



HAL
open science

Contribution à l'optimisation de la décision sous incertitudes : application à la maintenance

Bruno Castanier

► **To cite this version:**

Bruno Castanier. Contribution à l'optimisation de la décision sous incertitudes : application à la maintenance. Autre. Université de Nantes, 2012. tel-00862426

HAL Id: tel-00862426

<https://theses.hal.science/tel-00862426>

Submitted on 16 Sep 2013

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

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

UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE

**SCIENCES ET TECHNOLOGIES
DE L'INFORMATION ET DES MATÉRIAUX**

Année : 2012

Dossier présenté en vue de l'obtention du diplôme d'

Habilitation à Diriger des Recherches

Proposé par

Bruno CASTANIER

**Contribution à l'optimisation
de la décision sous incertitudes :
application à la maintenance**

Jury

Président	Philippe Castagliola, Professeur des Universités, Université de Nantes
Rapporteurs	Olivier Gaudoin, Professeur des Universités, Grenoble INP Benoit Iung, Professeur des Universités, Université de Lorraine Daniel Noyes, Professeur des Universités, ENIT Tarbes
Examineurs	Christophe Bérenguer, Professeur des Universités, Grenoble INP Fabrice Guérin, Professeur des Universités, ISTIA Angers Marvin Rausand, Professeur, NTNU (Norvège)

Habilitation préparée au sein de
l'Institut de Recherche en Communications et Cybernétique de Nantes (UMR CNRS 6597)

Table des matières

Notice	3
1 Activités d'enseignement	5
1.1 Activités à l'École des Mines de Nantes	5
1.2 Activités dans des formations hors-EMNantes	8
2 Activités de recherche	9
2.1 Rapport d'activités scientifiques	9
2.2 Participation à l'animation scientifique	15
3 Publications scientifiques	16
Synthèse scientifique	25
1 Introduction	25
2 Approches paramétriques	27
1 Introduction : Contexte et objectifs	27
2 Méthodologie	29
2.1 Choix des modèles de dégradation du système	32
2.2 Les hypothèses de maintenance et choix de la structure	33
2.3 Le critère décision : construction et évaluation du critère	33
2.4 Optimisation du critère	34
3 Des approches multi-seuils pour un système à un composant	34
3.1 Un schéma conditionnel d'inspection/remplacement	34
3.2 Prise en compte de l'efficacité et des durées de maintenance	37
3.3 Prise en compte de l'influence des conditions d'environnement sur le processus de dégradation	40
4 Systèmes à plusieurs modes de dégradation	44
4.1 Les dépendances dans le cadre multicomposants	44
4.2 Une structure multi-seuils pour un système à 2 composants	45
4.3 Politique conditionnelle pour un système à plusieurs modes de défaillance	47
5 Conclusions et perspectives	50
5.1 Contributions	50
5.2 Perspectives	51
3 Programmation Dynamique et entretien routier	53
1 Introduction	53
2 Méthodologie	54
2.1 Choix et construction du critère de décision :	54
2.2 Procédure de résolution et illustration numérique	55
2.3 Améliorations des performances algorithmiques et préconisations	55
3 Problématiques et positionnement	56
3.1 Cadre général du projet SBADFORM	56
3.2 Analyse bibliographique centrée sur les modèles de maintenance	57

4	Modélisation du processus de fissuration longitudinale	60
4.1	Processus gamma dépendant de l'état	62
4.2	Processus de détérioration en phase d'initiation	62
4.3	Processus de détérioration en phase de propagation	63
5	Modélisation de l'entretien des chaussées	64
5.1	L'inspection	64
5.2	Les actions d'entretien	65
5.3	Modélisation de l'effet d'entretien sur le processus (ρ, θ)	67
6	Modélisation des coûts de maintenance	68
7	Politiques de décaissements : problème restreint	69
7.1	Objectifs	69
7.2	Critère de coût	69
7.3	Procédure de résolution numérique	69
7.4	Application numérique	70
7.5	Discussions	71
8	Politiques de décaissement et de chargement	72
8.1	Objectifs	72
8.2	Modélisation de l'impact de l'épaisseur de la chaussée sur la fissuration	72
8.3	Prise en compte de la non-observabilité de θ	73
9	Conclusions et perspectives	74
9.1	Conclusions	74
9.2	Perspectives	75
4	Maintenance et Options réelles	77
1	Objectifs	77
2	Positionnement scientifique du problème	79
2.1	Maintenir ou investir ?	79
2.2	L'obsolescence matérielle	81
2.3	Options réelles : une approche pour l'optimisation de la maintenance ?	82
3	Définition et formulation du problème général	84
3.1	Modélisation de la performance des systèmes technologiques	84
3.2	Modélisation de l'évolution technologique	84
3.3	Description du processus de décision et définition des coûts associés	84
3.4	Choix et construction du critère de décision	85
3.5	Procédure de résolution et illustration numérique	85
4	Stratégies de maintenance sous séquence technologique	86
4.1	Objectifs	86
4.2	Description du processus de décision	86
4.3	Analyses des politiques	87
4.4	Extension pour la disponibilité des pièces de rechange	89
5	La décision de maintenance face au saut technologique	90
5.1	Objectifs	90
5.2	Construction du critère	91
5.3	Analyses de performance du modèle	92
5.4	Incertitudes sur la qualité de l'information du marché	93
6	Conclusion et perspectives	95
6.1	Conclusion	95
6.2	Perspectives	96

5	Conclusion et Perspectives Scientifiques	97
1	Conclusion	97
2	Perspectives scientifiques	97
2.1	Construction de méta-modèles pour l'optimisation de la maintenance des structures et infrastructures	98
2.2	Maintenance et performance industrielle	98
2.3	Nouveaux domaines d'application de la maintenance : les éoliennes off- shores	98
	Bibliographie	99
	Annexe : Sélection d'articles	107

Notice d'activités

Bruno CASTANIER

Né le 19 septembre 1970

École des Mines de Nantes
 Département Automatique et Productique
 La Chantrerie - 4, rue Alfred Kastler - BP 20722
 44307 NANTES CEDEX 03
 Tél : (33) 251.858.312 - Fax : (33) 251.858.349
 e-mail : Bruno.Castanier@mines-nantes.fr

SITUATION PROFESSIONNELLE

Maître Assistant au Département Automatique et Productique de l'École des Mines de Nantes depuis Septembre 2002

Équipe d'accueil : Équipe Systèmes Logistiques et de Production de l'Institut de Recherche en Communications et Cybernétique de Nantes - UMR CNRS 6597

FORMATION

1998 - 2001 <i>Université Technologie Troyes U.T.T.</i>	Thèse de Doctorat sous la direction de M. Christophe BÉ-RENGUER et M. Antoine GRALL : <i>Modélisation stochastique et optimisation de la maintenance conditionnelle des systèmes à dégradation graduelle</i> - Date de soutenance : 14 Décembre 2001
1997 - 1998 <i>Université de Montpellier II</i>	Diplôme d'Études Approfondies à l'Université des Sciences et Techniques du Languedoc Spécialité : Mathématiques Appliquées - option : Biostatistique
1996 - 1997 <i>Université de Montpellier II</i>	Maîtrise d'Ingénierie Mathématique à l'U.S.T.L Spécialité : Probabilité - Statistique
1994 - 1996 <i>Université de Montpellier II</i>	Licence de Mathématiques Fondamentales à l'U.S.T.L
1990 - 1993 <i>Université de Montpellier II</i>	Diplôme d'Études Universitaire Générales à l'U.S.T.L Spécialité : Mathématique - Sciences Physiques
1990	Baccalauréat C Mention Bien Spécialité : Mathématique - Sciences Physiques

EXPERIENCES PROFESSIONNELLES

Janv. 2002 - Sept. 2002 <i>EMNantes</i>	Maître Assistant Associé au Département Automatique Productique
Sept. 2001 - Déc. 2001 <i>U.T.T.</i>	Attaché Temporaire d'Études et de Recherche au Département Génie des Systèmes Industriels
Oct. 1998 - Sept. 2001 <i>U.T.T.</i>	Moniteur d'Initiation à l'Enseignement Supérieur

ENSEIGNEMENTS - RESPONSABILITÉS

• Enseignements :

Intitulé	Type	Niveau	Heures	Lieu	Année
Maintenance des systèmes industriels	CM	Bac +5	15 h	EMN	2002-...
	TP	Bac +5	15 h	EMN	2002- ...
Ingénierie de la fiabilité	CM	Bac +5	17 h	EMN	2003-...
Management des risques industriels	CM	Bac +5	8 h	EMN	2007-...
Notions élémentaires au SLI	CM	Bac +5	3 h	EMN	2008-...
Simulation Stochastique	TD	Bac +4	10 h	EMN	2010-...
Statistiques pour l'ingénieur	TD	Bac +4	10 h	EMN	2009-2010
Ingénierie de la Maintenance	CM	Bac +5	16 h	Polytech'Nantes	2006-2008
Maintenance	TD	Bac+2	2 × 12 h	IUT Nantes	2006-2008
Processus stochastiques et files d'attente	CM	Bac +3	12 h	U.T.T.	2001
	TD	Bac +3	12 h	U.T.T.	2001
Fonctions à plusieurs variables	TD	Bac +2	2 × 34 h	U.T.T	1998-2001

• Responsabilités :

Depuis Mai 2012	Adjoint Enseignement du <i>Département Automatique Productive</i>
Sept. 2003 à Sept. 2010	Responsable de l'option <i>Qualité et Sécurité de Fonctionnement</i>
Depuis sept. 2005	Responsable de la Formation Spécialisée double diplômante (Department Industrial and System Engineering, Virginia Tech, USA)

ENCADREMENTS - PRODUCTION SCIENTIFIQUE

• Encadrement :

Type	soutenue	en cours
Thèse de doctorat	3	3
Thèse de Master	6	0

• Publication :

Type	98 - 01	02 - 05	06	07	08	09	10	11	12	Total
Revue internationale	1	2	1		1	1	1	2	2 (+2)	11
Revue francophone	1									1
Chapitre de livre		1					1	2		4
Congrès international	3	5		1	5	2	3	3	3	25
Colloque sans acte			2	1	5		1	2	1	12
Congrès francophone	3	2	1	1			1	1	1	10
Rapport	2				2		1	1		6
Total	10	10	4	3	13	3	8	11	6	69

• Contrat :

2006	Étude : Caractérisation de données fiabilistes de bouteilles d'air comprimé - <i>Financier</i> : Ingénierie Services Développement (ISD) - <i>Montant</i> : 3 k €
2008-2012	Contrat de recherche : State-Based Decision For Road Maintenance - SBADFORM - <i>Financier</i> : Conseil régional des Pays de la Loire - <i>Montant</i> : 295 k €
2010-2013	Contrat de recherche : Pôle de compétence Evaluation des méthodes de Contrôle Non Destructif en Pays de la Loire - ECND-PdL - <i>Financier</i> : Conseil régional des Pays de la Loire - <i>Montant</i> : 495,600 k €

1 Activités d'enseignement

Mon implication au niveau des activités d'enseignement se traduit par la construction de Cours Magistraux, l'encadrement de Travaux Dirigés et de divers projets étudiants, l'animation pédagogique, la construction et la responsabilité de tout ou partie du cursus de formation. Mon activité d'enseignement a débuté en 1998 lors de mon monitorat à l'Université de Technologie de Troyes, s'est poursuivie en 2001 en poste d'Attaché Temporaire d'Etudes et de Recherche et depuis 2002 sur la région nantaise à l'École des Mines de Nantes et des participations à des formations en IUT QLIO et à Polytech'Nantes. Je propose de résumer mon activité en présentant en premier lieu celles effectuées à l'EMNantes puis mon implication dans les établissements extérieurs.

1.1 Activités à l'École des Mines de Nantes

Le profil des ingénieurs de l'École des Mines de Nantes se définit comme des *professionnels de la conduite de projets technologiques* autour de deux grands domaines : *Sciences de l'environnement et de l'énergie* et *Sciences et technologies de l'information*.

Le recrutement majoritaire de nos élèves s'effectuait jusqu'à l'année 2011 sur concours après une année de classes préparatoires en Math Sup. Le cursus se définissait en Cycle de Base (2 ans) pour l'acquisition des connaissances et compétences de base qui seront approfondies en cycle dit Graduate School (2 ans). Le cycle Graduate s'articulait autour d'options qui sont des parcours thématiques. L'École des Mines de Nantes propose 10 options. Aujourd'hui, le mode de recrutement majoritaire de nos élèves est un recrutement après l'année de Math Spé dans la banque dite *Mines-Pont*. L'organisation de la formation est réduite à un unique *cycle* dit de Formation Ingénieur.

L'organisation générale de l'enseignement à l'École des Mines de Nantes a évolué à la fois en termes de structure avec le passage au standard européen L.M.D., d'où la décomposition en cycles, et un schéma pédagogique basé sur l'acquisition de compétences reposant sur la notion d'Unité de Valeurs (UV). Une option se définit comme un parcours plus ou moins fixé préalablement associant UV Scientifiques et Techniques, de Sciences Sociales et de Gestion, de Méthodes Génériques pour l'Ingénieur ainsi que les UV de formation à l'entreprise, le choix d'UV est généralement très restreint en opposition aux formations construites sur la base de catalogues d'UV.

Le projet *Compétences* de l'École des Mines de Nantes repose principalement sur un référentiel présentant les compétences transversales à l'ensemble des options et les compétences propres à chacune des options. L'objectif est ensuite de mesurer le niveau et la progression de l'élève au travers d'exercices adaptés et d'en faire une restitution individuelle en fin de formation.

J'ai pris part à la construction et la mise en place progressive de 2004 à 2008 de ces nouvelles orientations tout en cherchant à développer mes activités d'enseignement. Aujourd'hui, au vu de mes nouvelles responsabilités et du nouveau processus de formation en 3 années, je participe activement à la construction et à la mise en place du nouveau programme pédagogique essentiellement pour les activités du département Automatique Productive (3 options de fin d'étude et une forte participation dans les enseignements de Tronc Commun).

Les enseignements

À mon arrivée à l'EMNantes, j'ai intégré le Département Automatique Productive, département d'enseignement et de recherche, en charge ou participant en Cycle de Base aux formations en Mathématiques, Électronique et assurant la responsabilité de trois options : *Automatique, Informatique Industrielle, Gestion des Opérations en Production et Logistiques et Qualité et Sécurité de Fonctionnement*. Mon investissement s'est concentré dans cette dernière option dont j'ai eu très rapidement la responsabilité. La responsabilité d'option associée

à la responsabilité d'Unité de Valeur consiste à définir et mettre en place un programme pédagogique en relation avec les thèmes généraux de l'option. Ces thèmes peuvent évoluer en fonction des besoins industriels, actuels et futurs. Une évaluation de ces besoins est réalisée approximativement tous les 4 ans dans le contexte d'un processus d'actualisation d'option conduit par le responsable de formation par le biais d'analyses du devenir des anciens optionnaires, de l'expertise d'industriels et d'universitaires. J'ai mené ou participé à 2 actualisations de l'option Qualité et Sécurité de Fonctionnement.

L'évaluation de la charge d'enseignement à l'École des Mines de Nantes diffère du système universitaire. Toute activité relative à l'enseignement (cours, responsabilités, encadrements, recrutements, ...) est comptabilisée sous forme d'*Unité Pédagogique*. Une charge globale est attribuée au niveau du département qui se répartit ensuite, suivant les intérêts et motivations de chacun, au niveau des Enseignants Chercheurs. Ma charge moyenne par an est de plus de 310 UP (la moyenne cible par EC est de 180 UP) avec des années à plus de 400 UP.

Je décris par la suite les cours que j'estime majeurs en termes de responsabilité, de volume horaire ou d'innovation pédagogique.

Maintenance des Systèmes Industriels : Ce cours de 30 heures est dispensé aux élèves en dernière année de l'option *Qualité et Sécurité de Fonctionnement* (QSF). Son objectif général est de présenter les approches classiques pour l'organisation de la maintenance. Il sert aussi de cours d'intégration des différentes méthodes d'analyse des systèmes abordées dans d'autres modules combinées à des approches d'évaluation quantitative, partant de l'estimation des paramètres reposant sur un retour d'expérience jusqu'à l'optimisation en passant par des concepts de probabilité. Un sujet de Travaux Pratiques permet à la fois la mise en application des concepts abordés et une introduction à des problématiques de recherche.

Fiabilité des Systèmes : Ce cours de 17 heures fait suite au cours *Théorie de la Fiabilité*. Il associe méthodes quantitatives et qualitatives rencontrées par l'ingénieur Fiabiliste sur tout le cycle de vie produit en présentant les domaines d'actualité comme la sécurité des systèmes instrumentés de sécurité. Son format pédagogique est classique et associe cours magistral, exercices en séance et exercices d'évaluation.

Gestion des Risques Industriels : Cette intervention de 8 heures se situe dès le début du semestre. Elle sert d'introduction aux définitions, méthodes et objectifs liés à la Gestion des Risques qui seront par la suite approfondis dans les divers cours. Il s'organise sous forme d'un cours magistral, sans évaluation.

Management de la Qualité : L'objectif de cette auto-formation est la connaissance de la norme ISO 9001 v 2000. Elle s'organise sous forme d'un jeu informatique *Horizon Qualité en Ligne* (HQeL) durant lequel l'élève doit mettre en place l'ensemble des outils de la norme dans son entreprise virtuelle, l'efficacité de ses outils est directement proportionnelle à la qualité des réponses lors de Q.C.M. Cet outil a été développé sous la direction de l'École des Mines d'Alès et en partenariat avec les Écoles des Mines de Douai et Nantes au sein d'un projet UNIT <http://www.unit.eu/ori-oai-search/notice/view/unit-ori-wf-1-1021>.

Les projets

L'encadrement de projet élève est une part importante de l'activité d'un enseignant chercheur à l'EMNantes. Pour ma part, elle se décline en 3 volets :

Les projets d'Orientations Scientifique et technique ont pour objectif d'amener les élèves de première année de Graduate School (GS1) à *apprendre à résoudre un problème d'ingénieur* ou à *réaliser une initiation à la recherche*. Ces projets, en groupe de 3 ou 4, se

déroulent sur une bonne partie de l'année universitaire. Chaque année depuis 2007, je propose 1 à 2 sujets autour des méthodes d'évaluation de la fiabilité de systèmes complexes.

Le Projet en Contexte International est un projet se déroulant sur le semestre de formation en GS2 en partenariat avec une université étrangère. Son objectif est double avec l'apprentissage de la gestion de projet à distance et la mise en application des concepts et méthodes du domaine QSF. Notre partenaire est le département Industrial and Systems Engineering de Virginia Tech. Les équipes projet formées de 8 étudiants français et américains se doivent de répondre à une problématique posée par un industriel. En début de projet, les étudiants se retrouvent physiquement pendant une semaine dans le pays porteur des projets pour définir le cahier des charges pour ensuite définir et mettre en place la solution. Mon implication se situe sur l'ensemble de l'exercice avec sa responsabilité (définition du partenariat, du cahier des charges, organisation de la semaine de *recouvrement* et des soutenances à distance), la recherche de *sponsors* industriels et le suivi d'au moins deux projets par an.

Les Projets de Fin d'Études sont d'une durée de 6 mois en entreprise. Mon implication dans le processus de cet exercice est le soutien aux élèves pour le choix de leur sujet, la validation de leur mission (ces deux points restent essentiels pour la projection de l'élève dans son devenir professionnel), le suivi de 3 à 4 élèves en moyenne par année et l'organisation des soutenances pour l'option QSF.

Les responsabilités et animation

Outre ma participation continue au *Comité Pédagogique*, instance de réflexion pédagogique, j'ai été représentant élu (2009-2011) des enseignants chercheurs au *Comité de l'Enseignement*, instance décisionnelle. Par ailleurs, j'ai et je continue d'occuper un ensemble de responsabilités au niveau pédagogique, responsabilités que je vais détailler dans les paragraphes suivants (je ne décrirai pas ici ma dernière responsabilité d'adjoint enseignement au DAP).

Responsable d'option (RO) : Après une année d'*adjoint au responsable d'option QSF*, j'ai pris l'entière responsabilité de ce programme de fin d'études. Initialement concentrée sur la totalité de la dernière année, la GS2, (un semestre de formation et projets de fin d'étude), le périmètre de la responsabilité a progressivement évolué avec une plus grande implication sur la GS1 et un rôle accru en tant que relais avec les anciennes promotions. L'activité du RO se décline depuis le recrutement des élèves au sein de son option en deuxième année du Cycle de Base jusqu'à la remise du diplôme en passant bien entendu par la définition et la mise en place du programme pédagogique, assurer les recrutements et la gestion (hors administrative) des intervenants extérieurs. Il est le relais privilégié des élèves, de l'adjoint-enseignement au niveau des départements et de la Direction des Études. Mon souci au cours de mes années de responsabilité a été de proposer constamment des améliorations de programme afin d'assurer l'adéquation entre formation et besoins actuels et futurs des entreprises, de mettre en place les différentes réformes de notre système de formation, notamment la déclinaison du programme sous forme d'UV et le projet *Compétences*, d'assurer un suivi particulier de chacun des élèves (une dizaine en moyenne par an) pour les aider dans leur réflexion de projection professionnelle (choix et validation des parcours à l'étranger, projection personnelle, etc.), d'assurer le recrutement des élèves dans l'option (recrutement très aléatoire pour l'option pouvant aller de 4 élèves en 2002 à 21 élèves pour 2011). Durant mes 8 années de responsabilité, j'ai pu participer à deux recrutements de Maître Assistant (2006 et 2010) pour former l'équipe pédagogique autour de QSF.

Responsable d'UV (RUV) : Comme je l'ai souligné, l'Unité de Valeur est un ensemble de cours et d'exercices pédagogiques d'un volume de 90 heures à l'emploi du temps. Le rôle du RUV est d'en définir les objectifs et les enseignements qu'il synthétisera dans une fiche mise à disposition des élèves dont il assurera le suivi des résultats au sein de son UV. J'ai actuellement la responsabilité des 4 UV Scientifiques et Techniques de GS2 de l'option QSF :

- *UV - Stratégies et Organisation de la Qualité :* Elle a pour but d'introduire les principales définitions et outils ayant trait à la qualité et à la maîtrise des risques et d'en souligner les organisations et les méthodes existantes qui en découlent. À l'issue de cette UV, l'élève sera à même de définir les objectifs stratégiques de l'entreprise déclinés tant au niveau de son système de management qu'au niveau de la conception de systèmes technologiques, de mobiliser les hommes et ressources (ressources matérielles et méthodologiques) nécessaires pour atteindre les différentes exigences découlant de ses objectifs stratégiques. Des thèmes comme le management de la qualité, management des risques industriels et professionnels, ingénierie système et sûreté de fonctionnement y sont abordés.
- *UV - Méthodes d'Analyse en Qualité :* L'objectif principal de l'UV MAQ peut s'exprimer comme suit : *Permettre à l'élève QSF de conduire une analyse complète sur un système quelconque pour justifier de ses performances en fonction des exigences du client, et, le cas échéant, de définir les actions correctives pertinentes.* Cette UV combine les approches qualitatives d'analyse et d'amélioration des systèmes classiques en qualité et en sûreté de fonctionnement ainsi que les méthodes d'évaluation quantitative des risques.
- *UV - Maintien en Condition Opérationnelle :* L'objectif global peut s'exprimer comme suit : *Concevoir et gérer un système de mesures pour rationaliser la prise de décision en phase d'exploitation.* Plus spécifiquement, les interventions principales sont la *Maîtrise Statistique des Procédés et Plans d'Expérience*, la *méthode Six Sigma*, le *Soutien Logistique Intégré* et la *Maintenance des Systèmes Industriels*.
- *UV - Projet en Contexte International :* Outre l'aspect international du projet, l'objectif de cette UV est la conduite d'un projet dans sa globalité depuis l'analyse du besoin client jusqu'à la mise en place de la solution sur des thèmes de l'option. Cet exercice est support à la mise en pratique des enseignements théoriques dispensés tout au long de la formation sur un sujet porté par un industriel, sujet souvent propice à la réflexion sur les méthodologies et problématiques futures.

Responsable de la Formation Spécialisée : Le programme en Français de formation spécialisée *Gestion des Opérations de Production et Logistiques* et en *Qualité et Sûreté de Fonctionnement* est ouvert aux étudiants de Virginia Tech intégré dans le cursus double-diplômant avec l'École des Mines de Nantes. La responsabilité concerne aussi le programme double-diplômant ouvert à nos élèves à Virginia Tech. Depuis 2005 et le départ des cadres EMNantes de la collaboration, j'assume cette responsabilité. Mes actions ont essentiellement portées sur la redynamisation du programme avec l'extension à une année complète d'échange, la définition des programmes en fonction des attentes des étudiants, de la formation suivie dans leur établissement d'origine et les spécificités de chacun des établissements et, enfin, l'accueil et le suivi des étudiants. Le stage de fin d'études peut prendre deux formes : soit un parcours classique dit *industriel*, soit une initiation à la recherche dans un laboratoire de recherche. Dans le contexte d'un stage recherche, l'encadrement s'effectue en collaboration entre les deux établissements.

1.2 Activités dans des formations hors-EMNantes

Mes activités de formation ont débuté lors de mon monitorat à l'Université de Technologie de Troyes. Pendant mes trois années de thèse, j'ai participé à l'enseignement en Mathématiques par l'encadrement de Travaux Dirigés dispensés à des groupes d'environ 20 étudiants en

deuxième année de cycle de base dans le contexte de l'UV MT22 - *Fonctions de plusieurs variables et applications*.

Mon monitorat s'est poursuivi par un poste d'Attaché Temporaire d'Études et de Recherche dans le département du Génie des Systèmes Industriels. Durant la durée de mon contrat (3 mois), j'ai construit et mis en place le cours magistral (12 heures) *Processus stochastiques et Files d'attente*, les travaux dirigés associés (12 heures) et l'évaluation finale. Ce cours est partie intégrante de l'UV SY18 *Outils de modélisation et d'évaluation des performances* offerts aux étudiants de première année de Branche (Bac +3). Environ 75 étudiants y participaient. J'ai également participé à l'encadrement de projets de fin d'étude.

Sur Nantes, j'ai effectué pendant trois années des activités contractuelles. J'ai assuré l'encadrement de travaux dirigés dans les départements QLIO (Qualité, Logistique Industrielle et Optimisation) et GSM (Génie des Systèmes Mécaniques) de l'I.U.T. de Nantes pour des élèves de première et deuxième année. Cet enseignement a pour objectif d'initier aux méthodes quantitatives classiques d'évaluation de la fiabilité et optimisation de la maintenance.

Je me suis investi pendant trois années dans la formation MORI (Management des Organisations et des Risques), option de dernière année offerte à une quinzaine d'étudiants de l'École Polytechnique Universitaire de Nantes. J'ai proposé un cours magistral et l'évaluation associée en *Ingénierie de la maintenance* (16 heures) ayant pour objectif de familiariser les étudiants avec la fonction maintenance, les organisations qui en découlent et les méthodes d'optimisation de la maintenance industrielle.

2 Activités de recherche

2.1 Rapport d'activités scientifiques

L'objectif principal de mes recherches se définit dans la construction de modèles décisionnels sous incertitude, l'incertitude pouvant être liée à une connaissance partielle des données d'entrée du problème et aux comportements aléatoires des phénomènes étudiés. La motivation de mes recherches est le développement d'outils d'aide à la décision permettant l'étude et le contrôle de systèmes industriels afin de satisfaire au mieux les exigences d'un utilisateur. Ces développements concernent *la définition de nouvelles stratégies ou politiques* pour la structuration des règles de décision, *la construction de modèles d'évaluation de performance* reposant sur des formalismes mathématiques empruntés à la théorie des probabilités ou bien des approches de simulations stochastiques, et la proposition d'algorithmes spécifiques pour l'optimisation de ces stratégies suivant la complexité des contextes à prendre en compte.

Ma stratégie de recherche a été jusqu'ici de consolider mes activités autour de mon thème principal d'application à savoir l'optimisation de la maintenance, et d'investiguer de nouveaux champs en relation avec les sujets développés au sein de mon équipe d'accueil *Systèmes Logistiques et de Production* de l'IRCCyN. C'est le plan que j'ai choisi pour présenter mes activités, depuis mon stage de D.E.A. réalisé en 1998 à l'*Institut National de Recherche en Informatique et Automatique*, INRIA Rhône-Alpes, sous la direction de Gilles Celeux jusqu'à aujourd'hui.

Optimisation de la maintenance

L'optimisation de la maintenance connaît aujourd'hui un fort regain d'intérêt tant ce domaine est clairement identifié comme un levier d'amélioration des performances industrielles. Ses impacts sur le plan économique se traduisent, par exemple, par l'amélioration de la productivité et plus particulièrement de la disponibilité des moyens de production ou encore la définition de leurs plans de renouvellement et d'investissement. Sur des aspects sécuritaires et gestion des risques, son intérêt se situe dans l'organisation des actions de prévention de défaillances critiques pour des structures ou infrastructures à forte importance en terme de risque. Par ailleurs, aux vues des objectifs divers et complexes, des contraintes multiples et des aspects stratégiques que revêt l'organisation de la maintenance, le manager se doit de

reposer sa décision sur des outils rationnels et performants d'aide à la décision. La complexité des critères de décision est renforcée par la diversité des actions de maintenance à prendre en compte et surtout la nature stochastique des phénomènes de dégradation rencontrés. Ainsi, le challenge de l'optimisation de la maintenance est de proposer le meilleur cadre pour la décision dans un contexte fortement complexe.

Les verrous scientifiques visés se concentrent autour de la proposition de nouveaux modèles technico-économiques pour l'optimisation de la décision de maintenance qui prennent en compte de manière simultanée les objectifs économique-stratégiques du responsable maintenance et de l'efficacité dans le contexte opérationnel de l'activité de maintenance. Ceci se traduit par la proposition de politiques qui reposent sur des indicateurs reflétant au mieux la performance du système à maintenir ainsi que l'évolution des mécanismes de dégradation associés, l'intégration d'hypothèses plus réalistes tant au niveau de l'action de maintenance à proprement parler qu'au niveau du processus plus global de décision considérant par exemple l'accessibilité ou la complexité du système à maintenir, ou encore la disponibilité des pièces de rechange ou l'obsolescence du système.

Les travaux présentés dans cette section sont le fruit de collaborations avec des membres des établissements suivants : l'IUT de Nantes, l'Université de Technologie de Troyes, la Norwegian University of Science and Technology, le Laboratoire Central des Ponts et Chaussées et la Shanghai Jiao Tong University.

- *Pour une meilleure prise en compte des mécanismes de défaillance.*

L'objectif des travaux présentés dans cette section est l'amélioration des performances *locales* des politiques de maintenance, d'une part en reposant la décision sur des indicateurs d'état du système plus représentatifs que les indicateurs traditionnels tels que l'âge des systèmes, par exemple, et, d'autre part, en explicitant le comportement et l'évolution de ces indicateurs dans divers contextes d'exploitation. Les premiers travaux développés dans ce contexte ont été effectués et présentés dans la deuxième partie de ma thèse [6] encadrée par C. Bérenguer et A. Grall au sein l'axe *Fiabilité et Maintenance* du Laboratoire de Modélisation et Sécurité des Systèmes (LM2S) de l'Université de Technologie de Troyes. Ils consistent à définir et évaluer des politiques de maintenance pour des systèmes dits à dégradation graduelle pour lesquels le niveau de dégradation n'est pas directement observable. Les stratégies d'inspection/remplacement que nous avons proposées [50, 26, 11] sont des politiques paramétriques à seuils dictant l'action à réaliser en fonction du niveau de dégradation observé (Ne rien faire ou Remplacer) et de la date de prochaine inspection. Les avantages de cette proposition sont une amélioration des performances de la politique au regard des politiques classiques basées sur le temps ou l'âge du système et une simplicité de mise en œuvre dans un contexte opérationnel. Le modèle de dégradation proposé est un modèle de distribution de *type phase discrète* défini par une loi exponentielle de paramètre donné sur une période de temps discrète. Ce modèle a l'avantage de bien représenter les cinétiques de dégradation de systèmes soumis à une usure au cours du temps (corrosion, fissuration, vieillissement, fatigue, etc.) pour lesquels peu d'observations sont disponibles, d'offrir des propriétés mathématiques permettant de conduire un formalisme avancé et une résolution du critère de décision par des approches numériques classiques.

Dans un second temps, nous avons étendu ces politiques dans plusieurs directions [3]. Nous avons tout d'abord proposé un modèle de maintenance lorsque la variable de décision devient multidimensionnelle. Elle peut ainsi soit représenter l'état de dégradation de chacun des composants formant le système soit différents modes de vieillissement du système qui ne peuvent pas être résumés par une unique variable. La règle de décision de type limite de contrôle est étendue pour offrir l'opportunité d'une réparation partielle du système par le remplacement d'un seul des composants. La maintenance combinée sur deux composants entraîne des dépendances économiques sur le système, rendant le processus d'évolution d'un composant non-markovien et augmentant de facto la difficulté de résolution du critère de

décision. Nous avons proposé de reposer cette évaluation sur des approches classiques de simulations de Monte Carlo [24, 49] et l'application de résultats sur les processus stochastiques de renouvellement markovien [27, 13]. L'une des limites que nous pouvons souligner dans ce contexte multi-dimensionnel est que chaque indicateur montre des comportements analogues (de type dégradation graduelle) et indépendants l'un vis-à-vis de l'autre. Dans le cadre de la thèse d'Estelle Deloux [7], nous avons étudié des systèmes soumis à des modes de dégradation différents et dépendants. Le modèle de dégradation [32, 34] est un processus bivarié représentant la dégradation graduelle, modélisée par un processus gamma stationnaire, et une défaillance de type choc dont la fonction intensité est dépendante du niveau de dégradation courant et d'un niveau de stress lié à un signal vibratoire. Chacun des modes de défaillance est défini par son propre indicateur : le niveau de dégradation observable par le biais d'inspections périodiques et l'intensité du processus vibratoire cause de la défaillance de type choc. Une défaillance entraîne une perturbation dans le fonctionnement du système qui se traduit ici par une augmentation progressive des vibrations. Nous avons défini une politique de maintenance reposant sur ces indicateurs combinant une politique conditionnelle classique d'inspection/remplacement et une carte de contrôle. L'avantage de notre approche se traduit par une meilleure détectabilité de la défaillance permettant ainsi de mieux répartir les efforts sur la maintenance préventive.

Par ailleurs, nous avons cherché à différencier les deux phases de dégradation dans le modèle de maintenance : la phase d'initiation durant laquelle la dégradation n'est pas observable et la phase de propagation. Cette modélisation a été appliquée pour répondre à des problématiques de fissuration longitudinale liée à la fatigue des chaussées causée par l'action répétitive des charges de trafic [33, 53, 37]. Nous avons développé des travaux à finalités similaires dans le contexte de la maintenance de tuyauteries de centrales nucléaires en partant de l'expérimentation pour la collecte de données jusqu'au développement du modèle de maintenance (stage de Master of Science de Ziqing Zhai en collaboration avec le Pr. Zhang du département nucléaire de la Shanghai Jiao Tong University).

Plus récemment, dans le cadre de la thèse de Mariem Zouch [9] intégrée dans le projet SBADFORM - State-Based Decision For Road Maintenance - financé par la Région Pays de la Loire, nous avons cherché à intégrer dans le modèle de maintenance conditionnelle une information supplémentaire pour rendre compte de l'état de fissuration sous-jacent. Cette intégration reposant sur des considérations réelles des processus d'initiation et propagation des fissures longitudinales en sous-couche d'une chaussée permet d'améliorer la prédiction de l'évolution de l'indicateur de fissuration en surface. Cette proposition s'est faite en deux phases avec la considération, en premier lieu [38, 60, 61], d'une chaîne de Markov bivariée composée du pourcentage linéaire de fissuration d'une section de route et de son taux d'accroissement espéré puis, en second lieu [55, 63, 22], d'un processus de Markov bivarié non stationnaire en l'état de dégradation. La dépendance en l'état nous a conduit à étendre des résultats théoriques classiques sur les processus de Lévy. De plus, nous avons reposé la modélisation du taux d'accroissement espéré du pourcentage de fissuration après fissuration sur le processus gamma bilatéral, généralisation du processus gamma pouvant prendre des valeurs positives ou négatives qui n'avait, à notre connaissance, jamais été utilisé dans un contexte d'optimisation de la maintenance.

Enfin, nous avons cherché à intégrer des perturbations de comportements de dégradation liées, par exemple, à une variabilité des contextes opérationnels et étudier ces impacts sur l'évolution de la décision de maintenance. Ces travaux ont fait l'objet de la thèse d'Estelle Deloux [7] dans laquelle nous avons étudié différentes formes d'impacts. Une covariable ou stress est introduite pour modéliser cet impact. Ce stress peut conduire à augmenter une probabilité de défaillance [32, 34] ou encore accélérer en moyenne les processus de dégradation. Les approches proposées reposent sur des modèles de type *Durée de Vie Accélérée*. Plusieurs cas ont été étudiés suivant que l'impact soit ponctuel ou permanent, discret [15, 35, 54, 16] ou continu [36, 18]. Bien que l'importance de prendre en compte les effets d'un environnement

opératoire sur la dégradation soit intuitivement évidente, la question du comment l'intégrer dans le modèle de décision reste délicate au vu des nombreuses alternatives et la complexification associée de la résolution du critère d'optimisation [4, 16, 2].

- *Pour une meilleure prise en compte de l'efficacité des actions de maintenance.*

L'objectif de ce paragraphe est de souligner notre contribution sur le plan de la modélisation de l'efficacité des actions de maintenance. Nous avons appliqué nos modèles pour des politiques de maintenance basées sur l'âge ou conditionnelles. Les actions en tant que telles peuvent être classées en fonction de leur objectif. L'inspection permet de rendre compte de l'état du système alors que les réparations et les entretiens ont pour objectif de remettre le système dans un état *conforme*.

Nous avons effectué une étude sur l'impact de la prise en compte de la qualité des inspections sur la périodicité de celles-ci dans un contexte d'exploitation pétrolier. Une inspection permet de détecter une éventuelle fissure dans la protection interne d'un pipeline. Sans cette protection, le pipeline est sujet à un processus de corrosion important largement favorisé par des conditions environnementales très agressives : très forte pression et mélange liquide (pétrole, gaz, eau salée et sable) extrêmement corrosif. La qualité de l'inspection a été modélisée par une fonction de probabilité de non-détection de la fissure dans la protection décroissante depuis le temps d'initiation [31, 14]. Un modèle étendu pour la comparaison des coûts liés à la maintenance de pipeline en fonction de différentes technologies de contrôle non destructif est discuté dans [14].

Concernant les actions de remise en état, il est clair que les modèles sont dictés par le choix de l'indicateur d'état du système. De plus, l'efficacité de l'action se caractérise par l'état du système à l'issue de la maintenance mais aussi par son impact sur le processus d'évolution de la dégradation après maintenance. Nous avons proposé un modèle de type Kijima ou modèle de réduction de l'âge pour la maintenance imparfaite dans des contextes basés sur l'âge où le vieillissement du système est résumé par un taux de défaillance uniquement fonction de sa date de mise en service [51, 52]. L'originalité provient principalement de la composante aléatoire de la maintenance préventive et du choix d'un critère d'évaluation des performances de la politique reposant sur la période de rentabilité économique du système en exploitation. Par ailleurs, durant ma thèse de doctorat [6], nous avons étendu la modélisation de type réduction de l'âge pour des systèmes à dégradation cumulée. L'efficacité d'une *réparation* est modélisée par une variable aléatoire dite de recouvrement [28] définie par la différence entre le niveau de dégradation avant et après réparation. Les politiques proposées sont des politiques d'inspection/remplacement/réparation pour lesquelles la durée de maintenance est non négligeable. Nous avons aussi montré l'intérêt d'optimiser la qualité de régénération d'une réparation en introduisant un nouveau seuil de décision [12]. L'évaluation numérique de la fonction objectif a nécessité la construction d'un algorithme spécifique reposant sur le théorème du point fixe. Nous avons par ailleurs [33, 53, 37] considéré un impact sur la qualité de service du système après entretien, ce modèle a été développé pour l'entretien des chaussées en considérant qu'une réparation ponctuelle d'une fissure ou *scellement* présente des inconvénients en termes de propriétés d'adhérence, de surépaisseur et donc d'altération du confort de conduite et d'élévation des nuisances sonores. Ainsi, un remplacement à neuf est à reconsidérer après un certain niveau d'entretien.

Les travaux présentés à ce stade ne considèrent pas l'influence sur les paramètres du processus de dégradation après entretien. Dans notre contexte de l'entretien routier, cette hypothèse est extrêmement restrictive. En effet, l'indicateur de décision initialement utilisé est le niveau de fissuration en surface qui revient à zéro après entretien - l'entretien consistant à mettre une couche d'enrobé d'une épaisseur et qualité variables sur toute la surface de la chaussée considérée. Il ne reflète pas le vrai niveau de dégradation de la couche de fond qui est clairement une composante à prendre en compte pour expliquer le vieillissement en surface de la chaussée. Nous avons étudié plusieurs approches dans le projet SBADFoRM pour

répondre à cette problématique. La principale repose sur l'introduction du processus du taux d'accroissement espéré du pourcentage de fissuration sur un intervalle de temps défini dans le paragraphe précédent. Ce taux d'accroissement représente en quelque sorte le potentiel de fissuration de la chaussée et donc donne une information sur le niveau de dégradation de la couche de fond. Un entretien va alors modifier son évolution en fonction de son efficacité (que nous avons résumé par l'épaisseur de la nouvelle couche) mais aussi en fonction de l'état observé avant entretien. Les premières approches [38, 60, 61, 55, 63, 22] définissent l'impact en fonction de la nature de l'entretien (définition de 4 classes d'entretien). Nos derniers modèles [42, 43, 20] cherchent à optimiser la décision en décomposant un entretien en actions de décaissement (enlèvement d'une épaisseur d'enrobé) et de rechargement par une nouvelle couche, la décision étant la définition des épaisseurs respectives sous contrainte d'épaisseur totale de la chaussée.

- *Pour une meilleure prise en compte du contexte décisionnel.*

L'objectif de ce paragraphe est de présenter les travaux que nous avons réalisés pour prendre en compte un contexte plus global de décision, en considérant que les modèles présentés jusqu'ici sont principalement centrés sur les performances *intrinsèques* du système à maintenir.

Les premiers modèles que nous avons développés dans le cadre de ma thèse à l'UTT offrent la possibilité de combiner des approches conditionnelles et des politiques basées sur le temps ou l'âge, politiques motivées par exemple par des contraintes externes ou plus globales à l'exploitation isolée du système étudié [25, 10]. La motivation de ces recherches est alors de bénéficier des intérêts potentiels d'une approche conditionnelle dans un cadre de politiques périodiques imposées, sachant que dans les cas limites, la politique conditionnelle optimale vient se superposer aux politiques périodiques.

Nous avons cherché à intégrer dans les modèles de maintenance classiques des pratiques industrielles plus réalistes comme les délais d'acheminement des pièces de rechange, les contraintes d'exploitation ou obligations économiques en cas de défaillance. Nous avons motivé notre problème sur l'analyse d'un processus d'entretien de pipelines soumis à corrosion en exploitation off-shore [31, 14]. Outre la prise en compte des aspects qualité de l'inspection décrite dans le paragraphe précédent, le modèle de maintenance repose sur le *modèle à délai* plus connu dans le monde industriel et présenté dans la méthodologie *Optimisation de la Maintenance par la Fiabilité* par la terminologie *intervalle P-F* (pour Potential Failure). L'intégration de l'ensemble de ces considérations nous a conduit à proposer une modélisation simplifiée offrant un bon compromis entre complexités théoriques et applicabilité en contexte industriel.

Par ailleurs, les travaux de thèse de Phuong Khanh Nguyen Thi s'intéressent au développement de modèles associant efficacité opérationnelle et préoccupations tactiques ou stratégiques. Dans ce cadre, nous avons proposé un modèle de maintenance conditionnelle pour organiser les opérations dites *réparation imparfaite* et de remplacement en prenant en compte les investissements possibles dans des technologies améliorant les performances moyennes du système courant [40]. La prise en compte de ces aspects permet de considérer la maintenance comme un levier d'amélioration des processus industriels, un potentiel d'investissement pour accroître les performances de l'entreprise et non plus comme un *mal nécessaire*. Par ailleurs, dans le souci de globalisation de la décision par intégration des processus industriels connexes, nous étendons le modèle précédent pour intégrer dans la décision de maintenance et d'investissement le niveau de stock des pièces de rechange, mesure à prendre en considération dans la mesure d'*obsolescence économique* du système en place [41].

Enfin, nous venons avec David Lemoine, Maître Assistant à l'École des Mines de Nantes, de débiter deux thèses dont une CIFRE en collaboration avec General Electric Healthcare pour une meilleure prise en compte des contextes décisionnels en maintenance. La thèse CIFRE a pour objectif d'améliorer le système de gestion et distribution des pièces de rechange pour du matériel médical dans des zones à faible demande. La seconde thèse, plus théorique, cherche à améliorer la planification tactique de la production par une meilleure prise en compte des

aspects maintenance, notamment conditionnelle [44, 64, 66].

Travaux dans le contexte de l'optimisation

L'objectif de cette section est de présenter brièvement les travaux que nous avons développés dans le contexte plus général de l'optimisation, travaux qui peuvent être sans relation directe avec l'optimisation de la maintenance.

- *Amélioration de la qualité des estimateurs*

L'objet de ce travail est de proposer une méthode de calculs d'intervalles de confiance des paramètres estimés pour une loi de Weibull dans un contexte fortement censuré [68]. La technique développée s'appuie sur une intégration de Monte-Carlo se fondant sur la reconstruction des données manquantes et intégrée dans l'algorithme SEM (Stochastic Expectation Maximisation). Ce travail a été intégré dans le logiciel WEIBULL développé par l'équipe Is2 de l'INRIA.

- *Algorithme d'optimisation pour la maintenance*

Le développement et la mise en place industrielle des politiques de maintenance sont confrontés à la complexification de l'expression du critère que l'on cherche à optimiser (due à la construction de politiques de plus en plus générales sur des systèmes de plus en plus complexes). Ce problème ne permet plus de justifier l'utilisation de méthodes exactes pour la recherche des paramètres de décision optimaux dans des temps de calcul *raisonnables*. Nous avons concentré nos recherches sur l'amélioration des performances des algorithmes d'évolution (AE) et plus particulièrement des algorithmes génétiques (AG). Bien que de nombreux travaux utilisent ces algorithmes pour optimiser leur fonction objectif, peu d'entre eux s'intéressent à la validité des résultats proposés. Reposant sur la modélisation de l'évolution de la recherche des AG par une approche markovienne, il nous a été possible de caractériser la non-convergence d'AG fréquemment utilisés dans le cadre du problème d'allocation de la fiabilité pour des systèmes complexes. Issu de ce constat, nous avons prouvé la nécessité d'intégrer un paramètre permettant de concentrer très rapidement la recherche des solutions optimales dans des régions de l'espace contenant des solutions réalisables. Ce résultat théorique a été illustré numériquement sur plusieurs problèmes classiques d'allocation de la fiabilité [30, 29].

- *Optimisation des tournées de véhicules dans un contexte incertain*

Les travaux présentés dans ce paragraphe sont issus de la thèse de Jorge Mendoza [8], thèse réalisée en cotutelle avec l'équipe COPA de l'Université de Los Andès (Colombie). Nous avons cherché à intégrer une composante stochastique sur la demande des clients dans un problème de tournées de véhicules comportant plusieurs compartiments. L'objectif principal de la thèse étant de proposer des algorithmes performants dans ce contexte, nous avons proposé d'étendre des approches classiques en recherche opérationnelle d'optimisation pour la prise en compte de la stochasticité de notre problème. La complexité du problème nous a conduit à nous focaliser sur des approches de type heuristiques [58, 62] et méta-heuristiques [59, 39, 17]. La prise en compte des aspects stochastiques entraîne notamment une difficulté dans l'évaluation de la fonction objectif pour chacune des solutions visitées. Pour contourner ce problème, nous avons reposé l'évaluation de cette fonction par une forme approchée [69] ou par estimation avec des valeurs simulées pour les demandes. Nous avons par la même occasion développé des approches de type multi-objectif pour conjuguer dans la prise de décision performance économique et aversion au risque [39].

2.2 Participation à l'animation scientifique

Encadrements de thèse et contrat de recherche

1. *Titre* : Politiques de Maintenance Conditionnelle pour un Système à Dégradation Continue soumis à un Environnement Stressant
 - *Doctorant* : Estelle Deloux
 - *Financement* : MESR
 - *Directeur* : Christophe Bérenguer
 - *Taux d'encadrement* : 70%
 - *Durée de thèse* : 3 ans
 - *Devenir du doctorant* : Maître de Conférences (UTT - Troyes)
2. *Titre* : Solving Real-World Vehicle Routing Problems in Uncertain Environments
 - *Doctorant* : Jorge Mendoza
 - *Financement* : Cotutelle Los Andè - Mines
 - *Directeur* : Christelle Guéret - Andrès Médaglia (Los Andès)
 - *Taux d'encadrement* : 30 %
 - *Durée de thèse* : 3 ans
 - *Devenir du doctorant* : Maître de Conférences (IMA - Angers)
3. *Titre* : Décision dans l'incertain : Application à la gestion de l'entretien routier
 - *Doctorant* : Mariem Zouch
 - *Financement* : Contrat Région
 - *Directeur* : Philippe Castagliola
 - *Taux d'encadrement* : 30%
 - *Durée de thèse* : 3 ans
 - *Devenir du doctorant* : Chercheur (TNO - Pays-Bas)
4. *Titre* : Politiques de maintenance de systèmes soumis à évolution technologique incertaine
 - *Doctorant* : Thi Phuong Khanh Nguyen
 - *Financement* : MESR
 - *Directeur* : Philippe Castagliola
 - *Taux d'encadrement* : 30 %
 - *Date de soutenance* : 6 Novembre 2012 - *Durée* : 3 ans
5. *Titre* : Optimisation jointe de la planification tactique de production et de maintenance
 - *Doctorant* : Martin Géhan
 - *Financement* : MESR
 - *Directeur* : Nathalie Bostel
 - *Taux d'encadrement* : 30%
 - *Début de thèse* : Octobre 2011
6. *Titre* : Gestion des stocks à plusieurs niveaux de service avec demande incertaine
 - *Doctorant* : Adnane Lazrak
 - *Financement* : CIFRE
 - *Taux d'encadrement* : 30 %
 - *Directeur de thèse* : Pierre Dejax
 - *Début de thèse* : Avril 2012

Au niveau de mes activités contractuelles :

- **Responsable du contrat de recherche SBADFORM** avec le Laboratoire Central des Ponts et Chaussées (Bouguenais) sur des problèmes d'optimisation de l'entretien des routes soumises à fissuration longitudinale. *Durée* : 5 ans (2008 - 2012) - *Montant* : 295 k €

- **Participant à l'action collective** ECND-PdL avec 18 établissements de la Région Pays de Loire sur des problématiques d'évaluation des méthodes de contrôle non destructif. Durée : 4 ans (2010-2013) - Montant : 495 k €
- **Co-responsable du contrat de recherche** Thèse CIFRE entre ARMINES et General Electric HealthCare. Durée : 3 ans (2012-2015) - Montant 60 k €

Évaluation et animation scientifique

Enfin, pour conclure cette partie de présentation de mes activités, ma participation à la vie scientifique peut être résumée par :

- cinq invitations à participer à des jurys de thèse en tant qu'examinateur :
 - A. Ferreira Porto Rosa - Maîtrise Statistique de Procédés par lots à temps variable - décembre 2005 ;
 - H. Khraibani - Modélisation statistique de données longitudinales sur un réseau routier entretenu - octobre 2010 ;
 - Y. Zhang - Maîtrise Statistique des Procédés avec paramètres estimés - août 2011 ;
 - K. T. Huynh - Quantification de l'apport de l'information de surveillance dans la prise de décision en maintenance - novembre 2011 ;
 - C. Jimenez - Modélisation, optimisation et évaluation des impacts de l'introduction des technologies d'auto-identification (RFID) dans un aéronef - mars 2012.
- des invitations pour relectures dans les revues internationales pour les thèmes "Maintenance", "Fiabilité", "Statistiques" et "Modélisation aléatoire" :
 - Reliability Engineering and System Safety
 - European Journal of Operational Research
 - IMA Journal of Management Mathematics
 - Journal of Risk and Reliability
- des invitations pour présentation de travaux (workshops, séminaires industriels ou universitaires) :
 - dans des congrès internationaux : INFORMS (2009 à 2012), Applied Reliability Symposium (San Diego 2009)
 - dans des journées scientifiques : SMAI (Bordeaux 2010), JMFS (Chambéry 2012)
 - dans des groupes de travail industriel : S3 de l'IMdR-SdF (Paris 2011)
 - dans des séminaires universitaires : FIMA (Grenoble 2012), LASQUO (Angers 2012)
- des invitations pour relecture de communications pour les conférences :
 - MOSIM'04 (Conférence Francophone de Modélisation et Simulation, Nantes)
 - ESREL 2009, 2010, 2011 et 2012 (European Conference on Safety and Reliability)
 - Lambda-Mu 7 et 8 (Conférence Francophone de Maîtrise des Risques et Sécurité de Fonctionnement)
- une participation à l'organisation pour la conférence MOSIM'04
- des invitations à expertiser des dossiers scientifiques pour la Région Champagne-Ardenne et un projet porté par Erasmus University (Rotterdam, Pays-Bas).

3 Publications scientifiques

Bibliographie

Chapitres de livre

- [1] M. Rausand, and A. Hoyland. Chapter 11 : Reliability of Maintained Systems. In *System Reliability Theory : Models, Statistical Methods, and Applications*. Wiley, 2003.
- [2] E. Deloux, B. Castanier, and C. Bérenguer. An Adaptative Condition-Based Maintenance Policy with Environmental Factors. In M. J. Kallen and S. P. Kuniewski, editors, *Risk and Decision Analysis in Maintenance Optimization and Flood Management*, pages 137–148. IOS Press, 2009.
- [3] C. Bérenguer, A. Grall, and B. Castanier. Stochastic Modelling and Performance Evaluation of Condition-Based Maintenance for Gradually Deteriorating Systems. In Andrews et al. [5], pages 309–330.
- [4] E. Deloux, B. Castanier, and C. Bérenguer. Environmental Factors and Maintenance Decision Model. In Andrews et al. [5], pages 331–360.
- [5] ESReDA Maintenance Modelling and Applications. *J. Andrews, C. Bérenguer, and L. Jackson, editors. Det Norske Veritas*, 2011.

Thèses de doctorat

- [6] B. Castanier. *Modélisation Stochastique et Optimisation de la Maintenance Conditionnelle des Systèmes à Dégradation Graduelle*. PhD thesis, Université de Technologie de Troyes - France, 2001.
- [7] E. Deloux. *Politiques de Maintenance Conditionnelle pour un Système à Dégradation Continue soumis à un Environnement Stressant*. PhD thesis, Université de Nantes - France, 2008.
- [8] J. E. Mendoza. *Solving Real-World Vehicle Routing Problems in Uncertain Environments*. PhD thesis, Université de Nantes, France, 2009.
- [9] M. Zouch. *Décision dans l'Incertain : Application à la Gestion de l'Entretien Routier*. PhD thesis, Université de Nantes, France, 2011.

Articles publiés dans des revues

- [10] B. Castanier, A. Grall, and C. Bérenguer. A Stochastic Model for Hybrid Maintenance Policies Evaluation and Optimization. *International Journal of Reliability, Quality and Safety Engineering*, pages 233–248, September 2001.
- [11] B. Castanier, A. Grall, and C. Bérenguer. Une Structure Multi-Seuils pour l'Optimisation de la Maintenance Conditionnelle des Systèmes à Détérioration Markovienne Continue. *Journal Européen des Systèmes Automatisés*, 35(10) :1235–1262, 2001.
- [12] B. Castanier, C. Bérenguer, and A. Grall. A Sequential Condition-Based Repair/Replacement Policy with Non-Periodic Inspections for a System subject to Continuous Wear. *Appl. Stochastic Models in Bus. Ind.*, 19 :327–347, 2003.

- [13] B. Castanier, A. Grall, and C. Bérenguer. A Condition-Based Maintenance Policy with Non-Periodic Inspections for a Two-Unit Series System. *Reliability Engineering and System Safety*, 87(1) :109–120, 2005.
- [14] B. Castanier and M. Rausand. Maintenance Optimization for Subsea Oil Pipelines. *International Journal of Pressure Vessels and Piping*, 83(4) :236–243, 2006.
- [15] E. Deloux, B. Castanier, and C. Bérenguer. Maintenance Policy for Deteriorating System Evolving in a Stressful Environment. *Journal of Risk and Reliability*, 222(4) :613–622, 2008.
- [16] E. Deloux, B. Castanier, and C. Bérenguer. Predictive Maintenance Policy for a Gradually Deteriorating System Subject to Stress. *Reliability Engineering and System Safety*, 94(2) :418–431, 2009.
- [17] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. A Memetic Algorithm for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. *Computers and Operations Research*, 37(11) :1886–1898, 2010.
- [18] E. Deloux, B. Castanier, and C. Bérenguer. Environmental Information Adaptive Condition-Based Maintenance Policies. *Structure and Infrastructure Engineering*, 8(4) :373–382, 2011.
- [19] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. Constructive Heuristics for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. *Transportation Science*, 45(3) :335–345, 2011.
- [20] M. Zouch, T. G. Yeung, and B. Castanier. Optimizing Road Milling and Resurfacing Actions. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 226(2) :156–168, 2012.
- [21] T. P. K. Nguyen, B. Castanier, and T. G. Yeung. Optimal Maintenance and Replacement Decisions under Technological Change with Consideration of Spare Parts Inventories. In revision in *International Journal of Production Economics*, 2012.
- [22] M. Zouch, T. G. Yeung, and B. Castanier. A Two-Phase State-Dependent Deterioration Model for Maintenance Optimization. *Naval Research Logistics*, submitted in 2011.
- [23] T. P. K. Nguyen, B. Castanier, and T. G. Yeung. Maintenance Investment Problem under Technological Change with Forecast Horizon Optimal Decisions. *European Journal of Operational Research*, submitted in 2012.

Articles publiés dans des actes de conférences internationales

- [24] C. Bérenguer, A. Grall, and B. Castanier. Simulation and evaluation of condition-based maintenance policies for multi-component continuous-state deteriorating systems. In M.P. Cottam, D.W. Harvey, R.P. Pape, and J. Tait, editors, *Proceedings of the European Safety and Reliability Conference - ESREL'2000 - 15-17 may 2000, Edinburgh, Scotland*, pages 275–282, 2000.
- [25] B. Castanier, A. Grall, and C. Bérenguer. Combining Regulation-Based and Condition-Based Maintenance Policies : a Stochastic Approach. In *Proceedings of the Seventh ISSAT International Conference on Reliability and Quality in Design*, pages 184–189, August 2001.
- [26] B. Castanier, C. Bérenguer, and A. Grall. A Model for Maintenance Cost Optimization under Availability Constraint. In *European Safety and Reliability Conference Proceedings - ESREL'2001 - 16-20 september 2001, Torino, Italy*, 2001.
- [27] B. Castanier, C. Bérenguer, and A. Grall. A (n, N) Maintenance Strategy with Non-Periodic Inspections for a Multi-Component System. In *Proceedings of Mathematical Models for Reliability Conference, MMR 2002*, June 2002.

- [28] B. Castanier, C. Bérenguer, and A. Grall. Stochastic Maintenance Planning for a Repairable System which is Inoperative during Maintenance Operations. In *Proceedings of six PSAM (Probability and Safety Assessment Management) Conference, San Juan, Puerto Rico, USA*, volume 2, pages 1365–1370, June 2002.
- [29] L. Rigal, B. Castanier, and P. Castagliola. Introduction of a New Selection Parameter in Genetic Algorithm for Constrained Reliability Design Problems. In *Proceedings of Genetic and Evolutionary Computation Conference - Seattle, USA*, volume 2, pages 90–101, June 2004.
- [30] L. Rigal, B. Castanier, and P. Castagliola. Introduction of a New Selection Parameter in Genetic Algorithm for Constrained Reliability Design Problems with Several Failure Modes. In *Proceedings of the 10th International Conference on Reliability and Quality in Design - ISSAT - La Vegas, USA*, pages 40–44, August 2004.
- [31] B. Castanier and M. Rausand. Optimization of the Maintenance Decision for a Pipeline in Oil Production. In *Proceedings of the European Safety and Reliability Conference - ESREL'2005, Gdansk, Poland*, volume 1, pages 305–311, June 2005.
- [32] E. Deloux, B. Castanier, and C. Bérenguer. Combining Statistical Process Control and Condition-Based Maintenance for Gradually Deteriorating Systems subject to Stress. In *Risk, Reliability and Societal Safety, Proceedings of ESREL 2007- European Safety and Reliability Conference, Stavanger, Norway*, volume 1, pages 265–272, June 2007.
- [33] B. Castanier and T. G. Yeung. Optimal Highway Maintenance Policies Under Uncertainty. In *Proceedings of the Annual Reliability and Maintainability Symposium. Las Vegas, Nevada, United States*, pages 25–30, January 2008.
- [34] E. Deloux, B. Castanier, and C. Bérenguer. Maintenance Policy for Non-Stationary Deteriorating System. In IEEE, editor, *Annual Reliability and Maintainability Symposium, Las Vegas, USA*, January 2008.
- [35] E. Deloux, B. Castanier, and C. Bérenguer. Comparison of Health Monitoring Strategies for a Gradually Deteriorating System in a Stressfull Environment. In *International Conference on Probabilistic Safety Assesment and Management, Hong Kong, Chine*, May 2008.
- [36] E. Deloux, B. Castanier, and C. Bérenguer. Condition-Based Maintenance Approaches for Deteriorating System Influenced by Environmental Conditions. In *ESREL 2008- European Safety and Reliability Conference, Valencia, Spain*, volume 1, pages 469–475, September 2008.
- [37] T. G. Yeung and B. Castanier. Dynamic Maintenance Ppolicies for Civil Infrastructure to Minimize Cost and Damage Safety Risk. In *Proceedings of the ESREL 2008-European Safety and Reliability Conference. Valencia, Spain*, volume 4, pages 3171–3176, September 2008.
- [38] M. Zouch, T. G. Yeung, and B. Castanier. A Condition-Based Imperfect Maintenance Model with Action Dependent Deterioration. In *Proceedings of the 2009 Industrial Engineering Research Conference - Miami, USA*, June 2009.
- [39] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. A moea for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. In *VIII Metaheuristic International Conference (MIC) - Hamburg, Germany*, 2009.
- [40] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Optimal Maintenance and Replacement Decisions under Technological Change. In *Proceedings of the European Safety and Reliability Conference - ESREL'2010 - Rhodes, Greece*, September 2010.
- [41] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Optimal Maintenance and Replacement Decisions under Technological Change with Consideration of Spare Parts Inventories. In *Proceedings of the 16th International Symposium on Inventories - Budapest, Hungary*, August 2010.

- [42] M. Zouch, T. G. Yeung, and B. Castanier. Optimal condition-based resurfacing decisions for roads. In *Proceedings of the European Safety and Reliability Conference - ESREL'2010 - Rhodes, Greece*, pages 1379–1384, 2010.
- [43] M. Zouch, T. G. Yeung, and B. Castanier. Optimal Resurfacing Decisions for Road Maintenance : A POMD Perspective. In *Proceedings of the Annual Reliability and Maintainability Symposium (RAMS). Lake Buena Vista, Florida, United States*, January 2011.
- [44] B. Castanier and D. Lemoine. A Preliminary Integrated Model for Optimizing Tactical Production Planning and Condition-Based Maintenance. In *Proceedings of the International Conference on Industrial Engineering and Systems Management - IESM 2011 - Metz, France*, pages 998–1007, May 2011.
- [45] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Impact of Maintenance on Replacement Investment under Technological Improvement. In *Proceedings of the European Safety and Reliability Conference - ESREL'2011 - Troyes, France*, pages 869–876, September 2011.
- [46] M. Zouch, T. G. Yeung, B. Castanier, and T. Lorino. Application of a bivariate deterioration model for a pavement management optimization. In *Proceedings of the Transport Research Arena - TRA 2012 - Athens, Greece*, April 2012.
- [47] M. Zouch, T. G. Yeung, and B. Castanier. Bayesian updating for road maintenance optimization. In *Proceedings of the joint Probabilistic Safety Assessment Management Conference - PSAM 11 - and the European Safety and Reliability Conference - ESREL'2012 - Helsinki, Finland*, June 2012.
- [48] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Information Acquisition of New Technology Performance for Maintenance/Investment Decisions. In *proceeding of 3rd IEEE International Conference on Emergency Management and Management Sciences, Beijing, China*, pages 31–34, August 2012.

Articles publiés dans des actes de conférences francophones

- [49] B. Castanier, A. Grall, and C. Bérenguer. Comparaison de politiques de maintenance conditionnelle pour des systèmes multi-composants à dégradation continue. In *Actes des XXXIIe Journées de Statistique - 15-19 may 2000, Fès, Maroc*, pages 199–202, 2000.
- [50] B. Castanier, C. Bérenguer, and A. Grall. Modèle d'évaluation des Performances d'une Politique de Maintenance Conditionnelle avec Prise en Compte de l'Indisponibilité. In *IV-ième Congrès International de Génie Industriel - june, 11-15, 2001, Marseille, France*, pages 459–467, 2001.
- [51] B. Castanier and A. Mohafid. Optimisation de la Maintenance Préventive pour Augmenter la Durée de Rentabilité d'un Système. In *Actes du 6ème Congrès international pluridisciplinaire Qualité et Sûreté de Fonctionnement - Qualita 2005- Bordeaux, France*, volume 1, pages 55–62, March 2005.
- [52] A. Mohafid and B. Castanier. Optimisation de la Maintenance Imparfaite à Efficacité aléatoire. In *Actes de MOSIM'06, Rabat, Maroc*, pages 1319–1325, April 2006.
- [53] B. Castanier and T. G. Yeung. Application des Processus Gamma pour la Gestion de l'Entretien des Routes. In *Workshop 3SGS'08 Surveillance, Sûreté et Sécurité des Grands Systèmes. Troyes, France*, June 2008.
- [54] E. Deloux, B. Castanier, and C. Bérenguer. Prise en Compte de Facteurs de Stress Environnementaux pour l'Optimisation de la Maintenance. In *Workshop 3SGS'08 Surveillance, Sûreté et Sécurité des Grands Systèmes. Troyes, France*, June 2008.

- [55] M. Zouch, T. G. Yeung, and B. Castanier. Un Processus de Décision Bivarié pour l'Optimisation de l'Entretien de Chaussées Soumises à Fissuration. In *Proceedings of $\lambda - \mu 07$ Conference, Institut pour la Maîtrise des Risques et de la Sécurité de Fonctionnement, La Rochelle, France*, October 2010.

Communications dans des conférences sans actes ou à diffusion restreinte

- [56] E. Deloux, B. Castanier, and C. Bérenguer. A Predictive Maintenance Policy Based on Two Explicative Variables. In *INFORMS Annual Meeting. Pittsburgh, Pennsylvania, USA*, November 2006.
- [57] E. Deloux, B. Castanier, T. G. Yeung and C. Bérenguer. A Predictive Maintenance Policy Combining Statistical Process Control and Condition-Based Approaches. In *Proceedings of the 31st ESReDA seminar. Slovakia*, pages 210–220, November 2006.
- [58] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. Construction Heuristics for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. In *CLAIO'08 - Cartagena, Colombia*, pages 113–113, 2008.
- [59] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. A Genetic Algorithm for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. In *EU/MEeting 2008 On Metaheuristics for Logistics and Vehicle Routing - Troyes, France*, 2008.
- [60] M. Zouch, T. G. Yeung, and B. Castanier. Condition-Based Maintenance Models for Civil Infrastructure Subject to Cracking. In *Applied Reliability Symposium, invited paper. San Diego, USA*, June 2009.
- [61] M. Zouch, T. G. Yeung, and B. Castanier. A Conditioned-Based Imperfect Maintenance Model with Action Dependent Deterioration. In *INFORMS Annual Meeting, invited paper. San Diego, USA*, October 2009.
- [62] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. New Constructive Heuristics for the Multi-Compartment Vehicle Routing Problem with Stochastic Demands. In *ALIO-INFORMS Joint International Meeting. Buenos Aires, Argentina*, 2010.
- [63] M. Zouch, T. G. Yeung, and B. Castanier. Développement d'une Approche Conditionnelle pour l'Entretien de Chaussées Soumises à Fissuration. In *Journées Modélisation Aléatoire et Statistique de la Société de Mathématiques Appliquées et Industrielles, invited paper. Bordeaux, France*, September 2010.
- [64] B. Castanier and D. Lemoine. Optimisation conjointe planification tactique et maintenance pour un système à dégradation stochastique continue. In *Actes du 12ème congrès annuel de la société française de recherche opérationnelle et d'aide à la décision - ROA-DeF 2011 - Saint-Etienne, France*, March 2011.
- [65] M. Zouch, T. G. Yeung, and B. Castanier. How to take into account potential change of the deterioration mode in Condition-Based Maintenance decision rule. In *ES-REDA Conference 2011 Advances in Reliability-based Maintenance Policies. La Rochelle, France*, October 2011.
- [66] M. Gehan, B. Castanier, and D. Lemoine. Comparaison de deux règles pour l'intégration de la maintenance conditionnelle dans un plan de production. In *Actes du 13ème congrès annuel de la société française de recherche opérationnelle et d'aide à la décision - ROA-DeF 2012 - Angers, France*, April 2012.
- [67] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Information acquisition of new technology performance for maintenance/investment decisions. In *25th European Conference on Operational Research, Vilnius, Lithuania*, 2012.

Rapports techniques

- [68] B. Castanier. Estimation des Variances pour des Modèles de Durées de Vie Censurées. Technical report, Rapport DER-EDF/Is2-INRIA, 1998.
- [69] J. E. Mendoza, B. Castanier, C. Guéret, A. L. Medaglia, and N. Velasco. Approximating the Expected Cost of Recourse on a Multi-Compartment Vehicle Routing Problem with Stochastic Demands. Technical report, 08/3/AUTO, Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN). Ecole des Mines de Nantes, France, 2008.

Synthèse scientifique

Chapitre 1

Introduction

L'intérêt principal de mes recherches se définit dans la construction de modèles décisionnels en contexte incertain qu'il soit dû au comportement aléatoire des systèmes étudiés, des environnements dans lesquels ils évoluent ou encore à la connaissance et la caractérisation de ces modes d'évolution. Les premiers points réfèrent plus à l'analyse des comportements stochastiques des phénomènes étudiés alors que les deux derniers portent essentiellement sur l'analyse statistique des données collectées et l'information disponible.

Dans ce contexte général, l'objectif de mes recherches est alors de proposer des modèles de maintenance qui permettent de répondre aux préoccupations industrielles de plus en plus complexes par une meilleure prise en compte tant de la réalité industrielle que du contexte décisionnel tout en assurant une simplicité de mise en œuvre des règles de gestion dans le contexte opérationnel. Par ailleurs, cette complexité ne permettant plus l'utilisation d'approches de résolution classiques, la recherche de nouvelles méthodes d'optimisation est devenue nécessaire. Mes contributions se situent alors tant sur le champ de la modélisation que sur le domaine de l'optimisation en soi.

Deux schémas de pensées peuvent être déroulés dans la construction de modèles de décision optimaux. La première approche consiste à proposer a priori une structure décisionnelle généralement paramétrique qui régit la décision en fonction du contexte opérationnel courant. L'optimisation de cette structure repose alors sur la détermination de ses paramètres qui optimisent un critère de décision, critère généralement moyen par unité de temps. La difficulté majeure dans cette approche est l'évaluation numérique de la fonction objectif qui repose sur l'étude du comportement stationnaire du processus stochastique étudié. Mes travaux développés dans ce contexte se basent sur des résultats avancés de la théorie des processus de renouvellement, renouvellement classique ou renouvellement markovien. Même si les aspects d'optimalité au sens large des politiques envisagées ne sont pas démontrés, cette approche m'a clairement permis d'étudier les relations entre modélisation de l'indicateur d'état du système étudié et processus de décision. Par ailleurs, cette approche présente l'avantage de pouvoir évaluer des politiques pour lesquelles les paramètres choisis ne sont pas les paramètres optimaux mais adaptés par le décideur, par exemple, au regard d'éléments contextuels non modélisés dans le critère de décision.

La seconde approche ne préjuge pas d'une structure de décision a priori mais, si elle existe, sera déduite de la procédure d'optimisation. Elle consiste en l'étude directe du processus de décision auquel on associe une fonction gain ou récompense. Le processus de décision est composé en chaque instant d'un état du système étudié et d'une action. La politique est alors définie par l'ensemble des actions associées à chacun des états du système. L'optimisation consiste à déterminer de manière itérative la meilleure politique (i.e. celle qui génère le maximum de récompense) en fonction du critère de décision, généralement un critère escompté. Mes travaux reposent ici sur des modélisations de type Processus de Décision Markoviens. Le premier avantage d'une telle approche est le caractère optimal de la politique proposée qui peut ne pas être intuitive. Ceci m'a notamment permis d'étudier les performances de

politiques intégrant de nouveaux indicateurs de détérioration. Le second avantage est la seule caractérisation d'une fonction de transition d'états entre deux instants -en opposition à la caractérisation du comportement stationnaire du processus dans la première approche- qui permet alors d'étendre le contexte décisionnel par la prise en compte de nouvelles hypothèses, par exemple, ou l'élargissement de l'ensemble des actions disponibles. La difficulté est ici plutôt liée au choix de l'algorithme d'optimisation en soi et la stratégie d'exploration de l'espace des solutions qui est généralement très grand.

La présentation à suivre de mes contributions reprend ce schéma en différenciant les approches que j'appellerai paramétriques aux approches non paramétriques, la notion de paramètre est clairement à associer à la structure décisionnelle. Ainsi, dans le premier chapitre, la discussion est essentiellement centrée sur des approches d'optimisation de la maintenance conditionnelle qui reposent sur la construction et l'optimisation de seuils de décision. Le chapitre suivant reprend l'ensemble des travaux que nous avons développés dans le cadre d'un projet en collaboration avec le département ESAR de l'IFSTTAR (anciennement LCPC. Liée à une restructuration, la structure de ce département a beaucoup évolué). Ces travaux concernent l'optimisation de la gestion de l'entretien des chaussées soumises à fissuration longitudinale. Ce contexte nous a permis de développer un certain nombre d'approches de type Programmation Dynamique Stochastique. Le troisième chapitre cherche à répondre à des questions de gestion de la maintenance lorsque les systèmes technologiques subissent des évolutions dans le temps. La problématique d'obsolescence et les incertitudes sur les générations à venir sont ainsi à rajouter dans la définition de ces stratégies de maintenance. Les approches proposées sont elles aussi développées dans un contexte non paramétrique et reposent sur le concept *des options réelles*. La conclusion de mon mémoire permettra de positionner les axes de mes futurs travaux. Enfin, une sélection d'articles est proposée à la fin de ce manuscrit.

Je tiens dès à présent à souligner que j'ai pris le parti de ne pas présenter l'ensemble de mes contributions, contributions rapidement mentionnées dans la première partie de ce document, notamment sur la résolution de problèmes de tournées de véhicules soumis à des aléas. Ce choix est lié à la fois à un aspect de concision mais aussi de cohérence de présentation vis-à-vis du thème de recherche principal qu'est la maintenance. Je m'en excuse au près des personnes avec qui j'ai pu travailler.

Chapitre 2

Approches paramétriques pour l'optimisation de la maintenance conditionnelle

1 Introduction : Contexte et objectifs

Outre les objectifs classiques d'assurance ou maintien des performances des systèmes actuels définis dans les normes industrielles telles que, par exemple, la norme CEI 60300-3-16 [4], la maintenance industrielle¹ se positionne aujourd'hui comme un véritable levier d'amélioration des performances d'entreprises en tant que l'une des activités clés de l'*Excellence Opérationnelle*² ou de *Lean Manufacturing*. Par ailleurs, la prise en considération de l'ensemble des coûts liés au cycle de vie des produits dès la phase de conception a, elle aussi, eu pour effet de positionner la maintenance comme un réel centre d'intérêt et de bénéfice. Enfin, associées à la maximisation de la disponibilité des systèmes, la prise de conscience collective et les responsabilités des entreprises vis-à-vis des risques qu'elles génèrent ont fortement contribué à une meilleure prise en compte des facteurs de maintenance. Cette prise en compte a conduit au développement de méthodes rigoureuses pour l'analyse des systèmes et de définition de plans d'amélioration efficaces (la maintenance pouvant être considérée ici comme source potentielle de gain mais aussi de danger).

Adossée à ces motivations industrielles, on trouve une très forte activité de recherche académique pour la proposition de modèles mathématiques de planification de la décision de maintenance. Sans revenir sur les modèles et typologies de maintenance, on peut cependant noter que la grande majorité des travaux de la seconde moitié du siècle dernier se concentrent sur les approches quantitatives basées sur l'analyse statistiques des durées de vie des systèmes, bénéficiant d'une recherche intensive du domaine de la Fiabilité des systèmes (le lecteur pourra se référer aux revues des modèles de maintenance [88, 106, 32, 113]). Ces modèles reposent sur des formalismes mathématiques avancés qui permettent d'intégrer dans les critères de décision de nombreuses hypothèses rendant compte de la complexité du contexte d'exploitation des systèmes : les effets de l'environnement impactant les performances des systèmes, la prise en compte de nouvelles données, l'efficacité aléatoire de la maintenance, les risques concurrents liés à la multiplicité des modes de défaillance dépendants, etc.. Cependant, on peut adresser plusieurs critiques à ces modèles. La première et non des moindres rejoint l'analyse de

1. **Maintenance** : Ensemble des actions techniques et administratives correspondantes, y compris les opérations de surveillance et de contrôle, destinée à maintenir (maintenance préventive) ou à rétablir (maintenance corrective) une entité dans un état spécifié ou dans des conditions données de sûreté de fonctionnement (disponibilité, fiabilité, maintenabilité et sécurité) lui permettant d'accomplir une fonction requise

2. **Excellence Opérationnelle** : une démarche systématique et méthodique menée en entreprise pour maximiser les performances en matière de productivité, de qualité des produits et de réduction des coûts, et plus généralement de performance.

P. Scarf [95] sur l'applicabilité des modèles et l'importance de collaborer entre industriels et chercheurs académiques. Il est clair que trop souvent les modèles développés par des mathématiciens chevronnés, renforcés par l'usage de terminologies scientifiques mal adaptées et de concepts théoriques avancés ne permettent pas d'en assurer leur compréhension et leur prise en main par les utilisateurs dans des contextes pratiques et techniques de planification de la maintenance [19]. Le second est lié au grand nombre de données à collecter pour assurer la qualité d'ajustement des modèles de fiabilité, données pour lesquelles la nature n'est pas toujours disponible dans le contexte opérationnel. Dans la même lignée, on notera que les données recherchées ne reposent pas forcément sur des informations correspondantes à des phénomènes connus de dégradation, en opposition aux modèles classiques de fiabilité structurelle développés dans les domaines de l'ingénierie mécanique et du génie civil, reposant principalement sur l'explication de la relation résistance-contrainte. Or, ces informations du type amplitude et fréquence de vibrations, niveau de corrosion, etc., par le biais des développements récents des capteurs et de techniques d'auscultation associées, sont de plus en plus disponibles. Plus récemment, dans la suite de Makis et Jardine [60] dans lesquels ils démontrent les conditions d'optimalité de leur politique de maintenance, on trouve de nombreux travaux qui intègrent ces informations, appelées ici *covariables*, dans la décision de maintenance en se basant sur le modèle des hasards proportionnels [28]. Cependant, la construction du modèle de défaillance repose toujours sur des considérations purement statistiques avec des lois de fiabilité *générales* nécessitant un grand nombre d'observations et de défaillances associées.

Afin de contourner la problématique du manque de données, il peut être intéressant d'intégrer de la connaissance, lorsqu'elle est disponible, sur les phénomènes étudiés. Nicolaï [83] définit une classification des modèles de dégradation des systèmes allant du tout statistique pour les modèles de type *boite noire* aux modèles de type *boite blanche* dont la motivation principale est l'explication et la modélisation du processus de défaillance au travers de calculs de structures empruntés au domaine de la Mécanique et du Génie Civil. Les modèles développés sont généralement des modèles spécifiques à des modes de dégradation bien identifiés tels que la fatigue, l'érosion, la corrosion, etc.. Par ailleurs, ces modèles de prédiction reposent sur de nombreux paramètres souvent difficiles à estimer dans des contextes opérationnels pour lesquels de nombreuses incertitudes et de variabilité des conditions opératoires sont à intégrer. Ce point peut être vu comme une des explications quant à la qualité prédictive de ces modèles. De plus, leur intégration dans des critères décisionnels étendus reste délicate et extrêmement consommatrice de ressources, notamment numériques. Une approche proposée pour réduire ces efforts est l'utilisation de simulations stochastiques de type Monte carlo. Les approches de type *boîtes grises* reposent sur l'utilisation de processus stochastiques pour la modélisation de l'évolution de l'indicateur de dégradation pouvant prendre plusieurs valeurs, ne se limitant plus ainsi à la description binaire *en fonction* ou *en panne*. Le choix des processus stochastiques, la caractérisation des *états du système* et les fonctions de transition associées doivent être motivés par la connaissance du phénomène physique de dégradation et leur caractérisation repose sur des considérations statistiques liées aux observations. Ces approches sont particulièrement adaptées à des processus de dégradation *lents* ou *progressifs* pour lesquels il est possible d'obtenir des informations entre l'initiation du *symptôme* ou *défaute* et la défaillance du système. Avant de centrer ma discussion sur ces dernières approches, section 2.1, je voudrai souligner que le lecteur peut trouver de nombreuses discussions sur l'intérêt de chacune des approches [102, 105] mais que, même si certaines divergences peuvent être constatées, les points de vue se rejoignent dans la motivation d'intégrer un ensemble de connaissances et de données dans les modèles décisionnels.

L'intérêt pour les approches de maintenance conditionnelle peut être jugé comme relativement récent. Jusque dans les années 2000, les recherches sur la maintenance conditionnelle se concentraient essentiellement sur la construction de modèles mathématiques pour l'évaluation et l'optimisation de critères de décision assez simples et les efforts se concentraient majoritairement autour de la modélisation des phénomènes de dégradation et l'étude des propriétés

asymptotiques des processus stochastiques associés. Bien que motivés par un intérêt certain pour leurs performances supposées dans des contextes opérationnels, on trouve peu de travaux d'application aux problématiques industrielles. Cette absence de *transfert scientifique* peut certes s'expliquer par un certain manque de *vulgarisation* des résultats théoriques mais aussi par la limitation des processeurs informatiques ne permettant pas d'offrir la puissance nécessaire à la conduite de nombreuses expérimentations numériques pour la mise en évidence de la flexibilité et les gains de telles approches ainsi que l'obtention de résultats dans des durées *convenables* dès que l'on cherche à étendre les hypothèses pour une meilleure prise en compte du contexte décisionnel. Ce n'est que vers la fin des années 2010 que l'on trouve des travaux d'application de ces modèles, principalement sur des problèmes où la maintenance joue un rôle prédominant sur la sûreté d'exploitation des systèmes.

Dans ce contexte de recherche, les travaux que je vais présenter dans ce chapitre se situent essentiellement sur le point de vue théorique pour lequel nous avons cherché à améliorer l'applicabilité des approches conditionnelles soit par le biais de modélisations de dégradation et de l'influence des contextes opératoires fluctuants tels que l'environnement, soit sur le processus même de maintenance et la nature des opérations de maintenance ou encore par l'extension du contexte décisionnel en intégrant diverses contraintes telles que disponibilité minimale ou schéma d'inspection périodique.

Plus précisément, nous pouvons souligner les contributions principales suivantes :

- la construction d'un schéma conditionnel d'inspection permettant d'adapter les intervalles inter-inspection en fonction de l'observation ;
- la prise en compte des effets d'indisponibilité des matériels, indisponibilité liée à la perte de fonction du système ou liée aux interventions de maintenance ;
- la modélisation et l'optimisation de maintenances imparfaites ;
- la considération de défaillances multiples dépendantes ou indépendantes ;
- la modélisation et l'étude des effets environnementaux sur la décision de maintenance.

Avant de détailler l'ensemble de ces contributions ainsi que les perspectives, je rappellerai la démarche générale pour la construction d'un modèle de maintenance allant de la spécification des indicateurs de performance jusqu'à l'analyse des bénéfices attendus.

2 La méthodologie : de la modélisation à l'optimisation de la décision

L'objectif de cette section est de présenter l'approche générique que nous avons développée pour la construction et l'analyse des performances des différents modèles de maintenance. Cette présentation est ici restreinte aux approches conditionnelles pour un système dit mono-composant (les hypothèses spécifiques aux systèmes multi-composant seront brièvement introduites dans la section concernée, Section 4). Cette discussion permettra dans un même temps de poser des hypothèses générales que j'utiliserai dans la suite de ce chapitre.

Choix et construction de l'indicateur de performance : L'indicateur est une valeur mesurable qui doit permettre de rendre compte de l'état réel de santé du système et de sa capacité à assurer dans le temps la ou les fonctions pour lesquelles il est conçu. La difficulté est alors de construire des modèles associant ces mesures généralement liées à des quantités physico-chimiques telles que le niveau d'érosion, de fissuration, de vibration, la concentration ionique dans les matériaux etc. aux niveaux de satisfaction des fonctions du système étudié et donc à la défaillance proprement dite. L'étude des relations entre évolution des pathologies ou mécanismes de dégradation et les modes de défaillance associés généralement attribuée au Génie Mécanique et Matériaux et, plus spécifiquement, au domaine de la fiabilité des structures n'entre pas dans le cadre de ce chapitre. Nous considérons ici que l'indicateur de performance est donné et qu'il peut s'apparenter à une mesure de vieillissement du système

qui croît de manière continue en fonction de son usage -la motivation et la modélisation de cet indicateur seront discutées dans la section 2.1-. Le mode de défaillance retenu ici est l'accumulation excessive de dégradation définie par un seuil que l'on supposera connu.

Acquisition et traitement des observations : La caractérisation des mécanismes de dégradation repose sur l'étude d'observables pour lesquels des mesures ou observations sont réalisées de manière continue ou ponctuelle. Accessibilité et qualité de ces observations définie en tant que distance entre information acquise et l'état réel du système sont dépendantes des techniques et technologies utilisées. Ce dernier point est actuellement source de très nombreux travaux de recherche tant au niveau des solutions technologiques qu'au niveau des outils d'analyse et de traitement statistique de l'information. Le traitement de ces données consiste d'une part à étudier l'adéquation entre le phénomène recherché et son observation (ceci intègre aussi les problématiques de reconstruction des phénomènes à partir des observations) et d'autre part à construire, à partir de ces observations complétées éventuellement par de l'expertise, des modèles prédictifs d'évolution de ces observables. Ces derniers points conduisent à des études spécifiques en fonction des systèmes étudiés. Nous considérons ici que les observations sur les observables sont directement accessibles et rendent parfaitement l'information sur l'état du système. Par ailleurs, les aspects estimation reliant les données aux modèles de prévision ne seront pas ici directement développés. Somme toute, je discuterai brièvement des méthodes classiques présagées pour l'estimation des paramètres sans chercher à prouver leur efficacité, ce champ relève plus de la statistique et nécessite une mise en œuvre sur des données réelles.

Étude des effets de la maintenance : Une action de maintenance permet de restaurer le système dans un état fonctionnel satisfaisant. L'effet de cette action se mesure généralement au travers de la variation des observables avant et après maintenance et il est alors possible d'en estimer le *gain fonctionnel* associé (en considérant toujours l'existence de la relation injective *observables* et *niveau de satisfaction de la fonction*).

Les hypothèses classiques se définissent par des maintenances dites *aussi bien que neuf* qui considèrent des effets parfaits équivalents à un remplacement par une entité neuve présentant les caractéristiques de dégradation identiques au système dégradé. Comme rapidement discutée dans l'introduction pour des modèles de durée de vie, cette hypothèse est très restrictive et ne rejoint pas l'idée de *réparation* qui peut aller jusqu'à considérer une dégradation supplémentaire en cas de mauvaise maintenance. Nous poserons dans la suite de ce manuscrit qu'une maintenance entraîne de facto une réduction des effets de dégradation sur le système.

Par ailleurs, un effet important à étudier est la réduction de la disponibilité des systèmes liées à la maintenance par la prise en considération de durées non négligeables au regard de l'horizon d'exploitation qui sont plus ou moins bien maîtrisées. Finalement, peu d'études sont disponibles dans la littérature sur ce sujet. Les hypothèses classiques peuvent être, de mon point de vue, particulièrement renforcées dans des contextes de sécurité pour lesquels la connaissance de l'indisponibilité du matériel liée à la maintenance peut permettre des procédures de substitution à la fonction principale. Par contre, dans des contextes de production pour lesquels les conséquences d'un arrêt se traduisent directement en termes de pertes économiques par exemple, une meilleure caractérisation de ces temps de maintenance peut être d'intérêt, d'autant que les modes de défaillance que nous étudions ici entraînent des interventions relativement importantes.

Enfin, dans l'ensemble des actions de maintenance disponibles, les inspections permettent de rendre compte de l'état courant du système sans pour autant contribuer à améliorer ou à dégrader l'état du système -on se limitera ici à des inspections de type non destructif.

Construction et résolution du critère de décision : Ce point est, de mon point de vue, prépondérant dans le processus global car il se doit de préjuger de l'intégration des

objectifs opérationnels du gestionnaire et de la faculté propre au modèle mathématique d'optimisation à être résolu. Sa résolution devra permettre de définir les décisions optimales en fonction des observations. La complexité du critère peut conduire à de nouvelles hypothèses de modélisation dont la validité dans le contexte opérationnel se doit d'être discutée. Sur le plan de la résolution, de nombreuses méthodes existent. L'une des premières classifications peut être liées aux aspects statiques et dynamiques des problèmes de décision. Alors que les modélisations graphiques et les outils associés offrent des solutions très prometteuses dans un contexte statique pour des systèmes complexes (on peut mentionner ici, par exemple, les réseaux bayésiens), leur application dans un contexte dynamique repose essentiellement sur de la simulation stochastique (par exemple, les réseaux de Petri) engendrant alors des difficultés au niveau de l'optimisation et des temps de résolution, difficultés renforcées lorsque les modèles sont continus (par rapport au temps mais aussi aux états). Dans la suite de ce manuscrit, ces outils ne seront pas étudiés. Je me concentrerai sur la proposition de modèles nécessitant des formalismes mathématiques plus *conventionnels* reposant sur des procédures d'optimisation numériques exactes, ces modèles proposant finalement des optimisations locales au vu d'intégration dans des modélisations de systèmes plus complexes.

Comme nous l'avons évoqué dans l'introduction, l'une des approches que l'on peut mettre en place pour optimiser la maintenance repose sur la proposition d'une structure définie a priori régissant la décision en fonction d'un contexte opérant observé. La mise en place de cette approche peut se décliner en 4 étapes.

1. **Choix d'une structure décisionnelle** : Une fois les aspects modélisation des performances du système et des actions de maintenance définis, l'approche consiste à déterminer une structure décisionnelle a priori permettant d'organiser la décision en fonction de l'information disponible. L'un de nos objectifs est de proposer des règles de décision suffisamment simples pour s'assurer de leur applicabilité dans le contexte opérationnel tout en assurant un caractère optimal.
2. **Construction du critère d'optimisation** : L'étape suivante consiste à construire le critère d'optimisation. Dans ce chapitre, seuls des critères asymptotiques par unité de temps seront traités. Il est clair que l'impact de l'état du système ainsi que la décision se doivent d'être définis en fonction de la nature du critère. Par exemple, pour un critère purement économique, on définira des coûts pour les différentes actions de maintenance ainsi que d'exploitation du système fonction de l'état de disponibilité du système.
3. **Évaluation de la fonction objectif** : La résolution du critère nécessite l'évaluation de la fonction objectif qui reste en soi le processus critique de l'approche, surtout lorsqu'on cherche à complexifier l'environnement décisionnel. Il est ainsi nécessaire de s'assurer de l'extractabilité du problème par le biais d'hypothèses sur le problème décisionnel (lois de dégradation du système, effets des actions de maintenance et règles de décision). Ici, nous nous placerons dans un contexte de renouvellement markovien qui nous permettra de caractériser le comportement stationnaire du système soumis à la politique de maintenance et ainsi réduire l'étude asymptotique du système à un horizon de temps fini. La loi stationnaire du système maintenu repose sur la résolution de l'*équation du renouvellement* pour laquelle il n'est pas toujours possible de trouver une solution analytique. Des méthodes classiques d'intégration numérique sont alors utilisées pour l'approcher.
4. **Optimisation paramétrique** : Dans ce chapitre, l'ensemble des problèmes d'optimisation présentés sont des problèmes paramétriques continus contraints ou non contraints. Je n'ai pas cherché à développer des méthodes appropriées d'optimisation et les résultats présentés sont directement issus d'approches d'optimisation numérique classiques de type *descente de gradient* disponibles dans les logiciels Matlab® et Scilab®. L'utilisation de ces programmes itératifs repose sur la construction de fonctions appropriées pour l'évaluation des différents termes constituant la fonction objectif. Là aussi, des méthodes classiques d'intégration numérique seront utilisées.

La décision : Le résultat du problème d'optimisation confronté à la donnée d'expert offre au gestionnaire de maintenance une aide à la décision. Dans le contexte de ce mémoire, seules des études reposant sur des données génériques sont réalisées. Je ne chercherai donc pas à mesurer la pertinence et la qualité des résultats dans un contexte opérationnel mais proposerai des comparaisons avec les stratégies dites classiques. Par ailleurs, des analyses de sensibilité en fonction des paramètres d'entrée sont systématiquement conduites, généralement par le biais d'expérimentations numériques. Ces analyses permettent d'identifier des comportements généraux des politiques et leur traduction en *bonnes pratiques* se définit clairement comme support à l'aide à la décision dans le contexte opérationnel.

2.1 Choix des modèles de dégradation du système

Les systèmes d'intérêt sont des systèmes pour lesquels la propagation de la dégradation est lente, assimilable à un processus cumulatif au cours du temps. Ce processus peut soit être continu dans le temps, soit assimilable à un processus ponctuel pour lequel l'évolution se fait par sauts successifs. À titre d'exemple, on peut citer des dégradations telles que la fissuration de structures ou d'infrastructures, l'érosion d'une digue ou d'engrenages, la corrosion d'un béton ou d'une surface métallique, l'augmentation de phénomènes vibratoires liée à la perte progressive de calage et d'usure pour des machines tournantes ou des problèmes d'écaillage pour des roulements à billes. On peut associer à ces modes de dégradation des modes de défaillance liée à une accumulation excessive. La cessation de la fonction à assurer peut être déterministe ou stochastique, connaissant le niveau de dégradation. Sans trop revenir sur les motivations pour la modélisation de la dégradation par le biais des processus stochastiques, motivations énoncées en introduction de ce chapitre, les choix sont de plusieurs types [14] :

- L'introduction d'un seuil délimitant zone de fonctionnement et zone de défaillance. Ce seuil peut être connu ou donné par une distribution de probabilité.
- La définition de distributions de durées de vie pour tout niveau de dégradation. Cette modélisation est généralement liée à une discrétisation de l'espace d'état. Ainsi, l'évolution du système depuis un état de fonctionnement donné est soit une défaillance pour un temps de séjour trop long, soit le passage dans un état plus dégradé.
- Lorsque la discrétisation n'est pas pertinente (pour des phénomènes continus, par exemple), on peut chercher à définir directement une distribution de durée de vie en fonction des observations. On retrouve notamment les modèles de hasards proportionnels.

Tous mes travaux présentés ici sont centrés sur des modélisations déterministes de la défaillance via l'introduction d'un seuil de défaillance supposé donné.

Les premiers modèles que j'ai étudiés considèrent que, sur une grille de temps $\{t_j, j \in \mathbb{N}\}$, l'évolution de la dégradation est donnée par la somme d'incrément positifs de dégradation mutuellement indépendants et identiquement distribués. La distribution d'un incrément ΔX sur une période de temps fixe δt est une loi exponentielle de paramètre fixé α . Outre les aspects cumulatifs, la motivation d'un tel processus se retrouve à la fois sur la simplicité de mise en œuvre, notamment au niveau de l'estimation statistique pour laquelle seules les valeurs des incréments de dégradation entre deux périodes sont nécessaires, et les propriétés mathématiques attractives de stabilité par convolution (la convoluée d'une loi de k -Erlang de paramètre α par une autre k' -Erlang de même paramètre est une $(k+k')$ -Erlang de paramètre α). Cette dernière propriété permet l'atteinte de l'expression analytique de la fiabilité.

Les seconds modèles utilisés dans le contexte de maintenance sont les processus gamma stationnaires pour lesquels les incréments de dégradation sur un intervalle de temps sont distribués suivant une loi gamma pour laquelle le paramètre de forme est linéaire en fonction de l'intervalle de temps considéré alors que le paramètre d'échelle reste constant. Je ne chercherai pas à justifier ici de l'utilisation de tels processus dans un contexte applicatif car c'est le sujet de nombre d'introductions de travaux sur ce thème (voir [109] pour un état de l'art récent sur le domaine). Par contre, concernant l'extension du premier modèle, l'intégration

d'un paramètre d'échelle permet d'en accroître la flexibilité et l'adéquation inférentielle. Par ailleurs, la continuité (par rapport à la mesure de Lebesgue) d'un tel processus lui confère un réel intérêt pour l'optimisation des périodes d'inspection.

2.2 Les hypothèses de maintenance et choix de la structure

Pour rappel, nous considérons ici deux types d'action de maintenance à savoir les inspections supposées parfaites et non destructives et les opérations de maintenance en tant que telles qui permettent de rétablir le système dans un meilleur état.

L'action d'inspection permet de récupérer de l'information sur le système à savoir son état de fonctionnement (marche ou défaillant) mais aussi le niveau de dégradation courant, $X_t = x$. L'hypothèse d'inspection parfaite considère une information exacte sur cet état. On les suppose non destructives dans le sens où l'état du système après une inspection reste inchangé.

À l'issue d'une inspection, une décision correspondant à la mise en place immédiate ou non d'une maintenance est prise. Un ensemble d'actions est disponible. Une des contributions de nos travaux se situe par rapport à la modélisation de l'efficacité de celles-ci (et les coûts associés pouvant être relatifs à cette efficacité). L'efficacité se mesure directement par la réduction du niveau de dégradation avec comme efficacité maximale une remise à zéro du niveau de dégradation. Cette dernière opération à efficacité maximale permet alors de considérer le système dans un état dit *aussi bon que neuf*, elle est par ailleurs assimilable à un remplacement par un système neuf et identique. Notons enfin qu'une action de maintenance influe uniquement sur le niveau de dégradation et aucunement sur sa loi d'évolution.

En termes de coût, je positionnerai nos travaux dans les hypothèses les plus classiques à savoir des coûts de mise en place d'opération constants et croissants en fonction de l'efficacité et gravité des situations. On obtient ainsi l'ordre suivant : $C_i \leq C_p \leq C_c$ où C_i est relatif à une inspection, C_p à une action préventive et C_c à un remplacement correctif (on note ici qu'une défaillance entraîne de facto une remise à neuf du système). Le choix de coûts unitaires constants, bien que peu réalistes, permet de simplifier les quantités mathématiques nécessaires à l'évaluation des critères d'optimisation et d'en assurer l'expression analytique. Nous avons considéré un coût d'immobilisation constant par unité de temps d'indisponibilité, indisponibilité liée à une défaillance ou bien à une immobilisation pour entretien. Une telle modélisation permet ainsi d'assurer aux politiques de maintenance les meilleurs compromis entre risque de défaillance et sur-maintenance.

Il est maintenant nécessaire de définir comment organiser la décision de maintenance, à savoir quand inspecter et quand entretenir. Pour se faire, nous imposons une structure décisionnelle à seuils multiples qui, fonction de l'observation courante, préconise l'opération à réaliser ainsi que la date de prochaine intervention. La structure peut alors être représentée par une fonction paramétrique dont l'expression est donnée mais dont les paramètres seront optimisés dans un second temps. Bien que l'optimalité de telles politiques à seuil ne soit pas toujours validée d'un point de vue mathématique, son intérêt réside essentiellement par une facilité d'implémentation et d'acceptation -celle-ci rejoignant les idées intuitives de règles de maintenance- dans un contexte industriel.

2.3 Le critère décision : construction et évaluation du critère

On rappelle ici que le critère de décision choisi est la minimisation de l'espérance des coûts de maintenance par unité de temps sur un horizon de temps infini. La construction repose sur l'identification de l'ensemble des contributions du processus de maintenance à savoir, au vu de nos hypothèses, les coûts d'inspection, les coûts de remplacement préventif ou correctif ainsi que les coûts d'indisponibilité, l'ensemble des coûts cités étant des coûts cumulés sur l'horizon d'étude, et l'identification de l'ensemble des scénarios d'évolution du

système après maintenance. La caractérisation de ces coûts moyens cumulés repose sur l'étude du comportement stationnaire du système soumis à la politique de maintenance.

Les hypothèses du modèle dégradation et des effets de la maintenance sur celle-ci assurent la propriété markovienne pour le système maintenu. Par ailleurs, l'hypothèse d'une action *parfaite* mise en place notamment après l'observation d'une défaillance assure *presque sûrement* le retour du système dans son état initial et donc la propriété de renouvellement markovien au processus décrivant la dégradation du système maintenu. L'utilisation de cette propriété sera le *cœur* de notre approche d'évaluation du comportement stationnaire. En effet, le théorème clé du renouvellement [26, 93] permet de restreindre l'étude asymptotique au comportement moyen du système défini sur un cycle de renouvellement, cycle défini par exemple, dans le cas du renouvellement classique, par l'intervalle de temps entre deux remises à neuf successives. Même si cette simplification permet d'en améliorer la résolution, la multiplication des scénarios d'évolution dans un cycle de renouvellement liée au relâchement d'une hypothèse (par exemple, la prise en compte de l'efficacité aléatoire d'une maintenance 3.2), pousse à l'utilisation de résultats avancés sur le renouvellement Markovien au détriment du renouvellement simple. L'obtention de la loi de transition dans un cycle de renouvellement markovien dit aussi *de semi-régénération* est aisée mais la difficulté provient alors de la construction de la loi d'entrée dans chaque cycle. L'utilisation de résultats classiques liés à l'étude de la loi stationnaire d'une chaîne de Markov ou encore sur la résolution de l'équation de renouvellement [26] ne permet pas toujours l'obtention de telles lois. Des approximations numériques sont alors proposées. Ces approximations entraînent d'une part de forts temps de calcul et d'autre part il est nécessaire de bien mesurer les effets des erreurs numériques sur les conclusions. Pour contrôler ce dernier point, la démarche suivie est une validation des résultats numériques par le biais de simulations stochastiques.

2.4 Optimisation du critère

Une fois la procédure d'évaluation de la fonction de renouvellement donnée, on doit chercher à déterminer les valeurs des paramètres de décision optimales, ces valeurs minimisant le critère de coût moyen par unité de temps à long terme. Les problèmes d'optimisation rencontrés ici mixent nombres entiers (nombre de seuils de décision) et variables réelles (valeurs des seuils de décision). La procédure utilisée est une procédure itérative qui cherche à déterminer pour chaque nombre n_i de seuils leurs valeurs respectives par le biais d'un algorithme de type quasi-Newton. Le critère d'arrêt de l'algorithme itératif est dès que la différence des coûts optimisés obtenus respectivement pour n_i et n_{i+1} devient négative.

D'une manière générale, si on note $\xi_{n_1}^*$ et $\xi_{n_2}^*$ les valeurs optimisées des seuils de remplacement préventif (dans le cas mono-composant) respectivement aux nombres n_1 et n_2 de seuils ($n_1 < n_2$), alors on a $\xi_{n_1}^* \geq \xi_{n_2}^*$. Ceci permet notamment de réduire le champ d'exploration de l'espace des solutions, d'autant que les courbes de coût moyen sont généralement très plates au niveau des optimums.

3 Des approches multi-seuils pour un système à un composant

L'objectif de cette section est de présenter brièvement le schéma décisionnel à seuils multiples et ses extensions ainsi que les conclusions majeures.

3.1 Un schéma conditionnel d'inspection/remplacement

Nous rappelons que l'objectif de nos travaux s'oriente sur la définition des dates d'inspection et des maintenances à effectuer en fonction du niveau de dégradation observé. Nous proposons ici de traiter le cas *simple* d'un système mono-composant avec des maintenances parfaites (inspection et remplacement). Les états de dégradation successifs à temps discret

forment un processus stochastique $\{X_k, k \in \mathbb{N}\}$ dont la loi d'évolution sur la grille de temps $\{t_k, k \in \mathbb{N}\}$ est donnée par une k -Erlang de paramètre α fixé. On rappelle alors qu'un incrément de dégradation sur une période de temps de longueur fixée $t_{k+1} - t_k$ est une variable aléatoire exponentielle de paramètre α .

La structure proposée est composée de n seuils, définissant ainsi $n + 1$ zones de fonctionnement. La règle de décision se décompose comme suit :

- Le système est inspecté et son niveau de dégradation mesuré se situe dans la zone de décision définie par l'intervalle $[\xi_r, \xi_{r+1}[$ avec $r \in \{0, \dots, n-1\}$. Le système est considéré comme peu dégradé et laissé en l'état. La prochaine inspection est programmée $n - r$ périodes de temps plus tard.
- La zone définie par $[\xi_n, L[$ où L est le seuil donné de défaillance représente la zone de remplacement préventif et la zone $[L, \infty)$ la zone de défaillance. Un système inspecté dans cette zone sera alors immédiatement remplacé. À l'issue de ce remplacement, le système est considéré comme neuf et son état de dégradation nul. La prochaine inspection est alors programmée n périodes de temps plus tard.

Ce schéma décisionnel est représenté dans la Figure 2.1 dans laquelle l'intervalle de programmation de la nouvelle inspection après éventuel remplacement est donné en abscisses et les valeurs de dégradation observées lors de l'inspection courante en ordonnées.

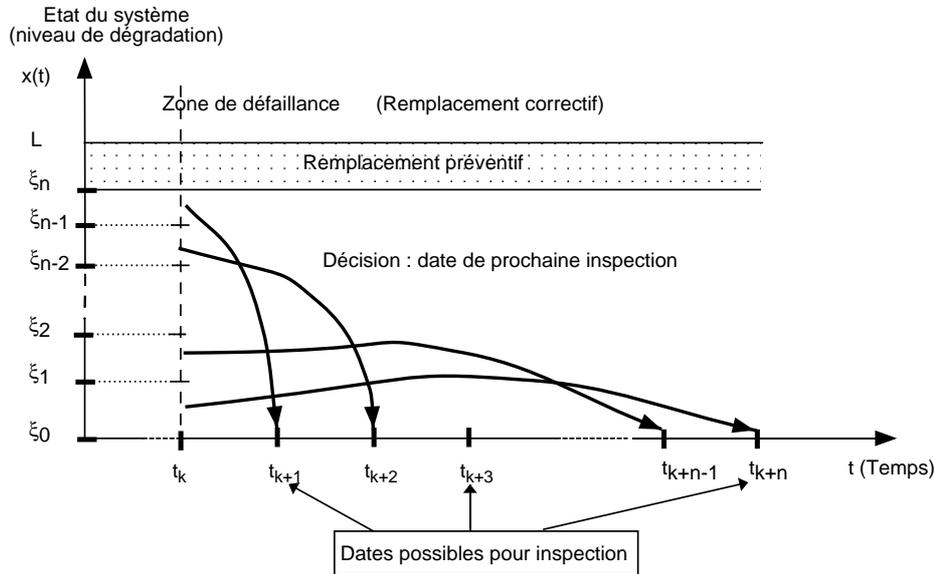


FIGURE 2.1 – Représentation de la structure décisionnelle multi-seuils pour un système mono-composant - cas : maintenance parfaite

Comme énoncé précédemment, le processus de dégradation du système maintenu est un processus de renouvellement. On définit un cycle de renouvellement comme l'intervalle de temps allant de la remise du système à neuf $X_0 = 0$ jusqu'à la date T de prochain remplacement, $X_T \geq x_n$. Le coût de maintenance moyen à long terme par unité de temps $C_\infty(\Xi)$ avec Ξ l'ensemble des seuils de décision $\xi_r, r \in \{0, \dots, n\}$ peut s'écrire :

$$C_\infty(\Xi) = \lim_{t \rightarrow \infty} \frac{C(t; \Xi)}{t} = \frac{\mathbb{E}(C(T; \Xi))}{\mathbb{E}(T)} \quad (3.1)$$

Sans chercher à en donner les expressions complètes [15], on peut dire en premier lieu que toutes les espérances mathématiques s'écrivent en fonction de la loi stationnaire définie sur un cycle de renouvellement $[0, T]$, et deuxièmement que le coût moyen de maintenance sur un cycle de renouvellement s'écrit comme la somme pondérée du nombre moyen d'inspections réalisées dans un cycle, de la probabilité de défaillance en fin de cycle et du temps moyen de

panne $\mathbb{E}(D_u)$, soit :

$$\mathbb{E}(C(T; \Xi)) = C_i \mathbb{E}(N_i(T)) + C_p (1 - Pr(X_T \geq L)) + C_c Pr(X_T \geq L) + C_u \mathbb{E}(D_u) \quad (3.2)$$

Nous avons conduit nombre d'expérimentations numériques pour mettre en évidence les diverses contributions des données sur les politiques optimisées. En voici quelques conclusions.

Influence des vitesses de dégradation : Outre le résultat évident d'un coût de maintenance réduit pour des systèmes à dégradation lente, nous avons mesuré l'importance des seuils de décision (en nombre et en valeur) pour des systèmes à dégradation rapide ou lente. Ainsi, nous avons pu mettre en évidence l'intérêt d'une surveillance continue pour des systèmes à dégradation rapide (sous condition de coût d'inspection modéré) et la réduction du nombre moyen d'inspections par cycle pour des dégradations lentes, l'accumulation des coûts d'inspection amenant alors une contribution non négligeable dans le critère.

Par ailleurs, nous avons cherché à mesurer la robustesse de la politique multi-seuils aux erreurs d'estimation des données de dégradation. Les résultats obtenus, illustrés dans la Figure 2.2, montrent une certaine robustesse de la politique notamment lors d'une sur-estimation de la vitesse de dégradation (qui revient à une sous-estimation du paramètre α). En effet, une sous-estimation de cette vitesse peut entraîner l'introduction d'un seuil de décision supplémentaire et ainsi augmenter le risque de défaillance avant la première inspection.

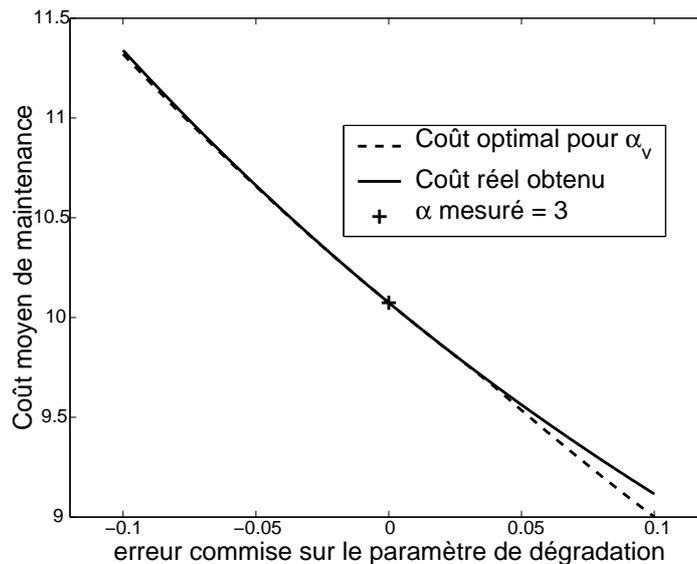


FIGURE 2.2 – Étude de robustesse de la politique multi-seuils aux erreurs d'estimation de dégradation

Généralisation des politiques classiques : La politique conditionnelle multi-seuils généralise des politiques de remplacement classiques basées sur l'âge (purement correctives ou bien remplacement préventif systématique), les politiques d'inspection continue ou périodique et remplacement conditionnel. La Figure 2.3 illustre l'évolution de la politique optimale en fonction des coûts unitaires d'inspection. On retrouve, suivant les configurations de coût, numériquement les conclusions énoncées ci-dessus. Par exemple, lorsque les coûts d'inspection sont élevés, l'ensemble des seuils de décision s'aggrègent et tendant vers 0, la politique optimale tend vers une politique de remplacement systématique (remplacement dès la première inspection).

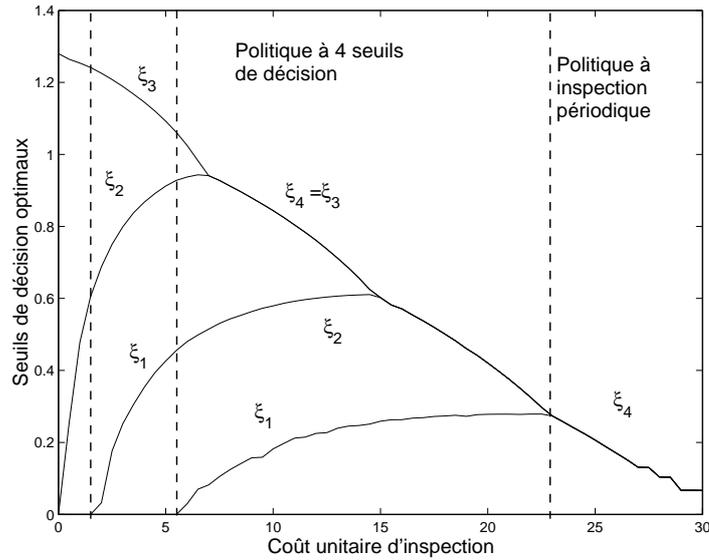


FIGURE 2.3 – Études du comportement de la politique optimisée en fonction du coût unitaire d'inspection

Amélioration des politiques classiques : Nous avons étendu le modèle multi-seuils pour pouvoir prendre en compte des contraintes de *maintenance réglementaire* de type maintenance systématique ou par bloc. Cette étude a notamment été motivée pour étudier les bénéfices attendus de l'intégration d'une politique multi-seuils dans les schémas classiques de maintenance industriel (plutôt orientés sur le temps que vers des approches conditionnelles). Nous avons mis en évidence les contributions positives ou négatives des maintenances réglementaires par l'introduction d'un coût de pénalité sur ce type d'opération. Un coût de pénalité négatif permet de mettre en évidence l'intérêt de renforcer la maintenance dans les schémas classiques par une approche conditionnelle alors qu'un coût positif augmente les effets *contraintes réglementaires*, la politique multi-seuils optimisée cherchant alors à anticiper les actions réglementaires basées sur l'âge.

Alors que la construction du modèle étendu pour intégrer une maintenance réglementaire systématique basée sur l'âge peut reposer sur les résultats du renouvellement classique, la difficulté pour les approches de type bloc provient de la caractérisation des instants réglementaires (la périodicité de ces événements étant indépendantes des décisions sur le système).

En conclusion : La politique multi-seuils offre un schéma purement conditionnel tant sur la définition des instants d'intervention que sur la nature même de l'intervention à réaliser. Les règles sur lesquelles elle se repose rejoignent les idées du praticien à savoir accroître les visites d'un système vieillissant pour ajuster au mieux la décision de remplacement. De plus, la formulation mathématique du problème et son optimisation permettent d'en ajuster cette décision en fonction des paramètres de dégradation du système mais aussi de son contexte économique, ces ajustements étant dans la pratique réalisés de manière trop empirique. Enfin, le formalisme mathématique issu de l'utilisation de la théorie du renouvellement permet d'envisager une complexification des environnements d'exploitation du système. C'est ce que nous allons nous attacher de démontrer dans les paragraphes suivants.

3.2 Prise en compte de l'efficacité et des durées de maintenance

Les premiers points d'intérêt rejoignent des hypothèses plus réalistes sur la caractérisation des actions de maintenance. Les hypothèses classiques posent que l'effet des maintenances est parfait dans le sens où après un remplacement le système peut être considéré comme

neuf et que toute durée d'action est considérée comme négligeable au regard des temps de fonctionnement.

L'approche que nous allons décrire se décompose en deux temps.

Maintenance préventive imparfaite : Le premier temps consiste à lever l'hypothèse de remplacement aussi bon que neuf et de considérer que l'efficacité de maintenance préventive définie par la différence entre le niveau observé avant et après maintenance est une variable aléatoire. Nous avons conduit l'ensemble des calculs du critère de décision en considérant une loi uniforme pour cette quantité aléatoire, supposant ainsi aucune connaissance sur l'aptitude à réaliser des améliorations sur le système actuel. La construction de la loi stationnaire repose sur l'identification et la caractérisation des scénarios de maintenance ainsi que l'évolution de la dégradation résultante sur un cycle de renouvellement markovien défini depuis l'état avant la décision de maintenance jusqu'à la date de la prochaine intervention. Il est alors nécessaire de considérer les 3 étapes suivantes :

1. l'identification du niveau de dégradation y dans le cycle fournie par la loi d'entrée $\pi(y)$;
2. le niveau de dégradation z à l'issue de la maintenance : $z = y$ si simple inspection, $z = \Delta(0, \xi_n)$ où $\Delta(0, \xi_n)$ est une variable aléatoire uniformément distribuée sur l'intervalle $(0, \xi_n)$ et $z = 0$ si $y \geq L$ (remplacement correctif) ;
3. la loi d'évolution du niveau z vers x qui est donnée directement par la $(n - k)$ -ième convoluée de f , f étant la loi de dégradation sur un intervalle de temps unitaire et $n - k$ le nombre d'intervalles de temps avant la prochaine intervention fourni par la règle de décision, $z \in [\xi_k, \xi_{k+1})$, $k \in \{0, \dots, n - 1\}$.

La loi stationnaire de dégradation du système maintenu est alors solution de l'équation de renouvellement suivante :

$$\begin{aligned} \pi(x) = & \sum_{k=0}^{n-1} \int_{\xi_k}^{\xi_{k+1}} \pi(y) f^{(n-k)}(x - y) dy \\ & + \frac{1}{\xi_n} \int_{\xi_n}^L \pi(y) \sum_{k=0}^{n-1} \int_{\xi_k}^{\xi_{k+1}} f^{(n-k)}(x - z) dz dy + \int_L^{\infty} \pi(y) dy f^{(n)}(x) \end{aligned} \quad (3.3)$$

Jusqu'à présent, la résolution de l'équation de renouvellement reposait sur un schéma récursif qui ne peut être mis en place ici. Nous avons proposé un schéma itératif de type point fixe pour approximer la loi stationnaire. Ce dernier schéma est inspiré d'une méthode numérique pour l'approximation des équations différentielles de Fredholm de deuxième espèce [45]. Les démonstrations de convergence de l'algorithme n'ont pas été établies mais de nombreux essais numériques indiquent que l'algorithme semble converger très rapidement (moins de 5 itérations) quelle que soit la loi de départ.

Les comparaisons numériques proposées des politiques de maintenance montrent l'obtention de performance optimiste dans le cas d'un remplacement parfait au regard de la maintenance imparfaite. Ceci est grandement lié au choix de la loi uniforme pour l'état de remise en service ainsi que la forme affine des coûts de réparation associés qui ne permettent pas, par exemple, de privilégier de petites réparations pour maintenir régulièrement le système dans des états peu dégradés.

Optimisation de l'efficacité de maintenance et de disponibilité : Nous avons choisi d'étudier ces deux aspects simultanément car ils nous semblent offrir une réponse, du moins partielle, à la problématique de la recherche de l'efficacité opérationnelle alliant dans un même temps respect des exigences de fonctionnement et minimisation des temps d'inactivité des systèmes de production en phase de maintenance.

Dans cette partie, nous avons considéré que l'efficacité de maintenance pouvait être contrôlée et qu'à l'issue d'une telle réparation, l'état du système est une variable aléatoire à support

$[0, \zeta)$ où ζ est une variable de décision. Nous avons à nouveau considéré la possibilité de remplacement préventif, conduisant ainsi à introduire un nouveau seuil de décision ξ_{n+1} . Par ailleurs, nous avons supposé que la durée de maintenance est une variable aléatoire dépendant du niveau d'efficacité de cette dernière. Nous avons supposé que cette durée est une variable poissonnienne de paramètre $\lambda(\Delta_x)$ où Δ_x est le niveau de réparation effectué durant l'opération sachant x le niveau de dégradation en début de maintenance. La Figure 2.4 illustre la structure de décision et schématise l'évolution de la dégradation d'un système soumis à une telle politique.

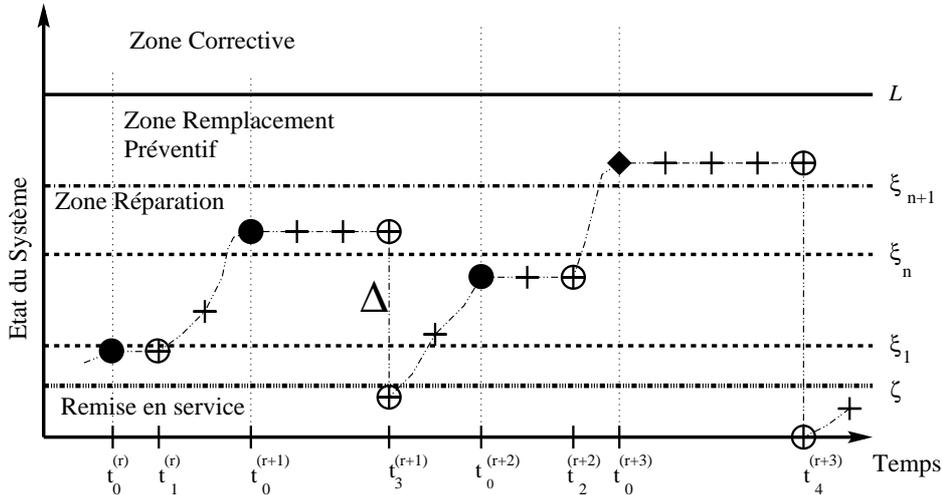


FIGURE 2.4 – Schématisation de l'évolution de dégradation d'un système soumis à la politique de réparation préventive imparfaite sous hypothèse d'indisponibilité de maintenance

La prise en compte des durées d'intervention augmente la difficulté d'évaluation de la loi stationnaire. Pour contourner cette difficulté, nous avons posé l'hypothèse de non dégradation durant la durée de la maintenance, hypothèse réaliste au vu du non fonctionnement du système. Associé à ceci, exprimer la durée de maintenance en fonction de son efficacité et non l'inverse permet d'assurer que la loi d'évolution de la dégradation dans un cycle semi-régénératif est uniquement fonction du niveau de dégradation en entrée du cycle et donc indépendant de la durée de maintenance. On peut ainsi, par analogie avec la méthode itérative présentée dans le paragraphe précédent, approcher la loi stationnaire $\pi(x)$. Une fois $\pi(x)$ obtenue, il est possible de caractériser la loi stationnaire des durées de maintenance, $P^i(k, x)$ définie comme la probabilité qu'une intervention dure au moins k unités de temps pour un niveau de dégradation initial x .

L'approche présentée ici permet d'exprimer le critère de décision de la manière suivante :

$$C_\infty(\Xi) = \lim_{t \rightarrow \infty} \frac{C(t)}{t} = \frac{\mathbb{E}(C_1(T_f)) + \mathbb{E}(C_2(T_i))}{MTBR} \quad (3.4)$$

avec le temps moyen entre réparation défini comme la somme des temps moyen dit de de fonctionnement, T_f intégrant les éventuels temps d'insdisponibilité de panne et les temps moyens d'intervention T_i , i.e. $MTBR = \mathbb{E}(T_f) + \mathbb{E}(T_i)$. $C_1(T_f)$ représente les coûts classiquement évalués dans les sections précédentes et $C_2(T_i)$ intègre les coûts d'indisponibilité liés aux interventions.

Les essais numériques que nous avons conduits ont souligné, sous certaines configurations de coûts et de données de dégradation, l'intérêt d'une telle politique. Toutefois, nombre de cas ont montré que la politique optimisée revenait à une des politiques limites présentées ci-dessus (politique de remplacement préventif ou politique de réparation préventive imparfaite). Ceci est, de notre point de vue, lié à plusieurs facteurs dont le choix des formes linéaires des fonctions coût d'entretien et durée de maintenance.

Par ailleurs, il peut être d'intérêt de contrôler le niveau minimal d'entretien, ζ_1 (ζ_2 étant le seuil supérieur). Est-il, en effet, important de s'imposer une opportunité de remise à neuf lors d'une réparation? Cette addition du seuil ζ_1 pourrait aussi permettre d'étudier les effets de réduction de la variabilité de l'efficacité d'une réparation en limitant la zone de remise en service $[\zeta_1, \zeta_2]$.

Une autre extension envisageable est la modélisation de l'efficacité d'une réparation en fonction de la durée de l'intervention. Une telle approche permettrait de rejoindre d'éventuelles préoccupations industrielles qui autorisent l'exercice de la maintenance dans des périodes de temps très limitées. Je suppose toutefois à ce niveau quelques complications dans la caractérisation de la loi de dégradation puisque maintenant fonction de la durée de maintenance et donc perte de l'avantage souligné ci-dessus.

Avant de clore ce paragraphe, je voudrai souligner que même si la modélisation développée ici reste très empirique et certaines hypothèses difficiles à valider dans un environnement opérationnel, la modélisation de la maintenance imparfaite dans un contexte conditionnel me semble être plus facile à justifier que les approches basées sur des concepts non mesurables tels que le taux de défaillance, notamment au vu des difficultés de l'ajustement du facteur de réduction d'âge, par exemple. En effet, la différence entre mesures amont et aval de l'intervention permet d'estimer directement l'efficacité de l'intervention sur les niveaux de dégradation. La maintenance imparfaite reste de mon point de vue un des sujets les plus prometteurs en termes de modélisations et d'applications. En effet, mesurer l'impact au niveau des lois de dégradation (ici, seul l'impact sur le niveau de dégradation est étudié) et donc de leur comportement futur devrait permettre la proposition de politiques plus adaptées. Une hypothèse de maintenance parfaite offre une vision trop optimiste de la situation pouvant entraîner un réel gap entre prescrit et observé.

3.3 Prise en compte de l'influence des conditions d'environnement sur le processus de dégradation

La prise en compte des effets de l'environnement sur les performances de dégradation est un champ bien investi dans le cadre de la fiabilité sous le vocable de *Tests de durées de vie accélérées*. L'un des objectifs majeurs est la caractérisation des quantités fiabilistes de systèmes à longue durée de vie, la mise sous stress de tels systèmes peut permettre l'observation de défaillances. Les connaissances ont été transférées au contexte industriel et intégrées dans des normes internationales [1, 2]. Dans le contexte de la maintenance et particulièrement la maintenance conditionnelle, peu d'études sont encore disponibles sur ce champ. La non prise en compte des effets environnementaux rejoint l'une des hypothèses suivantes :

- l'environnement n'a pas ou peu d'influence sur les caractéristiques de dégradation ;
- les fluctuations de l'environnement ne sont pas mesurées et aucune connaissance sur ces fluctuations n'est disponible ;
- les fluctuations sont très rapides au vu des périodes de décision et une approche *moyenne* peut alors être préconisée.

Quand une de ces conditions n'est pas vérifiée, et ceci représente bien des cas, la maintenance peut être améliorée par la prise en compte de cette nouvelle information. Cette conclusion bien qu'évidente pose le problème du comment intégrer la connaissance de ces aspects environnementaux dans un modèle de maintenance. Nous allons nous attacher dans ce paragraphe à proposer différentes approches. Avant cela, une classification des effets des conditions environnementales sur les performances du système est discutée.

Classification des effets de l'environnement sur le système : En tout premier lieu, nous rappelons que l'étendue de notre étude se restreint aux systèmes à dégradation continue et que le mode de défaillance est liée à une variable de dégradation $X(t)$ et un seuil de défaillance L . L'influence de l'environnement peut alors être mesurée de différentes ma-

nières suivant la nature et les modes des dommages occasionnés : permanents ou temporaires, ponctuels ou continus.

- **L'impact ponctuel** est défini lorsque l'influence de l'environnement sur le mode de dégradation est assimilée à un choc. Cette représentation est particulièrement adaptée à des systèmes soumis à des phénomènes vibratoires pour lesquels l'amplitude vibratoire peut affecter leur performance. Dans ce contexte, on peut distinguer différents types d'impact :
 - L'amplitude du choc est telle qu'elle peut causer une cessation immédiate de la fonction. La résistance au choc peut être dépendante du niveau de dégradation du système, ceci permet par ailleurs de modéliser les effets répétés de tels chocs. Peu de travaux dans un contexte d'optimisation de la maintenance ont été proposés sur ce champ [110]. On peut assimiler cette modélisation à un modèle de type risques compétitifs. Ce cas sera présenté dans le cadre de l'étude des systèmes multi-composants présentant des dépendances stochastique, sous-section 4.3.
 - L'amplitude du choc n'est pas suffisante pour causer une défaillance immédiate mais peut entraîner des dommages irrémédiables sur la loi d'évolution de la dégradation du système. On peut par exemple considérer qu'à l'issue d'un tel choc, la vitesse moyenne de dégradation est augmentée de manière irréversible. L'accélération peut être fonction de l'intensité du choc et aussi du niveau de dégradation du système.
 - L'amplitude du choc n'est pas suffisante pour des dommages irréversibles mais crée un effet ponctuel. La loi de dégradation n'est pas changée mais un saut de dégradation à forte amplitude est observable. Par exemple, ce saut peut être expliqué dans le cas d'une structure métallique soumise à érosion ou corrosion à une perte de matière.
- **L'impact continu** est à considérer lorsque la fréquence des changements de condition reste faible et qu'ainsi les cycles d'exposition aux différentes conditions ne sont pas négligeables devant l'horizon d'exploitation. Par exemple, on peut retenir les cycles climatiques et leurs impacts sur une structure béton. La discussion sur les effets permanents ou temporaires reste pertinente dans ce contexte. Pour chacun des effets, il peut par contre être intéressant de rajouter à l'intensité de la contrainte le temps d'exposition du système à cette même contrainte qui sera pour nous directement lié aux temps de séjour dans chacun des niveaux de sollicitation.

Dans la suite de ce paragraphe, nous présentons diverses stratégies de maintenance intégrant l'information sur l'évolution des conditions environnementales suivant la nature ponctuelle ou continue de ses effets.

Stratégies d'actualisation pour des contraintes ponctuelles : Afin de modéliser les effets ponctuels de l'environnement sur la dégradation, nous avons considéré un processus Gaussien à temps discret $(Y_k)_{k \in \mathbb{N}}$ représentant son intensité ponctuelle. Un seuil de stress λ permet de définir le niveau de sollicitation. Un niveau de sollicitation faible, $Y_k < \lambda$, n'entraîne aucune incidence sur la dégradation supposée donnée par un processus gamma de loi $g(\cdot, \alpha \delta t, \beta)$. Lorsque $Y_k \geq \lambda$, l'incrément de dégradation est alors donné par une loi gamma dont le paramètre de forme est augmenté d'un facteur $\delta \alpha$ (le paramètre d'échelle β reste constant). Dans le cas d'un effet non permanent, la loi de dégradation revient à son régime nominal dès que $Y_{k_1} < \lambda$ avec $k_1 > k$, à savoir la loi gamma avec un paramètre de forme α . Pour un effet permanent, le paramètre de forme est augmenté de manière irréversible à chaque nouveau stress, $Y_k \geq \lambda$. Ainsi, après n chocs, la loi de dégradation est la loi gamma de paramètres $((\alpha + n\delta\alpha)\delta t, \beta)$. Cette modélisation est à rapprocher de celle présentée dans [35, 94]. Enfin, lorsque le système est défaillant, on considère qu'il continue de se dégrader jusqu'à la nouvelle intervention. Par contre, on considère un changement de mode au niveau de l'environnement et une dérive est introduite au niveau du processus Gaussien. L'amplitude est alors un processus Brownien. La Figure 2.5 propose une illustration de chacun des effets.

Nous avons montré le bénéfice d'adapter successivement la période d'inspection en fonction

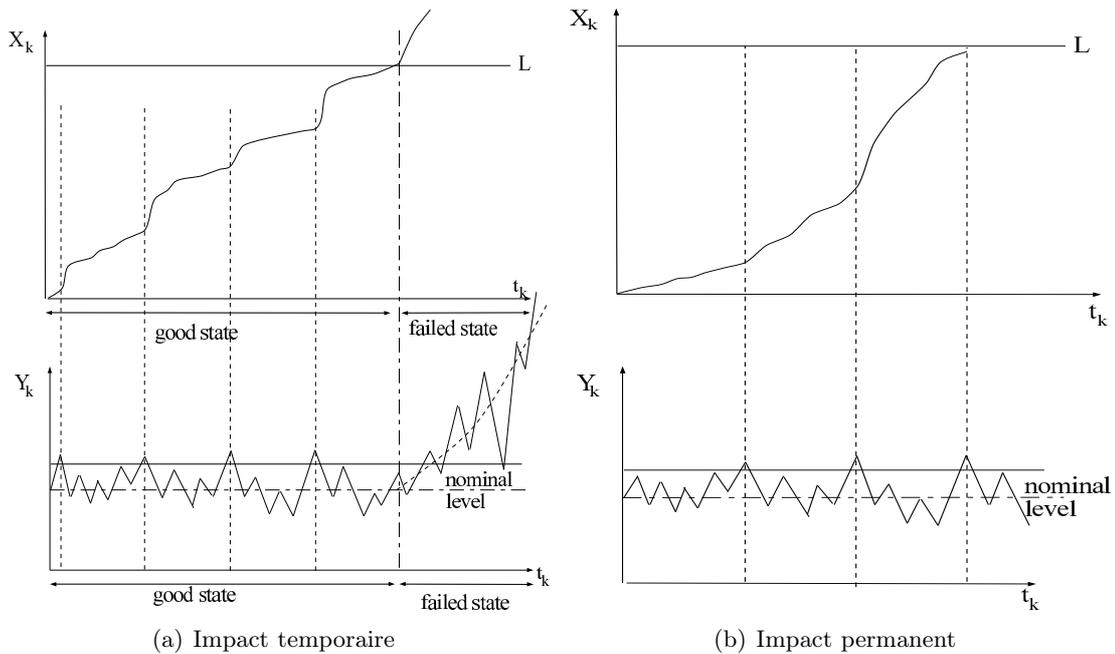


FIGURE 2.5 – Illustration de l'influence du processus environnemental sur la dégradation - Effets ponctuel et permanent

du nombre de contraintes reçues par le système. Pour se faire, nous avons considéré une politique de remplacement préventif de type seuil noté ξ pour laquelle la mesure de dégradation, appelée les X -inspections, est réalisée périodiquement. L'acquisition des Y_k est immédiate et ne nécessite aucune intervention. Nous avons intégré les règles de réduction de l'intervalle d'inspection suivantes pour deux politiques :

- Une inspection est réalisée à une date t_k et $X_k < \xi$. La future inspection est planifiée n_τ intervalles de temps plus tard ;
- Dès qu'un stress est mesuré, $Y_s \geq \lambda$ avec $s \in \{k, \dots, k + n_\tau\}$, soit seule la date de la prochaine inspection (le schéma périodique initial est repris dès la prochaine inspection) est avancée d'une unité de temps, soit c'est le schéma périodique jusqu'au prochain remplacement qui est réduit d'une unité.
- La règle de remplacement préventif reste inchangée.

Les résultats numériques ont clairement montré le bénéfice de telles approches avec notamment la possibilité d'allonger les périodes d'inspection en début de cycle, lorsque le système est neuf. Ce dernier point est d'autant plus vrai dans le cas de la seconde politique préconisant les changements définitifs de période (jusqu'au prochain remplacement). On peut aussi noter que cette politique tend donc à proposer des schémas non périodiques d'inspection, rejoignant en quelque sorte les politiques multi-seuils présentées précédemment.

Il est clair que les propositions que nous avons faites ici ne sont que prospectives et qu'il est nécessaire de chercher à améliorer les politiques, notamment par la détermination d'un critère de changement de politiques d'inspection qui pourrait être, par exemple, fonction de l'intensité cumulée des stresses (notre choix de réduire d'une unité de temps pour chaque stress reste empirique). Nous tenons cependant à souligner les difficultés numériques de résolution du critère de décision qui nous ont poussés à utiliser des approches de simulation de type Monte Carlo pour évaluer les différentes solutions.

Politiques de maintenance pour des effets continus : Nous avons modélisé le processus de contrainte comme une chaîne de Markov à temps continu, $(Y_t)_{t \geq 0}$. Ces contraintes s'appliquent sur un système à dégradation continue que nous modélisons ici par un processus gamma stationnaire à *temps continu*. L'effet de la contrainte excessive, $Y(t) = 1$, se traduit

directement sur la vitesse de dégradation que nous modélisons comme un facteur multiplicatif, e^γ , sur le paramètre de forme de la loi gamma [6, 55]. Dans ce cadre, la fonction de forme de la loi gamma sur un intervalle de longueur t s'écrit $\alpha(t, Y(t)) = \alpha_0 e^{\gamma Y(t)}$ où α_0 est la valeur du paramètre de forme sans stress, $Y(t) = 0$. Une représentation du processus d'évolution de la dégradation soumise au stress est proposée Figure 2.6.

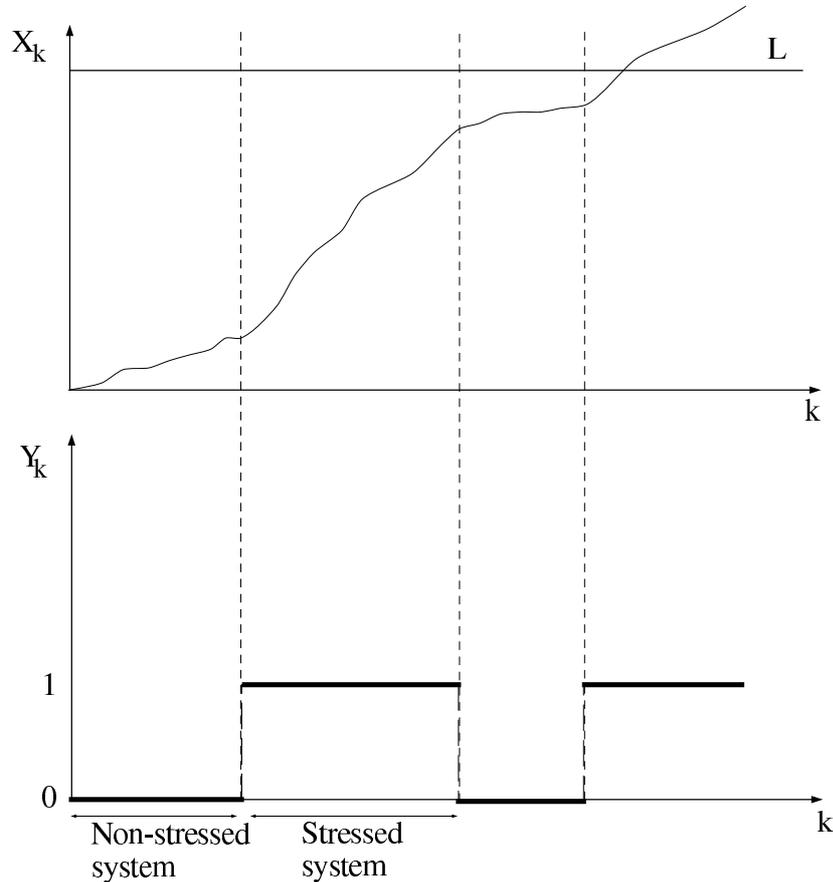


FIGURE 2.6 – Evolution de la dégradation en fonction d'une contrainte continue

Dans ce contexte, nous avons proposé 4 nouvelles politiques de maintenance qui vont d'une politique purement stationnaire pour laquelle les paramètres de décision (période d'inspection et seuil de remplacement préventif) sont optimisés sur les caractéristiques moyennes de dégradation jusqu'à une politique dynamique pour laquelle la règle de décision est actualisée continuellement en fonction de la proportion de temps passé dans l'état stressé. Les deux politiques *intermédiaires* permettent de limiter le nombre de changements de règle de décision.

Nous avons conduit plusieurs expérimentations numériques pour mettre en évidence les bénéfices liés à l'intégration de la connaissance du processus d'évolution des contraintes environnementales. Les résultats confirment l'avantage d'adapter la règle de décision en fonction du ratio de temps passé dans chacun des états, surtout si on ne considère aucune pénalité liée à un éventuel changement d'organisation. La Figure 2.7 illustre les bénéfices de la flexibilité de la décision. La politique 0 représente la politique stationnaire, la politique 1 est la politique actualisée continuellement alors que la politique 2 peut être vue comme une forme discrétisée de la politique 1 avec des changements de stratégie uniquement lors de passage de seuils. Ainsi, la politique 2 tend vers la politique 1 qui minimise le coût moyen de maintenance quand le nombre de seuils augmentent.

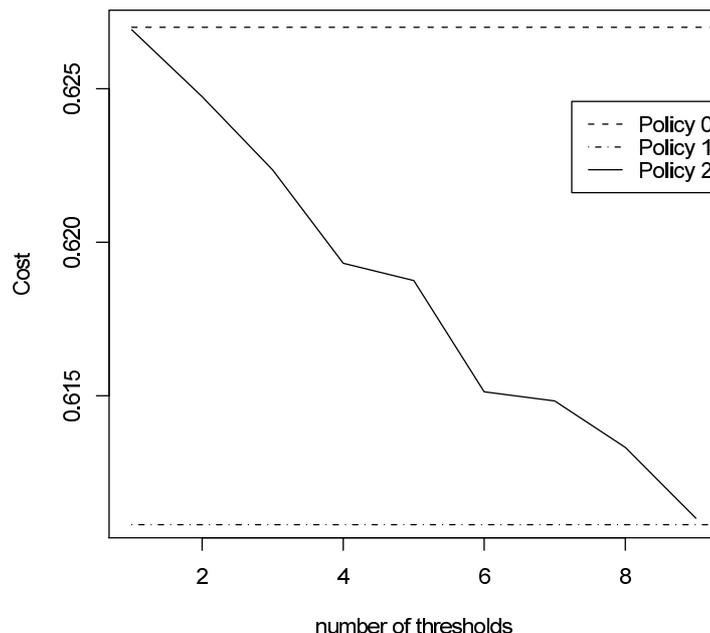


FIGURE 2.7 – Mise en évidence des bénéfices liés à la flexibilité des politiques pour des contraintes continues

4 Politiques conditionnelles pour des systèmes à plusieurs modes de dégradation

Jusqu'à présent, nous nous sommes concentrés sur des politiques dites mono-composant. L'utilisation de tels modèles se doit d'être restreinte à des systèmes à unique mode de défaillance ou du moins lorsque les effets de cette dégradation sont critiques devant les autres modes de dégradation. Dans ce contexte mono-composant, nous avons par ailleurs convenu que l'indicateur de dégradation pouvait être une mesure agrégée, synthèse de plusieurs modes de dégradation. Cependant, une telle mesure n'est pas toujours pertinente, notamment si les modes de défaillance sont indépendants.

Nous proposons dans cette section d'aborder la problématique de la maintenance de systèmes présentant plusieurs modes de défaillance, systèmes pouvant être considérés comme formés de plusieurs composants indépendants ou encore d'un seul composant présentant plusieurs modes de défaillance dépendants.

Après avoir présenté brièvement les différents types de dépendance pour des systèmes multi-composants, je présenterai comment nous avons cherché à étendre le modèle multi-seuils pour exploiter le bénéfice de maintenance jointe, ces travaux constituant la dernière partie de ma thèse. Ensuite, je présenterai un modèle de maintenance pour un système soumis à deux modes de défaillance dépendants, le premier étant un mode de type dégradation cumulée alors que le second est du type choc.

4.1 Les dépendances dans le cadre multicomposants

Différents types d'interactions entre les composants sont à distinguer, [101] : les dépendances économiques, structurelles et stochastiques.

- Une dépendance stochastique entre des composants peut intervenir lorsque le comportement d'un composant du système influence les états des autres composants, ou bien lorsque tous les composants sont soumis aux mêmes contraintes environnementales ou à des conditions de dégradation identiques.
- Une dépendance structurelle existe dans le cas où les composants sont physiquement reliés entre eux, par exemple un système formé par une arborescence d'entités montées

en série ou en parallèle. La maintenance d'un composant peut affecter les comportements des entités voisines.

- Une dépendance économique existe lorsqu'une maintenance sur tout le système permet de réduire un coût d'opération par rapport à une maintenance effectuée sur chaque entité séparément.

Dans cette section, seules les dépendances stochastiques et économiques seront abordées au travers de deux problèmes. Nous proposons en premier lieu d'étendre la structure multi-seuils pour un système formé de deux composants indépendants (les effets d'un composant sur l'autre sont négligeables). La dépendance économique est modélisée par l'intégration de coûts d'opération mutualisés (ces coûts peuvent par exemple représenter les coûts de montage et démontage du système). Il est alors intéressant de définir une nouvelle structure décisionnelle permettant d'éventuels regroupements d'opération pour les 2 composants, une simple juxtaposition de modèles mono-composant. La politique proposée fait partie de la classe des politiques *stationnaires* et plus particulièrement à la classe des politiques *opportunistes* [32] pour lesquelles une intervention sur une partie du système crée une opportunité de maintenance sur les autres parties. Lors d'une opportunité, il n'est pas obligatoire d'intervenir sur les autres entités mais si tel est le cas, il est nécessaire de quantifier le gain de l'option d'anticipation de cette intervention. Wildeman *et al.* [115, 116] proposent la construction d'une fonction de pénalité permettant d'évaluer les pertes d'un décalage d'une date pour la maintenance d'un composant par rapport à la date optimisée dans le cas mono-composant. Cependant, les modèles présentés reposent sur des maintenances de remplacement pour lesquelles l'état du système est résumé par son fonctionnement en $(0,1)$ et le nombre de composants défectueux. Cette modélisation classique de la maintenance [9, 39, 86] ne permet pas l'utilisation des informations sur le niveau de dégradation pour construire les règles de décision. D'autres modélisations existent. Sans chercher à proposer un état de l'art exhaustif en matière de regroupement d'opérations [32, 29, 30, 107], on peut toutefois souligner que peu de ces travaux abordent la maintenance conditionnelle et qu'ils se restreignent à des systèmes à espace d'états de dégradation discret [7, 108, 114].

Les modélisations des dépendances stochastiques sont plus délicates dans le sens où, d'une part, leur calibration nécessite un grand nombre de données et, d'autre part, une part intuitive liée à une connaissance des mécanismes de défaillance est nécessaire pour expliquer les interactions entre défaillances. Les approches classiques se réfèrent généralement à de la maintenance de systèmes multi-composants pour lesquels des défaillances interagissent. Nicolai et Dekker [84] propose de réduire à deux le nombre de classes de modèles de dépendance stochastique : les interactions de type I et II. Les interactions de type I supposent qu'une défaillance d'un composant entraîne la défaillance des autres composants avec une probabilité donnée alors qu'ils définissent les interactions de type II de la façon suivante : *Le système est formé de plusieurs composants et la défaillance d'un composant affecte soit le taux de défaillance des autres composants soit cause un niveau (aléatoire) de dommage sur l'état d'un ou plusieurs autres composants.* Initialement, Thomas [101] différenciait le dernier type. Les articles identifiés dans cette revue considèrent des modes de défaillances identiques pour chacun des composants. Dans notre cas, nous nous intéressons à un système formé d'un unique composant présentant deux modes de défaillance différents. L'approche de modélisation des dépendances s'inspire des dépendances de type II et rejoint les modélisations proposées dans le cadre des risques compétitifs considérant les probabilités conditionnelles jointes [72].

4.2 Une structure multi-seuils pour un système à 2 composants

Le système d'intérêt est formé par deux composants indépendants dont l'évolution des niveaux respectifs de dégradation est modélisée par une k -Erlang de paramètres respectifs α_i , $i \in \{1, 2\}$. On néglige ainsi les interactions entre ces deux composants. La panne d'un composant n'entraîne pas l'arrêt systématique du système mais ne permet plus le respect

des spécifications (le système est donc défaillant) et l'observation d'une panne nécessite une inspection.

Le schéma décisionnel est défini comme suit :

1. Une structure multi-seuils classique permet de régir la décision de manière individuelle, en fonction de l'état propre de chacun des composants (seuls les remplacements sont ici considérés).
2. Une zone d'anticipation de remplacement préventif est définie. Elle permet d'effectuer un remplacement sur un composant alors que celui-ci n'était pas initialement prévu et qu'un remplacement de l'autre composant est effectué. La zone d'anticipation pour le composant i est définie par (ζ_i, ξ_n^i) .
3. La date de prochaine inspection est donnée par la première date fournie pour chacun des composants.

L'évolution de l'état du système maintenu est ainsi modélisé par un processus bivarié $(X_k)_{k \in \mathbb{N}} = (X_k^{(1)}, X_k^{(2)})_{k \in \mathbb{N}}$. Comme précédemment, ce processus sera analysé au travers de ses propriétés semi-régénératives. Après une inspection complète, l'évolution de chaque composant et ainsi du système global dépend uniquement de l'état courant. Ainsi, associée à l'hypothèse de remplacement parfait, le processus à temps discret (Y_r) restreint aux instants d'intervention est un processus de renouvellement markovien. On peut alors identifier l'ensemble des scénarios de maintenance sur un cycle de semi-régénération et construire la loi de transition associée.

$$\begin{aligned} \pi(x_1, x_2) = & \left(\int_{\xi_n^1}^{\infty} \int_{\zeta_2}^{\infty} \pi(y_1, y_2) dy_2 dy_1 + \int_{\xi_n^2}^{\infty} \int_{\zeta_1}^{\xi_n^1} \pi(y_1, y_2) dy_1 dy_2 \right) f_1^{(n)}(x_1) f_1^{(n)}(x_1) \\ & + \sum_{\substack{i=1 \\ j=3-i}}^2 \sum_{l=0}^{n-1} \left(\int_{\xi_l^i}^{\min(\xi_{l+1}^j, \zeta_j)} \left(\int_{\xi_n^i}^{\infty} \pi(y_1, y_2) dy_i \right) f_j^{(n-l)}(x_j - y_j) dy_j \right) f_i^{(n-l)}(x_i) \\ & + \sum_{k,l=1}^{n-1} \int_{\xi_k^1}^{\xi_{k+1}^1} \int_{\xi_l^2}^{\xi_{l+1}^2} \pi(y_1, y_2) f_1^{(n-\max(k,l))}(x_1 - y_1) f_2^{(n-\max(k,l))}(x_2 - y_2) dy_1 dy_2 \end{aligned} \quad (4.1)$$

La résolution de l'équation différentielle de la seconde espèce (4.1) repose ici aussi sur un schéma numérique itératif :

Étape 0 : Choisir les fonctions arbitraires pour $w_1^{(0)}(y_1)$, $w_2^{(0)}(y_2)$ et $w_{1,2}^{(0)}$ qui représentent respectivement les lois marginales associées au remplacement du composant 2, du composant 1 et la probabilité de remplacement du système entier ;

Étape s : Déterminer l'approximation au rang s en remplaçant les lois marginales dans l'équation (4.1) par $w_1^{(s-1)}(y_1)$, $w_2^{(s-1)}(y_2)$ et $w_{1,2}^{(s-1)}$. On évalue ensuite :

$$\begin{aligned} w_1^{(s)}(y_1) &= \int_{\xi_n^2}^{\infty} \pi^{(s)}(y_1, y_2) dy_2 \\ w_2^{(s)}(y_2) &= \int_{\xi_n^1}^{\infty} \pi^{(s)}(y_1, y_2) dy_1 \\ w_{1,2}^{(s)} &= \int_{\xi_n^1}^{\infty} \int_{\zeta_2}^{\infty} \pi^{(s)}(y_1, y_2) dy_2 dy_1 + \int_{\xi_n^2}^{\infty} \int_{\zeta_1}^{\xi_n^1} \pi^{(s)}(y_1, y_2) dy_1 dy_2 \end{aligned}$$

La Figure 2.8 illustre la forme que peut prendre la densité stationnaire du processus bivarié.

Ici encore, l'analyse exhaustive des scénarios de maintenance sur un cycle de renouvellement markovien permet la construction du critère de coût. Je ne présenterai pas ici les conclusions des analyses numériques, l'objectif étant de montrer le potentiel de notre approche. On peut noter cependant les efforts à consentir pour l'évaluation et l'optimisation du

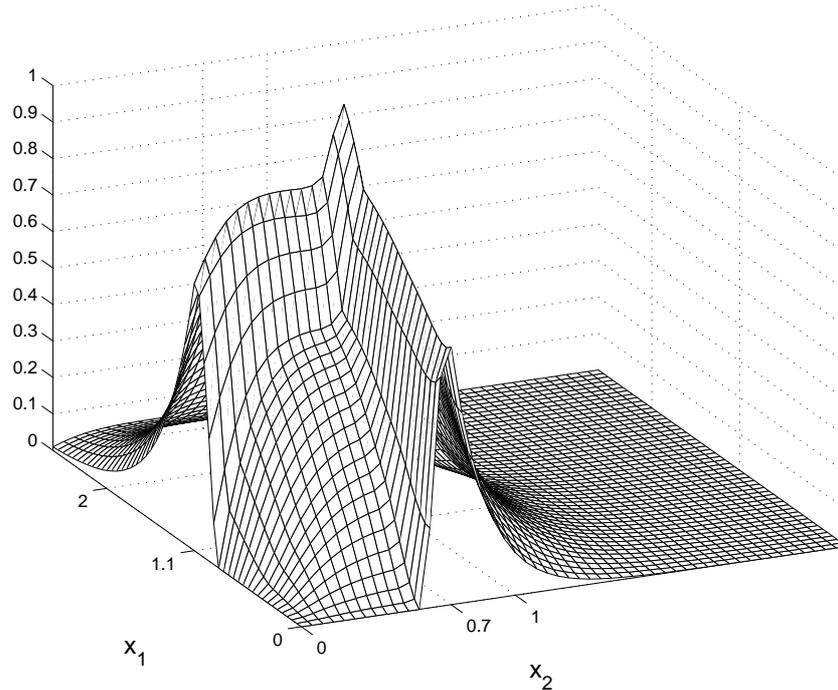


FIGURE 2.8 – Exemple de la forme de la fonction de densité de probabilité stationnaire du système maintenu

critère dans le cas où le système est formé de deux composants indépendants. La généralisation de cette approche pour la prise en compte de nouveaux composants ou encore l'introduction de nouvelles dépendances semble délicate. On peut toutefois retenir l'intérêt de l'introduction du seuil de remplacement anticipé et penser à sa généralisation pour des systèmes complexes.

4.3 Politique conditionnelle pour un système à plusieurs modes de défaillance

Modèle de défaillance : Alors que dans le cas précédent, les défaillances sont supposées indépendantes, nous considérons ici un système soumis à deux modes distincts mais dépendants. Le premier mode de défaillance est un mode cumulatif lié à un excès de dégradation. La modélisation de la dégradation utilisée ici repose sur les hypothèses classiques présentées dans les sections précédentes d'un seuil déterministe de défaillance et d'incréments de dégradation par unité de temps exponentiellement distribués. Le second mode de défaillance est de type choc ou contrainte. Nous avons associé cette défaillance à un processus pouvant être lié à une variable d'environnement externe (type température, humidité, etc.) ou une conséquence directe des modes opératoires du système (vibrations internes, température interne, etc.). Notons $(Y_k)_{k \in \mathbb{N}}$ le processus associé à l'intensité de la contrainte à chaque instant t_k . On supposera que $(Y_k)_{k \in \mathbb{N}}$ forme un processus gaussien de moyenne m et écart-type σ donnés.

La défaillance liée à la contrainte est de type choc, c'est-à-dire que le système peut défaillir avec une probabilité donnée $1 - q$. Il est clair que cette probabilité peut être fonction de l'intensité de la contrainte Y_k . Afin de modéliser cette dépendance, on introduit un seuil γ . Si le niveau de contrainte Y_k reste inférieur à ce seuil γ , le système est considéré dans un état nominal d'opération et la contrainte a un faible impact sur la probabilité de défaillance. Par contre, dès que la contrainte dépasse γ , le système est considéré stressé et induit une augmentation de la probabilité de panne. Cette augmentation est modélisée par l'introduction d'un facteur d'accroissement F_A . Par ailleurs, comme la résistance du système à la contrainte peut varier avec le temps et l'usage [98, 118], on propose de modéliser la probabilité d'une

défaillance liée à la contrainte par une fonction dépendant du niveau courant de dégradation X_k . On pose la probabilité de non défaillance en régime nominal liée à un choc connaissant la valeur de dégradation x après k unités de temps depuis sa mise en service : $q(x, k) = \exp(-ax - bk)$. On note $(Z_k)_{k \in \mathbb{N}}$ le processus formé des niveaux de fonctionnement, 0 pour un fonctionnement et 1 pour la défaillance.

Outre la probabilité de défaillance par choc croissante en l'état de dégradation, la dépendance entre modes de défaillance s'exprime aussi par l'introduction d'une augmentation linéaire en moyenne dans l'intensité de la contrainte lors d'une défaillance. Ainsi, après défaillance, le processus gaussien classique $(Y_k)_{k \in \mathbb{N}}$ évolue en mouvement Brownien avec dérive [93].

Le modèle à risques compétitifs est représenté dans la Figure 2.9.

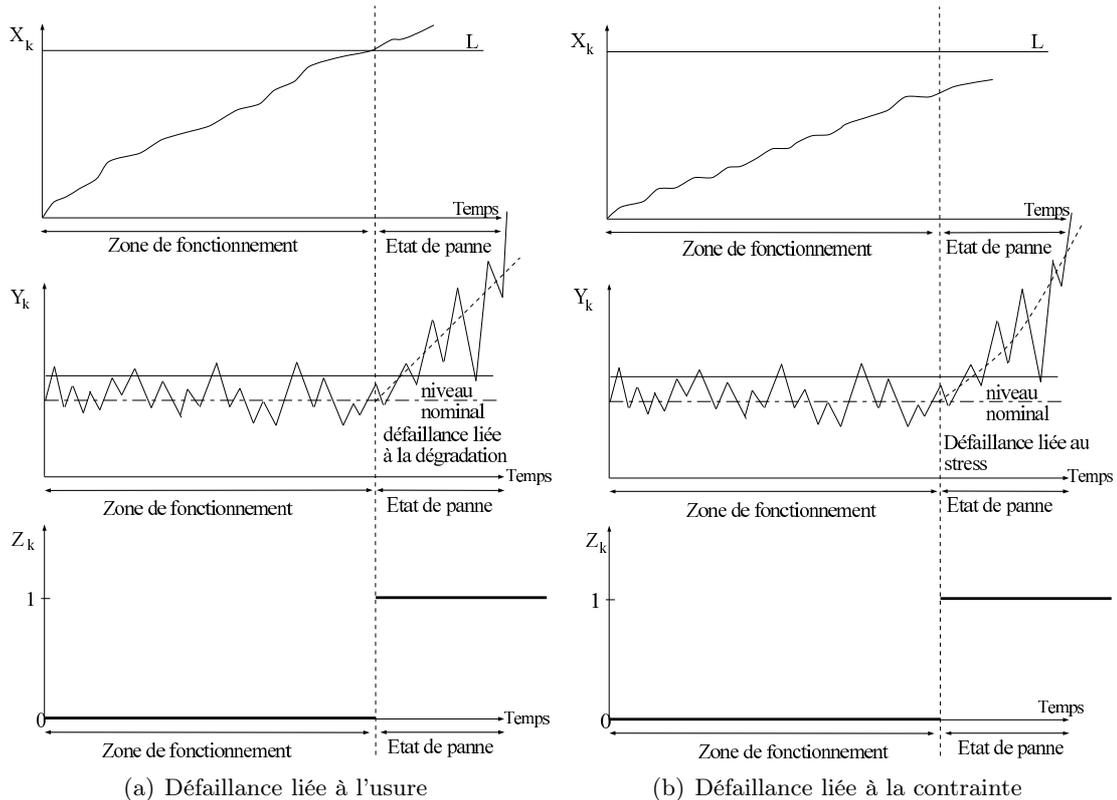


FIGURE 2.9 – Evolution du processus de dégradation dans le cas du modèle à risques compétitifs

Structure décisionnelle de maintenance : La démarche que nous avons proposée est la superposition de politiques de maintenance définies pour chacun des deux modes. En terme de politique d'inspection, nous proposons deux niveaux : des inspections dites *avancées* ou *parfaites* et des inspections *minimales*. Une inspection avancée planifiée périodiquement permet de rendre compte de l'état du système et de son niveau de dégradation courant alors que l'inspection minimale permet uniquement de savoir si le système est en fonctionnement ou non. Une carte de contrôle classique pour la surveillance d'observations individuelles [13], l'intensité de la contrainte Y_k , est mise en place pour la détection d'une éventuelle dérive synonyme de défaillance. Si la valeur de Y_k est supérieure à la *limite de contrôle supérieure*, UCL , de la carte alors on effectue une inspection minimale. La valeur de la limite UCL doit ainsi offrir le meilleur équilibre entre fausses alarmes (une inspection alors que le système est en fonction) et détection au plus tôt d'éventuelles défaillances. Afin de limiter le nombre de paramètres de décision mais aussi dans un souci de simplification de mise en place opérationnelle, nous proposons un schéma périodique pour les inspections parfaites.

En termes d'actions de maintenance à proprement parler, nous proposons deux niveaux de maintenance, à savoir une maintenance dite *de routine* et les classiques remplacements qui peuvent être préventifs ou correctifs. Un schéma conditionnel basé sur un seuil ξ est proposé pour le remplacement préventif, $X_k \in [\xi, L)$. La maintenance de routine, pouvant aussi être appelée *maintenance minimale* par analogie à la classification des modèles de maintenance imparfaite, cherche à améliorer la résistance du système à la contrainte en réduisant les effets temporels sur celui-ci. Nous avons proposé à titre d'illustration d'associer cette action à une opération de graissage des voies de guidage d'un tramway, graissage qui permet de limiter les effets vibratoires sur le système embarqué de guidage sans pour autant avoir un impact sur son niveau de dégradation. Un schéma purement périodique de période δ est proposé pour la mise en place de la maintenance de routine. La Figure 2.10 illustre l'influence de la valeur de la période de maintenance de routine δ sur un ensemble de grandeurs fiabilistes.

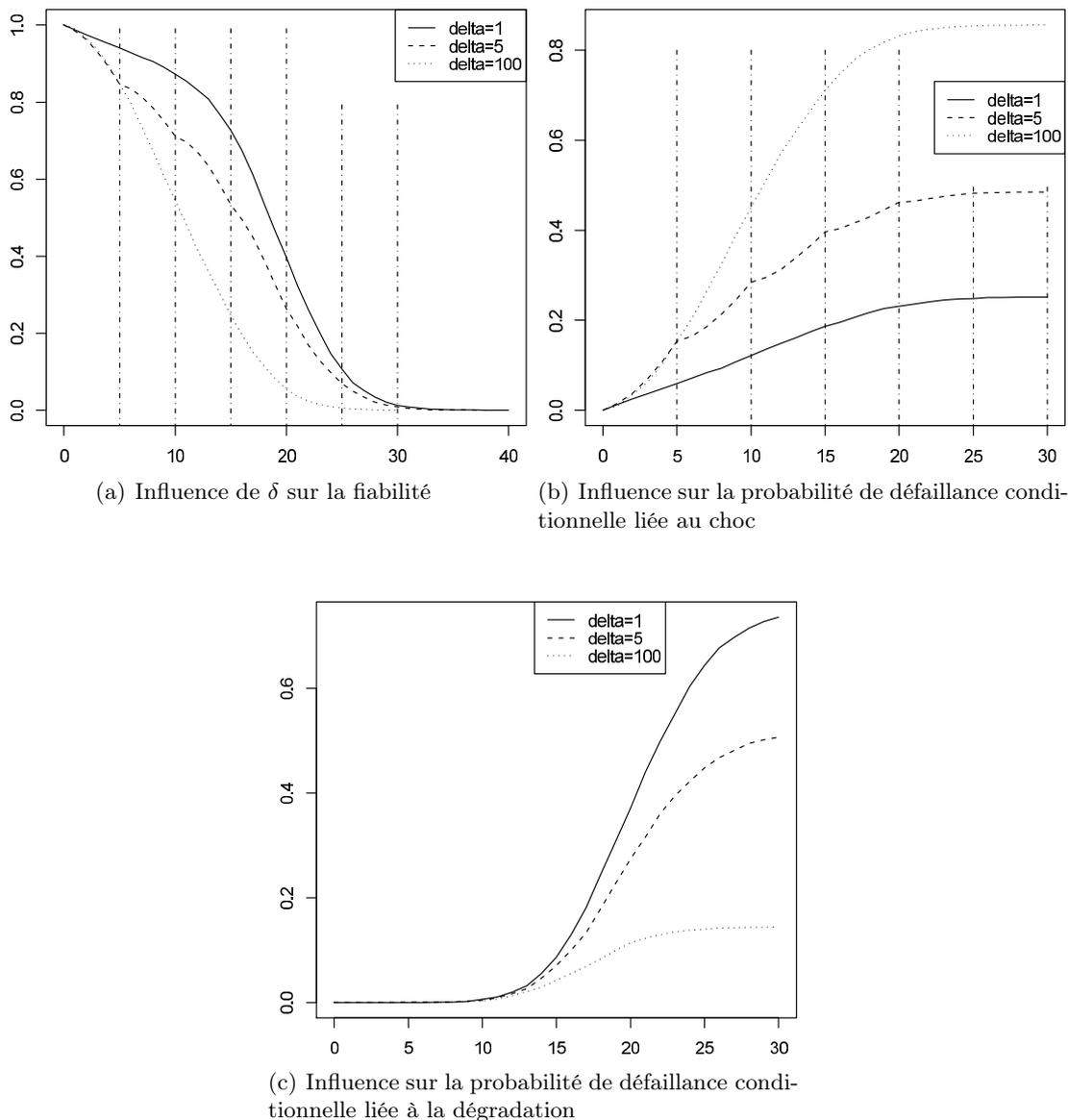


FIGURE 2.10 – Illustration de la périodicité de la maintenance de routine sur la fiabilité et diverses probabilités de défaillance

En résumé, les paramètres de décision que l'on doit optimiser sont :

- τ , la période d'inspection qui permet d'équilibrer les coût cumulés d'inspection et d'indisponibilité liée aux défaillances ;

- ξ , le seuil de maintenance préventive qui permet de gérer le ratio remplacements préventif et correctif;
- UCL , la limite de contrôle supérieure qui gère le compromis fausse alarme et détection de la défaillance;
- δ , la période des maintenances de routine pour la réduction des effets temporels sur la probabilité de défaillance de type choc.

La caractérisation du comportement du système maintenu ainsi que la construction du critère de décision reposent ici aussi sur l'analyse d'un cycle de renouvellement défini depuis la mise en service du système jusqu'à la date de remplacement, remplacement préventif ou correctif que ce soit par l'observation de la défaillance lors d'une inspection minimale ou bien parfaite.

Contributions : Les contributions se positionnent de notre point sur deux niveaux. Le premier est la proposition d'un modèle de défaillance à risques compétitifs mixant dégradation continue et défaillance soudaine. Le premier modèle reflète une très bonne connaissance du mécanisme de défaillance alors que le second permet d'intégrer des incertitudes dans ce dernier, notamment dans la relation entre niveau de dégradation et défaillance. La définition d'un taux de défaillance dépendant à la fois du niveau de dégradation et du temps de mise en service reflète de notre point de vue une certaine réalité pour laquelle les instants de défaillance peuvent s'expliquer en fonction du temps et de l'usage du système. Nous sommes toutefois conscients de la difficulté d'identifier et de caractériser chacune des contributions mais nous invitons les exploitants en maintenance à enrichir leur base de données en fonction. Par ailleurs, une telle définition d'un taux de défaillance nous a permis de modéliser les effets d'une maintenance de routine qui, dans un contexte général, n'impacte pas significativement la dégradation du système mais garde un rôle prépondérant dans son bon fonctionnement.

Le second niveau de contribution se situe sur le champ de la construction de politiques de maintenance. Afin de bénéficier de toute l'information disponible, nous avons proposé de combiner deux approches classiques rencontrées en maintenance ou en surveillance des procédés. À notre connaissance, nous avons été précurseurs dans ce sens.

5 Conclusions et perspectives

5.1 Contributions

L'un des premiers points que j'aimerais souligner dans cette conclusion est relatif à l'intérêt qu'ont suscité nos travaux, sans s'en octroyer l'exclusivité, autour de la modélisation de la maintenance conditionnelle pour des systèmes dits à dégradation graduelle. Nos travaux ont été à l'initiative de nombreuses thèses de doctorat dont l'objectif est d'améliorer l'applicabilité des modèles en introduisant de nouvelles hypothèses et de prouver l'intérêt de ces approches dans des contextes industriels. Une réelle communauté scientifique internationale s'est créée autour de ces problématiques de maintenance conditionnelle de systèmes à dégradation continue et notamment l'application des processus gamma dans ce contexte de maintenance. On pourrait citer près d'une dizaine de contributions sur ces problématiques au dernier congrès PSAM-ESREL 2012 alors que, dans ce même congrès, les sessions *Maintenance* y étaient quasiment inexistantes encore en 2001.

Afin d'étudier des situations de plus en plus complexes, nous avons reposé la résolution des critères de décision sur des résultats classiques de la théorie du renouvellement et renouvellement markovien. Ici aussi, nous avons contribué à la mise en évidence de l'avantage de ces résultats théoriques qui trouvent, dans le contexte de la maintenance, un réel champ d'application.

Pour clore ce paragraphe de contributions, je soulignerai l'impact des travaux que nous avons menés pour la prise en compte dans les modèles de maintenance conditionnelle des effets

de variables environnementales sur les phénomènes de dégradation et des modèles multicomposants. Ces deux thèmes avaient été particulièrement bien abordés dans la littérature dans un contexte de durée de vie mais très peu en maintenance conditionnelle. Pour les premières études, un avantage certain que nous voyons dans un cadre de maintenance conditionnelle est une facilité de modélisation des effets de l'environnement soit par le biais d'observations, soit par le biais d'expertise, facilité due au caractère mesurable de la dégradation. Nous restons persuadés de l'applicabilité de nos travaux et leur intérêt sur des problématiques industrielles mais aussi auprès de la communauté scientifique (notre travail [31] a été cité près d'une vingtaine de fois depuis 2009 -source Web of Knowledge). Pour les systèmes multicomposants, cet intérêt nous semble encore plus fort au vu des statistiques de citation, à savoir plus de 30 fois depuis 2005 pour l'article [18].

5.2 Perspectives

Je ne pourrais pas conclure ce chapitre sans souligner quelques perspectives à donner à nos travaux.

La première et non des moindres est de renforcer la justification de l'utilisation de tels processus stochastiques pour la modélisation de dégradation. Plusieurs voies complémentaires pourront contribuer à ceci.

- L'une des limites principales est la stationnarité des processus stochastiques qui ne permet pas de prendre en compte des comportements moyens non linéaires. Plusieurs travaux ont d'ores-et-déjà proposé des approches non stationnaires. On retiendra la généralisation du processus gamma caractérisée par van Noortwijk [110] dans laquelle fonctions de forme et d'échelle ne sont pas linéaires en fonction du temps et très récemment la définition d'un processus gamma dépendant de l'état présentée au dernier PSAM7-ESREL2012 par Vatn [111]. La fonction de forme du processus gamma dépendant de l'état est elle-même une variable aléatoire distribuée suivant une loi gamma avec comme paramètre de forme dépendant de la valeur courante de dégradation une fonction non linéaire en temps donnée. Cette approche est particulièrement pertinente lorsque l'évolution de la dégradation est fonction de l'état et pas forcément du nombre de cycles de sollicitations (par exemple, l'accroissement d'une fissure est directement fonction de sa taille, loi de Paris-Erdogan [87]). Ces derniers travaux sont à rapprocher du modèle que nous allons définir dans le chapitre suivant. L'une des limites du modèle proposé par Vatn est la définition de la fonction de forme de la loi gamma, cet aspect n'étant pas discuté dans son papier.
- Une autre limite que nous pouvons souligner est liée au manque d'identification du modèle sans réelle caractérisation de ses paramètres au regard des connaissances ou expertises sur les mécanismes de dégradation. Ce point rejoint partiellement la discussion motivant la construction de *boites grises* p. 28. Un travail dans ce sens devrait à la fois permettre d'améliorer les qualités prédictives du modèle lorsque les données restent peu nombreuses et aussi faciliter son acceptation par la communauté scientifique de Génie Civil et Mécanique (notamment française).

Une seconde voie d'extension de nos travaux est l'intégration d'hypothèses plus réalistes concernant par exemple le processus d'acquisition de l'information. En effet, nous avons jusqu'ici considéré l'information obtenue comme parfaite, dans le sens où elle représente complètement l'état du système. Il serait intéressant de considérer un ensemble d'incertitudes sur le processus de mesure, incertitudes liées à l'aptitude de l'appareil de mesure à détecter les défauts et à la variabilité de celle-ci ainsi que l'incertitude sur la relation entre l'indicateur d'état du système et son état réel, i.e. la pertinence de la mesure par rapport à la dégradation même.

Une troisième voie d'intérêt qui rejoint notre travail sur le couplage carte de contrôle et schéma d'inspection périodique est la combinaison de divers signaux ou sources d'information

sur la dégradation. Je pense ici, par exemple, à un couplage de méthodes de type Health Monitoring relevant d'un outillage de surveillance *en mode continu* de la structure et des méthodes de type Contrôle Non Destructif pour lesquelles la mesure est ponctuelle dans le temps (on parle d'inspection).

Enfin, pour conclure cette liste non exhaustive de perspectives, il serait intéressant de généraliser les règles de décision dans le cas multicomposant pour des systèmes complexes formés d'un plus grand nombre de composants. Il est clair qu'il sera nécessaire de trouver des alternatives à notre programme de résolution numérique et d'évaluation de la loi stationnaire associée.

Chapitre 3

Application de la Programmation Dynamique à la gestion de l'entretien routier

1 Introduction

Dans le chapitre précédent, l'organisation de la maintenance repose sur une structure paramétrique de décision définie a priori, les valeurs des paramètres étant par la suite optimisées, par exemple, pour minimiser un coût de maintenance moyen à long terme. On peut soulever ici deux remarques à cette approche.

- La première est que la définition a priori de la structure n'offre aucune garantie du caractère optimal de la décision, même si la proposition de la structure peut reposer sur un ensemble d'expertises et d'analyses de pratiques industrielles. L'hypothèse de structure optimale est relativement forte et se doit d'être vérifiée par une analyse fine du problème d'optimisation. Par ailleurs, le choix même de la structure et des paramètres associés peut être particulièrement délicat lorsque, par exemple, l'état du système est défini par plusieurs variables corrélées.
- La deuxième remarque est liée aux difficultés numériques rencontrées dans nos approches paramétriques. La résolution du problème d'optimisation repose sur l'évaluation de la loi stationnaire d'évolution de l'état du système maintenu, cette loi étant ici formulée comme une équation de renouvellement complexe ne présentant pas de solution analytique. Il est alors nécessaire de l'évaluer numériquement point par point pour chaque itération de l'algorithme d'optimisation des paramètres de décision. Outre l'explosion des temps de calcul, ceci étant vu comme une limite pour étendre nos approches mais aussi on peut s'interroger sur la qualité et la précision des résultats numériques offerts dans des contextes décisionnels plus complexes (dans nos études précédentes, nous avons cherché à valider ces résultats en les comparant à des estimations obtenues par le biais de simulations stochastiques).

Nous proposons ici de ne pas considérer de structure a priori et de formuler le problème comme un problème de programmation dynamique stochastique. Un problème de programmation dynamique [10] permet de caractériser la décision optimale, si elle existe, pour chacun des états du système en *minimisant un coût* ou respectivement en *maximisant une valeur*, cette optimisation étant faite sur le gain immédiat de l'action et l'espérance des gains futurs. Il devient alors nécessaire de caractériser la transition entre états du système sur une période de décision, sans chercher à caractériser le comportement stationnaire du système maintenu, ce qui laisse présager de fortes économies sur ce champ.

Les travaux que je vais présenter ici reposent sur la programmation dynamique stochastique. L'objectif principal de ce chapitre est de montrer comment nous l'avons mise en œuvre

sur une problématique issue d'un contexte industriel réel. Ces travaux ont été développés dans le cadre d'un projet financé par la Région Pays de la Loire et en collaboration avec une équipe de recherche de l'ESAR du LCPC, nouvellement IFSTTAR (Nantes-Bouguenais). Ils reprennent l'ensemble des travaux de thèse [121] de l'IRCCyN défendue en novembre 2011 par Mariem ZOUCHE sous la direction de Philippe CASTAGLIOLA et co-encadrée par Thomas YEUNG et moi-même. L'objectif global de ce projet (2007-2012) est la proposition de nouvelles approches pour la gestion de l'entretien des chaussées qui prennent en compte les spécificités techniques des modes de dégradation des routes en associant les performances économiques de l'aspect gestion. La motivation première en était l'étude de l'applicabilité des approches processus stochastiques et particulièrement des processus gamma dans ce contexte du génie civil.

Finale­ment, les contributions sur ce projet se situent plus, du moins sur un positionnement à court terme, sur le champ théorique. En effet, la base des relevés de dégradation des chaussées pressentie pour cette étude a été jugée par les experts au vu de son analyse, peu fiable et mal renseignée pour justifier de l'apport de nos propositions. D'une manière générale, les contributions présentées sur ce chapitre sont :

- l'introduction d'un *potentiel de fissuration* dans le modèle de dégradation à deux phases : initiation et propagation ;
- la définition du processus gamma dépendant de l'état pour la prise en compte des effets non stationnaires de l'évolution de l'indicateur de dégradation après observation du défaut ;
- l'applicabilité d'une approche type programmation dynamique sur un problème de maintenance inspiré du réel. L'approche nous a ainsi permis d'étendre les résultats sur le champ global de l'optimisation de la maintenance par :
 - la proposition d'un schéma de maintenance conditionnelle facilement opérationnel basé sur le niveau de dégradation du système et sa vitesse d'évolution ;
 - une meilleure considération des effets d'une maintenance imparfaite sur les lois d'évolution du système maintenu ;
 - un modèle de maintenance multi-action complexe optimisant le compromis traitement correctif du symptôme et mesure palliative en fonction de l'évolution de l'état observé avant maintenance.

L'organisation de ce chapitre est la suivante. Avant de situer le problème de gestion de l'entretien routier par rapport aux problématiques générales en modélisation de la maintenance, je présenterai la méthodologie générale que nous avons cherché à appliquer pour conduire notre étude. Les dernières parties se focaliseront sur la présentation des contributions scientifiques et je conclurai par une mise en évidence des perspectives de ces travaux.

2 Méthodologie : de la formulation du problème en programmation dynamique à la décision de maintenance

La démarche que je vais présenter ici reste très classique dans son déroulement. Sa présentation est plutôt à vocation pédagogique et ne se situe pas réellement dans le champ des contributions. Par ailleurs, je ne reviendrai pas sur les grandes étapes pour la construction d'un modèle de maintenance présentée en première partie de la section 2. À partir de la définition des états du système et la caractérisation de leur loi d'évolution, la méthodologie mise en place se décline en 3 étapes.

2.1 Choix et construction du critère de décision :

Définition de la fonction objectif : Durant cette étude, nous sommes concentrés sur une optimisation économique de la maintenance. Le critère choisi est la minimisation d'un coût de maintenance. Ici aussi, bien qu'une modélisation de type programmation dynamique facilite

la résolution d'un critère moyen sur un horizon de temps fini, nous avons choisi de résoudre un critère moyen sur un horizon infini. Une des spécifications de conception d'une chaussée dite à *longue durée* est la durée de vie à 30 ans, i.e. sans réparation majeure sur cette durée [3] mais la durée de son exploitation reste, elle, indéfinie.

Dans le contexte stationnaire, la résolution d'un problème de programmation dynamique cherche à déterminer la *décision* ou *action* optimale pour chaque état x , l'ensemble de ces décisions formant la politique optimale dénotée par la suite π^* . La décision optimale offre le meilleur compromis entre le coût immédiat de la mise en place de cette décision a et l'espérance du coût à venir après cette mise en place, espérance pondérée par un facteur d'escompte, $\lambda < 1$ (cette pondération assure la convergence du critère), équation de Bellman que l'on peut écrire de la façon suivante :

$$V_\tau(x) = \min_{a \in A(x)} \{C(x, a) + \lambda E_G [V_\tau(y) \mid x, a]\} \quad (2.1)$$

avec

- τ : l'intervalle de décision séparant deux inspections successives que l'on considèrera constant, on pourra considérer τ comme variable de décision à optimiser ;
- $A(x)$: l'ensemble des décisions possibles connaissant l'état x ;
- $C(x, a)$: le coût immédiat engendré par la décision a connaissant l'état x ;
- E_G : l'espérance conditionnelle suivant la loi associée à l'état suivant la décision.

Le facteur d'escompte : Outre l'intérêt mathématique du facteur d'escompte, on peut justifier l'introduction de ce facteur sur des considérations purement économiques en y associant les dépréciations économiques de la monnaie ou, pour ce qui me concerne, une notion plus intéressante de *mesure d'aversion au risque* du décideur : “ λ proche de 0” sera synonyme d'une décision sans prise en considération du futur alors qu'un “ λ voisin de 1” induira une décision *équilibrée* entre coût immédiat et coûts futurs.

2.2 Procédure de résolution et illustration numérique

Dans notre cas, et nous le définirons plus tard, l'état du système est infiniment mesurable et l'équation de Bellman dans sa forme continue ne peut être résolue directement. Nous avons alors proposé, après avoir formulé cette équation, de discrétiser l'ensemble des états et de considérer un ensemble dénombrable de décisions. Le problème se ramène alors à un problème de décision markovien classique [90] et nous avons utilisé les algorithmes classiques de résolution tels que le *Policy Iteration Algorithm* (PIA) [41]. Outre la facilité de mise en place et la production rapide de résultats, cette discrétisation permet, entre autres, d'étudier les conditions d'optimalité.

Afin de mettre en évidence l'intérêt de notre approche, nous avons développé des codes informatiques sous R© qui, après définition de l'ensemble des paramètres de dégradation et des coûts, évaluent et mettent en forme les politiques optimales¹. Un autre avantage dans la discrétisation du problème se trouve dans l'ergonomie de la présentation du résultat sous forme de tables de décision faciles à mettre en œuvre dans un contexte opérationnel.

2.3 Améliorations des performances algorithmiques et préconisations

Il est sûr que notre première approche de résolution est perfectible, notamment en termes de sensibilité au pas de discrétisation et de puissance et temps de calcul. L'approche que nous avons choisie de mettre en place pour l'amélioration de ces performances est la mise en évidence de propriétés structurelles du problème d'optimisation. Ceci repose sur la caractérisation des ordres de grandeur des gains associés aux différentes décisions en fonction des états. La notion

1. Un stagiaire de Master informatique financé par le projet SBADFORM réalise de Mars 2012 à Juillet 2012 une interface web pour la diffusion de nos résultats.

d'ordre peut être difficile à définir dans un cadre stochastique surtout lorsque l'état du système est multi-dimensionnel. Une fois cette notion d'ordre définie, une analyse systématique et rigoureuse des fonctions de coût est à réaliser pour définir les conditions d'existence, ces conditions devant par ailleurs être réalistes par rapport au contexte d'exploitation.

Dans les cas les plus simples, les conditions d'existence des structures rejoignent des idées intuitives alors que, dans des contextes plus complexes, celles-ci peuvent paraître très restrictives.

Les propriétés de structures offrent un double avantage. Du point de vue algorithmique, elles permettent d'accélérer les méthodes d'exploration de l'espace des solutions en évitant certaines régions. Par exemple, sous condition d'existence, soient 2 états s_1 et s_2 où s_1 est défini comme *moins dégradé* que s_2 , alors si l'action a_1 est optimale pour l'état s_1 , on peut dire que l'action optimale pour s_2 sera au mieux a_1 , évitant alors à l'algorithme d'explorer des solutions considérant des actions *moins efficaces* que a_1 pour l'état s_2 . Le second avantage est d'ordre opérationnel. En effet, de telles structures rendent l'information plus compréhensible et lisible sur les tables de décision. De plus, elles permettent de préjuger de la pertinence des actions : doit-on conserver une action a_k jugée efficace au vu des différentes données du problème sur un petit intervalle de dégradation ?

3 Optimisation de l'entretien routier : problématiques et positionnement

3.1 Cadre général du projet SBADFORM

Les problématiques d'optimisation de gestion de l'entretien routier sont complexes car elles se doivent d'intégrer des dimensions économiques (dimensionnement et disponibilité des budgets), des aspects politiques, sociaux et stratégiques (impacts non négligeables sur la définition des priorités), le dimensionnement et la disponibilité des équipes de réalisation des travaux (appels d'offres, etc.), les aspects de réseau (niveau décisionnel stratégique) et bien entendu, plus au niveau de la section, l'état de la chaussée et ses caractéristiques d'évolution qui diffèrent suivant le type de chaussée à entretenir (niveau opérationnel). Toutes ces dimensions sont à conjuguer avec les aspects réglementaires, environnementaux (dont les nuisances sonores), sécuritaires et confort de l'utilisateur. Le décideur est donc face à un problème d'optimisation multi-critères très complexe d'autant que les données à prendre en compte ne sont pas du tout homogènes et, pour certaines, entièrement non maîtrisables et purement subjectives (notamment les données politiques et sociales). Nous proposons ici de centrer les critères en fonction des données actuellement disponibles et fournies par les partenaires du projet, données de nature plutôt technico-économiques.

Nous avons choisi de respecter le mode décisionnel actuel à savoir une notation globale d'un réseau gérant les activités au niveau stratégique construite à partir de notes de sections équivalentes à 200 mètres de chaussée pour la gestion au niveau opérationnel. La première partie du projet SBADFORM se centre uniquement sur la gestion de l'entretien d'une section indépendamment du réseau. Étant actuellement seulement initiée, la problématique de la maintenance conditionnelle au niveau du réseau ne sera pas abordée dans le contexte de ce manuscrit.

Afin de positionner nos travaux dans un contexte de recherche académique, nous proposons de présenter dans le paragraphe suivant une analyse bibliographique des modèles de maintenance relative à nos intérêts, à savoir les modèles d'inspection, les modèles de maintenance imparfaite ainsi que les modèles multi-actions.

3.2 Analyse bibliographique centrée sur les modèles de maintenance

L'objectif de notre étude n'est pas directement lié à la proposition de nouveaux modèles performants dans le sens où ils rejoignent, par exemple, les caractéristiques spécifiques des fissures de fatigue, d'autant qu'une modélisation aussi précise qu'elle soit ne permettra pas de retrouver la loi d'évolution du pourcentage de fissuration $\rho(t)$. Aussi, nous ne cherchons pas ici à cataloguer l'ensemble des modèles de fissuration fournis par la littérature. Par contre, nous concentrerons notre analyse sur les modèles qui répondent à nos préoccupations soit sur le champ de l'inspection, soit sur l'aspect de modélisation et prise en compte de l'impact d'un entretien dans le processus de décision ou enfin pour offrir la possibilité d'intégrer plusieurs actions de maintenance dont l'efficacité et les coûts associés peuvent différer.

Afin de rendre cette analyse la plus concise, seules les approches dites mono-composant sont ici présentées, ceci étant justifié par notre choix de concentrer notre étude sur un unique mode de dégradation.

On mettra en opposition les approches dites basées sur l'âge qui reposent essentiellement sur les distributions de durée de vie aux approches conditionnelles pour lesquelles on cherche aussi à modéliser un phénomène de dégradation cumulée.

Les modèles d'inspection

L'inspection permet de rendre compte de l'état d'un système. Elle peut être parfaite ou imparfaite, continue, périodique ou séquentielle. L'objectif de cette section est d'en décrire brièvement les approches en dégageant les avantages et les limites.

Castanier [14] propose une classification des modèles d'inspection en *politiques d'inspection correctives* et *politiques d'inspection préventives*.

Les premières cherchent à minimiser les impacts d'une défaillance qui n'est pas *auto-décelable* se traduisant par l'objectif de réduction des temps de panne. La majorité des travaux proposés dans la littérature s'inspire du modèle de Munford [70]. L'état du système est généralement réduit à un état de fonctionnement et un état de panne et les modèles reposent sur la connaissance d'une distribution de durée de vie. L'inspection est supposée parfaite. Dans le cas où elle est périodique, la résolution numérique des critères de coûts associés permet l'obtention de la périodicité optimale. Si l'inspection est séquentielle, des politiques sous-optimales sont proposées, notamment lorsque la distribution de durée de vie est une loi de Weibull à 2 paramètres. L'approche classique pour modéliser les inspections imparfaites est l'introduction d'une probabilité d'erreur. Cette probabilité peut être constante [47] ou fonction du temps passé depuis la défaillance [19].

Les politiques d'inspection préventives cherchent à identifier la présence d'un symptôme ou défaut précurseur de la défaillance ou bien s'assurer que le système est dans un état appartenant à un intervalle *proche* de la zone de défaillance. Cet intervalle peut être défini par un seuil déterministe [22], un seuil d'alerte aléatoire [103] ou encore à optimiser pour réduire, par exemple, des coûts d'exploitation, cet intervalle optimisé offrant alors le meilleur compromis entre temps de fonctionnement et indisponibilité de panne. Les modélisations associées sont fonction de la modélisation des états de fonctionnement du système. On peut rencontrer les modèles à délai introduits par Christer [21], des distributions sur les temps d'atteinte d'un seuil d'alarme, ou bien, par des processus dit de saut généralement Markovien pour la modélisation de la dégradation cumulée (discrète ou continue). La grande majorité des travaux suppose une inspection parfaite permettant ainsi d'observer le vrai niveau de dégradation du système. Les approches considérant que l'inspection ne peut rendre qu'une information partielle reposent :

- dans le cadre des politiques basées sur l'âge, sur une probabilité de détection équivalente aux politiques correctives ;
- dans un contexte de dégradation cumulée, sur :
 - l'introduction d'une erreur modélisée par un bruit blanc ;

- l'étude des processus de Markov cachés.

Par rapport au schéma d'inspection, ici aussi, les trois approches continue, périodique et séquentielle sont considérées. Alors que les deux premières n'engendrent pas de complexification du problème, le problème séquentiel est difficile à résoudre numériquement sans hypothèse ou structure décisionnelle supplémentaires. Sous une hypothèse de systèmes à taux de défaillance croissant, on peut montrer que la séquence des inspections se doit d'être décroissante en fonction de l'âge ou bien de l'état du système.

Les modèles de maintenance imparfaite

L'objectif de ce paragraphe est de présenter des modèles de maintenance imparfaite qui nous semblent pertinents pour notre étude. Les approches classiques en maintenance supposent une réparation parfaite, dans le sens où celle-ci permet de restituer le système dans son état *aussi bon que neuf*. Bien qu'elle puisse se justifier dans un contexte de systèmes non réparables et, dans ce cas, on associe réparation à remplacement total, nombre de travaux ont cherché à modéliser l'efficacité de l'action de maintenance [113]. On distingue les modèles de maintenance minimale et modèles de maintenance imparfaite.

Les premiers supposent qu'une réparation permet uniquement de rétablir la fonction du système sans aucun impact sur son âge dans un état dit *aussi mauvais que vieux*. On peut associer ces actions à des actions du type *rétablissement d'un courant électrique disjoncté* ou, pour des systèmes complexes, des actions dont l'impact de cet entretien sur le niveau de vieillissement du système est négligeable (changement de fusibles, d'ampoules électriques, etc.). Les modélisations sont proposées dans un contexte de maintenances basées sur l'âge et dans un contexte purement correctif. Ces modèles sont peu pertinents dans notre contexte de fissuration.

Les seconds modèles permettent de rétablir la fonction du système mais aussi ont un impact sur le niveau de vieillissement. On trouve alors les modèles de *réduction de l'intensité des défaillances*, les modèles de *réduction de l'âge* ou encore les modèles de réduction du taux de fiabilité [74] pour lesquels les performances du système en terme de dégradation sont modélisées au travers d'un index de fiabilité [56]. Les modèles de réduction de l'intensité se basent sur une réduction du taux de défaillance d'une quantité proportionnelle à sa valeur juste avant la réparation lorsque les défaillances sont modélisées par des processus de Poisson non homogènes [33]. Les modèles de maintenance basés sur la réduction de l'âge utilisent la notion "d'âge virtuel" proposée par Malik [61] et développée par Kijima *et al.* [51] ou Last [53]. Dans ces modèles, l'action de maintenance a pour effet de rajeunir le système qui se comportera comme un système neuf ayant vécu un âge plus petit que celui dans lequel il se trouvait juste avant l'opération de maintenance. Quelques modèles ont été proposés dans ce cadre et nous citerons pour exemple ceux de Kijima de type I et II [50] et celui de Doyen [33] de réduction arithmétique de l'âge de mémoire. L'impact d'un entretien est directement introduit au niveau du taux de défaillance du système. Les limites des approches présentées peuvent se situer à deux niveaux. Tout d'abord, étant donnée la non-observabilité d'un taux de défaillance, il devient extrêmement délicat de bien modéliser l'impact de l'entretien ou du moins il est nécessaire de consolider cette modélisation sur un très grand nombre de données. D'autre part, il est supposé que la forme de la loi de durée de vie reste inchangée après entretien, on considère que le système réparé est identique dans son mode de vieillissement au système non réparé. On ne peut pas alors prendre en compte des effets soit d'accélération du vieillissement liée à l'usage depuis la date de première mise en service, soit d'amélioration liée par exemple à l'utilisation de technologies plus performantes. Bien que non orientés vers les applications en maintenance, les modèles de durée de vie accélérée ou encore les modèles à hasard proportionnels permettent de modéliser les variations des lois associées par le biais de facteurs d'accélération ou de l'intégration de différentes covariables. On ne cherchera pas ici à étudier l'ensemble de la littérature vouée à ce domaine fiabiliste mais on pourra se référer à [72] qui en présente les

principales modélisations.

Dans un contexte de maintenance conditionnelle, les travaux développés pour modéliser l'efficacité d'un entretien sont plus rares. La notion de maintenance minimale ne tient que si la défaillance n'est pas uniquement liée à une détérioration excessive par dépassement d'un seuil déterministe. Les premières approches développées [16, 17] sont à mettre en parallèle avec les premiers modèles de réduction de l'âge dans le sens où une maintenance dite imparfaite permet de diminuer le niveau de dégradation à une valeur pouvant être aléatoire comprise entre le niveau observé avant la maintenance et le niveau parfait sans changer la loi d'évolution de la dégradation. Il est à noter que, dans [17], les auteurs proposent un modèle de maintenance permettant d'optimiser l'efficacité de la maintenance en fonction du niveau de dégradation observé. Sous leurs hypothèses, le processus de dégradation du système maintenu conserve des propriétés d'ergodicité et, même si la complexité du problème impose la construction d'algorithmes appropriés, la loi stationnaire d'un tel processus est encore accessible. Ainsi, la résolution et l'optimisation d'un critère de décision asymptotique sont rendues possibles notamment grâce à l'application de la théorie du renouvellement markovien qui généralise la théorie du renouvellement en introduisant la notion de cycle de semi-régénération [26, 27]. Si l'évolution future de la dégradation est altérée après une maintenance, hypothèse réaliste dans le sens où un système même réparé risque de vieillir plus vite qu'un système neuf, la complexité du problème fait qu'il est alors difficile, voire impossible de caractériser numériquement la loi stationnaire du système maintenu. Dans sa thèse [83], Nicolaï discute de l'utilisation d'approches type Programmation Dynamique pour la résolution du problème d'optimisation de la maintenance lorsque celle-ci permet de modifier les paramètres du processus γ , processus modélisant l'évolution de la dégradation. L'avantage d'une telle approche est la possibilité, sous certaines hypothèses, de déterminer la structure de maintenance optimale sous condition d'existence, la majorité des travaux développés dans le cadre de l'optimisation de la maintenance reposant sur une structure de maintenance paramétrique dite à seuils ou politique limite de contrôle.

Les modèles de maintenance multi-action

On appelle maintenance multi-action les politiques pour lesquelles plusieurs actions permettant de remettre le système en condition sont disponibles. Il est bien entendu que l'efficacité des actions diffèrent ainsi que leur coût. Bien que la problématique d'optimisation de telles maintenances trouve largement son intérêt dans un contexte industriel, elle reste très peu abordée dans la littérature. En général, les approches proposées sont liées aux politiques de maintenance imparfaite pour lesquelles on différencie réparation et remplacement, la réparation étant la maintenance imparfaite et le remplacement la remise à neuf du système. Pour les stratégies conditionnelles, la règle de décision généralement adoptée [17] est de type contrôle limite :

- une zone de remplacement préventif est définie entre le seuil de remplacement préventif (généralement une variable de décision à optimiser) et le seuil de défaillance (généralement une donnée du problème car définie par les propriétés physiques et mécaniques du système considéré) ;
- une zone de réparation préventive est définie entre les seuils de réparation et de remplacement préventif.

Par conséquent, pour des systèmes peu dégradés, on privilégiera des réparations permettant de conserver le système dans des états acceptables et ainsi étendre le cycle de vie du système en opposition à des systèmes trop dégradés pour être maintenus. Lorsque le nombre d'actions augmente, la proposition d'une structure paramétrique n'est plus évidente surtout si leur efficacité est mesurable au niveau de la loi de dégradation. Ici aussi, les approches d'optimisation de type Programmation Dynamique Stochastique [10] trouvent tout leur intérêt. L'avantage de telles approches reste dans leur mise en oeuvre qui ne nécessite ni de la définition d'une

structure de décision a priori, ni de la caractérisation d'une loi d'évolution asymptotique mais de la seule transition sur une période de temps donnée et les gains associés à cette transition. Sans entrer trop dans les détails (voir [90] pour de plus amples informations sur les Processus de Décision Markovien (MDP)), la méthodologie consiste, pour une politique donnée (un ensemble de règles de décision fixées), à évaluer le coût d'une transition d'un état de dégradation vers un autre pour chacune des actions. On associera ensuite un algorithme de recherche pour la détermination de la politique optimale, si elle existe. Sous certaines conditions sur les paramètres de dégradation, d'efficacité de maintenance et de coûts, on pourra retrouver une structure de type seuil pour les règles de décision. À notre connaissance, aucun travail n'a encore été proposé sur ce dernier champ.

4 Modélisation du processus de fissuration longitudinale

Sans chercher à reprendre ici l'ensemble des éléments mécaniques de la chaussée et des facteurs contribuant associés, nous rappelons brièvement que le processus de fissuration longitudinale est un processus en deux phases. La première phase constitue la phase d'initiation de fissures qui prennent naissance sous forme de micro-fissures dans la couche de fond de la chaussée soumise à des contraintes répétitives liées essentiellement au trafic et qui remontent jusqu'à la surface. La seconde phase est la phase dite de propagation. Les fissures ont la particularité de se développer en surface principalement au niveau des passages de roues et même si elles peuvent se ramifier, on peut les considérer uniquement dans leurs longueurs. Actuellement, les campagnes d'auscultation des chaussées sont planifiées tous les trois ans et rendent, pour le mode de détérioration fissures longitudinales, un pourcentage linéaire de fissuration par section de chaussées (une section représente 200 mètres) -une illustration de la construction et de l'évolution de cet indicateur est proposée Figure 3.1. Ces informations sont enregistrées dans la base IQRN. On notera qu'aucune autre information n'est disponible entre deux inspections.

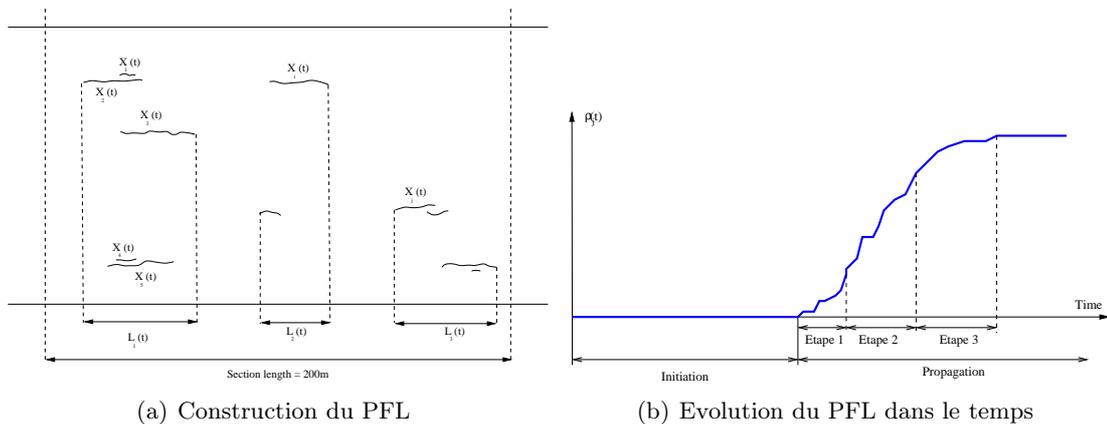


FIGURE 3.1 – Illustration de l'indicateur de fissuration longitudinale fourni par l'IQRN

L'indicateur $\rho(t)$ est compris entre $[0, 1]$. Son initialisation est liée à l'occurrence de la première fissure à la surface de la chaussée. Cette occurrence étant liée à un processus de fatigue dont les sollicitations peuvent être associées à un phénomène aléatoire, il serait intéressant de définir la loi d'initiation par l'étude des dates de premières apparitions des fissures sur des sections de chaussées homogènes, $T_{(1),j} = \min\{T_{i,j}, i \in \mathbb{N}, j \in \mathbb{N}\}$ avec $T_{i,j}$ la date d'apparition de la i ème fissure sur le tronçon j .

Concernant la caractérisation de la phase de propagation de $\rho(t)$, $\rho(t) > 0$, on peut la décomposer en 3 étapes illustrées sur la Figure 3.1(b) :

- L'étape 1 est le premier tronçon caractérisé par une accélération du pourcentage de fissuration. Il est principalement lié à l'apparition et au développement des premières fissures. Celles-ci sont en moyenne peu nombreuses et de faibles tailles. Ainsi, tout accroissement ou toute nouvelle fissure a une forte probabilité de contribuer à l'accroissement de $\rho(t)$;
- L'étape 2 est la phase de transition comprenant un point d'inflexion ;
- L'étape 3 voit l'accélération de $\rho(t)$ diminuer. Le nombre de fissures étant très important, tout développement de fissures existantes ou toute nouvelle fissure a une très forte probabilité de venir se superposer aux fissures existantes et ne contribue donc peu à l'accroissement de la fonction. Avec le temps, sans intervention, l'ensemble de la section va être fissurée et $\rho(t) = 1$.

La forme générale de l'évolution de $\rho(t)$ est une courbe dite en " S " et fonction des caractéristiques intrinsèques de la chaussée (composition de l'enrobé, épaisseur des couches de fond, géologie des sous-sols, ...) et des données extérieures (charge du trafic, climat, ...). Enfin, les données de la base IQRN montrent des comportements présentant de grandes variabilités suivant les sections, ceci pouvant facilement être expliqué au regard de la grande variabilité des conditions de construction et d'utilisation des sections. Il est clair qu'une grande part d'incertitude est à prendre en compte dans les processus de dégradation due notamment à des zones de non connaissance sur le processus physique de dégradation lui-même, la grande variabilité des propriétés des matériaux utilisés, des conditions environnementales et de leur impact sur la dégradation, les conditions de trafic et les contraintes résultantes [117]. Ce constat est sûrement à rapprocher aux limites des approches fiabilistes reposant sur les modèles physiques de dégradation.

Par ailleurs, le processus de naissance et de développement des fissures longitudinales en sous-couche souligne clairement la nécessité d'analyser la relation entre le phénomène réel de détérioration et son mode par lequel cette détérioration est observée. On rejoint l'idée de *Markers* développée initialement dans le contexte bio-médical et reprise dans un contexte de systèmes industriels par Singpurwalla [97]. Dans ces travaux, Singpurwalla associe à chaque observation une mesure du *potentiel de danger* défini par la fonction de hasard cumulée résiduelle. Les modèles développés dans ce cadre, si l'on ajoute les modèles de type *Modèles de Hasard Proportionnel*[28], cherchent à mettre en relation covariables observées et défaillance. Or, dans notre contexte, c'est directement le processus de fissuration sous-jacent de la couche de fond et l'indicateur de fissuration en surface que nous devons prendre en compte. D'autres approches comme les modèles de Markov cachés [112] supposent une loi connue reliant observation et état réel, l'objectif étant alors d'estimer l'état réel en fonction des observations. Après avoir proposé une classification de processus associés définis par une fonctionnelle de processus sous-jacent, Newby, dans [75], propose une structure unifiée pour la décision lorsque la dégradation décrite par un processus X_t peut être associée à un processus Y_t observée ou reconstruite.

Notre approche consiste à modéliser la détérioration par un processus bivarié dépendant. La première composante, ρ_t représente la valeur de l'indicateur en surface alors que la seconde, notée θ_t , reflète son potentiel de fissuration instantané. Cette dernière variable permet de rendre compte d'un niveau estimé de fissuration de la sous-couche de la chaussée considérée et donc de définir la vitesse espérée d'évolution de la fissuration en surface. On peut ainsi différencier 2 sections considérées initialement identiques qui auraient même valeur d'indicateur de fissuration mais des vitesses d'évolution différentes. On définit plus précisément θ_t comme le taux espéré d'accroissement de $\rho(t)$. Cette modélisation de la dégradation ne s'intègre pas directement dans la classification proposée dans [75] car le processus ρ_t peut être certes considéré comme un processus associé à θ_t mais θ_t est lui aussi dépendant de ρ_t .

Soient $\Delta\rho_{(\rho,\theta)}(t, t + \tau)$ et $\Delta\theta_{(\rho,\theta)}(t, t + \tau)$ les incréments respectifs de fissuration et de potentiel sur l'intervalle $(t, t + \tau)$ connaissant les valeurs des processus à l'instant t . On note $h_\tau(x, y | \rho, \theta)$ la densité de transition sur l'intervalle de longueur τ . On peut alors montrer que cette densité s'exprime comme le produit des densités de probabilité conditionnelle respectives

de chacun des incréments $g_\tau(x | \rho, \theta)$ et $f_\tau(y | \rho, \theta, x)$, i.e. :

$$h_\tau(x, y | \rho, \theta) = g_\tau(x | \rho, \theta) \cdot f_\tau(y | \rho, \theta, x) \quad (4.1)$$

On se propose de ramener l'expression de la variation de potentiel $\Delta\theta$ en fonction de l'état passé (ρ, θ) et du nouveau niveau de fissuration $\rho + x$, à la seule expression de $\Delta\theta$ en fonction des niveaux de fissuration passé ρ et actualisé $\rho + x$ (l'information θ est contenue dans la valeur x). On considère alors $f_\tau(y | \rho, x)$ pour les incréments de potentiel d'accroissement de fissuration observable.

On notera que la modélisation proposée ici permet d'actualiser la loi du processus de fissuration en fonction des observations. Ceci permet donc d'intégrer les particularités de chacune des sections par rapport à un comportement moyen. A contrario, il faut être conscient qu'un tel choix conduit à accroître les difficultés d'estimation des lois de probabilité notamment par le nombre de données nécessaires.

L'objectif des paragraphes suivants est la construction des densités de probabilité conditionnelle $g_\tau(x | \rho, \theta)$ et $f_\tau(y | \rho, x)$ pour le processus bivarié de fissuration.

4.1 Processus gamma dépendant de l'état

Avant de définir les lois du processus bivarié $\{(\rho_t, \theta_t), t > 0\}$ à chacune des phases de fissuration, nous allons définir un nouveau processus : le processus gamma dépendant de l'état. La motivation de ce nouveau processus est liée à la possibilité d'adapter sa loi d'évolution, définie généralement d'une manière globale sur l'échantillon complet, au comportement singulier de chacun des systèmes étudiés. Dans le cas de la gestion des chaussées, même au vu du grand nombre de sections considérées, il est difficile d'établir des échantillons homogènes tant les aspects dégradation sont liées à de très nombreux facteurs qui peuvent facilement être différenciés d'une section à une autre (par exemple, les conditions de construction ou d'entretien ou bien encore la géologie, etc.). On définit le processus gamma dépendant de l'état comme un processus gamma pour lequel sa fonction de forme est dépendante de l'état courant. Plus formellement, on donne la définition suivante :

Définition 1 : *Un processus gamma dépendant de l'état G est un processus stochastique qui vérifie :*

- (i) $G_0 = 0$;
- (ii) $G_1 - G_0, G_2 - G_1, G_3 - G_2, \dots$ sont indépendants ;
- (iii) Pour un intervalle de temps fixé de longueur $\tau > 0$, la distribution des incréments $G_{t+\tau} - G_t$ dépend uniquement de G_t et est une distribution gamma de paramètre d'échelle β et de paramètre de forme dépendant de G_t et de τ : $\gamma(\tau, G_t) = \int_0^\tau \alpha(s, G_t) ds$ où $\alpha(s, G_t)$ est une fonction non décroissante en s .

Ce processus est stationnaire par rapport au temps mais par contre il n'est pas infini divisible, propriété des processus de Lévy. La preuve de ces lemmes est proposée dans [127].

Afin de rendre le processus infini divisible, on propose de considérer le paramètre de forme de G fonction de la variation à droite espérée de G_t , i.e. $\gamma(\tau, G_t, \theta_t) = \int_0^\tau \alpha(s, G_t, \theta_t) ds$ où $\alpha(s, G_t, \theta_t)$ reste une fonction croissante en s . La distribution conditionnelle de $G_{t+s} - G_t$ sachant G_t et θ_{t+u} , $u \in [0, s]$ est infinie divisible. En choisissant une distribution infinie divisible pour le processus $\{\theta_t, t > 0\}$, la distribution du processus bivarié est alors infinie divisible.

4.2 Processus de détérioration en phase d'initiation

Durant la phase d'initiation, aucune observation en surface n'est observable, i.e. $\rho_t = 0$. Par contre, liée aux contraintes répétitives sur la chaussée dues au trafic, la valeur de θ définie ici comme potentiel de fissuration augmente de manière continue tant qu'il n'y a

pas relâchement par effet de fissuration. On peut supposer que la distribution de la date de première fissuration est fonction de ce potentiel de fissuration. On propose de modéliser le processus d'apparition des fissures en surface par un processus de Poisson homogène par morceaux de fonction d'intensité $\lambda(\tau, \theta)$ où τ est la longueur de l'intervalle de temps considéré et θ la valeur du potentiel au début de cet intervalle. On modélise l'évolution du potentiel avant fissuration par un processus gamma stationnaire, i.e. l'accroissement du potentiel sur un intervalle de temps de longueur τ est distribué suivant une loi gamma de paramètres de forme $\tilde{\gamma}_p(\tau)$ et d'échelle β donnés (ces données doivent être issues d'un retour d'expérience et renforcées par analyses d'experts. Nous ne chercherons pas dans ce travail à définir les méthodes d'estimation associées à ces paramètres²).

4.3 Processus de détérioration en phase de propagation

Une fois la fissure initiée, l'indicateur de fissuration de surface ne peut que croître au cours du temps jusqu'à la valeur 1 représentant la fissuration totale. On propose de modéliser cet indicateur par un processus gamma dépendant de l'état (ρ, θ) . Ainsi, si à t , on connaît les valeurs de ρ_t et θ_t , l'accroissement du pourcentage de fissuration sur l'intervalle de longueur τ est distribué suivant une loi gamma tronquée de densité de probabilité :

$$g(x, \tau; \rho_t, \theta_t) = \left(\frac{\beta \gamma(\tau; \rho_t, \theta_t)}{\Gamma(\gamma(\tau; \rho_t, \theta_t))} x^{\gamma(\tau; \rho_t, \theta_t) - 1} e^{-\beta x} \right) \mathbb{I}_{\{x < 1 - \rho_t\}} + \left(1 - \int_0^x \frac{\beta \gamma(\tau; \rho_t, \theta_t)}{\Gamma(\gamma(\tau; \rho_t, \theta_t))} u^{\gamma(\tau; \rho_t, \theta_t) - 1} e^{-\beta u} du \right) \mathbb{I}_{\{x = 1 - \rho_t\}} \quad (4.2)$$

La fonction $\gamma(\tau; \rho, \theta)$ définit au coefficient $1/\beta$ près la valeur moyenne d'un incrément de fissuration sur un intervalle de longueur τ et possède les propriétés suivantes :

- $\gamma(\tau; \rho, \theta) = 0$ pour $\rho = 1$;
- $\gamma(\tau; \rho, \theta) \leq 1, \forall (\rho, \theta)$;
- Elle est croissante puis décroissante en fonction de ρ

Contrairement à l'indicateur de fissuration de surface ρ , les variations $\theta_{t+\tau} - \theta_t, \forall t > 0$ ne sont pas strictement positives, la vitesse d'un intervalle de temps à l'autre peut décroître tout en restant positive. Nous avons proposé de modéliser le processus $\{\theta_t, t > 0\}$ par un processus bilatéral gamma [48]. Ce processus permet de capturer ces variations positives et négatives et peut-être défini, sous condition, comme la différence de deux processus gamma que nous noterons G_n et G_p . En ce sens, la variation des θ_t et $\theta_{t+\tau}$ peut être vue comme la différence des vitesses de deux processus gamma définis sur deux périodes de temps consécutives de paramètres de forme respectifs $\gamma(\tau; \rho_t, \theta_t)$ et $\gamma(\tau; \rho_{t+\tau}, \theta_{t+\tau})$. On définit alors les paramètres de forme γ_n et γ_p des processus G_n et G_p par les vitesses moyennes de dégradation espérées respectives sur la période de temps considérée vérifiant :

$$\gamma_p(\tau; \rho_{t+\tau}) - \gamma_n(\tau; \rho_t, \rho_{t+\tau}) = \frac{\mathbb{E}(\Delta\rho(t + \tau, t + 2\tau) \mid \rho_t, \rho_{t+\tau}) - \mathbb{E}(\Delta\rho(t, t + \tau) \mid \rho_t, \rho_{t+\tau})}{\tau} \quad (4.3)$$

L'évaluation des espérances de l'équation (4.3) ne peuvent pas être directement évaluées. Nous proposons de les approximer en définissant les fonctions γ_n et γ_p sachant les valeurs ρ_t et $\rho_{t+\tau}$ respectivement par :

$$\gamma_n(\tau, \rho_t, \rho_{t+\tau}) = \frac{\gamma(\tau, \rho_t, \tilde{\theta}(\rho_t, \rho_{t+\tau}))}{\tau} \quad (4.4)$$

$$\gamma_p(\tau, \rho_{t+\tau}) = \frac{\gamma(\tau, \rho_{t+\tau}, \tilde{\theta}(\rho_{t+\tau}))}{\tau} \quad (4.5)$$

2. Cette étude est d'ores-et-déjà initiée et a fait l'objet de deux communications [124, 125]. Cependant, la pertinence des propositions n'a pas pu être validée sur les données réelles.

où $\tilde{\theta}(\rho_t, \rho_{t+\tau}) = (\rho_{t+\tau} - \rho_t)/\tau$ est la valeur actualisée de θ_t sachant l'accroissement de fissuration sur l'intervalle considéré et $\bar{\theta}(\rho_{t+\tau}) = \int_{\theta_0}^{\theta_{\max}} \gamma(\tau, \rho_{t+\tau}, \theta) d\theta / (\beta \cdot (\theta_{\max} - \theta_0))$ est la vitesse moyenne de fissuration en supposant aucune information a priori sur la loi de θ .

La Figure 3.2 permet de synthétiser l'ensemble des lois nécessaires au modèle de fissuration. À t_1 , sachant une non-fissuration, le potentiel croissant est modélisé par le biais d'une loi gamma. À t_2 , sachant un niveau de fissuration et la valeur du potentiel d'accroissement, on modélise l'incrément de fissuration par le biais de la loi gamma dépendante de l'état puis on définit la nouvelle valeur θ_{t_2+v} pour l'intervalle de temps suivant à l'aide de la loi gamma bilatérale fonction uniquement des valeurs de fissuration observées ρ_2 et $\rho_2 + \Delta\rho$.

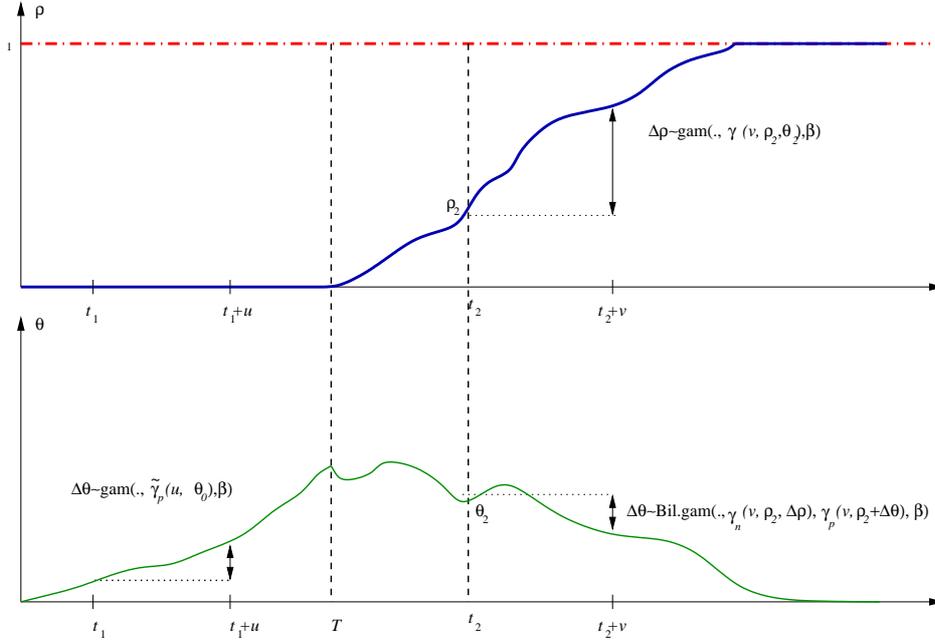


FIGURE 3.2 – Illustration du modèle de fissuration (ρ, θ)

5 Modélisation de l'entretien des chaussées

Après avoir explicité la pathologie retenue et l'indicateur associé, nous proposons dans ce paragraphe de discuter des actions d'entretien qui permettent de remettre la chaussée dans des états plus conformes. Les approches classiques rencontrées dans le contexte de l'optimisation de la maintenance se concentrent majoritairement sur des actions dites de renouvellement. À l'issue d'une telle action, le système maintenu est considéré comme neuf et ses caractéristiques moyennes de dégradation se retrouvent à l'identique. Dans le contexte que nous allons décrire et au vue de la multiplicité et diversité des actions d'entretien, il sera nécessaire d'étendre cette hypothèse et rejoindre les approches dites *maintenance imparfaite* [113], généralement présentées dans un cadre durées de vie et peu pour les approches conditionnelles.

Nous proposons ici une première classification qui correspond soit aux actions de mesure de l'état de la chaussée, soit aux actions de remise en l'état de celle-ci. Dans chacune de ces classes, nous proposerons d'identifier des actions type qui permettront de réduire le nombre d'actions à prendre en compte dans notre étude.

5.1 L'inspection

Les inspections sont définies par la mise en place de techniques de mesure permettant d'avoir un retour d'information quant au niveau de fissuration de la section de route. On se limite dans cette étude aux techniques visuelles qui permettent de détecter d'éventuelles

fissures de surface et d'en mesurer leur taille. Seul le PFL est actuellement fourni. On considère que cette technique non destructive visuelle est parfaite dans le sens où elle rend l'exacte valeur du PFL. La mise en place des techniques de mesure est réalisée lors de campagnes communes d'auscultation dont il reste nécessaire d'en estimer le coût unitaire que l'on supposera constant (du moins dans un premier temps). Ce coût unitaire ne doit pas uniquement comporter le simple coût d'acquisition de l'information mais doit aussi intégrer les aspects disponibilité du matériel et faisabilité du programme d'inspection -on entend par là que, puisque l'acquisition de l'information de dégradation continue dans le temps n'est pas envisageable, son coût doit alors, dans cet exemple, être infini.

5.2 Les actions d'entretien

L'objectif de l'entretien est la conservation des performances fonctionnelles (structurelles et de surface) de la route dans le temps. On peut résumer cet objectif par la remise de la chaussée dans un état jugé conforme. Une action d'entretien pour une pathologie donnée est caractérisée par son efficacité et son coût. On propose une classification en quatre groupes en fonction de l'efficacité des techniques utilisées : ponctuels, légers, semi-lourds et lourds. Dans notre étude, on ne retiendra qu'une seule technique par classe à savoir :

- Le **colmatage** des fissures (catégorie entretiens ponctuels = catégorie 1) est effectué, en l'absence de désordre secondaire, pour des fissures dont la longueur reste faible. En terme d'efficacité, le colmatage permet de rétablir la chaussée dans un état neuf, sans fissure de surface, d'une manière relativement temporaire. On peut alors considérer que, dans un premier temps, la fissure colmatée disparaît et que seules ses extrémités continuent de se développer, puis, dans un second temps, la partie centrale de la fissure réapparaît dans sa totalité.
- Le **recouvrement par un BBTM** (catégorie entretiens légers = catégorie 2) consiste à mettre une couche d'enrobé inférieure à 6 cm. Il a comme conséquence de faire disparaître, au moins d'une manière temporaire, les fissures de surface. Cette action est bien entendu plus performante que le colmatage dans le sens où le délai d'apparition des fissures en surface est significatif.
- Le **recouvrement par un BB de plus de 6 cm** (catégorie entretiens semi-lourds = catégorie 3) possède les mêmes propriétés qu'un recouvrement par un BBTM quant à l'effet sur les fissures de surface. Il permet aussi de rendre plus rigide la structure. Lors de la répétition de cet entretien, la chaussée devenant de plus en plus rigide, n'est plus sujette à la fissuration longitudinale. Cependant des contraintes de dépassement d'épaisseur cumulée de la chaussée ne permettent pas de multiplier à l'infini cette action. Il est à souligner ici que les effets de réapparition des fissures ne sont pas linéaires et que l'impact d'un recouvrement par un BB d'épaisseur donnée est meilleur que celui de plusieurs recouvrements par BB d'épaisseur cumulée équivalente répartis dans le temps.
- Le **décaissement** (catégorie entretiens lourds = catégorie 4) consiste à remplacer une épaisseur donnée de l'enrobé par une nouvelle couche d'épaisseur donnée. Ainsi, toutes les fissures existantes disparaissent en surface mais le niveau de fissuration du fond de couche non réparé reste inchangé et va donc continuer son processus à une vitesse fonction de l'épaisseur de la nouvelle couche (si celle-ci reste inchangée, la fissuration en fond de couche se développera à la même vitesse qu'auparavant). Pour la vitesse de remontée au travers de la nouvelle couche, celle-ci est fonction du niveau de fissuration du fond de couche et de l'épaisseur de la nouvelle couche. Ainsi, les vitesses d'initiation et de propagation de la fissuration en surface augmentent, pour des couches d'entretien d'épaisseur constante, après chacun des entretiens successifs (on pourra cependant considérer des vitesses limites propres à la mécanique déterministe des processus de remontée des fissures).

Afin de préciser les effets du décaissement sur l'évolution de la fissuration dans le temps, on se propose de commenter un scénario d'entretien. La Figure 3.3 représente l'évolution de la fissuration au cours du temps pour une chaussée d'épaisseur initiale e et pour laquelle un décaissement d'une largeur e_1 est réalisée (sous-figure 3.3(c)). Dans ce cas, la fissuration en fond de couche reste à son niveau et va ensuite se propager dans la nouvelle couche jusqu'à la surface.

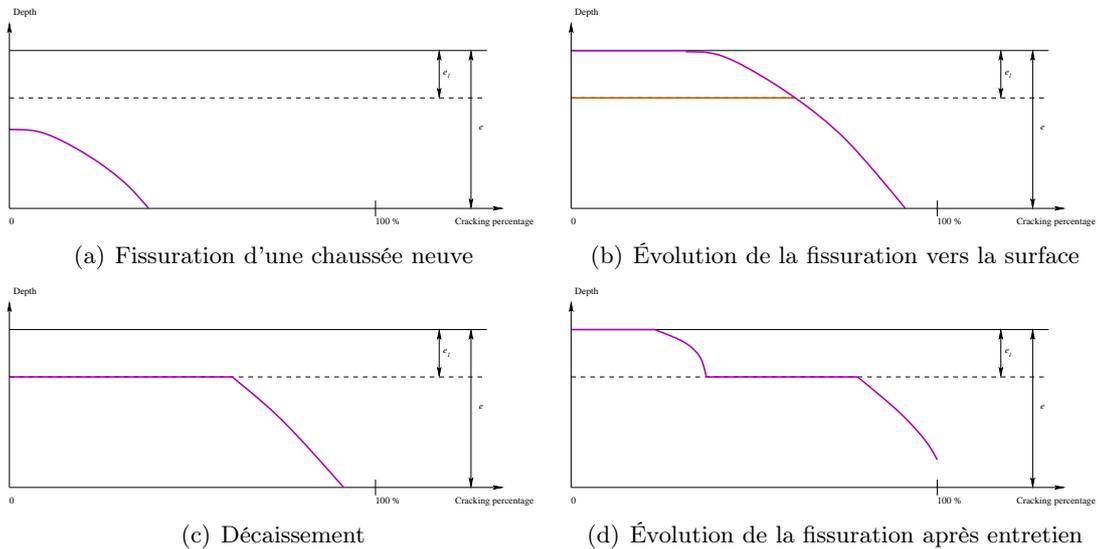


FIGURE 3.3 – Schématisation de l'évolution du pourcentage de fissuration du fond de couche vers la surface de la chaussée avec un décaissement d'épaisseur fixée

Par ailleurs, pour chacune des actions considérées, on suppose que leur mise en place est effective sur la longueur totale du tronçon considéré. De plus, le temps d'entretien est considéré négligeable devant le temps d'exploitation de la chaussée. Enfin, la caractérisation des coûts associés à chacun des entretiens n'est pas simple car ils doivent comprendre l'ensemble des coûts directs -coûts de matière, coûts de mise en place et d'équipements, etc., et les coûts indirects intégrant notamment les coûts d'immobilisation et de gêne aux usagers. Néanmoins, on se propose de considérer des coûts unitaires constants et croissants en fonction du type d'entretien.

On peut souligner que l'efficacité d'un entretien ne doit pas uniquement se mesurer par rapport à la qualité " immédiate " de la chaussée après entretien (hypothèse classique dans le cadre de l'optimisation de la maintenance) puisque l'ensemble de ceux-ci conduit à la disparition temporaire de la fissuration. L'efficacité d'une action va donc dépendre à la fois de la catégorie d'entretien réalisé et aussi du niveau de fissuration, et même du nombre et de la taille des fissures existantes, auquel il aura été réalisé, puisqu'il conditionne l'évolution future de la fissuration de la chaussée. Cette remarque est illustrée dans la Figure 3.4 qui simule l'évolution de l'indicateur de fissuration $\rho(t)$ pour une section de route donnée soumise à différentes actions d'entretien dans le temps.

On considère ici, à l'instant $t = 0$, que la section est neuve. Une première fissure apparaît ce qui a comme incidence d'augmenter le pourcentage de fissuration $\rho(t)$. Sa vitesse de propagation peut évoluer en fonction notamment de sa taille. Une deuxième fissure apparaît. Elle peut avoir comme conséquence sur $\rho(t)$ d'augmenter sa vitesse. À T_1 , le niveau de fissuration est considéré suffisamment important pour mettre en place une action. Elle permet de rétablir la condition en surface $\rho(t) = 0$ mais en conservant le processus de fissuration du fond de couche qui va continuer à évoluer (ceci est illustré par les courbes discontinues sur la figure du haut). Une fissure existante peut réapparaître. Sa vitesse de propagation est généralement très élevée avant de retrouver une valeur plus classique. Ceci est approximé sur la figure par l'hypothèse de l'apparition à sa longueur supposée si aucune action n'avait été réalisée. Cette

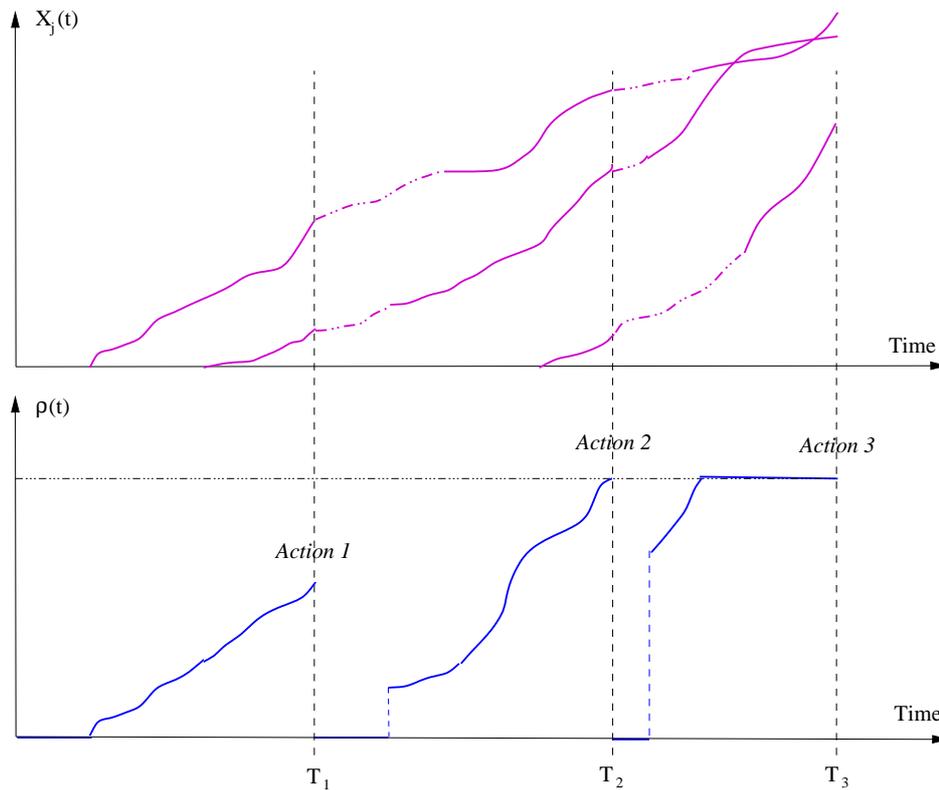


FIGURE 3.4 – Schéma d'évolution du processus de fissuration et du pourcentage de fissuration longitudinale pour une section de route maintenue

hypothèse peut être justifiée si les périodes inter-inspections sont grandes.

5.3 Modélisation de l'effet d'entretien sur le processus (ρ, θ)

Dans un contexte maintenu, le processus de fissuration $\{(\rho_t, \theta_t), t > 0\}$ est uniquement observé lors des inspections. On ne retiendra dans ce travail qu'un schéma périodique d'inspection, de période fixe que l'on cherchera à optimiser. Cette hypothèse, même si elle conduit à des stratégies de décision sous-optimales par rapport, par exemple, à des stratégies conditionnelles pour lesquelles la périodicité s'adapte en fonction de l'état, permet d'une part de réduire le nombre de paramètres de décision et d'autre part de faciliter la planification dans des contextes décisionnels plus larges (par exemple, au niveau du réseau). On note τ la période d'inspection. De plus, on suppose l'inspection parfaite dans le sens où elle est non destructrice -inspection visuelle- et rend la vraie valeur de ρ . On note par contre que l'observation de θ n'est pas possible. À l'issue d'une inspection, le manager peut décider de ne rien faire, et donc de laisser la chaussée en l'état, soit de mettre en place un des entretiens.

Si on ne considère pas les différents enrobés, un entretien est caractérisé par son épaisseur de décaissement e_1 (e_1 pouvant être nul dans le cas des actions de recouvrement - classes 2 et 3 définies dans le paragraphe 5.2) et de rechargement e_2 . Un entretien, quelle que soit sa nature, permet de remettre la surface de la chaussée à neuf, i.e. $\rho = 0$. Par contre, seule l'épaisseur de la sous-couche traitée peut être considérée comme neuve. Ceci a comme conséquence de réduire le potentiel de fissuration sans obligatoirement le ramener à sa valeur lors de la mise en service de la chaussée, notée θ_0 , des valeurs plus petites que θ_0 pouvant être permises. En effet, à ce stade, aucune contrainte d'épaisseur n'étant donnée, le potentiel pourrait tendre vers zéro. Par ailleurs, on peut supposer une dépendance de l'efficacité d'un entretien en fonction du niveau de l'indicateur de fissuration de surface avant sa mise en place. En effet, pour des θ du même ordre mais de niveaux de fissuration en surface différents, on s'attend à ce que deux

chaussées possédant les mêmes caractéristiques de dégradation soumises au même entretien ne montrent pas le même comportement.

Pour modéliser les effets de l'entretien sur l'état de dégradation, on introduit une fonction *maintenance* $\phi_m(\rho, \theta)$ qui détermine l'état de dégradation à l'issue d'un entretien de nature m ayant observé un niveau de dégradation (ρ, θ) avant entretien, la nature de l'entretien étant définie par les épaisseurs respectives de décaissement et de rechargement effectuées et l'épaisseur résiduelle à l'issue de l'entretien. On peut montrer que les épaisseurs de rechargement et résiduelle, associées à l'état de la chaussée, caractérisent complètement le processus d'entretien. On définit $\phi_m(\rho, \theta)$ comme suit :

$$\phi_m(\rho, \theta) = \begin{cases} (\rho, \theta) & \text{si } action = DN, \\ (0, \theta_m(\rho, \theta)) & \text{si } action \in \{MX_m\} \end{cases} \quad (5.1)$$

où DN correspond à l'action "Ne rien faire" et MX_m à des actions d'entretien. On choisira $\theta_m(\rho, \theta)$ une fonction à valeurs dans $(0, \theta)$ et croissante en ρ et θ . Par souci de simplification, on choisira un nombre fini m d'actions d'entretien et une fonction $\theta_m(\rho, \theta)$ déterministe.

Par ailleurs, afin de modéliser un vieillissement lié à l'entretien de la chaussée, nous avons introduit un facteur d'accélération de dégradation r_m , fonction de l'entretien. Le facteur d'accélération est un paramètre multiplicatif introduit au niveau des fonctions de forme des lois gamma pour la fonction d'intensité du processus de Poisson dans la phase d'initiation et de l'incrément de fissuration (ceci a pour conséquence de changer les fonctions de forme des lois respectives des accroissements de potentiel, Eq. (4.4) et Eq. (4.5)). On peut alors montrer que le processus $\{(\rho_t, \theta_t, m), t > 0\}$ où m est la valeur du dernier entretien effectué sur la chaussée, $m \geq 1$, est un processus markovien.

6 Modélisation des coûts de maintenance

En termes de coûts, nous distinguons les coûts liés à l'entretien en soi et les coûts que l'on peut associer à des coûts d'exploitation. Les coûts d'entretien intègrent les coûts de mise en œuvre d'une action intégrant la mobilisation des moyens humains et opérationnels, les coûts de matières premières et d'opération mais aussi des coûts indirects liés à la gêne des usagers pour cause de travaux. Même si ces derniers peuvent être assez difficiles à estimer, les informations liés aux coûts d'entretien peuvent être justifiées. On supposera unitaires l'ensemble de ces coûts et fonctions de la nature de l'action envisagée.

Le coût d'inspection sera noté c_i . Dans le contexte d'exploitation, bien qu'à la vue des grandeurs des autres coûts engagés, le coût d'inspection peut paraître négligeable si l'on ne considère que l'acquisition des données. On peut cependant y intégrer les coûts d'acquisition de la technologie, du traitement de l'information. Par ailleurs, ce coût doit refléter l'ensemble des contraintes liées à l'inspection qui entraîne une limitation de sa mise en œuvre comme la disponibilité des moyens.

On suppose que les coûts d'entretien, notés c_m , sont décomposables en une partie fixe et une partie variable croissante en fonction de la nature de l'entretien.

Les coûts d'exploitation ne sont pas directement mesurables. Ils doivent représenter le niveau de satisfaction des usagers et des autorités locales, le niveau de sécurité mais aussi les surcoûts liés à un niveau de fissuration élevé comme la surconsommation des véhicules, les nuisances sonores, etc. Bien que difficiles à estimer, des études ont été réalisées pour essayer de caractériser ces coûts en fonction du niveau de dégradation des chaussées, notamment dans le cadre de l'estimation de la valeur économique du patrimoine routier. On suppose donnée la fonction de coût c_q par unité de temps d'usage et par unité de niveau de fissuration observé, nommée coût qualité. Ainsi, sur une période de temps τ , si ρ_1 et ρ_2 sont les niveaux de fissuration observés en début et fin de période, le coût qualité cumulé peut être approximé par $C_q(\tau, \rho_1, \rho_2) = \tau(c_q(\rho_1) + c_q(\rho_2))$. Nous soulignons ici l'importance de ce coût pour imposer à la politique optimale la préconisation des actions d'entretien. En effet, une des particularités

de notre problème est que la notion de défaillance n'existe pas : même très endommagée, on peut circuler sur une chaussée. Il est donc important de pénaliser le critère afin de contraindre la politique de maintenance.

7 Politiques de décaissements : problème restreint

7.1 Objectifs

On cherche à répondre ici à l'optimisation de la politique de maintenance multi-action dans un contexte restreint. On suppose en effet que les actions d'entretien sont limitées aux seuls décaissements pour lesquels l'épaisseur de rechargement en enrobé est égale à celle enlevée. L'épaisseur totale de la chaussée reste ainsi constante. Par ailleurs, on considère un nombre fini, d'entretiens différenciés par l'épaisseur entretenue, MX_k , $k \in \{1, \dots, 4\}$ classés par ordre croissant d'efficacité. Tout ceci nous permet de simplifier certaines notations, notamment les coûts c_m que nous considérerons constants en fonction de la nature ceux-ci. L'action qui consiste à décaisser l'ensemble de la chaussée permet de remettre la chaussée dans son état initial, dit *aussi bon que neuf* pour lequel $\rho = 0$ et $\theta = \theta_0$.

7.2 Critère de coût

Comme stipulé dans la section 2, le critère de décision choisi est la minimisation d'un coût moyen escompté sur un horizon infini. Cette optimisation conduira à la proposition de la politique optimale π^* , politique étant définie, pour chaque état (ρ, θ, m) de la chaussée, par l'épaisseur de la sous-couche à traiter en fonction de la période optimale d'inspection τ^* . Sous les hypothèses présentées, la fonction de coût escompté s'écrit, pour un facteur d'escompte $\lambda \in (0, 1)$:

$$V_\tau^{\pi^*}(\rho, \theta, m) = c_i + \min(DN(\rho, \theta, m), MX_k(\rho, \theta, m), k \in \{1, \dots, 4\}) \quad (7.1)$$

où

$$\begin{cases} DN(\rho, \theta, m) = \int_0^{1-\rho} \int_{-\theta}^{\theta_{\max}-\theta} h_\tau(x, y \mid \rho, \theta, m) [C_q(\rho, \rho + x, \tau) + \lambda V_\tau^\pi(\rho + x, \theta + y, m)] dy dx \\ MX_k(\rho, \theta, m) = c_k + DN(0, \phi_k(\rho, \theta), k), \forall k \in \{1, \dots, 4\} \end{cases} \quad (7.2)$$

Les équations (7.2) donnent les coûts respectifs des actions *Ne rien faire* et des décaissements d'épaisseur k , le nombre d'épaisseurs considérés étant volontairement limité à 4. La fonction coût *Ne rien faire* est composée de la loi d'évolution de la dégradation sur l'intervalle de temps à venir associée au coût qualité respectif auquel on ajoute le coût escompté à venir. La fonction de coût lié à un décaissement est exprimée ici comme la somme du coût unitaire de l'action, fonction de l'épaisseur traitée, et des coûts d'évolution en considérant la chaussée dans son nouvel état après entretien.

On souligne ici que l'action MX_4 correspond à un renouvellement total de la chaussée pour lequel l'épaisseur de chaussée traitée est égale à l'épaisseur totale de la chaussée.

7.3 Procédure de résolution numérique

La procédure de résolution numérique proposée repose sur la discrétisation des espaces des solutions associée à un algorithme classique de recherche itérative des solutions, le *Policy Iteration Algorithm*, *PIA* [90].

On effectue une discrétisation des variables d'intérêt, à savoir le temps sur un horizon suffisamment grand, l'intervalle des niveaux de fissuration de surface entre $0 \times (0, 1]$ et l'intervalle du potentiel de fissuration que nous proposons de borner par θ_{max} qui correspond

à la valeur maximale de fissuration tolérée sur un intervalle de temps. On note l le pas de discrétisation. Nous ne détaillerons pas ici les formules d'approximation pour l'évaluation des probabilités de transition $P_{ss'}$ [127]. L'idée repose sur la décomposition de $h_\tau(x, y \mid \rho, \theta)$ par la loi de Bayes proposée dans l'Equation (4.1), l'identification des lois associées aux scénarios conduisant à l'observation ou non à une fissuration. À titre d'illustration, on donne ici la première probabilité associée au scénario 1 " Aucune fissure n'apparaît dans l'intervalle de décision " :

$$\begin{aligned} P_{ss'} &= Pr(\Delta\rho = 0, \Delta\theta = \theta' - \theta \mid \rho = 0, \theta') \\ &= \exp(-\Lambda(\tau, \theta, m)) \cdot \int_{\theta' - \theta}^{\theta' - \theta + l} \tilde{g}(y; \tilde{\gamma}_p(\tau; \theta_0, m), \beta) dy \end{aligned} \quad (7.3)$$

où $\tilde{g}(y; \tilde{\gamma}_p(\tau; \theta_0, m), \beta)$ est la densité de probabilité normalisée de la loi gamma de paramètre de forme $\tilde{\gamma}_p(\tau; \theta_0, m)$ et de paramètre d'échelle β . L'évaluation du scénario 2 " la première fissure apparaît dans l'intervalle, à la date $T \in (0, \tau)$ " requiert l'analyse du scénario 1 jusqu'à la date T , puis l'analyse du scénario d'accroissement de la fissuration associée à l'évolution du potentiel. Bien que consommatrices en temps de calcul, les différentes intégrales peuvent être évaluées par des méthodes d'intégration numérique classiques.

L'algorithme d'optimisation que nous proposons consiste à rechercher pour une période d'inspection fixée les actions optimales à mettre en place en fonction de l'état de dégradation par l'application directe du *PIA*. La première valeur de τ est la première valeur fournie par la discrétisation et on incrémente sa valeur tant que le maximum de gain obtenu pour tout état croît en fonction de la période de décision, i.e. $\max\{V_\tau^{\pi^*}(s), s \in [0, 1] \times [\theta_0, \theta_{max}] \times [1, 4]\} \geq \max\{V_{\tau+l}^{\pi^*}(s), s \in [0, 1] \times [\theta_0, \theta_{max}] \times [1, 4]\}$. L'arrêt de l'algorithme fournit alors directement la période d'inspection optimale associée aux décisions d'entretien optimisées.

7.4 Application numérique

Nous proposons ici d'appliquer l'ensemble des modèles mathématiques sur un cas générique d'entretien routier. L'application sur des données réelles n'a pas été possible au vu de la qualité de l'information contenue dans la base IQRN. Par contre, au regard des modèles choisis, nous avons développé des modèles d'estimation paramétrique reposant sur les approches classiques du Maximum de Vraisemblance [124, 125]. Afin d'illustrer la politique de décision, nous présentons ici un exemple de matrices de décision obtenues pour les données résumées dans la Table 7.4.

		Dégradation		Maintenance		Coût	
ρ	Ini.	$\Lambda(\tau; \theta, m)$	$r_m(b_1\tau + \theta)$	k	4	c_i	100
		b_1	1.2	r_1	3	c_1	200
	Propag.	γ	$r_m \left(\frac{a_1\theta\tau}{(1-\theta) + \left(\frac{\theta}{\theta} - (1-\theta)\right)^2} + a_2\tau^2 \right)$	r_2	1.75	c_2	450
		(a_1, a_2)	(1.5, 0.9)	r_3	1.33	c_3	1080
		β	2.5	r_4	1	c_4	2950
θ	θ_0	0.03	$\theta_m(\rho, \theta)$	$r_m \left(\theta + \frac{\rho}{\tau} \right)$	c_q	c_q	600
	θ_{max}	1			λ	0.97	

TABLE 3.1 – Données génériques de dégradation, maintenance et coût pour l'obtention d'une politique optimale

Discussion sur les données : L'introduction de ces données dans les modèles de dégradation présentés dans la sous-section 4 permet de retrouver la forme dite en "S" de la loi de

fissuration. Celle-ci se traduit par des courbes spécifiques des fonctions de forme pour ρ et θ pour chacune des phases de propagation. Nous avons proposé dans [126] une discussion sur un ensemble de fonctions de forme et des performances attendues en termes de simulations de courbes de dégradation.

Politique optimisée et matrices de décision : La discrétisation proposée pour résoudre le problème d'optimisation est un pas $l = 0.1$ et les valeurs visitées pour la détermination de la période d'inspection optimale sont restreintes à l'intervalle $[1, 3]$ avec un pas de 0.2. La période optimale fournie par l'algorithme est $\tau^* = 1.8$ et la règle de décision optimale est donnée par les 4 matrices de décision définies en fonction du dernier entretien effectué sur la chaussée. À titre d'exemple, voici la matrice $\pi^*(s)$:

$$\pi^*(\rho, \theta, 1) = \begin{matrix} \rho \backslash \theta & 0.03 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \\ 0.9 & \left(\begin{array}{c} 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{array} \right. & \left(\begin{array}{c} 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{array} \right. & \left(\begin{array}{c} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \end{array} \right. & \left(\begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 2 \\ 2 \end{array} \right. & \left(\begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 2 \end{array} \right. & \left(\begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 2 \end{array} \right. & \left(\begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{array} \right. & \left(\begin{array}{c} 4 \\ 4 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{array} \right. & \left(\begin{array}{c} 4 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{array} \right. & \left(\begin{array}{c} 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 3 \\ 3 \end{array} \right) \end{matrix}$$

Chacune des matrices de décision est définie en fonction de la nature du dernier entretien, $m = \{1, 2, 3, 4\}$. La décision de maintenance est déterminée en fonction de la valeur de l'indicateur de fissuration observée ρ et la valeur θ estimée par la vitesse moyenne de dégradation observée depuis la dernière inspection, on rappelle que cette estimation est directement liée à la fonction de forme de la loi gamma associée lorsque aucune fissuration n'est observée. Les valeurs des matrices renvoient à l'action de maintenance à mettre en place, à savoir 0 pour " Ne rien faire ", 1 jusqu'à 3 pour des entretiens imparfaits MX_{1-3} et 4 pour un remplacement à neuf.

7.5 Discussions

Le premier constat que nous pouvons tirer de cet exemple est l'intérêt d'intégrer la notion de potentiel de fissuration dans notre modèle décisionnel. En effet, la décision d'entretien peut être différente pour une même valeur de l'indicateur de fissuration en surface, pouvant par exemple aller de ne rien faire jusqu'à un entretien relativement lourd.

Le second constat est relatif à une structure décisionnelle. En effet, alors qu'aucune hypothèse préalable n'a été posée en ce sens, les matrices présentent une certaine structure organisant la décision en fonction du niveau de dégradation. Bien que cela semble naturel, il resterait difficile de proposer une structure paramétrique pour la définition de la décision. Toutefois, nous soulignons que l'existence et l'identification d'une structure sont conditionnées aux valeurs des paramètres choisis.

Le troisième constat est la possibilité de faire des entretiens sur une chaussée ne présentant aucune fissure en surface. On peut noter un caractère préventif à ce type de décision.

Le quatrième constat est lié à l'évolution des décisions en fonction de la nature du dernier entretien. La décision de maintenance est sensible au dernier entretien effectué surtout pour des niveaux de dégradation intermédiaires : pour une chaussée considérée comme neuve, $m = 4$, les décisions de mettre en place une action d'entretien lourd MX_{1-4} sont repoussées par rapport à des chaussées ayant subi des maintenances au préalable.

Nous pouvons souligner que ces résultats ont été observés sur l'ensemble des expérimentations numériques que nous avons conduites (hors le quatrième constat). Nous avons par ailleurs cherché à déterminer les conditions pour l'obtention de ces propriétés structurelles pour les politiques optimales [127].

8 Politiques de décaissement et de chargement

8.1 Objectifs

Afin de rejoindre les pratiques usuelles de l'entretien routier, l'hypothèse de chargement est introduite. Cette hypothèse se décline sous deux formes, à savoir un chargement de la chaussée à épaisseur donnée avec ou sans décaissement préalable. L'épaisseur d'un rechargement à l'issue d'un décaissement reste à déterminer. On peut souligner que, dans la pratique, cette épaisseur est au moins celle du décaissement. Outre le rapprochement en termes d'actions d'entretien, cette hypothèse nous entraîne à considérer l'effet de l'épaisseur totale de la chaussée sur le processus d'évolution de la fissuration. L'amortissement de l'onde de choc entraînant la fatigue de la structure est proportionnel à l'épaisseur de la chaussée. Une conséquence sur l'aspect maintenance est la possibilité d'entretien qui améliore les performances du système par rapport à son état initial. Ainsi, on peut penser à une épaisseur limite de rechargement qui permettrait de ne plus considérer la fissuration longitudinale comme une pathologie d'intérêt sur la durée de vie de la chaussée. Dans notre étude, nous considérons que cette épaisseur limite ne peut pas être atteinte pour des raisons financières. Par ailleurs, il est clair que des contraintes d'épaisseur de rechargement maximal sont à prendre en compte en fonction des autres infrastructures (ponts, trottoirs, rails de sécurité, etc.). Nous noterons h_{max} cette contrainte d'épaisseur.

Outre l'optimisation de l'entretien en fonction des épaisseurs de décaissement et de chargement en enrobé, un autre objectif de cette partie est une meilleure prise en compte de l'incertitude sur le paramètre θ , taux d'accroissement instantané de fissuration espéré. Dans le premier temps, l'approche d'estimation pour θ consistait à considérer le taux d'accroissement moyen sur l'intervalle de décision mais un ensemble d'essais numériques nous a montré la sensibilité des politiques à la valeur de cette approximation. Son avantage résidait cependant dans la formulation du problème de maintenance par Processus de Décision Markovien. L'approche que nous utiliserons dans ce paragraphe consiste à étendre notre étude aux Processus de Décision Markovien Partiellement Observable (POMDP) [90].

Dans cette partie, nous focaliserons notre travail sur la modélisation de ces deux nouvelles hypothèses en considérant en premier lieu l'approche Processus Décision Markovien pour mettre en évidence les nouvelles hypothèses d'entretien. Ces travaux ont été développés dans [122, 126]. Dans un second temps [123], les effets liés à l'intégration d'incertitudes sur θ sur la base d'un Processus de Décision Partiellement Observable seront étudiés.

8.2 Modélisation de l'impact de l'épaisseur de la chaussée sur la fissuration

Jusqu'à présent, les variables d'intérêt pour la modélisation de l'évolution de la fissuration sont le niveau actuel de fissuration ρ , le potentiel instantané de fissuration, θ , l'épaisseur totale de la chaussée h ainsi que l'épaisseur traitée lors du dernier entretien, définie en fonction de sa nature m . Dorénavant, une action a est définie par le couple (ζ, ξ) représentant les épaisseurs respectives de décaissement et de rechargement. On notera que l'action $(0, 0)$ représente l'action *Ne rien faire*. À chaque opération (décaissement ou chargement), on associe un coût composé de coûts fixes propres à chacune des actions et des coûts variables en fonction des épaisseurs. Le coût d'une action peut s'écrire, pour des épaisseurs ζ et ξ données :

$$C(\zeta, \xi) = c_n \cdot \mathbb{I}_{\{\zeta > 0\}} + c_p \cdot \mathbb{I}_{\{\xi > 0\}} + c_{un} \cdot \zeta + c_{up} \cdot \xi.$$

Si on note l l'épaisseur totale de la chaussée, l'évolution de la fissuration dépendra de cette épaisseur (définissant notamment les propriétés mécaniques de rigidité et d'amortissement de la structure) ainsi que, pour une chaussée entretenue, de l'épaisseur traitée ξ . Les expressions régissant l'évolution de la fissuration restent alors inchangées (on remplace seulement m par le couple (l, ξ)).

La fonction efficacité $\phi(\cdot)$ doit permettre de retrouver les hypothèses suivantes :

1. Si aucune action n'est décidée, les valeurs de fissuration et de potentiel restent inchangées.
2. Si la chaussée est totalement décaissée, $\zeta = l$, l'évolution future de la fissuration ne dépend que des caractéristiques de la nouvelle couche.
3. Un entretien permet toujours d'améliorer les performances d'une chaussée.
4. Un décaissement d'épaisseur $\zeta \in (0, l)$ aura pour première conséquence, i.e. avant la mise en place d'un potentiel décaissement, de *dégrader* la chaussée. Cette *dégradation provisoire* est croissante en ζ .
5. Un chargement permet de réduire les effets de fissuration proportionnellement à l'épaisseur de chargement.

Nous proposons la forme suivante pour la fonction $\phi(\cdot)$, $\forall s = (\rho, \theta, l, \xi_m) \in S$:

$$\phi(s, a) = \mathbb{I}_{\{\zeta < l, l - \zeta + \xi \geq h_{max}\}} \frac{l}{l - b_1 \zeta + b_2 \xi} \theta + \mathbb{I}_{\{\zeta = l, \xi \geq h_{max}\}} \frac{h_{max}}{\xi} \theta_0 \quad (8.1)$$

où $0 < b_1 < b_2 < 1$ et θ_0 est la valeur minimale de potentiel estimée pour une épaisseur de chaussée maximale. La détermination des constantes b_1 et b_2 reste délicate et se doit de reposer sur des analyses d'experts (b_2 doit par exemple refléter le temps nécessaire à la remontée des fissures de la couche non traitée à la nouvelle couche).

Analyse de performance de la politique : La contribution majeure de la décomposition de l'entretien en deux phases est la mise en évidence des comportements des politiques en fonction des choix en conception en termes d'épaisseur initiale de la chaussée. En résumé des conclusions présentées dans [122, 126], outre la sensibilité des résultats aux fonctions d'efficacité d'entretien et de coût qualité, nous noterons, sous certaines hypothèses, la non convergence des résultats d'optimisation de notre approche vers les pratiques usuelles reposant sur l'avis d'experts. En effet, il n'est pas toujours optimal de renforcer la chaussée lors d'entretien, i.e. d'accroître graduellement l'épaisseur totale de la chaussée, mais pour des chaussées à épaisseur initiale médiane, sous les hypothèses de nos travaux, il peut s'avérer optimal de réduire cette épaisseur.

8.3 Prise en compte de la non-observabilité de θ

Afin de prendre en compte l'incertitude sur θ liée à son caractère non observable, on introduit une distribution de probabilité \mathbf{b}^θ sur les valeurs possibles de θ appelées *belief-states*. Les valeurs possibles de θ forment un espace discret et la distribution de probabilité un vecteur de probabilité. On définit alors l'état de la chaussée par $s = (\rho, \mathbf{b}^\theta, h, m)$ et le processus formé par ces états reste markovien. Cette modélisation est empruntée aux chaînes de Markov Cachées. En associant la décision de maintenance, on définit le Processus de Décision Markovien Partiellement Observable (POMDP).

L'avantage de cette approche est de moduler la décision en intégrant une connaissance d'expert définissant le vecteur de probabilités \mathbf{b}^θ pour tout ρ observé.

Sur le plan de la modélisation, il convient d'explicitier la relation entre la valeur réelle du potentiel θ et son estimation. On suppose cette relation, notée q_{so} , est donnée par une loi normale dont la moyenne est la vraie valeur θ et la variance est définie par la distance $|\theta - \tilde{\theta}|$. On construit alors la loi de probabilité de transition reliant un état initial possible

$\mathbf{b} = (\rho, b_\theta)$ à un état final observé $o = (\rho', \tilde{\theta})$, notée $\Omega(o|\mathbf{b})$. Le graphique de la Figure 3.5 permet d'expliquer la transition de \mathbf{b} vers l'état o par le biais de la transition de la *vraie* détérioration partiellement observée, i.e. de l'état $s = (\rho, \theta)$ vers l'état $s' = (\rho', \theta')$.

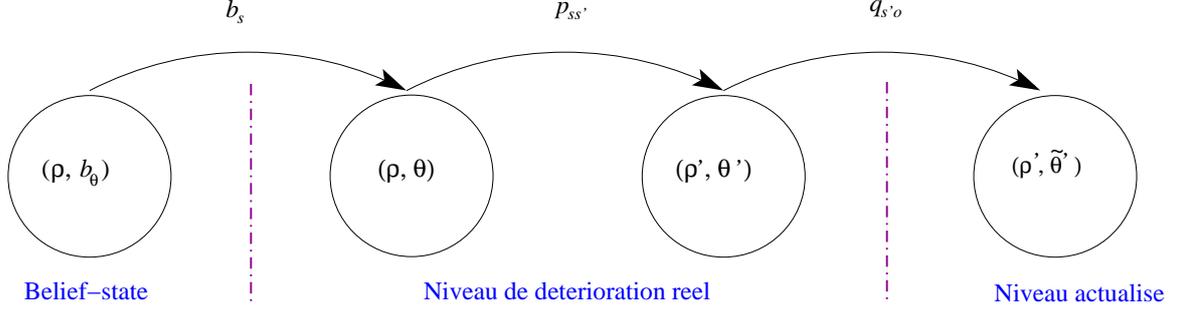


FIGURE 3.5 – Représentation d'une transition du processus de décision partiellement observable

On obtient alors la fonction de transition suivante :

$$\Omega(o|\mathbf{b}) = \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} b_s p_{ss'} q_{s'o} \quad (8.2)$$

La probabilité de la valeur actualisée du belief-state b' ayant observé l'état o alors que l'état est s' est donnée par application de la formule de Bayes :

$$\mathbf{b}'_{s'}(o, \mathbf{b}) = \frac{\sum_{s \in \mathcal{S}} b_s p_{ss'} q_{s'o}}{\sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} b_s p_{ss'} q_{s'o}} \quad (8.3)$$

L'expression du coût (7.1) est fonction de l'état défini, entre autre, par \mathbf{b}^θ et intègre, au niveau du coût de l'action *Ne rien Faire*, Equation (7.2), la probabilité de transition associée $\Omega(o|\mathbf{b})$.

D'un point de vue algorithmique, les difficultés principales sont liées à l'explosion combinatoire du problème d'optimisation due notamment à l'introduction des états de croyance. Des algorithmes non exacts permettent de contourner cette difficulté en restreignant l'évaluation de ces valeurs actualisées uniquement en certains points de croyance. On peut noter 2 classes d'algorithmes : les approximations de type à base de grille [58, 120] et à base de point [89, 119]. Après la sélection des points de croyance, la procédure d'actualisation des valeurs est classique. Dans notre travail [123], nous avons utilisé l'algorithme d'itération de valeur avec l'approximation régulière à base de grille proposée par Lovejoy [58].

9 Conclusions et perspectives

9.1 Conclusions

Je ne vais pas rappeler explicitement les contributions citées dans la section introductive de ce chapitre. Cependant, je voudrais souligner ici l'importance de confronter une recherche académique à vocation plutôt théorique à des problèmes complexes liés à une réalité industrielle. Sur ce champ, le projet SBADFORM a été une grande satisfaction personnelle traduite par des avancées intéressantes sur le champ théorique des processus stochastiques, sur le domaine de la modélisation et l'optimisation de la maintenance mais aussi consolide mon opinion sur l'intérêt de méta-modèles pour la maintenance des structures (je reviendrai sur ce point dans les perspectives ci-après) et le rapprochement d'intérêts de deux communautés scientifiques. Outre la production de modèles, les conclusions du projet soulèvent, sans en donner une réponse, la question de la pertinence de l'information contenues dans les bases actuelles de l'IFSTTAR pour la gestion de l'entretien des routes nationales en France ainsi que le maintien des campagnes d'auscultation de ces mêmes routes.

9.2 Perspectives

Les travaux actuels permettent de répondre très partiellement à la problématique de gestion de l'entretien routier en se focalisant sur un unique mode de dégradation, la fissuration longitudinale, et la restriction du domaine de décision à une seule section de chaussée. Ceci permet de définir les perspectives dans le contexte de l'application SBADFORM sur deux axes.

Le premier axe de recherche est, comme nous l'avons citée précédemment, l'étude d'applicabilité de notre modèle de maintenance. L'approche envisagée se décline en deux étapes :

1. la caractérisation statistique du modèle de fissuration que nous avons développé ainsi que la production des outils d'estimation associés. La validation du modèle sera menée sur des données issues de modèles mécaniques numériques ainsi que sur des données terrain (essentiellement béton) ;
2. une étude de performance de notre modèle en termes de flexibilité et de qualité d'ajustement. Cette étude reposera sur des données simulées (la base IQRN n'étant pas suffisamment renseignée). Les premières conclusions de notre étude amènent des questions sur la définition même du taux d'accroissement de dégradation. Nous avons proposé dans une étude préliminaire de définir cette variable comme un facteur de correction entre la dégradation estimée par le biais d'observations successives d'une section et la dégradation espérée [124, 125]. Par ailleurs, il est important de chercher à identifier ce paramètre par rapport aux phénomènes étudiés pour s'assurer de sa pertinence et ainsi renforcer l'intérêt de la communauté Génie Civil et Mécanique pour la maintenance des structures, une telle approche se positionnant en rupture par rapport aux pratiques actuelles qui reposent sur l'extension de modèles de dégradation multi-physiques et multi-paramétrés peu robustes aux aléas et incertitudes. Pour aller dans ce sens, nous avons déposé un projet *Stratégies d'Identification d'un Méta-Modèle pour la Maintenance des structures (SI3M)* ayant de fortes chances de financement par la Région Pays de Loire.

Le deuxième axe de recherche est l'optimisation de la décision dans le contexte du réseau routier. Nous considérerons dans ce premier temps uniquement les dépendances économiques qui sont liées d'une part à des économies d'échelle lors d'entretiens simultanés sur des sections contiguës et d'autre part au respect des contraintes d'utilisation du réseau comme l'obligation de pouvoir joindre à tout moment un point du réseau. Pour se faire, nous avons défini avec l'IFSTTAR un réseau virtuel constitué de routes nationales à 2 ou 4 voies. La première étape de cette étude consistera en la formulation du modèle mathématique représentant l'état de disponibilité et de maintenance du réseau en chaque période de décision (décision annuelle). La deuxième étape sera la construction d'un algorithme efficace permettant au vu des contraintes budgétaires de planifier les actions à réaliser sur chacune des sections en fonction de leur niveau de dégradation dans l'année à venir. La complexité du problème abordé nous laisse présager outre la nécessité de développer nos propres algorithmes d'optimisation, la nécessité de simplifier l'espace des états des chaussées.

Chapitre 4

Extension de l'environnement décisionnel : plans d'investissement et stocks de pièces de rechange

1 Objectifs

Les résultats présentés dans cette section sont issus de travaux effectués dans le cadre de la thèse de Phuong Khanh Nguyen Thi dirigée par Philippe Castagliola, encadrée par Thomas Yeung et moi-même depuis octobre 2009 (la soutenance est planifiée le 06 novembre 2012). Ils ont fait l'objet de deux publications en revue (dont une est dans le processus de révision) [76, 77] et de 6 conférences (dont 3 avec actes et une invitée) [81, 82, 78, 79, 80]. La motivation principale de cette thèse est liée aux questions suivantes : lorsque l'innovation technologique n'est pas prise en compte dans le processus de maintenance, on peut estimer que, globalement, l'objectif de l'optimisation de la maintenance se cantonne à la maximisation des temps de bon fonctionnement pour retarder un éventuel remplacement, modulo un ensemble de contraintes inhérentes au problème de décision, mais qu'en est-il avec une prise en compte de l'innovation technologique ? Ne doit-on pas bénéficier au plus tôt d'une technologie, certes plus onéreuse à l'achat, performante en termes de productivité mais aussi moins consommatrice de ressources de maintenance ? Ou a contrario, ne doit-on pas chercher à maintenir coûte que coûte le matériel actuel pour espérer une technologie supérieure ?

On voit bien les relations entre stratégies de maintenance et stratégies d'investissement économique, investissement défini ici comme renouvellement de matériel. Cependant, notre vision est que les courants de recherche actuels ainsi que l'organisation de l'entreprise forcent à les opposer ou du moins à les découpler. En effet, la question de l'investissement est majoritairement motivée par des préoccupations d'ordre stratégique économique alors que l'organisation de la maintenance se décline plutôt au niveau opérationnel en cherchant à minimiser son impact sur les performances des processus en interaction. Nous voyons un double intérêt au rapprochement de ces deux domaines. Le premier, classique, est l'optimisation de la décision dans un contexte global et les intérêts réciproques que peuvent s'apporter ces deux domaines [62]. Le second est d'ordre plus philosophique dans le sens où il positionne le domaine de la maintenance comme un levier d'amélioration des performances de l'entreprise et non plus comme, encore trop souvent, un processus support d'assurance de ses performances. Ce dernier aspect a tendance à restreindre de manière naturelle la fonction maintenance aux activités de maintien des systèmes sans pour autant chercher à mettre au premier plan son potentiel d'amélioration des performances de l'entreprise, et de facto, à positionner la maintenance et les acteurs associés dans une perception péjorative voire négative de leurs activités vis-à-vis des organisations.

Afin de justifier une meilleure intégration de la maintenance au contexte plus global d'en-

treprise, nous proposons ici de quantifier l'influence de la décision de maintenance sur une stratégie d'investissement dont les critères de décision reposent, entre autre, sur l'obsolescence des matériels et le séquençement ou la probabilité d'occurrence de nouvelles technologies sur un horizon de temps défini. Afin de limiter l'étendue de notre étude, nous concentrons notre argumentation sur des aspects purement technologiques indépendamment des différents marchés, notamment l'offre et la demande. Par ailleurs, nous n'intégrerons pas ici la composante financière de l'investissement en considérant la disponibilité à tout moment des fonds monétaires pour la réalisation d'un investissement. Cette hypothèse peut paraître forte au vu des contextes économiques mais reste cependant en cohérence avec le point de vue de toute une communauté économiste considérant que, si l'évolution technologique laisse supposer l'assurance de compétitivité de l'entreprise, cette dernière sera apte à justifier la mise en place des capitaux pour la réalisation de tels investissements. Il reste cependant d'intérêt d'étudier la rentabilité d'un tel investissement pour la prise de décision.

Plus précisément, l'objectif principal de ce travail consiste à optimiser de manière conjointe les intérêts opérationnels de la maintenance et la stratégie d'investissement/remplacement de matériel dans un contexte d'évolution technologique. Les préoccupations les plus courantes dans le domaine de l'optimisation de la maintenance sont le maintien et la garantie de performance du système actuel. Dans ce cadre, les décisions de maintenir ou de remplacer reposent classiquement sur des ratios de coûts et de performances "fonctionnelles" du système modélisées par le biais d'indicateurs de dégradation ou encore des taux de défaillance. La décision d'investissement sera elle motivée par, d'une part, le niveau d'obsolescence du système actuel par rapport aux performances connues ou estimées des technologies disponibles ou à venir et, d'autre part, l'estimation du retour d'investissement en intégrant la vente de l'actif et d'un éventuel stock de pièces de rechange. Il est clair que le niveau de dépréciation de l'actif est aussi fonction de son âge ou de son niveau de dégradation. Ceci permet de faire le lien avec des objectifs de maintenance. Par ailleurs, l'un des challenges pour la prise de décision en investissement est clairement l'acquisition d'information sur les futures évolutions à la fois en termes de performances technologiques mais aussi des dates de mise sur le marché. Il est clair que les problématiques d'investissement diffèrent en fonction de la fréquence estimée d'apparition de ces nouvelles technologies [23] :

1. La fréquence d'apparition des nouvelles technologies est très élevée. Cette hypothèse est généralement liée à des systèmes technologiques peu onéreux et non réparables pour lesquels l'évolution technologique est plutôt synonyme de petites améliorations, notamment fonctionnelles. Les modélisations d'évolution technologique rencontrées sont généralement des modèles linéaires ou géométriques continus et la question de stratégie d'investissement se pose a priori par rapport à un niveau opérationnel. Dès son amortissement, le composant pourra être remplacé par un composant à niveau de performance à définir.
2. La fréquence d'apparition des nouvelles technologies est du même ordre de grandeur que la durée de vie d'un composant. Dans ce cas, il est intéressant de considérer une séquence discrète de dates d'arrivée pour laquelle les durées inter-événements sont des variables aléatoires distribuées par une loi de probabilité supposée donnée. Règles de maintenance et d'investissement sont à définir simultanément.
3. Les durées d'évolution technologique sont longues par rapport à la durée de vie des composants. La préoccupation principale du décideur peut se ramener à la définition d'une politique optimale de maintenance en intégrant la probabilité d'apparition d'une unique nouvelle technologie sur un horizon d'étude bien défini.
4. La nouvelle technologie est déjà présente sur le marché et ses performances sont connues. La problématique principale est liée à l'obsolescence matérielle et la gestion du compromis entre anticipation de renouvellement pour bénéficier de la nouvelle technologie et maximiser la vie résiduelle du composant actuel.

Le rapprochement entre performances opérationnelles et décisions stratégiques impose, au vu des diverses incertitudes, la construction d'un modèle décisionnel flexible, dynamique et réactif fonction du niveau de l'information courant disponible. L'analyse des options réelles, concept inspiré des options financières [100], nous semble être une approche prometteuse pour l'atteinte de ces objectifs.

Les travaux que nous allons présenter dans ce chapitre se positionnent sur les 4 points précédemment cités. Plus précisément, en voici les contributions plus significatives :

- une prise en compte de l'état de dégradation du système pour améliorer l'évaluation courante du coût de l'actif et s'assurer du meilleur compromis actions de maintenance et investissement ;
- l'application des options réelles dans le contexte de la maintenance ;
- l'extension du contexte décisionnel en maintenance intégrant disponibilité et coûts des moyens de maintenance par le biais de la gestion des stocks de pièces de rechange ;
- la formulation d'hypothèses plus réalistes telles que des lois non stationnaires pour la construction de la séquence technologique ou encore des coûts d'achat décroissants à partir de leur date de mise sur le marché ;
- l'impact de la connaissance ou niveau d'information relative aux nouvelles technologies.

L'organisation à suivre de ce chapitre est la suivante. Nous chercherons dans un premier temps à positionner le problème de maintenance/investissement par le biais d'une analyse bibliographique. Cette section nous permettra par ailleurs d'introduire la notion d'*options réelles* qui reste novatrice dans le contexte classique de l'optimisation de la maintenance. La deuxième partie de ce chapitre sera centrée sur les modèles de performance de la maintenance dans un environnement global intégrant stock de pièces de rechange et investissement dans des nouvelles technologies. La partie suivante abordera la problématique maintenance/investissement lorsque la fréquence d'apparition de nouvelles technologies est relativement faible avec, entre autre, une étude de l'horizon d'étude à prendre en compte pour garantir une robustesse à la décision et l'introduction d'une nouvelle option d'achat d'information sur les technologies à venir. En guise de conclusion, les perspectives à donner à ces travaux seront discutées.

2 Positionnement scientifique du problème

En introduction, nous avons mentionné très brièvement la problématique générale par rapport à son intérêt dans un contexte décisionnel industriel. L'objectif de cette section est de présenter un rapide horizon des orientations de recherche sur les problématiques d'optimisation de la maintenance dans un contexte d'évolution technologique. Nous proposons de les aborder sous le biais de 3 champs : les modèles de maintenance et d'investissement, l'obsolescence matérielle et enfin une réflexion sur l'intérêt des options réelles pour l'optimisation de la maintenance.

2.1 Maintenir ou investir ?

La vision économiste du problème

La question du maintien ou de l'investissement a été initialement discutée d'un point de vue macro-économique au vu, d'une part, des volumes financiers croissants engendrés par la maintenance et, d'autre part, comme alternative de cette dernière à l'investissement dans l'attente de perspectives économiques favorables [62]. On retrouve cette vision économiste dans de nombreux travaux que l'on peut classer comme le souligne [92] en fonction de la nature continue ou discontinue des modèles de changements technologiques, l'aspect continu fait référence à des évolutions séquentielles et discontinu à des sauts technologiques. L'une des premières critiques que l'on peut énoncer est relative à la modélisation de la performance des stratégies de maintenance uniquement basée sur des indicateurs économiques sans réelle prise en compte

des indicateurs de performance technique et fonctionnelle des systèmes ou encore de la diversité des actions de maintenance. Ainsi, les processus étudiés dans, par exemple, [36, 37, 85], sont l'évolution des coûts de maintenance et de production directement modélisée par un mouvement brownien géométrique pour lequel le taux de dérive instantané représentant le facteur de diminution des revenus est constant. Ces approches, notamment [85], permettent de définir le seuil de décision d'un remplacement ou d'investissement dans une nouvelle technologie en fonction de la dépréciation économique de l'actif et de la probabilité d'apparition d'une nouvelle technologie au cours du temps. Cette approche trouvant son intérêt dans un changement discontinu a été largement étudiée et plusieurs extensions sont proposées, notamment des stratégies en fonction de plusieurs technologies disponibles ou futures [91, 44]. Dans ce cadre, les différentes technologies sont caractérisées par une amélioration connue des flux moyens de profit.

Dans le contexte d'évolution technologique continue, les problématiques se rapprochent des situations rencontrées dans [91, 44]. Il est en effet nécessaire de faire son choix entre investir immédiatement dans la technologie courante ou bien différer cet investissement dans l'attente d'une technologie plus performante [73]. Par contre, les modélisations de l'évolution des gains associées aux technologies successives ainsi que leurs instants d'apparition sur le marché deviennent ici très importantes. L'hypothèse initiale d'une évolution déterministe et linéaire des technologies au cours du temps proposée dans [34] est trop restrictive. Des modèles plus récents [63, 20, 11] considèrent des évolutions géométriques pour les dates d'apparition. Par ailleurs, les différences entre évolutions successives ne sont plus seulement caractérisées par des améliorations en termes de revenus attendus par unité de temps mais intègrent aussi des notions de coût de maintenance et de prix d'achat [20, 42] et aussi de valeur résiduelle de l'actif courant [96]. Cependant, ces notions ne font pas référence directement aux caractéristiques fiabilistes classiquement utilisées pour l'optimisation de la maintenance. Hritonenko et Yatsenko dans [43] proposent une formulation joignant l'ensemble des performances précédentes à la durée de vie. Cette équation leur permet de déterminer les conditions d'ordre entre les coûts d'achat et les coûts de fonctionnement/maintenance explicitant les résultats paradoxaux de [20] et de [42] démontrant respectivement qu'une accélération du changement technologique permet de différer ou d'anticiper la décision d'investir dans une nouvelle technologie.

L'ensemble des modèles que nous avons cités dans le paragraphe ci-dessus considèrent des évolutions déterministes des technologies. Certains auteurs [63, 11] proposent des approches stochastiques par le biais de mouvement brownien géométrique par exemple. Ceci entraîne par contre des difficultés de résolution liées à l'explosion combinatoire des états. Par ailleurs, les lois de ces processus stochastiques reposent sur des paramètres supposés estimés sur la base d'observations. De notre point de vue, l'acquisition de ces données peut être extrêmement délicate et les valeurs des paramètres se doivent d'intégrer une connaissance assez subjective des caractéristiques (dates de mise sur le marché et performance) des technologies à venir. La qualité de l'information et la définition d'un processus d'acquisition d'information sont ainsi des challenges à relever dans la prise de décision de maintenance sous évolution technologique. Cette problématique a été finalement peu étudiée dans un tel contexte. On pourra citer les travaux [69, 59] et [104] proposant des modèles d'acquisition d'information. Cette information est principalement relative aux performances de systèmes déjà disponibles sur le marché. La décision associée se résume ainsi à investir ou non en l'état de connaissance ou bien devoir accroître cette connaissance avant de décider.

Enfin, plusieurs travaux [8, 64] s'intéressent au problème récurrent de la définition d'un horizon d'étude suffisant pour fournir une solution optimale robuste.

De la maintenance à l'investissement

Les approches classiques en optimisation de la maintenance reposent sur la caractérisation des performances fonctionnelles des systèmes et les effets de vieillissement, soit par le biais de distributions de durée de vie soit par des états de dégradation. L'amélioration technologique est peu prise en compte et les approches alors préconisées sont d'évaluer les nouvelles stratégies de maintenance en considérant l'indépendance des règles. Ceci revient à ne pas considérer l'investissement comme endogène au processus de décision mais typiquement comme une contrainte extérieure. Les travaux que nous avons rencontrés pour la prise en compte d'amélioration technologique se limitent au problème de saut technologique (une seule nouvelle technologie est à considérer) au niveau du système [12, 40] ou bien au niveau du composant [66, 67, 25, 24]. Seuls les travaux présentés dans [40] considèrent l'arrivée de la nouvelle technologie sur le marché comme aléatoire, les autres considérant cette technologie disponible.

Hopp et Nair [40] considèrent une séquence d'apparition technologique donnée par une probabilité conditionnelle non stationnaire d'apparition de la future technologie. Leur modèle de maintenance conditionnelle pour un système à états discrets de dégradation permet d'étudier la sensibilité du seuil de remplacement optimal en fonction des paramètres des lois non stationnaires de mise sur le marché, de coûts d'acquisition et de revenus attendus. Bogonovo *et al.* [12] propose de prendre en compte plusieurs types de maintenance, maintenance minimale, maintenance imparfaite et remplacement. L'impact de ces actions est modélisé directement sur le taux de défaillance d'une loi de Weibull. Le modèle géométrique d'évolution technologique est défini en termes de décroissance exponentielle du taux de défaillance en référence à un taux nominal. On a ainsi la technologie à l'instant t par rapport à un instant t_0 caractérisée par le taux de défaillance suivant :

$$\lambda(t) = \lambda(t_0) \cdot e^{-\sigma(t-t_0)}. \quad (2.1)$$

Leur travail permet d'aborder la question de l'obsolescence que nous nous proposons de discuter dans le paragraphe suivant.

Lorsque l'évolution technologique touche le composant plutôt que le système en soi, les problématiques se focalisent sur la décision de remplacement d'un ou d'un ensemble composants [66], par exemple, et du problème de compatibilité ou d'obsolescence des composants [67, 25, 24]. Mercier [66, 65] modélise les composants par un ensemble de caractéristiques dont notamment les taux de défaillance non constants. Elle propose d'étudier l'optimalité de politiques préventives consistant à remplacer systématiquement les composants défaillants par des composants de nouvelle technologie déjà disponible sur le marché et à définir le nombre K de remplacements correctifs à partir duquel on remplace préventivement tous les composants de l'ancienne technologie. En conclusion, elle souligne que les conditions d'optimalité sont difficiles à dériver dans le cadre d'un horizon d'étude fini mais que le décideur peut toutefois se baser sur les résultats prouvés dans le cadre stationnaire. Les auteurs dans [67, 25, 24] étendent cette approche sur de multiples champs dont la prise en compte des dépendances entre composants ou encore la possibilité de réparation par un modèle de type *réduction d'âge* avec des durées non négligeables. La complexité du modèle ne permet plus une analyse analytique de ses performances et une estimation par le biais de simulations de Monte Carlo associées à des réseaux de Petri est présentée. Il reste difficile d'en tirer des conclusions générales sur les politiques optimales.

2.2 L'obsolescence matérielle

La notion d'obsolescence est primordiale pour la décision de la maintenance. L'obsolescence est définie comme la perte de valeur d'un composant, système ou usine, non liée à son état ou encore à son histoire passée mais à des changements liés à des scénarios exogènes d'évolution technologique ou commerciale [99].

L'obsolescence peut être par nature objective par manque de disponibilité de la technologie courante, du service de maintenance associé ou d'incompatibilité avec des nouveaux systèmes. On parle respectivement d'*obsolescence indirecte* et d'*obsolescence par incompatibilité*. Les problématiques de décision, bien illustrées dans les travaux de [66, 65, 25, 24] par exemple, sont essentiellement liées à la définition de stratégies de remplacement préventive des items sachant les performances des nouvelles technologies. Il est cependant clair de la nécessité de chercher à estimer les gains des diverses stratégies traduits dans ces articles en termes de fiabilité, de coûts d'énergie (ou de production) et d'éventuels revenus par unité de temps de fonctionnement. Clavareau et Labeau [25, 24] intègrent par ailleurs la notion de stock de pièces de rechange qui permet éventuellement de différer un investissement.

La notion d'obsolescence peut être plus subjective en relation avec des questions de performances actuelles et espérées en considérant la nature incertaine des technologies à venir. On peut aussi avoir des technologies passées qui restent toujours disponibles et assimiler ceci, par exemple, à un marché concurrentiel présentant des améliorations successives de technologie. Outre les décisions classiques de maintenance pour la réparation ou le remplacement de systèmes, il devient nécessaire de choisir la technologie la plus appropriée, sachant que ce choix devra porter sur le meilleur ratio valeur de l'actif courant intégrant sa dépréciation physique et performances attendues en fonction des diverses technologies. L'intérêt de tels modèles est la mise en relation entre décision de maintenance et une notion économique de saut technologique (association des performances technologiques et économiques). Dans un tel contexte, on ne connaît pas l'ensemble des caractéristiques des technologies à venir. Ce point rejoint les modèles discutés précédemment [40, 12, 63, 11].

Je voudrais souligner ici qu'à part [25, 24] la problématique des stocks de pièces de rechange dans ce contexte d'évolution technologique n'est pas explicitement abordée. Dans leur état de l'art des modèles de gestion des stocks de pièces de rechange, Kennedy *et al.* [49] soulignent l'importance du problème d'obsolescence des pièces de rechange au vu des coûts engendrés (coûts de stockage et coûts d'obsolescence). Par contre, ils ne soulignent pas l'influence d'un état de stock sur la décision même d'investissement, se ramenant dans leur analyse aux objectifs de mission de la gestion des pièces de rechange sur les objectifs de disponibilité des systèmes.

2.3 Options réelles : une approche pour l'optimisation de la maintenance ?

L'objectif de cette sous-section n'est pas de fournir un état de l'art approfondi sur les approches en termes de stratégies ou de modèles abordés dans chacun des contextes mais de recenser les travaux étudiant leurs relations. Cependant, dans un souci pédagogique, j'en rappellerai brièvement la définition et les méthodes d'évaluation associées, rejoignant la programmation dynamique discutée dans le chapitre précédent.

Définition et intérêt des options réelles

Les options réelles est un concept dans le domaine économique et financier, directement dérivé de la théorie des options financières. Il a été formalisé en 1977 par Myers [71]. La différence majeure est la notion d'investissement dans des actifs réels et non dans des produits financiers et n'impose pas une obligation d'achat ou de vente [54].

Les options réelles s'appliquent à tout type de projet vérifiant les conditions générales suivantes, [52] :

- *L'incertitude* qui rend le résultat du projet non connu. Les incertitudes peuvent être traduites sous forme de probabilités ou croyances ;
- *La flexibilité* qui induit la possibilité de décision à tout moment en fonction de la situation ;
- *L'irréversibilité* d'un investissement pour lequel la dépense initiale ne peut être récupérée ;

- *La révélation d'information* qui permet une réduction de l'incertitude tout au long du déroulement d'information. Le mécanisme de révélation peut être passif (le coût de l'investissement est donné lors de la mise sur le marché) ou actif (mise en place de tests marketing ou d'études de marché).

L'intérêt des options réelles par rapport notamment aux méthodes traditionnelles de type cash-flows actualisés est essentiellement lié à la flexibilité des décisions pour une meilleure évaluation des opportunités d'investissement avec mesure de risque et incertitude. La flexibilité des décisions est l'ajout d'options telles que *retarder ou avancer l'investissement, le modifier ou l'annuler, etc.*, en fonction de la connaissance du marché. La variable de décision pour les options réelles est la *Valeur Actuelle Nette Augmentée* qui prend en compte la Valeur Actuelle Nette de l'investissement et le coût des options.

Les approches classiques pour l'évaluation des options réelles sont [38] :

- L'approche analytique ou résolution d'équations aux dérivées partielles qui est très limitée du fait d'hypothèses très restrictives ;
- Les arbres de décision qui consistent à discrétiser l'évolution du cours de l'actif sous l'hypothèse d'un mouvement brownien géométrique introduite comme la solution discrète de la formule de Black et Scholes. Par contre, cette approche n'est pas compatible pour une intégration de plusieurs options ;
- Les méthodes de simulation de Monte Carlo qui offrent une grande flexibilité mais requièrent beaucoup de temps de calcul. Par ailleurs, Mercier [66, 65] souligne le problème de dépendance de la décision optimale à la longueur de l'horizon d'étude qui demande alors une augmentation du nombre de trajectoires à simuler.
- La programmation dynamique stochastique qui s'intègre bien dans l'esprit des options réelles. Cependant, elle est particulièrement sensible à la qualité des informations disponibles, surtout lorsque l'acquisition reste difficile, et ne s'adapte pas bien dans des contextes où les marchés sont fortement risqués.

Application des options réelles pour la maintenance

Aujourd'hui, à notre connaissance, très peu de travaux [38, 46, 68] abordent l'optimisation de la maintenance par une modélisation de type options réelles, bien que la problématique se prête bien au contexte d'application. Aucun de ces travaux n'intègre la question de l'investissement au travers de l'évolution technologique.

Sans chercher à en expliquer la mise en œuvre, objet des chapitres suivants, nous proposons ici de montrer en quoi cette approche nous semble d'intérêt pour le rapprochement des problématiques maintenance et investissement. Les options réelles, initialement développées pour répondre à des questions stratégiques, par leur flexibilité fonction de l'évolution du contexte décisionnel, permettent d'y intégrer des préoccupations plus opérationnelles. Plus précisément, en définissant l'investissement dans une nouvelle technologie comme un projet, l'option *réparer* permet de différer un investissement notamment si le marché est peu favorable et attendre, par exemple, une technologie plus performante. Par ailleurs, cette option peut éventuellement être un facteur d'anticipation ou consolidation de la décision d'investir en assurant un niveau de prix de revente de l'actif plus élevé, en supposant que ce coût soit fonction du niveau de fonctionnement ou de vieillissement.

Dans le contexte que nous allons développer, nous chercherons à formuler le problème sous forme d'un programme dynamique stochastique. En effet, au vu de l'analyse [38], la programmation dynamique est particulièrement recommandée pour des projets où le risque technique est relativement élevé et pour lesquels leur valeur objective dépend étroitement des décisions successives.

3 Définition et formulation du problème général

Performance économique du système courant, processus de dégradation, coûts d'achat et de revente, coûts de maintenance, caractérisation du marché, etc. sont un ensemble de quantités qu'il va falloir modéliser. Ces modélisations vont engendrer un grand nombre de données qu'il pourra être difficile de caractériser dans un contexte opérationnel. Nous proposons d'adopter une démarche similaire au chapitre précédant à savoir de considérer, lorsque ceci est possible, les quantités comme continues puis d'étudier le problème discrétisé. Considérer le problème continu permet de mettre en avant les propriétés des processus étudiés (par exemple, la croissance d'un processus de dégradation) et aussi le nombre de paramètres associés à la dynamique ou transition des systèmes et événements. Nous considérons par ailleurs un intervalle de décision constant.

3.1 Modélisation de la performance des systèmes technologiques

Les deux aspects de la performance d'un système que nous allons considérer sont :

- la performance fonctionnelle qui peut être décrite par un indicateur de dégradation croissant au cours de son temps d'utilisation que nous modélisons par un taux de dégradation par unité de temps supposé constant et un seuil de panne donné. L'évolution de cette dégradation sera modélisée par un processus gamma stationnaire, noté X_t ;
- la performance économique qui est la combinaison de l'aptitude à produire (et donc de la performance fonctionnelle) et de l'attractivité commerciale du produit ou service fourni (qui est plutôt fonction du marché). On appellera $G(x)$ la fonction profit cumulé dans une période de temps à un niveau de dégradation donné x .

Cette distinction met en évidence les indicateurs d'intérêt suivant le décideur. Un système à tout moment a une valeur économique qui décroît en fonction du temps et de son usage. On définit un processus de dévaluation du prix d'achat d'un système depuis sa date de mise sur le marché et de sa valeur résiduelle à un instant donnée en fonction du coût d'achat du matériel neuf à cette date et de son niveau de dégradation.

3.2 Modélisation de l'évolution technologique

Une évolution technologique comprend deux paramètres, à savoir la fréquence d'apparition ou de mise à disposition des nouvelles technologies et les performances de ces nouvelles technologies.

On suppose que la fréquence d'apparition est donnée par une probabilité d'apparition d'une nouvelle technologie croissante par période de temps. L'amélioration technologique conduit à une réduction de la vitesse moyenne de dégradation (on se propose d'introduire un facteur multiplicatif dépendant de la technologie sur le paramètre α du processus gamma) et une augmentation du profit réalisé. Dans notre étude, nous avons supposé que cette augmentation du profit était directement intégrée au niveau de la modélisation de la dégradation.

3.3 Description du processus de décision et définition des coûts associés

Comme nous l'avons souligné, la valeur d'une option est constituée de la valeur nette augmentée dans le sens où elle prend en compte le coût de l'investissement ou décision courante et des conséquences économiques associées, conséquences définies sur un horizon d'étude donné. On considère un ensemble fini et constant de décisions ou actions en fonction de l'état du système. Ces décisions sont

- *Ne rien faire* qui consiste à laisser le système en l'état ;
- *Réparer* qui permet de restaurer le système dans un état de fonctionnement plus élevé, sans que ce dernier puisse être considéré comme neuf ;

- *Remplacer* qui se décompose en deux étapes à savoir la vente de l'actif courant et son renouvellement par un système neuf de même génération technologique.
- *Investir* qui est équivalente à la décision précédente pour laquelle un choix de changement de génération technologique est possible. Il est clair que ce choix se réalisera parmi les technologies disponibles sur le marché lors de la prise de décision. Par ailleurs, nous considérons un marché cohérent dans le sens où une amélioration de performance est constatée à chaque nouvelle génération.

Chacune de ces décisions engendre un coût supposé immédiat. Dans le cas d'une réparation, il consiste à un coût fixe d'opération alors que pour le remplacement et l'investissement, le coût sera la différence entre le prix d'achat du nouveau système et le coût de vente du matériel actuel.

3.4 Choix et construction du critère de décision

Comme nous l'avons rappelé, la valeur d'une option à un instant donné sera évaluée par le biais de la programmation dynamique. L'une des différences principales dans l'utilisation de la programmation dynamique par rapport au chapitre précédent est l'association d'un horizon d'étude fini. La longueur de cet horizon représente in fine la durée de validité de l'option, durée qui est soit déterminée par le décideur lui-même ou représentant l'atteinte d'un état stationnaire, ou considéré comme tel, soit optimisée pour s'assurer de la robustesse de la décision.

Nous ne reviendrons pas ici sur la définition d'un programme dynamique, principalement discret, sachant que la différence avec le critère (2.1) page 55 présenté dans le contexte stationnaire est l'introduction du temps, $n < N$ dans un horizon d'étude fini de longueur N :

$$V_n^N(x) = \max_{a \in A(x)} \{C(x, a) + \lambda E_G [V_{n+1}^N(y) \mid x, a]\} \quad (3.1)$$

avec

- $V_n^N(x)$: la fonction de valeur à l'instant n pour un état donné x ;
- $A(x)$: l'ensemble des décisions possibles connaissant l'état x ;
- $C(x, a)$: le coût immédiat engendré par la décision a connaissant l'état x ;
- E_G : l'espérance conditionnelle suivant la loi associée à l'état suivant la décision.

L'hypothèse d'un horizon d'étude infini est particulièrement mal adaptée au contexte d'évolution technologique. En effet, les questions de modélisation de la séquence technologique sur un horizon infini est délicate dans le sens où il serait alors nécessaire de considérer une infinité de nouvelles générations qui, à chaque évolution, améliorent les performances de la précédente. Par ailleurs, de notre point de vue, la notion de risque de décision erronée, liée par exemple à une mauvaise estimation de la condition du marché, sera limitée devant les performances du processus à long terme. Cet aspect rejoint en quelque sorte l'étude des régimes transitoires généralement occultée dans un contexte d'optimisation sur des horizons infinis).

3.5 Procédure de résolution et illustration numérique

Ici aussi, la discrétisation en état et en temps nous permet l'utilisation d'algorithmes classiques tels que le *Backward Induction Algorithm* [90] dans sa version classique sur un horizon de temps fini ou étendue pour l'étude de situations stationnaires (paragraphe 5.4).

L'objet principal de ce travail étant la mise en évidence des relations entre maintenance et investissement, nous n'avons pas cherché de manière systématique à améliorer les performances de ces algorithmes. Nous avons toutefois conduit des études sur l'existence de propriétés structurelles sur les politiques optimales pour la mise en évidence de résultats présentés, dans la mesure du possible, sous forme de *bonnes pratiques*.

4 Stratégies de maintenance sous séquence technologique

4.1 Objectifs

Cette section vise principalement à mettre en évidence l'apport de politiques de maintenance sur les stratégies de renouvellement des actifs sous l'hypothèse d'une évolution continue des technologies, continuité au sens de la définition donnée dans la partie précédente. Nous rappelons que dans ce contexte la décision de renouvellement consiste à définir la date d'investissement et à choisir la technologie la plus appropriée au contexte.

Nous considérons ici comme incertaine cette évolution technologique, l'information sur les nouveaux systèmes n'étant disponible que lors de leur apparition sur le marché. Par ailleurs, nous supposons que le marché est cohérent dans le sens où les performances de tout nouveau système sont supérieures aux technologies passées. La performance d'un système de génération j est résumée par son taux de dégradation instantanée α_j , le paramètre d'échelle β restera inchangé. Pour construire une séquence technologique, on combine un modèle géométrique pour l'amélioration des performances et des probabilités d'apparition non stationnaires pour les technologies successives.

Le processus de décision que nous allons préciser dans le paragraphe suivant consiste à déterminer les dates d'intervention sur le système, l'intervention étant soit une maintenance M permettant de restaurer partiellement la fonction du système, soit un investissement I^h dans la génération technologique h offrant le meilleur ratio performance attendue et coût d'achat.

4.2 Description du processus de décision

On se propose d'organiser la décision à chaque période de temps n en fonction de la valeur courante de dégradation x du système de génération technologique j sachant disponible un ensemble de technologies $j, j+1, \dots, k$. On définit l'état du système à l'instant n par le triplet $s_n = (x, j, k)$. Comme nous l'avons souligné en introduction de ce chapitre, nous supposons le marché cohérent ce qui implique $k \geq j$. L'ensemble des actions possibles en fonction de s_n est $A(s_n) = \{DN, M, I\}$. La maintenance M est une maintenance partielle qui remet le système dans l'état $\max(0, x - d)$ où d est le paramètre donné d'efficacité de la maintenance.

On estime ici que les temps de mise en place d'une décision et de réalisation de l'action associée sont négligeables devant le temps d'opération. Par ailleurs, on souligne que l'état défaillant, m , n'entraîne pas une intervention automatique de remise en état du système et qu'une option d'attente d'une meilleure technologie est offerte par le choix de l'action DN . Par contre, afin de pénaliser cet état de défaillance, on supposera un profit négatif sur la période, $G(m) < 0$. La Figure 4.1 illustre ce processus de décision.

Sous ces conditions, le problème de maintenance revient à déterminer la politique qui maximise le profit sur un horizon de temps donné, N . On note $V^\pi(s)$ la valeur actualisée estimée maximale sur l'horizon du temps fini pour l'état s du système. Si la dernière période de décision est N , à l'époque de décision $N + 1$, on ne fait rien et la condition terminale est $V_{N+1}(s) = 0$ ($\forall s \in S$: l'espace d'état du système).

Soit $V_i(x, k, j)$ la valeur actualisée maximale depuis l'instant de décision i jusqu'à la dernière période N . Alors, $V_1(s) = V^\pi(s)$.

$$V_i(x, k, j) = \max\{DN_i(x, k, j); M_i(x, k, j); I_i^h(x, k, j)\} \quad (4.1)$$

où DN_i, M_i, I_i sont respectivement le choix de ne rien faire, maintenir et remplacer au début

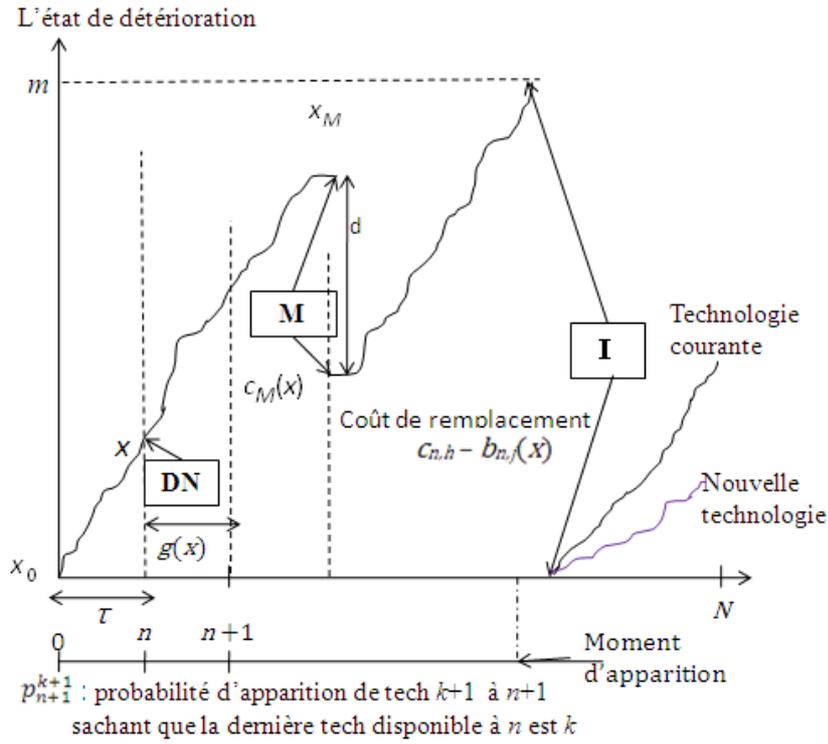


FIGURE 4.1 – Illustration des effets des décisions de maintenance

de la période de décision i . On a alors :

$$DN_i(x, k, j) = G(x) + \lambda \sum_{\forall k'} \sum_{\forall x'} p_i(x', k', j | x, k, j) V_{i+1}(x', k', j) \quad (4.2)$$

$$M_i(x, k, j) = -c_M(x) + DN_i(\max(0, x - d), k, j) \quad (4.3)$$

$$I_i^h(x, k, j) = -c_{i,h} + b_{i,j}(x) + DN_i(0, k, h) ; \quad \forall h : j \leq h \leq k. \quad (4.4)$$

avec

- λ le facteur d'actualisation, $\lambda \in [0, 1]$;
- $k' \in k, k + 1$.

La valeur associée à l'option *Ne rien faire*, équation (4.2), est fonction de la valeur actualisée à l'époque suivante pour toute valeur de dégradation $x' > x$ du système courant j (on ne change pas de technologie) et qu'éventuellement une nouvelle technologie apparaît (on passerait de k à $k+1$ dans ce cas). Pour la valeur associée à l'option *Maintenir*, équation (4.3), on retrouve bien le coût propre de l'action en fonction du niveau de réparation à laquelle on rajoute la valeur de l'option *Ne rien faire* avec le système j restauré dans son nouvel état de dégradation. Enfin, la valeur de l'option *Investir dans la technologie h* , équation (4.4), est donnée par le coût d'investissement propre, le gain lié à la revente du système courant j et la valeur de l'option *Ne rien faire* avec la nouvelle technologie h toute neuve, $x = 0$.

La forme d'une politique est illustrée dans la Table 4.1

4.3 Analyses des politiques

Nous avons cherché sur la base d'exemples numériques à mettre en évidence un ensemble de propriétés que nous pouvons décrire comme suit.

Influence du prix d'achat : Nous avons pu vérifier les résultats intuitifs d'investir dans les technologies les plus avancées lorsque l'évolution de leurs coûts d'acquisition reste faible

TABLE 4.1 – Exemple de table de décision pour les 3 premières périodes n où les intervalles représentent les valeurs de dégradation du système de technologie j sachant la disponibilité de la génération k .

i	k, j	Ne rien faire	Maintenance	Remplacement
1	1, 1	[0, 1.5)	[1.5, 7.3)	[7.3, 20]
2	1, 1	[0, 1.5)	[1.5, 7.3)	[7.3, 20]
	2, 1	[0, 1.5)	[1.5, 7.3)	[7.3, 20]
3	1, 1	[0, 1.5)	[1.5, 7.1)	[7.1, 20]
	2, 1	[0, 1.5)	[1.5, 7.1)	[7.1, 20]
	3, 1	[0, 1.5)	[1.5, 7.1)	[7.1, 20]

et bénéficier au maximum des effets de décôte dans le cas contraire. Pour des valeurs intermédiaires, nous avons noté des effets plus nuancés :

- l’investissement dans une nouvelle technologie est soumis à un intervalle d’amortissement suffisant. On pourra ainsi en fonction du temps résiduel avant la fin de l’horizon d’étude choisir lors d’un remplacement une technologie qui n’est pas forcément la plus avancée ;
- l’option *retarder* peut s’avérer très intéressante pour bénéficier de l’arrivée d’une nouvelle technologie. Nous avons ainsi pu observer des cas de perte de monotonie dans les politiques optimales : *Ne rien faire* sur un intervalle de dégradation, *Maintenir* sur un système plus dégradé puis de nouveau *Ne rien faire* pour préconiser un investissement pour un système très dégradé.
- la disponibilité d’une technologie plus avancée conduit à anticiper l’investissement.

Influence des écarts de technologie : Ici aussi, nous avons retrouvé des résultats intuitifs tels que l’investissement dans la meilleure technologie pour de fortes différences de performance entre technologies, le retour d’investissement étant alors très court. A contrario, on cherchera à bénéficier de la décôte des générations les plus vieilles pour de faibles évolutions. Pour des écarts plus nuancés, on retrouve des résultats analogues à l’étude précédente à savoir :

- la décision de remplacement par l’identique est repoussée pour pouvoir bénéficier d’une éventuelle nouvelle génération technologique pour des séquences à évolution technologique intermédiaire (ceci se traduit par un seuil de remplacement plus élevé pour des facteurs d’amélioration de performance plus forts). Cette propriété est de moins en moins vraie au cours du temps ;
- en cas d’obsolescence avérée (disponibilité d’une technologie plus performante), l’investissement sera anticipé ;
- on peut retrouver la perte de la monotonie de la politique optimale.

Cette discussion entre gros ou faibles écarts fait référence au phénomène de rupture technologique que nous proposons d’étudier dans la section suivante.

Influence de la taille de l’horizon d’étude : Nous avons déjà pu discuter des effets de l’horizon d’étude sur l’évolution des décisions dans les deux paragraphes précédents. Nous avons de la même façon fait varier la longueur de cet horizon pour, d’une part, illustrer la recherche de garantie de rentabilité économique d’un investissement sur le long terme (on cherchera à investir dans la technologie la plus chère si l’horizon de temps est suffisant) et, d’autre part, pour mettre en évidence l’existence d’une taille d’horizon minimale garantissant l’optimalité de la décision d’investir dans la dernière génération.

Discussion : Il est clair que les résultats présentés dans le contexte de notre étude sont fortement liés à nos choix de modélisation et des valeurs des multiples paramètres d’entrée. Il

est par ailleurs difficile d'estimer finement certaines de ces quantités sur la base d'observations. De notre point de vue, elles doivent principalement refléter la connaissance du décideur vis-à-vis du marché et des caractéristiques des systèmes à venir. Dans un tel contexte, l'avantage principal de notre modèle réside dans sa faculté à fournir des informations quantitatives quant aux intervalles de décision, et donc par le biais de conduite de simulations en faisant varier la valeur des données d'entrée de consolider les valeurs d'option et donc la décision.

4.4 Extension pour la disponibilité des pièces de rechange

Nous avons cherché à appliquer le modèle précédent dans un contexte de gestion de pièces de rechange. Cette étude présente le double avantage suivant. Le premier est relatif à la définition de l'obsolescence qui a été étudié dans une première partie sous l'unique aspect économique au travers de la valeur résiduelle. Dans un contexte de pièces de rechange, on peut aborder l'obsolescence matérielle définie ici par l'incompatibilité des pièces de rechange d'une génération à une autre. Le second avantage est l'intégration d'un processus industriel de gestion des stocks dans le processus maintenance.

Le problème de maintenance est défini comme suit. On s'intéresse à un système k sur n pour lequel la durée de vie de ses composants, supposés identiques, est exponentielle de paramètre donné. Le rendement du système est fonction du nombre de composants actifs. Lors d'une défaillance d'un composant, un remplacement de ce dernier peut éventuellement être effectué sous condition de disponibilité dans le stock de pièces de rechange, pièces fournies uniquement lors de l'acquisition d'un nouveau système.

Le processus de décision est analogue à celui de l'étude précédente dans la mesure où on différencie la maintenance imparfaite consistant à ne remplacer qu'un certain nombre d'items défectueux de la maintenance parfaite pour laquelle tous les items défectueux sont remis à neuf. Par ailleurs, des coûts de stockage sont intégrés dans le modèle afin de limiter les effets de surstockage. La notion d'état pour le problème décisionnel est alors étendu au quadruplet $s = (x, k, j, n_s)$ avec x le nombre de composants défectueux, k la dernière génération disponible sur le marché, j la technologie courante et n_s le nombre de pièces de rechange disponible en stock. La formulation dynamique du problème reste proche de l'équation (4.1).

Nous proposons de faire évoluer la représentation de la politique sous forme de diagrammes donnés pour chaque période n , pour toute technologie utilisée et disponible, Figure 4.2. À chaque état de dégradation x (axe horizontal) et pour un niveau donné de stock s (axe vertical) correspond une seule action. Les actions ne rien faire, maintenance imparfaite, maintenance parfaite et remplacement sont respectivement représentées en jaune, bleu clair, bleu foncé et en rouge. La politique optimale cherche à maintenir à moindre coût en fonction de la disponibilité des pièces de rechange pour attendre l'apparition éventuelle d'une nouvelle technologie.

En termes d'analyses des politiques optimales, outre l'intérêt de l'option d'attente des maintenances imparfaites pour bénéficier des dernières améliorations technologiques, voici les règles de gestion que nous avons pu déduire de nos expérimentations numériques.

L'investissement pour le réapprovisionnement en pièces de rechange : L'une des premières règles est relative au maintien des performances maximales du système courant par le biais de maintenance parfaite si le stock le permet ou imparfaite. La décision d'investissement est essentiellement liée à la question de réapprovisionnement en pièces de rechange, surtout pour des générations technologiques plus ou moins équivalentes. Cette règle tend à se substituer à des investissements anticipés au fur et à mesure de l'importance des écarts entre technologies.

Niveau de stock comme variable d'ajustement pour l'investissement : La décision d'investissement dans une technologie plus performante peut être anticipée pour bénéficier du

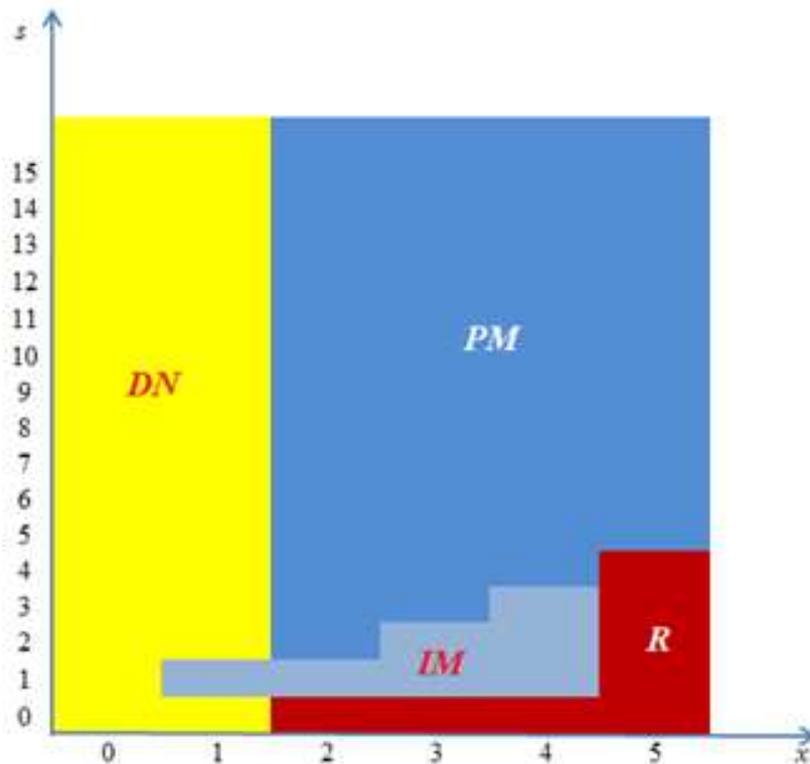


FIGURE 4.2 – Politique optimale dans le cas non-obsolescence ($k = j = 1$)

meilleur coût de revente du système actuel (système peu dégradé et niveau de stock élevé). Cet aspect est illustré par la zone d'investissement en haut à gauche de la Figure 4.3.

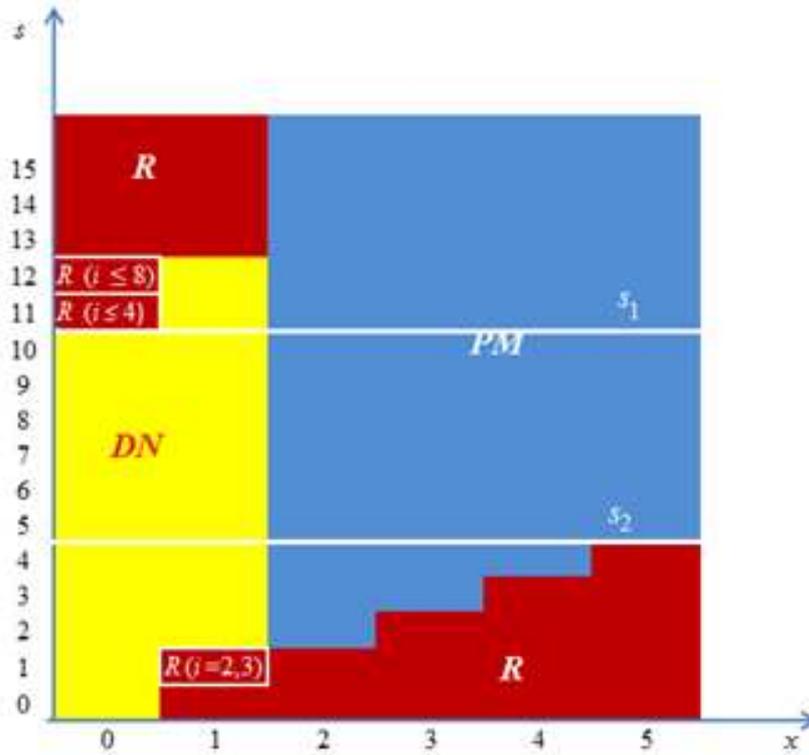
5 La décision de maintenance face au saut technologique

5.1 Objectifs

Dans la section précédente, nous avons cherché à analyser la relation entre politiques de maintenance et d'investissement lorsque l'évolution technologique se traduit de manière séquentielle. Dans ce contexte, nous avons construit ces séquences en 3 points : une probabilité non stationnaire d'apparition de la génération suivante, une construction empirique et déterministe des caractéristiques des nouvelles générations. Les conclusions que nous avons dégagées n'ont pas pu être généralisées. Par ailleurs, ces hypothèses n'ont pas permis de bien mettre en évidence les effets de l'incertitude du marché sur les décisions. Nous proposons de centrer cette section sur cet aspect en développant le modèle sur deux champs.

Le premier champ consiste à ne plus considérer une séquence technologique mais une seule évolution technologique sur un horizon de temps. Ce dernier cas est à rapprocher des études pour lesquelles le temps du changement technologique est plus long en moyenne que la durée de vie des actifs. L'analyse pourra ainsi mettre en évidence l'impact d'un unique saut technologique à l'instar des effets cumulés dans le temps des différentes générations.

La seconde orientation concerne l'incertitude du marché et sa flexibilité. Nous proposons dans un premier temps de revenir sur la construction des caractéristiques technologiques d'une part en renforçant les incertitudes sur le prix d'acquisition de la nouvelle technologie et, d'autre part, en proposant une modélisation plus globale des incertitudes du profit réalisé dans chacune des périodes de temps. Pour cela, nous proposons d'étudier le profit généré dans une période de décision sans chercher à expliciter le processus de dégradation. Dans un second temps, on s'attachera à étudier la relation entre incertitude sur le marché et décisions et

FIGURE 4.3 – Politique optimale dans le cas d'obsolescence ($k = 2, j = 1$)

intégrer au problème décisionnel une nouvelle option d'achat d'information, option assimilée à la réduction d'incertitude.

5.2 Construction du critère

Les données d'intérêt pour caractériser une nouvelle technologie sont sa date de mise sur le marché ainsi que ses performances attendues, déterminant ainsi la notion de saut technologique. On note p_{n+1, θ_i} la probabilité supposée connue que la nouvelle technologie ait le caractère θ_i à la période de décision $n + 1$. Dans un premier temps, on considère que la caractéristique affectée par le changement technologique est uniquement le prix d'achat de la nouvelle technologie. On se propose de modéliser ce prix d'achat par un mouvement Brownien géométrique. Un tel modèle permet une prise en compte de l'incertitude du temps d'apparition de la nouvelle technologie ainsi que de la volatilité de la prévision du coût au cours du temps.

Par ailleurs, le problème d'optimisation diffère des précédents dans le sens où on peut différencier des scénarios en fonction de l'apparition ou non de la nouvelle technologie sur le marché. En effet, par exemple, une fois apparue, on connaît ses performances et son prix d'achat. Le problème de décision revient à décider de maintenir ou de changer de technologie. La valeur associée à cet état sera notée $\hat{V}(\theta_i, x)$. On suppose de plus que, si la nouvelle technologie n'est pas apparue sur l'horizon d'intérêt, alors elle n'apparaîtra plus. Cette hypothèse entraîne donc la définition de la valeur finale $V_N^N(x)$ uniquement en fonction des performances du système courant, valeur notée $\tilde{V}(x)$.

Nous ne donnerons pas les expressions du critère de maximisation de la valeur $V_n^N(x)$ fonction des valeurs $\hat{V}(\theta_i, x)$ et $\tilde{V}(x)$, [77].

5.3 Analyses de performance du modèle

Nous avons cherché à conduire des études analytiques pour l'identification de structures décisionnelles pour les politiques optimales, sous condition d'existence. La conduite de telles études est classique et repose sur un ensemble de théorèmes dont les démonstrations sont basées sur des récurrences. Par contre, les difficultés sont clairement les définitions des assertions à prouver. De plus, une difficulté majeure de notre problème est liée à la combinaison des fonctions de valeur $\hat{V}(\theta_i, x)$ et $\tilde{V}(x)$ dans la fonction de valeur $V_n^N(x)$. Nous avons alors proposé une démonstration en deux étapes suivant la considération ou non du problème d'évolution technologique.

Une autre contribution que nous avons apportée sur ce champ est la définition d'une méthode d'identification de l'horizon d'étude pour garantir la robustesse de l'optimalité des premières décisions.

La propriété principale que nous avons pu démontrer est le potentiel qu'offre la maintenance pour retarder un investissement lorsque le prix d'acquisition de la nouvelle technologie est élevé. On cherche ainsi à tirer le plus de bénéfice possible de la technologie courante avant d'investir. Bien que ce résultat soit intuitif, la démonstration permet de définir les conditions d'obtention d'une décision optimale en terme de taille d'horizon d'étude.

À titre illustratif, pour des valeurs de paramètres d'entrée données, voici les résultats et conclusions obtenues sur les politiques optimales lorsqu'on fait varier l'horizon d'étude.

TABLE 4.2 – Décision optimale à $n = 1$ pour différentes longueurs d'horizon

Action optimale				
N	Inconnu	Investissement	Maintenance	Ne rien faire
6	$\forall g \in \Psi$			
8	$g \in [0.32, 7.99]$	$g \in [0, 0.31]$		$g \in [8, 10]$
12	$g \in [5.01, 5.11]$ $g \in [7.3, 7.34]$	$g \in [0, 5]$	$g \in [5.12, 7.29]$	$g \in [7.35, 10]$
15		$g \in [0, 5.06]$	$g \in [5.07, 7.31]$	$g \in [7.32, 10]$
24		$g \in [0, 5.06]$	$g \in [5.07, 7.31]$	$g \in [7.32, 10]$

Le Tableau 4.2 présente les intervalles de décision optimaux, intervalles de profit cumulé g par unité de temps, pour la première décision, $n = 1$ pour différentes longueurs de l'horizon d'étude $N \in \{6, 8, 12, 15, 24\}$. La lecture des résultats est la suivante :

- Pour tout N inférieur à 6, on ne peut pas déterminer l'option optimale pour tous les états du système. En effet, aucune des conditions d'optimalité n'est satisfaite.
- Lorsque l'horizon d'étude est $N = 8$, l'option optimale prescrit un remplacement par le nouveau système si $g \in [0, 0.31]$, de ne rien faire pour $g \in [8, 10]$, tandis que pour tout $g \in [0, 32, 7, 99]$, l'information n'est pas suffisante pour prendre une décision. La politique optimale pour tous les états du système n'est donc pas complètement déterminée et il est nécessaire d'augmenter N .
- À partir de $N = 15$, la politique optimale est déterminée en tout point g et ne change pas $\forall N > 15$. Elle prescrit le remplacement pour $g \in [0, 5.06]$, la réparation pour $g \in [5.07, 7.31]$ et ne rien faire pour $g \in [7.32, 10]$. On retrouve une politique optimale monotone en g .

Nous avons par ailleurs conduit des études de sensibilité numérique sur les paramètres comme la vitesse du marché (probabilité d'apparition de la nouvelle technologie) ou encore l'efficacité de la maintenance.

Les conclusions principales sont :

- Les bénéfices de la maintenance sont particulièrement importants pour attendre l'arrivée, lorsqu'elle est fortement probable, d'une technologie plus performante ;

- Lorsque la nouvelle technologie n’est pas encore apparue, l’intérêt d’une maintenance imparfaite croît en fonction de son efficacité de restauration ;
- Par contre, lorsque la nouvelle technologie est disponible, on perd la propriété d’extension de la durée de vie économique du système actuel grâce à l’option maintenance au bénéfice d’un investissement dans la nouvelle technologie. Ce résultat est nuancé lorsque le ratio prix d’achat et performance attendue est fort.

5.4 Incertitudes sur la qualité de l’information du marché

Nous venons de souligner l’importance de l’impact du changement technologique sur les politiques de maintenance et d’investissement. Nous avons cherché à illustrer cet impact par la conduite d’expérimentations numériques sur un ensemble de paramètres caractérisant principalement la différence entre technologie actuelle et générations futures ainsi que la vélocité du marché représentée par une probabilité d’apparition. Ces informations sont de nature très incertaine et intègrent une grande part de subjectivité de la part du décideur. Des actions telles que des études de marché peuvent être commandées afin d’en améliorer la connaissance sans toutefois la rendre parfaite. Ces études ont un coût non négligeable qui dépend fortement de la difficulté d’analyse du marché et de la fiabilité de l’information rendue. Par ailleurs, aussi riches qu’elles soient, ces études n’offrent qu’une vision temporaire du marché. Il est ainsi clair de l’importance de bien déterminer les instants de mises en place de telles mesures pour renforcer la connaissance du marché et alors de définir la décision la plus appropriée.

L’objectif de notre travail est clairement lié à l’impact de la qualité de l’information disponible pour la prise de décision en maintenance et la construction de politiques d’acquisition de nouvelle information sur la nouvelle technologie. Cet apport permettra d’améliorer la justesse de la décision de planification de la maintenance ainsi que les décisions d’investissement en réduisant notamment le risque d’erreur d’un renouvellement anticipé de l’actif par la même technologie en fait quasi-obsolète ou a contrario le report d’un projet de remplacement en attendant vainement une meilleure technologie.

Nous avons proposé une approche en deux étapes. La première étape consiste à estimer le besoin d’acquisition d’une information supplémentaire en fonction de l’état de la situation globale. La seconde étape consiste à la décision sur le système en fonction du niveau d’information classique ou renforcé avec la problématique dans le cas d’ajout d’information de son intégration dans le modèle courant.

Ceci se traduit au niveau décisionnel par une nouvelle option : *Acquérir de l’information* notée A . Le processus d’acquisition de l’information est modélisé comme un processus d’actualisation bayésienne, ceci pouvant être vu comme une extension des travaux de [57, 104]. L’état du système est alors défini par le couple (x, π^p) où π^p représente la connaissance a priori sur la distribution de la caractéristique de la nouvelle technologie. La fonction de valeur à maximiser sur un horizon de temps fini peut s’écrire de la façon suivante :

$$V_n^N(x, \pi^p) = \max \{O_n^N(x, \pi^p), A_n^N(x, \pi^p)\} \quad (5.1)$$

avec

- la valeur $O_n^N(x, \pi^p)$ de l’option *agir directement sans information complémentaire* qui revient à choisir entre les actions *Ne rien faire*, *Maintenir* et *Investir* ;
- la valeur de l’option d’acquisition $A_n^N(x, \pi^p) = -c_s + \sum_j f(s_j; \pi^p) O_n^N(x, \pi'(\pi^p, s_j))$ où
 - c_s est le coût direct d’acquisition de l’information,
 - $f(s_j; \pi^p)$ représente la probabilité prédictive de l’information s_j avec un a priori π^p ,
 - $O_n^N(x, \pi'(\pi^p, s_j))$ la valeur de l’option d’action immédiate avec la nouvelle connaissance actualisée $\pi'(\pi^p, s_j)$. Cette connaissance actualisée deviendra ensuite la connaissance a priori jusqu’à la prochaine acquisition d’information complémentaire.

La conduite de preuves formelles nous a permis de définir les conditions sur les données d’entrée d’obtention de la monotonie de la politique optimale et la convexité des fonctions de

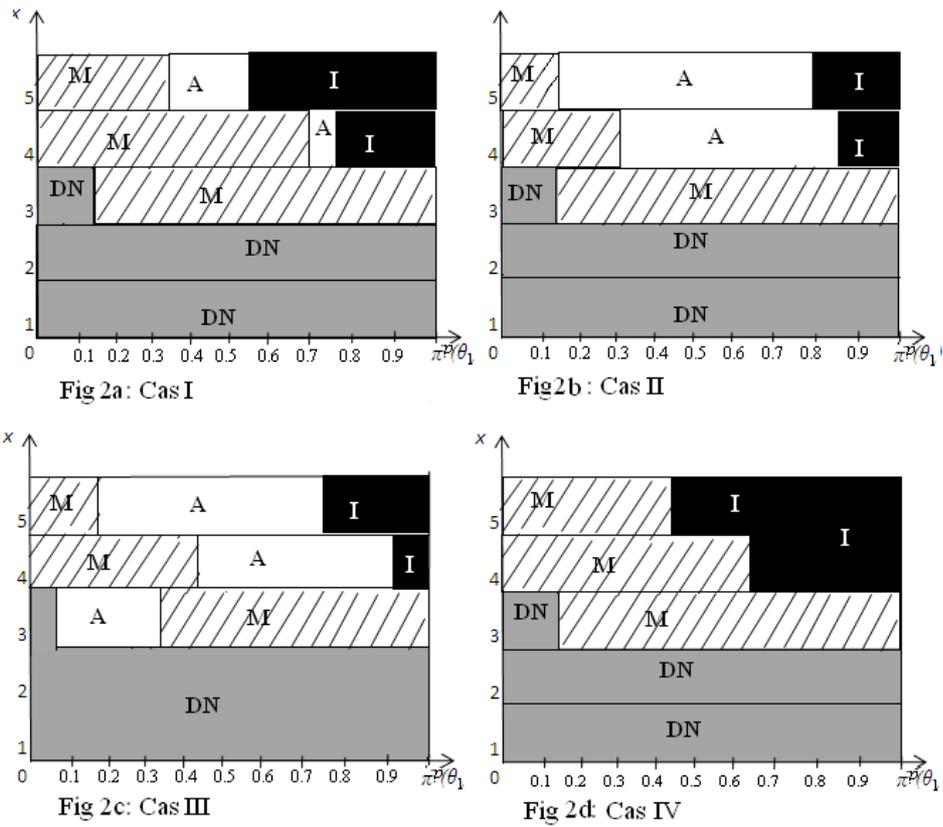


FIGURE 4.4 – Représentation des politiques optimales pour différents coûts et fiabilité de l'information obtenue - cas 1 : information assez fiable et coût d'achat élevé - cas 2 : information très fiable et coût élevé - cas 3 : information assez fiable et coût très faible - cas 4 : information peu fiable.

valeur. Cette convexité s'exprime en fonction de l'ordre du rapport de vraisemblance pour la distribution de probabilité de la caractéristique de la nouvelle technologie. Il a été en effet nécessaire d'introduire une fonction score pour la définition d'ordre entre les états (x_1, π_1^p) et (x_2, π_2^p) , $(x_1, \pi_1^p) \leq_{LR} (x_2, \pi_2^p)$.

Outre ces conditions, nous avons mis en évidence par le biais d'expérimentations numériques d'autres propriétés sur les politiques optimales. On retrouve des résultats intuitifs tels qu'un bénéfice de l'option d'acquisition de l'information décroissant dans le coût d'achat de cette information et du niveau de connaissance a priori sur l'apparition d'une technologie élevée ou croissant dans la qualité de l'information obtenue. Ces résultats peuvent être déduits des diagrammes de décision, Figure 4.4.

Par contre, nous avons trouvé un résultat non trivial concernant l'effet de la fréquence d'apparition des nouvelles technologies sur la politique optimale. Si aucune information complémentaire ne peut être donnée sur la probabilité d'apparition p_n , l'augmentation de cette dernière n'implique pas nécessairement un comportement monotone sur la décision d'acquisition. En effet, on aurait pu penser que, par exemple, lorsqu'une augmentation de cette probabilité entraînerait un renforcement (ou inversement) de la décision d'acquisition d'information sur la caractéristique de cette nouvelle technologie. On peut voir sur la Figure 4.5 que, pour des états très dégradés $x = 5$, on observe ce comportement monotone en la probabilité d'apparition p_n alors que ce n'est pas le cas pour des états moins dégradés $x = 2$ surtout lorsqu'on a un fort a priori sur l'efficacité de la nouvelle technologie.

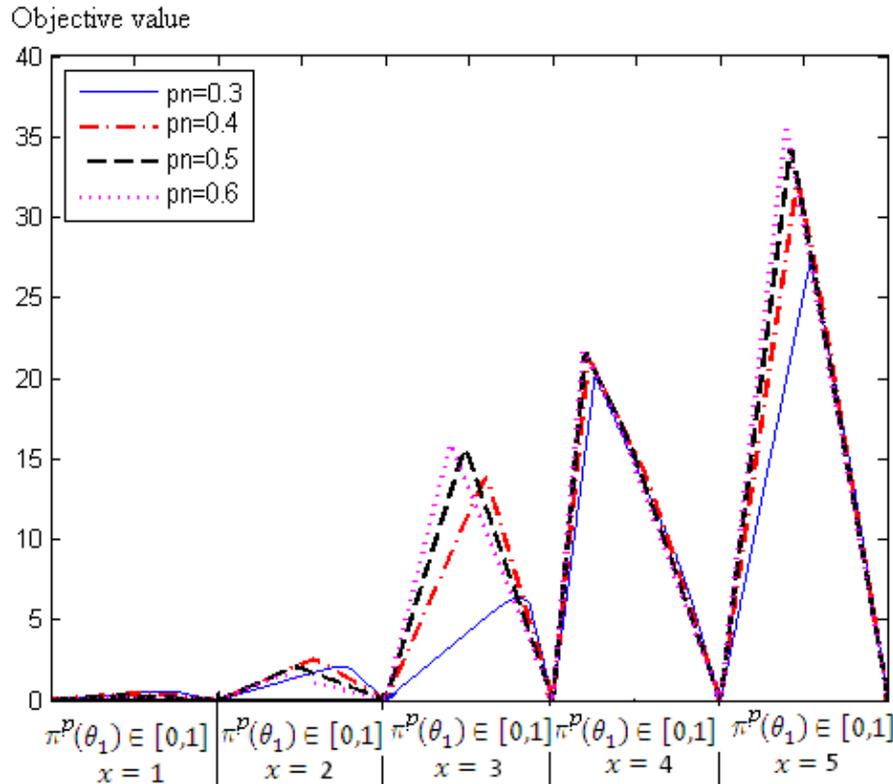


FIGURE 4.5 – Impact de la probabilité d’apparition de la nouvelle technologie sur la valeur de l’option d’acquisition de l’information A

6 Conclusion et perspectives

6.1 Conclusion

En terme de conclusion, je vais revenir rapidement sur les contributions principales de notre travail, ces contributions principalement liées au domaine de l’optimisation de la décision de maintenance dans un contexte d’évolution technologique incertaine.

La plupart des travaux rencontrés dans le contexte de l’optimisation de la maintenance repose sur une hypothèse de stationnarité des performances des systèmes induisant par la suite des renouvellements par des items de même génération et de disponibilité à tout moment de ces ressources. Cette problématique d’évolution technologique est abordée par ailleurs dans des travaux orientés Sciences Économiques et de Gestion. Ces travaux se focalisent sur la détermination de stratégies d’investissement sans tenir compte d’éventuels avantages liés à des stratégies de maintenance. De mon point de vue, la contribution principale de ce travail est donc l’étude conjointe de modèles décisionnels alliant préoccupations opérationnelles de maintenance et objectifs stratégiques d’investissement. Nous avons pu, en abordant un ensemble de problématiques dont la prise en compte de la disponibilité des moyens de maintenance et leur compatibilité avec les systèmes courants, mettre en évidence les interactions entre ces deux champs.

L’analyse des modèles mathématiques que nous avons proposés permet met en évidence des règles de gestion dans le domaine ciblé. Il est bien évident que les conclusions dépendent de la qualité des modèles et des hypothèses associées. Je suis tout à fait conscient des limites de notre approche mais, comme je l’ai marqué précédemment, je reste persuadé de l’utilité de notre démarche pour consolider l’opinion d’un décideur en offrant une estimation du coût des différentes options dans un contexte particulièrement incertain. Dans ce contexte de forte

incertitude, l'évaluation du niveau de connaissance pour une prise de décision adéquate offre un réel intérêt et ouvre de bonnes perspectives notamment vis-à-vis de la question d'estimation de l'état de santé d'un système par le biais de plusieurs modes d'auscultation. Ce point sera discuté dans le paragraphe suivant.

6.2 Perspectives

Les premières perspectives qui pourraient être abordées sont liées aux hypothèses pouvant paraître trop restrictives au vu du contexte décisionnel réel. Dans ce sens, il serait par exemple intéressant d'améliorer les relations entre la perte de productivité du système à maintenir et les fluctuations du marché. Une approche envisagée est l'utilisation de processus stochastiques *composés* tels que le processus Variance Gamma associant processus Gamma et mouvement Brownien, le processus subordonneur Gamma caractérisant l'évolution de la dégradation et le mouvement Brownien les fluctuations du marché. Une telle modélisation permettrait de différencier les effets des actions s'appliquant sur le système ou sur le marché.

Une autre piste de travail est, dans le cadre d'une évolution technologique séquentielle, de combiner le modèle d'amélioration technologique, caractérisée par de petites évolutions à des fréquences élevées, au modèle dit de saut technologique. Une telle approche pourrait être modélisée sur le plan de la gestion des pièces de rechange par la possibilité pour le premier cas de compatibilité avec les nouvelles générations alors qu'un saut majeur entraînerait l'obsolescence de ces rechanges.

Comme souligné dans le paragraphe précédent, je trouve les questions autour de la qualité de l'information et de l'acquisition de connaissance fort intéressantes au vu des intérêts manifestes des communautés scientifiques à la convergence des problématiques du *Structural Health Monitoring* ou encore du *Prognostic Health Management* et celles de l'optimisation de la maintenance conditionnelle. Sur le plan de la qualité d'information, les modèles de dégradation sont trop souvent supposés connus alors que le nombre et la qualité des données sont généralement insuffisantes pour assurer une bonne estimation des lois surtout dans des contextes opératoires particuliers. Dans un contexte de durée de vie, de nombreux travaux sont développés pour améliorer ces modèles basés sur l'intégration de nouvelles observations reposant sur le principe de l'actualisation bayésienne mais aucun, à ma connaissance, ne l'est dans un contexte de processus de dégradation cumulé. Dans le même ordre d'idée, bien que les techniques de surveillance permettent de renvoyer une information en continue sur l'état du système par rapport à un mode de défaillance donné, l'inspection in situ par des opérateurs de maintenance reste de rigueur. La définition de politiques d'inspection combinant ces deux modes d'acquisition reposant sur l'estimation de la qualité d'information à chaque instant me semble une piste d'intérêt.

Enfin, si on revient sur la méthodologie employée, notre travail peut être vu comme une première approche pour l'introduction du concept d'options réelles pour l'optimisation de la maintenance. Il pourrait être très intéressant de chercher à intégrer de nouvelles options permettant ainsi de mieux refléter le comportement du décideur. Ces intégrations pourraient, par exemple, permettre d'intégrer de nouvelles préoccupations comme la capacité de l'entreprise à investir, la constitution de fonds propres pour l'investissement et la maintenance et ainsi, de mon point de vue, renforcer l'intérêt de la maintenance dans les stratégies d'investissement.

Chapitre 5

Conclusion et Perspectives Scientifiques

1 Conclusion

Sans chercher à rappeler l'ensemble des contributions que j'ai mentionné tout au long de ce manuscrit, je les positionnerai, dans cette conclusion, par rapport à une analyse des attentes réciproques entre problématiques industrielles et recherche académique dans les domaines appliqués. Il est clair que chacun est dans le même temps prescripteur et client de l'autre renforçant naturellement les liens entre domaines. Dans ce contexte, la maintenance est reconnue comme l'une des activités industrielles pour laquelle le gap entre besoins industriels et production scientifique est le plus large (on peut reprendre l'analyse de Scarf [95] à cet égard). Vis-à-vis de ce constat, ma recherche se positionne à deux niveaux :

- Le premier niveau est lié à une recherche de type académique pour la construction de nouvelles approches en maintenance. Cette recherche a donc comme finalité de faire évoluer les pratiques courantes. Sur ce point, je pense que mes travaux contribuent dans une certaine mesure à cette finalité par la proposition de nouvelles politiques de maintenance conditionnelle, l'amélioration d'approches existantes par intégration de nouvelles contraintes, de variables d'influence de processus de décision ou d'hypothèses moins restrictives. La méthodologie générale passe par les étapes de construction de nouveaux problèmes qui semblent tout à fait réalistes et de résolution en proposant des règles de décision cohérentes déclinées sous forme de bonnes pratiques.
- Le second niveau s'apparente à une approche motivée par le problème industriel. L'étape de définition du problème permet une appropriation du problème par chacun suivi d'une convergence qu'elle se passe au niveau du vocabulaire comme des intérêts. L'analyse et la formulation suivantes permettent de définir l'ensemble des hypothèses confrontées à l'expertise métier. La formulation de ces hypothèses est en soi une contribution et participe grandement à l'identification des verrous scientifiques associés. Enfin, l'analyse des résultats ainsi que leur représentation se doivent d'être orientés utilisateur afin d'en assurer l'appropriation et l'acceptabilité pour une efficacité maximale.

Les contributions des chapitres 1 et 3 se positionnent plutôt sur le premier point alors que celles du chapitre 2 sont principalement sur le second.

2 Perspectives scientifiques

Outre les perspectives scientifiques liées aux extensions des méthodes discutées auparavant dans ce manuscrit, les thèmes et l'organisation de mes recherches futures peuvent être résumés au travers des projets, qu'ils soient bien initiés, en phase de lancement ou encore en phase de dépôt.

2.1 Construction de méta-modèles pour l'optimisation de la maintenance des structures et infrastructures

Ce thème est lié au projet Stratégie d'Identification de Méta-Modèle pour la Maintenance de structures -SI3M- qui vient d'être accepté pour financement par la Région Pays de la Loire. Il est porté par Franck Schoefs, Professeur au laboratoire de Génie Mécanique et Génie Civil de Nantes, et est en collaboration avec l'IFSTTAR et le Department of Engineering du Trinity College of Dublin.

Le projet vise à proposer une nouvelle stratégie de maintenance pour les structures dégradées à partir de résultats de Contrôle Non Destructif (CND) via des méta-modèles. On entend par méta-modèles des modèles à faible nombre de paramètres reposant sur l'expertise physique et la pertinence probabiliste d'une part et sur les indicateurs de dégradation et de durabilité directement accessibles à partir de contrôles non destructifs CND d'autre part. Ils ont donc la particularité d'avoir en entrée à la fois des indicateurs de dégradations et des indicateurs de durabilité, les deux étant présentés sous forme de processus stochastiques et ayant en commun de reposer sur deux états : initiation et propagation. Pour être complet et pertinent, on intégrera à la fois la variabilité du matériau et celle de la mesure (erreur de mesure) dans le projet. Ce projet s'inscrit donc en rupture avec les recherches existantes où la mesure in situ est considérée comme parfaite et où l'optimisation repose sur des modèles de dégradations multi-physiques et multi-paramétrés qui ne peuvent actuellement être alimentés par des CND sur structures existantes. Il a été rendu possible par la mise en réseau de compétences d'horizons divers au sein du projet d'émergence collective des Pays de la Loire : ECND-PdL.

2.2 Maintenance et performance industrielle

Fortement liée à mes activités actuelles au sein de l'équipe SLP et principalement avec David Lemoine, l'objectif de ces recherches est d'améliorer l'intégration et la prise en compte de la maintenance dans de nombreux processus industriels notamment avec la prise en compte de processus connexes tels que la gestion des pièces de rechange ou encore les aspects planification de la production.

Optimisation conjointe planification tactique de la production et maintenance : Cette étude initiée en 2010 a pour objectif de renforcer l'applicabilité et la robustesse des méthodes de construction de plans de production face aux aléas, principalement liés à la maintenance.

Maintenance et chaîne logistique L'objectif de ce thème est l'exploration et le développement de nouveaux modèles d'optimisation et de méthodes pour construire une chaîne robuste d'approvisionnement de pièces de rechange (c'est-à-dire qui minimise le coût de distribution et d'exploitation tout en assurant les objectifs de qualité de service définis contractuellement, et prenant en compte les incertitudes sur les données). D'un point de vue maintenance, on cherchera à rapprocher stratégies de gestion des pièces de rechange et politiques de maintenance. Ce thème est actuellement motivé au travers d'une collaboration avec General Electric Healthcare initiée par une thèse CIFRE qui a débuté en avril 2012.

2.3 Nouveaux domaines d'application de la maintenance : les éoliennes offshore

Ce thème est fortement motivé par l'émergence de l'activité *éoliennes offshore* dans les régions Bretagne et Pays de Loire. Dans ce cadre, nous sommes en voie de dépôt d'un projet *Hyperwind* avec un ensemble d'industriels (Keops Performance, JSPM AREVA, ASTRIUM, etc.) et le LINA (Laboratoire d'Informatique de Nantes - Philippe Leray).

Au niveau de la thématique maintenance, il est attendu plusieurs contributions qu'elles se situent au niveau de la problématique d'exploitation ou des outils de modélisation et d'optimisation. Concernant les résultats sur le champ exploitation, on décomposera la production de méthodes de

1. Planification robuste et dynamique des maintenances correctives et préventives pour une éolienne avec données de dégradation hétérogènes sous conditions opératoires et intervalles de réalisation incertains.
2. Synchronisation des politiques de maintenance pour la gestion du champ d'éoliennes à ressources partagées en fonction des objectifs de production.

Sur le champ des outils, les développements devraient porter sur les aspects modélisation et plus particulièrement sur l'intégration d'approches dynamiques stochastiques dans des méthodes graphiques telles que les réseaux bayésiens.

Bibliographie

- [1] Mil-hdbk 217 f : Reliability Prediction of Electronic Equipment. Technical report, Military Handbook - Department of Defense - USA, 1991.
- [2] Ute c 80-810 : Recueil de Données de Fiabilité - modèle universel pour le calcul de la fiabilité prévisionnelle des composants, cartes et équipements électroniques. Technical report, Norme internationale - Union Technique de l'Electricité, 2005.
- [3] Des Chaussées à Longue Durée de Vie pour Routes à Forte Circulation. Technical report, Rapport du Centre de Recherche sur les Transports - OCDE/FIT - ISBN 978-92-821-0161-2, 2008.
- [4] Gestion de la Sûreté de Fonctionnement - partie 3-16 : Guide d'application - lignes directrices pour la spécification des services de support de maintenance. Technical report, Norme internationale - Commission Electrotechnique Internationale - ISBN 978-2-88910-100-9, 2008.
- [5] S. Özekici, editor. *Reliability and Maintenance of Complex Systems*. NATO ASI Series - Series F : Computer and Systems Sciences - vol. 154. Springer, 1996.
- [6] V. Bagdonavicius and M. Nikulin. Estimation in Degradation Models with Explanatory Variables. *Lifetime Data Analysis*, 7 :85–103, 2000.
- [7] F. Barbera, H. Schneider, and E. Watson. A Condition Based Maintenance Model for a Two-Unit Series System. *European Journal of Operational Research*, 116(2) :281–290, 1999.
- [8] J.C Bean, J.R Lohmuann, and R.L. Smith. Equipment Replacement under Technological Change. *Naval Research Logistics*, 41 :117–128, 1994.
- [9] E. Berk and K. Moinzadeh. Analysis of Maintenance Policies for m Machines with Deteriorating Performance. *IIE Transactions*, 32 :433–444, 2000.
- [10] D. P. Bertsekas. *Dynamic Programming : Deterministic and Stochastic Models*. Prentice-Hall, 1987.
- [11] G. Bethuyne. The timing of technology adoption by a cost-minimizing firm. *Journal of Economics*, 76(2) :123–154, 2002.
- [12] E. Borgonovo, M. Marseguerra, and E. Zio. A monte carlo methodological approach to plant availability modeling with maintenance, aging and obsolescence. *Reliability Engineering and System Safety*, 67(1) :61–73, 2000.
- [13] I.W. Burr. *Statistical Quality Control Methods*. Marcel Dekker, 1976.
- [14] B. Castanier. *Modélisation Stochastique et Optimisation de la Maintenance Conditionnelle des Systèmes à Dégradation Graduelle*. PhD thesis, Université de Technologie de Troyes - France, 2001.
- [15] B. Castanier, C. Bérenguer, and A. Grall. Modèle d'évaluation des Performances d'une Politique de Maintenance Conditionnelle avec Prise en Compte de l'Indisponibilité. In *IV-ième Congrès International de Génie Industriel - june, 11-15, 2001, Marseille, France*, pages 459–467, 2001.

- [16] B. Castanier, C. Bérenguer, and A. Grall. Stochastic Maintenance Planning for a Repairable System which is Inoperative during Maintenance Operations. In *Proceedings of six PSAM (Probability and Safety Assessment Management) Conference, San Juan, Puerto Rico, USA*, volume 2, pages 1365–1370, June 2002.
- [17] B. Castanier, C. Bérenguer, and A. Grall. A Sequential Condition-Based Repair/Replacement Policy with Non-Periodic Inspections for a System subject to Continuous Wear. *Appl. Stochastic Models in Bus. Ind.*, 19 :327–347, 2003.
- [18] B. Castanier, A. Grall, and C. Bérenguer. A Condition-Based Maintenance Policy with Non-Periodic Inspections for a Two-Unit Series System. *Reliability Engineering and System Safety*, 87(1) :109–120, 2005.
- [19] B. Castanier and M. Rausand. Maintenance Optimization for Subsea Oil Pipelines. *International Journal of Pressure Vessels and Piping*, 83(4) :236–243, 2006.
- [20] T. Cheevaprawatdomrong and R. L. Smith. A paradox in equipment replacement under technological improvement. *Operations Research Letters*, 31(1) :77–82, 2003.
- [21] A.H. Christer. Delay Time Models of Industrial Inspections Maintenance Problems. *Journal of the Operational Research Society*, 35 :401–406, 1984.
- [22] A.H. Christer and A. Wang. A Simple Condition Monitoring Model for a Direct Monitoring Process. *European Journal of Operational Research*, 82 :258–269, 1995.
- [23] J. Clavareau. *Modélisation des stratégies de remplacement de composants et de systèmes soumis à obsolescence technologique*. PhD thesis, Université Libre de Bruxelles, Faculté des Sciences Appliquées, 2008.
- [24] J. Clavareau and P. E. Labeau. Maintenance and replacement policies under technological obsolescence. *Reliability Engineering and System Safety*, 94 :370–381, 2009.
- [25] J. Clavareau and P. E. Labeau. A petri net-based modeling of replacement strategies under technological obsolescence. *Reliability Engineering and System Safety*, 94 :357–369, 2009.
- [26] C. Coccozza-Thivent. *Processus Stochastiques et Fiabilité des Systèmes*. Springer-Verlag, 1997.
- [27] C. Coccozza-Thivent. Convergence de Fonctionnelles de Processus Semi-Régénératifs. *Prépublication de l'Université de Marne-La-Vallée*, 02/2000, March 2000.
- [28] D.R Cox. Regression Models and Life Tables. *Journal of the Royal Statistics*, B(34) :187–202, 1972.
- [29] R. Dekker, H. Frenk, and E. Wildeman. How to Determine Frequencies for Multicomponent Systems? a General Approach. In Özekici [5], pages 239–280.
- [30] R. Dekker, F. van der Duyn Schouten, and E. Wildeman. A Review of Multicomponent Maintenance Models with Economic Dependence. *Mathematical Methods of Operations Research*, 45 :411–435, 1997.
- [31] E. Deloux, B. Castanier, and C. Bérenguer. Predictive Maintenance Policy for a Gradually Deteriorating System Subject to Stress. *Reliability Engineering and System Safety*, 94(2) :418–431, 2009.
- [32] D.I. and M.A. Parlar. Survey of Maintenance Models for Multi-Unit Systems. *European Journal of Operational Research*, 51 :1–23, 1991.
- [33] L. Doyen and O. Gaudoin. Classes of Imperfect Repair Models based on Reduction of Failure Intensity or Virtual Age. *Reliability Engineering and System Safety*, 84(1) :45–56, 2004.
- [34] E. J. Elton and M. J. Gruber. On the Optimality of an Equal Life Policy for Equipment subject to Technological Improvement. *Operational Research Quarterly*, 27(1) :93–99, 1976.

- [35] M. Fouladirad, A. Grall, and L. Dieulle. On the Use of On-Line Detection for Maintenance of Gradually Deteriorating Systems. In Taylor & Francis, editor, *European Safety and Reliability Conference - ESREL 2006*, September, 18-22 2006.
- [36] T. Goldstein, S. P. Ladany, and A. Mehrez. A Dual Machine Replacement Model : A Note on Planning Horizon Procedures for Machine Replacements. *Operations Research*, 34 :938–941, 1986.
- [37] Z. Goldstein and A. Mehrez. Replacement of Technology when a new Technological Breakthrough is Expected. *Engineering Optimization*, 27(3) :265–278, 1996.
- [38] G. Haddad, P. Sandborn, and M. Pecht. Using Real Options to Manage Condition-Based Maintenance Enabled by phm. In *IEEE International Conference on Prognostics and Health Management - Denver, Colorado*, June 2011.
- [39] A. Haurie and P. L'Ecuyer. A Stochastic Control Approach to Group Preventive Replacment in a Multicomponent System. *IEEE Tansactions on Automatic Control*, 27(2) :387–393, 1982.
- [40] W. J. Hopp and S. K. Nair. Markovian deterioration and technological change. *IIE Transactions*, 26 :74–82, 1974.
- [41] R. A. Howard. *Dynamic Programming and Markov Decision Processes*. MIT Press, Cambridge, Massachusetts, 1960.
- [42] N. Hritonenko and Y. Yatsenko. Optimal Equipment Replacement without Paradoxes : a Continuous Analysis. *Operations Research Letters*, 35(2) :245–250, 2007.
- [43] N. Hritonenko and Y. Yatsenko. The dynamics of asset lifetime under technological change. *Operations Research Letters*, 36 :565–568, 2008.
- [44] K. J. M. Huisman and P. M. Kort. Strategic Technology Adoption taking into account Future Technological Improvements : A Real Options Approach. *European Journal of Operational Research*, 159(3) :705–728, 2004.
- [45] K. Itô. *Encyclopedic Dictionary of Mathematics*, volume 1. second edition. MIT Press, Massachussets, and London, England, 1996.
- [46] X. Jin, L. Li, and J. Ni. Options model for joint production and preventive maintenance system. *International Journal of Production Economics*, 119 :347–353, 2009.
- [47] N. Kaio and S. Osaki. Inspection Policies : Comparisons and Modifications. *Operations Research*, 22(4) :387–400, 1988.
- [48] U. Küchler and S. Tappe. Bilateral Gamma Distributions and Processes in Financial Mathematics. *Stochastic Processes and their Applications*, 18 :261–283, 2008.
- [49] W. J. Kennedy, J. W. Patterson, and L. D. Fredendall. An Overview of Recent Literature on Spare Parts Inventories. *International Journal of Production Economics*, 76 :201–215, 2002.
- [50] M. Kijima. Some Results for Repairable Systems with General Repair. *Journal of Applied Probability*, 26 :89–102, 1989.
- [51] M. Kijima, H. Morimura, and Y. Suzuki. Periodical Replacement Problem Without Assuming Minimal Repair. *European Journal of Operational Research*, 37 :194–203, 1988.
- [52] C. Krychowski. *Apport et Limites des Options Réelles à la Dcision d'Investissement Stratégique : Une Étude Appliquée dans le Secteur des Télécommunications*. PhD thesis, Ecole des Hautes Études Commerciales - France, 2007.
- [53] G. Last and R. Szekli. Stochastic Comparison of Repairable Systems by Coupling. *Journal of Applied Probability*, 35 :348–370, 1998.

- [54] D. Lautier. Les Options Réelles : une idée séduisante - un concept utile et multiforme - un instrument facile à créer mais difficile à valoriser. Technical report, Université Paris IX, 1ère version, 2001.
- [55] A. Lehmann. Joint Modeling of Degradation and Failure Time Data. In *Accelerated Life Testing - ALT 2006- Angers, France*, 2006.
- [56] M. Lemaire. *Fiabilité des Structures*. Hermès-Lavoisier, 2005.
- [57] M. Lévesque and M. L. Maillart. Business opportunity assessment with costly, imperfect information. *IEEE Transactions on Engineering Management*, 55 :279–291, 2008.
- [58] W.S. Lovejoy. Computationally feasible bounds for partially observed markov decision processes. *Operations Research*, 39(1) :162–175, 1991.
- [59] K. E. Mac Cardle. Information acquisition and the adoption of new technology. *Management Science*, 31 :1372–1389, 1985.
- [60] V. Makis and A.K.S. Jardine. Optimal Replacement in the Proportional Hazards Model. *Information Systems and Operations Research*, 30(1) :172–183, 1992.
- [61] M.A.K. Malik. Reliable Preventive Maintenance Scheduling. *AIIE Transactions*, 11(3) :221–228, 1979.
- [62] E. R. Mc Grattan and J. A. Schimtz. Maintenance and Repair : Too Big to Ignore. *Federal Reserve Bank of Minneapolis Quarterly Review*, 23 :2–13, 1999.
- [63] K. E. McKone, R. G. Schroeder, and K. O. Cua. An overhaul-replacement model for equipment subject to technological change in an inflation-prone economy. *International Journal of Production Economics*, 56 :291–301, 1998.
- [64] K. E. McKone, R. G. Schroeder, and K. O. Cua. Total Productive Maintenance : A Contextual View. *Journal of Operations Management*, 17 :123–144, 1999.
- [65] S. Mercier. Optimal Replacement Policy for Components with General Failure rates submitted to Obsolescence. In *ESREL 2008-European Safety and Reliability Conference, Valencia, Spain*, pages 603–610, September 2008.
- [66] S. Mercier. Optimal replacement policy for components with general failure rates submitted to obsolescence. *Applied Stochastic Models in Business and Industry*, 24 :221–235, 2008.
- [67] O. Michel, P.E. Labeau, and S. Mercier. Monte carlo optimization of the replacement strategy of components subject to technological obsolescence. In *Proceedings of the joint Probabilistic Safety Assessment Management Conference - PSAM 7 - and the European Safety and Reliability Conference - ESREL'2004, Berlin, Germany*, volume 6, pages 282–297, 2004.
- [68] L. Miller and C. Park. Economic analysis in the maintenance, repair, and overhaul industry : an options approach. *the Engineering Economist*, 49 :21–41, 2004.
- [69] G. E. Monahan. Optimal stopping in a partially observable binary- valued markov chain with costly perfect information. *Journal of Applied Probability*, 19 :72–81, 1982.
- [70] A.G. Munford. Comparison Among Certain Inspection Policies. *Management Science*, 27(1) :260–267, March 1981.
- [71] S.C. Myers. Determinants of Corporate Borrowing. *Journal of Financial Economics*, 5 :147–175, 1977.
- [72] J.A. Nachlas. *Reliability Engineering : Probabilistic Models and Maintenance Methods*. Taylor & Francis, 2006.
- [73] S. H. Nair. Modeling Strategic Investment Decisions under Sequential Technological Change. *Management Science*, 41 :282–297, 1995.

- [74] L. C. Neves, D. M. Frangopol, and P. S. Cruz. Cost of Life Extension of Deteriorating Structures under Reliability-Based Maintenance. *IMA Journal of Management Mathematics Computers and Structures*, 82 :1077–1089, 2004.
- [75] Martin Newby. Bivariate decision processes. In Mikhail Nikulin, Daniel Commenges, and Catherine Huber, editors, *Probability, Statistics and Modelling in Public Health*, pages 364–377. Springer US, 2006.
- [76] T. P. K. Nguyen, B. Castanier, and T. G. Yeung. Optimal Maintenance and Replacement Decisions under Technological Change with Consideration of Spare Parts Inventories. In revision, 2012.
- [77] T. P. K. Nguyen, B. Castanier, and T. G. Yeung. Maintenance Investment Problem under Technological Change with Forecast Horizon Optimal Decisions. *European Journal of Operational Research*, Submitted in 2012.
- [78] T. P. K. Nguyen, T. G. Yeung, and B. Castanier. Impact of Maintenance on Replacement Investment under Technological Improvement. In *Proceedings of the European Safety and Reliability Conference - ESREL'2011 - Troyes, France*, pages 869–876, 2011.
- [79] T. P. K. Nguyen, T. G. Yeung, and B. Castanier. Information Acquisition of New Technology Performance for Maintenance/Investment Decisions. In *Proceedings of the 3rd IEEE International Conference on Emergency Management and Management Sciences - Beijing, China*, pages 31–34, August 2012.
- [80] T. P. K. Nguyen, T. G. Yeung, and B. Castanier. Information Acquisition of New Technology Performance for Maintenance/Investment Decisions. In *25th European Conference on Operational Research - Vilnius, Lithuania*, 2012.
- [81] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Optimal Maintenance and Replacement Decisions under Technological Change. In *Proceedings of the European Safety and Reliability Conference - ESREL'2010 - Rhodes, Greece*, September 2010.
- [82] P. K. Nguyen Thi, T. G. Yeung, and B. Castanier. Optimal Maintenance and Replacement Decisions under Technological Change with Consideration of Spare Parts Inventories. In *Proceedings of the 16th International Symposium on Inventories - Budapest, Hungary*, August 2010.
- [83] R. P. Nicolai. *Maintenance Models for Systems subject to Measurable Deterioration*. PhD thesis, Erasmus Rotterdam University - The Netherlands, 2008.
- [84] R.P Nicolai and R. Dekker. Optimal Maintenance of Multi-component System : A Review. In K.A.H Kobbachi and D.N Prabhakar Murthy, editors, *Complex System Maintenance Handbook*, pages 263–286. Springer Series in Reliability Engineering, 2008.
- [85] S. H. Ott and D. C. Mauer. Investment under Uncertainty : The Case of Replacement Investment Decisions. *The journal of Financial and Quantitative Analysis*, 30 :581–605, 1995.
- [86] S. Özekici. Optimal Periodic Replacement of Multicomponent Reliability Systems. *Operations Research*, 36(4) :542–552, 1988.
- [87] P. Paris and F. Ergodan. A Critical Analysis of Crack Propagation Laws. *Journal of Basic Engineering*, 85 :528–534, 1963.
- [88] W.P. Pierskalla and J.A. Voelker. A Survey of Maintenance Models : The Control and Surveillance of Deteriorating Systems. *Naval Research Logistics Quarterly*, 23 :353–388, 1979.
- [89] J. Pineau, G. Gordon, and S. Thrun. Point-Based Value Iteration : An Anytime Algorithm for pomdps. In *International Joint Conference on Artificial Intelligence (IJCAI). Acapulco, Mexico*, pages 1025–1032, 2003.
- [90] M.L. Puterman. *Markov Decision Process : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1994.

- [91] S. Rajagopalan. Adoption Timing of New Equipment with Another Innovation Anticipated. *IEEE Transactions on engineering management*, 46(1) :14–25, 1999.
- [92] J. L. Rogers and J. C. Hartman. Equipment Replacement under Continuous and Discontinuous Technological Change. *IMA Journal of Management Mathematics*, 16(1) :23–36, 2005.
- [93] S.M. Ross. *Stochastic Processes*. Wiley, Second Edition, 1996.
- [94] B. Saassouh, L. Dieulle, and A. Grall. Online Maintenance Policy for a Deteriorating System with Random Change of Mode. *Reliability Engineering and System Safety*, 92 :1677–1685, 2007.
- [95] P.A. Scarf. On the Application of Mathematical Models in Maintenance. *European Journal of Operational Research*, 99(3) :493–506, June 1997.
- [96] I. E. Schochetman and R. L. Smith. Infinite horizon optimality criteria for equipment replacement under technological change. *Operations Research Letters*, 35(4) :485–492, 2007.
- [97] N.D. Singpurwalla. The Hazard Potential : Introduction and Overview. *Journal of the American Statistical Society*, 101(476) :1705–1717, 2006.
- [98] N.D. Singpurwalla and S. Wilson. Failure Models Indexed by Two Scales. *Advanced in Applied Probability*, 30 :1058–1072, 1998.
- [99] J. S. Song and P. H. Zipkin. Managing Inventory with the Prospect of Obsolescence. *Operations Research*, 44 :215–224, 1993.
- [100] C. M. Stewart. Determinants of Coporate Borrowing. *Journal of Financial Economics*, 5 :147–175, 1977.
- [101] L.C. Thomas. A Survey of Maintenance and Replacement Models for Maintainability and Reliability of Multi-Item Systems. *Reliability Engineering*, 16 :297–309, 1986.
- [102] T. Tinga. Application of Physical Failure Models to Enable Usage and Load Based Maintenance. *Reliability Engineering and System Safety*, 95 :1061–1075, 2010.
- [103] F. Turco and P. Parolini. A Nearly Optimal Inspection Policy for Productive Equipment. *International Journal Production Research*, 22(3) :515–528, 1984.
- [104] C. Ulu and J. E. Smith. Uncertainty, information acquisition and technology adoption. *Operation Research*, 57 :740–752, 2009.
- [105] P. Vaidya and M. Rausand. Remaining Useful Life, Technical Health, and Life Extension. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 225(2) :219–231, 2011.
- [106] C. Valdez-Flores and R.M. Feldman. A Survey of Preventive Maintenance Models for Stochastically Deteriorating Single-Unit Systems. *Naval Research Logistics*, 36 :419–446, 1989.
- [107] F.A van der Duyn Schouten. Maintenance Policies for Multicomponent Systems : An Overview. In Özekici [5], pages 118–136.
- [108] F.A. van der Duyn Schouten and S.G Vanneste. Analysis and Computation of (n, n) -Strategies for Maintenance of a Two-Component System. *European Journal of Operational Research*, 48 :260–274, 1990.
- [109] J.M. van Noortwijk. A Survey of the Application of Gamma Processes in Maintenance. *Reliability Engineering and System Safety*, 94(1) :2–21, 2009.
- [110] J.M. van Noortwijk, J.A.M van der Weide, M.J. Mallen, and M.D Pandey. Gamma Processes and Peaks-Over-Threshold Distributions for Time-Dependent Reliability. *Reliability Engineering and System Safety*, 92(12) :1651–1658, 2007.

- [111] J. Vatn. A state based model for opportunity based maintenance. In *Proceedings of the joint Probabilistic Safety Assessment Management Conference - PSAM 11 - and the European Safety and Reliability Conference - ESREL'2012 - Helsinki, Finland*, June 2012.
- [112] P. Vignat, M. Avila, F. Duculty, S. Aupetit, M. Slimane, and F. Kratz. Maintenance policy : degradation laws versus hidden markov model availability indicator. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 226(2) :137–155, 2012.
- [113] H. Wang. A Survey of Maintenance Policies of Deteriorating Systems. *European Journal of Operational Research*, 139 :469–489, 2002.
- [114] D.J.D. Wijnmalen and J.A.M. Hontelez. Coordinated Condition-Based Repair Strategies for Components of a Multi-Component Maintenance System with Discount. *European Journal of Operational Research*, 98 :52–63, 1997.
- [115] R.E. Wildeman. *The Art of Grouping Maintenance*. PhD thesis, Erasmus University Rotterdam, 1996.
- [116] R.E. Wildeman, R. Dekker, and A.C.J.M. Smit. A Dynamic Policy for Grouping Maintenance Activities. *European Journal of Operational Research*, 99 :530–551, 1997.
- [117] J. M. Yang, M. Gunaratne, B. Lu, and B. Dietrich. Use of Recurrent markov Chains for Modeling Crack Performance of Flexible Pavements. *Journal of Transportation Engineering*, 131 :861–872, November 2005.
- [118] S.C. Yuang. *A Bivariate Renewal Process and its Application in Maintenance Policies*. PhD thesis, Faculty of Virginia Polytechnic Institute and State University, 1999.
- [119] N.L. Zhang and W. Zhang. Speeding Up the Convergence of Value Iteration in Partially Observable Markov Decision Processes. *Journal of Artificial Intelligence Research*, 14 :29–51, 2001.
- [120] R. Zhou and E.A. Hansen. An Improved Grid-Based Approximation Algorithm for pomdps. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 707–716, 2001.
- [121] M. Zouch. *Décision dans l'Incertain : Application à la Gestion de l'Entretien Routier*. PhD thesis, Université de Nantes, France, 2011.
- [122] M. Zouch, T. G. Yeung, and B. Castanier. Optimal condition-based resurfacing decisions for roads. In *Proceedings of the European Safety and Reliability Conference - ESREL'2010 - Rhodes, Greece*, pages 1379–1384, 2010.
- [123] M. Zouch, T. G. Yeung, and B. Castanier. Optimal Resurfacing Decisions for Road Maintenance : A POMD Perspective. In *Proceedings of the Annual Reliability and Maintainability Symposium (RAMS). Lake Buena Vista, Florida, United States*, January 2011.
- [124] M. Zouch, T. G. Yeung, and B. Castanier. Application of a bivariate deterioration model for a pavement management optimization. In *Proceedings of the Transport Research Arena - TRA 2012 - Athens, Greece*, April 2012.
- [125] M. Zouch, T. G. Yeung, and B. Castanier. Bayesian updating for road maintenance optimization. In *Proceedings of the joint Probabilistic Safety Assessment Management Conference - PSAM 11 - and the European Safety and Reliability Conference - ESREL'2012 - Helsinki, Finland*, June 2012.
- [126] M. Zouch, T. G. Yeung, and B. Castanier. Optimizing Road Milling and Resurfacing Actions. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 226(2) :156–168, 2012.
- [127] M. Zouch, T. G. Yeung, and B. Castanier. A Two-Phase State-Dependent Deterioration Model for Maintenance Optimization. *Naval Research Logistics*, submitted in 2011.

Annexe : Sélection d'articles

A condition-based maintenance policy with non-periodic inspections for a two-unit series system

B. Castanier^a, A. Grall^{b,*}, C. Bérenguer^b

^aIRCCyN/École des Mines de Nantes, Département Automatique et Productique, 4 rue Alfred Kastler, F-44307 Nantes, France

^bUniversité de Technologie de Troyes, ISTIT-CNRS FRE 2732, Équipe Modélisation et Sécurité des Systèmes, 12, rue Marie Curie, BP 2060-10010 Troyes Cedex, France

Received 17 December 2002; received in revised form 30 July 2003; accepted 23 April 2004

Abstract

This paper considers a condition-based maintenance policy for a two-unit deteriorating system. Each unit is subject to gradual deterioration and is monitored by sequential non-periodic inspections. It can be maintained by good as new preventive or corrective replacements. Every inspection or replacement entails a set-up cost and a component-specific unit cost but if actions on the two components are combined, the set-up cost is charged only once. A parametric maintenance decision framework is proposed to coordinate inspection/replacement of the two components and minimize the long-run maintenance cost of the system. A stochastic model is developed on the basis of the semi-regenerative properties of the maintained system state and the associated cost model is used to assess and optimize the performance of the maintenance model. Numerical experiments emphasize the interest of a control of the operation groupings.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Condition-based maintenance; Wear; Multi-unit system; Economic dependences; Stochastic model; Markov renewal process

1. Introduction and problem statement

For a system which gradually deteriorates, it is natural in an attempt to avoid failure and its possible negative effects to consider the system condition in order to carry out preventive maintenance actions. The problem of modeling and optimization of a maintenance policy has been widely approached in the literature when the system degradation can be modeled by a single observable variable [1–3]. However the failure characteristics of a system often need several system variables to be described especially in the case of multi-unit systems. The multi-unit maintenance problem cannot be reduced to single-unit maintenance problems, except if all units are independent of one another.

In case of interactions between components namely economic, structural and/or stochastic dependencies,

the optimal decision on one unit is not necessarily optimal for the whole system [4–6]. Stochastic dependence means that the state of a system component (e.g. its age, degradation rate, degradation level...) influences the states of others whereas structural dependence exists e.g. in case of physically interconnected components when the maintenance of a component affects the state of others. Economic dependence implies the opportunity to reduce an operation cost by grouping a maintenance on several components compared to a maintenance made on every entity separately (economy of scale). This paper focus on condition-based maintenance for multi-component systems in the context of economic dependences. Most of the existing multi-component maintenance models allow groupings of the maintenance tasks [5,7–9], but few of them are proposed in the context of the condition-based maintenance.

The extension of a mono-component maintenance policy to a multi-component system is neither unique nor evident and should not limit itself to a simple juxtaposition of the mono-component models. It can quickly end in complex models and one has to resort to simulation tools [10].

* Corresponding author. Fax: +33-3-2571-5699.

E-mail addresses: bruno.castanier@emn.fr (B. Castanier), antoine.grall@utt.fr (A. Grall), christophe.berenguer@utt.fr (C. Bérenguer).

It becomes very difficult to characterize the resultant policies optimality even if the number of possible maintenance actions on the system is reduced [11–13]. The multi-component maintenance models with economic dependences can be classified on the basis of the planning aspect [5]. From this point of view dynamic models which generate dynamic decisions that may change over the planning horizon differ from stationary models which assume a long-term stable situation and mostly an infinite planning horizon. In a dynamic planning framework Wildeman et al. [6,14] propose a rolling-horizon approach and derive penalty functions which express the costs of shifting a preventive maintenance from a tentatively planned time on the basis of individual periodic preventive maintenance planning per component.

Our study is limited to the class of stationary models and deals more precisely with condition-based opportunist maintenance, [7] which does not necessarily plan maintenance actions in advance. Due to set-up savings, an intervention on a part of the system yields an opportunity for maintenance on the other system parts. The decision rule has to control the influence of the operation groupings in order to reduce the individual operating costs for each component and the global average cost of exploitation. The promising performance of condition-based maintenance policies for single-component systems leads to base the maintenance decision rules for the whole system on the knowledge of the deterioration level of each component. Knowing the deterioration state of a system's component and the maintenance actions made on the rest of the system, the decision rule should allow the determination of the nature of the current operation on this component (replacement or left as it is) as well as the date of next inspection on the system. We consider a maintenance model in the class of the control-limit policies for which the maintenance decision is made by comparison of deterioration level to critical thresholds [15–17]. The classical (n, N) replacement policy presented in Ref. [18] for which the state of the system is continuously observed is extended in this paper to a decision rule which allows aperiodic inspection times.

The objective of this work is to propose a parametric structure of condition-based maintenance rule for a two-unit system and to develop the associated cost model. The suggested methodology which allows the evaluation of the long-run criterion is based on the semi-regenerative properties of the deterioration process of the maintained system. Section 2 is devoted to the description of the system deterioration model and of the possible maintenance actions. The maintenance decision framework is proposed in Section 3 and the associated cost model is developed in Section 4. It is used to assess and optimize the performance of the maintenance model. Numerical examples given in Section 5 allow to emphasize the interest of decision parameters to control the operation groupings.

2. Description of the system

2.1. Stochastic deterioration model

The literature on multicomponent systems has dealt mainly with multi-unit systems with lifetime models for the system components, and failure dependences or interactions [19–22]. In this paper, the system we consider is composed of two components which stochastically, independently and gradually deteriorates. Each component is subject to a continuous accumulation of wear in time which is assumed to be perfectly described by a scalar random variable. The deterioration level of a component cannot influence the way the other deteriorates.

Let $X^{(i)} = (X_k^{(i)})_{k \in \mathbb{N}}$ be the stochastic process describing the deterioration process of the component i on an infinite discrete time grid $(t_k = k\delta t)_{k \in \mathbb{N}}$. The elementary random deterioration increments in a time interval (i.e. between t_k and $t_{k+1} = t_k + \delta t$) are supposed to be non-negative, stationary and exchangeable. Natural candidates for the associated probability density function can be obtained in the class of infinitely divisible distributions, e.g. gamma distributions [23–25]. The exponential and Erlang laws are special cases of gamma distributions and the use of the exponential law makes it easier to further investigate some analytical developments for the evaluation of the maintenance policy. The deterioration increment of component i between t_k and t_{k+1} is modeled by an exponential random variable $\Delta_{(k,k+1)}X^{(i)}$ with parameter α_i . For all $k \in \mathbb{N}$, the probability density function of the elementary deterioration increment for component i is hereafter denoted f_i with $f_i(x) = \alpha_i e^{-\alpha_i x}$. $f_i^{(l)}$ is the l -th convolution of f_i , i.e. the density function of a l -Erlang law with parameter α_i . Discussions and examples of such deterioration models based on damage accumulation are proposed in Refs. [26,27].

The component i is considered as failed as soon as its deterioration level exceeds a critical level L_i . Fig. 1 illustrates the deterioration process of a single component

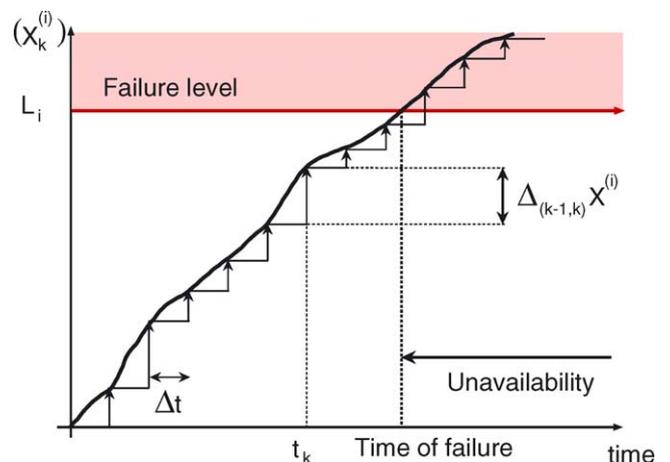


Fig. 1. Example of the evolution of the deterioration process for the component i without maintenance.

assuming it is new at time $t_0=0$, i.e. $X_0^{(i)} = 0$. The failure of a component cannot be directly observed. The component deterioration goes on until the next intervention on the system. Actually a component can be declared ‘failed’ if a pending failure or an important defect or deterioration is present. The system is supposed to be failed as soon as at least one component is declared ‘failed’ even if it does not lead to an obvious complete failure of the system. In practice such a failure model can describe, e.g. complex civil engineering structures (if the resistance of one part of the structure drops below a limit threshold because of a deterioration, the entire structure no more fulfills safety requirements and is declared failed) or production lines (if a system is made of several parts produced by different machines and one of these machines produces a bad-quality component, the system at the end of the production line is a bad-quality one and the production line can be considered as failed).

The global system deterioration evolution is modeled by the bivariate stochastic process $(X_k)_{k \in \mathbb{N}} = (X_k^{(1)}, X_k^{(2)})_{k \in \mathbb{N}}$.

2.2. Maintenance operations and associated costs

The available actions for the system maintenance are the inspection of a component which is the only way to reveal its actual deterioration level and the replacement of a component which can be either a true physical replacement or a repair but always leads to restore a ‘good as new’ component. The time elapsed by the system in failed state incurs a cost at a cost rate c_d . The inspection of a component is supposed to be perfect and non-destructive and every maintenance operation is assumed to be instantaneous. A given component can be replaced independently of the others but an intervention on a component always causes a system stop leading to an inspection hereafter referred as ‘complete inspection of the system’ of all the other components, even if not initially scheduled.

Each intervention performed on the system entails a corresponding unit cost which consists of a specific unit cost and a set-up cost in case of preventive or corrective maintenance operation. Let respectively the specific unit costs be c_i for an inspection of the whole system, $c_p^{(i)}$ for a preventive replacement of the component i and $c_c^{(i)}$ for a corrective replacement of the component i . The set-up cost c_s brings together the inspection cost of the whole system in case of at least one replacement as well as all the additional costs going with the intervention implementation. An intervention on the system potentially requires, e.g. a dismantling and a reassembly of the system, incurs production losses, etc... The set-up cost is assumed to be independent of the operation nature and it is incurred only once in case of several maintenance operations performed at the same time. Hence performing maintenance jointly on several units costs less than on each unit separately. To sum up, the different unit costs corresponding to each intervention on the system are:

- c_i if no replacement is performed;
- $c_s + c_p^{(i)}$ if only the component i is replaced and this replacement is a preventive replacement;
- $c_s + c_p^{(1)} + c_p^{(2)}$ if the entire system is preventively replaced;
- $c_s + c_c^{(i)}$ if only the component i is replaced and this replacement is a corrective replacement;
- $c_s + c_p^{(i)} + c_c^{(j)}$ if a preventive replacement is performed on component i and a corrective replacement on component j ;
- $c_s + c_c^{(1)} + c_c^{(2)}$ if the entire system is correctively replaced.

The individual operating costs by component and the global average cost of exploitation can be reduced by control of the operation groupings. The existence of economic dependencies has to be taken into account by the decision rule which allows to determine simultaneously the nature of the operation to be performed on the system (left as it is, replace the whole system or only replace one component) and the date of the next inspection. The objective is to propose a decision framework which makes it possible to balance operation groupings, preventive and corrective maintenance.

3. Maintenance decision framework

3.1. Structure of the maintenance policy

The parametric structure of the condition-based maintenance policy proposed in this section extends to multi-component systems the multi-threshold maintenance policy proposed in [17,28] for mono-component systems. For each component of the system a condition-based rule defines several regions for the maintenance decisions on the basis of a multi-threshold structure. In the sequel, the decision threshold values for the maintenance of component i ($i \in \{1,2\}$) are denoted $\xi_k^{(i)}$, $k=0, \dots, n_i$ with $\xi_0^{(i)} = 0$. Another threshold ζ_i is added for each component i which defines a zone of ‘opportunistic replacement’. At time t_k , assume x_i denote the observed deterioration level of component i . The maintenance decision has to be taken according to three successive following steps (component level, then system level for maintenance decision and then next inspection time for the whole system) which are:

- Step 1 *Component level*: a first maintenance decision is taken separately for each component according to the multi-threshold control-limit policy. The component i is left as it is if its state belongs to the inspection zone, i.e. $x_i \in [0, \xi_{n_i}^{(i)})$. It is preventively replaced if $x_i \in [\xi_{n_i}^{(i)}, L_i)$ and correctively replaced if $x_i \geq L_i$.
- Step 2 *System level*: the new ‘opportunistic replacement’ threshold ζ_i introduced for the component i allows to

adjust the former decision. If $x_i \geq \zeta_i$ and if a replacement is scheduled on a component $j(j \neq i)$ then a replacement of the component i is simultaneously triggered.

Step 3 Next inspection: finally, the next a priori scheduled intervention date for component i depends on its deterioration level at the end of the maintenance operation denoted x_i^+ . If $x_i^+ \in [\xi_k^{(i)}, \xi_{k+1}^{(i)})$ with $k = 0, \dots, n_i - 1$, it is planned $n_i - k$ time units later. The date of the next inspection on the system is given by the minimum of all the ‘a priori’ units dates. If $x_1^+ \in [\xi_k^{(1)}, \xi_{k+1}^{(1)})$ and $x_2^+ \in [\xi_l^{(2)}, \xi_{l+1}^{(2)})$, the next intervention on the system is planned $\min(n_1 - k, n_2 - l)$ decision periods later.

As a first consequence of the maintenance decision rule structure, taking different numbers of thresholds for the two components is useless. Suppose actually $n_1 \neq n_2$ and for example $n_1 = n_2 + \delta_n$ with $\delta_n > 0$. Depending on its deterioration level at the end of the maintenance operation, the next inspection for component 2 should be scheduled $n_2 - l$ periods later ($l \geq 0$) and at the latest n_2 periods later. Assuming $x_1^+ < \xi_{k+1}^{(1)}$ the next inspection for component 1 should be scheduled at least $n_1 - k$ periods later. As a consequence, the next intervention time for the system is finally planned $\min(n_1 - k, n_2 - l) \leq n_2$ periods later. Due to the decision rule structure, it is impossible to overpass the duration of n_2 periods between two successive inspections of the whole system.

To simplify notations in the sequel, the deterioration level after an operation is no more referred to x_i^+ but directly to its actual deterioration level which can be 0 in case of a replacement and x_i otherwise.

An example of the deterioration evolution of a two-unit system subject to a multi-threshold policy with $n_1 = n_2 = 2$ is proposed in Fig. 2. The different possible decisions which can arise after an inspection are summarized:

- At time S_0 , the system is in the new state ($x_1(S_0) = x_2(S_0) = 0$). An inspection is scheduled $\min(n_1, n_2) = 2$ periods later at time S_1 .
- At time S_1 , the system is inspected. The deterioration levels $x_1(S_1)$ and $x_2(S_1)$ of the two components belong to their respective second inspection zone. No replacement is required and the next inspection date for each component is scheduled one time unit later (at time S_2).
- At time S_2 , a preventive replacement is required for the component 1 ($x_1(S_2) \in [\xi_2^{(1)}, L_1)$). The deterioration level $x_2(S_2)$ of the component 2 is lower than its preventive replacement threshold $\xi_2^{(2)}$ and no natural replacement is required but $x_2(S_2)$ is greater than its opportunistic replacement threshold ζ_2 . The replacement of component 2 is anticipated. After replacement, the system is new and the next intervention is scheduled $n = 2$ time units later.
- At time S_4 , a preventive replacement is required for the component 1 and $x_2(S_4)$ is lower than ζ_2 . Hence, only

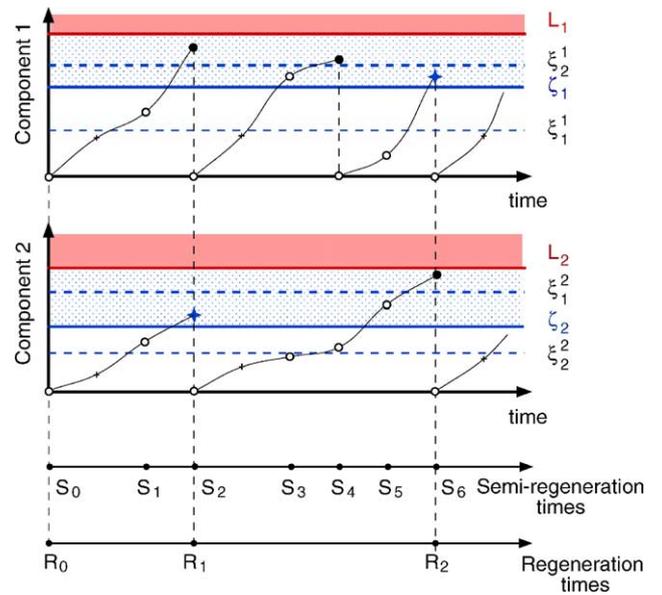


Fig. 2. Example of the deterioration evolution of a two-unit system subject to a multi-threshold policy with $n_1 = n_2 = 2$.

the component 1 is replaced. The next inspection dates are respectively scheduled for component 1 two time units later and for component 2 one-time unit later. Finally, the next intervention on the system is planned one time unit later.

3.2. Comparison with classical multi-component maintenance policies

For some limit threshold configurations, the proposed ‘opportunistic replacement’ maintenance policy can be reduced to more classical maintenance policies. Such a possibility to emulate other policies structures illustrates the adaptability of the structure described in this paper. Examples of limit cases are given hereafter:

- If $\zeta_1 = \zeta_2 = 0$, a complete system replacement is performed as soon as a component is replaced. Hence, the evolution of each component tends to be identical (same operations at the same time). A strong economic dependence lead to have $\zeta_1^* = \zeta_2^* = 0$. Two kinds of policy can be encountered according to the ξ_k^i values configuration:
 - If $n = 1$ and $\xi_1^1 = \xi_2^1$, the system is periodically and systematically inspected and the complete system replacement is performed as soon as a component requires a replacement, i.e. as soon as its age exceeds the threshold value $\xi_1^1 = \xi_2^1$. This model is the ‘modified block replacement model’, [29,30].
 - If $\xi_n^{(i)} = 0$ (hence $\xi_k^{(i)} = 0, \forall k$) the system is completely replaced at each intervention time. The policy is a ‘periodic replacement policy’ and the classical problem associated with this kind of policies is to

determine the complete system replacement date T , i.e. n in the case of the proposed policy in order to optimize a given criterion.

- If $\varsigma_i = \xi_n^i$, the policy is the simple juxtaposition of two mono-component policies with the same number of thresholds. The evolution processes of the two maintained components are mutually independent. Such a policy comes out in the optimal case (i.e. if $\varsigma_1^* = \xi_n^1$ and $\varsigma_2^* = \xi_n^2$) of very weak economic dependences. For each component the proposed maintenance policy can be reduced to more classical one-unit maintenance policies by adopting some specific threshold configurations, e.g. periodic inspection/replacement ($\xi_1 = \dots = \xi_{n-1} = 0$ and $\xi_n \neq 0$), periodic block replacement ($\xi_1 = \dots = \xi_n = 0$), etc. [16].

4. Definition and evaluation of the maintenance policy performance criterion

A cost model based on the long-term average operating cost per unit of time is proposed to assess and optimize the performance of the multi-threshold policy. The long-run average operating cost has been chosen as an optimality criterion and is firstly detailed in this section. The evaluation of such a criterion requires the determination of the stationary laws associated to the deterioration evolution of the maintained system at steady state. It is developed in the second paragraph.

4.1. Long-run average operating cost per time unit

The cumulative operating cost up to time t gathering the running cost and the maintenance costs incurred by the interventions on the system is:

$$C(t) = C_{\text{INSP}}(t) + \sum_{i=1}^2 C_{\text{PREV}}^{(i)}(t) + \sum_{i=1}^2 C_{\text{CORR}}^{(i)}(t) - c_s N_{\text{SR}}(t) + C_{\text{FU}}(t) \quad (1)$$

where:

- $C_{\text{INSP}}(t)$ is the cumulative cost associated with the system inspections on a horizon of length t ;
- $C_{\text{PREV}}^{(i)}(t)$ is the cumulative cost associated with the preventive replacements of the component i independently on the operation performed on the component j (i.e. it includes the cumulative set-up costs);
- $C_{\text{CORR}}^{(i)}(t)$ is the cumulative cost associated with the corrective replacements of the component i independently on the operation performed on the component j (i.e. it includes the cumulative set-up costs);
- $N_{\text{SR}}(t)$ is the total number of complete system replacements performed between 0 and t ;
- $C_{\text{FU}}(t)$ is the cumulative cost incurred by the time elapsed by the system in the failed state (i.e. when at least one

component has failed). As a result of the deterioration measurement at discrete time t_k , the failure time of component i hence of the system is unknown. It is lower bounded by the time t_k just before the failure which allows the evaluation of an upper bound of the time the system has spent in the failed state. More details are given in Appendix A.

The long-run average operating cost per unit of time is defined as follows:

$$C_\infty = \lim_{t \rightarrow \infty} \frac{\mathbb{E}(C(t))}{t} = \lim_{t \rightarrow \infty} \left(\frac{\mathbb{E}(C_{\text{INSP}}(t))}{t} + \sum_{i=1}^2 \frac{\mathbb{E}(C_{\text{PREV}}^{(i)}(t))}{t} + \sum_{i=1}^2 \frac{\mathbb{E}(C_{\text{CORR}}^{(i)}(t))}{t} - c_s \frac{\mathbb{E}(N_{\text{SR}}(t))}{t} + \frac{\mathbb{E}(C_{\text{FU}}(t))}{t} \right) \quad (2)$$

The degradation of the unmaintained system state is described by the stochastic process $(X_k)_{k \in \mathbb{N}}$. Let in the sequel the process $(\tilde{X}_k)_{k \in \mathbb{N}}$ describe the evolution of the maintained system state. It can be analyzed through its regenerative characteristics: after a complete replacement of the system (all the system components are simultaneously replaced), it is in the ‘good as new’ initial state and its future evolution does not depend any more on the past. These complete system replacement times are regenerations points for the process describing the evolution of the global maintained system state. But describing the system state on a renewal cycle can be rather tedious and it can be interesting to take advantage of the semi-regenerative properties of the process $(\tilde{X}_k)_{k \in \mathbb{N}}$. After a global inspection of the system, the evolution of every component and hence of the global system only depends on the revealed deterioration levels. As a consequence every intervention date S_r is a semi-regeneration point for the process describing the maintained system state. Under some mild additional conditions on $(\tilde{X}_k)_{k \in \mathbb{N}}$ given in Ref. [31], the study of the asymptotic behavior of such a semi-regenerative process can be restricted to a semi-regenerative cycle defined by two successive intervention dates.

Let define the discrete time process $(Y_r)_{r \in \mathbb{N}} = (\tilde{X}_{S_r})_{r \in \mathbb{N}}$ describing the maintained system state at the beginning of the maintenance operations. For each $r > 0$, $Y_r = (\tilde{X}_{S_r}^{(1)}, \tilde{X}_{S_r}^{(2)})$ where $\tilde{X}_{S_r}^{(i)}$ describes the state the i th maintained unit at maintenance time. The defined bivariate process $(Y_r)_{r \in \mathbb{N}}$ is an embedded Markov chain of the semi-regenerative process $(\tilde{X}_k)_{k \in \mathbb{N}}$. Let $\pi(x_1, x_2)$ denote in the sequel its stationary law.

From the semi-regenerative properties of the deterioration process $(\tilde{X}_k)_{k \in \mathbb{N}}$, the limit at infinity in Eq. (2) can be changed into a ratio of expectations with respect to the stationary law π over one semi-renewal cycle at steady state:

$$C_\infty = \frac{\mathbb{E}_\pi(C_{INSP}(S))}{\mathbb{E}_\pi(S)} + \sum_{i=1}^2 \frac{\mathbb{E}_\pi(C_{PREV}^{(i)}(S))}{\mathbb{E}_\pi(S)} + \sum_{i=1}^2 \frac{\mathbb{E}_\pi(C_{CORR}^{(i)}(S))}{\mathbb{E}_\pi(S)} - c_s \frac{\mathbb{E}_\pi(N_{SR}(S))}{\mathbb{E}_\pi(S)} + \frac{\mathbb{E}_\pi(C_{FU}(S))}{\mathbb{E}_\pi(S)} \quad (3)$$

where:

- S is the length of one semi-renewal cycle at steady state, i.e. the time between two successive semi-regenerative points of the deterioration process at steady state;
- $\mathbb{E}_\pi(S)$ is the expected length of one semi-regenerative cycle at steady state with respect to π given by:

$$\mathbb{E}_\pi(S) = \sum_{k=1}^n k \left[\sum_{i \neq j} \int_{\xi_{n-k}^i}^{\xi_{n-k+1}^i} \left(\int_0^{\xi_{n-k+1}^j} \pi(x_1, x_2) dx_j \right) + \int_{\xi_n^j}^{\xi_{n-k}^j} \pi(x_1, x_2) dx_j \right] dx_i + \mathbb{1}_{\{k=n\}} \int_{\xi_n^1}^{\xi_n^2} \pi(x_1, x_2) dx_2 dx_1 + \int_{\xi_1}^{\xi_n^1} \int_{\xi_n^2}^{\infty} \pi(x_1, x_2) dx_2 dx_1 \quad (4)$$

- $\mathbb{E}_\pi(C_{INSP}(S))$ is the expected cumulative cost incurred by inspections on the first semi-regenerative cycle (related to the probability to have an inspection at time S)

$$\mathbb{E}_\pi(C_{INSP}(S)) = \int_0^{\xi_n^1} \int_0^{\xi_n^2} c_i \pi(x_1, x_2) dx_2 dx_1; \quad (5)$$

- $\mathbb{E}_\pi(C_{PREV}^{(i)}(S))$ is the expected cumulative cost incurred by the preventive replacement of the component i (component i preventively replaced at S):

$$\mathbb{E}_\pi(C_{PREV}^{(i)}(S)) = \int_{\xi_n^i}^{L_i} \int_0^{\infty} (c_s + c_p^{(i)}) \pi(x_1, x_2) dx_j dx_i + \int_{\xi_i}^{\xi_n^i} \int_{\xi_n^i}^{\infty} (c_s + c_p^{(i)}) \pi(x_1, x_2) dx_j dx_i; \quad (6)$$

- $\mathbb{E}_\pi(C_{CORR}^{(i)}(S))$ is the expected cumulative cost incurred by the corrective replacement of the component i (component i correctively replaced at S):

$$\mathbb{E}_\pi(C_{CORR}^{(i)}(S)) = \int_{L_i}^{\infty} \int_0^{\infty} (c_s + c_c^{(i)}) \pi(x_1, x_2) dx_j dx_i; \quad (7)$$

- $\mathbb{E}_\pi(N_{SR}(S))$ is the probability of a global system replacement (the two components are replaced at S):

$$\mathbb{E}_\pi(N_{SR}(S)) = \mathbb{E}_\pi(\mathbb{1}_{\{\text{Global system replacement at } S\}}) = \int_{\xi_n^1}^{\infty} \int_{\xi_2}^{\infty} \pi(x_1, x_2) dx_2 dx_1 + \int_{\xi_1}^{\xi_n^1} \int_{\xi_n^2}^{\infty} \pi(x_1, x_2) dx_2 dx_1; \quad (8)$$

- $\mathbb{E}_\pi(C_{FU}(S))$ is the expected cumulative cost incurred by the time elapsed by the system in the failed state (at least one component has failed). The mathematical expression of this expectation is given in Appendix A.

4.2. Stationary law of the maintained system state

The maintained system behavior can be characterized by the stationary law π of the Markov chain $(Y_r)_{r \in \mathbb{N}}$. The evaluation of this stationary probability density can be obtained from an exhaustive analysis of all the possible maintenance scenarios on a semi-regenerative cycle. Let (y_1, y_2) and (x_1, x_2) be the system deterioration levels observed at the beginning of two successive maintenance operations, i.e. at two successive semi-regeneration points. The possible scenarios are the followings:

- Scenario 1: $y_1 \geq \xi_2^1$ and $y_2 \geq \zeta_2$ or $y_1 \geq \zeta_1$ and $y_2 \geq \xi_n^2$. The two components are replaced and the next inspection time is scheduled n time units later. The deterioration laws of the two components until the next inspection are respectively $f_1^{(n)}(x_1)$ and $f_2^{(n)}(x_2)$.
- Scenario 2: $y_1 \geq \xi_n^1$, $y_2 \in [\xi_l^2, \xi_{l+1}^2)$, $l \in \{0, \dots, n-1\}$ and $y_2 < \zeta_2$. The component 1 is replaced and the component 2 is left in its state y_2 . The next inspection is scheduled $n-l$ time units later. The deterioration increments x_1 and $x_2 - y_2$ follow the respective deterioration laws $f_1^{(n-l)}(x_1)$ and $f_2^{(n-l)}(x_2 - y_2)$.
- Scenario 3: $y_2 \geq \xi_n^2$, $y_1 \in [\xi_k^1, \xi_{k+1}^1)$, $k \in \{0, \dots, n-1\}$ and $y_1 < \zeta_1$. The component 2 is replaced and the component 1 is left in its state y_1 . The next inspection is scheduled $n-k$ time units later. The respective deterioration laws until the next inspection are $f_1^{(n-k)}(x_1 - y_1)$ and $f_2^{(n-k)}(x_2)$.
- Scenario 4: $y_1 \in [\xi_k^1, \xi_{k+1}^1)$ and $y_2 \in [\xi_l^2, \xi_{l+1}^2)$, k and $l \in \{0, \dots, n-1\}$. No component is replaced. The next inspection time is scheduled $n - \max(k, l)$ time units later. The respective deterioration laws until the next inspection are $f_1^{(n-\max(k,l))}(x_1 - y_1)$ and $f_2^{(n-\max(k,l))}(x_2 - y_2)$.

By integration on the whole state space, the description of the maintenance scenarios can lead to the following expression of the stationary probability density for the deterioration process at operation times:

$$\begin{aligned} \pi(x_1, x_2) = & \underbrace{\left(\int_{\xi_n^1}^{\infty} \int_{\zeta_2}^{\infty} \pi(y_1, y_2) dy_2 dy_1 + \int_{\xi_n^2}^{\infty} \int_{\zeta_1}^{\xi_n^1} \pi(y_1, y_2) dy_1 dy_2 \right)}_{\text{Scenario 1}} f_1^{(n)}(x_1) f_2^{(n)}(x_2) \\ & + \underbrace{\sum_{i \neq j} \sum_{l=0}^{n-1} \left(\int_{\xi_l^i}^{\min(\xi_{l+1}^i, \zeta_j)} \left(\int_{\xi_n^i}^{\infty} \pi(y_1, y_2) dy_i \right) f_j^{(n-l)}(x_j - y_j) dy_j \right)}_{\text{Scenarios 2 and 3}} f_i^{(n-l)}(x_i) \\ & + \underbrace{\sum_{k,l=1}^{n-1} \int_{\xi_k^1}^{\xi_{k+1}^1} \int_{\xi_l^2}^{\xi_{l+1}^2} \pi(y_1, y_2) f_1^{(n-\max(k,l))}(x_1 - y_1) f_2^{(n-\max(k,l))}(x_2 - y_2) dy_2 dy_1}_{\text{Scenario 4}}. \end{aligned} \tag{9}$$

The evaluation of the probability density function $\pi(x_1, x_2)$ is tricky and requires to solve a bi-dimensional one-sided integral equation of the second kind (Eq. (9)). From a numerical point of view, due to the regularity of the densities $f_i^{(p)}$ the integral equation can be solved using an iterative method based on the method of successive approximations (see [32,33] for more details). As an example Fig. 3 illustrates the stationary probability density surface of the deterioration process at operation times for a system with two different components subject to the ‘opportunistic replacement’ maintenance policy. Fig. 3 shows that the pdf increases as the system state deteriorates as long as it remains below the last preventive maintenance threshold, i.e. as the inspection frequency increases. As illustrated, the structure of the decision rule (changes in the inspection frequency) induces a non-differentiability of the stationary pdf.

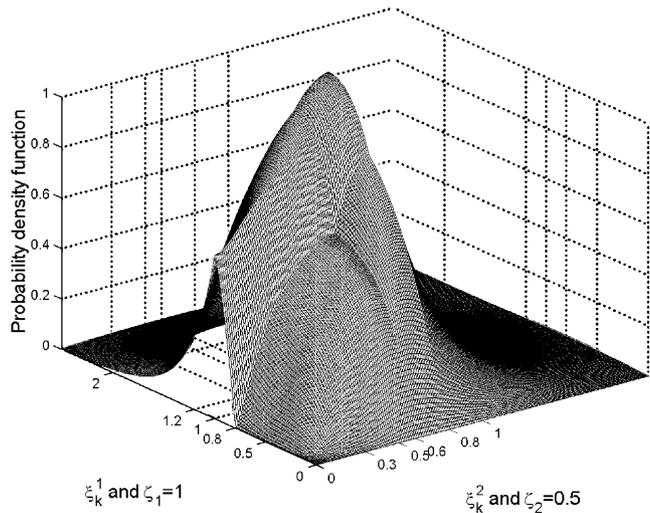


Fig. 3. Representation of the probability density function $\pi(x_1, x_2)$ —Parameters: $\alpha_1=4, L_1=2, \Xi^{(1)}=\{0.5, 0.8, 1.2\}, \zeta_1=1, \alpha_2=5, L_2=1, \Xi^{(2)}=\{0.3, 0.6, 0.8\}, \zeta_2=0.5$.

The computation of the stationary law of the maintained system state, hence of the performance criterion C_∞ are numerically achievable. The performance of the proposed maintenance policy and the influence of some characteristic

parameters are going to be evaluated especially considering the long-run operating cost per unit of time as a function of the opportunistic replacement thresholds.

5. Performance evaluation of the maintenance policy

In this section, some illustrations of the performance of the proposed ‘opportunistic replacement’ maintenance policy are presented to emphasize the need for controlling the balance of opportunistic replacements (i.e. the replacement of a component is anticipated to perform simultaneously two replacements and reduce set-up costs) and separated maintenances. To make the numerical computations easier, the thresholds $\xi_k^{(i)}$ are first optimized separately for each component i as if it was a single one with the unit costs c_i (inspection of component i), $(c_s + c_p^{(i)})$ (preventive replacement of component i), $(c_s + c_c^{(i)})$ (corrective replacement of component i) and c_d (unavailability cost of the system per unit of time). The opportunistic replacement thresholds ζ_1 and ζ_2 which take respectively their values in $[0, \xi_n^1]$ and in $[0, \xi_n^2]$ are optimized in a second step. In the sequel C_{ind} denote the cost obtained when both mono-component policies are simply juxtaposed and C_{opt} denote the minimal average cost observed for the ‘opportunistic replacement’ maintenance policies (with optimal values ζ_1^* and ζ_2^*).

The proposed numerical results show the evolution of the long-term average operating cost per unit of time as a function of the opportunistic replacement thresholds ζ_1 and ζ_2 successively for different set-up costs configurations and for different deterioration parameters.

5.1. Influence of the set-up cost

In this section, in order to highlight the influence of the set-up cost c_s the unit operation costs are assumed to be

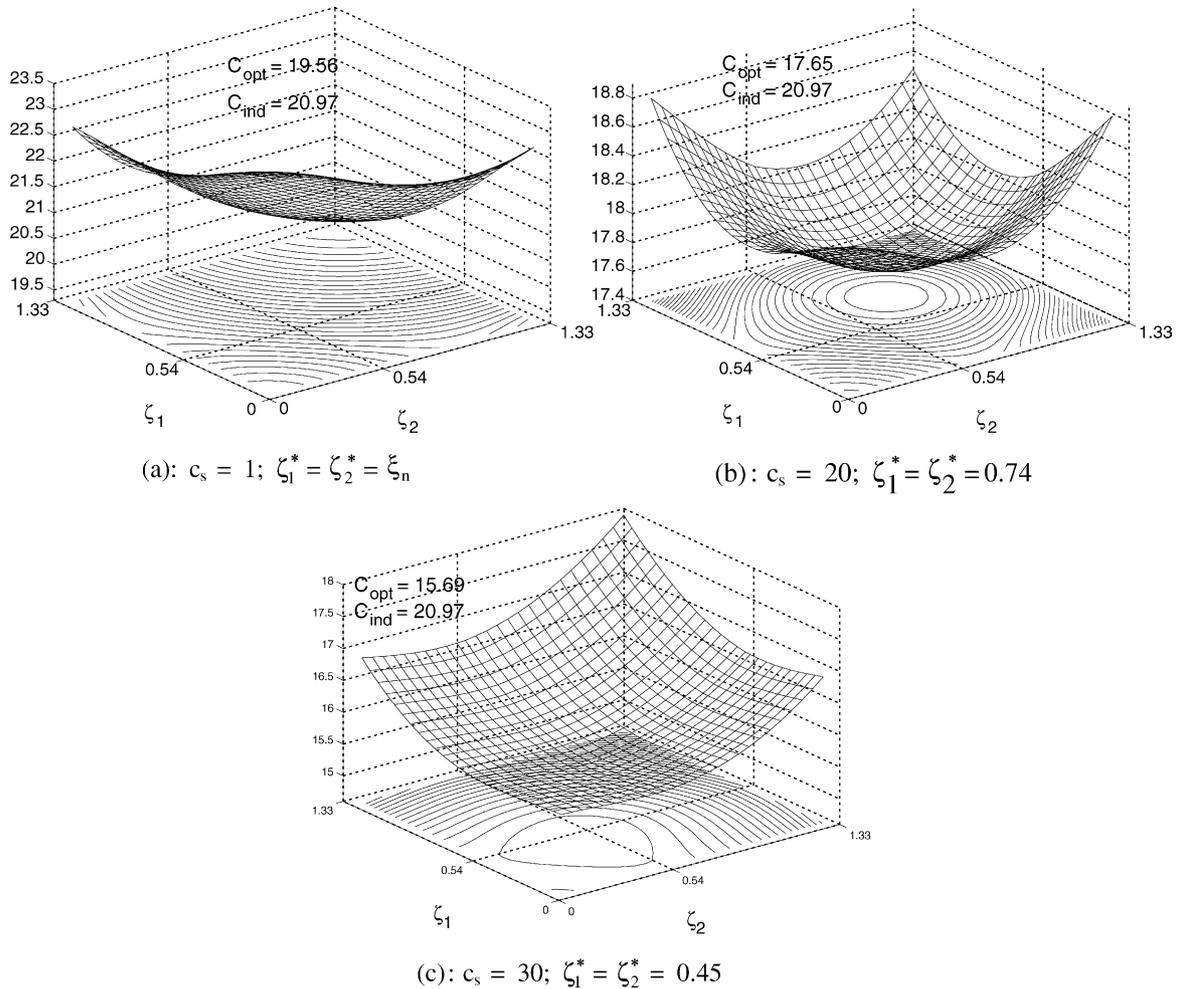


Fig. 4. Maintenance cost surfaces obtained for two identical components with different set-up costs-Parameters: $\alpha_1 = \alpha_2 = 3, L_1 = L_2 = 2; c_i = 1, c_p = 40, c_c = 100, c_d = 10$.

the same for the two components of the system which are supposed to be identical, i.e. to have the same deterioration parameters. As a consequence the optimal number of thresholds and the optimized threshold values are the same for the two components, i.e. $n_1^* = n_2^*, \zeta_1^{(1)*} = \zeta_2^{(1)*}$ and $\zeta_2^{(1)*} = \zeta_2^{(2)*}$. For such a system and for three different increasing set-up cost values ($c_s = 1, c_s = 20$ and $c_s = 30$), Fig. 4 illustrates the evolution of the cost surface as a function of the opportunistic replacement thresholds ζ_1 and ζ_2 with $\alpha_1 = \alpha_2 = 3, c_i = 1, c_p = 40, c_c = 100$ and $c_d = 10$.

All the surfaces are convex shaped which leads to a straightforward optimization procedure. Furthermore, due to the identical components, the surfaces are symmetric with respect to the line $\zeta_1 = \zeta_2$. The optimal number of thresholds is $n_1^* = n_2^* = 2$ and the optimized values for the two thresholds are respectively $\zeta_1^{(1)*} = \zeta_2^{(1)*} = 0.54, \zeta_1^{(1)*} = \zeta_2^{(2)*} = 0.81$ and $\zeta_2^{(1)*} = \zeta_2^{(2)*} = 1.33$ for the three set-up cost values.

When the set-up costs are null or close to zero (Fig. 4a), the optimal values of thresholds ζ_i tend toward the values of the preventive replacement thresholds $\zeta_n^{(i)}$. In that case,

the joint-policy and two independent mono-component policies almost give equal performance ($C_{opt} = 19.56$ et $C_{ind} = 2 \times 10.48 = 20.97$). The difference between C_{opt} and C_{ind} only depends on the inspection unit cost. For a joint-policy, c_i represents actually the unit cost of a global system inspection (both components) whereas for the juxtaposed policies, c_i is the unit cost of an inspection for a single component and the inspection cost of the complete system is $2c_i$.

When the set-up cost increases (Fig. 4b and c), the optimal values of $(\zeta_i)_{i=1,2}$ decrease and lead to strengthen the effects of groupings between the maintenance policies for both components. For a very high set-up cost the economies of scale are of some importance and a global system replacement is preferred as soon as a component needs to be replaced. Furthermore, the minimal average maintenance cost ($C_{opt} = 15.69$) becomes meaningfully lower than $C_{ind} = 20.97$. Several numeric attempts with various configurations of unit costs and two independent identical components lead to the same conclusions.

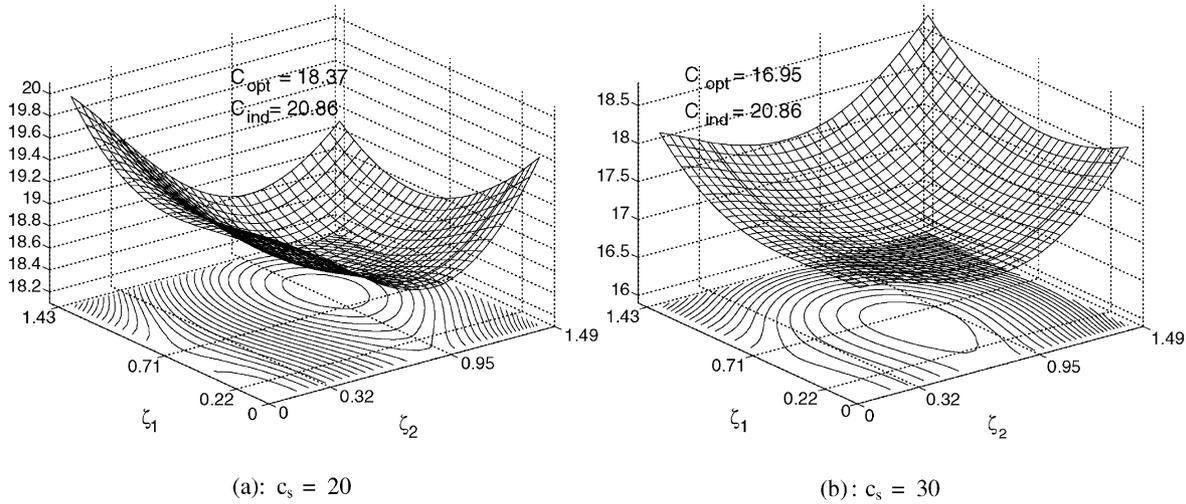


Fig. 5. Maintenance cost surfaces obtained for two slightly different components when opportunistic replacement thresholds ζ_i vary from 0 to $\xi_n^{(i)}$ —Deterioration: $\alpha_1=3, \alpha_2=4, L_1=L_2=2$; —Maintenance costs: $c_i=1, c_p=50, c_c=100, c_d=10$.

5.2. Influence of the deterioration parameters

For two independent but not identical components, Figs. 5 and 6 illustrate the evolution of the maintenance cost surface as a function of the opportunistic replacement thresholds ζ_1 and ζ_2 . Different set-up cost values have been considered respectively for two slightly (Fig. 5) and two strongly (Fig. 6) different components. The other maintenance costs remain constant and are the same for the two components, i.e. $c_i=1, c_p=50, c_c=100$ and $c_d=10$.

Fig. 5a and b (slightly different components) are obtained for $\alpha_1=3$ and $\alpha_2=4$. The optimal number of thresholds is $\min(n_1, n_2)=3$ and the obtained thresholds values are, respectively $\xi_1^{(1)*} = 0.22, \xi_2^{(1)*} = 0.71, \xi_3^{(1)*} = 1.43$ and $\xi_1^{(2)*} = 0.32, \xi_2^{(2)*} = 0.95, \xi_3^{(2)*} = 1.49$. Fig. 6a and b consider stronger differences in the deterioration characteristics with $\alpha_1=2$ and $\alpha_2=6$. The optimal number of thresholds is $\min(n_1, n_2)=2$ and the obtained thresholds values are

respectively $\xi_1^{(1)*} = 0.46, \xi_2^{(1)*} = 1.38$ and $\xi_1^{(2)*} = 1.22, \xi_2^{(2)*} = 1.58$. For both cases, the cost surfaces are convex, which lead to prove the practical ability to find the optimal opportunistic replacement thresholds. Maintenance cost surfaces proposed in Fig. 6 show that the opportunistic replacement threshold value ζ_1 for the faster deteriorating component has a very weak influence on the average operating cost compared to the opportunistic replacement threshold value ζ_2 for the component 2. In this case, the decision of replacements grouping is only determined according to the evolution of the fast deteriorating component (component 1): a high maintenance frequency is required by the component 1, contrary to the component 2. It becomes more interesting to add some replacements for the component 1 simultaneously with the replacement of the component 2 only for a high set-up cost $c_s=47 \approx c_p$, see Fig. 6b. The greater the difference between the deterioration parameters $(\alpha_i)_{i=1,2}$ increases the greater the influence of

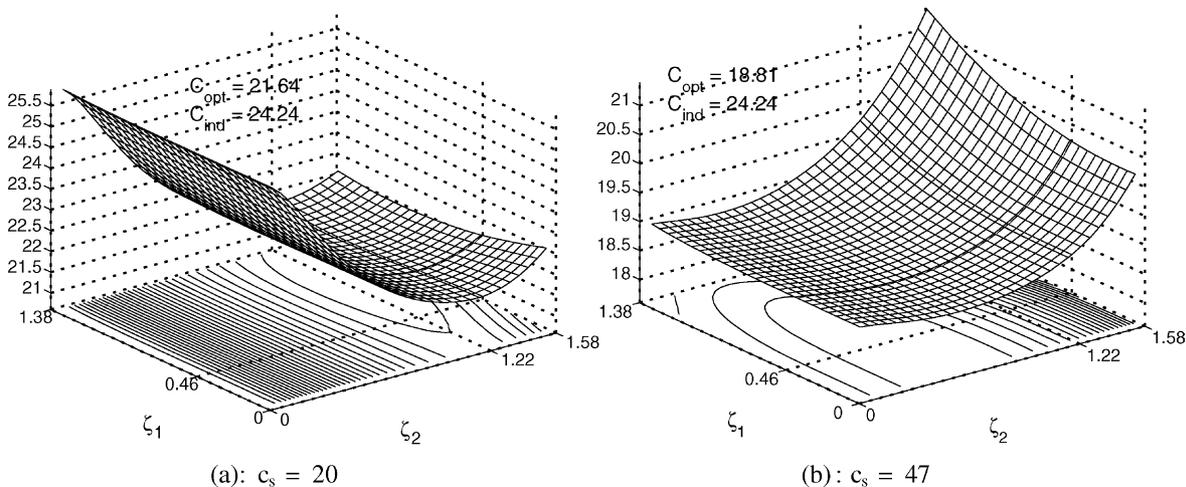


Fig. 6. Maintenance cost surfaces obtained for two very different components when opportunistic replacement thresholds ζ_i vary from 0 to $\xi_n^{(i)}$ —Deterioration: $\alpha_1=2, \alpha_2=6, L_1=L_2=2$; —Maintenance costs: $c_i=1, c_p=50, c_c=100, c_d=10$.

the opportunistic replacement threshold corresponding to the faster (resp. slower) deteriorating component increases (resp. decreases).

When the deterioration characteristics of the two components are close one to each other, their behavior is naturally similar (without opportunistic replacement). For the ‘opportunistic replacement’ maintenance policy, the threshold values ζ_i , $i=1,2$ are very close, see Fig. 5. The distribution of the maintenance influence of each component for the replacements groupings is well balanced.

6. Conclusion

We have proposed in this paper a simple decision rule structure for condition-based maintenance of a two-unit system with independent stochastically deteriorating units. The construction of the proposed maintenance decision rule is based at a first step on a mono-component condition-based policy which allows to propose simultaneously and dynamically the maintenance action to perform for each unit independently and the next system intervention time. New thresholds have been introduced to control groupings of the maintenance tasks in order to reduce the global long-run average cost thanks to economies of scale.

This maintenance model should be extended to multi-unit systems composed of more than two units. Such an extension lead to a complex stationary law and a high number of maintenance decision parameters. The numerical solution rapidly becomes intractable. It could be interesting to have a coupled approach based on both the analytical/numerical approach and the use of Monte Carlo simulation to tackle with more complex systems.

Appendix A

A.1. Evaluation of the expected cost of unavailability

The cumulative cost of unavailability up to time t , $C_{FU}(t)$ is supposed to be proportional to the time elapsed by the system in failed state $D_U(t)$. The given cost rate is c_d . The expected cumulative cost $\mathbb{E}_\pi(C_{FU}(S))$ incurred by the time the system passes in failed state over $[0,S]$ at steady state is given by:

$$\mathbb{E}_\pi(C_{FU}(S)) = c_d \mathbb{E}_\pi(D_U(S))$$

The system is considered as failed as soon as a component deterioration level exceeds its critical level L_i . Due to the discrete time grid for inspections the exact failure time of component i is unknown and the time $D_U^{(i)}(S)$ the component passes in failed state over $(0,S)$ can only be approximated, e.g. by its upper bound $\bar{D}_U^{(i)}(S)$:

if a failure occurs at time $t_f^{(i)}$ with $t_f^{(i)} \in (t_k, t_{k+1})$ and $t_{k+1} \leq S$,

$$D_U^{(i)}(S) \approx \bar{D}_U^{(i)}(S) = S - t_k$$

Let $H_i(k|y;l)$ define the probability that the failure time of component i is k time units after the last inspection time given that the last inspection revealed a deterioration level equal to y and the next inspection was scheduled l time units later. In the case of an exponential law with parameter α_i for the deterioration increment of component i it comes if $y \leq \zeta_{n-l+1}^i$:

$$H_i(k|y;l) = \begin{cases} 0 & \text{if } l < k \\ e^{-\alpha_i(L-z)} & \text{if } l = k \\ \frac{(\alpha_i(L_i - z))^{l-k}}{(l-k)!} e^{-\alpha_i(L-z)} & \text{if } l > k. \end{cases} \quad (A1)$$

The expected unavailability duration over a semi-regenerative cycle is given by:

$$\begin{aligned} \mathbb{E}_\pi(\bar{D}_U(S)) &= \sum_{k=1}^n k \mathbb{P}(\bar{D}_U(S) = k) = \sum_{k=1}^n \mathbb{P}(\bar{D}_U(S) \geq k) \\ &= \sum_{k=1}^n \left(\sum_{i=1}^2 \mathbb{P}(\bar{D}_U^{(i)}(S) \geq k) - \mathbb{P}(\bar{D}_U^{(1)}(S) \geq k, \bar{D}_U^{(2)}(S) \geq k) \right) \\ &= \sum_{k=1}^n \sum_{r=k}^n \left(\sum_{i=1}^2 \mathbb{P}(\bar{D}_U^{(i)}(S) = r) - \mathbb{P}(\bar{D}_U^{(1)}(S) = r, \bar{D}_U^{(2)}(S) = r) \right) \end{aligned} \quad (A2)$$

Eq. (A2) is based on the unavailability duration of each component considered separately (positive part) and takes into account the time both components are simultaneously in failure state (negative part) which is counted twice in the first part of the equation.

The probabilities can be evaluated by an analysis of all the scenarios leading to the system failure. For a given component i ($i \in \{1,2\}$), let in the sequel $P_{pq}^{(i)}(r)$ define the probability that $\bar{U}^{(i)}(S) = r$ conditionally to the last operation(s) performed on the system and let $P_{pq}(r)$ define the probability that $\bar{D}_U^{(1)}(S) = r$ and $\bar{D}_U^{(2)}(S) = r$ conditionally to the last operation(s) performed on the system, where:

- $p=0$ stands for ‘component i has been replaced’,
- $p=1$ stands for ‘component i has not been replaced’,
- $q=0$ stands for ‘component $j(j \neq i)$ has been replaced’,
- $q=1$ stands for ‘component j ($j \neq i$) has not been replaced’.

We have respectively:

- if the component i has just been replaced, the next inspection time depends on the deterioration level of the component j and:

$$P_{01}^{(i)}(r) = \sum_{l_j=k}^N \int_{\xi_N^{l_j}}^{\infty} \int_{\xi_{N-l_j}^{l_j}}^{\xi_{N-l_j+1}^{l_j}} \pi(y_1, y_2) dy_j dy_i H_i(r|0; l_j) \quad (A3)$$

$$P_{01}(r) = \sum_{l_j=k}^N \int_{\xi_N^{l_j}}^{\infty} \int_{\xi_{N-l_j}^{l_j}}^{\xi_{N-l_j+1}^{l_j}} \pi(y_1, y_2) H_j(r|y_j; l_j) dy_j dy_i H_i(r|0; l_j) \quad (A4)$$

- if the component i has just been replaced, the next inspection time dependson the deterioration level of the component j and:

$$P_{01}^{(i)}(r) = \sum_{l_i=k}^N \int_{\xi_N^{l_i}}^{\xi_{N-l_i+1}^{l_i}} \int_{\xi_{N-l_i}^{l_i}}^{\infty} \pi(y_1, y_2) dy_j H_i(r|y_i; l_i) dy_i \quad (A5)$$

$$P_{01}(r) = \sum_{l_i=k}^N \int_{\xi_N^{l_i}}^{\xi_{N-l_i+1}^{l_i}} \int_{\xi_{N-l_i}^{l_i}}^{\infty} \pi(y_1, y_2) dy_j H_i(r|y_i; l_i) dy_i H_j(r|0; l_i) \quad (A6)$$

- if the two component have just been simultaneously replaced:

$$P_{00}^{(i)}(r) = \int_{\xi_N}^{\infty} \int_{\xi_N}^{\infty} \pi(y_1, y_2) dy_j dy_i H_i(r|0; N) \quad (A7)$$

$$P_{00}(r) = \int_{\xi_N}^{\infty} \int_{\xi_N}^{\infty} \pi(y_1, y_2) dy_j dy_i H_i(r|0; N) H_j(r|0; N) \quad (A8)$$

- if no replacement has been performed:

$$P_{11}^{(i)}(r) = \sum_{l_i, l_j=1}^N \int_{\xi_{N-l_i}^{l_i}}^{\xi_{N-l_i+1}^{l_i}} \left(\int_{\xi_{N-l_j}^{l_j}}^{\xi_{N-l_j+1}^{l_j}} \pi(y_1, y_2) dy_j \right) H_i(r|y_i; \min(l_i, l_j)) dy_i \quad (A9)$$

$$P_{11}(r) = \sum_{l_i, l_j=1}^N \int_{\xi_{N-l_i}^{l_i}}^{\xi_{N-l_i+1}^{l_i}} \int_{\xi_{N-l_j}^{l_j}}^{\xi_{N-l_j+1}^{l_j}} \pi(y_1, y_2) H_i(r|y_i; \min(l_i, l_j)) \times H_j(r|y_j; \min(l_i, l_j)) dy_j dy_i \quad (A10)$$

As a conclusion it comes:

$$\mathbb{P}(\bar{D}_U^{(i)}(S) = r) = P_{00}^{(i)}(r) + P_{10}^{(i)}(r) + P_{01}^{(i)}(r) + P_{11}^{(i)}(r) \quad (A11)$$

$$\mathbb{P}(\bar{D}_U^{(1)}(S) = r, \bar{D}_U^{(2)}(S) = r) = P_{00}(r) + P_{10}(r) + P_{01}(r) + P_{11}(r) \quad (A12)$$

and from Eq. (2) the final expression of $\mathbb{E}_{\pi}(\bar{D}_U(S))$ is obtained.

References

- [1] Barlow R, Proschan F. Mathematical theory of reliability, vol. 17. Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics; 1996.
- [2] Scarf P. On the application of mathematical models in maintenance. Eur J Oper Res 1997;99(3):493–506.
- [3] Wijnmalen D, Hontelez J. Review of a Markov Decision Algorithm for Optimal Inspections and Revisions in a Maintenance System with Partial Information. Eur J Oper Res 1992;62:96–104.
- [4] Dekker R, Smith M. New developments in the optimisation of maintenance. In: Lydersen S, Hansen G, Sandtorv H, editors. Proceedings of the European Safety and Reliability Conference—ESREL’98—16–19 june, Trondheim, Norway, 1998. p. 731–7.
- [5] Dekker R, van der Duyn Schouten F, Wildeman R. A review of multi-component maintenance models with economic dependence. Math Methods Oper Res, 1997;45:411–35.
- [6] R. Wildeman, The Art of Grouping Maintenance, Ph.D. thesis, Erasmus University Rotterdam (1996).
- [7] Cho D, Parlar M. A survey of maintenance models for multi-unit systems. Eur J Oper Res 1991;51:1–23.
- [8] Dekker R, Frenk H, Wildeman E. In: Özekici S, editor. How to determine frequencies for multi-component systems ? A general approach Reliability and maintenance of complex systems, NATO ASI Series—Series F: computer and systems sciences, vol. 154, 1996. p. 239–80.
- [9] van der Duyn Schouten F. In: Özekici S, editor. Maintenance policies for multicomponent systems : an overview Reliability and maintenance of complex systems, NATO ASI Series—Series F: computer and systems sciences, vol. 154. Berlin: Springer; 1996. p. 118–36.
- [10] Marseguerra M, Zio E, Podofillini L. Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation. Reliab Eng Syst Saf 2002;77:151–66.
- [11] Aven T, Dekker R. A useful framework for optimal replacement models. Reliab Eng Syst Saf 1997;58:61–7.
- [12] Sandve K, Aven T. Cost optimal replacement of monotone, repairable systems. Eur J Oper Res 1999;116(2):235–48.
- [13] Wildeman R, Dekker R, Smit A. A dynamic policy for grouping maintenance activities. Eur J Oper Res 1997;99:530–51.
- [14] Wildeman R, Dekker R. Dynamic influences in multi-component maintenance. Qual Reliab Eng Int 1997;13:199–207.
- [15] Kopnov V. Optimal degradation processes control by two-level policies. Reliab Eng Syst Saf 1999;66:1–11.
- [16] Grall A, Bérenguer C, Dieulle L. A condition-based maintenance policy for a stochastically deteriorating system. Reliab Eng Syst Saf 2002;76(02):167–80.
- [17] Castanier B, Grall A, Bérenguer C. Une structure multiseuil pour l’optimisation de la maintenance des systèmes à détérioration continue. APII-JESA—J Eur Syst Automat 2001;35(10):1235–62 in French.
- [18] van der Duyn Schouten F, Vanneste S. Analysis and computation of (n,N)-strategies for maintenance of a two-component system. Eur J Oper Res 1990;48:260–74.
- [19] Murthy D. Analysis and design of unreliable multi-component systems with modular structure. Large Scale Systems 1983;245–54.
- [20] Murthy D, Nguyen D. Study of two-component system with failure interaction. Naval Res Log 1985;32:239–47.
- [21] Scarf P, Deara M. Block replacement policies for a two-component system with failure dependence. Naval Res Log 2003;50:70–87.
- [22] Scarf P, Deara M. On the development and application of maintenance policies for a two-component system with failure dependence. IMA J Math Appl Business Ind 1998;9:91–107.
- [23] Feller W. An introduction to probability theory and its applications, 2nd ed Wiley series in probability and mathematical statistics. New York: Wiley; 1971.
- [24] Park K. Optimal continuous-wear limit replacement under periodic inspections. IEEE Trans Reliab 1988;37(1):97–102.

- [25] Park K. Optimal Wear-Limit Replacement with Wear-Dependent Failures. *IEEE Transact Reliab* 1988;37(1):293–4.
- [26] Cooke R, Mendel M, Vrijling H. In: Cooke R, Mendel M, Vrijling H, editors. *Engineering probabilistic design and maintenance for flood protection*. Dordrecht: Kluwer; 1997.
- [27] van Noortwijk J, Klatter H. Optimal inspection decisions for the block mats of the eastern-scheldt barrier. *Reliab Eng Syst Saf* 1999; 65:203–11.
- [28] Castanier B, Bérenguer C, Grall A. In: Zio E, Demichela M, Piccinini N, editors. *A model for maintenance cost optimization under availability constraints Safety and reliability, ESREL 2001—Toward a safer world*, vol. 2, 2001. p. 1313–20.
- [29] Archibald T, Dekker R. Modified block-replacement for multiple-components systems. *IEEE Trans Reliab* 1996;45(1):75–83.
- [30] Dekker R. Integrating optimisation, priority setting, planning and combining of maintenance activities. *Eur J Oper Res* 1995;82:225–40.
- [31] Coccozza-Thivent C. Convergence de fonctionnelles de processus semi-régénératifs, prépublications de l'Université de Marne la Vallée 02/2000; March 2000. [in French]
- [32] Polyanin AD, Manzhirov AV. *Handbook of integral equations*. Boca Raton: CRC Press; 1998.
- [33] Castanier B. *Modélisation Stochastique et Optimisation de la Maintenance Conditionnelle des Systèmes à Dégradation Graduelle*, Thèse de doctorat, Université de technologie de Troyes; 2001. [in French].

Maintenance optimization for subsea oil pipelines

Bruno Castanier^{a,*}, Marvin Rausand^b

^a IRCCyN/Ecole des Mines de Nantes, La Chantrerie, 4 rue A. Kastler, Nantes, France

^b Department of Production and Quality Engineering, Norwegian University of Science and Technology, Trondheim, Norway

Abstract

In this paper, the classical PF interval model, proposed for reliability centered maintenance, is used to optimize the preventive replacement policy of a subsea oil pipeline. The pipeline is subject to failure modes that may lead to oil release and pollution of the sea. The proposed replacement policy is based on regular condition measurements of the quality of an internal pipeline coating (lining). Holes in the coating will expose the pipe material (steel) to a corrosive environment. It is, therefore, important to identify holes in the coating and take action before the corrosion penetrates the pipe wall. The model proposed in the paper can be used to determine the inspection interval that minimizes the average maintenance cost per time unit. The model formulation is outlined and illustrative examples are solved. The mathematical evaluation of the replacement policy requires numerous assumptions. These assumptions have to be tailored to the specific application. The assumptions and the data provided are included for illustrative purposes to exemplify potential results that may be obtained by using the replacement policy.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Preventive maintenance; Optimization; Replacement policy; Oil pipeline; PF model

1. Introduction

A wide range of preventive maintenance (PM) optimization models have been proposed in the literature [1–4]. Many of the models have been developed by highly skilled mathematicians, but often without too much regard to the practical applicability of the models. Some of the models are based on assumptions that can never be fulfilled, and on data that are very difficult to acquire. Another problem is that the models are presented with a mathematical terminology that is very difficult to understand for people working with practical maintenance planning. One of the nestors within maintenance optimization has claimed that there are no other scientific discipline where the gap between theory and practice is bigger than for maintenance optimization.

An objective of this paper is to try to build a small bridge between theory and practice by starting with a practical example and showing how a maintenance optimization models can be established based on realistic assumptions. Most realistic situations will, however, be rather complex, and it is seldom possible to solve complex problems with simple tools.

A certain level of mathematical complexity is, therefore, unavoidable.

As a practical example we will consider a steel pipeline that is located on the sea bottom. The pipeline is used to transport a mixture of oil, gas, and water from a subsea production unit to a process plant that is located either on a floating production unit, or on land. The length of the pipeline may be several kilometres.

The fluid in the pipeline is corrosive and may contain sand particles that will cause abrasive erosion. The amount of sand in the fluid is not constant and may vary with time, with some shorter periods with rather high sand content. The combined effect of corrosion and erosion is hostile, and the corrosion will be significantly more aggressive when erosion is present. The pipeline considered in this paper is internally protected by a coating or lining. As long as this coating is intact, the steel pipe is protected from internal corrosion. The coating will, however, be worn by abrasive erosion, mainly caused by the sand in the fluid. When the coating is penetrated the steel pipe is exposed to the corrosive fluid.

When a hole in the coating is present, an erosive corrosion of the steel will take place, finally resulting in a leakage of the fluid to the sea. When a leakage to the sea is detected, the production through the pipeline has to be closed down, and a corrective maintenance (CM) action has to be initiated. This repair operation will require that a semi-submersible rig or a similar large vessel is brought to the location of the leakage. The time from the leakage is detected until the pipeline is

* Corresponding author. Tel.: +33 25 1858312; fax: +33 25 1858349.
E-mail address: bruno.castanier@emn.fr (B. Castanier).

repaired may be several weeks, and the repair operation will imply a very high cost.

To prevent leakage to the sea, the pipeline is inspected at regular intervals by an inspection tool that is sent through the pipeline. If a coating failure is detected a preventive maintenance (PM) action is initiated. The PM action will require the same intervention vessel as a CM action, but the production through the pipeline does not have to be closed down except during the active repair operation. A PM action will, therefore, imply a significantly lower total cost than a CM action, and in addition, it will not give any pollution. The inspection, however, implies a rather high cost, and can, therefore, not be carried out very frequently. The inspection is never fully reliable, and a coating failure may, therefore, be overseen.

The objective of this paper is to establish a stochastic model that can be used to determine a cost-optimal inspection frequency of the pipeline in an erosive–corrosive environment. Hence, the proposed solution should determine the best balance between prevention of oil pollution and inspection cost minimization. It should further be possible to apply the stochastic model to compare different types of inspection tools and inspection strategies. The problem may be considered as a so-called PF interval optimization problem, as discussed by Rausand and Høyland [5]. The PF interval concept is often presented in reliability centered maintenance [6,7], but too few optimization models are available [8,9]. The model is also similar to the delay time model discussed by Christer [10].

The structure of the remaining part of the paper is as follows. In Section 2 we introduce the notation and terminology, and describe the assumptions of the model in a probabilistic framework. The PF interval model is described and discussed in Section 3. Section 4 provides a step by step description of the construction of the cost function by determining each quantity that characterises the decision criterion. In Section 5, numerical experiments are presented to illustrate the behaviour of the maintenance decision parameter for given data. In the last section, we discuss future challenges based on the restrictions of the model.

2. Model assumptions

Let T_1 denotes the time from start-up until the first detectable penetration P of the coating. The failure may be caused by a combination of abrasive erosion and random shocks caused by excessive sand content. A realistic model for the failure rate function of T_1 might be $z_1(t) = \lambda_1 + \beta t$. The survivor function of the coating is, therefore:

$$R_1(t) = \exp \left[- \int_0^t z_1(u) du \right] = \exp \left[- \left(\lambda_1 t + \frac{\beta t^2}{2} \right) \right] \quad (1)$$

The model implies a linear degradation of the coating. The rate of degradation is determined by the parameter β . If we believe that the degradation is more aggressive than linear, this may be incorporated into the model by introducing an extra

parameter, say $\alpha \geq 1$ such that $z_1(t) = \lambda_1 + (\beta t)^\alpha$. If random shocks can be prevented, we may choose $\lambda_1 \approx 0$.

The pipeline is intended to be in operation for a specified number of years, that is determined by the production life of its upstream systems (i.e. the producing wells). The intended service life of the pipeline will vary from installation to installation depending on the reservoir characteristics. To simplify the mathematical calculations, we assume that the intended service life for our example is infinity. This assumption may seem unrealistic, but we will show that the obtained result is a good approximation when the service life is rather long, i.e. 10–15 years.

When there is a hole in the coating the steel pipe will be exposed to corrosion, mainly pitting corrosion. Pitting corrosion has been extensively studied and several models have been suggested for the time to penetrate steel of different qualities. In this paper, we will use a rather simple model, with a constant corrosion rate. Several researchers have concluded (e.g. see [5], p. 58) that the time T_{PF} to penetrate the pipe wall will be distributed according to a truncated Gumbel distribution of the smallest extreme, that is, a distribution with an exponentially increasing failure rate function $z_2(t) = \gamma v e^{vt}$. The parameters, γ and v , of this model depend on the thickness of the pipe wall, the steel quality of the pipeline, and the corrosiveness and erosiveness of the fluid in the pipeline. Initial estimates of these parameters may be deducted from tables in the scientific literature on corrosion.

The total time T from start-up till the pipe is penetrated and we have a leakage to sea, will therefore, be $T = T_1 + T_{PF}$. The probability distribution of T may be found by standard convolution formulas. Note that T_1 is generally much longer than T_{PF} , and that when a coating failure has occurred, we do not have very long time available to intervene to avoid a leakage to sea.

Let τ denote the time between two consecutive inspections of the pipeline. We assume that all the intervals between inspections have the same length. The length of τ is our main decision variable, that is, the quantity we want to optimize. Let $\theta_1(\tau)$ denote the probability that a detectable coating failure P is not detected during the first inspection after P has occurred. The coating failure will grow with time and it is assumed that the probability $\theta_2(\tau)$ to not detect the failure in the next inspection is lower than $\theta_1(\tau)$, and so on. Modeling the imperfection of an inspection according to the number of performed inspections seems to be a more easy-implemented approach than the construction of an error function directly proportional to the size of the hole in the coating. Note at this stage, that the inspection strategy can be seen as a pipe age-based strategy. So, the first inspection after a potential replacement of failed pipes will be scheduled τ time units after the restarting time.

When a coating failure is detected during an inspection, an intervention vessel is called upon to replace the worn pipe segment. The vessel must be ordered and brought out to the location, and the pipe segment must be replaced. The mobilization of the vessel and its transport out to the location takes a time T_M that is generally a random variable. Most

uncertainty is connected to the availability of a suitable intervention vessel. In some situations, we may have to wait several weeks/months to get a vessel on location. An average time from the failure is detected until the pipe segment has been replaced may be around 4 weeks.

The production must be closed down during the active repair time, that is, the time required to cut the pipe and install a new pipe segment. An average active repair time may be around 10 h. In most cases this is a standard operation and the uncertainty of the repair time estimate will be rather insignificant. We may, therefore, assume that the active repair time is a deterministic time t_R .

If a leakage to sea from the pipeline is detected, the production in the pipeline has to be closed down immediately, and an intervention vessel must be called upon to repair the failure. The leakage will be detected by observing oil on the sea. The time required to detect the leakage will be dependent on the specific situation, where the pipeline is running and also on the weather conditions when the leakage occurs. If no inspection allows to detect a hole in the pipe, we assume that the required time T_D to detect a leakage is a random variable. In general, the pollution and the clean-up cost will be a function of the length of this detection time T_D . Finally, the production will be closed down from the oil pollution detection until the end of the repair action.

We consider a perfect repair in the way that no coating failure is present after this action. Furthermore, because of the coating reliability, we can assume that only one segment can be damaged between two consecutive maintenance actions. Finally, we will consider the pipeline to be as good as new after a repair.

The following costs are defined:

- c_i , the cost of an inspection of a pipe segment, which is equal to the ratio of the total inspection cost and the number of pipe segments.
- c_r , the unit cost related to a replacement of a failed pipe segment.
- c_u , the production losses due to downtime.
- c_m , the cost per time unit for the intervention vessel mobilization.
- c_n , cost per time unit of pollution, and damages due to pollution and clean-up.

All these costs are subject to random variation, but we will in this paper assume that they are fixed values.

3. PF interval model

The situation described above is illustrated in Fig. 1. We start observing the pipeline when it has been installed at time $t=0$. A detectable coating failure P will occur at the random time T_1 . The probability that P will be detected during the first inspection after P has taken place is $1-\theta_1(\tau)$. The probability that the potential failure will be detected in the second inspection is $\theta_1(\tau)(1-\theta_2(\tau))$, and so on. We assume that the conditional probability that P is not detected in the k th

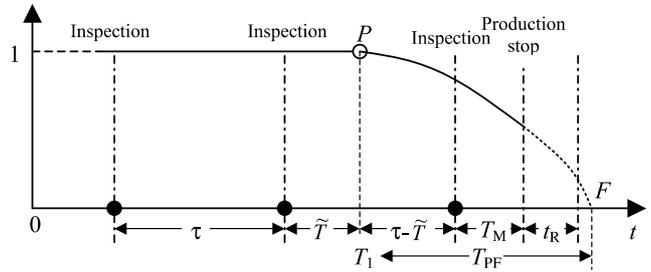


Fig. 1. Average behavior and concepts used in the PF interval model.

inspection carried out after P given that P has not been detected before, $\theta_k(\tau)$, is given by

$$\theta_k(\tau) = 1 - (p_\infty - (p_\infty - p_0)e^{-\lambda_0(k-1)\tau}) \tag{2}$$

where the parameters p_0 and p_∞ may be considered as intrinsic inspection quality parameters (function of the inspection device) and λ_0 a parameter that is a function of the hole propagation speed (i.e. a function of the quality of the coating). The potential detection function and the associated bounding parameters p_0 and p_∞ are illustrated in Fig. 2.

If P is detected at time t after its occurrence, a new segment is ordered. A replacement will take place as soon as the new segment is in situ. From Fig. 1 we notice that the replacement is considered as a preventive replacement if no failure occurs before the production stop at time $\tau - t + t_R + T_M$.

Let \tilde{T} be the time between the last inspection before P and T_1 , and $F_{\tilde{T}}(t)$ the associated probability distribution function for $0 < t < \tau$, so that we for all $t \in (0, \tau)$ have:

$$F_{\tilde{T}}(t) = \sum_{n=0}^{\infty} \Pr(n\tau < T_1 < n\tau + t) = \frac{1 - e^{-\lambda_1 t}}{1 - e^{-\lambda_1 \tau}} \tag{3}$$

The PF interval, T_{PF} , is assumed to be a random variable with a Gumbel probability distribution function $F_{PF}(t)$ with two parameters. Let $T_F = \tilde{T} + T_{PF}$, be the time from the last inspection before P until a (possible) critical failure. Let $F_F(t)$ be the probability distribution function of T_F . For mathematical convenience and to simplify the expression of $F_F(t)$, we assume that the occurrence of a hole in the coating is a pure random event, and so $z_1(t) = \lambda_1$, i.e. that $\beta = 0$. Hence, because of the markovian properties of the coating crack process, we can show (e.g. see [5], p. 399) that the probability distribution function $F_F(t)$ can be found by the convolution of the

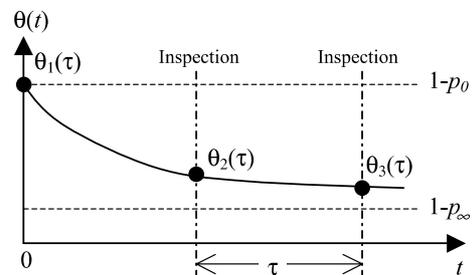


Fig. 2. Decreasing conditional probability function of no detection of P at time t .

probability distribution functions $F_{\bar{T}}(t)$ and $F_{PF}(t)$:

$$F_F(t) = \Pr(T_F \leq t) = \int_0^{\tau} F_{PF}(t-u) dF_{\bar{T}}(u) \\ = \frac{\lambda_1}{1 - e^{-\lambda_1 \tau}} \int_0^{\tau} \left(1 - e^{-\gamma(e^{v(t-u)} - 1)}\right) e^{-\lambda_1 u} du \quad (4)$$

The restoration time, with no leakage to sea, is defined by $T_M + t_R$ where t_R is given (deterministic). The probability distribution function of the restoration time is given by the probability distribution function of T_M , $F_M(t)$. With $\bar{F}_M(t) = 1 - F_M(t)$, its expectation is given by:

$$E(T_M) = \int_0^{\infty} \bar{F}_M(t) dt + t_R \quad (5)$$

In case of pollution, the detection time T_D , with probability distribution function $F_D(t)$, has to be added to the restoration duration.

Finally, in this section, the different quantities, which characterize a stochastic failure of a pipeline and the associated time for repairing a failed segment have been formulated.

4. Construction of the model of the mean cost per unit of time

This section describes how to construct, step by step, the objective function and the mean maintenance cost of one segment per unit of time, $C_{\infty}(\tau)$. Because of the perfect repair assumption and the age-based inspection strategy, the classical results of the renewal theory can be utilized [5]. Hence, it is possible to restrict the maintained system evolution to a single renewal cycle defined by the interval between two successive repairs. The function of this maintenance cost can be written as follows

$$C_{\infty}(\tau) = \frac{E(C(\tau))}{MTBR(\tau)} \quad (6)$$

where $E(C(\tau))$ is the expected maintenance cost for one segment accumulated between two successive repairs, and $MTBR(\tau)$ is the mean time between repairs, when the inspection interval is τ .

The accumulated cost $C(\tau)$ is the sum of the contributions of each maintenance operation from the pipeline start-up until the production is restarted after the repair. The next subsections show how to define all these contributions and the mean time between repairs.

4.1. Inspection cost

The proposed pipeline inspection scheme is an age-based periodic inspection policy. Let $N(\tau)$ denote the number of inspections carried out before the potential failure P . Let $Z(\tau)$ be the number of inspections carried out between P and the

observation of the potential failure, with detection of pollution at time $T_F + T_D$. Because the unit cost of an inspection c_i is assumed constant, the expected cumulated inspection cost $E(C_i(\tau))$ is:

$$E(C_i(\tau)) = c_i(E(N(\tau)) + E(Z(\tau))) \quad (7)$$

To determine the expected numbers $E(N(\tau))$ and $E(Z(\tau))$ we use the memory-less property of the exponential distribution. This leads to

$$E(N(\tau)) = \frac{\exp(-\lambda_1 \tau)}{1 - \exp(-\lambda_1 \tau)} \quad (8)$$

and

$$E(Z(\tau)) = \sum_{k=1}^{\infty} \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \bar{F}_{FD}(k\tau) \quad (9)$$

where $\bar{F}_{FD}(t) = 1 - F_{FD}(t)$ is the reliability function of the pollution detection time.

4.2. Preventive replacement cost

The preventive cost $C_P(\tau)$ is related to a preventive replacement after the detection of a hole in the coating; this operation has to be completely finished before the time of failure T_F . The cost $C_P(\tau)$ is the sum of the cost to bring the intervention vessel to the location $C(T_M) = c_m T_M$, the cost of the damaged segment replacement action c_r , and the production loss that is proportional to the operation duration, $c_u t_R$. Because the costs c_m , c_r , and c_u are assumed to be constant, the expected preventive cost $E(C_P(\tau))$ is

$$E(C_P(\tau)) = (c_r + c_u t_R + c_m E(T_M)) P_P(\tau) \quad (10)$$

where $P_P(\tau)$ is the probability of a preventive repair, when the inspection interval is τ .

To determine the probability $P_P(\tau)$, we describe the maintenance scenario that leads to a preventive replacement:

- A hole in the coating is detected during the k th inspection after the occurrence of the (detectable) hole P , i.e. after time T_1 .
- No critical failure occurs before the production stop when the new segment is available on place for replacement, i.e. $T_F > k\tau + T_M$.

Finally, the probability $P_P(\tau)$ can be written as:

$$P_P(\tau) = (1 - \theta_1(\tau)) \Pr(T_F > \tau + T_M) + \theta_1(\tau) (1 - \theta_2(\tau)) \\ \Pr(T_F > 2\tau + T_M) + \dots \\ = \sum_{k=1}^{\infty} (1 - \theta_k(\tau)) \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \int_0^{\infty} \bar{F}_F(k\tau + u) dF_M(u) \quad (11)$$

4.3. Corrective replacement cost

The corrective replacement cost is a function of the failure occurrence time T_F and the time to the detection, T_D , of this

failure. The failure F can occur before the detection of the potential failure P , or during the transfer of the intervention vessel when a failure P has already been detected and before the production stop to replace preventively the damaged segment. Let $F_{FD}(t)$ denote the probability distribution function of the random variable $T_F + T_D$ defined by the convolution of the respective distribution functions. We can distinguish three scenarios that lead to a failure and, to pollution.

Let $P_1(\tau)$ denote the probability of the segment failure is detected without observing the potential failure P , i.e. $T_F + T_D \in ((k-1)\tau, k\tau]$ if $Z(\tau) \geq k$. We have:

$$P_1(\tau) = \sum_{k=1}^{\infty} \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) (F_{FD}(k\tau) - F_{FD}(k-1)\tau) \quad (12)$$

Let $P_2(\tau)$ denote the probability that the segment failure is detected after the observation of P , but before the arrival of the vessel, and such that the production is not yet stopped. Hence, $Z(\tau)\tau < T_F + T_D < Z(\tau)\tau + T_M$ and we have:

$$P_2(\tau) = \sum_{k=1}^{\infty} (1 - \theta_k(\tau)) \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \left(\int_0^{\infty} F_{FD}(k\tau + u) dF_M(u) - F_{FD}(k\tau) \right) \quad (13)$$

Let $P_3(\tau)$ denote the probability that the segment failure occurs before the production stop that is planned after the observation of P , but it is not detected before the arrival of the vessel. We have $T_F < Z(\tau)\tau + T_M$ and $T_F + T_D > Z(\tau)\tau + T_M$ and so we can write:

$$P_3(\tau) = \sum_{k=1}^{\infty} (1 - \theta_k(\tau)) \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \int_0^{\infty} (F_F(k\tau + u) - F_{FD}(k\tau + u)) dF_M(u) \quad (14)$$

We can easily verify that the total probability is equal to one, i.e. $P_P(\tau) + P_1(\tau) + P_2(\tau) + P_3(\tau) = 1$.

Let $C_1(\tau)$, $C_2(\tau)$, and $C_3(\tau)$ denote the costs associated with these three scenarios. Let Y be the additional downtime due to emergency stop when a critical failure occurs. For scenarios 1 and 2, Y is defined from the detection of the critical failure until the beginning of the replacement action. For scenario 3, the additional downtime Y is 0 because the replacement is started when the potential failure P is detected and no pollution will be detected, thus not causing any premature stop. Let X be the

pollution time from the critical failure T_F until the production stop. A detailed description of the scenarios leads to the results in Table 1.

Let $E_j(Y)$ and $E_j(X)$ denote the conditional expectation of $E(Y)$ and $E(X)$, respectively, given the scenario j . We can now find the expressions for the different conditional expectations $E_j(Y)$:

$$E_1(Y) = P_1(\tau) \int_0^{\infty} \bar{F}_M(t) dt \quad (15)$$

$$E_2(Y) = E(E(Y|Z(\tau), T_M, \text{scenario } 2)) = \sum_{k=1}^{\infty} (1 - \theta_k(\tau)) \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \int_0^{\infty} \int_{k\tau}^{k\tau+u} (k\tau + u - t) dF_{FD}(t) dF_M(u) \quad (16)$$

and the conditional expectations $E_j(X)$:

$$E_j(X) = P_j(\tau) \int_0^{\infty} \bar{F}_D(t) dt, \quad j = 1, 2 \quad (17)$$

$$E_3(X) = \sum_{k=1}^{\infty} (1 - \theta_k(\tau)) \left(\prod_{j=0}^{k-1} \theta_j(\tau) \right) \int_0^{\infty} \int_{k\tau}^{k\tau+u} (k\tau + u - t) \bar{F}_D(k\tau + u - t) dF_F(t) dF_M(u) \quad (18)$$

When the costs c_u , c_n , c_m and c_r are included, the expected corrective cost $E(C_C(\tau))$ becomes:

$$E(C_C(\tau)) = (c_r + c_u t_R + c_m E(T_M))(1 - P_P(\tau)) + \sum_{j=1}^3 (c_u E_j(Y) + c_n E_j(X)) \quad (19)$$

4.4. Mean time between repairs

The mean time between repairs $MTBR(\tau)$ is defined from the pipeline start-up until the production is restarted after the damaged pipe segment has been replaced. To evaluate the $MTBR(\tau)$, we distinguish between two scenarios: observation or not of the potential failure P . Then, we have:

$$MTBR(\tau) = E(N(\tau) + Z_P(\tau))\tau(P_P(\tau) + P_3(\tau)) + \left(\frac{1}{\lambda_1} + E(T_{PF}) + E(T_D) \right) (P_1(\tau) + P_2(\tau)) + E(T_M) \quad (20)$$

where $Z_P(\tau)$ denotes the number of inspections carried out to observe the coating failure, given that a replacement will be implemented before the observation of an potential pollution. Its associated conditional expectation $E(Z_P(\tau))$ is:

Table 1
Determination of the downtime Y and the pollution time X given the different three corrective scenarios

	Y	X
Scenario 1	T_M	T_D
Scenario 2	$Z(\tau)\tau + T_M(T_F + T_D)$	T_D
Scenario 3	0	$Z(\tau)\tau + T_M - T_F$

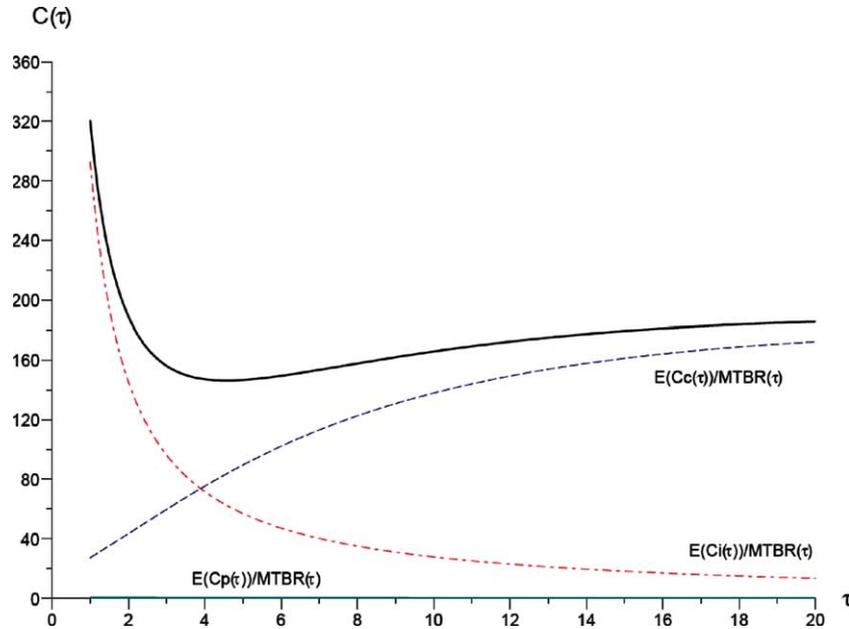


Fig. 3. Contributions of each maintenance actions on the overall mean maintenance cost per unit of time when the inspection interval varies—data: cost = { $c_i = 300$; $c_r = 10$; $c_u = 100$; $c_m = 1$; $c_n = 10,000$ }; degradation = { $\lambda_1 = 0.01$; $\gamma = 10$; $\nu = 0.01$ }; operation = { $p_0 = 0$; $p_\infty = 1$; $\lambda_0 = 1$; $t_R = 0.2$ }; transport = lognormal(0.5; 1); pollution detection = Weibull(1.1; 2.1).

$$E(Z_p(\tau)) = \sum_{k=1}^{\infty} \Pr(Z_p(\tau) \geq k) = \sum_{k=1}^{\infty} \prod_{j=0}^{k-1} \theta_j(\tau) \quad (21)$$

5. Numerical evaluation

In the previous section, we described mathematical formulation of the different quantities needed to determine the cost criterion. In this section, we describe different numerical studies to illustrate the performance of our maintenance policy and the existence of an optimal inspection interval. The proposed results are obtained by using the SCILAB program.

The curves in Fig. 3 illustrate the expected contributions of each maintenance actions (inspections, preventive and corrective replacements) on the overall average maintenance cost per unit of time, when the inspection interval τ varies from 0 to infinity. The maintenance data values (costs, degradation parameters, maintenance actions properties, transport and pollution parameters) are here arbitrarily chosen. The shape of the overall criterion is classical and presents a minimum value for an optimum inspection interval $\tau^* = 4.6$ and the associated minimum cost $C_\infty(\tau^*) = 146.26$. The contribution of the preventive repair is here very close to 0 and has a low impact on the overall criterion because of the low replacement cost compared with the corrective penalties. In our case, with our specific data, we notice that the optimal value of τ belongs to the PF interval, $E(T_{PF}) = 9.16$. A discussion about such an optimal condition was proposed in [4] for a deterministic PF interval value.

In Table 2, we study the behavior of the maintenance policy when the quality of inspections varies. For that, only p_0 and p_∞ take different values ($\lambda_0 = 1$). The data in Table 2 represent the

optimal inspection period τ^* and its associated average maintenance cost per unit of time $C_\infty(\tau^*)$ for a given configuration. The quality of the inspection has a big impact on the maintenance decision parameter. A perfect inspection allows to reduce the number of inspections and hence to reduce the associated inspection costs and the probability of a potential leakage. We can also discuss the impact of the value p_∞ when p_0 is fixed. When p_∞ decreases, the associated cost increases but, contrary to an intuitive idea, the associated optimal inspection interval also increases (we can easily imagine that the fact of not being able to detect rapidly a potential failure should lead to inspect the system more frequently). In fact, this point is not a general property of the influence of p_∞ on τ^* : If we increase the mean time to detect a pollution after a rupture, $E(T_D)$ ($\eta_D = 5.1$ instead of 2.1, e.g.), for a given value $p_0 = 0.5$, we have the decreasing of τ^* from 3.8, when $p_\infty = 1$ –3.4 when $p_\infty = 0.5$.

The last numerical experiment illustrates the applicability of our maintenance model in the definition of the inspection strategy. In practice, different non-destructive testing (NDT) techniques (ultrasonic or magnetic) can be used to detect any measurable defects in the wall of the pipe: corrosion, cracks, or other types of anomalies and especially in our case a hole in the coating. The efficiency of an inspection is dependent on the

Table 2
Optimal inspection period and associated maintenance cost { τ^* ; $C_\infty(\tau^*)$ } for different values of p_0 and p_∞

P_∞	p_0			
	0	0.5	0.75	1
0.5	{4.7–172.7}	{6.4–131.1}	–	–
0.75	{4.6–156.8}	{5.8–124.8}	{6.6–107.2}	–
1	{4.6–146.3}	{5.7–119.1}	{6.4–104.7}	{7.2–88.6}

Table 3
Estimation of the intrinsic quality parameter p_0 as a function of expert judgement

	Expert judgement			
	Very good	Good	Low	Very low
p_0	0.9	0.6	0.3	0.1

NDT technique and none of the techniques can be regarded as perfect. Hence, for different NDT techniques, we will have different values for the intrinsic inspection quality parameters, p_0 and p_∞ . For the sake of simplicity and because of a relative small influence of the p_∞ value on the optimal criterion, we assume that only the p_0 value will vary as a function of the efficiency of the NDT technique. Because the evaluation of the p_0 value given the NDT technique can be rather difficult, we propose to estimate p_0 with expert judgements. An example of p_0 estimation is illustrated in Table 3.

Furthermore, to optimize the efficiency of an inspection, a utility tool should be launched just before the inspection to clean the pipeline by removing dirt, rust and mill scale but also the wax that is stuck to the pipeline wall. The quality of cleaning before the inspection has a big impact on the probability of non-detection a potential failure. We assume that, for each given NDT techniques, the intrinsic sensitivity—including the quality of the cleaning—varies and depends only on the throughput speed of the Inspection In-line (ILI) instrument in the pipe. This throughput speed can be directly controlled by the flow velocity in the pipe.

Finally, an inspection strategy is defined by the choice of a NDT technique—NDT 1 and NDT 2—and the production flow rate, r . Different inspection costs are introduced to compare the inspection strategies. The cost for the NDT technique j , $c_i^{(j)}$, is fixed and the cost for the control of the intrinsic sensitivity is evaluated as a function of the economic losses due to the reduction of the production flow, $(1-r)c_u$. The costs, $c_i^{(j)}$, for NDT techniques and the associated sensitivities p_0 as a function of the production flow rates r are given in Table 4. A cost, c_i , for one inspection, given the NDT technique j and the production flow rate r is the sum $c_i = c_i^{(j)} + (1-r)c_u$.

To determine the best inspection strategy, the economic criterion (6) is optimized with the new inspection costs evaluated for each strategy. The optimal inspection intervals and the associated maintenance costs per unit of time are presented in Table 5 for the inspection data in Table 4. The general conclusion of this numerical experiment highlights the necessity to prioritize the intrinsic quality of the inspection. For a given NDT technique, the increase in the production flowrate leads to reduced quality of the sensitivity of the instrument and,

Table 4
Cost $c_i^{(j)}$ for two NDT techniques and the associated sensitivities p_0 as a function of the production flow rates r

	$c_i^{(j)}$	Flow rate r		
		0.3	0.6	0.9
NDT 1	300	0.9	0.6	0.3
NDT 2	200	0.6	0.3	0.1

Table 5
Optimal inspection period and associated maintenance cost $\{\tau^*; C_\infty(\tau^*)\}$ for different two NDT techniques and the production flow rates $r-p_\infty=0.95; \lambda_0=1$

r	0.3	0.6	0.9
NDT 1	{7.9–104}	{6.6–120}	{5.6–137}
NDT 2	{5.6–109}	{4.4–119}	{3.8–122}

thus, to decrease the optimal inspection interval and to increase the average maintenance cost per unit of time. The comparison between the two NDT techniques—especially the two last columns—shows that it can be interesting, under some production flow constraints, to privilege the increase in the inspection frequency instead of the sensitivity of the NDT technique.

6. Discussion

In this paper, we have proposed a new approach to maintenance optimization applied to a subsea pipeline exposed to internal corrosion–erosion. The optimization is based on a minimum average cost per time unit criterion. The cost function is defined by a decision parameter τ , the inspection interval. The optimization allows to balance between inspection cost and the effects of a potential oil pollution. Because of the complexity of the cost expression, the main difficulty in the optimization problem is the numerical evaluation of the cost function (6), the determination of the optimal value τ^* is a classical minimization of a single-parameter continuous function.

We have tried to establish the mathematical assumptions in accordance with the industrial context of the system, but realize that all facets of the system and its application have not been realistically covered. We are, therefore, convinced that these justifications need to be improved and tailored to the specific application. We are, however, confident that our approach is an important step to optimize pipeline replacement policies.

We have discussed the problem of an infinite study horizon. It would be interesting to numerically justify our choice by including a limited pipe life length function. The knowledge of the remaining time between the observation of a potential failure and the expected stop of the exploitation should be an additional relevant measure for the implementation of a preventive repair.

In our approach, we are only considering one single failure on one pipe segment. Even if our approach should be extended to tackle the problem of multiple failures on several pipe segments, a justification of this assumption can be related to the compromise between theoretical difficulties of our model and its applicability in an industrial context. Furthermore, because of the fast speed of the steel corrosion after a hole in the coating of a given pipe segment, we can assume that the probability of multiple failures in the given segment is negligible. Lastly, a justification concerning the assumption of restriction on a single segment is as follows. We have, in the same time, a very low probability of failure occurrence and a strong dependence between the failures related to the pipeline geometry and the operating conditions. Hence, if multiple failures occur during a

given time interval, they have a high probability to be in a restricted neighboring. The replacement of a given number of successive pipe segments would be a good approach and allows to justify our assumptions. To be completely validated, these considerations should be altogether confronted with the specific application.

References

- [1] Pierskalla WP, Voelker JP. A survey of maintenance models: the control and surveillance of deteriorating systems. *Naval Res Logist Q* 1979;23:353–88.
- [2] Valdez-Flores C, Feldman RM. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Res Logist Q* 1989;36:419–46.
- [3] Cho DI, Parlar M. A survey of maintenance models for multi-unit systems. *Eur J Oper Res* 1991;51:1–23.
- [4] Wang H. A survey of maintenance policies of deteriorating systems. *Eur J Oper Res* 2002;139:469–89.
- [5] Rausand M, Høyland A. *System reliability theory: models, statistical methods, and applications*. 2nd ed. New York: Wiley; 2004.
- [6] Moubray J. *Reliability-centered maintenance*.: Industrial Press Incorporation; 2001.
- [7] Zwingelstein G. *La maintenance basée sur la fiabilité: guide pratique d'application de la RCM*. Paris: Hermès; 1998.
- [8] Podofilini L, Zio E, Vatn J. Modelling the degrading failure of a rail section under periodic inspection. In: Spitzer C, Schmocker U, Dang VN, editors. *Probabilistic safety assessment and management. Proceedings of the international symposium, Berlin, Verlag; 14–18 June 2004*. p. 2570–75.
- [9] Castanier B, Rausand M. Application du modèle de l'intervalle PF pour l'optimisation de la politique de maintenance d'un pipeline en exploitation off-shore. In: Dolgui A, Dauzère-Péres S, editors. *Modélisation et simulation pour l'analyse et l'optimisation des systèmes industriels et logistiques. Proceedings of the international symposium, Nantes, Lavoisier; 1–3 September, 2004*. p. 797–803.
- [10] Christer AH. A review of delay time analysis for modelling plant maintenance. In: Osaki S, editor. *Stochastic models in reliability and maintenance*. Berlin: Springer; 2002. p. 89–117.



Predictive maintenance policy for a gradually deteriorating system subject to stress

E. Deloux^a, B. Castanier^{a,*}, C. Bérenguer^b

^aIRCCyN/Ecole des Mines de Nantes, Nantes, France

^bUniversité de Technologie de Troyes/CNRS, Troyes, France

ARTICLE INFO

Article history:

Received 2 October 2007

Received in revised form

4 April 2008

Accepted 23 April 2008

Available online 4 May 2008

Keywords:

Predictive maintenance

Stochastic modeling

Control chart

Economic performance

ABSTRACT

This paper deals with a predictive maintenance policy for a continuously deteriorating system subject to stress. We consider a system with two failure mechanisms which are, respectively, due to an excessive deterioration level and a shock. To optimize the maintenance policy of the system, an approach combining statistical process control (SPC) and condition-based maintenance (CBM) is proposed. CBM policy is used to inspect and replace the system according to the observed deterioration level. SPC is used to monitor the stress covariate. In order to assess the performance of the proposed maintenance policy and to minimize the long-run expected maintenance cost per unit of time, a mathematical model for the maintained system cost is derived. Analysis based on numerical results are conducted to highlight the properties of the proposed maintenance policy in respect to the different maintenance parameters.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Many manufacturing processes or structural systems suffer increasing wear with usage or age and are subject to random failures resulting from this deterioration [1] and most of them are maintained or repairable systems. Moreover, for some systems, such as aircrafts, submarines, military systems, and nuclear systems, it is very important to avoid failure during actual operation because it can be dangerous or disastrous. Therefore, maintenance of them is necessary since it can improve reliability. The growing importance of maintenance has generated an increasing interest in the development and implementation of optimal maintenance strategies for improving system availability, preventing the occurrence of system failures, and reducing maintenance costs of deteriorating systems. A lot of models for optimizing the maintenance scheduling and planning have been proposed [2]. Maintenance can be categorized into two major classes: corrective and preventive. Preventive maintenance itself can be categorized into two classes: systematic preventive maintenance and condition-based maintenance (CBM). CBM approaches are usually more efficient than systematic preventive maintenance policies based on the *a priori* statistical knowledge of the system lifetime [3]. CBM policies are thus particularly justified for critical systems and the collected information on the system

condition can be exploited by the maintenance decision-maker in order to maximize the availability of the system and to minimize its long-run expected cost. The price for this higher efficiency is an enhanced modeling work of the maintained system behavior in order to support correctly the maintenance decisions [4].

A classical assumption in CBM modeling is that the system failure can be explained by a deterioration process. One way to model continuous and gradual deterioration process due to wear and tear, e.g. systems subject to erosion (hydraulic structures, dikes) [5], consumption [6], cumulative wear (cutting tools) [7], is to represent the time-dependent deterioration by single continuous state stochastic process. The failure occurs when this deterioration exceeds a failure threshold [8]. But, in fact, because of the complexity of the systems and the influence of different variables and stresses [9,10] on the failure mechanism, it is difficult, in most of the cases, to reduce failure mechanisms to one deterioration process. Nevertheless, partial information on the system state can be tackled by the monitoring of easily observable covariates (e.g. vibration, temperature, humidity, etc.). Some works focused on non-maintained system reliability, propose relevant degradation models [11] such as hazard rate processes [12–15], competing risk processes [16] for example. However, only few are developed in such a context for CBM applications [17].

Moreover, a system failure can also induce some changes in the environmental conditions and especially in the previously mentioned easily observable covariates. For example, the internal vibration signal can be influenced by the present of springlers in a ball bearing. Hence, the observation of such covariates should

* Corresponding author. Tel.: +33 2 51 85 83 12; fax: +33 2 51 85 83 49.
E-mail address: bruno.castanier@emn.fr (B. Castanier).

Nomenclature			
$(X_k)_{k \in \mathbb{N}}$	discrete time process describing the deterioration at time t_k	τ	X-inspection period
$(Y_k)_{k \in \mathbb{N}}$	discrete time process describing a stress at time t_k	δ	routine maintenance action period
$(Z_k)_{k \in \mathbb{N}}$	discrete time process describing the system state at time t_k	θ	maintenance data vector
Δt	unit time length	Ξ	maintenance decision parameter vector
t_k	discrete time	c_{ix}	cost per inspection of X_t
a	scale parameter	c_{iz}	minimal inspection cost
F_A	failure acceleration factor due to stress	c_c	cost per corrective replacement
b	sensitivity to stress	c_p	cost per preventive replacement
L	cumulative degradation threshold	c_m	greasing cost
m	nominal level of Y_t	c_u	unavailability cost
α	deterioration parameter	$N_{ix}(t_k)$	number of planned inspections
σ	standard deviation of Y_t	$N_{iz}(t_k)$	number of unplanned inspections
μ	drift coefficient	$N_m(t_k)$	number of planned routine maintenance
λ	stress threshold	$N_p(t_k)$	number of preventives replacements
ξ	preventive replacement threshold	$N_c(t_k)$	number of correctives replacements
UCL	upper control limit	$D_u(t_k)$	cumulative unavailability duration
		$f_\lambda^{(k)}$	probability density function of the deterioration increment after k periods of time of a non-failed system
		r_k	time since the last routine maintenance

provide some partial information of the system state. Condition monitoring (CM) received a great deal of attention in the literature [18,19]. CM methods, e.g. statistical process control (SPC), applied to systems are used to detect potential failures or to determine updated failure rate values of sensitive components. For example, vibration analysis, oil analysis, humidity, temperature and so forth are variables allowing the detection of a failure, because the stress intensity enhances when a failure has occurred, but also the acceleration of the failure mechanisms due to stress. The use of CM techniques will generally improve plant production availability, and reduces downtime costs. Finally, modern maintenance paradigm consists of two major elements: preventive maintenance and CM decision analysis. Both of these have received significant development in recent decades but a longstanding gap continues to exist between these two elements. However, coordinate SPC and preventive maintenance presents economic performance [20–23]. Cassidy et al. [21] performed a preliminary investigation to model and analyze the relationship between maintenance and quality combining classical age-based approaches and SPC. Ben-Daya and Rahim [20] provide a framework for capturing maintenance and quality control interactions leading to models for their joint optimization. Linderman et al. [23] develop an analytical model to demonstrate this economic benefit.

In this study, we propose to answer to the different points presented previously, in particular the proposition of a model for a deteriorating system subject to stress and the construction and evaluation of a maintenance policy for such a model. More precisely, we consider a system with two failure mechanisms which are due to an excessive deterioration level and a shock due to a stressful environment. We propose here to model the failure process by the introduction of two explicative variables, a cumulative deterioration process, which is modeled by a non-decreasing stochastic process, and a stress covariate, which modeled by a stochastic process fluctuating around a given mean. To optimize the maintenance policy of the system, a combined SPC and CBM approach is proposed. CBM is used to inspect and replace the system according to the observed deterioration level. SPC is used to monitor the stress covariate. In this study, we develop a model which allows us to investigate the influence of the maintenance decision parameters on the total cost of the maintained system and we show that the long-run expected maintenance cost per unit of time can be minimize by an

appropriate joint choice of these decision variables. Furthermore, we perform a sensitivity analysis to highlight the properties of the optimal maintenance policies when maintenance data and system deterioration characteristics change.

This paper is organized as follows. In Section 2, the failure process and the associated variables are described. Section 3 is devoted to the presentation of the maintenance policy. In Section 4, results from numerical experiments illustrate and analyze the behavior of the proposed maintenance policy. Finally, in Section 4, conclusions are drawn from the work.

2. Description of the combined stress-degradation failure process

This section is devoted to describe the system failure process, the relationship between the stress and the system degradation behavior. Then, the associated mathematical model is derived.

2.1. Description of the system failure process

We consider a single-unit system subject to two failure mechanisms. The first failure process results from an excessive deterioration level. The second one is a lethal shock failure process which can be explained, at the same time, by the deterioration level and the stressful environment [14].

In order to illustrate our model, we will rest our various assumptions on the analysis of the particular failure mechanisms of a guiding system of a tramway due to chippings and vibrations.

2.1.1. Cumulative deterioration failure

The first failure mechanism results from an excessive deterioration level which can be described by a measurable variable. The system stops to fulfill its function as soon as the aging variable is greater than a pre-determined level L . In this case, either a “hard” failure or a “soft” failure has occurred, viz. an important deterioration is present which reduces the system performance too much and the system is considered as failed, even if it is still working from, e.g., a mechanical point of view. For example, a production system will continue to produce even if its performance level is not good enough. However, the resulting products will be of poor quality and should be scrapped and in

this case the system can be considered in the failed state. The threshold L , in this example, represents the minimum level of system performance in order to obtain products conform to requirements. The determination of this threshold rests on a level of expertise and data related to non-conformities. Within the framework of the safety, L is a tolerance which can be imposed directly by a regulation, for example, a wear threshold for tyres and brake pads. Furthermore, a failure is assumed to be not self-announcing, for example, a bridge, a dike, etc. may not be able to sustain its nominal load any longer, but an inspection is necessary to reveal it. Obviously, it is not reasonable to wait for the structure to collapse before declaring that it failed; it is in a failed state as soon as it can no longer fulfill its intended function under a given performance level. The threshold L can thus be seen as a deterioration level which must not be exceeded for economical or security reasons. An inspection is needed in order to detect if the deterioration level is greater than this threshold L .

Let $(X_k)_{k \in \mathbb{N}}$ be the continuous increasing stochastic process describing the deterioration process on an infinite discrete time $(t_k = k\Delta t)_{k \in \mathbb{N}}$. Later in this paper, the value of the unit time length Δt will be fixed to 1. We assume that $(X_k)_{k \in \mathbb{N}}$ is modeled by a phase-type distribution with parameter α , i.e. an increment over one fixed time unit, $X_{k-1} - X_k$, is exponentially distributed with parameter α and a probability density function $f(x) = \alpha e^{-\alpha x}$. This model is classical and generally well adapted to cumulative wear, fatigue, crack growth, corrosion, etc., with a minimal information on the considered deterioration process [24].

The process $(X_k)_{k \in \mathbb{N}}$ represents in the example of the rollers, the propagation of the chipping.

2.1.2. “Stress-failure” process

The second failure mechanism is related to a stress-failure process. Let assume the system is subject to an environmental stress that can be external to the system (temperature, humidity, etc.) or a direct consequence of the system operating mode (internal vibrations, internal temperature, etc.). Let $(Y_k)_{k \in \mathbb{N}}$ be the associated stress process where Y_k represents the stress intensity at time t_k . Y_k is assumed to be a classical Gaussian process with a given mean m and standard deviation σ .

The stress-failure process is modeled by a classical shock process, a failure can occur because of the stress with a given probability $1 - q$. The associated stress-deterioration failure probability $1 - q$ depends on the stress intensity Y_k . To model the influence of the stress on this probability, a stress threshold λ is introduced. If $Y_k < \lambda$, the system is considered in the nominal operating state and the stress has a low impact on the probability of a failure. On the other hand, when $Y_k \geq \lambda$, the system is considered stressed and this increases the probability of a failure due to the stress. Let F_A be the associated failure acceleration factor due to stress. Moreover, because the sensitivity of the system to stress can differ with time and usage [25,26], the probability of a failure due to stress is a function of the current deterioration level X_k and the time t_k .

Let Z_k be the system state indicator. If $Z_k = 0$, the system is in a good state and $Z_k = 1$ indicates a failed system. The failure due to shock for a non-maintained system is given by the following conditional probability of a shock given the value of X_k , the magnitude Y_k , $t_k \geq 1$:

$$\begin{cases} P(Z_k = 1 | X_k = x < L, Y_k < \lambda) = 1 - e^{-ax - bt_k} \\ P(Z_k = 1 | X_k = x < L, Y_k \geq \lambda) = 1 - F_A e^{-ax - bt_k} \end{cases} \quad (1)$$

where a is the scale parameter and b represents the sensitivity of the system to vibrations.

As mentioned previously, the system condition can have some effects on the environmental stress, especially on the internal operating conditions. Let us assume that a system failure due to an excessive deterioration or a shock increases continuously the stress intensity in average. In case of failure, a drift coefficient μ is introduced in the intensity stress process. Hence, the new process $(Y_k)_{k \in \mathbb{N}}$ after a failure is a Brownian motion with drift [27].

$(Y_k)_{k \in \mathbb{N}}$ represents in the example of the rollers the vibration amplitude. In fact, an increase in vibration can be associated with deteriorated states and thus a detection of a drift in vibration could prevent failures.

The two failure mechanisms are sketched in Fig. 1. The effect of a failure in the two cases on the stress intensity process is also illustrated in the representation of Y_k . Finally, a system with a competing failure process is considered, the conditional probability of failure given the value of X_k and/or Y_k for $t_k \geq 1$ is obtained by adding to the set of equations (1) the probability of failure due to an excessive deterioration level:

$$\begin{cases} P(Z_k = 1 | X_k = x \geq L) = 1 \\ P(Z_k = 1 | X_k = x < L, Y_k < \lambda) = 1 - e^{-ax - bt_k} \\ P(Z_k = 1 | X_k = x < L, Y_k > \lambda) = 1 - F_A e^{-ax - bt_k} \end{cases} \quad (2)$$

2.2. Reliability function

Let $f_\lambda^{(k)}$ be the probability density function of the deterioration increment after k periods of time of a non-failed system subject to deterioration and stress. $f_\lambda^{(k)}$ is given by

$$\begin{aligned} f_\lambda^{(k)}(x) &= ((1 - p)F_A + p)^k \frac{\alpha^k}{a^{k-1}(k-1)!} \\ &\times (1 - e^{-ax})^{k-1} e^{-\alpha x} \prod_{j=1}^k e^{-bj} \end{aligned} \quad (3)$$

with $p = P(Y_k < \lambda) = \Phi((\lambda - m)/\sigma)$; $\Phi(\cdot)$ is the cumulative distribution function of the normal distribution.

The reliability of the unmaintained system at time $t_k = k\Delta t$ can be obtained directly by integrating $f_\lambda^{(k)}$ between 0 and L . The reliability $R(t_k)$ is given by

$$\begin{aligned} R(t_k) &= \int_0^L f_\lambda^{(k)}(x) e^{-ax} dx \\ &= ((1 - p)F_A + p)^k \int_0^L \frac{\alpha^k}{a^{k-1}(k-1)!} e^{-(\alpha+a)x} \\ &\times (1 - e^{-ax})^{k-1} \prod_{j=1}^k e^{-bj} dx \end{aligned} \quad (4)$$

The curves in Fig. 2 represent the reliability function (cf. Eq. (4)), the associated probability of failure due to the different failure mechanisms (shock and excessive deterioration level) and the reliability without any stress so with a probability of failure due to shocks equal to 0. This figure illustrates the impact of the “stress-failure” process on the reliability of the system.

3. Definition and evaluation of the maintenance policy

This section presents the maintenance decision framework which corresponds to a “control limit policy” [28,29]. First, the structure of the maintenance policy is presented to define when an inspection or a replacement should be implemented. Then, the mathematical expressions of the associated long-run maintenance cost per unit of time are developed to optimize the maintenance decision.

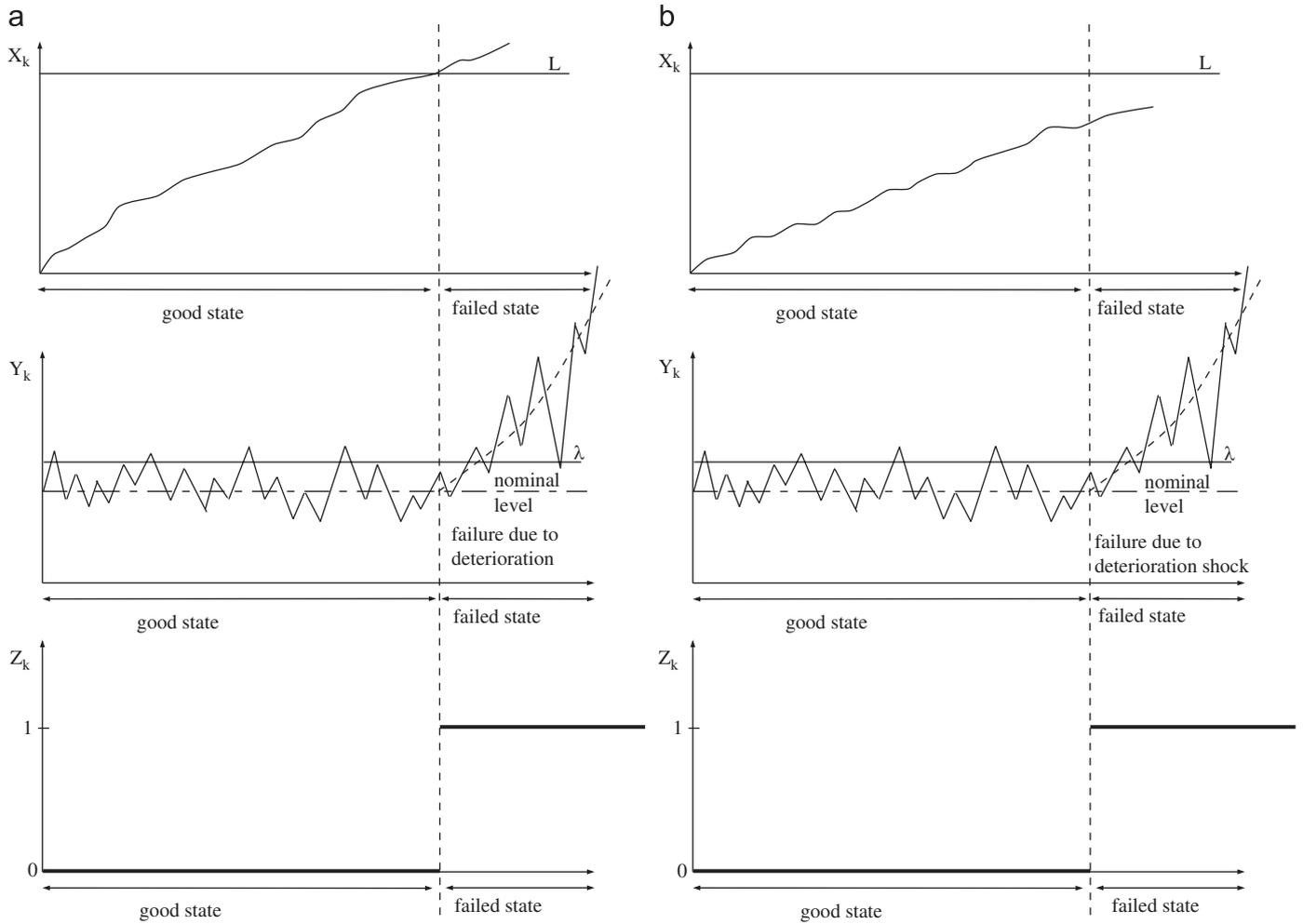


Fig. 1. Evolution of the deterioration process, stress process and system state of a non-maintained system. (a) Failure due to an excessive deterioration level, $X(t) > L$. (b) Failure due to shock.

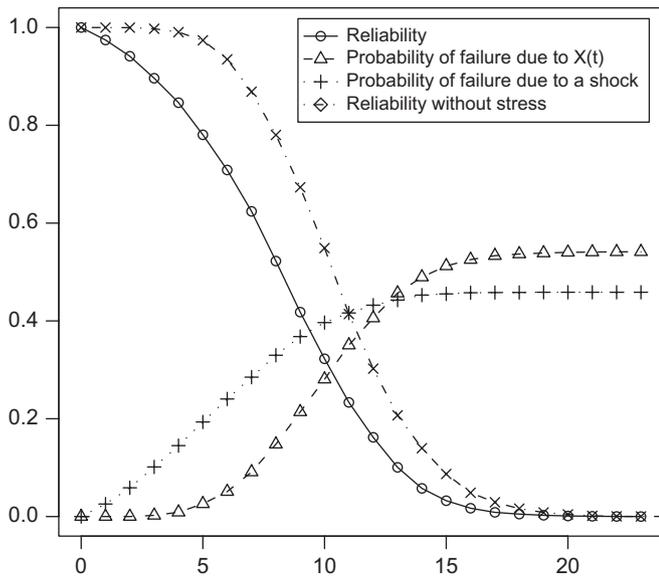


Fig. 2. Reliability and probability of failure due to an excessive deterioration or due to a shock with $\alpha = 5$; $a = 0.05$; $F_A = 0.96$; $b = 0.025$; $L = 2$.

3.1. Structure of the maintenance policy

The maintenance policy is based on the failure model described in the previous section. Moreover, failures are assumed to be non-obvious, an inspection is required to know the system state. Two types of inspection are available, an inspection of the deterioration level (X -inspection) and an inspection of the system state (Z -inspection). The maintenance policy also offers the possibility of replacements either preventive or corrective. Finally, a routine maintenance action can be performed to reduce the impact of the stressful environment.

3.1.1. X -inspection

The cumulative deterioration level X_k can be observed only through costly inspections. Let c_{ix} be the unitary inspection cost. Even if non-periodic inspection strategies are optimal [24] from a mathematical point of view, their implementation in an industrial context remains difficult and a periodic strategy is more frequently used, thus a periodic age-based (versus time-based) strategy is proposed. The first inspection after a potential replacement of failed system is scheduled τ time units after the restarting time and then the system is inspected at regular time

intervals $(\tau, 2\tau, \dots)$ until the next replacement. The benefit of such an assumption is a reduced number of the decision parameters, only the inspection period τ . This inspection is assumed to be perfect in the sense that it reveals the exact deterioration level X_k .

3.1.2. Z-inspection

The stress variable Y_k is continuously monitored, i.e. at each time t_k . A classical control chart for individual observations [30] is implemented to detect a potential failure by the observation of a drift in Y_k . If the value Y_k is greater than the upper control limit, UCL, a minimal inspection is performed at a given unitary cost c_{iz} . This minimal inspection reveals only if the system is failed or not, i.e. $Z_k = 0$ or $Z_k = 1$, but delivers no information on X_k . A minimal inspection is less expensive than an inspection on X_k , $c_{iz} < c_{ix}$. In this paper, the value of UCL is not evaluated in order to minimize the number of false alarms ($Y_k > \text{UCL}$ but the system is still functioning) and non-detections (the system is in the failed state but $Y_k < \text{UCL}$), but the value of UCL should be optimized to balance the cost incurred by the unavailability time due to a failure (incurred by a non-detection) and the cost incurred by false alarms (the cost of a false alarm corresponds to the cost of the system state inspection, c_{iz}). The computations of the number of false alarms and non-detections are developed in Appendix A.1. It should be noted that if the system is in the failed state during this inspection and thus a replacement is implemented, the next X-inspection is planned τ time units after.

3.1.3. Preventive and corrective replacement

A replacement can take place to renew the system when it is failed (corrective replacement) or to prevent the failure (preventive replacement). A corrective replacement is performed when the system is observed in the failed state during an X-inspection or during a minimal Z-inspection. The unitary cost of a corrective replacement c_c includes all the direct and indirect costs incurred by this maintenance action. Only the unavailability cost c_u per unit of time the system is failed has to be added to c_c . The decision rule for a preventive replacement is a classical control limit rule: if ξ is the preventive replacement threshold, a preventive replacement is performed during the inspection on X_k if the deterioration level belongs to the interval (ξ, L) . Let c_p be the preventive replacement cost ($c_p < c_c$). Replacements are assumed to be instantaneous.

3.1.4. Routine maintenance action

A routine maintenance action has to be implemented to reduce the impact of the stress on the system modeled by a decreasing in the failure hazard function. This maintenance action would not correspond to a minimal repair. Routine maintenance is assumed to be instantaneous. In the example of the tramway, this operation can be, e.g., the lubrication of the rails. The effectiveness of this action decreases with time and a new operation has to be scheduled. A periodic scheme is proposed for this operation, δ is the time between two successive actions. Let c_m be the unitary operation cost (in the example, this cost would also include the tramway immobilization due to the unavailability of the track). Let r_k be the time passed since the last action. We have $r_k = t_k - \delta \cdot [t_k/\delta]$ where $[\cdot]$ is the integer part function. To model the effect of this maintenance on the system, r_k is introduced in the conditional probability equations (1) instead of t_k and thus the following system of equations is obtained:

$$\begin{cases} P(Z_k = 1 | X_k = x < L, Y_k < \lambda) = 1 - e^{-ax - br_k} \\ P(Z_k = 1 | X_k = x < L, Y_k > \lambda) = 1 - F_A e^{-ax - br_k} \end{cases} \quad (5)$$

Fig. 3 represents the influence of the routine maintenance period, δ , on the reliability and on the different probabilities of failure.

Fig. 3(a) highlights the impact of the routine maintenance period on the reliability. The different curves are superimposed until $t_1 = 1$ the time of the routine maintenance in the case $\delta = 1$. At $t_2 = 2$ the impact of the routine maintenance ($\delta = 1$) on the reliability is visible and leads to an increase in the reliability of 4.9% when compared with the cases $\delta = 100$ and $\delta = 3$. The curves for $\delta = 3$ and $\delta = 100$ are superimposed until $t_3 = 3$ the time of the first routine maintenance in the case $\delta = 3$. At $t_4 = 4$ the reliability on the case of $\delta = 3$ is improved of 17.4% when compared to $\delta = 100$. The routine maintenance action slows down the decrease of the reliability each time it is performed, see the reliability curve on Fig. 3(a), at times t_3, t_6, \dots

The curves presented in Fig. 3(b) illustrate the impact of the routine maintenance on the probability of failure due to shock. As previously mentioned, the three curves are superimposed until $t_1 = 1$ and then the curves which correspond to $\delta = 3$ and 100 are superimposed until $t_3 = 3$. On the curves which correspond to $\delta = 3$, irregularities are visible at $t_3 = 3$; $t_6 = 6$; $t_9 = 9$ times of the routine maintenance.

The curves presented in Fig. 3(c) illustrate the impact of the routine maintenance on the probability of failure due to an excessive deterioration level. The routine maintenance has an impact only on the probability of failure due to shock (cf. Eq. (1)), the probability of failure due to an excessive deterioration level is not influenced by this maintenance action, the shape of the curves is always the same (all the curves are smooth) and only the proportion of failure due to this failure mechanism changes.

3.1.5. Maintenance decision variables

Finally, the maintenance decision parameters which should be optimized are:

- τ : the inspection period which allows to balance the cumulative inspection cost and the early detection or the prevention of a failure;
- ξ : the preventive maintenance threshold which allows to set the ratio between preventive and corrective replacements;
- UCL: the upper limit control which balances false alarms and failure detection;
- δ : the maintenance action period which allows for a reduction of the probability of shock.

An illustration of the evolution of a maintained system and of the maintenance decision is presented in Fig. 4.

3.2. Long-run average cost per time unit

This section is presents the mathematical framework for evaluating the performance of the proposed maintenance policy with respect to the long-run maintenance cost per unit of time. The optimization criterion is a function of the maintenance decision parameter vector $\Xi = (\tau, \xi, \text{UCL}, \delta)$, the system deterioration characteristics $\theta = (a, b, F_A, \alpha, m, \mu, \sigma, \lambda, L)$ and the different maintenance costs.

The expected cumulative maintenance cost up to time t_k incurred by the successive actions on the system knowing θ the system characteristics vector is given by the following equation:

$$\begin{aligned} \mathbb{E}(C(t_k)) = & c_{ix} \cdot \mathbb{E}(N_{ix}(t_k)) + c_{iz} \cdot \mathbb{E}(N_{iz}(t_k)) \\ & + c_p \cdot \mathbb{E}(N_p(t_k)) + (c_c - c_p) \cdot \mathbb{E}(N_c(t_k)) \\ & + c_u \cdot \mathbb{E}(D_u(t_k)) + c_m \cdot \mathbb{E}(N_m(t_k)) \end{aligned} \quad (6)$$

where $\mathbb{E}(\cdot)$ represents the expected value knowing θ ; $N_{ix}(t_k)$ (respectively, $N_{iz}(t_k)$) is the number of X-inspections (respectively, Z-inspections) performed before t ; $N_m(t_k)$ is the number of routine maintenance performed before t_k ; $N_p(t_k)$ (respectively, $N_c(t_k)$) is

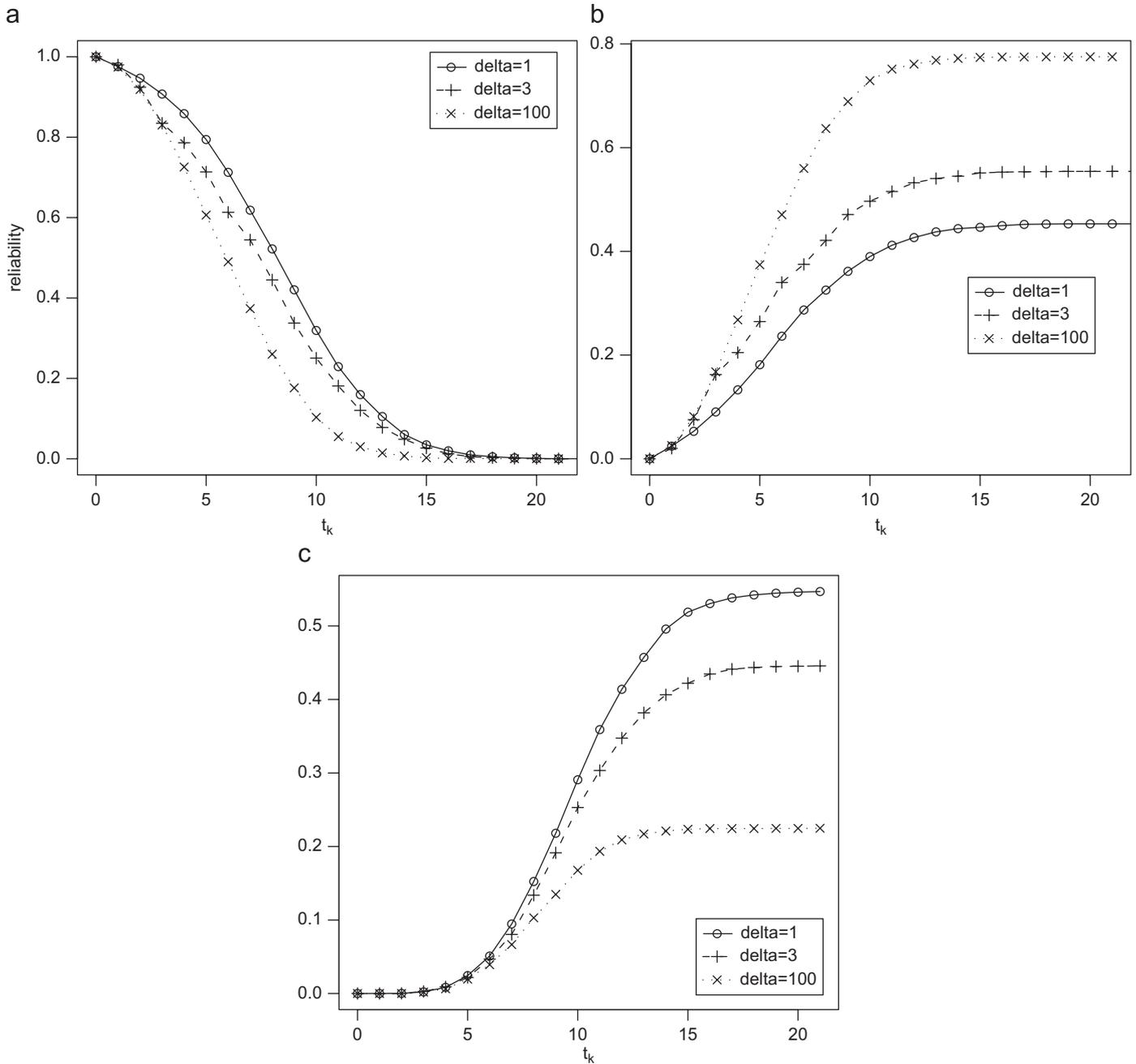


Fig. 3. Illustration of the influence of routine maintenance period δ on the reliability and the different probabilities of failure. (a) Influence of δ on the reliability. (b) Influence of δ on the conditional probability of failure due to shock at t_k given $Z_{t_{k-1}} = 0$. (c) Influence of δ on the conditional probability of failure due to X_{t_k} at t_k given $Z_{t_{k-1}} = 0$.

the number of preventive (respectively, corrective) replacement performed before t_k ; $D_u(t_k)$ is the time that the system passed in failed state over $(0, t_k)$ knowing the system is considered as failed as soon as the system deterioration level exceeds its critical level L .

By using classical renewal arguments [31], this maintenance criterion can be expressed on a renewal cycle S defined between two consecutive replacements because the system is considered as good as new after a replacement. Then, the expression of the long-run average maintenance cost per unit of time is:

$$C_\infty(\Xi; \theta) = \lim_{t_k \rightarrow \infty} \frac{C(t_k)}{t_k} = \frac{\mathbb{E}(C(S))}{\mathbb{E}(S)}$$

$$= \frac{c_{ix}\mathbb{E}(N_{ix}(S)) + c_{iz}\mathbb{E}(N_{iz}(S)) + c_p + c_m\mathbb{E}(N_m(S)) + c_u\mathbb{E}(D_u(S))}{\mathbb{E}(S)}$$

$$+ \frac{(c_c - c_p)\mathbb{E}(N_c(S))}{\mathbb{E}(S)} \quad (7)$$

The different expectations in Eq. (7) can be evaluated by a function of the reliability of the maintained system at steady-state, $R_m(t_k)$. The difference between $R(t_k)$ (the reliability of the unmaintained system) and $R_m(t_k)$ results in the possibility of preventive replacement after inspection and thus reduces the possible states from which failure can occur. In order to determine $R_m(t_k)$ for all $t_k \in (0, S)$, two cases are identified:

- **Case 1:** No inspection on the system has been made before t_k ($t_k < \tau$). The probability density function for the system to be in a good state is only a function of the observed deterioration

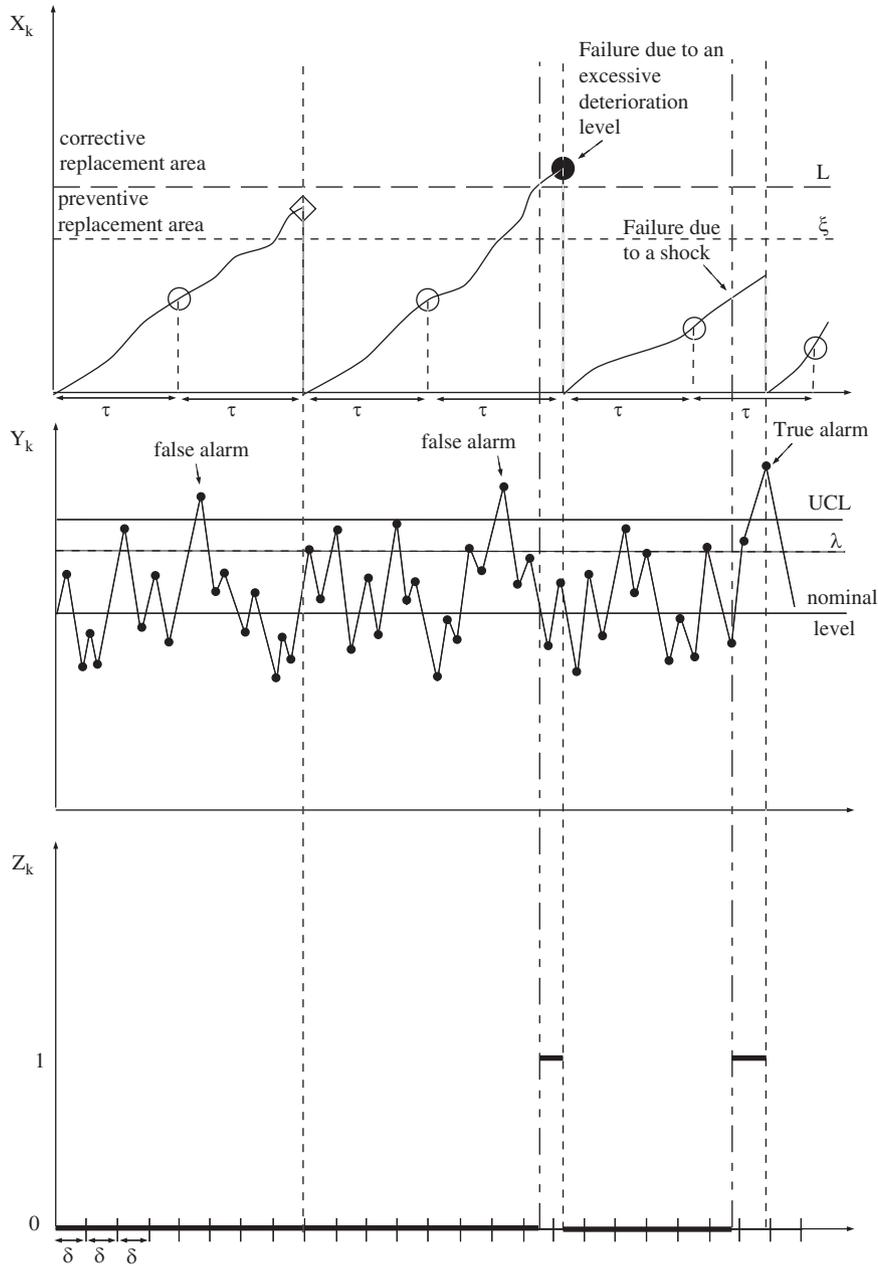


Fig. 4. Example of the evolution of the deterioration process, the stress process and the system state of a maintained system: ○ X-inspection; • Z-inspection; ◇ preventive replacement; • corrective replacement.

level $x \in (0, L)$ at time t_k and the consecutive shocks in $(0, t_k)$. Hence for $t_k < \tau$, we have

$$R_m(t_k) = \int_0^L f_\lambda^{(k)}(x) e^{-ax} dx \tag{8}$$

- Case 2: At least one maintenance action has been made at t_k ($t_k \geq \tau$). The probability density function for the system to have been never replaced is a function of the observed deterioration level $x \in (0, \xi)$ during the last inspection at time $([t_k/\tau], \tau)$, (where $[\cdot]$ is the integer part function), the deterioration level reached since the last inspection $y \in (x, L)$ and the consecutive shocks in $(0, t_k)$. Hence for $t_k \geq \tau$, we have

$$R_m(t_k) = \int_0^\xi \int_x^L f_\lambda^{(k-[k/\tau], \tau)}(y-x) f_\lambda^{([k/\tau], \tau)}(x) e^{-a(y+x-[k/\tau], \tau)} dy dx \tag{9}$$

The expected length of a regenerative cycle $\mathbb{E}(S)$ can be expressed by the four following exclusive scenarios:

- Scenario 1: A failure occurs between $(n\tau - \Delta t; n\tau)_{n \in \mathbb{N}}$ (see Fig. 5) and is detected during an X-inspection at time $n\tau$, the system is replaced at this time.
- Scenario 2: A failure occurs between two inspections at time $t_i \in ((n-1)\tau; n\tau - \Delta t)_{n \in \mathbb{N}}$ (see Fig. 5) and is not detected by the control chart but only during the following X-inspection at $n\tau$.
- Scenario 3: A failure occurs between two inspections at time $t_i \in ((n-1)\tau; n\tau - \Delta t)_{n \in \mathbb{N}}$ (see Fig. 5) and is detected by the control chart (a Z-inspection) before the next X-inspection.
- Scenario 4: A cycle ends by a preventive replacement during an X-inspection

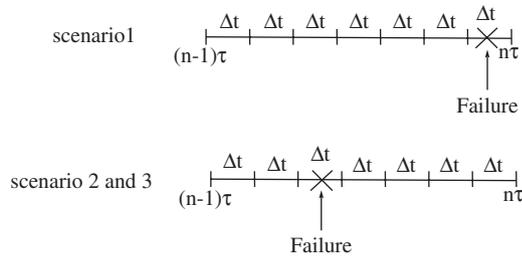


Fig. 5. Scale allowing to locate the three scenarios.

Hence we have

$$\begin{aligned}
 \mathbb{E}(S) = & \sum_{i=0}^{\infty} \left[\underbrace{\mathbb{1}_{\{i>0\}} i \tau (R_m(i\tau - 1) - R_m(i\tau))}_{\text{Scenario 1}} \right. \\
 & + \underbrace{\sum_{j=i\tau+1}^{(i+1)\tau-1} i \tau (R_m(j) - R_m(j-1)) \prod_{k=j}^{(i+1)\tau-1} \mathbb{P}(Y_k < \text{UCL} | \text{fail.})}_{\text{Scenario 2}} \\
 & + \underbrace{\sum_{j=i\tau+1}^{(i+1)\tau-1} j \mathbb{P}(Y_j > \text{UCL} | \text{fail.}) (R_m(j) - R_m(j-1))}_{\text{Scenario 3}} \\
 & + \underbrace{\mathbb{1}_{\{j>i\tau+1\}} \sum_{k=i\tau+1}^{j-1} j (R_m(k) - R_m(k-1)) \prod_{w=k}^{j-1} \mathbb{P}(Y_w < \text{UCL} | \text{fail.})}_{\text{Scenario 3}} \\
 & \left. + \underbrace{\mathbb{1}_{\{i>0\}} i \tau \left(\int_0^{\xi} f_{\lambda}^{((i-1)\tau)}(x) dx - \int_0^{\xi} f_{\lambda}^{(i\tau)}(x) dx \right)}_{\text{Scenario 4}} \right) \quad (10)
 \end{aligned}$$

The computations of the other quantities are obtained similarly and are presented in Appendix A.

4. Performance evaluation of the maintenance policy

This section is devoted to illustrate the benefits of the proposed maintenance policy and compare its performance to classical policies. Numerical experiments are first proposed to highlight the convexity of the cost function and to analyze the influence of each maintenance decision parameter on the criterion. The economical performance of the proposed policy is illustrated by a comparison to classical approaches. Then, numerical studies are developed to illustrate the robustness of the policy when there are uncertainties on the system characteristics. All the numerical results are provided here by the use of R (www.r-project.org) software, specific programs are developed to numerically evaluate each expectation in Eq. (7) and the classical numerical gradient procedure provided by the R software is used.

4.1. Influence of the maintenance parameters

To highlight the impact of the four decision parameters (τ , ξ , UCL, and δ) and the benefit of the combination of SPC and CBM, the performance of the maintenance policy is discussed when the decision parameters values change for a given set of maintenance data. This set of maintenance data is here a set of generic values which are arbitrarily fixed: $\alpha = 5$, $a = 0.05$, $b = 0.025$, $F_A = 0.96$, $L = 2$, $m = 10$, $\mu = 0.22$, $\sigma = 1$, $\lambda = 10.5$, $c_c = 100$, $c_p = 30$, $c_{ix} = 2$, $c_{iz} = 0.5$, $c_u = 25$, $c_m = 4$. The cost curves presented in Fig. 6 are obtained when two of the four maintenance decision parameters change and the two other parameters are fixed to their optimal value.

4.1.1. Influence of τ

The influence of τ is illustrated in Figs. 6(b), (d) and (e). An inspection allows for prevention and detection of a failure. When the inspection interval is too long, the main decision rule is restricted to the corrective replacement detected by the control chart. This information is illustrated in the flatness of the curve presented in Fig. 6(e), the valley-shape of the curve in Fig. 6(d) and in the increase in the cost value when UCL increases in Fig. 6(b). In the last case, the detection of a failure is thus delayed because of the lack of information on the system. When the inspection interval is small, a lot of information is available on the system with brief unavailability periods due to a failure. A monitoring policy is not relevant in this case (see the flatness of the curve when τ is small in Fig. 6(b)).

4.1.2. Influence of ξ

The influence of ξ is illustrated in Figs. 6(a), (e) and (f). The preventive replacement threshold allows to prevent failure in case of sufficient information on the system. If this information is limited, the value of ξ has no more influence on the criterion (see Fig. 6(e), e.g., when $\tau > 6$). When the frequency of the inspection is very high, the shape of the curve in Fig. 6(e) illustrates the excessive cost of a too early preventive replacement. Moreover, the robustness of the optimal threshold value respectively with the control chart policy and the routine maintenance action period is illustrated by the valley-shape curves in Fig. 6(a) respectively 6(f).

4.1.3. Influence of UCL

The influence of UCL is illustrated in Figs. 6(a)–(c). As mentioned previously in Section 4.1.1, the control chart allows to detect the failures. This conclusion is clearly illustrated in Fig. 6(b) for high values in τ and UCL. Another aspect of the control chart illustrated in the decrease in the first part (UCL < 10) of the curve in Fig. 6(b) when τ is very high ($\tau = 10$) is the balance with the detection of a failure and false alarms. The two other curves do not provide any relevant information on the impact of UCL because, finally, a high level of information on the system state is available due to a relative low optimal value of the inspection interval $\tau^* = 25$.

4.1.4. Influence of δ

The influence of δ is illustrated in Figs. 6(c), (d) and (f). The routine maintenance action allows to reduce the probability of failure due to a shock, thus it influences directly the reliability of the system. This information is illustrated in Figs. 6(c), (d) and (f), whatever the value of the other parameters, the optimized value of δ is always the same. So, later in this paper, the value of this parameter will be fixed to this optimal value, $\delta^* = 2$.

4.2. Influence of the unit maintenance costs

The global maintenance cost gathers inspection, preventive replacement, corrective replacement, routine maintenance and unavailability costs, see Eq. (7). In this section, the consequences of the variations of those costs on the optimal maintenance policy parameters are investigated. Optimal maintenance parameters have been computed for several unit costs configurations.

An example of the evolution of the optimal preventive threshold, the optimal upper control limit and the optimal inspection period as a function of the cost of the inspection on X is given in Fig. 7. A low cost of the inspection on X causes a low inspection period ($\tau = 1$ for $c_{ix} = 1$). Hence as long as no replacement is triggered, inspections are scheduled as frequently as possible. The proposed policy is closed to a systematic inspection/replacement one (cf. Section 3.4.2).

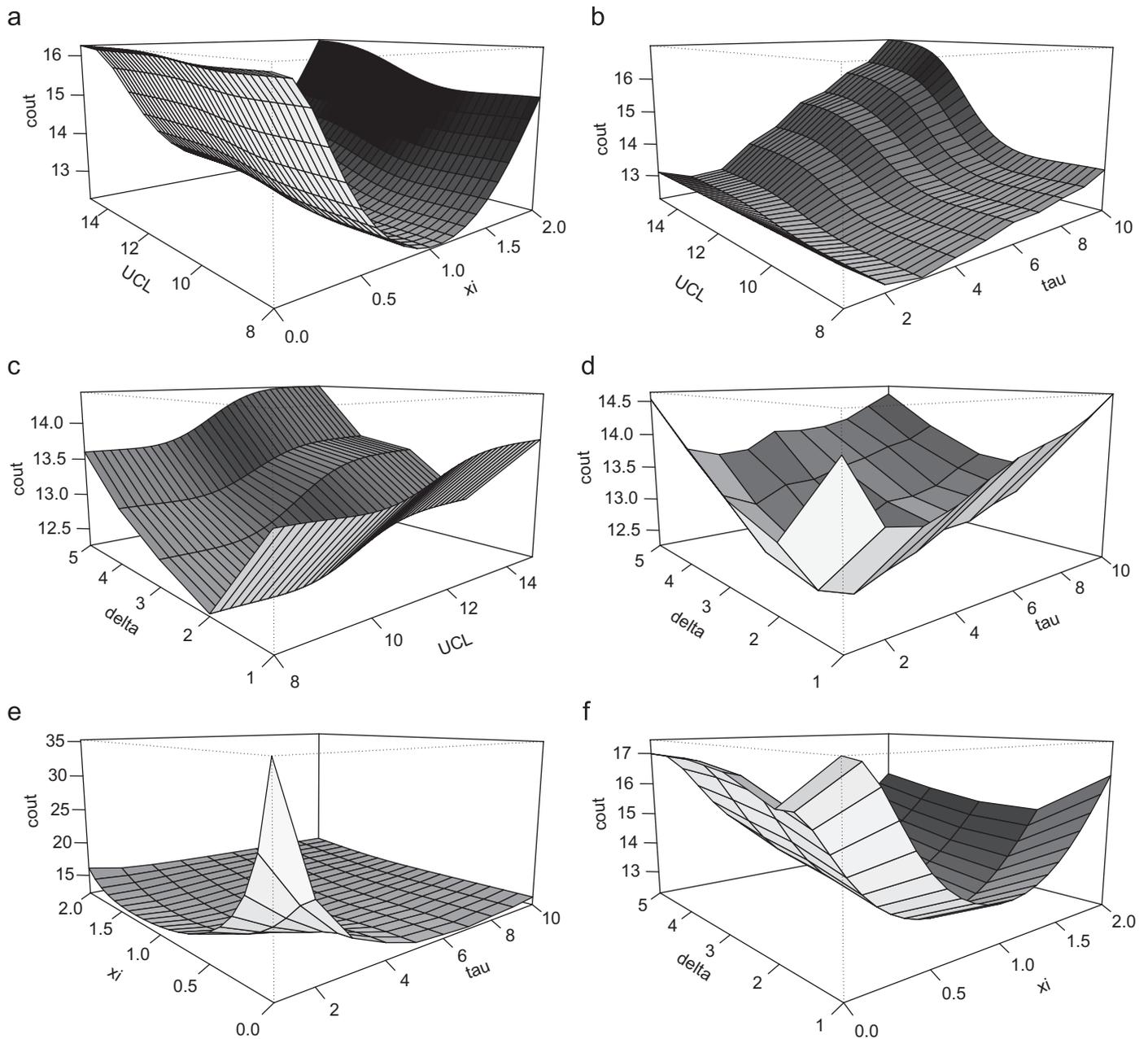


Fig. 6. Cost evolution according to ξ , τ , UCL and δ . (a) Influence of UCL and ξ on the cost (τ and δ fixed). (b) Influence of UCL and τ on the cost (ξ and δ fixed). (c) Influence of δ and UCL on the cost (τ and ξ fixed). (d) Influence of δ and τ on the cost (ξ and UCL fixed). (e) Influence of ξ and τ on the cost (δ and UCL fixed). (f) Influence of δ and ξ on the cost (τ and UCL fixed).

The value of the inspection period τ increases with c_{ix} , but in order to balance the lost due to the increasing value of τ , ξ and UCL decrease. When τ becomes really high, the upper limit control allows for detection of a failure between two X-inspections and thus, this proposed policy is closed to a corrective replacement one: only a corrective replacement can be implemented as soon as a failure is detected by the monitoring of the stress covariate Y_k (cf. Section 3.4.2).

Other numerical experiments have revealed the influence of the other maintenance costs. The inspection period τ and the preventive threshold ξ increase with c_p , whereas UCL decreases. A corrective replacement has to be preferred to a costly preventive replacement provided that the unavailability time is weakest as possible thanks to the value of the upper limit control. The upper control limit increases with c_c and c_{iz} so, in both of these cases the proposed policy tends to an inspection/replacement policy. The

three maintenance parameters decrease with c_u and the policy tends again to an inspection/replacement policy.

In a general way, it is noteworthy that an empirical choice of the maintenance parameters as an exclusive function of the unit costs is quite difficult to obtain. The separation of their influence remains possible only with some particular configurations. The last point emphasizes the need for an optimization procedure.

4.3. Comparison with classical policies

The economical performance of the proposed policy (hereafter denoted Policy 0) is illustrated by a comparison to classical approaches. The following three classical policies are considered.

- **Policy 1—periodic inspection/replacement.** The deterioration level of the system is periodically inspected with a periodicity

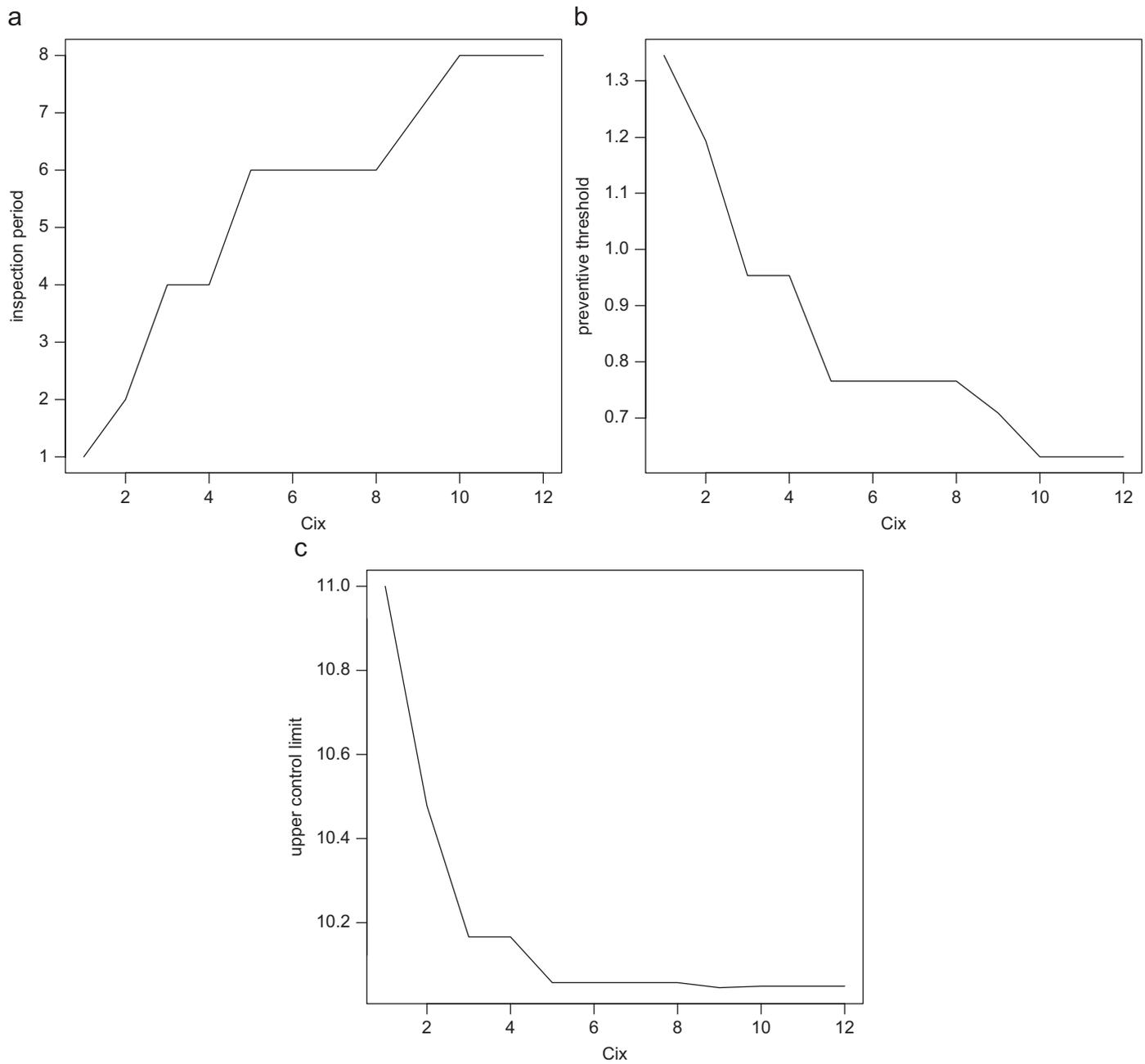


Fig. 7. Influence of the cost of the inspection of X_t . (a) Influence on the inspection period. (b) Influence of the preventive threshold. (c) Influence of the upper limit control.

τ and the system is preventively replaced if its deterioration is found to be greater than ζ . In case of failure, the system is correctively replaced as soon as the failure is detected by the inspection on X .

- **Policy 2—control chart.** Only a corrective replacement (cost c_c) could be implemented as soon as a failure is detected when the stress covariate Y_k becomes greater than UCL.
- **Policy 3—block replacement policy.** No inspection is performed. The system is replaced every time T . The replacement time has to be optimized to balance the system unavailability and the maintained system lifetime.

4.3.1. Performance of the proposed maintenance policy when system characteristics vary

The economic performance of each policy is evaluated when the system characteristics vary. All the results presented in this

study are obtained for the optimized values of the decision parameters, only the maintenance action period δ is fixed (cf. results Section 3.1). Let us denote $(\tau_0, \zeta_0, \text{and } \text{UCL}_0)$ (respectively, $(\tau_1, \zeta_1), \text{UCL}_2$ and T_3) the optimized decision parameters for Policy 0 (respectively, Policies 1, 2 and 3).

The curves in Fig. 8(a) are the respective representation of the optimized cost criterion calculated for the different policies. They are obtained when the scale parameter varies from 0 to 0.09. All the other system characteristics are fixed. From the analysis of the curves, we can conclude that:

- Policy 0 is always the policy which minimizes the cost criterion.
- When a tends to 0, Policy 0 tends to Policy 1. The gap between Policy 0 and 1 when $a = 0$ is due to UCL which permits to detect failure between the inspections on X .

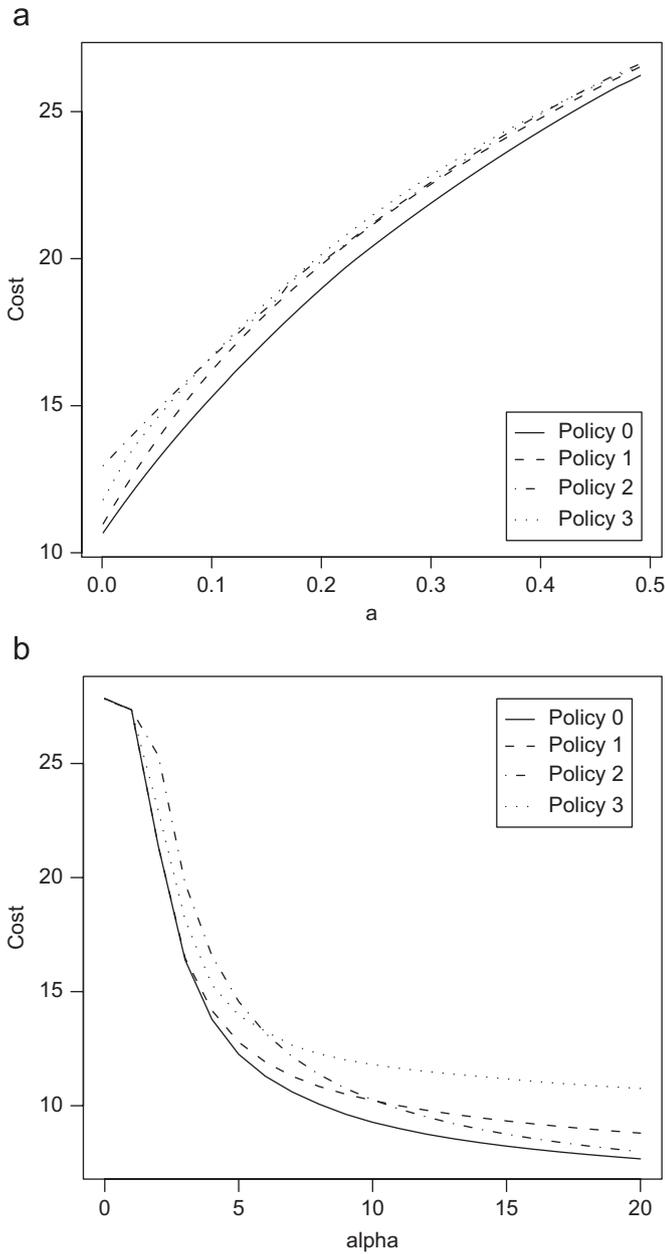


Fig. 8. Influence of the system characteristics. (a) Influence of the scale parameter. (b) Influence of the deterioration speed.

- When a increases Policy 0 tends to Policy 2. If the scale parameter would continue to increase, Policy 0 would tend to Policy 2. The probability of failure due to a shock becomes so big that only corrective replacements should be performed when a failure is detected by the drift in Y_k .

The curves in Fig. 8(b) represent, respectively, the optimized cost criterion calculated for the different policies when the deterioration speed varies from 1 to 10. For a low deterioration speed, the probability of failure is mainly due to an excessive deterioration level, and the conditional maintenance policy used for X_k allows to prevent failure. So Policy 0 tends to Policy 1 for a low deterioration speed. When the deterioration speed increases, the probability of failure due to a shock increases to, and Policy 0 tends to Policy 2.

The results are not presented here, but when the sensitivity to stress parameter and acceleration factor vary, Policy 1 is still the

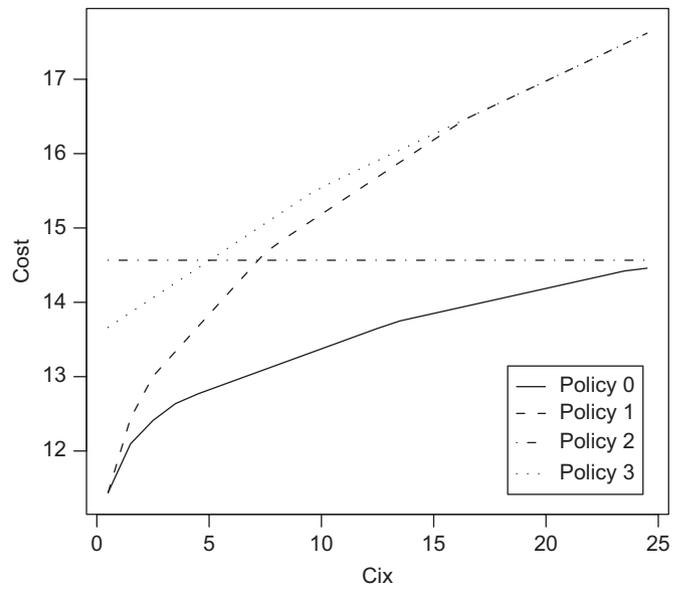


Fig. 9. Influence of the system characteristics: cost variation when c_{ix} increases.

policy which minimizes the cost criterion and tends to one of the other policy in some extreme configurations.

4.3.2. Performance of the proposed maintenance policy when maintenance costs vary

The economic performance of each policy is evaluated when the unit maintenance cost vary. The curves in Fig. 9 are the respective representation of the optimized cost criterion calculated for the different policies. They are obtained when the cost of the inspection on X varies from 0 to 12. All the other costs are fixed. Policy 0 is again the policy which minimizes the cost criterion. When the inspection cost c_{ix} tends to 0 the inspection period decreases and tends to 1. The failure is only detected during an inspection on X , thus Policy 0 tends to Policy 1. When the inspection cost c_{ix} increases, Policy 0 tends to Policy 2. In order to reduce the influence of the increase in the inspection cost, the optimal inspection period increases whereas, in the same time, the optimal preventive threshold decreases and the optimal upper limit control decreases to reduce the unavailability of a failed system. For a high c_{ix} , the inspection period should be so long that only corrective replacements should be performed when a failure is detected by the drift in Y_k .

The same analysis has been made when the other maintenance costs vary and the results are the same: the proposed policy is always the policy which minimizes the cost criterion or in particular configuration the proposed policy tends to a classical policy.

The observations made in this section show the adaptability of the proposed policy to very different system characteristics. The proposed maintenance policy is demonstrated to be more general than the three others which can be, however, emulated by the proposed structure and can be considered as “limit” cases. As a result, the proposed policy shows all its interest in non-extreme configurations, i.e. when the behavior of the system is less predictable and when none of the unitary costs is negligible: in these cases, it is necessary to combine the CBM policy and the SPC.

4.4. Sensitivity analysis

Collecting data in practical situation can be quite difficult in respect with the lack of data and the measurement variability.

Hence, the quality of the different estimations is directly influenced by the data collection. The objective of this section is to analyze the robustness of the proposed maintenance policy when the estimations of the system characteristics are biased.

4.4.1. Sensitivity analysis when the failure threshold (L) varies

In practical application, it is not a simple task to identify with precision the value of the failure threshold, some errors of measurement or estimation can be made.

Fig. 10 represents the influence of the error made on the failure threshold. Fig. 10(a) proposes to compare the average cost obtained with the optimized maintenance policy from the value $L = 2$ while an uncertainty on this parameter has been made (the full line) and the average optimized cost that we would have obtained if any uncertainty had been made (the dotted lines).

Fig. 10(b) illustrates the relative loss, expressed as a percentage, to this uncertainty. The proposed maintenance policy seems

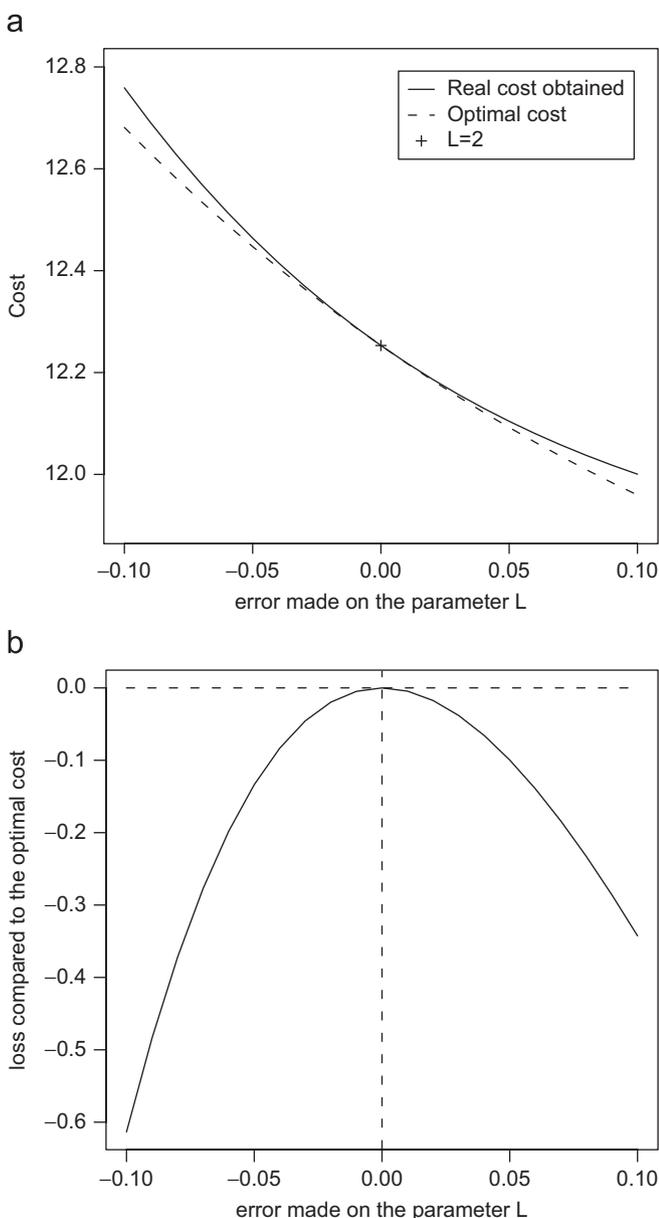


Fig. 10. Influence of the error made on the failure threshold: $c_c = 100$; $c_p = 30$; $c_{ix} = 2$; $c_{iz} = 0.5$; $c_u = 25$; $c_m = 4$; $\alpha = 5$; $a = 0.05$; $F_A = 0.96$; $b = 0.025$. (a) Comparison of the average cost obtained and the optimized cost. (b) Loss compared to the optimal cost.

Table 1

Impact of the error made on the system characteristics vector on the maintenance cost

Error made on the system characteristics vector	-10%	0	+10%
Cost with uncertainties	9.814	12.253	15.050
Optimal cost	9.729	12.253	14.885
Relative cost compared to the optimal cost	0.866%	0	1.096%

to be sufficiently robust with the uncertainty on the parameter L : less than 0.65% of relative loss is obtained in the worst case for an uncertainty of -10% on the parameter L .

4.4.2. Sensitivity analysis of other system characteristics with the “one-factor-at-a-time” method

The same studies as previously for the failure threshold have been made for the other system characteristics: the deterioration rate (α), the failure acceleration factor due to stress (F_A), the scale parameter (a), the sensitivity to stress (b), the stress threshold (λ), the nominal level of the stress (m), the standard deviation of the stress (σ) and the drift of the stress (μ). In all the cases, when an uncertainty is made on one parameter, we obtained less than 0.45% of relative loss in the worst case. By using the “one-factor-at-a-time” method, the robustness of the model is highlighted when only one factor at a time varies for the given system characteristics vector ($a = 0.05$, $b = 0.025$, $F_A = 0.96$, $\alpha = 5$, $m = 10$, $\mu = 0.22$, $\sigma = 1$, $\lambda = 10.5$, $L = 2$).

4.4.3. Sensitivity analysis of the system characteristics vector

Previously, the “one-factor-at-a-time” method has been used, when one parameter varies, others remain constant. But in a practical case, the parameters vary simultaneously and the effect of each input on the output is not necessarily linear. In order to move away from the hypothesis of the linearity and to evaluate the effect an uncertainty of the system characteristics vector, the worst cases are considered, i.e. each parameter is set to its value corresponding to a relative error of -10% or +10%. The results presented in Table 1 are obtained by using a Monte Carlo approach. Less than 1.1% of relative loss is obtained in the worst case.

Finally, given the results, the model seems to be robust to uncertainties made on the estimation of the system characteristics vector.

5. Conclusion

The main interests of this work are the two following points. Firstly, the construction of a failure model for continuously deteriorating systems subject to environmental influence which combines a degradation model well controlled, known and a model by shock for which uncertainties on the failure mechanism can be integrated. Secondly, the construction and the evaluation of maintenance policies for this model is proposed. We provide a simple and relevant maintenance policy which takes advantage of all the information available for the system state description. The combination of SPC and CBM provide an accurate decision-making tool for a system with several failure mechanisms. The associated stochastic model of the maintenance cost has been developed in order to assess the policy performance. The development of the cost model has been supported by the evaluation of the reliability of the maintained system. Numerical experiments have demonstrated the good performance of the proposed maintenance structure, namely its adaptability to different system characteristics, its lower cost when compared with classical maintenance

policies and its robustness when system characteristics are estimated with uncertainty. Even if the proposed structure for maintenance decision has shown interesting performance when compared with classical maintenance policies, a lot of research work remains to be done. It could be interesting to relax some assumptions in order to consider more complicated situations, such as dependencies between the deterioration process and the stress covariate. Moreover, in general the system is not subject to a single stress phenomenon, but to several stress sources, so, it could be interesting to integrate several stresses in our model. Furthermore, actually the stress is modeled by a discrete process but it could be interesting to model it with a continuous-time process.

Appendix A

A.1. Evaluation of the expected number of inspections

- $N_{iz}(S)$ is the number of inspections of Z_{t_k} . The expected number of inspections of Z_{t_k} is given by the expected number of false alarms and the expected number of true alarms.

$$\mathbb{E}(N_{iz}(S)) = \sum_{i=0}^{\infty} \left[\underbrace{\sum_{j=i\tau+1}^{(i+1)\tau-1} [\mathbb{P}(Y_j > UCL | \text{no failure}) R_m(j)]}_{\text{false alarms}} + \underbrace{\mathbb{P}(Y_j > UCL | \text{fail.}) (R_m(j-1) - R_m(j)) \prod_{w=j}^{i\tau-1} \mathbb{P}(Y_w < UCL | \text{failure})}_{\text{true alarms}} \right] \tag{A.1}$$

- $(N_{ix}(S))$ is the expected number of inspections of X_t before S , it is given by

$$\mathbb{E}(N_{ix}(S)) = \sum_{i=0}^{\infty} \left[\mathbb{1}_{\{i>0\}} R_m(i-1) \sum_{j=(i-1)\tau+1}^{i\tau-1} (R_m(j-1) - R_m(j)) \times \prod_{k=j}^{i\tau-1} \mathbb{P}(Y_k < UCL | \text{failure}) \right] \tag{A.2}$$

A.2. Evaluation of the expected number of maintenance actions

- $\mathbb{E}(N_m(S))$ the expected number of routine maintenance

$$\mathbb{E}(N_m(S)) = \sum_{i=1}^{\infty} \left[\prod_{k=j}^{i\tau-1} \mathbb{1}_{\{i=\delta_{\lfloor \frac{i}{\tau} \rfloor}\}} (R_m(i-1) + \sum_{j=(i-1)\tau+1}^{i\tau-1} (R_m(j-1) - R_m(j)) \times \prod_{k=j}^{i\tau-1} \mathbb{P}(Y_k < UCL | \text{failure}) \right] \tag{A.3}$$

- $\mathbb{E}(N_c(S))$ is the expected number of a corrective replacements, it is given by

$$\mathbb{E}(N_c(S)) = \sum_{i=0}^{\infty} \left[\mathbb{1}_{\{i>0\}} (R_m(i\tau-1) - R_m(i\tau)) + \sum_{j=(i-1)\tau+1}^{(i\tau-1)} (R_m(j-1) - R_m(j)) \times \prod_{k=j}^{i\tau-1} \mathbb{P}(Y_k < UCL | \text{failure}) \right]$$

$$+ \sum_{j=i\tau+1}^{(i+1)\tau-1} \mathbb{P}(Y_j > UCL | \text{failure}) * \left[(R_m(j-1) - R_m(j)) + \mathbb{1}_{\{j>i\tau+1\}} \sum_{k=i\tau+1}^{j-1} (R_m(k-1) - R_m(k)) \times \prod_{w=k}^{j-1} \mathbb{P}(Y_w < UCL | \text{failure}) \right] \tag{A.4}$$

A.3. Evaluation of the expected cumulative unavailability duration

- $\mathbb{E}(D_u(S))$, the expected value of the cumulative unavailability duration before S is given by

$$\mathbb{E}(D_u(S)) = \sum_{i=0}^{\infty} \left[\sum_{j=i\tau+1}^{(i+1)\tau-1} (R_m(j-1) - R_m(j)) \times \prod_{k=j}^{(i+1)\tau-1} \mathbb{P}(Y_k < UCL | \text{failure}) \right] \tag{A.5}$$

References

- [1] Valdez-Flores C, Feldman RM. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Res Logistics* 1989; 36:419–46.
- [2] Wang H. A survey of maintenance policies of deteriorating systems. *Eur J Oper Res* 2002;139:469–89.
- [3] Gertsbakh I. *Reliability theory with applications to preventive maintenance*. Berlin: Springer; 2000.
- [4] Scarf PA. On the application of mathematical models in maintenance. *Eur J Oper Res* 1997;99(2):493–506.
- [5] van Noortwijk JM. A survey of the application of gamma processes in maintenance. *Reliab Eng Syst Safety*; 2007. doi:10.1016/j.ress.2007.03.019.
- [6] Hong H. Inspection and maintenance planning of pipeline under external corrosion considering generation of new defects. *Struct Safety* 1999;21: 203–22.
- [7] Jeang A. Tool replacement policy for probabilistic tool life and random wear process. *Qual Reliab Eng Int* 1999;15:205–12.
- [8] van Noortwijk JM, Kok M, Cooke R. Optimal maintenance decisions for the sea-bed protection of the eastern-scheldt barrier. *Eur J Oper Res* 1997;82(2): 25–56.
- [9] Deloux E, Castanier B, Bérenguer C. Combining statistical process control and condition-based maintenance for gradually deteriorating systems subject to stress. In: Aven, Vinnem, editors. *Risk, reliability and societal safety, proceedings of ESREL 2007—European safety and reliability conference 2007*, Stavanger, Norway, 25–27 June 2007. London: Taylor & Francis; 2007. p. 265–72.
- [10] Meeker WQ, Escobar LA. *Statistical methods for reliability data*. New York: Wiley; 1998.
- [11] Singpurwalla ND. *Reliability and risk A Bayesian Perspective*. New York: Wiley; 2006.
- [12] Bagdonavicius V, Nikulin M. Estimation in degradation models with explanatory variables. *Lifetime Data Anal* 2000;7:85–103.
- [13] A. Lehmann, Joint modeling of degradation and failure time data, In: *Degradation, damage, fatigue and accelerated life models in reliability testing*, ALT2006, Angers, France; 2006. p. 26–32.
- [14] Singpurwalla ND. Survival in dynamic environments. *Stat Sci* 1995;10(1): 86–103.
- [15] Yashin AI, Manton KG. Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. *Stast Sci* 1997;12(1):20–34.
- [16] Elsayed EA, Chan CK. Estimation of thin-oxide reliability using proportional hazards models. *IEEE Trans Reliab* 1990;39(3):329–35.
- [17] van Noortwijk JM, Kallen MJ, Pandey MD. Gamma processes for time-dependent reliability of structures, In: Kolowrocki K. editor, *Advances in safety and reliability, proceedings of ESREL 2005—European safety and reliability conference 2005*, Tri City (Gdynia-Sopot-Gdansk), Poland, 27–30 June 2005; 2005.
- [18] Wang W. A two-stage prognosis model in condition based maintenance. *Eur J Oper Res* 2007;182(3):1177–87.
- [19] Zhan Y, Mechefske CK. Robust detection of gearbox deterioration sing compromised autoregressive modeling and Kolmogorov–Smirnov test statistic. Part II: experiment and application. *Eur J Oper Res* 2007;21:1983–2011.
- [20] Ben-Daya M, Rahim MA. Effect of maintenance on the economic design of x-control chart. *Eur J Oper Res* 2000;120:131–43.
- [21] Cassady CR, Bowden RO, Liew L, Pohl EA. Combining preventive maintenance and statistical process control: a preliminary investigation. *IIE Trans* 2000;32: 471–8.

- [22] Yeung TG, Cassady CR, Schneider K. Simultaneous optimization of \bar{x} -bar control chart and age-based preventive maintenance policies under an economic objective. Working paper; 2006.
- [23] Linderman K, McKone-Sweet KE, Anderson JC. An integrated systems approach to process control and maintenance. *Eur J Oper Res* 2005;164: 324–40.
- [24] Castanier B, Bérenguer C, Grall A. A sequential condition-based repair/replacement policy with non-periodic inspections for a system subject to continuous wear. *Appl Stochastic Models Business Industry* 2003;19(4):327–47.
- [25] Singpurwalla ND, Wilson S. Failure models indexed by two scales. *Adv Appl Probab* 1998;30:1058–72.
- [26] Yang SC. A bivariate renewal process and its application in maintenance policies, Ph.D. thesis, Faculty of Virginia Polytechnic Institute and State University; 1999.
- [27] Ross SM. *Stochastic processes*. 2nd ed. New York: Wiley; 1996.
- [28] Park SK. Optimal wear-limit replacement with wear-dependent failures. *IEEE Trans Reliab* 1988;27:293–4.
- [29] Zuckerman D. Optimal replacement policy for the case where the damage process is a one-side Lévy process. *Stochastic Processes Appl* 1978;7:141–51.
- [30] W Burr I. *Statistical quality control methods*. New York: Marcel Dekker; 1976.
- [31] Asmussen S. *Applied probability and queues*. In: *Wiley series in probability and mathematical statistics*. New York: Wiley; 1987.



A memetic algorithm for the multi-compartment vehicle routing problem with stochastic demands

Jorge E. Mendoza^{a,b}, Bruno Castanier^a, Christelle Guéret^{a,*}, Andrés L. Medaglia^b, Nubia Velasco^b

^aÉquipe Systèmes Logistiques et de Production, IRCCyN (UMR CNRS 6597), École des Mines de Nantes, B.P. 20722, F-44307 Nantes Cedex 3, France

^bCentro para la Optimización y Probabilidad Aplicada (COPA), Industrial Engineering Department, Universidad de los Andes, Cr 1 Este No. 19A-10 Bogotá, Colombia

ARTICLE INFO

Available online 25 June 2009

Keywords:

Multi-compartment vehicle routing problem
Stochastic demands
Evolutionary algorithms
Memetic algorithms

ABSTRACT

The multi-compartment vehicle routing problem (MC-VRP) consists of designing transportation routes to satisfy the demands of a set of customers for several products that, because of incompatibility constraints, must be loaded in independent vehicle compartments. Despite its wide practical applicability the MC-VRP has not received much attention in the literature, and the few existing methods assume perfect knowledge of the customer demands, regardless of their stochastic nature. This paper extends the MC-VRP by introducing uncertainty on what it is known as the MC-VRP with stochastic demands (MC-VRPSD). The MC-VRPSD is modeled as a stochastic program with recourse and solved by means of a memetic algorithm. The proposed memetic algorithm couples genetic operators and local search procedures proven to be effective on deterministic routing problems with a novel individual evaluation and reparation strategy that accounts for the stochastic nature of the problem. The algorithm was tested on instances of up to 484 customers, and its results were compared to those obtained by a savings-based heuristic and a memetic algorithm (MA/SCS) for the MC-VRP that uses a spare capacity strategy to handle demand fluctuations. In addition to effectively solve the MC-VRPSD, the proposed MA/SCS also improved 14 best known solutions in a 40-problem testbed for the MC-VRP.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

The multi-compartment vehicle routing problem (MC-VRP) consists of designing a set of minimal-cost routes to serve the demands for multiple products of a set of customers. The distinguishing feature of the problem is that products are incompatible and must be transported in independent vehicle compartments. The MC-VRP naturally arises in several practical situations. For instance, dairies often use vehicles with multiple compartments to collect milk of different types (e.g., from cows and goats) and qualities (e.g., different suckling dates), petroleum companies deliver different types of fuel to outlet retailers using multi-compartment tankers, public utilities use trucks with compartments to perform selective waste collection, and food companies distribute in compartmentalized vehicles groceries that require different levels of refrigeration.

Despite its wide practical applicability, the MC-VRP has received little attention in the literature (see [12] for a recent survey on the VRP). Early work on routing problems with multiple compartments

addressed an MC-VRP variant where all compartments of a vehicle can be used to transport any product [9,8]. To solve the problem, these authors split the resolution into two steps. First, they allocate orders to vehicles and compartments; then, they service the set of customers by solving a sequence of traveling salesman problems (TSPs). More recently, researchers [3,1,33] have proposed solution methods that tackle the product/compartment allocation and routing issues simultaneously. Van der Bruggen et al. [41] solved a real-world MC-VRP as part of the distribution system design of a petroleum company. Their construction heuristic (CH) forces the insertion of customers with large demands or located far from the depot in early iterations, thus minimizing both distance and the number of used vehicles. Lately, El Fallahi et al. [14] proposed a savings-based heuristic, a tabu search (TS), and a memetic algorithm (MA) for a MC-VRP variant in which each compartment is dedicated solely to one product. Under this scenario, no allocation decisions are needed since the compartment capacity for each product is given. The authors tested two different service policies: in the first one all products must be delivered to the customer by the same route; in the second one each product can be delivered by a different route.

The body of work on the MC-VRP has been restricted to the case in which demands are known with certainty. In practice, however, customer demands are frequently stochastic, meaning that their exact value is not known at the moment when routes are planned.

* Corresponding author.

E-mail addresses: je.mendoza45@egresados.uniandes.edu.co (J.E. Mendoza), bruno.castanier@emn.fr (B. Castanier), christelle.gueret@emn.fr (C. Guéret), amedagli@uniandes.edu.co (A.L. Medaglia), nvelasco@uniandes.edu.co (N. Velasco).

As pointed out by Christiansen and Lysgaard [10], neglecting the stochastic nature of the demand may lead to a substantial cost increase, thus there is a strong incentive to extend the MC-VRP to handle some type of uncertainty. This paper incorporates uncertainty into the demands of the MC-VRP leading to the MC-VRP with stochastic demands, henceforth called the MC-VRPSD. In the MC-VRPSD, the demand of each customer for each product is modeled by a random variable with a known probability distribution.

Since the MC-VRPSD is a variant of the stochastic VRP, different solution frameworks can be applied to solve the problem [36]. A taxonomy of these frameworks classifies them into *dynamic* or *static* [35]. Routing decisions in dynamic approaches are made in multiple stages and based on demand realizations (at former stages). Depending on the remaining vehicle capacity and the set of unvisited customers, routes are *re-optimized* at each stage. Some examples of this category are the Markov decision process (MDP) approach of Secomandi and Margot [35] and the neuro-dynamic programming approach of Secomandi [34]. On the other hand, routing decisions in static approaches, once made remain unchanged regardless of the demand realizations. Some of the approaches that fall into this category include chance constraint programming [36], robust optimization [39], and stochastic programming with recourse [10]. From a computational point of view static strategies are preferred over their dynamic counterparts, since the problem, known to be computationally intractable, is solved only once [21]. Static strategies are also specially useful when a stable solution is desired [6] or when re-optimization during route execution is impossible due to the lack of information.

Among the static approaches, the two-stage stochastic programming (or stochastic programming with recourse) has been widely used for the VRPSD, a particular case of the MC-VRPSD in which each customer has a random demand for a single product. The two-stage stochastic programming framework solves the problem in two phases. In the first phase a set of routes is planned, while in the second phase the planned routes are executed. However, because of the demand uncertainty, while servicing a given customer the remaining capacity of the vehicle may not be enough to satisfy the whole customer demand; in such a case, a *route failure* is said to occur. In case of a failure, a problem-dependent predefined corrective action, known as *recourse*, is taken to recover the solution feasibility. In general, the route failures generate an extra cost whose value is only known after the second phase. Thus, the objective of the problem is to design during the first phase a set of routes that minimizes the sum of the cost of the planned routes and the expected cost of the route failures.

Exact methods based on two-stage stochastic programming formulations of the VRPSD include that of Laporte et al. [21] who proposed a branch-and-bound method based on the integer L-shaped algorithm. Similarly, Christiansen and Lysgaard [10] proposed a column generation approach that proved to be competitive in tight instances of the problem. More recently, Rei et al. [32] tackled the single VRPSD (SVRPSD), a variant where only one route is to be designed. Their method consists of using local branching [15] to generate optimality cuts on an integer L-shaped algorithm. Although successful, these approaches are limited to solve instances of up to 100 customer nodes. Consequently, industrial-scale problems are out of reach and should be targeted using heuristics [19].

Early construction heuristics for the VRPSD include the savings-based algorithm of Stewart and Golden [38] and that of Dror and Trudeau [13]. More recently, Yang et al. [42] tackled a variant of the problem with *restocking*. The principle behind the restocking policy is to decide upon the remaining capacity after completing the service of each customer and perform preventive trips to the depot to avoid future route failures. These authors introduced two construction heuristics based on insertion procedures and a dynamic

programming algorithm for tour splitting. Gendreau et al. [20] proposed *tabustoch*, a TS algorithm designed to tackle a variant of the problem where in addition to the demands customers are also stochastic, that is, they are present (or not) with a given probability. These authors tested *tabustoch* on a large set of instances and reported average deviation gaps of 0.38% with respect to the optimal solutions achieved by a branch-and-bound algorithm [18]. A different solution strategy was introduced by Fleury et al. [17] who addressed the stochastic capacitated arc routing problem (SCARP), a related problem where the demands are spread along the edges rather than concentrated in the nodes. Their approach consists of solving the deterministic counterpart of the problem (the CARP) using the expected demands and preserving some capacity in the vehicles by means of an MA. The intuition behind their strategy is to avoid route failures by using some extra capacity in case of high demand realizations; nonetheless, if route failures occur during the execution, a recourse action takes place to recover the feasibility as in the two-stage stochastic programming approach. Later, Fleury et al. [16] proposed what probably is the first evolutionary approach specially tailored for routing problems with stochastic demands, namely, a memetic algorithm for the SCARP. Besides the expected cost of the routes, the authors also tested an objective function that minimizes both the expected value of the cost and its variability. Lately, Bianchi et al. [7] implemented several algorithms for the SVRPSD with restocking, including simulated annealing (SA), iterated local search (ILS), ant colony optimization (ACO), evolutionary algorithms (EA), and TS. They reported encouraging results against their own implementation of Bertsimas' cyclic heuristic [6] and two other CHs adapted to the problem. In sum, metaheuristics have shown to be an effective approach to tackle vehicle routing problems with uncertain demands and look promising on new variants such as the MC-VRPSD.

In contrast to the VRPSD, research to the MC-VRPSD is scarce. Tatarakis and Minis [40] tackled a single-route variant of the problem in which the sequence of customers (route) is fixed beforehand. The problem consists of selecting the optimal restocking points along the route. The authors propose a set of dynamic programming algorithms and solve to optimality problems of up to 15 customers. Mendoza et al. [24] proposed a set of construction heuristics comprising stochastic versions of the nearest neighbor [4], best insertion [4], and Clarke and Wright [11] heuristics, extended to the multi-compartment case. The latter proved to be the most competitive according to extensive computational experiments conducted on different types of instances. To the best of our knowledge, these are the few attempts made to solve the MC-VRPSD.

To expand the set of tools available for the MC-VRPSD this paper proposes a new memetic algorithm based on a stochastic programming with recourse formulation for the case in which each compartment is dedicated to one product. The algorithm couples components proven to be effective on classical vehicle routing problems with a mechanism tailored to manage the stochastic nature of demands in the complex multi-product scenario. The objective is to provide a solution approach, alternative to the existing construction heuristics, to tackle large-sized instances currently out of reach for exact methods.

The remainder of the paper is organized as follows. Section 2 formally states the problem and introduces the corresponding notation. Section 3 outlines the proposed MA. Section 4 presents the computational experiments conducted on a set of randomly generated realistic instances and a set of problems adapted from the MC-VRP literature. The results are compared to those obtained by a construction heuristic and a deterministic approach based on that proposed by Fleury et al. [17] for the SCARP. Finally, Section 5 concludes the paper and outlines future research perspectives.

2. Problem description

Formally, the MC-VRPSD is defined on a complete and undirected graph $G = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{0, \dots, n\}$ is the vertex set and \mathcal{E} the edge set. Vertices $v = 1, \dots, n$ represent the customers and vertex $v = 0$ represents the depot. A distance d_e is associated with edge $e = (u, v) = (v, u) \in \mathcal{E}$ and it represents the travel cost between vertices u and v . There exists a set $\mathcal{P} = \{1, \dots, p, \dots, m\}$ of products that must be transported in independent compartments of fixed capacity Q_p . All vehicles are identical and the fleet size unlimited. For product p customer v has an independent random demand $\xi_{v,p}$ following a known probability distribution with mean $\mu_{v,p}$ and standard deviation $\sigma_{v,p}$. The actual values of the demands (realizations) are nonnegative and less than the capacity of the corresponding compartment Q_p , yet only known upon the vehicle's arrival at the customer location. Each customer must be visited only once by exactly one vehicle (route) and the total length of each route cannot exceed a maximum distance L . Henceforth, without loss of generality the discussion is restricted to the case of collection routes, nonetheless the case of delivery routes is equivalent.

The MC-VRPSD formulation as a two-stage stochastic programming model follows. In the first stage, a set \mathcal{R} of *a priori* or *planned* routes is designed. Each route $r \in \mathcal{R}$ is a sequence of vertices $r = (0, v_1, \dots, v_i, \dots, v_{n_r}, 0)$, where $v_i \in \mathcal{V} \setminus \{0\}$ and n_r represents the number of customers serviced by the route. Depending on the context, we refer to the route $r \in \mathcal{R}$ also as an ordered set of arcs $r = \{(v_1, v_2), \dots, (v_{i-1}, v_i), \dots, (v_{n_r-1}, v_{n_r})\}$. In the second stage, each planned route is executed until a route failure occurs, that is, whenever the capacity of at least one compartment is exceeded. Upon failure, the compartment is loaded up to its capacity and the recourse action takes place. The recourse action is defined as a return trip to the depot to unload all the compartments, followed by a trip back to the customer location to complete the service. After service completion, the route is resumed from that point on as originally planned. It is assumed that (i) there is no practical way to keep track of the capacity available on each compartment, thus preventive trips to the depot (restocking) are not allowed and (ii) there is no communication between vehicles, thus cooperation between routes is impossible. The actual solution to the problem (or second-stage solution) is then the true set of routes traveled by the vehicles (planned routes plus recursive trips). Since the locations of the route failures depend on the demand realizations $\vec{\xi}$, the cost of the second-stage solution $C(\mathcal{R})$ is a random variable, with mean $\mu_{C(\mathcal{R})}$ and standard deviation $\sigma_{C(\mathcal{R})}$, given by

$$C(\mathcal{R}) = \sum_{r \in \mathcal{R}} C_r \quad (1)$$

$$= \sum_{r \in \mathcal{R}} l_r + \sum_{r \in \mathcal{R}} G_r(\vec{\xi}) \quad (2)$$

where $C_r (=l_r + G_r(\vec{\xi}))$ is the total travel cost, l_r denotes the planned length (planned cost), and $G_r(\vec{\xi})$ the length of the returning trips to the depot caused by route failures (cost of recourse) for each route $r \in \mathcal{R}$. Then, the problem is to determine in the first stage a set of planned routes \mathcal{R} that minimizes the expected cost of the transportation plan

$$E[C(\mathcal{R})] = \sum_{r \in \mathcal{R}} l_r + \sum_{r \in \mathcal{R}} E[G_r(\vec{\xi})] \quad (3)$$

Note that C_r , the total travel distance of route r , is a random variable which value is only known when the vehicle returns to the depot after completing the route. Thus, we model the distance constraint as

$$l_r + E[G_r(\vec{\xi})] \leq L \quad \forall r \in \mathcal{R} \quad (4)$$

Furthermore, the planned cost of route r (represented by l_r) is the sum of the distance covered by the arcs traversed by the route, given by $l_r = \sum_{(u,v) \in r} d_{(u,v)}$. On the other hand, the expected cost of recourse $E[G_r(\vec{\xi})]$ depends on two elements: the failure probability while servicing each customer $v_i \in r$, denoted by $Pr(v_i)$; and the distance between customer v_i and the depot. The expected cost of recourse of route r can be calculated as follows:

$$E[G_r(\vec{\xi})] = \sum_{i=1}^{n_r} 2 \times d_{v_i,0} \times Pr(v_i) \quad (5)$$

In the multi-compartment scenario, the failure probability while servicing a given customer v_i , depends on the location (customer) where the last failure occurred and the demands collected for each product since that last failure. The value of $Pr(v_i)$ is given by

$$Pr(v_i) = \sum_{j=0}^{i-1} \left[\prod_{p \in \mathcal{P}} Pr \left(\xi'_{v_j,p} + \sum_{k=j+1}^{i-1} \xi_{v_k,p} \leq Q_p \right) - \prod_{p \in \mathcal{P}} Pr \left(\xi'_{v_j,p} + \sum_{k=j+1}^i \xi_{v_k,p} \leq Q_p \right) \right] \times Pr(v_j) \quad (6)$$

where v_j is the customer being serviced when the last failure occurred and $\xi'_{v_j,p}$ is customer v_j 's unserved demand for product p due to the failure. Consequently, the first product denotes the probability that the collected demand for each product after the failure at customer v_j and up to customer v_{i-1} does not exceed the capacity of the corresponding compartment Q_p . Similarly, the second product denotes the probability that the capacities are not exceeded after servicing customer v_i . The exact calculation of $Pr(v_i)$ requires the evaluation of the probability functions of all unserved demands, which involves complex numerical approximations that are computationally expensive to be embedded within an optimization algorithm. Thus, the proposed MA uses an approximation based on the *take-all policy* (TAP) derived in [24]. The TAP approximation assumes that even if a route failure occurs at customer v_j the whole demand for all products is serviced before performing the return trip (i.e., $\xi'_{v_j,p} = 0 \forall p \in \mathcal{P}$). Nonetheless, after unloading the compartments at the depot, the vehicle must return to the customer being served when the failure occurred. Using the TAP approximation, $Pr(v_i)$ is given by the more convenient expression:

$$Pr(v_i) = \sum_{j=0}^{i-1} \left[\prod_{p=1}^m F(v_{j+1}, v_{i-1}, p) - \prod_{p=1}^m F(v_{j+1}, v_i, p) \right] \times Pr(v_j) \quad (7)$$

where $F(v_{j+1}, v_{i-1}, p)$ is the cumulative probability that the demand for product p collected between nodes in position $j+1$ and $i-1$ in route r does not exceed the capacity of the compartment Q_p .

Computational experiments conducted in [24] showed that if the capacity of the compartments is sufficiently large compared to the largest expected demand for the corresponding product, which is the case in many real-world applications, the true cost of recourse in (5) can be estimated within an average error of 3.00% by the TAP approximation.

3. Memetic algorithm

Memetic algorithms are genetic algorithms (GAs) that use local search procedures to intensify the search [28]. The proposed MA borrows the general structure, individual representation, and crossover operator from the two evolutionary algorithms for the distance-constrained VRP proposed in [27]. Starting from an initial population $\mathcal{P}(0)$ comprised of P individuals, the algorithm runs for T generations. At every generation t , crossover, mutation, and local search operators are independently applied with probabilities p_c , p_m , and

p_{ls} , to each individual in the current population $\mathcal{P}(t)$. During the application of the operators the distance constraint (4) is relaxed, thus a reparation strategy is run simultaneously with the evaluation after the generation of each individual. The individuals generated by the operators join the current population to form an expanded population $\mathcal{E}(t)$, from which the best P individuals are selected to become part of the next generation, namely $\mathcal{P}(t+1)$. Clones in the objective space, that is, individuals sharing exactly the same value of the objective function in (3), are completely forbidden in the population to foster diversification. Algorithm 1 presents the general structure of the proposed MA.

Algorithm 1. General structure of the MA

```

1:  $t \leftarrow 0$ 
2: initialize  $\mathcal{P}(t)$ 
3: while  $t < T$  do
4:    $\mathcal{E}(t) \leftarrow \mathcal{P}(t)$ 
5:   randomly select  $\lceil p_c \times P \rceil$  parents from  $\mathcal{P}(t)$  and generate  $\mathcal{C}(t)$ 
6:   while  $\mathcal{C}(t) \neq \emptyset$  do
7:     cross two parents from  $\mathcal{C}(t)$  and generate child  $c$ 
8:     evaluate and repair  $c$ 
9:     if  $c \in \mathcal{E}(t)$  then
10:       repeat crossover
11:     else
12:       add  $c$  to  $\mathcal{E}(t)$  and eliminate the two parents from  $\mathcal{C}(t)$ 
13:     end if
14:   end while
15: for each individual  $s \in \mathcal{P}(t)$  do
16:   mutate  $s$  with probability  $p_m$  and generate  $s'$ 
17:   evaluate and repair  $s'$ 
18:   if  $s' \in \mathcal{E}(t)$  then
19:     repeat mutation
20:   else
21:     add  $s'$  to  $\mathcal{E}(t)$ 
22:   end if
23:   apply to  $s$  with probability  $p_{ls}$  the local search procedure and generate  $s''$ 
24:   evaluate and repair  $s''$ 
25:   if  $s''$  improves  $s$  and  $s'' \notin \mathcal{E}(t)$  then
26:     add  $s''$  to  $\mathcal{E}(t)$ 
27:   end if
28: end for
29: select  $\mathcal{P}(t+1)$  from  $\mathcal{E}(t)$ 
30:  $t \leftarrow t + 1$ 
31: end while

```

3.1. Solution encoding and initial population

The proposed MA encodes the MC-VRPSD solutions into a multipermutation genotype known as the genetic vehicle representation (GVR) [29]. Specifically, each permutation contains an ordered set of customers representing a route. The GVR representation allows the straightforward application of the selected crossover, mutation and local search operators designed to work on independent routes.

To accelerate algorithmic convergence, the initial population is generated based on a stochastic best insertion (SBI) heuristic [26]. Starting from a route servicing a single randomly selected customer, the SBI heuristic evaluates at every iteration the insertion cost of each customer not included in the route. The customer is inserted into a position that is feasible in terms of the distance constraint at the lowest insertion cost. If no feasible insertions are found, a new route is created and the evaluation process repeated. The heuristic stops when all customers are allocated to a route.

In contrast to deterministic VRPs, the insertion cost calculation in the MC-VRPSD is not straightforward. The insertion cost of customer

v into position i of a given route r is denoted by $I_v(i, r)$ and given by

$$I_v(i, r) = c_v(i, r) + g_v(i, r) \tag{8}$$

where $c_v(i, r)$ is the increment on the planned distance of the route and $g_v(i, r)$ is the impact (positive or negative) on its expected cost of recourse. Let u and w be the customers currently in positions $i-1$ and i in route r . The value of $c_v(i, r)$, is computed in constant time by:

$$c_v(i, r) = d_{(u,v)} + d_{(v,w)} - d_{(u,w)} \tag{9}$$

On the other hand, the computation of $g_v(i, r)$ is far more complex because inserting a customer into the i -th position of a route generates an impact on the failure probability of all subsequent customers as stated in (6). The insertion recourse cost is given by

$$g_v(i, r) = E[G_r(\vec{\xi}')] - E[G_r(\vec{\xi})] \tag{10}$$

where r' is the route resulting from the insertion. Hence, the estimation of the cost of recourse $E[G_r(\vec{\xi}')] requires the recalculation of the last $n_r - i - 1$ terms in (5). In the worst case, when $i = 1$ and route r' services the n customers, the computation of $E[G_r(\vec{\xi}')] is done in $O(mn^2)$ using the TAP approximation.$$

To avoid the expensive computation of unpromising insertions, the SBI heuristic selects the best insertion in two phases. In the first phase, only $c_v(i, r)$ is evaluated for each insertion. The best K insertions, those with the lowest value of $c_v(i, r)$, advance to a second phase where their evaluation is completed by adding $g_v(i, r)$. Then, the best insertion point is selected among the K updated candidates from phase two. Previous work showed that low values of K led not only to faster executions but also to lower cost solutions [26]. Therefore, after some preliminary experimentation, a fixed value of $K = 10$ was chosen for the SBI heuristic embedded in the proposed MA.

Limited by n customers, the random initialization of the SBI heuristic is able to produce up to n distinct solutions. Therefore, if the population size is large (i.e., close to or larger than n), clone individuals are likely to be generated during the initialization of the algorithm. To guarantee a fully diversified initial population, clones are eliminated and replaced by randomly generated solutions.

3.2. Crossover operator

The crossover operator is based on the GVR crossover proposed by Pereira et al. [29], in which a child inherits all the traits (routes) from one parent (recipient) and a small portion of the genetic material (subroute) from the other parent (donor). A subroute is randomly selected from the donor and inserted into the position having the lowest insertion cost. To speed up the procedure, the insertion cost is calculated taking into account only the planned cost of the route. Hence, the cost of inserting a subroute $r' = (v, \dots, v')$, with initial and terminal vertices v and v' , into position i of a route r is denoted by $c_r(i, r)$ and calculated as follows:

$$c_r(i, r) = d_{(u,v)} + d_{(v',w)} - d_{(u,w)} \tag{11}$$

where u and w are the customers currently in positions $i-1$ and i in route r . After insertion, duplicate customers are eliminated from the child, preserving those in the inserting subroute. Fig. 1 illustrates the crossover procedure in the proposed MA.

3.3. Mutation operator

The mutation operator, known as *inversion mutation* [29], reverses the visit order of all vertices in a randomly selected subroute. The inversion mutation is a random 2-opt move that diversifies the population along two fronts. First, the route structure is altered by the

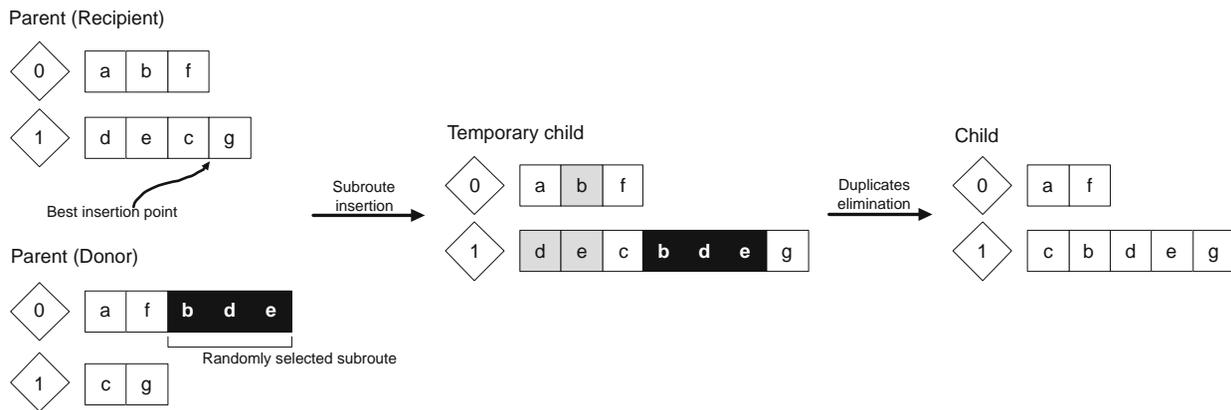


Fig. 1. Crossover procedure applied to the GVR genotype in the proposed MA.

exchange of two arcs; and second, the traveling direction of a subsequence of customers is changed. The latter is specially important since in the MC-VRPSD the cost of recourse is not symmetric. Note from (6) that the failure probability at customer v_i depends on the demands collected at the preceding customers. Because the demands collected from the depot to a given customer v are different from those that are to be collected from customer v to the depot, the failure probability at customer v differs upon the traveling direction, leading to different costs of recourse.

Fig. 2 illustrates the inversion mutation procedure and the impact of the traveling direction in the cost of recourse. For simplicity, the example assumes that the problem parameters are such that independence of the route sequence, the failure probability at a given customer increases by 0.20 with respect to the failure probability of the preceding customer. Note how reversing the traveling direction of the subsequence (e, c, g) impacts the failure probability at customers e and g and the total cost of recourse of route 1.

3.4. Local search operator

Often, classical GAs are not aggressive enough for combinatorial optimization problems [37] and should be enhanced with local search procedures to become competitive in problems like the VRP [30]. In the specific case of VRPs with stochastic demands, Bianchi et al. [7] showed how embedding local search procedures for deterministic routing problems into a GA for the SVRPSD improved the performance of the algorithm. There are two plausible explanations for this behavior. First, procedures for deterministic VRPs work only on the planned cost of the routes. By avoiding the expensive calculation of the cost of recourse during the local search phase, the algorithm is allowed to explore a wider range of solutions in equivalent execution times [7]. Second, because the planned cost of a solution normally dominates the cost of recourse [31], a good solution for the VRPSD is likely to be made up of good planned routes. Therefore, local search procedures that find better planned routes favor the overall performance of the algorithm. Based on these principles, the local search procedure embedded in the MA couples the following two neighborhood moves classically used in deterministic VRPs: *Relocate* and *2-opt*. The former extracts a customer from its current position and inserts it into a different route; while the latter replaces two non-adjacent edges on a given route.

Each iteration of the local search procedure works as follows. First, a feasible individual is selected from the current population $\mathcal{P}(t)$. Then, the relocate move extracts one customer from its current route and tries to insert it into a different route. Once an improving move is found, it is immediately executed. Then, 2-opt moves are tried in the

route where the customer was inserted. As soon as either a successful 2-opt move is found or there are no more two-arc exchanges, a new local search iteration begins. The whole procedure is repeated until no more improvements are found. Fig. 3 illustrates the local search procedure applied to the GVR genotype.

3.5. Reparation and fitness evaluation

In the MA, the reparation and evaluation of individuals is performed by *stochastic split* (*s-split*) a proposed extension of the chromosome partition mechanism by Prins [30]. *S-split* requires the GVR genotype to be transformed into a chromosome without route delimiters, that is, a giant tour with a single permutation of customers. From the chromosome, an auxiliary graph G' is built and used to find the optimal partition of the permutation into a set of feasible routes. The directed graph $G' = (\mathcal{V}', \mathcal{A})$ is composed of the vertex set $\mathcal{V}' = \{0, v_1, \dots, v_i, \dots, v_n\}$ and the arc set \mathcal{A} . Vertices $v_1, \dots, v_n \in \mathcal{V}' \setminus \{0\}$ and 0 is an auxiliary vertex. Arc $(v_i, v_{i+n_r}) \in \mathcal{A}$ represents a feasible route r with expected cost $E[C_r] = l_r + E[G_r(\vec{\xi})]$ starting and ending at the depot and traversing the sequence of customers from v_{i+1} to v_{i+n_r} . The *s-split* procedure consists of finding the set of arcs (i.e., routes) along the shortest path connecting 0 and v_n in G' . Fig. 4 illustrates the *s-split* procedure, where a GVR individual containing an infeasible route ($L = 70$) is transformed into a single chromosome, followed by a partition into two feasible routes of minimum expected cost. As in the example used to illustrate the mutation operator, it is assumed that the failure probability at a given customer increases by 0.20 with respect to the failure probability of the preceding customer.

To partition the chromosome, the MA uses Algorithm 2. The procedure, based on Bellman's algorithm for directed acyclic graphs, splits the chromosome without generating G' explicitly. For each vertex $v_i \in \mathcal{V}' \setminus \{0\}$ the algorithm maintains two labels, namely, Z_i and B_i . Label Z_i holds the expected cost of the shortest path from 0 to v_i and label B_i holds the predecessor of v_i in the shortest path. The loops in steps 3 and 7 generate all arcs (feasible routes) in \mathcal{A} . The expected cost of each route ($E[C_r]$) is evaluated (steps 8–21), and its value is used to update labels Z_i and B_i (steps 22–28). The value of Z_n is the fitness function of the individual ($E[C(\mathcal{P})]$) and the routes can be recoded into the GVR genotype using the labels in vector B . In the worst case, when $L = \infty$, Algorithm 2 runs in $O(n^3m)$. This worst-case running time is acceptable taking into account that the procedure: (i) evaluates a complex stochastic objective function, (ii) repairs individuals verifying a stochastic constraint, and (iii) often contributes to the convergence of the algorithm by improving input solutions. In practice, however, the *s-split*

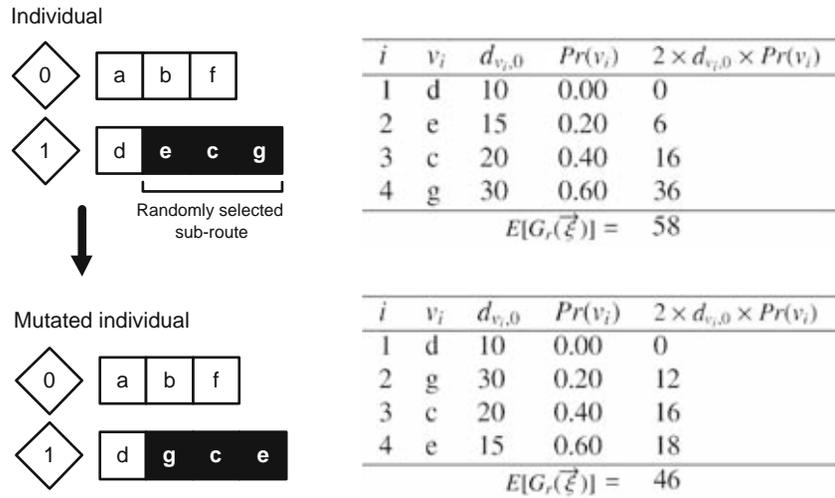


Fig. 2. Inversion mutation applied to the GVR genotype in the proposed MA.

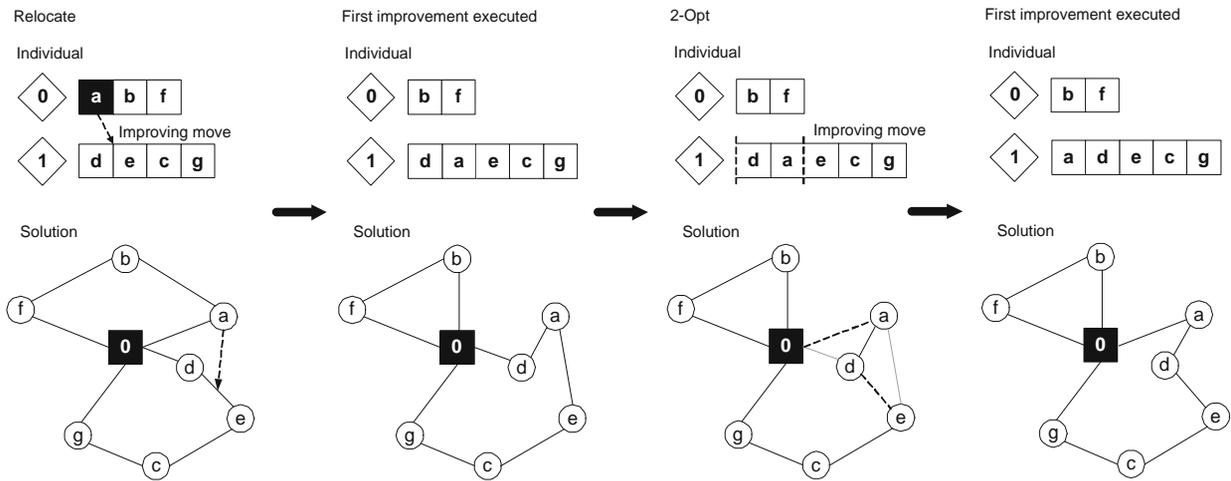


Fig. 3. Local search procedure applied to the GVR genotype in the proposed MA.

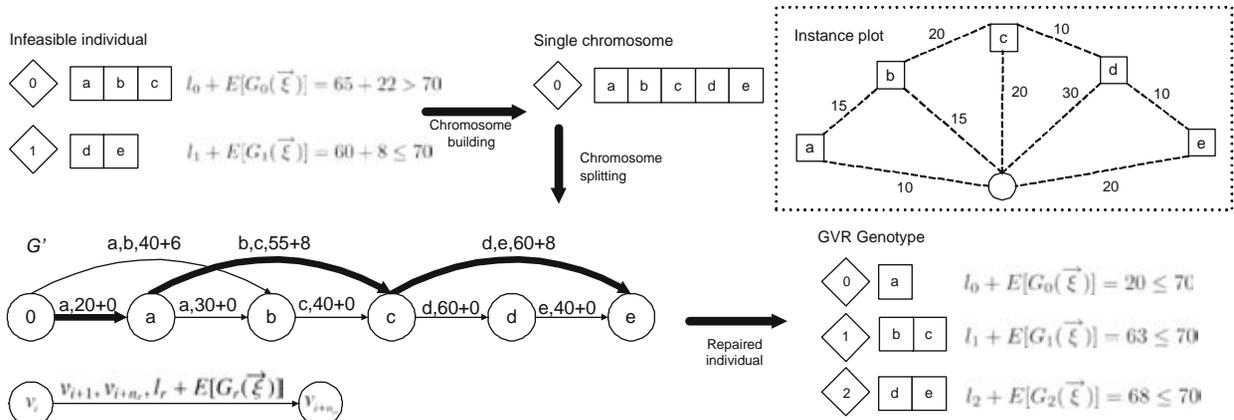


Fig. 4. s-Split procedure for individual evaluation and repair in the proposed MA.

procedure implements a data structure that allows for the recovery in constant time of previous arc evaluations, avoiding many expensive recalculations of the cost of recourse (steps 13–21) along the execution of the MA. Finally, it is worth mentioning that s-split inherits from the original Prins' proposition the flexibility to be easily extended to work on MC-VRPSD variants with richer side constraints [30].

Algorithm 2. S-split algorithm for individual evaluation and repair

```

1: set  $Z_0 = 0$  and  $Z_i = +\infty$ ,  $1 \leq i \leq n$ 
2: set  $B_i = 0$ ,  $1 \leq i \leq n$ 
3: for  $i = 1$  to  $n$  do
4:    $l_r = 0$ 
5:    $E[G_r(\vec{\xi})] = 0$ 
6:    $j = i$ 
7:   while  $j \leq n$  and  $l_r + E[G_r(\vec{\xi})] \leq L$  do
8:     if  $j = i$  then
9:        $l_r = d_{0,v_j} + d_{v_j,0}$ 
10:    else
11:       $l_r = l_r - d_{v_{j-1},0} + d_{v_{j-1},v_j} + d_{v_j,0}$ 
12:    end if
13:    for  $k = 1$  to  $j$  do
14:       $x = 1$ ,  $y = 1$ 
15:      for product  $p = 1$  to  $m$  do
16:         $x = x \times F(v_{k+1}, v_{j-1}, p)$ 
17:         $y = y \times F(v_{k+1}, v_j, p)$ 
18:      end for
19:       $Pr(v_j) = Pr(v_j) + (x - y) \times Pr(v_k)$ 
20:    end for
21:     $E[G_r(\vec{\xi})] = E[G_r(\vec{\xi})] + 2 \times d_{v_j,0} \times Pr(v_j)$ 
22:    if  $l_r + E[G_r(\vec{\xi})] \leq L$  then
23:      if  $Z_{i-1} + l_r + E[G_r(\vec{\xi})] \leq Z_j$  then
24:         $Z_j = Z_{i-1} + l_r + E[G_r(\vec{\xi})]$ 
25:         $B_j = i - 1$ 
26:      end if
27:       $j = j + 1$ 
28:    end if
29:  end while
30: end for

```

4. Computational experiments

4.1. Instances

Since the MC-VRPSD is a new problem in the vehicle routing literature, there is still a lack of benchmark instances to test the proposed MA. Hence, we constructed three new sets of problems with different characteristics. The first two sets, namely, *tuning set* and *test set*, are randomly generated instances built using the instance generator for VRPSDs reported in [25]. On the other hand, the third set, henceforth called *adapted MC-VRP instances*, is comprised of adaptations to the MC-VRPSD of the 40 problems proposed by El Fallahi et al. [14] for the MC-VRP.

The tuning set consists of 180 instances generated with the following characteristics. First, 50, 100, and 200 customers and the depot were randomly distributed over a 100×100 Euclidean space. Each client v has a demand for three different products ($p = 1, 2, 3$) following a normal distribution $N(\mu_{v,p}, \sigma_{v,p})$, where $\mu_{v,p}$ takes the value of 10 or 30 and $\sigma_{v,p}$ is set such that the coefficient of variation $CV = \sigma_{v,p}/\mu_{v,p}$ is either 0.1 or 0.3. The capacity of the compartments was fixed by defining a *tightness ratio* $(\sum_{v \in \mathcal{V} \setminus \{0\}} \mu_{v,p})/Q_p$ equal to 10 [7]. Finally, the maximum distance per route was set as $L = \beta \times \max_{v \in \mathcal{V} \setminus \{0\}} d_{0,v}$, where β is uniformly distributed between 3 and 4. Thirty instances were generated for each combination of

number of customers and coefficient of variation. The test set comprises 180 different instances generated with the same parameters, except for the mean values of the demands $\mu_{v,p}$, which were set using a uniform distribution $U[10, 100]$.

El Fallahi et al. [14] built two testbeds for the MC-VRP by adapting 20 classical instances for the capacitated VRP and the distance-constrained capacitated VRP from Beasley's OR library [2] to the multi-compartment scenario. The instances range from 50 to 484 customers. In 16 out of the 20 problems the customers are randomly distributed on the Euclidean space, while in the remaining 4 the customers are grouped into clusters; in 7 problems the customers have service times and there exists a distance constraint. The two sets of instances were generated as follows. Set 1 was obtained dividing by two the capacity of the vehicles and the customer demands from the original 20 instances. On the other hand, Set 2 was generated by randomly splitting each customer demand into two parts, and setting the capacity of each compartment as follows: $Q_1 = (Q \times \bar{q}_1)/(\bar{q}_1 + \bar{q}_2)$ and $Q_2 = (Q \times \bar{q}_2)/(\bar{q}_1 + \bar{q}_2)$, where \bar{q}_1 and \bar{q}_2 are the average demands generated for products 1 and 2, and Q is the capacity of the vehicles in the original instance.

We adapted to the stochastic scenario the MC-VRP instances from [14] by assuming normally distributed demands for each customer and product. For each demand, the mean $\mu_{v,p} \forall v \in \mathcal{V}, p \in \mathcal{P}$ is the deterministic demand on the MC-VRP instance, and the standard deviation is $\sigma_{v,p} = 0.3 \times \mu_{v,p} \forall v \in \mathcal{V}, p \in \mathcal{P}$. For those instances without a distance constraint, we made $L = 4 \times \max_{v \in \mathcal{V} \setminus \{0\}} d_{0,v}$ and set to 0 the service time for each customer.

4.2. Implementation and parameter tuning

The MA was coded using Java Grid-Enabled Genetic Algorithm (JG²A), an object-oriented framework for the rapid development and parallel execution of evolutionary algorithms in a grid environment [5]. Our current implementation supports normal distributions for the demands, but can be easily adapted to other probability distributions having the cumulative property, meaning that the sum of two or more independent and Ψ distributed random variables is also Ψ distributed (e.g., the normal, Poisson, and gamma distributions) [10,22]. The statistical support for the computation of probabilities is provided by the package Stochastic Simulation in Java (SSJ) [23].

The parameters for the MA were set via a two-phase study. The focus of this two-phase study was to find a good set of parameters for the MA without conducting an excessive number of experiments, thus reducing the overall computational time. In the first phase, the objective was to identify the values for the population size (P); and the crossover, mutation, and local search probabilities (p_c , p_m , and p_{ls}) that lead the MA to achieve a good balance between computational efficiency and solution accuracy. Then, the results of the first phase were used as input in the second phase to fine tune the number of generations (T).

In the first phase, a full 2⁴ experimental design was conducted for P , p_c , p_m , and p_{ls} . After preliminary testing, the following low and high levels were selected for each parameter: 30 and 80 for P , 0.5 and 0.9 for p_c , and 0.2 and 0.8 for p_m and p_{ls} . Each of the 16 ($= 2 \times 2 \times 2 \times 2$) combinations was tested on the 180 instances comprised on the tuning set. The running time for the MA was set proportional to the problem size as follows: for 50, 100, and 200 customers the allocated time was set to 15, 45, and 90 min, respectively. The execution of the 2880 ($= 16 \times 180$) experiments took a total CPU time of 2400 h (100 days). Thanks to the parallel execution of JG²A, the experiment was completed in just 68.5 h (nearly 3 days) on a computer grid with 35 homogeneous PCs, each one of them with a 1.85 GHz Intel Core 2 Duo processor, 4 GB of RAM, and running on Debian GNU/Linux 4.0.

To measure the impact of each combination of parameters on the accuracy of the algorithm, the best solution found on each

Table 1
Performance of MA/SCS on the MC-VRP instances.

Instance	n	MA/EPW		TS/EPW		MA/SCS		Gap MA/EPW	Gap TS/EPW	Gap best
		Cost	Time	Cost	Time	Cost	Time ^a			
<i>Set 1</i>										
vrpnc1	50	524.6	17.4	524.6	19.6	524.6	9.7	0.00	0.00	0.00
vrpnc2	75	855.8	25.4	850.0	56.3	857.3	16.3	0.18	0.86	0.86
vrpnc3	100	876.8	21.8	831.3	50.9	841.6	23.5	-4.01	1.24	1.24
vrpnc4	150	1089.6	93.9	1061.1	285.4	1045.9	46.8	-4.01	-1.43	-1.43
vrpnc5	199	1389.6	115.9	1348.3	403.6	1381.9	85.9	-0.55	2.49	2.49
vrpnc6	50	571.4	6.1	575.9	12.3	556.7	8.4	-2.58	-3.34	-2.58
vrpnc7	75	933.0	39.2	970.8	21.8	924.3	16.0	-0.93	-4.79	-0.93
vrpnc8	100	969.2	18.7	888.6	60.4	877.6	23.9	-9.46	-1.24	-1.24
vrpnc9	150	1230.9	98.7	1232.1	98.3	1223.1	47.1	-0.63	-0.73	-0.63
vrpnc10	199	1520.1	140.2	1538.6	35.2	1497.5	67.0	-1.49	-2.67	-1.49
vrpnc11	120	1046.1	47.8	1043.3	142.5	1046.9	38.2	0.08	0.35	0.35
vrpnc12	100	820.6	18.2	819.5	63.1	821.0	24.3	0.04	0.18	0.18
vrpnc13	120	1547.4	76.4	1582.2	44.3	1586.9	44.1	2.55	0.29	2.55
vrpnc14	100	866.9	23.3	868.6	43.5	868.1	23.8	0.14	-0.06	0.14
E072-04f	72	241.9	11.7	244.5	26.7	242.0	19.0	0.03	-1.03	0.03
E076-08s	76	761.9	15.3	748.5	24.7	757.1	16.0	-0.64	1.14	1.14
E076-07u	76	703.5	15.1	692.2	24.9	698.8	15.5	-0.67	0.96	0.96
E135-07f	135	1193.2	47.3	1174.4	95.9	1195.3	58.6	0.17	1.78	1.78
E241-22k	241	748.6	504.4	741.4	386.9	763.5	109.7	1.99	2.98	2.98
E484-19k	484	1250.1	1643.6	1149.5	3591.7	1250.9	1010.5	0.07	8.83	8.83
<i>Set 2</i>										
vrpnc1	50	558.8	17.4	556.1	15.3	550.2	9.8	-1.54	-1.07	-1.07
vrpnc2	75	888.6	25.5	863.6	13.9	884.3	17.2	-0.49	2.39	2.39
vrpnc3	100	878.4	21.8	837.6	39.8	852.7	25.0	-2.93	1.80	1.80
vrpnc4	150	1089.1	93.9	1070.7	109.7	1080.0	50.5	-0.84	0.87	0.87
vrpnc5	199	1408.5	115.9	1361.4	208.4	1403.5	91.4	-0.35	3.09	3.09
vrpnc6	50	569.4	16.5	563.4	10.2	562.7	8.6	-1.19	-0.13	-0.13
vrpnc7	75	955.1	39.2	949.0	22.0	946.3	16.4	-0.93	-0.29	-0.29
vrpnc8	100	958.9	18.7	916.2	18.3	874.7	24.1	-8.78	-4.53	-4.53
vrpnc9	150	1262.7	98.7	1290.8	8.6	1240.5	51.4	-1.76	-3.90	-1.76
vrpnc10	199	1509.1	140.2	1490.2	190.3	1515.6	77.7	0.43	1.71	1.71
vrpnc11	120	1122.9	47.8	1201.6	27.9	1162.4	47.7	3.52	-3.26	3.52
vrpnc12	100	926.5	18.2	934.1	15.8	916.6	23.5	-1.07	-1.87	-1.07
vrpnc13	120	1542.4	76.4	1582.3	21.9	1607.4	37.4	4.22	1.59	4.22
vrpnc14	100	966.5	23.3	1141.6	35.7	937.6	23.6	-2.99	-17.87	-2.99
E072-04f	72	263.6	11.7	262.3	5.6	262.5	22.2	-0.40	0.09	0.09
E076-08s	76	793.5	15.4	772.2	13.9	778.2	14.8	-1.93	0.77	0.77
E076-07u	76	702.2	15.1	697.8	16.5	704.1	15.7	0.28	0.91	0.91
E135-07f	135	1233.2	47.3	1235.2	51.9	1293.8	48.7	4.92	4.75	4.92
E241-22k	241	796.7	504.5	787.8	202.9	810.7	116.6	1.76	2.91	2.91
E484-19k	484	1240.9	1643.6	1177.3	2122.5	1269.0	782.3	2.27	7.79	7.79
Average								-0.69	0.04	0.96

Time (s): execution time; Gap MA/EPW (%): gap of MA/SCS relative to MA/EPW; Gap TS/EPW (%): gap of MA/SCS relative to TS/EPW; Gap Best (%): gap of MA/SCS with respect to the best among MA/EPW and TS/EPW.

^aRunning in a PC with a 2.4GHz Core 2 Duo processor, 4GB of RAM, and running Windows XP Professional.

experiment was compared to that obtained using the stochastic Clarke and Wright (SCW) heuristic proposed in [26]. The accuracy of MA was measured as the improvement gap $\Delta_{MA|SCW}$ with respect to the SCW heuristic given by

$$\Delta_{MA|SCW} = \frac{E[C(\mathcal{R})]_{MA} - E[C(\mathcal{R})]_{SCW}}{E[C(\mathcal{R})]_{SCW}} \quad (12)$$

where $E[C(\mathcal{R})]_{MA}$ and $E[C(\mathcal{R})]_{SCW}$ are the values for the total cost expressed in (3) for the MA and the SCW, respectively. Henceforth, we use Eq. (12) to calculate the gap $\Delta_{Alg1|Alg2}$ between two algorithms $Alg1$ and $Alg2$.

The results showed that among all parameters and their interactions, only the effect of the population size P is statistically significant to the performance of MA. Based on the fact that setting P at its high value has a large and positive impact on solution accuracy, the population size was set at $P = 80$. Based on statistical evidence and the impact on the execution time, the rest of the parameters were set at their low levels, namely, $p_c = 0.5$, $p_m = 0.2$, and $p_{ls} = 0.2$.

For the second phase of the parameter tuning, the 180 instances of the tuning set were solved using the parameters fixed in phase one. For each problem, the best solution and the execution time were measured at generations 500, 1000, 2000, and 3000. The results show that the marginal improvement on accuracy after 1000 generations does not pay off the computational cost. Indeed, the average improvement on the best solution between generations 1000 and 3000 is less than 1%, while the computational time increases twofold. Consequently, the number of generations T was set at 1000.

In sum, following this two-phase study, the values of the parameters were set at $T = 1000$, $P = 80$, $p_c = 0.5$, $p_m = 0.2$, and $p_{ls} = 0.2$. Under these parameters, the expected number of calls to the genetic operators is 20,000 for the crossover and 16,000 for the mutation and local search procedures.

4.3. Benchmark algorithms

Even though there are no previously published results, to assess the effectiveness of the MA we considered two alternative solution

Table 2

Performance metrics for the test set.

Customers	CV	MA			SCW			MA/SCS		
		Best sol.	Avg. gap	Avg. time	Best sol.	Avg. gap	Avg. time	Best sol.	Avg. gap	Avg. time
50	0.1	30	0.00	114.6	0	7.41	0.8	0	8.74	11.2
50	0.3	30	0.00	119.2	0	2.81	0.8	0	5.32	11.4
100	0.1	30	0.00	503.2	0	2.09	9.1	0	4.58	29.8
100	0.3	30	0.00	504.0	0	7.62	8.5	0	11.33	29.7
200	0.1	30	0.00	2553.9	0	3.41	99.8	0	6.47	116.1
200	0.3	30	0.00	2882.1	0	2.30	106.2	0	4.65	109.9
Average			0.00			4.27			6.85	

Best sol.: number of best solutions found; Avg. gap (%): average gap with respect to the best solution found; Avg. time (s): average execution time, on a PC with a 2.4GHz Core 2 Duo processor, 4GB of RAM, and running Windows XP Professional.

approaches: (i) the SCW heuristic [26] and (ii) a deterministic strategy based on the approach proposed by Fleury et al. [17] for the SCARP. For the sake of completeness the two approaches are briefly described below.

The SCW heuristic is an extension to the multi-compartment scenario of the savings algorithm proposed by Dror and Trudeau [13] for the VRPSD. Similar to its deterministic counterpart, the algorithm starts from a trivial solution made up of n round trips from the depot to each customer. At any given iteration two routes containing customers v and u as extreme nodes (those connected to the depot) are merged generating an expected saving $s_{v,u} = E[C_v] + E[C_u] - E[C_{u,v}]$, where $E[C_v]$ and $E[C_u]$ are the expected costs of the routes servicing customers v and u , and $E[C_{u,v}]$ is the expected cost of the merged route. In contrast to the classical savings algorithm, in the SCW the savings of all possible mergers must be recalculated at each iteration, because their values depend on the current configuration of the routes. The largest saving is selected and the routes merged only if the new route satisfies the distance constraint. The merging process continues until the largest saving is no longer positive.

The second alternative approach is based on the *spare capacity strategy* (SCS) introduced by Fleury et al. [17] to deal with the uncertain demands on the SCARP. The SCS strategy applied to the MC-VRPSD consists of solving the deterministic version of the problem (the MC-VRP) preserving some capacity $\alpha \times Q_p$ on each compartment, where $\alpha \in [0, 1]$.

We implemented a memetic algorithm for the deterministic MC-VRP applying the SCS. This memetic algorithm (henceforth labeled MA/SCS) uses the same building blocks and parameters settings ($T = 1000$, $P = 80$, $p_c = 0.5$, $p_m = 0.2$, and $p_s = 0.2$) as the stochastic version. However, the s -split procedure is replaced by the split mechanism proposed by El Fallahi et al. [14] for the service policy in which the split deliveries are forbidden. To measure the quality of the solutions provided by MA/SCS for the MC-VRPSD, the best individual in the final generation is evaluated in terms of the objective function (3) using the TAP approximation.

To set an appropriate value for α , we tested the algorithm on the tuning set using values of α of 0.3, 0.2, 0.1 and 0. The results presented in Table 5 (Appendix A) suggest that high values of α avoid route failures (high cost of recourse), but might end up using a large (an inefficient) number of routes (high cost of planned routes) that affect the overall cost. Based on these results, $\alpha = 0.1$ offers a good tradeoff between the two components of the total expected cost that translates into the best performance of the MA/SCS to solve the MC-VRPSD.

4.4. Results

4.4.1. MC-VRP instances

To estimate how effective the spare capacity approach is for the MC-VRPSD, we tried first MA/SCS on the MC-VRP. We tested MA/SCS

on the two sets of MC-VRP instances from [14] described in Section 4.1. For each of the 40 instances, we conducted a single run of the algorithm using the standard set of parameters reported in Section 4.3 and setting $\alpha = 0$. Table 1 presents the results and a comparison with the solutions provided by the memetic algorithm and the tabu search introduced in [14], denoted MA/EPW and TS/EPW, using as acronym the last names of the authors.

The results show an average gap of 0.96% with respect to the best solutions reported in [14], which to the best of our knowledge are the best known for these instances. Moreover, MA/SCS improved the best solutions in 14 out of the 40 instances. These figures are encouraging, specially since MA/SCS was not specifically designed to solve the MC-VRP, and suggest that the algorithm is competitive enough to be used as an alternative solution method for the MC-VRPSD using the spare capacity strategy.

4.4.2. Test set

To assess the benefits of using the proposed MA over the two alternative approaches for the MC-VRPSD, namely, the SCW heuristic and MA/SCS to solve the MC-VRPSD, we conducted a single run of each approach on the 180 instances in the test set. Table 2 summarizes the results, classifying the instances into six different groups of 30 problems according to their size (number of customers) and the coefficient of variation of the demands. For each group of instances, the three solution methods are compared under the light of three metrics: (i) the number of best solutions found in terms of objective function (3), (ii) the average gap with respect to the best solution found, and (iii) the average execution time.

In terms of solution quality, the results show that the proposed MA outperforms the two alternative approaches. As Table 2 illustrates, the best solution for the 180 instances was found solely by the memetic algorithm. The results indicate that the superior gap of MA against the alternative approaches decreases as the number of customers increases. In problems with 50 customers, the average gap to the best solution is 5.11% for SCW and 7.03% for MA/SCS. On the other hand, on instances with 200 customers the average gap reduces to 2.85% for SCW and 5.56% for MA/SCS. As expected, with regard to the computational efficiency, both SCW and MA/SCS outperform MA.

4.4.3. Adapted MC-VRP instances

From a more managerial perspective, we conducted a second set of experiments to study the behavior on the field of the solutions for the MC-VRPSD provided by the three methods. To do so, we conducted a single run of MA, SCW, and MA/SCS on the two sets of adapted MC-VRP instances described in Section 4.1; then, we evaluated each solution over 50,000 independent scenarios of demand realizations using stochastic simulation.

Table 3 reports the first set of performance metrics: (i) the planned cost ($= \sum_{r \in \mathcal{R}} l_r$), (ii) the estimated cost of recourse over the

Table 3
Performance metrics for the adapted MC-VRP instances.

Instance	n	MA					SCW					MA/SCS				
		Dist.	Rec.	Cost (gap)	CV (C(\mathcal{R}))	Time	Dist.	Rec.	Cost (gap)	CV (C(\mathcal{R}))	Time	Dist.	Rec.	Cost (gap)	CV (C(\mathcal{R}))	Time
<i>Set 1</i>																
vrpnc1	50	556.3	11.3	567.6 (0.00)	2.77	99.1	570.0	15.3	585.3 (3.13)	3.32	0.6	554.3	15.7	570.0 (0.42)	2.93	8.6
vrpnc2	75	869.2	77.3	946.6 (0.00)	2.53	163.1	857.4	106.1	963.5 (1.79)	2.66	2.2	912.2	86.5	998.7 (5.51)	4.62	15.7
vrpnc3	100	851.6	34.6	886.2 (0.00)	1.50	383.7	941.5	9.7	951.2 (7.34)	1.89	6.5	886.0	22.5	908.4 (2.51)	2.64	22.4
vrpnc4	150	1060.4	118.1	1178.5 (0.00)	2.31	775.6	1092.3	111.4	1203.7 (2.14)	2.76	26.7	1150.4	75.3	1225.7 (4.01)	4.00	54.0
vrpnc5	199	1212.6	255.5	1468.2 (0.00)	2.00	1261.4	1337.5	169.4	1506.9 (2.64)	2.18	66.2	1469.0	138.0	1607.0 (9.46)	4.42	74.5
vrpnc6	50	569.4	6.8	576.3 (0.81)	2.09	72.1	620.2	6.7	626.9 (9.68)	2.30	0.3	556.7	14.9	571.6 (0.00)	3.06	8.2
vrpnc7	75	946.9	34.7	981.6 (0.00)	2.73	110.9	987.7	11.5	999.2 (1.79)	1.73	0.8	951.5	36.1	987.6 (0.61)	2.81	16.1
vrpnc8	100	916.8	1.0	917.8 (0.09)	0.51	254.2	992.1	6.5	998.6 (8.90)	1.31	2.6	908.0	9.0	917.0 (0.00)	1.85	21.2
vrpnc9	150	1244.9	5.2	1250.1 (0.00)	0.91	449.4	1296.1	6.6	1302.7 (4.21)	0.95	8.7	1242.0	14.5	1256.5 (0.51)	1.50	43.5
vrpnc10	199	1506.3	24.1	1530.3 (0.00)	1.68	771.4	1677.9	13.1	1691.0 (10.50)	1.20	22.7	1527.0	51.9	1578.9 (3.17)	2.33	70.2
vrpnc11	120	994.7	255.9	1250.6 (0.00)	2.45	929.4	1236.6	89.9	1326.4 (6.06)	4.30	16.4	1288.3	90.7	1379.0 (10.26)	7.14	45.9
vrpnc12	100	683.5	280.9	964.4 (0.00)	5.12	506.6	1005.5	49.7	1055.2 (9.42)	3.10	7.3	958.5	85.2	1043.7 (8.23)	4.94	23.3
vrpnc13	120	1564.7	0.3	1565.0 (0.00)	0.25	548.8	1592.3	0.0	1592.3 (1.74)	0.10	3.8	1642.7	2.6	1645.3 (5.13)	0.93	36.3
vrpnc14	100	944.0	55.1	999.1 (0.00)	4.13	573.9	1048.6	19.6	1068.2 (6.92)	3.06	3.5	959.8	57.3	1017.1 (1.80)	4.44	23.5
E072-04f	72	226.3	46.4	272.7 (0.00)	4.49	309.6	232.0	46.5	278.5 (2.13)	4.44	2.3	263.5	17.8	281.3 (3.16)	5.22	19.1
E076-08s	76	716.6	100.6	817.2 (0.00)	2.98	179.8	814.6	31.6	846.2 (3.55)	3.46	2.1	788.9	34.2	823.1 (0.72)	3.57	14.2
E076-07u	76	702.6	27.1	729.6 (0.00)	3.19	199.0	761.3	15.6	776.9 (6.48)	2.97	2.2	720.5	24.4	745.0 (2.10)	3.24	16.1
E135-07f	135	1117.6	251.2	1368.8 (0.00)	5.82	4200.6	1241.3	191.5	1432.8 (4.68)	3.95	22.6	1297.2	121.5	1418.8 (3.65)	6.91	44.4
E241-22k	241	662.1	192.2	854.3 (0.00)	2.43	1512.9	735.9	141.1	877.1 (2.66)	2.22	110.0	814.6	109.6	924.2 (8.17)	4.27	97.9
E484-19k	484	1185.1	140.6	1325.7 (0.00)	1.55	13 491.6	1207.9	138.5	1346.4 (1.56)	1.24	2180.7	1322.4	16.1	1338.5 (0.97)	1.54	505.5
<i>Set 2</i>																
vrpnc1	50	554.9	13.8	568.7 (0.00)	3.13	210.2	579.1	31.3	610.4 (7.34)	3.55	0.6	557.0	13.1	570.1 (0.24)	3.19	8.4
vrpnc2	75	803.5	172.8	976.3 (0.00)	3.50	348.3	884.4	100.8	985.2 (0.91)	3.75	2.2	915.9	62.2	978.1 (0.19)	4.52	15.7
vrpnc3	100	875.6	16.2	891.8 (0.00)	2.40	1001.9	924.7	22.9	947.6 (6.25)	2.84	6.3	889.4	16.9	906.3 (1.62)	2.12	22.7
vrpnc4	150	1076.7	110.7	1187.3 (0.00)	3.07	2127.3	1165.9	44.6	1210.4 (1.95)	3.05	26.1	1194.6	66.8	1261.4 (6.24)	3.96	53.9
vrpnc5	199	1295.6	201.7	1497.3 (0.00)	2.52	3456.9	1341.9	181.5	1523.5 (1.75)	2.31	64.8	1493.7	98.9	1592.6 (6.37)	3.92	77.9
vrpnc6	50	572.2	13.6	585.8 (1.43)	3.10	126.3	619.7	6.6	626.3 (8.44)	2.51	1.0	562.3	15.3	577.6 (0.00)	3.21	9.0
vrpnc7	75	978.9	22.0	1000.9 (0.00)	2.46	145.2	1022.4	17.0	1039.4 (3.84)	1.80	0.8	976.9	30.3	1007.2 (0.62)	2.99	15.1
vrpnc8	100	916.9	5.2	922.1 (0.00)	1.29	472.8	1001.5	9.7	1011.2 (9.67)	1.79	2.7	911.6	23.0	934.5 (1.35)	2.76	22.7
vrpnc9	150	1247.6	11.3	1259.0 (0.26)	1.49	712.4	1289.1	23.1	1312.3 (4.50)	1.66	8.7	1238.6	17.2	1255.8 (0.00)	1.79	46.2
vrpnc10	199	1545.0	27.5	1572.5 (0.00)	1.82	1145.8	1660.6	26.1	1686.6 (7.26)	1.66	22.9	1567.8	33.8	1601.7 (1.85)	1.81	72.4
vrpnc11	120	1099.2	159.9	1259.1 (0.00)	2.57	2798.3	1172.8	165.1	1337.9 (6.26)	4.85	16.7	1303.7	56.5	1360.2 (8.03)	5.59	50.7
vrpnc12	100	695.0	304.1	999.1 (0.00)	4.85	1364.3	986.1	60.8	1046.9 (4.78)	3.79	7.2	1020.7	51.9	1072.6 (7.35)	4.77	24.6
vrpnc13	120	1558.6	0.2	1558.8 (0.00)	0.31	809.3	1592.3	0.3	1592.6 (2.17)	0.36	3.6	1597.8	7.5	1605.3 (2.98)	1.56	38.2
vrpnc14	100	975.4	52.0	1027.4 (0.00)	4.46	791.9	1035.7	22.9	1058.6 (3.04)	3.23	3.6	998.7	36.8	1035.5 (0.78)	3.88	22.5
E072-04f	72	227.0	55.4	282.4 (1.21)	4.52	1176.2	244.2	35.4	279.6 (0.21)	3.59	2.3	268.0	11.0	279.0 (0.00)	4.55	17.2
E076-08s	76	822.1	5.5	827.6 (0.00)	1.48	377.1	834.0	37.6	871.6 (5.32)	2.62	2.1	812.7	35.8	848.6 (2.53)	3.74	14.7
E076-07u	76	708.2	29.7	737.9 (0.00)	2.94	452.7	753.5	26.8	780.3 (5.74)	3.34	2.2	733.5	24.7	758.2 (2.75)	3.42	15.5
E135-07f	135	939.8	482.1	1421.9 (1.36)	6.21	15 919.7	1215.9	187.0	1402.8 (0.00)	4.29	24.1	1328.1	123.9	1452.0 (3.50)	7.41	56.2
E241-22k	241	636.0	234.5	870.5 (0.00)	2.43	3508.1	774.6	118.0	892.6 (2.54)	2.21	108.3	871.4	69.1	940.4 (8.03)	3.80	103.0
E484-19k	484	1152.9	186.7	1339.6 (0.00)	1.64	12 913.8	1214.3	142.7	1357.0 (1.30)	1.48	2131.1	1343.2	16.7	1359.9 (1.52)	1.56	900.7
				(0.13)	2.69				(4.51)	2.60				(3.16)	3.57	

Dist.: total planned cost; Rec.: estimated cost of recourse; Cost (gap): estimated total cost and gap (%) with respect to the best solution found; CV(C(\mathcal{R})): estimated coefficient of variation of the total cost; Time (s): execution time, on a PC with a 2.4GHz Core 2 Duo processor, 4GB of RAM, and running Windows XP Professional.

50,000 replications ($=\sum_{r \in \mathcal{R}} \hat{G}_r(\vec{\xi})$), (iii) the estimated total cost of the solution ($=\hat{C}(\mathcal{R})$), (iv) its gap with respect to the best solution found, (v) the estimated coefficient of variation of the total cost $CV(C(\mathcal{R})) = \hat{\sigma}_{C(\mathcal{R})} / \hat{\mu}_{C(\mathcal{R})}$, and (vi) the execution time.

In terms of solution accuracy, the results are consistent to those reported on Section 4.4.2 for the test set. In 34 out of the 40 instances (85.00%) the proposed memetic algorithm (MA) obtained the solution with the lowest average total cost, while SCW and MA/SCS only achieved 1 (2.50%) and 5 (12.05%), respectively. A closer look at Table 3 also shows that the largest gap to the best solution found is just 1.43% for MA (set2-vrpnc6), while for SCW and MA/SCS is as large as 10.50% (set1-vrpnc10) and 10.26% (set1-vrpnc10), respectively. This observation highlights the balance towards MA in terms of quality stability.

Another interesting observation arises from the data associated with the variability of the total cost of the solutions, $CV(C(\mathcal{R}))$. The results suggest that the two approaches that specifically incorporate the stochastic nature of the problem, namely, SCW and MA, produce solutions of lower variability: 2.60% and 2.69% (on average), respectively, vs. 3.57% for MA/SCS. However, a closer look to the results

shows that both, SCW and MA, generate solutions within a wide variability range. For instance, the value $CV(C(\mathcal{R}))$ varies along the testbed from a low 0.36% (set2-vrpnc13) to 4.85% (set2-vrpnc11) for SCW and from 0.25% (set1-vrpnc13) to 6.21% (set2-E135-07f) for MA. These results could be explained by the fact that neither SCW nor MA incorporate an explicit mechanism to control the variability of the objective function, hence solutions with low coefficient of variation cannot be guaranteed.

The simulation results also shed some light on the structure of the solutions. Table 4 compares each approach based on three performance metrics: (i) the average number of failures per route, (ii) the probability that one or more routes on the solution violate the distance constraint (event A), denoted as $Pr(A)$, and (iii) the average utilization per route of distance L , denoted by $\bar{\rho}$.

With regard to the average number of failures per route, the targeted purpose of the spare capacity strategy is precisely to avoid them. The results support this fact, showing that MA/SCS achieves the lowest average number of failures per route (0.12). Moreover, not even in one of the 40 instances, the solution provided by MA/SCS exceeds an average number of failures per route of 0.3. On the other

Table 4
Structure metrics for the adapted MC-VRP instances.

Instance	n	MA			$\bar{\rho}$	SCW			$\bar{\rho}$	MA/SCS			$\bar{\rho}$
		Routes	Avg. fail.	Pr(A)		Routes	Avg. fail.	Pr(A)		Routes	Avg. fail.	Pr(A)	
Set 1													
vrpnc1	50	6	0.09	0.00	0.54	6	0.11	0.02	0.55	6	0.13	0.00	0.54
vrpnc2	75	9	0.40	0.00	0.61	8	0.56	0.32	0.70	11	0.28	0.03	0.52
vrpnc3	100	8	0.17	0.09	0.55	10	0.03	0.02	0.48	9	0.11	0.03	0.50
vrpnc4	150	10	0.43	0.34	0.59	11	0.33	0.48	0.55	13	0.19	0.00	0.47
vrpnc5	199	12	0.66	0.14	0.61	15	0.33	0.00	0.50	18	0.20	0.14	0.45
vrpnc6	50	6	0.06	0.26	0.90	7	0.04	0.11	0.80	6	0.12	0.54	0.89
vrpnc7	75	12	0.16	0.71	0.90	13	0.06	0.19	0.84	12	0.13	0.67	0.91
vrpnc8	100	10	0.01	0.00	0.83	10	0.03	0.06	0.87	10	0.03	0.19	0.83
vrpnc9	150	15	0.02	0.16	0.92	15	0.03	0.14	0.93	15	0.05	0.37	0.92
vrpnc10	199	20	0.05	0.39	0.88	22	0.02	0.20	0.84	19	0.10	0.76	0.94
vrpnc11	120	5	0.77	0.08	0.63	6	0.44	0.22	0.56	8	0.13	0.00	0.43
vrpnc12	100	6	0.91	0.10	0.69	11	0.12	0.00	0.41	11	0.22	0.00	0.41
vrpnc13	120	11	0.00	0.01	0.96	11	0.00	0.00	0.96	11	0.00	0.03	0.97
vrpnc14	100	11	0.11	0.02	0.87	13	0.03	0.00	0.74	11	0.13	0.09	0.88
E072-04f	72	3	0.76	0.85	0.86	3	0.76	0.85	0.88	5	0.19	0.00	0.53
E076-08s	76	6	0.66	0.40	0.79	9	0.13	0.27	0.54	9	0.16	0.00	0.53
E076-07u	76	7	0.21	0.00	0.60	8	0.05	0.19	0.56	8	0.13	0.03	0.54
E135-07f	135	6	0.53	0.10	0.39	7	0.35	0.00	0.35	8	0.22	0.00	0.30
E241-22k	241	15	0.79	0.68	0.71	19	0.45	0.49	0.58	25	0.23	0.00	0.46
E484-19k	484	13	0.67	0.15	0.61	15	0.48	0.07	0.53	22	0.04	0.02	0.36
Set 2													
vrpnc1	50	6	0.10	0.00	0.54	6	0.21	0.01	0.58	6	0.10	0.00	0.54
vrpnc2	75	8	0.68	0.64	0.71	11	0.29	0.31	0.52	12	0.16	0.00	0.47
vrpnc3	100	9	0.08	0.00	0.50	9	0.08	0.06	0.53	9	0.09	0.00	0.50
vrpnc4	150	11	0.35	0.20	0.54	14	0.12	0.00	0.43	14	0.13	0.00	0.45
vrpnc5	199	15	0.35	0.37	0.50	15	0.35	0.24	0.51	19	0.13	0.02	0.42
vrpnc6	50	6	0.10	0.37	0.90	7	0.03	0.15	0.80	6	0.11	0.41	0.90
vrpnc7	75	13	0.08	0.48	0.84	14	0.08	0.15	0.80	13	0.09	0.52	0.84
vrpnc8	100	10	0.02	0.07	0.84	10	0.03	0.12	0.87	10	0.08	0.44	0.84
vrpnc9	150	15	0.03	0.15	0.92	15	0.08	0.59	0.94	15	0.04	0.28	0.92
vrpnc10	199	21	0.06	0.30	0.85	22	0.05	0.30	0.84	20	0.08	0.59	0.90
vrpnc11	120	6	0.49	0.05	0.53	6	0.52	0.33	0.56	8	0.09	0.00	0.43
vrpnc12	100	5	1.32	0.67	0.85	11	0.14	0.15	0.41	12	0.09	0.00	0.38
vrpnc13	120	11	0.00	0.00	0.95	11	0.00	0.00	0.96	11	0.01	0.08	0.96
vrpnc14	100	12	0.07	0.00	0.80	12	0.04	0.00	0.81	12	0.07	0.12	0.80
E072-04f	72	3	0.91	0.82	0.89	4	0.49	0.01	0.66	5	0.12	0.00	0.53
E076-08s	76	10	0.02	0.00	0.48	9	0.18	0.63	0.56	9	0.15	0.00	0.55
E076-07u	76	7	0.26	0.01	0.61	8	0.17	0.07	0.56	8	0.11	0.01	0.55
E135-07f	135	4	1.25	0.89	0.61	7	0.36	0.01	0.34	8	0.22	0.00	0.31
E241-22k	241	14	0.99	0.91	0.78	22	0.32	0.30	0.51	27	0.13	0.00	0.44
E484-19k	484	12	0.86	0.24	0.66	13	0.74	0.14	0.62	23	0.04	0.01	0.35
Average		9.73	0.39	0.27	0.72	11.13	0.22	0.18	0.65	12.10	0.12	0.13	0.61

Routes: total number of routes; Avg. fail.: average number of failures per route; Pr(A): probability that one or more routes violate the distance constraint; $\bar{\rho}$: average utilization of distance L per route.

hand, MA produces solutions with higher failure rates (0.39) that not necessarily translate into higher costs. For example, on instance set2-vrpnc12, the solution provided by MA is comprised of 5 routes with an average failure rate of 1.32 per route. In contrast, MA/SCS produces 12 routes with virtually no failures (0.09). In terms of cost of recourse, MA is nearly six times more expensive than MA/SCS (304.1/51.9), yet this high cost is compensated with a more efficient routing (less routes) that results on a lower overall cost of MA over MA/SCS (999.1 vs. 1072.6). The good tradeoff between cost of recourse and planned cost of MA is explained by the s-split procedure. Including the evaluation of the expected cost of recourse in the tour partitioning process allows the MA to anticipate the impact on adding a new customer to the route and eventually failing or closing early a route (and creating a new one). Besides the positive impact on the overall routing cost, a fewer number of routes is beneficial from an operational point of view, since there are always hidden costs associated with managing a larger fleet.

With regard to the probability of violating the distance constraint Pr(A), MA/SCS produces solutions that are likely to meet the distance constraint based on its capacity to avoid failures. On the other

hand, both MA and SCW explicitly handle the distance constraint (4) guaranteeing that on average every route will satisfy the maximum distance. Even though the simulation study confirms this fact, neither MA nor SCW can guarantee that they will satisfy the distance constraint for every single realization of demands. Consequently, there is a probability Pr(A) that one or more routes will violate the distance constraint during the execution phase. According to the results, MA has a larger (average) probability of failing the distance constraint (0.27) compared to SCW (0.18) and MA/SCS (0.13). This finding is explained by the third metric, namely, the average utilization of the distance constraint per route ($\bar{\rho}$). MA produces solutions that use a larger fraction of the maximum allowable distance (an average 0.72) than those generated by SCW and MA/SCS (0.65 and 0.61 respectively), ultimately building more efficient routes.

5. Concluding remarks and future research

This paper introduces the MC-VRPSD, a problem that despite its practical applicability, has not received the deserved attention in

the literature. The problem consists of designing collection routes to serve a set of customers who have independent stochastic demands for different products that must be transported on independent compartments due to product incompatibility constraints. To solve the MC-VRPSD this paper proposes a memetic algorithm based on a stochastic programming with recourse formulation. The MA couples genetic operators and local search procedures proven to be effective in deterministic VRPs with *s-split*, an evaluation and reparation mechanism that accounts for the complex multi-compartment stochastic scenario. To assess the performance of the MA, two alternative solution approaches for the MC-VRPSD are considered: the SCW heuristic [26], an extension to the multi-compartment scenario of the savings-based algorithm by Dror and Trudeau [13]; and MA/SCS, a memetic algorithm for the MC-VRP that generates solutions able to absorb the demand fluctuation based on the spare capacity strategy of Fleury et al. [17]. The three methods were tested on two sets of instances. In the first set, comprising 180 randomly generated instances, the results indicate that MA dominates both SCW and MA/SCS in terms of the total expected cost by 4.27% and 6.85% (on average), respectively. In the second set of instances comprising 40 problems adapted from the MC-VRP literature, the results also suggest that MA tends to generate solutions that use a reduced number of routes compared to those generated by SCW and MA/SCS. This observation tips the balance towards MA in terms of implementation in a practical setting because larger fleet sizes are often expensive to operate and maintain. A further analysis on the structure of the solutions showed that the proposed MA produces routes that tend to efficiently utilize the limits of the distance constraint (72% of its capacity), but at a greater risk of failing it than the alternative methods. In addition to the experimentation on the MC-VRPSD, we conducted experiments on the deterministic MC-VRP using MA/SCS. The algorithm was validated on a set of 40 instances proposed by El Fallahi et al. [14] and its results compared to those reported by these authors using an MA and a TS. The experiment showed that

our approach is quite competitive, since it unveiled solutions with an average gap of 0.96% with respect to the best result reported in [14] and improved 14 out of the 40 best known solutions.

Research currently underway includes the extension of *s-split* to incorporate a constraint that guarantees that the maximum distance is met with a given probability (i.e., chance constraint) and a new biobjective approach that simultaneously minimizes the total expected cost of the transportation plan and its coefficient of variation. Future work encompasses: the derivation of a lower bound for the MC-VRPSD to compare against the solutions provided by the MA, an extension of the proposed algorithm to the case in which preventive trips to the depot (restocking) are allowed, the adaptation to handle different recourse actions (e.g., cooperation between routes upon route failures), and the extension of the TAP approximation to the case in which the demands of a single customer for different products are correlated.

Acknowledgments

The authors would like to thank professor Christian Prins at the Université de Technologie de Troyes for providing us with the benchmark instances for the MC-VRP, José L. Walteros at the Universidad de los Andes for his technical assistance on JG²A, Harold Castro at the COMIT research group for helping us with the grid environment at the Universidad de los Andes, and the referees for their insightful comments and suggestions. This research was partially funded by Grant 471-2007 from COLCIENCIAS/SENA (Colombia) and project MILES funded by the Pays de la Loire Regional Council, France.

Appendix A. Impact of α on the performance of MA/SCS

The impact of α on the performance of MA/SCS is presented in Table 5.

Table 5
Impact of α on the performance of the spare capacity strategy for the MC-VRPSD.

Customers	CV	Instances	Metric	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
50	0.1	30	Best solutions	0	0	18	12
			Best recourse	30	8	0	0
			Best distance	0	0	0	30
			Avg. routes	18.73	16.07	13.97	12.40
			Avg. gap	27.51	12.97	1.09	4.78
50	0.3	30	Best solutions	0	6	24	0
			Best recourse	30	0	0	0
			Best distance	0	0	0	30
			Avg. routes	19.03	16.03	13.93	12.50
			Avg. gap	14.82	2.30	0.26	6.16
100	0.1	30	Best solutions	0	0	28	2
			Best recourse	30	30	0	0
			Best distance	0	0	0	30
			Avg. routes	17.20	14.87	12.97	11.70
			Avg. gap	22.08	9.44	0.14	6.68
100	0.3	30	Best solutions	0	6	24	0
			Best recourse	30	0	0	0
			Best distance	0	0	0	30
			Avg. routes	16.93	14.80	13.00	11.80
			Avg. gap	12.55	1.78	0.15	10.31
200	0.1	30	Best solutions	0	0	26	4
			Best recourse	30	30	0	0
			Best distance	0	0	0	30
			Avg. routes	16.20	14.23	12.67	11.20
			Avg. gap	17.72	8.40	0.36	7.82

Table 5 (continued)

Customers	CV	Instances	Metric	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
200	0.3	30	Best solutions	0	0	27	3
			Best recourse	30	0	0	0
			Best distance	0	0	0	30
			Avg. routes	16.20	14.20	12.70	11.17
			Avg. gap	14.34	4.63	0.03	11.89
			Average	Best solution	0.00	2.00	24.50
		Best recourse	30.00	11.33	0.00	0.00	
		Best distance	0.00	0.00	0.00	30.00	
		Avg. routes	17.38	15.03	13.21	11.79	
		Avg. gap	18.17	6.59	0.34	7.94	

Best solutions: number of best solutions found; Best recourse: number of solutions with lowest cost of recourse; Best distance: number of solutions with lowest planned distance; Avg. routes: average number of routes over 30 instances; Avg. gap (%): average gap with respect to the best solution found.

References

- [1] Avella P, Boccia M, Sforza A. Solving a fuel delivery problem by heuristic and exact approaches. *European Journal of Operational Research* 2004;152(1):170–9.
- [2] Beasley JE. OR Library: (<http://people.brunel.ac.uk/~mastjib/jeb/info.html>).
- [3] Ben Abdelaziz F, Roucaïrol C, Bacha C. Deliveries of liquid fuels to SNPD gas stations using vehicles with multiple compartments. In: IEEE international conference on systems, man and cybernetics, 2002.
- [4] Bentley JB. Fast algorithms for geometric traveling salesman problem. *INFORMS Journal on Computing* 1992;4:387–411.
- [5] Bernal A, Ramírez MA, Castro H, Walteros JL, Medaglia AL. JG²A: a grid-enabled object-oriented framework for developing genetic algorithms. In: Proceedings of IEEE systems and information engineering design symposium, 2009. p. 1–6.
- [6] Bertsimas D. A vehicle routing problem with stochastic demand. *Operations Research* 1992;40(3):574–85.
- [7] Bianchi L, Birattari M, Chiarandini M, Manfrin M, Mastrolilli M, Paquete L et al. Hybrid metaheuristics for the vehicle routing problem with stochastic demands. *Journal of Mathematical Modelling and Algorithms* 2006;5(1):91–110.
- [8] Brown G, Ellis C, Graves G. Real-time, wide area dispatch of Mobil tank trucks. *Interfaces* 1987;17(1):107–20.
- [9] Brown G, Graves G. Real-time dispatch of petroleum tanks trucks. *Management Science* 1981;27(1):19–32.
- [10] Christiansen C, Lysgaard J. A branch-and-price algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research Letters* 2007;35(6):773–81.
- [11] Clarke G, Wright JV. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research* 1964;12(4):568–81.
- [12] Cordeau JF, Laporte G, Savelsbergh M, Vigo D. Vehicle routing. In: Barnhart C, Laporte G, editors. *Handbooks in operations research and management science: transportation*, vol. 14. Amsterdam: Elsevier; 2006. p. 367–428.
- [13] Dror M, Trudeau P. Stochastic vehicle routing with modified savings algorithm. *European Journal of Operational Research* 1986;23(2):228–35.
- [14] El Fallahi A, Prins C, Wolfier Calvo R. A memetic algorithm and a tabu search for the multi-compartment vehicle routing problem. *Computers & Operations Research* 2008;35(5):1725–41.
- [15] Fischetti M, Lodi A. Local branching. *Mathematical Programming* 2003;98(1–3):23–47.
- [16] Fleury G, Lacomme P, Prins C. Evolutionary algorithms for stochastic arc routing problems. In: *EvoWorkShops 2004*. Lecture notes in computer science, vol. 3005. Berlin, Heidelberg: Springer; 2004. p. 501–12.
- [17] Fleury G, Lacomme P, Prins C, Ramdane-Chérif W. Improving robustness of solutions to arc routing problems. *Journal of the Operational Research Society* 2005;56(5):526–38.
- [18] Gendreau M, Laporte G, Séguin R. An exact algorithm for the vehicle routing problem with stochastic customers and demands. *Transportation Science* 1995;29(2):143–55.
- [19] Gendreau M, Laporte G, Séguin R. Stochastic vehicle routing. *European Journal of Operational Research* 1996;88(1):3–12.
- [20] Gendreau M, Laporte G, Séguin R. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research* 1996;44(3):469–77.
- [21] Laporte G, Loveaux F, Van Hamme L. An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research* 2002;50(3):415–23.
- [22] Law AM, Kelton WD. *Simulation modeling and analysis*. Singapore: McGraw-Hill; 2000.
- [23] L'Ecuyer P, Meliani L, Vaucher J. SSJ: a framework for stochastic simulation in Java. In: Yücesan E, Chen CH, Snowdon JL, Charnes JM, editors. *Proceedings of the 2002 winter simulation conference*. New York: IEEE Press; 2002. p. 234–42.
- [24] Mendoza JE, Castanier B, Guéret C, Medaglia AL, Velasco N. Approximating the expected cost of recourse on a multi-compartment vehicle routing problem with stochastic demands. Technical Report 08/03/AUTO, Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN). École des Mines de Nantes, France; 2008.
- [25] Mendoza JE, Castanier B, Guéret C, Medaglia AL, Velasco N. An instance generator for vehicle routing problems with stochastic demands. Technical Report 08/1/AUTO, Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN). École des Mines de Nantes, France; 2008.
- [26] Mendoza JE, Guéret C, Castanier B, Medaglia AL, Velasco N. Construction heuristics for the multi-compartment vehicle routing problem with stochastic demands. In: XIV Latin Ibero-American congress on operations research (CLAIO). Cartagena, Colombia; 2008.
- [27] Mendoza JE, Medaglia AL, Velasco N. An evolutionary-based decision support system for vehicle routing: the case of a public utility. *Decision Support Systems* 2009;46(3):730–42.
- [28] Moscato P. Memetic algorithms: a short introduction. In: Corne D, Dorigo M, Glover F, editors. *New ideas in optimization*. New York: McGraw-Hill; 1999. p. 219–34.
- [29] Pereira FB, Tavares J, Machado P, Costa E. GVR: a new genetic representation for the vehicle routing problem. In: *Proceedings of AICS 2002–13th Irish conference on artificial intelligence and cognitive science*, 2002.
- [30] Prins C. A simple and effective evolutionary algorithm for the vehicle routing problem. *Computers & Operations Research* 2004;31(12):1985–2002.
- [31] Rei W, Gendreau M, Soriano P. A hybrid Monte Carlo local branching algorithm for the single vehicle routing problem with stochastic demands. Technical Report CIRRELT-2007-24, Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT); 2007.
- [32] Rei W, Gendreau M, Soriano P. Local branching cuts for the 0–1 integer L-shaped algorithm. Technical Report CIRRELT-2007-23, Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT); 2007.
- [33] Repoussis P, Tarantilis C, Ioannou G. A hybrid metaheuristic for a real life vehicle routing problem. In: *Numerical methods and applications. Lecture notes in computer science*. Berlin, Heidelberg: Springer; 2007. p. 247–54.
- [34] Secomandi N. Comparing neuro dynamic programming algorithms for the vehicle routing problem with stochastic demands. *Computers & Operations Research* 2000;27:1201–25.
- [35] Secomandi N, Margot F. Reoptimization approaches for the vehicle-routing problem with stochastic demands. *Operations Research* 2008;1–17.
- [36] Shen Z, Ordóñez F, Dessouky MM. The minimum unmet demand stochastic vehicle routing problem. USC ISE Working paper #2006–07. University of Southern California.
- [37] Sörensen K, Sevaux M. MA|PM: memetic algorithms with population management. *Computers & Operations Research* 2006;33(5):1214–25.
- [38] Stewart W, Golden B. Stochastic vehicle routing: a comprehensive approach. *European Journal of Operational Research* 1983;14:371–85.
- [39] Sungur I, Ordóñez F, Dessouky MM. A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty. *IIE Transactions* 2008;40(5):509–23.
- [40] Tatarakis A, Minis I. Stochastic single vehicle routing with a predefined customer sequence and multiple depot returns. *European Journal of Operational Research* 2008;197(2):2.
- [41] Van Der Bruggen R, Gruson L, Salomon M. Reconsidering the distribution structure of gasoline products for a large oil company. *European Journal of Operational Research* 1995;81(3):460–73.
- [42] Yang WH, Mathur K, Ballou R. Stochastic vehicle routing with restocking. *Transportation Science* 2000;34(1):99–112.

Optimizing Road Milling and Resurfacing Actions

Mariem Zouch, Thomas Yeung and Bruno Castanier
Department of Automatic control & Industrial Engineering
Ecole des Mines de Nantes / IRCCyN , F-44307 NANTES Cedex, France

Abstract

A condition-based maintenance optimization approach is developed for the road-cracking problem in order to derive optimal maintenance policies that minimize a total discounted maintenance cost. The approach is based on a Markov decision process that takes into account multiple actions with varying effects on future road performance. Maintaining the road consists of adding a new asphalt layer; however, as resurfacing actions are constrained by a maximum total road thickness, the maintenance decision is not only how thick a layer to apply, but also how much old road to remove. Each combination of these actions leads to different maintenance costs and different future degradation behaviours. The road state is modelled by a dependent bivariate deterioration variable (the longitudinal cracking percentage and the deterioration growth rate), for taking these different changes in the cracking patterns into account. Moreover, the sensitivity to cracking for existing roads can be reduced with the addition of new layers, and thus actions that can lead to states better than good-as-new have to be considered. A numerical analysis is provided to illustrate the benefits of the introduction of the deterioration speed in the decision framework, as well as the belief that initially building a road to its maximum thickness is not optimal. The trade-offs in the design decisions and the exploitation/maintenance costs are also explored.

Keywords: road cracking, stochastic deterioration, maintenance optimization, Markov decision processes

Notations

c_0	inspection cost	$\alpha(\cdot)$	cumulative hazard function of the Poisson process
c_1	fixed costs of milling	ℓ	section thickness
c_2	fixed costs of resurfacing	ℓ_{max}	maximum section thickness
$c_a(\xi)$	variable cost of adding thickness ξ	$\gamma_1(\cdot)$	linear-form shape function of the state-dependent gamma process
$c_r(\zeta)$	variable cost of removing thickness ζ	$\gamma_2(\cdot)$	gaussian-form shape function of the state-dependent gamma process
$c_q(\cdot)$	quality cost	Δ_θ	length of discrete intervals of θ
$f(\cdot)$	density function of θ in both phases	Δ_ρ	length of discrete intervals of ρ
$g(\cdot)$	density function of ρ in both phases	ζ	milling thickness
$h(\cdot)$	joint (ρ, θ) density function in both phases	θ	deterioration growth rate
i_θ	discrete value of θ	θ_{max}	maximum growth rate
i_ρ	discrete value of ρ	λ	discount factor
N_θ	number of possible values of θ	ξ_m	last added thickness
N_ρ	number of possible values of ρ	π	maintenance policy
S	set of states	ρ	longitudinal cracking percentage
s	state of the section	τ	decision epoch length
V	Total expected discounted cost	ξ	resurfacing thickness
		$\phi(\cdot)$	maintenance effect function on the deterioration growth rate

1 Introduction

During recent decades, transportation administrations of several countries have made continuing efforts to develop and implement efficient pavement management systems (PMS) in order to define the best maintenance policies that ensure acceptable use and safety conditions of the road in the most cost-effective way possible. A survey of existing PMSs is presented in reference [2]. According to reference [9], a PMS consists of four main components: the inventory, the pavement condition evaluation, the performance prediction models, and the planning method. Definition of the deterioration model is one of the most important components, as it determines the quality and the efficiency of the optimal policy.

From a survey of maintenance optimization approaches [1, 12], deterioration models can be classified into three classes as a function of the consistency of information and knowledge on the deteriorating process. The first class consists of models based on a resistance-load relationship, such as the Paris-Ergodan equation [6]. They have proved their efficiency for reliability purposes, but because of their numerical complexity they become intractable in more maintenance decision frameworks. The second class consists of models based on Markov processes [3, 4], which are widely used in existing PMSs. These models require the determination of a transition probability matrix, usually obtained by statistical data analysis techniques [2] when large amounts of deterioration data are available, or by expert judgements when only a few data are available. The third class contains models based on Lévy processes, such as

the Brownian motion and gamma processes that are widely used for cumulative deterioration modelling [11]. Zouch *et al.* [13] proposed a deterioration model that is a combination of the two last model classes.

Moreover, existing cracking models are based on observable deterioration indexes that only model the observable deterioration through quality indexes such as the surface distress index, the distress manifestation index, or the pavement condition index [2], and do not take into account the underlying deterioration processes. In fact, most of the existing condition-based approaches utilize a restrictive definition of the system state classically defined by an observable and measurable metric such as the size of the road cracks. This can be restrictive, especially when modelling the effects of imperfect maintenance. Different imperfect actions can have the same effects on the observable deterioration, i.e. concealing the observable damage, but different effects on the level of both the underlying and future deterioration.

Few maintenance planning methods, as in references [8] and [10] consider imperfect maintenance actions in addition to minimal and perfect repairs. In these approaches, the effects of imperfect maintenance consist of a partial reduction in the deterioration level, but they do not allow for the change of the system deterioration law, as in reference [13]. In the latter paper, Zouch *et al.* present a condition-based maintenance optimization approach for road maintenance that takes into account multiple imperfect actions with different effects on the immediate state of the road, as well as its future deterioration law. This is made possible by considering a second deterioration parameter that models the underlying deterioration process in addition to the observable parameter. The two parameters are then used to define a “state-dependent” deterioration model, where the deterioration law in each decision period depends on the deterioration level just before maintenance and the performed action.

To derive the optimal maintenance policy that minimizes the discounted total cost over the infinite horizon, a classical dynamic programming formulated as a Markov decision process (MDP) [7] is solved. The state-dependent deterioration model is incorporated into the MDP framework to derive action-dependent transition matrices.

In reference [9], maintenance decisions relate to the road thickness to renew, i.e. remove and add. The total road thickness is therefore kept fixed. The objective of this paper is to extend the model presented in reference [13] to take into account a constraint of a maximum road thickness, in order to make the model more realistic and more applicable. The maximum thickness constraint renders the decision more complex, as it consists in determining not only the resurfacing layer thickness to add, but also the thickness to remove prior to resurfacing. A special feature of this model is that consecutive changes in road thickness and composition, i.e. in new and old layers, may result in a better-than-new road performance. The possibility to restore a system to a state better than as-good-as-new (AGAN) is a new aspect in the maintenance literature. Our model accounts for the reality that two roads of the same thickness are not equal, and that in general the road with more layers will be superior.

The remainder of the paper is organized as follows. Section 2 briefly presents the state-dependent deterioration model for the road longitudinal cracking process. The optimization problem is formulated as a Markov decision process in Section 3. Finally, Section 4 is dedicated to the solution procedure and sensitivity analysis from numerical examples.

2 The road cracking process

Road deterioration is essentially due to harsh environmental conditions and traffic loads. One of the most important deterioration modes is longitudinal cracking, as it represents the structural health of the road. Meyers *et al.* [5] reported that over 90 percent of road sections in Florida that are in need of repair have such cracks. The longitudinal cracking process operates as follows. A repetitive tensile stress is generated at the bottom of the asphalt layer, and leads the road's tensile strength to deteriorate over time. When the stress applied by the traffic load exceeds the tensile strength of the road, micro-cracks appear at weak spots in the base of the road. These micro-cracks then propagate through the inferior layers of the road until they reach the road surface to give way to surface cracks that continue to propagate. The road-cracking process is therefore a two-phase cumulative process. The first phase is the initiation phase, during which no deterioration can be observed on the surface, but the road cannot be considered to be in perfect condition. The second phase is the propagation phase, which begins with the arrival of the first observable crack.

The current metric used in France to measure the cracking level of a road section is the longitudinal cracking percentage (LCP), represented in Figure 1. Note that the LCP met-

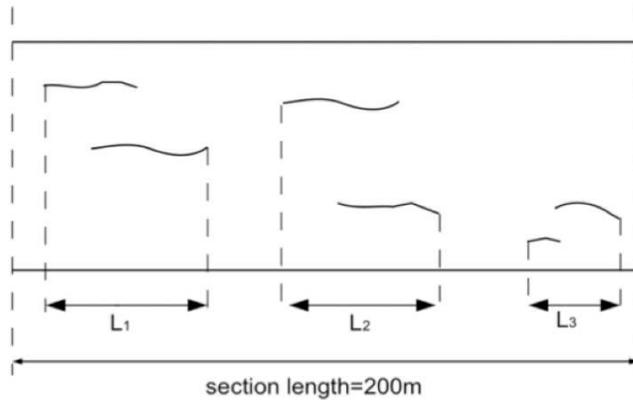


Figure 1: Longitudinal cracking percentage metric for 200m road section $LCP = \frac{L_1 + L_2 + L_3}{200}$

ric represents only the total longitudinal cracking, and does not take into account the number of cracks or overlapping cracks. Moreover, it is not suitable for the initiation phase where no cracks are observable. In reference [13], Zouch *et al.* have highlighted the benefits of introducing a new decision parameter, the deterioration growth rate (DGR), which models the underlying cracking process. More specifically, the DGR models a cracking potential in the first phase when the road is crack-free, and an instantaneous speed of deterioration in the second phase.

In this paper, the two-phase deterioration model developed in reference [13] is used. More specifically, the LCP and DGR are modelled as stochastic processes, denoted by $\{\rho_t\}_{t \geq 0}$ and $\{\theta_t\}_{t \geq 0}$, respectively. In the first phase, the cracking potential u and the first observable crack arrival are considered and modelled as a gamma process and a θ -dependent Poisson process, respectively. In the second phase, the DGR θ and the LCP ρ are modelled using a bilateral

gamma (BG) process and a state-dependent gamma (SDG) process, respectively. The proposed gamma process in reference [13] is called state-dependent, since its shape function in each decision epoch depends on the level of deterioration at its start, given by (ρ, θ) and on the last performed action. For more details of the deterioration model and the SDG process, the reader should refer to reference [13].

3 Model Formulation

We consider a road section, characterized by an initial thickness ℓ , that is continuously and stochastically deteriorating under the two-phase deterioration process [13] presented in the previous section. Periodic inspections are performed at the beginning of the decision epoch of length τ to yield perfect observations of the LCP ρ . Since θ cannot be measured, the problem is partially observed. However, the DGR θ can be approximated in the propagation phase using successive observations of ρ , and estimated in the initiation phase [13]. The problem can therefore be solved as a fully observable problem.

Based on the information yielded by inspection, the decision-maker should decide whether to do nothing (DN) and let the system deteriorate until the next decision epoch, or to maintain the section (MX). As the road thickness is constrained by a maximum thickness l_{max} , the maintenance consists of milling a thickness ζ before adding a new resurfacing layer of thickness ξ . Resurfacing the road section with a new layer conceals the longitudinal cracks, so that ρ is reset to zero, but also changes the thickness and the composition of the section, which makes it more rigid and reduces its DGR θ .

A maintenance decision (ξ, ζ) incurs an action cost $c_1 + c_2 + c_r(\zeta) + c_a(\xi)$, where c_1 and c_2 are the fixed costs of removing and adding, respectively, whereas $c_r(\zeta)$ and $c_a(\xi)$ are the variable costs of removing a thickness ζ and adding a thickness ξ .

Moreover, a quality cost $c_q(\cdot)$ is incurred for beginning a decision epoch in a given state (ρ, θ, ℓ) . This quality cost represents a penalty or a risk function for leaving the road in a deteriorated state for the current decision epoch.

Both decisions and maintenance actions are performed instantaneously at the beginning of each decision epoch, as decision and maintenance times are assumed to be negligible compared with the decision epoch length. The state of the road section is defined by $(\rho, \theta, \ell, \xi_m)$, where ρ and θ are the current LCP and DGR yielded by inspection, respectively; ℓ is the current thickness of the road section; and ξ_m is the last thickness added to the road section. Considering the last-performed maintenance action as a state parameter allows the decision-maker to know the current composition of the road, i.e. the old and new layer thicknesses. Therefore the state of the section is Markovian.

Let $S = [0, 1] \times [\theta_0, \theta_{max}] \times [0, l_{max}] \times [0, l_{max}]$ denote the state space. A DN action is assumed to have no impact onto the state of a road section. If a maintenance action $(\xi > 0, \zeta > 0)$ is selected, the state is transformed into $(0, \phi(s, \xi, \zeta), \ell + \xi - \zeta)$, where $\phi(s, \xi, \zeta)$ is the function that models the deterministic effect of the maintenance action on the DGR. Note that as maintenance actions may change the road section thickness and composition, it is possible to obtain a road that has better performance than the initial one. Thus it is possible to obtain a state better-than-new.

The objective is to derive an optimal maintenance policy p that minimizes the total discounted cost-to-go over the infinite horizon, denoted by V . Before presenting the MDP formulation, let $g(x; \rho, \theta, \ell, \xi_m)$ and $f(y; \rho, x, \theta, \ell, \xi_m)$ denote the density distributions of the LCP and DGR processes in the two phases, respectively, given the current levels of ρ and θ as well as the current section thickness ℓ and the last performed action type given by ξ_m .

More specifically, g represents the density of the Poisson process in the first phase and the density of the SDG process in the second phase, whereas f represents the density of the BG process in both phases. The joint density of the two deterioration parameters in both phases, denoted $h(x, y; \rho, \theta, \ell, \xi_m)$ can be expressed as follows:

$$\begin{aligned} h(x, y; \rho, \theta, \ell, \xi_m) &= g(x; \rho, \theta, \ell, \xi_m) f(y; \rho, x, \theta, \ell, \xi_m), \\ x &\in [0, 1 - \rho], y \in [-\theta, \theta_{max} - \theta] \end{aligned}$$

Explicit expressions of the density distribution functions g and f are given in reference [4]. Hence the MDP formulation of the maintenance optimization problem can be expressed as follows.

For $s = (\rho, \theta, \ell, \xi_m)$,

$$V^*(\rho, \theta, \ell, \xi_m) = c_0 + \min\{DN(\rho, \theta, \ell, \xi_m), MX(\rho, \theta, \ell, \xi_m)\} \quad (1)$$

where

$$DN(\rho, \theta, \ell, \xi_m) = \lambda \int_0^{1-\rho} \int_{-\theta}^{\theta_{max}-\theta} h(x, y; \rho, \theta, \ell, \xi_m) V(\rho + x, \theta + y, \ell, \xi_m) dy dx \quad (2)$$

and

$$MX(\rho, \theta, \ell, \xi_m) = \min_{(\xi, \zeta)} \left\{ c_1 I_{(\zeta > 0)} + c_2 I_{(\xi > 0)} + c_r(\zeta) + c_a(\xi) + DN(0, \phi(\rho, \theta, \ell, \xi_m, \xi, \zeta), \ell + \xi - \zeta, \xi) \right\} \quad (3)$$

subject to

$$\ell + \xi - \zeta < \ell_{max} \quad (4)$$

where $\lambda \in [0, 1]$ is the discount factor.

Note that the last added thickness jm is directly introduced as a parameter of the joint density because of its influence in the cracking process, and thus the associated definition of the road state. Equation (1) states that following the DN action when the current state is s incurs a quality cost plus the expected cost-to-go of the system deterioration from the state s to all possible states s' . The MX action in equation (2) incurs a maintenance cost, plus the quality cost and the cost-to-go of the the system beginning in the resulted state $(0, \phi(\rho, \theta, \ell, \xi_m), \ell + \xi - \zeta, \xi)$. The thickness constraint is formulated in (4).

4 Solution procedure and numerical examples

4.1 Solution procedure

We solve the MDP problem formulated above using the policy iteration algorithm (PIA) [7]. The following discretization of the decision problem is proposed. Assume that ρ and θ represent the first values in N_ρ and N_θ equal-sized discrete intervals of length Δ_ρ and Δ_θ on $[0, 1]$ and $[0, \theta_{max}]$, respectively, such that N_ρ and N_θ divide evenly into 100. As $\rho = 0$ is a key state, we consider it as a single state. Let N be the cardinality of the state space S , and $p_{ss'}^m$ denote the transition probability from state s to state s' when the last performed maintenance type is m .

$$\begin{aligned}
 p_{ss'}^m &= Pr\{s' | s\} \\
 &= Pr\{\rho' \in [i_{\rho'}, i_{\rho'} + \Delta_\rho], \theta' \in [i_{\theta'}, i_{\theta'} + \Delta_\theta] | \rho = i_\rho, \theta = i_\theta, \ell, \xi_m\} \\
 &= \int_{i_{\rho'} - i_\rho + \Delta_\rho}^{i_{\rho'} - i_\rho} g(x; i_\rho, i_\theta, \ell, \xi_m) \int_{i_{\theta'} - i_\theta + \Delta_\theta}^{i_{\theta'} - i_\theta} f(y; i_\rho, x, \ell, \xi_m) dy dx
 \end{aligned} \tag{5}$$

The set of possible maintenance actions in order to use the PIA is discretized as follows:

$$A = \{(\xi, \zeta) | \xi \in \{5, 10, 15, 20\}, \zeta \in [\min(0, \theta_{max} - \ell - \xi), \ell]\}$$

4.2 Characterization of the deterioration functions

We shall focus in this section on the construction of the deterioration function f for the ρ -process. The direct derivation of g for the BG process (θ_t) from the SDG law is proposed in reference [13]. The construction of these laws is expert-based knowledge on the cracking process rather than field data; one of the main motivations is the lack of confidence in the current French database for these longitudinal cracks. Recall that the ρ -process is modelled by an SDG process [13], where the deterioration function f on a time interval t is a gamma density function with a shape function, a function of the current state (ρ, θ) , and a given scale parameter β . Two choices of shape function for f are proposed here.

4.2.1 The shape function of the SDG process

For the cracking process (ρ_t) , two examples of the SDG process shape function are considered. The first shape function, γ_1 , has a linear form of the state parameters, whereas the second function, γ_2 , has a form inspired by the well-known Gaussian function:

$$\begin{aligned}
 \gamma_1(\tau; \rho, \theta, \ell, \xi_m) &= (a_0 + a_1\rho + a_2\theta + \frac{a_3}{\ell} + \frac{a_4}{\xi_m}) \\
 \gamma_2(\tau; \rho, \theta, \ell, \xi_m) &= \frac{b_1(\ell_{max} - \xi_m)^2}{\ell_{max} - \ell} \theta^2 \tau \exp\left\{\frac{(\rho - (1 - \theta))^2}{b_2(\ell_{max} - \xi_m)}\right\}
 \end{aligned}$$

Note that the two shape functions reflect both time and state effects, as they depend on the decision epoch length t as well as on the state parameters. However, they vary differently as

the state varies. Figure 2 shows the variations of the expected deterioration in ρ for different levels of θ given by the Gaussian form shape function. Note that at the beginning of the propagation phase the expected increase in deterioration increases as ρ increases, but with decelerated variation. However, when the DGR level becomes higher, the expected increase in cracking level increases and then decreases quickly, as the DGR is influential. This reflects the fact that, for a single road section, the probability of increasing the cracking level of the section is higher at the beginning of the propagation phase, and decreases as the section becomes more cracked (since overlapping cracks do not account for ρ). The main objective in considering these two shape functions is to highlight the state-dependent character of the deterioration model, as well as the effect of introducing the DGR as a deterioration parameter.

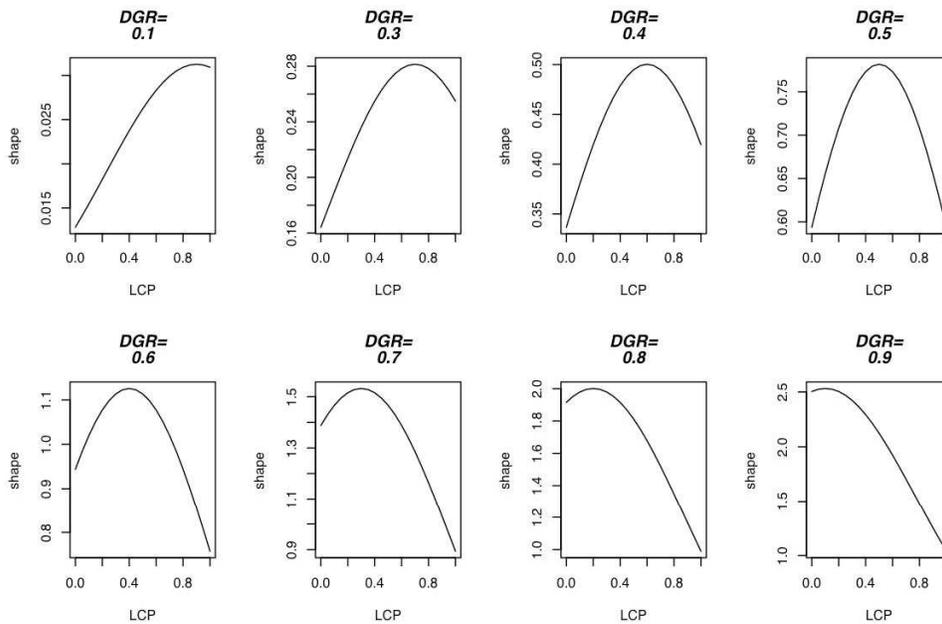


Figure 2: Variations of Gaussian form shape function in ρ for different levels of θ

Figure 3 reflects the fact that the expected crack growth obtained with the Gaussian form shape function increases with the DGR (a), and decreases as the total road thickness and the latest added thickness increase ((b) and (c)).

Figure 4 illustrates the BG process shape functions that represent the expected variation in the DGR θ as a function of the LCP. Note that the cracking process is accelerated for low LCP levels, i.e. the beginning of the cracking process, and is decelerated when the cracking level is high. This reflects the same behavior shown by the shape function γ_2 of a non-increasing probability of cracking propagation as the section becomes more cracked.

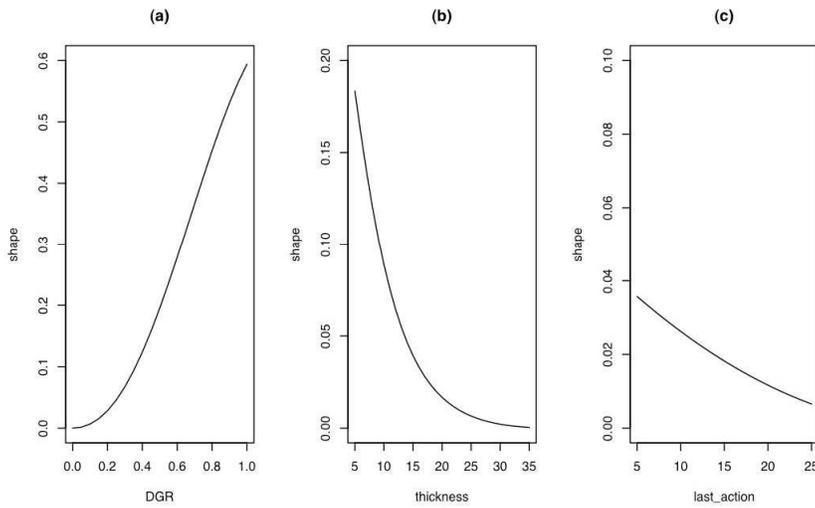


Figure 3: Variations of the Gaussian form shape function in θ , ℓ and ξ_m

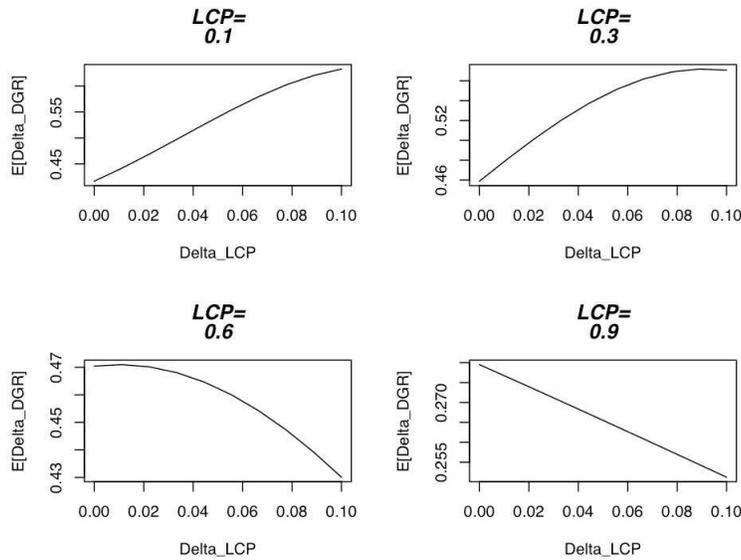


Figure 4: Variations of the DGR in increments of LCP for different levels of ρ

4.2.2 Comparison of the linear and Gaussian form shape functions

The main motivation behind comparing the linear and Gaussian form shape functions is to highlight the state-dependent character of the road-cracking process. In fact, in this model, the deterioration law changes in each decision epoch according to the road state at its beginning and the last performed maintenance action. Both linear and Gaussian form shape functions track this change, since their parameters are state and action dependent. The main difference between the two functions is the fact that, for a given DGR level, the linear shape function has the same variation rate for all LCP levels (the same slope), whereas the Gaussian form function has a variation rate that depends on the cracking level (the slope changes as the LCP changes) (Figure 2).

Moreover, as illustrated in Figure 2, for different values of θ , the linear shape function has a constant variation rate in ρ , whereas the Gaussian form function allows for different variation rates in ρ when the DGR level changes. These properties will be illustrated with the numerical analysis to show that the Gaussian form shape function takes future deterioration effects into account better, since it allows changing deterioration speeds for different cracking levels.

4.2.3 The crack occurrence rate function

The cumulative hazard function of the Poisson process of the initiation phase is given by the θ -dependent function $\alpha(\tau; \theta, \ell, \xi_m)$. Note that α is non-decreasing in the cracking potential θ and non-increasing in the total and new added thicknesses.

$$\alpha(\tau; \theta, \ell, \xi_m) = \left(d_0 + d_1\theta + \frac{d_2}{\ell} + \frac{d_3}{\xi_m} \right)$$

4.3 Numerical examples

In the following section, we present some numerical examples, and provide some analysis of the structural properties of the optimal policies.

4.3.1 Description of the case studies

We derive optimal policies for two types of section - sections with the current thickness less than the maximum thickness, and sections at maximum thickness - in order to show the effect of the thickness on the road performance as well, as to analyse the trade-off between thicker and newer roads. We consider three sections, which we denote Sec1, Sec2 and Sec3, and which have the following characteristics: $(\ell = 10, \xi = 10)$, $(\ell = 20, \xi = 10)$ and $(\ell = 30, \xi = 5)$, respectively, where each pair denotes the total section thickness ℓ and the newer layer thickness ξ .

For $\ell_{max} = 30$, consider the following set of available actions for each section:

$$\begin{aligned} A(\text{Sec1}) &= \{(\xi, \zeta) = \left\{ \begin{array}{l} (0, 0); (10, 5); (10, 10); (0, 5); (0, 10); (10, 15); \\ (0, 15); (10, 20); (0, 20); (5, 25); (10, 30) \end{array} \right\}\} \\ A(\text{Sec2}) &= \left\{ \begin{array}{l} (0, 0); (20, 5); (10, 5); (20, 10); (10, 10); (0, 5); (15, 15); \\ (0, 10); (20, 20); (5, 15); (10, 20); (15, 25); (20, 30) \end{array} \right\} \\ A(\text{Sec3}) &= \left\{ \begin{array}{l} (0, 0); (25, 5); (15, 5); (5, 5); (30, 10); (20, 10); (25, 15); \\ (10, 10); (30, 20); (15, 15); (20, 20); (25, 25); (30, 30) \end{array} \right\} \end{aligned}$$

Note that an action can decrease, increase or leave unchanged the total thickness of the road. The efficiency of a maintenance action is defined by a maintenance effect function that reflects the effect of different road compositions (i.e. new and old thick composition) on its performance. An example of a maintenance effect function is given by the following, where $(\rho, \theta, \ell, \xi_m)$ represents the state parameters just before maintenance, and (ζ, ξ) is the selected maintenance action.

$$\phi(\rho, \theta, \ell, \xi_m) = \begin{cases} 0.85A\theta & , \xi < \zeta, \xi \geq \xi_m \\ 0.75A\theta & , \xi < \zeta, \xi < \xi_m \\ 0.55A\theta & , \xi = \zeta, \xi \geq \xi_m \\ 0.45A\theta & , \xi = \zeta, \xi < \xi_m \\ 0.25A\theta & , \xi \geq \zeta, \xi \geq \xi_m \\ 0.1A\theta & , \xi \geq \zeta, \xi < \xi_m \end{cases}$$

where $A = \left(\rho + \frac{1}{(\ell + \xi - \zeta)^2} + \frac{2}{\xi} \right)$

Note that the maintenance actions in the proposed sets A(Sec1), A(Sec2) and A(Sec3) are ordered in terms of efficiency with respect to this function. We use different quality cost functions that are non-decreasing in deterioration parameters ρ and θ , and non-increasing in the total section thickness and in the last added layer thickness.

4.3.2 Numerical analysis

In the following, we present some optimal policies derived by the PIA for the three considered sections (Sec1, Sec2 and Sec3) using the different input functions presented above.

The deterioration parameters used for the numerical examples are given by

$$(a_0, a_1, a_2, a_3, a_4) = (1, 1.5, 1.2, 1, 1), \quad (b_1, b_2) = (1.2, 2)$$

$$(d_0, d_1, d_2, d_3) = (1, 1.2, 1, 1)$$

The set-up maintenance costs are

$$c_1 = 900, c_2 = 1700$$

$$c_r(\zeta) = c_{ur}\zeta \text{ with } c_{ur} = 90 \text{ the milling cost per thickness unit,}$$

$$c_a(\xi) = c_{ua}\xi \text{ with } c_{ua} = 210 \text{ the resurfacing cost per thickness unit.}$$

The policies presented in Tables 1-6 state, for each possible road section, the optimal maintenance action to perform. The policies are called Gaussian form and linear form with respect to the shape function used. For example, if the observed section state is ($\rho = 0.4, \theta = 0.7, \ell = 20, \xi_m = 10$), then the action advised by the Gaussian-form policy (Table 1(a)) is resurfacing the road section with a 10 cm layer without milling, whereas the policy advised by the linear-form policy (Table 1(b)) is resurfacing the section with a 5 cm layer without milling.

The discount factor used for the following numerical examples is 0.95. However, in order to highlight the difference between the two shape functions, some policies with different discount factors are compared. It is found that Gaussian-form policies with a low discount factor (0.3) may converge to linear-form policies obtained with higher discount factors (0.8). This shows that the Gaussian-form shape function takes into account the effects of future deteriorations better than the linear-form function does.

Tables 1-6 present decision matrices that detail the optimal maintenance action for each possible observed state. For example, from Table 1(a), if after inspection Sec2 is found with $\rho = 0.3$ and $\theta = 0.5$, then the optimal action given by the decision matrix for this state is to resurface the section with a thickness 5.

The obtained policies present some monotonic properties. More specifically, the efficiency of the recommended action is increasing in the deterioration level and the optimal policy is of the control limit type. For example, in Table 2(a), the DN action is the optimal decision for weak degradation levels (bottom left-hand side of the matrix). When degradation is growing (from the bottom left-hand to the upper right-hand side of the decision matrix), the recommended actions become stronger in terms of costs and efficiency.

The Gaussian-form policies promote predictive policies, whereas the linear form policies do not advise maintenance when the crack level is zero (see the first line of the two matrices in Table 1). This proves that the Gaussian-form shape function takes the effects of the DGR as a deterioration parameter into account better. To show the sensitivity of the optimal policies to the section thickness as well as to its composition, the policies of Tables 1 and 2 are compared, as follows.

1. For thinner sections such as Sec1, note that only actions that increase the section thickness are optimal (Table 2(a)).
2. When the sections are at their maximum thickness, it is not optimal to reduce their thickness (Table 2(b)). In fact, in this example (i.e. given the maintenance effect and cost parameters), actions that completely or partially renew sections with maximum thickness and reduce their thickness are less efficient than actions that keep the section thicknesses unchanged.
3. When the road section has a 'medium' thickness, there is a trade-off between making the section thicker or making it newer.

Results 1 and 2 confirm expert judgement stating that although the performance of a road is increasing in its thickness, it is not optimal to build the road at it is maximum thickness from the beginning; it is better to add new layers gradually. Moreover, results 1 and 2 indicate

Table 2 Optimal Gaussian-form policies with linear quality cost function

(a) Sec1										
ρ	θ									
	0.03	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.8	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.7	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.6	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.5	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.4	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)
0.3	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)
0.2	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)
0.1	(0, 0)	(0, 0)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)
0.03	(0, 0)	(0, 0)	(0, 0)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)
0.0	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)

(b) Sec3										
ρ	θ									
	0.03	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(30, 30)
0.8	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)
0.7	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)
0.6	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)
0.5	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)
0.4	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)
0.3	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)	(10, 10)
0.2	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)	(10, 10)
0.1	(0, 0)	(0, 0)	(0, 0)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(10, 10)
0.03	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)	(5, 5)
0.0	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(5, 5)	(5, 5)	(5, 5)	(5, 5)

that a control limit in thickness exists under which it is never optimal to reduce the section thickness.

Note that for given unit resurfacing and removing costs, some actions are never selected. This is because action efficiency is determined by the maintenance effect function. If the efficiency order is different from the total cost order, more expensive and less efficient actions will never be selected.

The trade-off between thicker and newer roads (3) is determined by the cost rates. Note that, especially for medium thickness sections, the policies are very sensitive to both maintenance and quality cost variations. The sensitivity of the model to the cost functions is illustrated by Tables 3 and 4, which present policies obtained using the Gaussian-form (a) and linear-form (b) shape functions with different quality cost functions. The policies in Table 3 are obtained using a quality cost function that increases more quickly in θ , whereas the policies in Table 4 are obtained using a quality cost function that increases more quickly in ρ .

From the numerical results, the policies are very sensitive to the variations of the quality cost function, especially the Gaussian-form policies. It is found that the Gaussian-form policies are much more sensitive to the quality cost variations in θ than the linear-form ones. For example, in Table 3, for $\rho > 0$, when θ increases from 0.4 to 0.5, the Gaussian-form policy (a) advises a stronger action than the linear-form one (b), which gives the same action. This confirms the fact that the Gaussian-form shape function takes the effects of the DGR into account better. Note that the actions in Table 3 are ones that make the road thicker, whereas most of the actions in Table 4 keep the section at the same thickness. This reflects the fact that to make roads more rigid, i.e. with lower DGR levels, it is better to make them thicker (Table 3). However, if costs are more sensitive to the LCP variations, it is better to conceal cracks by reducing or keeping the same thickness (Table 4). When the quality cost is not very important compared with the action costs, we note that DN action may become a better choice than maintenance actions that reduce the section thickness, especially for low levels of DGR.

Moreover, Table 4 shows that Gaussian-form policies are more sensitive to the quality cost variations in ρ than linear-form policies. For example, in Table 4 for $\theta > 0.5$, when ρ increases from 0.1 to 0.2, the Gaussian-form policy (a) advises stronger actions whereas the linear-form one (b) gives the same action. This can be explained by the fact that the Gaussian-form shape function represents the effect of the LCP variations on the evolution of the deterioration process better, and therefore reflects the deterioration evolution risk better. Thus the quality cost is a very important function in this model, as it represents a measure of cumulative risk for the decision epoch.

In Tables 5 and 6, the extended model is compared with the model in reference [13]. However, for the model of reference [13], only resurfacing actions that respect the total maximum thickness constraint are considered. Tables 5 and 6 present policies for low- and high-quality costs compared with maintenance costs, respectively. Note that for quality costs that are less important than maintenance costs, the two approaches converge to the policies presented in Table 5. These two policies have almost the same average values, the extended model allows the section to be renewed to the maximum thickness for maximum deterioration levels. When the deterioration risk is high, i.e. the quality cost is much more important than maintenance

costs, the two approaches give different policies (Table 6), but the difference in the policy average values is not very significant, i.e. $V(a) = 1.05V(b)$. This shows that the main interest from extending the possible maintenance actions to removing as well as adding different thicknesses is to improve prevention of the cracking risk.

5 Conclusion

This paper presents a condition-based maintenance optimization approach for the road-cracking problem based on a Markov decision process. This model extends a previous work [13] by taking into account the constraint of maximum road thickness. The special feature of this approach is that the MDP transition matrix is obtained using a state-dependent deterioration process based on two deterioration parameters, namely the LCP and the DGR, as well as on the road characteristics: the total thickness and the composition in new and old underlying layers. The optimal decision consists in defining both the optimal thickness to remove and the resurfacing thickness to add. This changes the composition of the underlying layers with different mixes of degraded and non-degraded layers and, finally, different cracking patterns. Moreover, we introduce the possibility of actions that can which lead to a state better than new. This approach can be implemented directly in a decision tool for designing new roads. The model provides the optimal maintenance costs, given the initial road reliability performance (here defined as a function of the initial road thickness). We have already highlighted the trade-off in initial maximum thickness and maintenance costs.

Moreover, numerical examples show the importance of defining and estimating cost parameters and functions, especially the quality cost function. In fact the quality cost function represents a measure of deterioration risk, and is very important in determining the optimal maintenance to perform. Finally, we provide different numerical analyses of the optimal maintenance policies for different road sections which lead to the identification. All of them lead to the identification of some structural.

Table 4 Optimal policies with quadratic increasing quality cost in ρ for Sec2

(a) Gaussian form										
ρ	θ									
	0.03	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(5, 15)	(10, 20)	(10, 20)	(10, 20)
0.8	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(10, 20)	(10, 20)	(10, 20)
0.7	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(5, 15)	(10, 20)	(10, 20)
0.6	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(5, 15)	(10, 20)	(10, 20)
0.5	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(15, 15)	(5, 15)	(10, 20)
0.4	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(15, 15)	(15, 15)	(5, 15)
0.3	(20, 5)	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(0, 10)
0.2	(20, 5)	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(15, 15)	(15, 15)	(0, 10)
0.1	(20, 5)	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(0, 5)
0.03	(20, 5)	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(0, 5)
0.0	(0, 0)	(20, 5)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(10, 10)	(0, 5)

(b) Linear form										
ρ	θ									
	0.03	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(5, 15)	(10, 20)	(10, 20)	(10, 20)
0.8	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(10, 20)	(10, 20)	(10, 20)
0.7	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(5, 15)
0.6	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.5	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.4	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.3	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.2	(20, 5)	(20, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.1	(20, 5)	(20, 5)	(0, 5)	(0, 5)	(0, 5)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.03	(0, 0)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)	(0, 10)
0.0	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 10)	(0, 10)	(0, 10)

References

- [1] D. M. Frangopol, M. J. Kallen, and J. M. van Noortwijk. Probabilistic models for life-cycle performance of deteriorating structures: review and future directions. *Structural Engineering and Materials*, 6:197–212, 2004.
- [2] N. Ismail, A. Ismail, and R. Atiq. An overview of expert systems in pavement management. *European Journal of Scientific Research*, 30(1):99–111, 2009.
- [3] M. J. Kallen and J. M. van Noortwijk. Optimal periodic inspection of a deterioration process with sequential condition states. *International Journal of Pressure Vessels and Piping*, 83(4):249 – 255, 2006. The 16th European Safety and Reliability Conference.
- [4] L. M. Maillart. Maintenance policies for systems with condition monitoring and obvious failures. *IIE Transactions*, 38:463–475, 2006.
- [5] L. Myers, R. Roque, and B. Ruth. Mechanisms of surface-initiated longitudinal wheel path cracks in high-type bituminous pavements. In *Proceedings of Asphalt Paving Technology Conference*, volume 67, pages 401–432, 1998.
- [6] P. Paris and F. Ergodan. A critical analysis of crack propagation laws. *Journal of Basic Engineering*, 85(4):528–534, 1963.
- [7] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley and Sons Inc., New York, 1994.
- [8] S. G. Ritchie, C. Yeh, J. P. Mahoney, and N. C. Jackson. Surface condition expert system for pavement rehabilitation planning. *Journal of Transportation Engineering*, 113(2):155–167, 1987.
- [9] M. Y. Shahin. *Pavement Management For Airport, Roads, and Parking Lots*. Chapman and Hall, New York, 1994.
- [10] M. Y. Shahin and J. A. Walter. Pavement maintenance management for roads and streets using the paver system. Technical Report technical report M-90/05, US Army Corps of Engineers (USA CERL), Champaign II, 1990.
- [11] J. M. van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1):2 – 21, 2009. Maintenance Modeling and Application.
- [12] H. Wang. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3):469 – 489, 2002.
- [13] M. Zouch, T. G. Yeung, and B. Castanier. Two-phase state-dependent deterioration model for maintenance optimization. *Submitted to Naval Research Logistics*, 2011.

**OPTIMAL MAINTENANCE AND REPLACEMENT DECISIONS
UNDER TECHNOLOGICAL CHANGE
WITH CONSIDERATION OF SPARE PARTS INVENTORIES**

**P.Khanh Nguyen Thi, phuong-khanh.nguyen-thi@emn.fr
Department of Industrial Engineering & Automatic control
École des Mines de Nantes / IRCCyN
Nantes 44307, France**

**Thomas G. Yeung*, thomas.yeung@emn.fr
Department of Industrial Engineering & Automatic control
École des Mines de Nantes / IRCCyN
Nantes 44307, France
+33 (0)2 51 85 86 45**

**Bruno Castanier, bruno.castanier@emn.fr
Department of Industrial Engineering & Automatic control
École des Mines de Nantes / IRCCyN
Nantes 44307, France
+33 (0)2 51 85 83 12**

*Corresponding Author

ABSTRACT: Classical spare parts inventory models assume that the same vintage of technology will be utilized throughout the planning horizon. However, replacement often occurs in the form of a new technology that renders existing spare parts inventories obsolete. This paper aims to study the impact of spare parts inventory on maintenance and replacement decisions under technological change via a Markov decision process formulation. The replacement decision is complex in that one must decide with which technology available on the market to replace the current asset. Under technological change, the do nothing and repair options have significantly more value as they allow the appearance of even better technologies in the future.

Key words: technology evolution, spare parts, equipment replacement, maintenance, Markov decision process

1. Introduction

Maintenance is key in ensuring the efficient use of equipment as well as an efficient production process. If assets often break down or operate in bad condition, this may cause lower product quality, increased energy consumption and reduced revenue. Therefore, the objective of maintenance is not simply to overcome failures, but to predict and prevent revenue loss at the management level. Managers must analyze all relevant information to assess the profitability of equipment, give sound investment decisions, and consider possible cost saving. In particular, high spare parts inventory is one important factor causing increased maintenance costs. Lack of inventory increases asset downtime due to waiting time when necessary spare parts are not available. Hence, consideration of the spare parts inventory problem is an essential task of managers.

There has been intensive research to study the different aspects of spare parts inventory problems such as management issues, multi-echelon problems, age-based replacement, repairable spare parts, problems involving obsolescence, etc. A review of the spare parts inventory's problem is presented in [13]. As the authors comment, spare parts inventories totally differ from other manufacturing inventories. Their function is to assist maintenance staff in keeping equipment in operating condition. The close relation between spare parts inventories and maintenance has been discussed in several articles. [11] studies the joint-optimization of age-based replacement and spare parts inventory policy (S, s) . Under a block replacement strategy, [24] utilizes dynamic programming to solve the spare parts ordering problem while [3] and [22] present management policies for a manufacturing system, aiming to optimize the maintenance strategy with continuous review spare items inventory jointly. [4] and [23] extend the problem by also studying minimal repair for minor failures while [8] considers repair capacity of degraded/failed units after they were replaced by spare parts.

However, all the above models are constructed on the assumption that the same vintage of technology will be utilized throughout the planning horizon. They do not allow us to take into account the appearance of new technology with performance improvement. This information is very important for managers who often confront the technology investment decision. They must weigh the benefits of utilizing the available equipment with their stock of spare parts and the revenues of investment in new technology. Nevertheless, there are a few articles that take into account the technological development in the spare parts inventory problem. Those that do are generally based on the introduction of an economical loss when

new technology appears by a cost of obsolescence. [14] does not explicitly consider an obsolete cost. They include it in the holding cost in a multi-echelon system. [7] develops models that can be seen as extensions of the EOQ formula for fast moving spare parts subject to sudden obsolescence risk. The authors examine the effects of obsolescence on costs under several different conditions: constant obsolescence risk, no shortages allowed; varying obsolescence risk, no shortages allowed and varying obsolescence risk, shortages allowed. However, they relax the interactive relation of spare parts inventory and maintenance strategy by considering a constant rate demand until the moment of obsolescence.

On the other end of the spectrum, models devoted to maintenance optimization involving technological change do not consider the spare part inventory impact [2, 5, 6, 9, 10, 12, 15]. This gap in the literature motivates us to develop an appropriate model to meet management's requirements: optimization of maintenance cost while simultaneously updating information on technological development and considering the impact of spare parts inventory levels to make sound investment decisions. With this model, we do not consider inventory optimization but focus on examining how spare parts inventory levels will influence the replacement decision and how much better a new technology must be in order to overcome the obsolescence of existing spare parts inventory.

We formulate a discrete-time, non-stationary Markov decision process (MDP) to determine the optimal action plan. To model technological evolution, we combine the uncertain appearance model wherein technology change is characterized by the uncertain arrival time of new technology and the geometric model. The geometric model presented by [1, 2, 12, 21, 18, 19, 20] is a model where geometric functions are utilized to characterize technology change, such as the maintenance/operation cost functions in vintage equipment or in time. Unlike these articles, we present technology change by the improvement of the expected deterioration rate and the expected profit function within a period. We also consider non-stationary likelihood of new technology's apparition over time. Thereby, we overcome the disadvantages of the geometric model proposed by [2]. Recall [16, 17] also consider the non-stationary probability of the appearance of new technologies. But in their model, [16, 17] focus on the problem of capital investment decisions due to technological change rather than physical deterioration of equipment. To simplify its exposition, he also does not consider salvage values while we establish a reasonable salvage value function which is proportional to the current purchase price of technology vintage, decreasing in the remaining life time and incurs an even greater reduction when it becomes obsolete due to new technology appearances.

The remainder of this paper is structured as follows. In Section 2, we present our mathematical formulation model and its assumptions. Next, in Section 3, the performance of our model is discussed through numerical examples. Finally, conclusions and future work are discussed in Section 4.

2. Model and Assumptions

2.1 Problem Statement

Consider a repairable asset which is accompanied by a cargo of n spare parts. They are utilized for the maintenance of the asset and are not sold separately in the market, i.e., we cannot replenish the spare parts store when it is empty. This is a common assumption for special spare parts because it can be difficult and costly to find original and compatible spare units. In making this assumption we do not consider the optimization of the inventory policy. Our primary goal is to study the impact of the spare parts inventory level on maintenance and replacement decisions under technological change rather than determining what should be the optimal order level/order quantity for the spare parts.

The asset operates continuously from the new state $x = 0$ and is characterized by its expected deterioration rate. In the failure state, denoted m , the asset continues to operate, but unprofitably. To reveal the deterioration and the spare parts inventory level, periodic inspections are performed. The inter-inspection interval τ defines the decision epochs.

We assume that only one new technology can appear in a decision interval τ . We introduce p_{i+1}^{k+1} as the non-stationary probability that technology $k + 1$ appears in the interval τ given the latest available technology at decision epoch i is k . This probability is normally non-decreasing in time. The difference in technological generations is modeled by an improvement factor on the expected instantaneous deterioration rate and the accrued profit within a decision period.

Let (x, k, j, s) be the system state of the system at the beginning of the i^{th} decision epoch with s spare parts in stock for maintenance of the asset having deterioration level x . The asset in use belongs among the technological generation j while the latest available technology in the market is k , $k \geq j$. In each decision epoch the possible actions are:

- 1) Do nothing: The asset continues to deteriorate until the next decision epoch and generates a profit $g_j(x)$. Note that $g_j(x)$ is the expected profit within a period when the deterioration state at the beginning of that period is x and the utilized technology is j . The

spare parts inventory level is not changed, so the holding cost within this period i is sc_s where c_s is the holding cost per spare part unit in a decision period.

2) Imperfect maintenance restores the asset to a given deterioration level, $\max(0, x - d)$ where d models the maintenance efficiency. An imperfect maintenance cost, c_1 is incurred and θ spare parts are utilized to replace the degraded units, thus the spare part inventory level is reduced by: $s - \theta$. As we assume that the maintenance time is negligible, in the next decision interval the asset deteriorates from the level $\max(0, x - d)$ and generates a profit $g(\max(0, x - d))$.

3) Perfect maintenance restores the asset to the initial deterioration level $x = 0$ (as-good-as-new) and the expected profit within next decision interval is $g(0)$. This action requires $\eta(x)$ spare parts that depends on the deterioration state of the asset. A perfect maintenance cost c_2 is also incurred ($c_2 > c_1$).

Note that as we assume the spare parts are only supplied when we buy a new asset, we can perform maintenance actions if and only if there are sufficient spare parts in stock.

4) Replace the asset by an available technology h in the market ($j \leq h \leq k$). The replacement time is also considered negligible. A cargo of n spare parts is supplied with the new asset. We assume the spares are only compatible with the same generation asset, hence, after replacement, the spare parts inventory level is n if we decide to replace by a newer generation asset and equals $n + s$ in the case of replacement by same-generation asset. The cost of such a replacement is given by the difference between the purchase price of the new asset $c_{i,h}$ and the salvage value $b_{i,jk}(x)$ of the existing asset. Note that the purchase price of a new technology asset can be estimated as well as the deterioration rates. This is realistic in cases where the technical parameters and specifications of future designs may be assumed reasonably well beforehand. To capture the characteristics of technology change, we assume that the purchase price of a new asset is decreasing over time after appearance and generally increasing over technological generation while the salvage value of the exiting asset is a function of its current purchase price, its mean residual life and the difference between technological generations. After replacement, the new asset generates a profit $g_h(0)$.

Note that we could easily take into account the salvage value of the remaining spare parts in our model. However, this salvage value is not important compared to the other values, especially in the obsolescence case with the common assumption that the spares are only

compatible with the same generation asset. Hence its influence on the optimal decision can be negligible and we do not consider it.

2.2 Model Formulation

In this paper, we use a non-stationary MDP formulation to find the optimal maintenance-replacement policy to maximize the expected discounted value-to-go over the finite horizon time denoted by $V^\pi(x, k, j, s)$.

Let $V_i(x, k, j, s)$ denote the maximum expected discounted value from the decision epoch $i, k \leq i$ to the last epoch N . Then, $V_i(x, k, j, s) = V^\pi(x, k, j, s)$.

We define $DN_i(x, k, j, s)$, $M_i(x, k, j, s)$, $PM_i(x, k, j, s)$, $R_i^h(x, k, j, s)$ to be the value functions of the Do Nothing, Imperfect Maintenance, Perfect Maintenance and Replacement options at the decision epoch i , when the system state is (x, k, j, s) . $V_i(x, k, j, s)$ is the maximal value of the option set $\{DN_i(x, k, j, s), M_i(x, k, j, s), PM_i(x, k, j, s), R_i^h(x, k, j, s)\}$. Note that the maintenance actions are realized if and only if there are sufficient spare parts in stock.

$$V_i(x, k, j, s) = \max \left\{ \begin{array}{l} DN_i(x, k, j, s), IM_i(x, k, j, s)1_{\{s>\theta\}}, \\ PM_i(x, k, j, s)1_{\{s>\eta(x)\}}, R_i^h(x, k, j, s) \end{array} \right\} \quad (1)$$

where:

$$DN_i(x, k, j, s) = g_j(x) - sc_s + \lambda \left[\sum_{\forall k' \in \{k, k+1\}} p_{i+1}^{k'} \sum_{\forall x' \in [x, m]} p_j(x'|x) V_{i+1}(x', k', j, s) \right] \quad (2)$$

$$IM_i(x, k, j, s) = -c_1 + DN_i(\max(0, x - d), k, j, s - \theta) \quad (3)$$

$$PM_i(x, k, j, s) = -c_2 + DN_i(0, k, j, s - \eta(x)) \quad (4)$$

$$R_i^h(x, k, j, s) = b_{i,j,k}(x) - c_{i,h} + DN_i(0, k, j, s + n)1_{\{h=j\}} + DN_i(0, k, j, n)1_{\{h \neq j\}} \quad (5)$$

λ is a discount factor, $\lambda \in [0, 1]$.

The value function of Do Nothing, $DN_i(x, k, j, s)$ is calculated from two parts. The first part is the reward action in the current period, given by the difference between the expected profit $g_j(x)$ and holding cost of spare parts c_s . The second part is the discounted revenue acquired by asset in future with discount factor λ , the transition probability of deterioration state $p_j(x'|x)$ and the appearance probability of next technology generation, p_{i+1}^{k+1} (when $k' = k + 1$). Note that if $k' = k$, next technology generation will not appear at next epoch with probability $p_{i+1}^k = 1 - p_{i+1}^{k+1}$.

The value function of imperfect/perfect maintenance or replacement is given by the action cost (imperfect/perfect maintenance or replacement cost) and the value acquired by

system after this action. Note that the value acquired by system after maintenance or replacement action is the value function $DN_i(\cdot)$ of the system state that is affected by the respective action.

3. Numerical Examples

In the previous section, we have developed a general model to find the optimal maintenance/replacement policy under technological change and also examine the impact of spare parts inventories on this policy. Now we apply this model to a k-out-of-n system for numerical illustration. The assumptions given in this section are reserved for the numerical examples to illustrate the performance of our general model and are not in anyway necessary for the general model.

3.1 Input Parameters

3.1.1 Probability of appearance of new technology

We define the appearance probability of new technology $k + 1$ at decision epoch $i + 1$, given the latest available technology at decision epoch i is k , as a time increasing function:

$$p_{i+1}^{k+1} = 1 - \delta \varepsilon^{i-k} \quad (8)$$

δ is the factor that reflects the non-appearance probability of the next generation at the next decision epoch when $k=i$. And the factor ε characterizes the increasing rate of the appearance probability of new technology over time. We have $\delta, \varepsilon \in [0, 1]$.

3.1.2 Deterioration Process

We consider a k-out-of-n system, specifically 2-out-of-5, where identical independent components have time to failure T that follows an exponential distribution with parameter α_j (j is the technological generation of system). The improvement of α_j is discussed in the next paragraph. We choose $\tau = 2$ weeks.

For a new system, all components are functional and the deterioration state of system is $x = 0$. When one component fails, the system degrades one unit, and it fails when at least 4 items fail; the system failure state is denoted by m ($m = 4, 5$). To perform IM, we utilize one spare part ($\theta = 1$) to replace the failed component and restore the system to the previous deterioration state ($d = 1$). With perfect maintenance, we replace all failed components

$\eta(x) = x$ to restore the system to its initial state. The transition probability of the deterioration state of the system is given by:

$$p_j(x'|x) = \binom{N-x}{x'-x} \left(1 - \exp(-\alpha_j \tau)\right)^{x'-x} \left(\exp(-\alpha_j \tau)\right)^{N-x} \quad (9)$$

$\forall x, x' \in X: \{0, 1, 2, 3, 4, 5\}$

As the components have an exponential life distribution, their conditional survivor function at age y is equal to the survivor function of a new component. If we do not consider the repair capacity of the system, the survivor function at time y , denoted by $R(x, t)$ depends only on the current deterioration state. (We have $R(m, t) = 0$).

$$R(x, t) = \sum_{z=k}^{N-x} \binom{N-x}{z} \left(\exp(-\alpha_j t)\right)^z \left(1 - \exp(-\alpha_j t)\right)^{N-x-z} \quad (10)$$

Therefore, the mean residual lifetime, $MRL(x)$ of system is:

$$MRL(x) = \int_0^{\infty} R(x, t) dt \quad (11)$$

3.1.3 Impact of Technological Evolution

We model the impact of technological evolution on the expected failure rate α_j of items with the following decreasing geometric function:

$$\alpha_j = ab^{-(j-1)} \quad (12)$$

Where a, b are constant. We choose arbitrarily values for a, b in Table 1 below.

The profit function is also improved by technological generations. Moreover, we know that the asset will operate less efficiently when its deterioration state is greater. Therefore, the expected profit function within a decision interval τ is decreasing by deterioration state and the greater the deterioration state is, the faster the decrease of the profit function. To reflect these characteristics, we use the following function to characterize the accrued profit.

$$g_{i,j} = (g_0 - a_1 \exp(x)(a_2 - \exp[r_1(1-j)])) \quad (13)$$

Where a_1, a_2, g_0, r_1 are constants, chosen arbitrarily in Table 1. In case of failure, the Do Nothing action is still allowed, but the profit in the next decision epoch $g(m)$ is assumed to be zero.

Additionally, under technological evolution, the purchase price of a new asset is assumed to be decreasing over time after appearance and normally increasing over technological generation as follows:

$$c_{i,k} = c_{1,1} v^{i-1} u^{k-1} \quad (14)$$

Where $c_{1,1}$ is the purchase price of the first of technological generation at the first decision epoch; v and u are constants characterizing the change of the purchase price over time and over technological generation. We choose arbitrarily $c_{1,1}, v, u$ in Table 1.

We assume the salvage value is a function of the current purchase price of the technology, the difference between technological generation and the MRL. Thus, we propose the following function for the salvage value, $\forall x \in [0, 5]$; h, r are constant, chosen arbitrarily in Table 1.

$$b_{i,j,k}(x) = hr^{k-j}c_{i,j} \frac{MRL(x)}{MRL(0)} \quad (15)$$

Table 1. The input parameters for the Example 1

Appearance probability	δ	ε	λ		
Discount factor	0.99	0.999	0.98		
Profit	g_0	r_1	a_1	a_2	
	415	0.05	15	1.5	
Maintenance, Holding cost, Failure rate	c_1	c_2	c_s	a	b
	80	120	1	2	0.9
Salvage value & Purchase price	h	r	$c_{1,1}$	v	u
	0.7	0.8	50	0.998	1.2

3.2 Numerical Analysis

For the numerical examples, we consider only five technological generations. We also assume the storage capacity for spare parts is limited to a maximal storage capacity of 20 units and the cargo of spare parts accompanied with new asset comprises 15 units. The optimal policy for the first ten decision epochs of Example 1 in cases where $k = 1, 2, 3$ is given in Table 2. Note that with time horizon $N \geq 190$, this policy does not change.

With the used technology j , given that generation k is the latest available technology, the optimal decision is defined according to the current deterioration x and stock level s . Each column in Table 2 represents a combination of k and j given by the vector (k, j) in the header. For each vector (k, j) the different combinations of stock level and system state are given in the left and right sides of the columns, respectively. The optimal action is given for each system state at each stock level.

We find that in the cases where there is no obsolescence problem i.e., $(k, j) = (1, 1); (2, 2); (3, 3)$; the replacement decision is made only at low stock levels ($s < 5$). For example, with $(k, j) = (1, 1)$ at stock level $s = 0$ when we cannot perform maintenance, thus the optimal policy prescribes replacement (R) for deterioration states greater than $x = 2$. This replacement threshold is non-increasing in the used technology (j) because the profit function and the failure rate of components are improved over technological generations, and the replacement option has more value than the do nothing option.

Table 3: The optimal policy for Example 1.

$(k, j) = (1, 1)$	$(k, j) = (2, 1)$	$(k, j) = (2, 2)$	$(k, j) = (3, 1)$	$(k, j) = (3, 2)$	$(k, j) = (3, 3)$
s Action	s Action	s Action	s Action	s Action	s Action
0 $x = 0, 1: DN$ $x \geq 2: R$	0 $x = 0: DN$ $x \geq 1: R$	0 $x = 0: DN$ $x = 1: R (i \leq 5)$ $DN (i \geq 6)$	$[0, 2] \forall x: R$	0 $x = 0: DN$ $x \geq 1: R$	0 $x = 0: DN$ $x \geq 1: R$
1 $x = 0: DN$ $1 \leq x \leq 4: IM$ $x = 5: R$	1 $x = 0: DN$ $x = 1: R (i = 2, 3)$ $IM (i \geq 4)$ $x \geq 2: R$	1 $x = 0: DN$ $1 \leq x \leq 4: IM$ $x = 5: R$	$[3, 4] 3 \leq x \leq s: PM$ $\forall x \in X \setminus [3, s]: R$	1 $x = 0: DN$ $1 \leq x \leq 3: IM$ $x = 4, 5: R$	1 $x = 0: DN$ $1 \leq x \leq 4: IM$ $x = 5: R$
$[2, 4] x = 0, 1: DN$ $2 \leq x \leq s: PM$ $s < x < 5: IM$ $x = 5: R$	$[2, 4] x = 0, 1: DN$ $2 \leq x \leq s: PM$ $x > s: R$	$[2, 4] x = 0, 1: DN$ $2 \leq x \leq s: PM$ $s < x < 5: IM$ $x = 5: R$	$s \geq 5 x \leq 2: R$ $x \geq 3: PM$	$[2, 4] x = 0, 1: DN$ $2 < x \leq s: PM$ $s < x < 4: IM$ $x = 4, 5: R$	$[2, 4] x = 0, 1: DN$ $2 \leq x \leq s: PM$ $s < x < 5: IM$ $x = 5: R$
$s \geq 5 x = 0, 1: DN$ $x \geq 2: PM$	$[5, 10] x = 0, 1: DN$ $x \geq 2: PM$ $s = 11 x = 0:$ $R (i \leq 4)$ $DN (\forall i \geq 5)$ $x = 1:$ $R (i \leq 3)$ $DN (\forall i \geq 4)$ $x \geq 2: PM$ $s = 12 x = 0:$ $R (i \leq 8)$ $DN (\forall i \geq 9)$ $x = 1:$ $R (i \leq 6)$ $DN (\forall i \geq 7)$ $x \geq 2: PM$ $s \geq 13 x = 0, 1: R$ $x \geq 2: PM$	$s \geq 5 x = 0, 1: DN$ $x \geq 2: PM$		$[5, 13] x = 0, 1: DN$ $x > 2: PM$ $s = 14 x = 0: R$ $x = 1:$ $R (i \leq 3)$ $DN (\forall i \geq 4)$ $x \geq 2: PM$ $s \geq 15 x = 0, 1: R$ $x \geq 2: PM$	$[5, 8] x = 0, 1: DN$ $x \geq 2: PM$ $s \geq 8 x = 0: DN$ $x = 1: IM$ $x \geq 2: PM$

As illustrated by Table 2, the replacement threshold is $x = 2$ for $j = 1$ and $x = 1$ for $j = 3$. When the stock level $s > 0$, imperfect/perfect maintenance (*IM* / *PM*) is performed by using spare parts to restore the asset to the previous/initial deterioration state. Consider the case $s = 2$, the policy dictates the *DN* action when $x = \{0,1\}$; *PM* when $x = 2$; *IM* with $x = \{3,4\}$; and finally replace with new asset at $x = 5$. With a high stock level, *PM* demonstrates its dominance and replacement is not necessary. For example, with $(k, j) = (2, 2)$; when the stock level is greater than 5, the optimal policy prescribes *DN* until $x = 2$; then *PM*.

Now we consider the influence of stock level on the optimal policy in the obsolescence cases: $(k, j) = \{(2, 1); (3, 1); (3, 2)\}$; where the available technology is greater than the one currently in use. With the chosen parameters in Table 1, the optimal decision is to replace the used asset by the latest available technology when replacement is prescribed.

We find that if the technological improvement is not significant, the firm tends to take advantage of the performance gained by available equipment and their spare parts in stock before investing in technological innovation. There exists an optimal inventory level to balance the loss due to later investment in new technology: replacement should be made when the stock level is lower than the defined threshold (s_1) or greater than (s_2). For example, with $(k, j) = (2, 1)$, the profit function and the failure rate are improved, i.e., $g_2(x)/g_1(x) = 1.01$ and $\alpha_2/\alpha_1 = 0.9$. If stock levels $\in [5, 10]$, the asset operates normally until deterioration state $x = 2$, then is maintained perfectly to restore in the initial state. At low stock levels, $s < 5$, compared with the cases $(k = j)$, the firm tends to replace earlier. For example, with $x = 2$, the replacement threshold in obsolescence case $(k, j) = (2, 1)$ is $s = 3$ while it is $s = 5$ in case $(k, j) = (1, 1)$. The greater the stock level is, the higher the holding cost is, so a high inventory level is not needed in the obsolescence case. As illustrated in Table 2, given $(k, j) = (2, 1)$ with $s \geq 13$, spare parts are utilized to perform *PM* when $x > 1$. This restores the asset to an initial state and a maximal profit $g(0)$ is incurred during the next period. If the deterioration state is $x = 0$ or $x = 1$, it is optimal to sell the asset in order to obtain the maximal residual values $b_{i,j,k}(0)$ or $b_{i,j,k}(1)$ and invest in new technology; else ($x > 1$) one continues to perform *PM* to reduce the deterioration state and stock level.

In the case where the latest available technology is a much greater improvement over the one currently in place, e.g., $(k, j) = (3, 1)$, the firm tends to invest early in new technology. It replaces the current asset by the latest available technology when the stock level is less than 2. If

the stock level $s \geq 3$, with $x \geq 3$, it utilizes spare parts in stock for PM to restore the asset to as-good-as-new before selling it at the next epoch (when its deterioration state $x \leq 2$ or stock level $s \leq 2$).

Table 3: The optimal policy for the first ten decision epochs in the obsolescence case (3, 1)

$\delta = 0.99$		0.8		0.6	
s	Action	s	Action	s	Action
[0, 2]	$\forall x: R$	0	$x < 2: DN$ $x \geq 2: R$	0	$x < 3: DN$ $x \geq 3: R$
		1	$x = 0: DN$ $x = 1, 2: IM$ $x \geq 3: R$	1	$x < 2: DN$ $x = 2, 3: IM$ $x \geq 4: R$
		2	$x < 2: DN$ $x = 2: PM$ $x = 3: IM$ $x \geq 4: R$	2	$x < 2: DN$ $x = 2: PM$ $x = 3, 4: IM$ $x = 5: R$
[3, 4]	$3 \leq x \leq s: PM$ $\forall x \in X \setminus [3, s]: R$	[3, 4]	$x = 0, 1: DN$ $2 \leq x \leq s: PM$ $x > s: R$	[3, 4]	$x = 0, 1: DN$ $2 \leq x \leq s: PM$ $s < x < 5: IM$ $x = 5: R$
$s \geq 5$	$x \leq 2: R$ $x \geq 3: PM$	$s \geq 5$	$x = 0, 1: DN$ $x \geq 2: PM$	$s \geq 5$	$x = 0, 1: DN$ $x \geq 2: PM$

Next we consider the rapid speed of technological evolution and its effects on the optimal strategy. Note that this speed is represented by the appearance probabilities of new technology in next the period p_{i+1}^{k+1} . As illustrated by the numerical examples in the case $(k, j) = (3, 1)$, when ε does not change, $\delta = 0.99, 0.8, 0.6\delta$, consider the first ten decision epochs (Table 3). We find that the smaller δ is, the greater the value of *DN* or *IM/PM* because the firm tends to keep the used asset for waiting the appearance of new technology when its appearance probability during the next period is high. Consider, for example, when $\delta = 0.99$, the replacement option has significantly more value and we replace immediately the asset at the low stock levels $s \leq 2$. Contrarily, the do nothing and maintenance options demonstrate their dominance when $\delta = 0.6, 0.8$.

4. Conclusion

In this paper, we proposed a model that allows us to take into account the spare part inventory in the maintenance/replacement problem of a stochastically deteriorating system under technological change. It determines the maintenance/replacement strategy based on the parametric performance of the system and the technological environment. We have combined several aspects never seen before in the same model: equipment replacement, maintenance optimization, technological evolution, and spare parts inventory.

Through our numerical examples we have shown the influence of the spare parts inventory level and technological change on the maintenance/replacement strategy. In the non-obsolescence case, it is obvious that replacement is done only at low stock levels; on the contrary, at high stock levels, the maintenance options demonstrate their dominance. In the obsolescence case, the replacement option with new technology is motivated, but the trade-off between the benefits of utilizing spare parts in store and the revenues of investment in new technology is also considered. Therefore, replacement is not done when the stock level is in the interval that is determined by the parameters of the model. The better the technological improvement is the greater value the replacement option has. However, in the case of rapid technological change, the do nothing and repair options have significantly more value as they allow the appearance of even better technologies in the future. These findings can greatly aid equipment managers in their maintenance and replacement decisions.

Some proposed assumptions can be seen as limitations of our model such as the expectation of purchase price and improvement of new technology. In fact, these can be stochastic and difficult to capture. An extension of our model could reflect the stochastic characteristics of these parameters. Furthermore, the stochastic efficiency of the imperfect maintenance action could also be included in our model or the non-stationary properties of the deterioration.

The principal objective of this paper is to consider simultaneously the influences of technological evolution and spare parts levels on the optimal maintenance/replacement strategy. Thus we have simplified the inventory problem by assuming that spare parts stock cannot be replenished without purchasing a new asset and that the quantity of spare parts supplied is determined by the manufacturer. In further research, we could take into account the replenishment capacity of stock and also examine the optimal inventory policy.

5. References

- 1) Bethuynne, G. 2002. The timing of technology Adoption by a cost-minimizing firm. *Journal of Economics* 76: 123-154.
- 2) Borgonovo, E; Marseguerra, M & Zio, E. 2000. A Monte Carlo methodological approach to plant availability modeling with maintenance, aging and obsolescence. *Reliability Engineering and System Safety* 67: 61-73.
- 3) Chelbi, A & Aït-Kadi, D. 2001. Spare provisioning strategy for preventively replaced systems subjected to random failure. *Int. J. Production Economics* 74: 183-189.
- 4) Chien, Y.H. 2009. A number-dependent replacement policy for a system with continuous preventive maintenance and random lead times. *Applied Mathematical Modelling* 33: 1708-1718.
- 5) Clavareau, J & Labeau, P.E. 2009a. A Petri net-based modeling of replacement strategies under technological obsolescence. *Reliability Engineering and System Safety* 94: 357-369.
- 6) Clavareau, J & Labeau, P.E. 2009b. Maintenance and replacement policies under technological obsolescence. *Reliability Engineering and System Safety* 94: 370-381.
- 7) Cobbaert, K & Oudheusden, D. V. 1996. Inventory models for fast moving spare parts subject to “sudden death” obsolescence. *Int. J. Production Economics* 44: 239-248.
- 8) De Smidt-Destombes, K. S; Van der Heijden, C.M & Hatem, A. V. 2006. On the interaction between maintenance, spare part inventories and repair capacity for k-out-of-N system with wear-out. *European Journal of Operational Research* 174: 182-200.
- 9) Dogramaci, A & Fraiman, N.M. 2004. Replacement decisions with maintenance under uncertainty: an imbedded optimal control Model. *Operations Research* 52: 785-794.
- 10) Hopp, W.J & Nair, S.K. 1994. Maintenance and replacement policies under technological obsolescence. *Reliability Engineering and System Safety* 94: 370-381.
- 11) Kabir, A.B.M.Z & Al-Olayan, A.S. 1996. A stocking policy for spare part provisioning under age based preventive replacement. *European Journal of Operational Research* 90: 171-181.
- 12) Karsak, E.E & Tolga, E. 1998. An overhaul-replacement model for equipment subject to technological change in an inflation-prone economy. *Int. J. Production Economics* 56-57: 291-301.

- 13) Kennedy, W.J; Wayne Patterson, J & Fredendall, L.D. 2002. An overview of recent literature on spare parts inventories. *Int. J. Production Economics* 76: 201-215.
- 14) Kim, J.S; Shin, K.C; Yu, H.K. 1996. Optimal algorithm to determine the spare inventory level for a repairable-item inventory system. *Computers Ops Res* 23: 289-297.
- 15) Mercier, S. 2008. Optimal replacement policy for obsolete components with general failure rates submitted to obsolescence. *Applied Stochastic Models in Business and Industry* 24: 221-235.
- 16) Nair, S.K. 1995. Modeling strategic investment decisions under sequential technological change. *Management Science* 41: 282-297.
- 17) Nair, S.K. 1997. Identifying technology Horizons for the strategic investment decisions. *IEEE Transactions on Engineering management* 44: 227-236.
- 18) Natali, H & Yatsenko, Y. 2007. Optimal equipment replacement without paradoxes: A continuous analysis. *Operations Research Letters* 35: 245-250.
- 19) Natali, H & Yatsenko, Y. 2008b. The dynamics of asset lifetime under technological change. *Operations Research Letter* 36: 565-568.
- 20) Natali, H & Yatsenko, Y. 2008a. Properties of optimal service life under technological change. *Int. J. Production Economics* 114: 230-238.
- 21) Torpong, C & Smith, R.L. 2003. A paradox in equipment replacement under technological improvement. *Operations Research Letters* 31: 77-82.
- 22) Saker, R & Haque, A. 2000. Optimization of maintenance and spare provisioning policy using simulation. *Applied Mathematical Modelling* 24: 751-760.
- 23) Sheu, H.S & Chien, Y.H. 2004. Optimal age-replacement policy of a system subject to shocks with random lead-time. *European Journal of Operational Research* 159: 132-144.
- 24) Vaughan, T.S. 2005. Failure replacement and preventive maintenance spare parts ordering policy. *European Journal of Operational Research* 161: 183-190.