochastic process for extrapolating concrete creep. *ASCE J Eng Mech Div*, 103(6), 1069–1088.
- Coble, J. & Hines, J. W. (2011). Applying the general path model to estimation of remaining useful life. *International Journal of Prognostics and Health Management*, 2(1), 71–82.
- Croteau, P., Cleroux, R. & Leger, C. (2005). Bootstrap confidence intervals for periodic preventive replacement policies. Dans *Statistical Modeling and Analysis for Complex Data Problems* (pp. 141–159). Springer.
- Crum, L. A. (1995). Comments on the evolving field of sonochemistry by a cavitation physicist. *Ultrasonics Sonochemistry*, 2(2), S147–S152.
- d'Agostino, L. & Salvetti, M. V. (2008). *Fluid dynamics of cavitation and cavitating turbopumps*. Springer Science & Business Media.
- Darling, D. A. (1957). The kolmogorov-smirnov, cramer-von mises tests. *The Annals of Mathematical Statistics*, 28(4), 823–838.
- Deloux, E., Fouladirad, M. & Bérenguer, C. (2016). Health-and-usage-based maintenance policies for a partially observable deteriorating system. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 230(1), 120–129.
- Deng, Y., Barros, A. & Grall, A. (2015). Degradation modeling based on a time-dependent ornstein-uhlenbeck process and residual useful lifetime estimation. *IEEE Transactions on*

Reliability, 65(1), 126–140.

Duan, F. & Wang, G. (2018). Exponential-dispersion degradation process models with random effects and covariates. *IEEE Transactions on Reliability*, 67(3), 1128–1142.

Dufresne, F., Gerber, H. U. & Shiu, E. S. (1991). Risk theory with the gamma process. *ASTIN Bulletin : The Journal of the IAA*, 21(2), 177–192.

EDF. (2019, December, 19). L'hydraulic en chiffre [Format]. Repéré à <https://www.edf.fr/groupe-edf/espaces-dedies/l-energie-de-a-a-z/tout-sur-l-energie/produire-de-l-electricite/l-hydraulique-en-chiffres>.

Efron, B. et al. (1979). Bootstrap Methods : Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1–26.

Efron, B. (1992). Bootstrap methods : another look at the jackknife. Dans *Breakthroughs in statistics* (pp. 569–593). Springer.

Efron, B. & Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator : Observed versus expected Fisher information. *Biometrika*, 65(3), 457–483.

Efron, B. & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Ellingwood, B. R. & Mori, Y. (1993). Probabilistic methods for condition assessment and life prediction of concrete structures in nuclear power plants. *Nuclear engineering and design*, 142(2-3), 155–166.

Engmann, S. & Cousineau, D. (2011). Comparing distributions : the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test. *Journal of applied quantitative methods*, 6(3), 1–17.

Escaler, X., Egusquiza, E., Farhat, M., Avellan, F. & Coussirat, M. (2006). Detection of cavitation in hydraulic turbines. *Mechanical systems and signal processing*, 20(4), 983–1007.

Escudero, D. S., Goudot, G., Vion, M., Tanter, M. & Pernot, M. (2018). 2D and 3D real-time passive cavitation imaging of pulsed cavitation ultrasound therapy in moving tissues. *Physics in Medicine & Biology*, 63(23), 235028.

Fletcher, R. & Powell, M. J. (1963). A rapidly convergent descent method for minimization. *The computer journal*, 6(2), 163–168.

Fouladirad, M., Paroissin, C. & Grall, A. (2018). Sensitivity of optimal replacement policies to lifetime parameter estimates. *European Journal of Operational Research*, 266(3),

963–975.

- Franc, J.-P., Riondet, M., Karimi, A. & Chahine, G. L. (2012). Material and velocity effects on cavitation erosion pitting. *Wear*, 274, 248–259.
- Franc, J.-P. D., Avellan, F., Belahadji, B., Billard, J.-Y., Briançon-Marjollet, L., Fréchou, D., Fruman, D., Karimi, A., Kueny, J.-L. & Michel, J.-M. (1995). *La cavitation : mécanismes physiques et aspects industriels*. Presses universitaires de Grenoble PUG.
- Genest, C., Rémillard, B. et al. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44(6), 1096–1127.
- Gindroz, B. (1991). *Lois de similitude dans les essais de cavitation des turbines Francis*. (Thèse de doctorat).
- Giorgio, M., Guida, M. & Pulcini, G. (2016). Modeling a degradation process with dependent increments in presence of random heterogeneity. *Proceedings 6th International Conference on Accelerated Life Testing and Degradation Models-ALT'2016*, pp. 51–60.
- Gohil, P. P. & Saini, R. (2015). Effect of temperature, suction head and flow velocity on cavitation in a Francis turbine of small hydro power plant. *Energy*, 93, 613–624.
- Guo, P., Wang, Z., Sun, L. & Luo, X. (2017). Characteristic analysis of the efficiency hill chart of Francis turbine for different water heads. *Advances in Mechanical Engineering*, 9(2), 1687814017690071.
- Hao, S., Yang, J. & Berenguer, C. (2019). Degradation analysis based on an extended inverse Gaussian process model with skew-normal random effects and measurement errors. *Reliability Engineering & System Safety*, 189, 261–270.
- Härdle, W., Horowitz, J. & Kreiss, J.-P. (2003). Bootstrap methods for time series. *International Statistical Review*, 71(2), 435–459.
- Henze, N., Meintanis, S. G. & Ebner, B. (2012). Goodness-of-fit tests for the gamma distribution based on the empirical Laplace transform. *Communications in Statistics-Theory and Methods*, 41(9), 1543–1556.
- Hoadley, B. (1971). Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case. *The Annals of mathematical statistics*, 1977–1991.
- Hongyi Li, G. & Maddala. (1996). Bootstrapping time series models. *Econometric reviews*, 15(2), 115–158.

- Huang, Z., Xu, Z., Wang, W. & Sun, Y. (2015). Remaining useful life prediction for a nonlinear heterogeneous wiener process model with an adaptive drift. *IEEE Transactions on Reliability*, 64(2), 687–700.
- Jian, W., Petkovšek, M., Houlin, L., Širok, B. & Dular, M. (2015). Combined numerical and experimental investigation of the cavitation erosion process. *Journal of fluids engineering*, 137(5).
- Jin, G. & Matthews, D. (2014). Reliability demonstration for long-life products based on degradation testing and a Wiener process model. *IEEE Transactions on Reliability*, 63(3), 781–797.
- Justel, A., Peña, D. & Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics & Probability Letters*, 35(3), 251–259.
- Kahle, W., Mercier, S. & Paroissin, C. (2016). *Degradation processes in reliability*. John Wiley & Sons.
- Kang, R., Gong, W. & Chen, Y. (2020). Model-driven degradation modeling approaches : Investigation and review. *Chinese Journal of Aeronautics*.
- Karlin, S. (2014). *A first course in stochastic processes*. Academic press.
- Kreiss, J.-P. & Lahiri, S. N. (2012). Bootstrap methods for time series. Dans *Handbook of statistics* (vol. 30, pp. 3–26). Elsevier.
- Krumenacker, L., Fortes-Patella, R. & Archer, A. (2014). Numerical estimation of cavitation intensity. *IOP Conference Series : Earth and Environmental Science*, 22(5), 052014.
- Kumar, P. & Saini, R. (2010). Study of cavitation in hydro turbines—A review. *Renewable and Sustainable Energy Reviews*, 14(1), 374–383.
- Lafleur, F. (2012). Vibratory detection system of cavitation erosion : historic and algorithm validation. *Proceedings of the Eighth international symposium on Cavitation*.
- Lawless, J. F. & Crowder, M. J. (2010). Models and estimation for systems with recurrent events and usage processes. *Lifetime data analysis*, 16(4), 547–570.
- Lawless, J. & Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime data analysis*, 10(3), 213–227.
- Le Son, K., Fouladirad, M. & Barros, A. (2016). Remaining useful lifetime estimation and noisy gamma deterioration process. *Reliability Engineering & System Safety*, 149, 76–87.

- Lebold, M. & Thurston, M. (2001). Open standards for condition-based maintenance and prognostic systems. *Maintenance and Reliability Conference (MARCON)*, 200.
- Li, J., Wang, Z., Zhang, Y., Liu, C. & Fu, H. (2018). A nonlinear Wiener process degradation model with autoregressive errors. *Reliability Engineering & System Safety*, 173, 48–57.
- Lorden, G. & Eisenberger, I. (1973). Detection of failure rate increases. *Technometrics*, 15(1), 167–175.
- McCulloch, C. E. & Neuhaus, J. M. (2005). Generalized linear mixed models. *Encyclopedia of biostatistics*, 4.
- Mecklin, C. J. & Mundfrom, D. J. (2004). An appraisal and bibliography of tests for multivariate normality. *International Statistical Review*, 72(1), 123–138.
- Meeker, W. Q. & Escobar, L. A. (2014). *Statistical methods for reliability data*. John Wiley & Sons.
- Muller, A., Suhner, M.-C. & Iung, B. (2008). Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system. *Reliability Engineering & System Safety*, 93(2), 234–253.
- Myung, I. J. (2000). The importance of complexity in model selection. *Journal of mathematical psychology*, 44(1), 190–204.
- Newby, M. & Barker, C. (2006). A bivariate process model for maintenance and inspection planning. *International Journal of Pressure Vessels and Piping*, 83(4), 270–275.
- Nguyen, K. T., Fouladirad, M. & Grall, A. (2018). Model selection for degradation modeling and prognosis with health monitoring data. *Reliability Engineering & System Safety*, 169, 105–116.
- Nicolai, R. P., Budai, G., Dekker, R. & Vreijling, M. (2004). Modeling the deterioration of the coating on steel structures : a comparison of methods. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, 5, 4177–4182.
- Oehlert, G. W. (1992). A note on the delta method. *The American Statistician*, 46(1), 27–29.
- Ohl, C.-D., Arora, M., Dijkink, R., Janve, V. & Lohse, D. (2006). Surface cleaning from laser-induced cavitation bubbles. *Applied physics letters*, 89(7), 074102.
- Pan, Z. & Balakrishnan, N. (2011). Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. *Reliability Engineering*

& *System Safety*, 96(8), 949–957.

Park, C. & Padgett, W. (2005). Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. *Lifetime Data Analysis*, 11(4), 511–527.

Paroissin, C. & Salami, A. (2014). Failure time of non homogeneous gamma process. *Communications in Statistics-Theory and Methods*, 43(15), 3148–3161.

Peng, W., Zhu, S.-P. & Shen, L. (2019). The transformed inverse Gaussian process as an age-and state-dependent degradation model. *Applied Mathematical Modelling*, 75, 837–852.

Peters, A., Lantermann, U. & el Moctar, O. (2018). Numerical prediction of cavitation erosion on a ship propeller in model-and full-scale. *Wear*, 408, 1–12.

Plastre, L., Devals, C., Gauthier, M., Nennemann, B., Trépanier, J.-Y. & Guibault, F. (2019). Cavitation erosion prediction based on a multi-scale method. *IOP Conference Series : Earth and Environmental Science*, 405(1), 012031.

Poznyak, A. S. (2009). *Advanced mathematical tools for automatic control engineers : Stochastic techniques*. Elsevier.

Rao, C. R. (1973). *Linear statistical inference and its applications*. Wiley New York.

Rosiński, J. (2014). Simulation of Lévy processes. *Encyclopedia of Statistics in Quality and Reliability*.

Ross, S. M., Kelly, J. J., Sullivan, R. J., Perry, W. J., Mercer, D., Davis, R. M., Washburn, T. D., Sager, E. V., Boyce, J. B. & Bristow, V. L. (1996). *Stochastic processes*. Wiley New York.

Scipy. (2019, December, 19). root(method='hydr') [Format]. Repéré à <https://docs.scipy.org/doc/scipy/reference/optimize.root-hybr.html>.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.

Si, X.-S., Wang, W., Hu, C.-H., Zhou, D.-H. & Pecht, M. G. (2012). Remaining useful life estimation based on a nonlinear diffusion degradation process. *IEEE Transactions on Reliability*, 61(1), 50–67.

Si, X.-S., Wang, W., Hu, C.-H., Chen, M.-Y. & Zhou, D.-H. (2013). A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. *Mechanical Systems and Signal Processing*, 35(1-2), 219–237.

- Singpurwalla, N. D. & Wilson, S. P. (1998). Failure models indexed by two scales. *Advances in Applied Probability*, 30(4), 1058–1072.
- Straub, D. (2014). Value of information analysis with structural reliability methods. *Structural Safety*, 49, 75–85.
- Straub, D., Chatzi, E., Bismut, E., Courage, W., Döhler, M., Faber, M. H., Köhler, J., Lombaert, G., Omenzetter, P., Pozzi, M. et al. (2017). Value of information : A roadmap to quantifying the benefit of structural health monitoring.
- Stute, W., Manteiga, W. G. & Quindimil, M. P. (1993). Bootstrap based goodness-of-fit-tests. *Metrika*, 40(1), 243–256.
- Sweeting, T. J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *The Annals of Statistics*, 1375–1381.
- Tomita, Y. & Shima, A. (1990). High-speed photographic observations of laser-induced cavitation bubbles in water. *Acta Acustica united with Acustica*, 71(3), 161–171.
- Van Noortwijk, J. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1), 2–21.
- Venkatasubramanian, V. (2005). Prognostic and diagnostic monitoring of complex systems for product lifecycle management : Challenges and opportunities. *Computers & chemical engineering*, 29(6), 1253–1263.
- Villaseñor, J. A. & González-Estrada, E. (2015). A variance ratio test of fit for gamma distributions. *Statistics & Probability Letters*, 96, 281–286.
- Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *The Annals of Mathematical Statistics*, 20(4), 595–601.
- Wang, X. & Xu, D. (2010). An inverse Gaussian process model for degradation data. *Technometrics*, 52(2), 188–197.
- Wang, X., Jiang, P., Guo, B. & Cheng, Z. (2014). Real-time reliability evaluation with a general Wiener process-based degradation model. *Quality and Reliability Engineering International*, 30(2), 205–220.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301), 236–244.

- Wasan, M. (1968). On an inverse Gaussian process. *Scandinavian Actuarial Journal*, 1968(1-2), 69–96.
- Xu, Z., Hong, Y. & Jin, R. (2016). Nonlinear general path models for degradation data with dynamic covariates. *Applied Stochastic Models in Business and Industry*, 32(2), 153–167.
- Yao, Z., Xiao, R., Wang, F. & Yang, W. (2015). Numerical investigation of cavitation improvement for a francis turbine. *Journal of Physics : Conference Series*, 656(1), 012075.
- Ye, Z.-S. & Chen, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics*, 56(3), 302–311.
- Ye, Z.-S. & Xie, M. (2015). Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry*, 31(1), 16–32.
- Ye, Z.-S., Chen, L.-P., Tang, L. C. & Xie, M. (2014a). Accelerated degradation test planning using the inverse Gaussian process. *IEEE Transactions on Reliability*, 63(3), 750–763.
- Ye, Z.-S., Xie, M., Tang, L.-C. & Chen, N. (2014b). Semiparametric estimation of gamma processes for deteriorating products. *Technometrics*, 56(4), 504–513.
- Yusof, N. S. M., Babgi, B., Alghamdi, Y., Aksu, M., Madhavan, J. & Ashokkumar, M. (2016). Physical and chemical effects of acoustic cavitation in selected ultrasonic cleaning applications. *Ultrasonics sonochemistry*, 29, 568–576.
- Zhang, C., Xiang, Y. & Shen, X. (2012). Some multivariate goodness-of-fit tests based on data depth. *Journal of Applied Statistics*, 39(2), 385–397.
- Zhang, M. & Revie, M. (2016). Model selection with application to gamma process and inverse Gaussian process. *Risk, Reliability and Safety*.
- Zhang, Z., Si, X., Hu, C. & Lei, Y. (2018). Degradation data analysis and remaining useful life estimation : A review on Wiener-process-based methods. *European Journal of Operational Research*, 271(3), 775–796.