Thèse de doctorat
de l'Université Louis Pasteur
Spécialité : Génie de l'environnement

présentée par

Patrick EISENBEIS

Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable

soutenue le 2 mai 1994

Commission d'examen :

M. André COLSON	**Président**
M. François LAISNEY	**Rapporteur interne**
M. Raimund HERZ	**Rapporteur externe**
M. Paul MOREL	**Rapporteur externe**
M. Henri REITZER	**Directeur de thèse**

Membres invités :

M. Bernard BREMOND

M. Daniel LOUDIERE

Thèse de doctorat
de l'Université Louis Pasteur
Spécialité : Génie de l'environnement

présentée par

Patrick EISENBEIS

Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable

soutenue le 2 mai 1994

Commission d'examen :

M. André COLSON	Président
M. François LAISNEY	Rapporteur interne
M. Raimund HERZ	Rapporteur externe
M. Paul MOREL	Rapporteur externe
M. Henri REITZER	Directeur de thèse

Membres invités :

M. Bernard BREMOND

M. Daniel LOUDIERE

AVANT-PROPOS

Cette étude n'a pu être réalisée que grâce à la participation conjointe des partenaires suivants :

- la société **Lyonnaise des Eaux-Dumez** qui a amené une participation technique et financière et qui a permis d'utiliser les archives concernant les travaux de maintenance du Service des Eaux de Bordeaux et de la CEB,

- la société **Pont-à-Mousson S.A.**, qui a amené une participation technique et financière et qui a donné de précieux renseignements sur les conduites d'eau potable et leur évolution technologique, ainsi que sur la corrosivité des sols,

- l'**Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg** (ENGEES), qui m'a accueilli dans ses locaux et m'a apporté une aide scientifique,

- le **Centre National du Machinisme Agricole du Génie Rural et des Eaux et Forêts** (CEMAGREF) et plus particulièrement la Division Ouvrages et Réseaux Hydrauliques du Groupement de Bordeaux qui a apporté un encadrement scientifique essentiel,

- l'**Université Louis Pasteur** et plus particulièrement l'Institut de Mécanique des Fluides pour sa participation scientifique.

REMERCIEMENTS

Je tiens à remercier les personnes suivantes qui ont participé à l'élaboration de cette thèse :

- M. Henri REITZER chargé de recherche CNRS de l'Institut de Mécanique des Fluides de l'ULP, directeur de thèse,
- M. Franck TRIOL de la Compagnie des Eaux et de la Banlieue de Paris ainsi que toute l'équipe du service Eau Potable de la CEB,
- M. Michel LANGENFELD de Pont-à-Mousson S.A.,
- MM. Pierre ACHART et MADIEC de la Lyonnaise des Eaux-Dumez Bordeaux, ainsi que l'équipe de M. PRIVAT,
- M. Daniel LOUDIERE, directeur de l'ENGEES,
- M. Jean DARBOUR de l'ENGEES,
- M. Daniel COMMENGES de l'INSERM Bordeaux,
- M. Christian GOURRIEROUX, de l'INSEE,
- M. MAHE du Service Départemental des Eaux et d'Assainissement du Bas-Rhin,
- M. Lahcen ZEDGUI, stagiaire en mastère à l'ENGEES,
- M. Pascal BOTZUNG, stagiaire de troisième année de l'ENGEES à la LED-Bordeaux,
- M. Olivier ALEXANDRE, du Laboratoire Gestion des Services Publics du CEMAGREF.

Je remercie également M. François LAISNEY Professeur à l'ULP, M. Paul MOREL Professeur à l'Université Bordeaux I et M. Raimund HERZ Assitant-Professeur à l'Université de Karlsruhe d'avoir accepté de prendre part à mon jury en tant que rapporteurs, ainsi que M. André COLSON, directeur de l'ENSAIS en tant qu'examinateur.

Mes remerciements vont également à toutes les personnes de l'équipe de la direction de la recherche de l'ENGEES, qui contribuent à son bon fonctionnement dans une ambiance chaleureuse.

Enfin je remercie tout particulièrement M. Bernard BREMOND, directeur de recherche à la division Ouvrage et Réseaux Hydrauliques du groupement de Bordeaux du CEMAGREF, qui a su répondre à mes nombreuses sollicitations, qui ont valu quelques Allers-Retours Strasbourg-Bordeaux. Ses précieux conseils ont fortement contribué à la réussite de cette étude.

RESUME

L'âge et surtout l'état actuel de certains réseaux d'eau potable laissent présager dans un proche avenir une nette augmentation des défaillances et risquent d'engendrer un surcoût économique lié à la nécessité de renouveler les conduites dégradées. Il convient dès à présent de mieux connaître le vieillissement des canalisations.

L'objectif de cette étude est d'élaborer un modèle de prévision des défaillances des conduites d'eau potable, critère prépondérant dans le choix des conduites à renouveler, en fonction de facteurs environnementaux propres à la conduite.

Des données provenant de services urbains et ruraux (Lyonnaise des Eaux-Dumez Bordeaux, Compagnie des Eaux de la Banlieue de Paris, Syndicats intercommunaux du Kochersberg et du Kronthal dans le Bas-Rhin) possédant des longs historiques de maintenance ont été utilisées. La durée de ces historiques est respectivement de 40, 53 et 17 ans, pour un linéaire total d'environ 700 Km. La spécificité de ces données (données censurées) a conduit à l'utilisation d'analyses statistiques particulières appelées analyse des données de survie.

Ce sont les modèles de Cox (semi paramétrique) et de Weibull (paramétrique) qui ont permis de bien caractériser l'apparition des défaillances des conduites d'eau potable. On a pu montrer que le facteur "nombre de défaillances déjà subies par la conduite" était prépondérant. Les autres facteurs significatifs sont le diamètre, puis le sol, l'année de pose (différenciation fonte ductile-fonte grise) et le facteur croisé trafic-chaussée. La comparaison de l'espérance du nombre de défaillances par rapport au nombre de défaillances réellement arrivées a permis de mettre en évidence un facteur propre à la conduite représentant sa disposition à un vieillissement accéléré.

Afin de pouvoir utiliser ce type de modèle dans un service ayant un historique de maintenance de durée faible, une méthode d'apprentissage a été proposée. Elle permet d'utiliser les résultats d'un modèle général pour créer un échantillon fictif de casses sur le service pour la période inconnue. Cet échantillon est alors rattaché aux défaillances connues pour créer un modèle propre au service.

Il est important que les services se dotent de bases de données informatisées répertoriant les conduites et leurs caractéristiques ainsi que les défaillances pour mettre en oeuvre efficacement ce modèle. On pourra alors mieux estimer la dégradation du réseau et, en relation avec une étude économique, il sera possible de prévoir à moyen ou long terme les dépenses futures en matière de renouvellement et d'entretien.

ABSTRACT

The age and the present state of some networks augur economical overcosts due to the fact that failures are increasingly arising. It is now advisable to be versed in the ageing of mains.

The object of this survey is the creation of a predicting model of water mains failures, which are the preponderant criterion for the selection of mains to replace, according to environmental factors.Data from differents urban and rural companies with long-standing water main repair (700 km) were collected.

The specificity of data (censured) prevail to use the analysis of survival data. In using both Cox and Weibull models, it is possible to characterize the emergence of failures. The preponderant factor is the number of failures already occured on the mains. The other significant factors are the diameter, followed by the nature of soil, the material (grey cast iron, ductile cast iron), motor traffic, and the location of the drinking water pipes in the street.

Confronting the number of estimated failures and the one actually occured, revealed a specific factor to each main : the main ability to an accelerated ageing process.

In order to use this model in a drinking water service with short-standing water main repair, a method of "apprentice-ship" is proposed. It use the results of a general model to create an artificial sample of failures on the network for the unknown period.Then, it is possible to create a proper model to the service.

For an efficiently application of this model, it is important to own a computerized data base indexing the mains, their characteristics and failures. Thus, it will be possible to estimate the deterioration of the main. An economical survey can easily complete this research, to predict the future expenses of renewal.

SOMMAIRE

VOLUME 1

AVANT-PROPOS 2

RESUME 4

ABSTRACT 5

SOMMAIRE 6

LISTE DES FIGURES 8

LISTE DES TABLEAUX 12

INTRODUCTION 15

CHAPITRE 1 : LA PROBLEMATIQUE DU RENOUVELLEMENT DES RESEAUX D'EAU POTABLE 16
1.1. LES POLITIQUES DE RENOUVELLEMENT MENEES ACTUELLEMENT 16
1.2. COMMENT LE VIEILLISSEMENT D'UN RESEAU SE MANIFESTE-T-IL? 21
1.3. CONCLUSION 26

CHAPITRE 2 : LES DEFAILLANCES DES CONDUITES D'EAU POTABLE : CAUSES ET MODELES DE PREVISION DEJA EXISTANTS 28
2.1. DESCRIPTION ET CAUSES DES DEFAILLANCES 28
2.2. LES MODELES DE PREVISION DES DEFAILLANCES DES CONDUITES D'EAU POTABLE 33
2.3. CHOIX DU MODELE 40

CHAPITRE 3 : LES DONNEES DISPONIBLES 42
3.1. LES DONNEES RECHERCHEES 42
3.2. PRESENTATION GENERALE DES RESEAUX 43
3.3. LE RECUEIL DES DONNEES 48
3.4. DESCRIPTION DES TRONÇONS ET DES DEFAILLANCES 59
3.5. LA SPECIFICITE DES DONNEES 61

CHAPITRE 4 : APPLICATION DE L'ANALYSE STATISTIQUE DES DONNEES DE SURVIE A L'ETUDE DES DEFAILLANCES 64
4.1. INTRODUCTION A L'ANALYSE DES DONNEES DE SURVIE 64
4.2. LES MODELES UTILISES EN ANALYSE DE SURVIE 69
4.3. UTILISATION DU MODELE DES RISQUES PROPORTIONNELS - ESTIMATION DES PARAMETRES 71

CHAPITRE 5 : APPLICATION DES MODELES DE COX ET DE WEIBULL A LA PREVISION DES DEFAILLANCES DES CONDUITES D'EAU POTABLE 76
5.1. PRESENTATION DU LOGICIEL EGRET 76
5.2. TRANSFORMATION DES DONNEES POUR L'ANALYSE DES DONNEES DE SURVIE 77
5.3. LE MODELE DE COX 79
5.4. APPLICATION DU MODELE DE WEIBULL 96
5.5. CONCLUSIONS SUR L'APPLICATION DES MODELES 102

CHAPITRE 6 : ADEQUATION DU MODELE 104
6.1. CALCUL DE L'ESPERANCE DU NOMBRE DE DEFAILLANCES 104
6.2. COMPARAISON AVEC LES DONNEES REELLES 106
6.3. MESURE DE L'INFLUENCE DE L'OUBLI D'UN FACTEUR 124
6.4. CONCLUSION 128

CHAPITRE 7 : L'APPRENTISSAGE 129
7.1. LA METHODE 129
7.2. RESULTATS 132
7.3. CONCLUSION 137

CHAPITRE 8 : UTILISATION DU MODELE PAR LES GESTIONNAIRES DE SERVICE 139
8.1. METHODOLOGIE POUR LA CREATION D'UN MODELE 139
8.2. UTILISATION POUR CONNAITRE L'ETAT DU RESEAU 144
8.3. UTILISATION POUR LA PLANIFICATION DU RENOUVELLEMENT 144
8.4. UTILISATION DANS L'ETUDE DE FIABILITE D'UN RESEAU 152

CONCLUSION 153

REFERENCES BIBLIOGRAPHIQUES 155

BIBLIOGRAPHIE 157

TABLE DES MATIERES 159

VOLUME 2

ANNEXES

LISTE DES FIGURES

		Pages
Figure 1 :	Démarche employée dans le cadre de l'approche curative [O'day et al., 1989]	20
Figure 2 :	Démarche employée dans le cadre de l'approche prédictive [O'day et al., 1989]	22
Figure 3 :	Taux de défaillance annuel par unité de longueur (données LED-Bordeaux, 1991) [Eisenbeis, 1991]	26
Figure 4 :	Différents types de ruptures des conduites d'eau potable [O'day et al., 1989]	29
Figure 5 :	Photographie représentant une conduite ayant subi une rupture.	29
Figure 6 :	Fréquence des types de ruptures en fonction du diamètre [O'day et al., 1989]	30
Figure 7 :	Répartition des fuites sur tuyau et sur joint en fonction du diamètre (d'après les données recueillies sur la LED-Bordeaux) [Eisenbeis, 1991]	30
Figure 8 :	Contraintes subies par une conduite d'eau potable [Moruzzi, 1984]	31
Figure 9 :	Causes de l'affaiblissement d'une conduite d'eau potable [O'day et al., 1989]	33
Figure 10 :	Courbe présentant le modèle physique de Vancouver [O'day et al., 1989]	35
Figure 11 :	Exemples de courbes de survie de conduites d'eau potable (New Haven) [Andreou et al., 1987]	40
Figure 12 :	Exemple de fiches de maintenance (LED-Bordeaux)	56
Figure 13 :	Répartition des conduites par année de pose par service	60
Figure 14 :	Répartition des diamètres étudiés sur les trois services	61
Figure 15 :	Proportion des conduites selon le nombre de défaillances subies	62
Figure 16 :	Répartition des défaillances en fonction du mois	62
Figure 17 :	Description de la spécificité des données	63
Figure 18 :	Distinction des canalisations en individus (ndef correspond au nombre de défaillance).	66
Figure 19 :	Exemple de courbe de survie.	67
Figure 20 :	Fonction de risque chez l'homme	68
Figure 21 :	Fonction de risque (chômage)	68
Figure 22 :	Fonction de risque, densité et fonction de survie du modèle exponentiel	69
Figure 23 :	Fonctions de risque du modèle de Weibull.	70
Figure 24 :	Outils probabilistes en analyse de données de survie	72
Figure 25 :	Définition des "individus" vis-à-vis des tronçons de canalisations	78
Figure 26 :	Ecart moyen entre deux défaillances	83
Figure 27 :	Fonction ln(H(t)) pour chacun des groupes de défaillance	84

Figure 28 : Vérification de l'hypothèse de proportionnalité des risques pour le facteur Trafic-Chaussée. 85

Figure 29 : Courbes de probabilité de non-défaillance (1ère défaillance) avec intervalle de confiance à 95 % pour une conduite de 500 m de long, de diamètre 100, posée après 1975, dans un sol corrosif, dans une rue à trafic fort et sous-chaussée. 88

Figure 30 : Courbes de probabilité de non-défaillance (1ère défaillance) avec intervalle de confiance à 95 % pour une conduite de 500 m de long, de diamètre 100, posée avant 1966 (fonte grise), dans un sol corrosif et dans une rue à trafic faible ou sous-chaussée. 88

Figure 31 : Courbe de risque de défaillance (Modèle de Cox) 89

Figure 32 : Courbes des risques de défaillance selon les groupes de défaillances 90

Figure 33 : Courbe de probabilité de non-défaillance (1ère défaillance d'une conduite de 150 mm, de 500 m de long, posée avant 1975, dans un sol corrosif, dans une rue à trafic élevé et sous chaussée) 90

Figure 34 : Courbe de probabilité de non-défaillance (2 à 4ème défaillance d'une conduite de 60 mm, de 500 m de long, posée avant 1966 (fonte grise), dans un sol corrosif, dans une rue à trafic fort, sous chaussée) 91

Figure 35 : Evolution du paramètre β_l en fonction de la longueur maximale des tronçons de l'échantillon initial. 92

Figure 36 : Valeurs des coefficients β_l en fonction de la longueur maximale des tronçons de l'échantillon élargi. 95

Figure 37 : Fonctions de risque de base pour la première défaillance. 100

Figure 38 : Fonctions de risque de base pour les 2 à 4èmes défaillances (2ème et + pour le Bas-Rhin). 100

Figure 39 : Fonctions de risque de base pour les 5èmes défaillances et +. 101

Figure 40 : Création d'un modèle sur un service à historique faible. 103

Figure 41 : Répartition de k défaillances entre 0 et t. 104

Figure 42 : Espérance de défaillance à partir de t_d à la suite de k défaillances. 106

Figure 43 : Démarche utilisée pour la comparaison des défaillances prévues et réelles 107

Figure 44 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons. 109

Figure 45 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (Hypothèse de non-stratification). 110

Figure 46 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (avec facteur "Vieillissement"). 113

Figure 47 : Comparaison des nombres de défaillances sur les 90 premières conduites (sans facteur "vieillissement") 114

Figure 48 : Comparaison des nombres de défaillances sur les 90 premières conduites (avec facteur "vieillissement") 115

Figure 49 : Méthode employée pour la comparaison des défaillances prévues et réelles entre 1983 et 1992 112

Figure 50 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (prévision entre 1983 et 1992). 117

Figure 51 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision sur 10 ans) 118

Figure 52 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (modèle non stratifié avec facteur "vieillissement", prévision entre 1983 et 1992). 121

Figure 53 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (modèle non stratifié avec facteur "vieillissement", prévision entre 1991 et 1992). 122

Figure 54 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision sur 2 ans) 123

Figure 55 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision entre 1951 et 1990; oubli du facteur trafic-chaussée). 126

Figure 56 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision entre 1951 et 1990; oubli du facteur sol). 127

Figure 57 : Description des données du service A. 130

Figure 58 : Création d'un modèle à partir des données du service A et du modèle général. 130

Figure 59 : Création d'un modèle à partir des données du service A après 1988 et d'un échantillon de défaillances construit à partir du modèle général pour les conduites posées avant 1988. 131

Figure 60 : Probabilités de non-défaillance pour une conduite n'ayant pas eu de défaillance, de diamètre 150, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage à partir de l'espérance). 133

Figure 61 : La technique d'anamorphose. 134

Figure 62 : Probabilités de non-défaillance pour une conduite n'ayant pas eu de défaillance, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose). 136

Figure 63 : Probabilités de non-défaillance pour une conduite ayant eu une à trois défaillances, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose). 136

Figure 64 : Probabilités de non-défaillance pour une conduite ayant eu au moins quatre défaillances, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose). 137

Figure 65 : Apprentissage année après année pour la création de modèles semi-locaux. 137

Figure 66 : Démarche à employer lors de l'élaboration d'un modèle de prévision des défaillances. 139

Figure 67 : Exemple de fiche de réparation. 142

Figure 68 : Etablissement du fichier en vue du modèle. 143

Figure 69 : Utilisation du modèle dans une politique curative de renouvellement 145

Figure 70 : Exemple de calcul de la date optimale de renouvellement (Méthode de Shamir et Howard) [Shamir et al., 1979] 146

Figure 71 : Recherche de la date optimale de renouvellement (Rue du Vieux Marché aux Poissons, Hypothèse non favorable, sans coût indirect) [El Naboulsi, 1993] 147

Figure 72 : Recherche de la date optimale de renouvellement (Rue du Vieux Marché aux Poissons, Hypothèse non favorable, avec coût indirect) [El Naboulsi, 1993] 148

Figure 73 : Utilisation du modèle dans la méthode de Shamir et Howard. 149

Figure 74 : Fonction de survie, de répartition et taux de défaillances pour les paramètres a=120, b=0,12, c=10. [Herz, 1992] 150

Figure 75 : Utilisation du modèle dans la méthode de Herz 151

LISTE DES TABLEAUX

		Pages
Tableau 1 :	Critères de renouvellement pris en compte dans les services américains [O'day et al., 1989]	19
Tableau 2 :	Manifestations et mesures de la dégradation des conduites d'eau potable	27
Tableau 3 :	Répartition des canalisations par diamètre (décembre 1990)	44
Tableau 4 :	Répartition des canalisations par secteur (décembre 1990)	45
Tableau 5 :	Répartition par diamètre sur le réseau de la CEB	47
Tableau 6 :	Répartition par matériau sur le réseau de la CEB	47
Tableau 7 :	Répartition des diamètres sur les syndicats du Bas-Rhin	48
Tableau 8 :	Répartition des matériaux sur les syndicats du Bas-Rhin	48
Tableau 9 :	Répartition des canalisations par commune du secteur 4 (décembre 1990)	49
Tableau 10 :	Répartition des conduites selon les communes étudiées sur la CEB	50
Tableau 11 :	Répartition des conduites selon le diamètre sur la CEB	50
Tableau 12 :	Répartition des conduites selon le matériau sur la CEB	50
Tableau 13 :	Caractéristiques des conduites prises en compte sur les différents services.	58
Tableau 14 :	Description des conduites recueillies par service.	60
Tableau 15 :	Nombre de défaillances recueillies par service.	61
Tableau 16 :	Résultats donnés par le logiciel EGRET	76
Tableau 17 :	Estimation des facteurs influents (test de Wald)	82
Tableau 18 :	Groupes de défaillances	84
Tableau 19 :	Résultats de l'estimation des β_i du modèle de Cox	86
Tableau 20 :	Bornes des intervalles de confiance (95 %)	87
Tableau 21 :	Coefficient β_l en fonction du groupe de défaillances.	93
Tableau 22 :	Valeurs du paramètre β_l de la longueur selon l'échantillon.	94
Tableau 23 :	Paramètres β_i du modèle de Weibull avec stratification	98
Tableau 24 :	Paramètres des fonctions de Weibull et exponentiel selon les services et les groupes de défaillances	99
Tableau 25 :	Paramètres β_i estimés dans l'hypothèse de non-stratification (** signifie que le facteur n'est pas significatif).	108
Tableau 26 :	Paramètres du modèle de Weibull stratifié avec le facteur correctif "Vieillissement".	111
Tableau 27 :	Paramètres du modèle de Weibull stratifié sans facteur vieillissement (données 1951-1982)	116

Tableau 28 : Coefficients estimés pour les différents modèles concernant la prévision sur 10 ans. 119

Tableau 29 : Nombre de défaillances prévues selon les modèles choisis 116

Tableau 30 : Nombre de défaillances prévues selon les modèles choisis (prévision sur deux ans. 120

Tableau 31 : Coefficients des modèles de Weibull selon les facteurs pris en compte dans le modèle. 124

Tableau 32 : Espérances estimées sur la période de 1951 à 1990 selon les facteurs pris en compte dans le modèle. 125

Tableau 33 : Coefficients du modèle semi-local en échantillonnant à partir de l'espérance 133

Tableau 34 : Coefficients du modèle semi-local en échantillonnant à partir de la technique d'anamorphose. 135

INTRODUCTION

Selon une étude de l'Union des Industries et Entreprises de l'Eau et de l'Environnement [Union des Industries et Entreprises de l'Eau et de l'Environnement, 1991], 15 % des conduites d'eau potable existant actuellement en France ont une ancienneté supérieure à 50 ans et 25 % supérieure à 40 ans. Il est donc certain qu'à moyen terme se posera le problème du remplacement de ces conduites déjà âgées.

Connaissant le coût au mètre linéaire de la pose d'une canalisation (500 à 2000 F/ml), les services d'eau verront leurs dépenses augmenter de manière importante. Ces dépenses seront d'autant plus difficiles à supporter que les services sont petits. Il convient donc dès à présent d'essayer de prévoir ces dépenses futures.

Beaucoup de services travaillent actuellement dans l'urgence et ne remplacent que les conduites d'eau déjà malades. Ceci entraîne un financement irrégulier d'année en année. Une prévision du renouvellement à moyen ou long terme permettrait au contraire un financement plus régulier.

Avant toute étude économique, les services doivent organiser la connaissance de leur réseau (nature de la conduite, date de pose, environnement, incidents). Le traitement de ces données leur permettra ensuite d'élaborer des outils utilisables pour la prévision du renouvellement des canalisations.

L'objectif de cette étude est de créer un modèle de prévision du vieillissement des conduites d'eau potable et plus particulièrement de leurs défaillances. Ce modèle doit permettre de définir quels sont les facteurs les plus influents sur les casses et de prévoir sur une période donnée le nombre de casses probables que les conduites pourront subir en fonction de ces facteurs. Bien entendu il est indispensable que ce modèle puisse être applicable le plus largement possible.

Dans le chapitre 1, sont présentés diverses politiques de renouvellement déjà mises en place, ainsi que les critères de vieillissement habituellement pris en compte dans le choix des conduites à renouveler. Le chapitre 2 récapitule les facteurs influençant les défaillances ainsi que les modèles de prévision déjà existants. Les données utilisées pour l'élaboration du modèle sont présentées dans le chapitre 3. Le chapitre 4 présente les outils statistiques utilisés, qui sont appliqués aux conduites d'eau potable au chapitre 5. Dans le chapitre 6, nous étudierons l'adéquation du modèle aux données réelles. Le chapitre 7 définit une méthode qui permettra aux services ayant de faibles historiques de maintenance d'utiliser le modèle. Enfin dans le chapitre 8, nous verrons comment les gestionnaires de services peuvent utiliser ce modèle pour la prévision du renouvellement en rapport avec des études économiques.

CHAPITRE 1 :
LA PROBLEMATIQUE DU RENOUVELLEMENT DES RESEAUX D'EAU POTABLE

Dans le paragraphe 1.1., sont présentées quelques politiques de renouvellement déjà menées en France et à l'étranger. Puis dans le paragraphe 1.2., sont définis les critères de vieillissement des conduites, critères souvent décisifs dans les stratégies de renouvellement.

1.1. LES POLITIQUES DE RENOUVELLEMENT MENEES ACTUELLEMENT

A l'heure actuelle le renouvellement des réseaux d'eau potable n'est pas une priorité pour tous les services. Parmi ceux-ci ce sont les services les plus importants en taille qui sont le plus sensibilisés. Cependant les démarches suivies dans le choix des conduites à renouveler sont différentes d'un réseau à l'autre, d'un pays à l'autre.

1.1.1. Les politiques de renouvellement menées en France

Une enquête de l'Association Générale des Hygiénistes et des Techniciens Municipaux menée auprès de 1105 services [AGHTM,1993] a permis de faire le point sur les expériences de politique de renouvellement menées en France.

Parmi les petits services, sur 238 réponses, on a pu tirer les enseignements suivants :

- 46% de ces services réalisent des opérations de renouvellement, mais peu opèrent suivant un plan à plus ou moins long terme.
- Les services ne réalisant pas d'opérations de renouvellement sont essentiellement des services de petite taille sans doute parce qu'ils n'ont pas la structure ni l'assise financière leur permettant d'engager ce type de travaux.
- Le taux moyen de renouvellement des conduites est de 0,64% par an.
- Les motifs prioritaires pour engager des opérations de renouvellement de conduites sont :
  * La réfection de voirie,
  * L'augmentation de diamètre pour renforcement,
  * La fréquence des fuites,
  * L'âge des conduites.

Sur les 13 gros services interrogés, seuls trois annoncent une véritable politique de renouvellement de leur réseau répondant à un ou plusieurs objectifs précis.

Ces services envisagent le problème de la façon suivante :

- Le premier a divisé son réseau en deux parties, le réseau urbain plus ancien (62 ans d'âge moyen) et le réseau périphérique (32 ans). L'objectif principal est de rajeunir le réseau urbain, sans laisser vieillir le second. Les taux de renouvellement annuel est depuis 5 ans de 1,9% dans le centre et de 1% dans la périphérie.

  L'installation d'un nouveau Système d'Information Géographique a permis de retrouver les âges des canalisations et donc de prévoir au rythme actuel l'âge moyen à moyen terme. Celui ci serait en 2016 de 24 ans pour le centre et de 30 ans pour la périphérie.

  D'autre part la sélection des opérations se fait selon les critères suivants : la fragilité des conduites, la sensibilité à l'environnement, les opérations de voirie, la nécessité de renforcer le réseau.

- Le deuxième service tient un raisonnement similaire au précédent, basé sur la moyenne d'âge du réseau. L'objectif ici est de maintenir l'âge moyen du réseau inférieur à 100 ans, tout en cherchant à diminuer les nuisances (remplacement des tronçons fuyards), à améliorer le service (renforcements, remplacement de tronçons entartrés) et à coordonner les travaux sur le réseau avec les autres travaux.

- La politique du troisième est un peu différente. Elle a été mise en oeuvre il y a une quinzaine d'années afin de relever le rendement de réseau qui était assez médiocre. Elle s'intégrait en fait dans un plan plus vaste visant à limiter ou inverser l'évolution croissante de la production d'eau pour retarder l'obligation de construire une deuxième usine de production.

En conclusion de l'enquête certains conseils sont donnés aux gestionnaires de service, dans l'objectif du renouvellement :

- Il est tout d'abord important de bien connaître le réseau et son histoire, en créant par exemple des fichiers d'intervention et des fichiers de conduites qui faciliteront le traitement des informations.

- Il faut ensuite analyser les circonstances qui peuvent influer sur la dégradation du réseau. Ce sont les facteurs de désordre, que nous verrons au chapitre 2.

- Puis il faut savoir ce qui doit déterminer la décision d'engager des travaux de renouvellement. Ce sont les facteurs déclenchants, qui correspondent aux critères de renouvellement.

Un autre service, le Syndicat des Eaux d'Ile de France [Halpern et Pascal, 1988], a défini cinq critères principaux pour évaluer ses besoins en renouvellement. Ce sont le maintien de la valeur

du patrimoine, la fréquence des fuites, l'insertion dans les programmes de travaux municipaux, l'âge des canalisations et la qualité de l'eau dans le réseau.

Ces méthodes ne permettent cependant pas d'anticiper sur une éventuelle dégradation du réseau.

1.1.2. Les politiques de renouvellement menées en Allemagne

Le DVGW (Deutsche Verein des Gas- und Wasserfaches E.V.) a sorti une fiche de recommandation en matière d'aide à la décision pour la planification du renouvellement et de la remise en état des conduites de distribution de gaz et d'eau [DVGW, 1989]. Cette fiche précise, entre autres, une méthode de planification du renouvellement.

Cette méthode se présente en deux étapes :

- détermination d'un taux de renouvellement annuel à partir de l'étude des données concernant les dommages et l'état des conduites, ainsi qu'à partir des ressources financières disponibles.

- puis prise en considération de critères dans le choix des conduites à renouveler à savoir :
  - le nombre de dommages, leur évolution, leurs conséquences économiques ;
  - le type et la cause des dommages ;
  - l'état de la conduite ;
  - la sécurité de la distribution ;
  - les mesures préventives (notamment pour ce qui concerne les risques tectoniques) ;
  - les pertes d'eau, sous l'aspect technique et sous l'aspect économique ;
  - la modification de la qualité de l'eau ;
  - les pertes de charge ;
  - les conditions économiques favorables en fonction d'autres travaux ;
  - les coûts de remplacement.

Cependant on se rend compte que, le plus souvent, la durée d'utilisation effective d'une conduite est déterminée par les travaux des tiers et rarement par des considérations de rentabilité.

Les Technische Werke der Stadt Stuttgart (TWS) [Kottmann, 1978] ont pu déterminer le taux de renouvellement nécessaire et les critères de choix des conduites à partir d'une banque de données. Ce sont :

- l'insuffisance de service,
- un nombre important de fuites aux raccordements des tuyaux,
- le nombre d'interventions pour la remise en état,
- la réfection de route,
- un mauvais état significatif de la conduite.

En résumé, les décisions de renouvellement sont prises selon quatre objectifs primordiaux :

- la rentabilité,
- la réputation du service auprès du public,
- la sécurité d'exploitation,
- le taux de renouvellement à long terme.

1.1.3. Les politiques de renouvellement menées aux États-Unis

Aux États-Unis, certains services ont déjà élaboré des politiques de renouvellement dès le début des années soixante, notamment en établissant des banques de données concernant les conduites et les dommages qu'elles ont pu subir. Deux approches différentes sont ou seront utilisées : une approche curative et une approche prédictive.

1.1.3.1. L'approche curative [O'day et al., 1989]

Cette approche s'apparente aux méthodes utilisées en France ou en Allemagne. Cependant l'existence de bases de données importantes permet de définir plus précisément les canalisations à problème. Les critères pris en compte dans l'évaluation des conduites à remplacer se trouvent dans le tableau 1.

Tableau 1 : Critères de renouvellement pris en compte dans les services américains [O'day et al., 1989]

Facteurs	East Bay MUD	Philadelphia	Denver	Louisville
Caractéristiques des conduites				
Age		X	X	X
Diamètre		X	X	X
Profondeur			X	
Corrosion		X	X	X
Conduite en antenne		X		X
Type de matériau ou de joint		X	X	X
Historique de casses ou de fuites	X	X	X	X
Caractéristiques du système				
Qualité de l'eau	X	X		X
Pression	X	X	X	X
Flexibilité opérationnelle	X	X		
Coûts de pompage	X			
Conduite renforcée		X		
Capacité de transport	X	X	X	
Caractéristiques environnementales				
Données de sol			X	X
Electrolyse		X		
Zone prioritaire	X			X
Zone de développement	X	X	X	X
Analyse du coût				
Coût de maintenance	X			
Coût de remplacement	X			
Taux de rendement des investissements	X			

On peut noter que les critères pris en compte sont différents selon les services, ceci étant dû pour une part à la spécificité géographique. Les services donnent ensuite un poids à chacun des critères de manière à déterminer une "note" pour chacune des conduites. Cette note permet alors de définir si la conduite doit être ou non remplacée.

L'approche curative générale suit en fait la démarche de la figure 1.

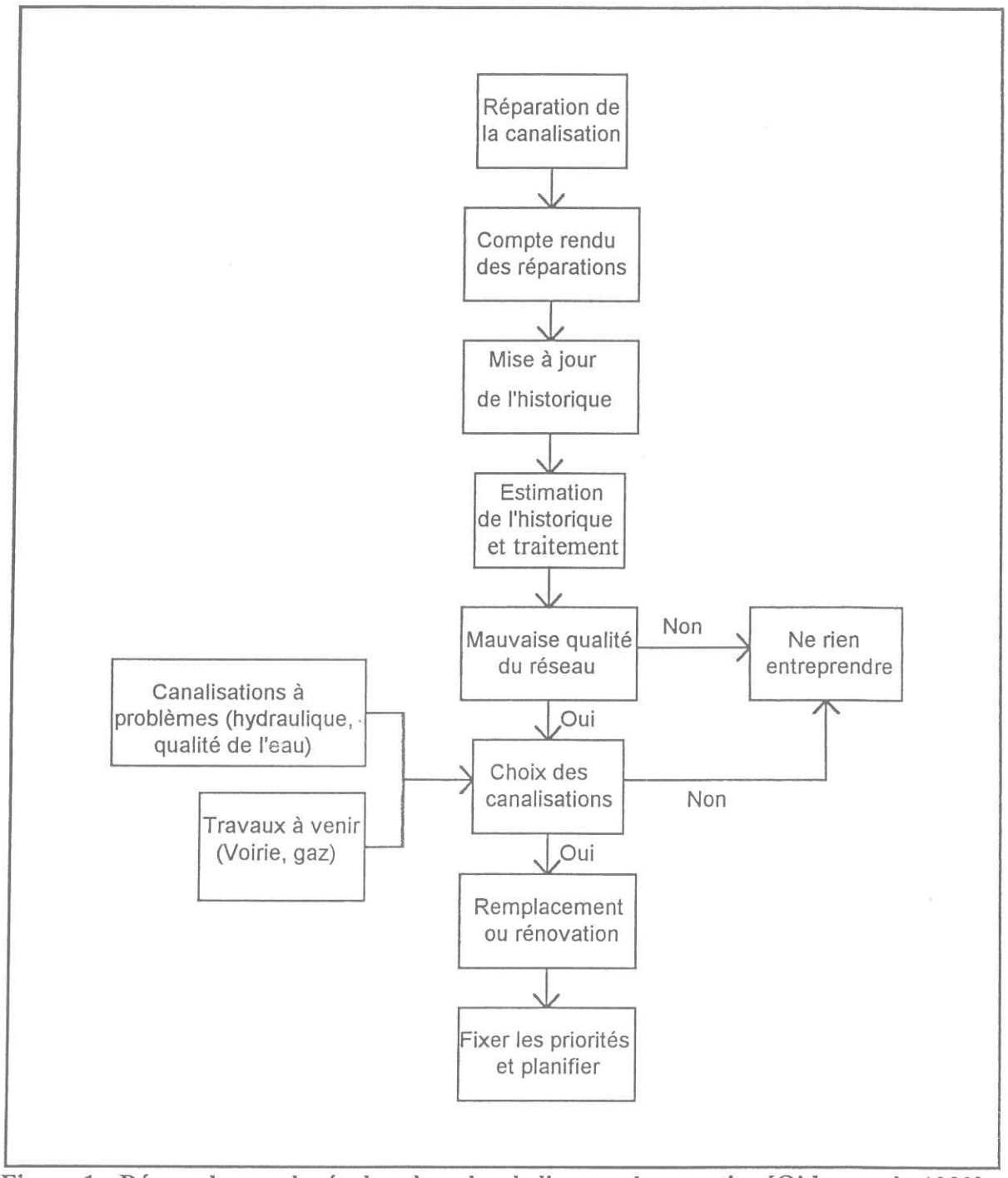


Figure 1 : Démarche employée dans le cadre de l'approche curative [O'day et al., 1989]

1.1.3.2. L'approche prédictive [O'day et al., 1989]

Certains services ne sont pas satisfaits de l'approche précédente car elle ne se limite qu'à l'évaluation des conduites ayant déjà subi des défaillances. Faisant face à une augmentation de la détérioration de leur système de distribution, ces services déjà anciens préfèrent mettre en place un système de planification prédictif, qui prendrait en compte la nature du procédé de détérioration.

Cette approche utilise un système d'inventaire et de bases de données concernant les conduites d'eau potable. La principale différence avec l'approche curative est qu'elle essaie de prédire les conditions physiques présentes et futures d'une canalisation ayant certaines caractéristiques propres et environnementales. Ensuite, une fois cette prédiction faite, il est possible de l'utiliser en y appliquant par exemple une méthode économique. Ceci permet donc aux services de mieux contrôler d'une année à l'autre la maintenance de leur réseau et de ne pas travailler uniquement en "urgence". La figure 2 montre la démarche employée pour cette approche.

Nous verrons dans le chapitre II quels types de modèles de prévision de la détérioration des conduites peuvent être utilisés.

1.2. COMMENT LE VIEILLISSEMENT D'UN RESEAU SE MANIFESTE-T-IL?

Comme on l'a vu au paragraphe 1.1.3.2. dans le chapitre précèdent, notamment pour ce qui concerne l'approche prédictive, il est important d'essayer de créer des modèles de vieillissement des réseaux d'eau potable. Avant de créer ce type de modèle, il faut définir ce que l'on appelle "vieillissement d'une conduite". Notons que l'on ne tient compte ici que du vieillissement des conduites de distribution, et non des branchements.

1.2.1. Définition du vieillissement d'une conduite d'eau potable [Kremer, 1990]

Le vieillissement d'une conduite correspond à sa dégradation dans le temps, celle-ci donnant lieu, soit à certains dommages, soit au mauvais fonctionnement hydraulique du réseau.

1.2.1.1. Mauvais fonctionnements hydraulique du réseau engendrés par le vieillissement d'une conduite

- **Chute de pression**, lorsque la section utile de la conduite diminue à cause de l'entartrage ou de protubérances dues à la corrosion.

- **Fuites diffuses**, dues à la détérioration des joints ou à la corrosion des tuyaux. Une forte augmentation de leur nombre peut avoir une incidence directe sur le réseau et diminuer le rendement.

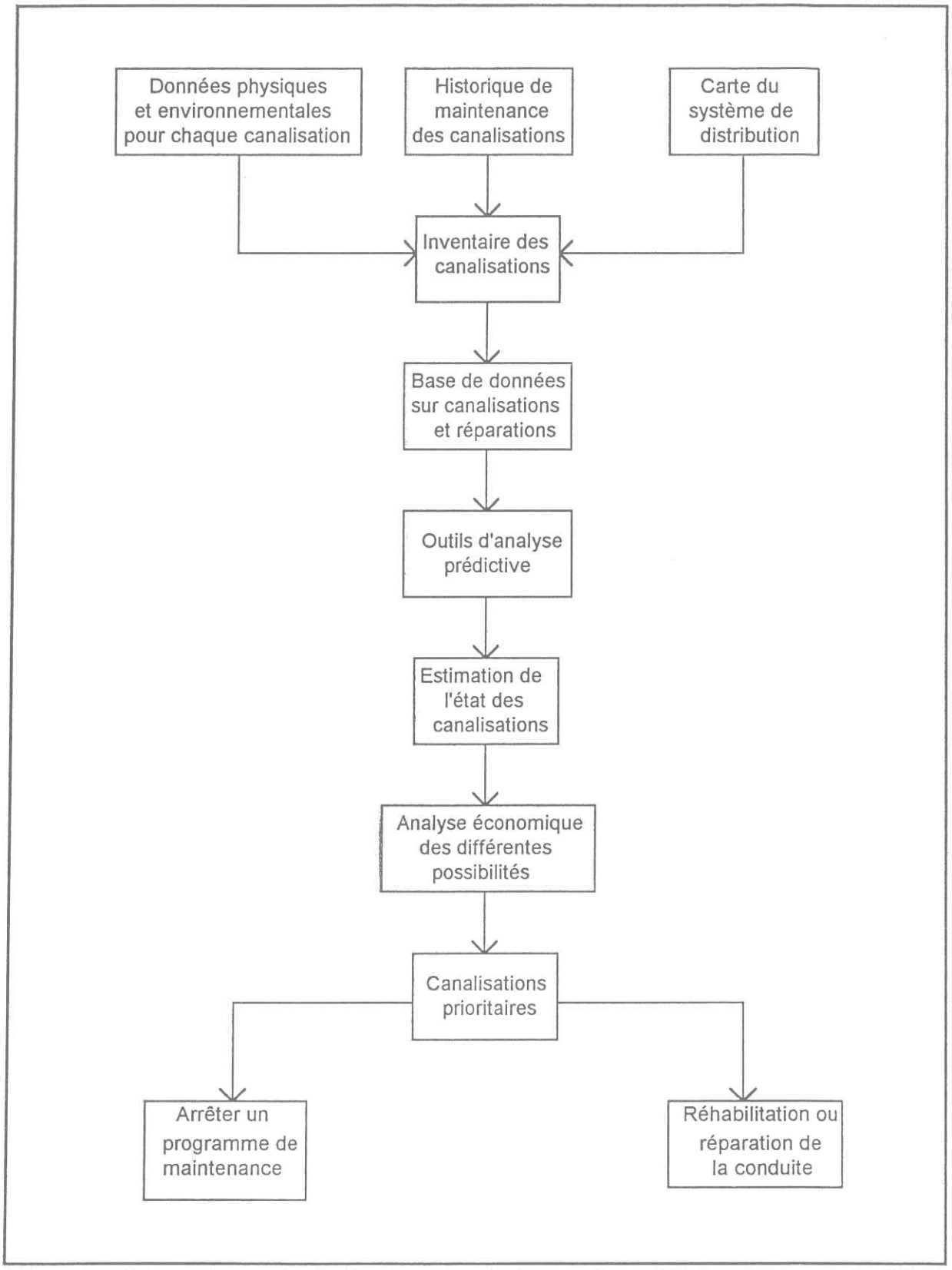


Figure 2 : Démarche employée dans le cadre de l'approche prédictive [O'day et al., 1989]

- **Ruptures** dues à l'action combinée de la corrosion sur la conduite et de mouvements de sols (vibrations, séismes, travaux divers). Une rupture peut entraîner une intervention sur le réseau de plusieurs heures, pendant laquelle les abonnés sont éventuellement privés d'eau ou subissent une chute de pression.

Chacune de ces détériorations engendre, outre certains dommages que nous verrons dans le paragraphe suivant :

* des pertes d'eau, d'où augmentation de la production,
* des pertes en énergie (augmentation du temps de pompage),
* des interventions sur le réseau.

1.2.1.2. Dommages divers engendrés par le vieillissement d'une conduite

- Détérioration de la qualité de l'eau :

On peut distinguer deux types de dommages liés à la dégradation de la qualité de l'eau. Le premier est celui qui engendre la non potabilité de l'eau. Il faut alors élaborer un nouveau traitement rendant cette eau potable. Le deuxième concerne l'augmentation du nombre de plainte des abonnés dû à une apparence négative de l'eau. Ceci entraîne alors une baisse d'image de marque du service exploitant.

- Fuites diffuses :

Elles peuvent indirectement déstabiliser la conduite en érodant le lit de pose, d'où rupture au niveau des points fragiles du tuyau.

- Ruptures :

Elles peuvent avoir des incidences indirectes :

- inondation, d'où coupure du trafic sur la chaussée concernée, ou dommage chez un particulier,
- coupure d'eau, donc dommages causés notamment aux industries ou aux centres de santé,
- déstabilisation du lit de pose,
- plaintes des abonnés.

Chacun de ces dommages engendre des coûts indirects, qui peuvent être souvent nettement supérieurs aux coûts directs de main d'oeuvre pour effectuer la réparation.

Comme on le voit, le vieillissement d'une conduite ne correspond pas à un seul phénomène ayant lieu sur la conduite mais plutôt à l'action de plusieurs facteurs propres à la conduite et à son environnement. Nous allons voir dans le paragraphe suivant comment on peut évaluer effectivement le vieillissement à partir des différents phénomènes décrits précédemment.

1.2.2. Évaluation du vieillissement d'une conduite

Le phénomène de vieillissement ne peut être évalué en tant que tel, mais à partir des phénomènes par lesquels ils se manifestent.

1.2.2.1. La qualité de l'eau

La mesure de la dégradation de la qualité de l'eau peut permettre de caractériser l'état de la ou des conduites proches du point de mesure ou de l'endroit où a eu lieu la plainte d'un abonné.

Elle peut renseigner sur l'état de corrosion de la conduite pour ce qui concerne les conduites métalliques. C'est le cas lorsque des produits de corrosion sont transportés à travers le réseau. On a par exemple le phénomène d'eau rouge, la substance rouge correspondant à l'oxydation d'ions ferreux en ions ferriques principalement dans les eaux douces, acides ou désaérées. Certaines autres substances comme le manganèse peuvent également entraîner la coloration de l'eau (en noir cette fois).

On peut également être renseigné sur la corrosion bactérienne, en examinant les teneurs en bactéries du fer ou sulfatoréductrices.

Un mauvais état des joints ou des conduites en matière plastique peut également altérer la qualité de l'eau, du fait de l'introduction de substances provenant de l'extérieur, comme des hydrocarbures.

Afin de bien connaître le vieillissement, il est important d'étudier la corrélation entre l'âge de la conduite et l'apparition d'une mauvaise qualité de l'eau. Cependant il est très difficile d'identifier la conduite "malade", car on ne peut dire avec précision quelle est celle qui a détérioré la qualité de l'eau.

1.2.2.2. La diminution de la capacité de transport

Cette valeur correspond à la diminution du diamètre interne de la conduite ou à l'augmentation du coefficient de rugosité. Ceci est engendré par l'entartrage calcaire sur les conduites non-revêtues d'un revêtement interne mais aussi par l'apparition de protubérances dues à des produits de corrosion.

On peut la caractériser de différentes façons :

- soit par la mise en évidence de l'augmentation de pertes de charges sur le réseau ou sur une partie, augmentation non induite par l'augmentation de la consommation. Ceci peut se faire par exemple à partir de modélisation hydraulique du réseau.

- soit à partir d'échantillons de canalisations relevés directement sur le réseau, échantillons étudiés par la suite. Cependant ces relevés demeurent ponctuels et leur interprétation est difficilement généralisable.

Le nouveau diamètre ou le nouveau coefficient de rugosité estimés ou mesurés sur une conduite donne alors un état de son entartrage. Si cela est possible on peut essayer là aussi de définir la corrélation pouvant exister entre la dégradation de la conduite et certains facteurs comme son âge, sa nature, la qualité de l'eau à l'entrée dans le réseau.

1.2.2.3. Les fuites diffuses

Les fuites diffuses sont mises en évidence lors de mesures bien spécifiques ou lorsque la tranchée d'une conduite est ouverte. Elles n'entraînent pas en général de réparations sur la conduite. Ce sont elles qui en général abaissent le rendement d'un réseau et ne lui permettent pas d'avoir une valeur supérieure à 90 ou 95%.

Elles caractérisent, soit une fragilisation du tuyau par de petites ouvertures , soit un mauvais état des joints qui deviennent alors poreux. On peut les constater de différentes manières :

- soit par le constat d'une diminution importante du rendement de réseau,

- soit par une augmentation croissante de la consommation de nuit,

- soit par des campagnes de mesures sur le terrain, par des techniques de corrélation acoustique. Ce genre de mesures permet alors de retrouver précisément quelles sont les conduites les plus dégradées.

L'état de la conduite est alors déterminé par le nombre et la taille des fuites que l'on a pu détecter.

1.2.2.4. Les ruptures ou les fuites apparentes

Ces ruptures sont celles qui entraînent automatiquement une intervention sur le réseau. Dans toute la suite du mémoire c'est ce qu'on appelle "défaillance". Elles sont mises en évidence, soit parce qu'elles entraînent au niveau de la chaussée une inondation plus ou moins importante, soit parce que l'on constate une augmentation brutale de la consommation de nuit. Dans ce cas on peut alors retrouver avec précision l'endroit de la rupture ou de la fuite par un système de jeu de vannes.

Ces casses caractérisent bien sûr un mauvais état de la conduite et une certaine fragilisation. Elles peuvent avoir lieu à la suite d'une corrosion dans le temps de la conduite corrélée à un mouvement de sol ou à une augmentation de la pression interne. Elle correspond soit à une diminution de la paroi de la conduite, soit à la formation de trous.

En général ces ruptures sont différentes selon le diamètre. Elles peuvent également concerner les fuites au niveau des joints qui sont assez importantes pour devenir apparentes.

L'état de la conduite et son vieillissement peuvent être caractérisés par un taux de fuites sur un an ou sur une période plus longue, ou bien par un certain nombre de fuites pendant une année (Cf. figure 3).

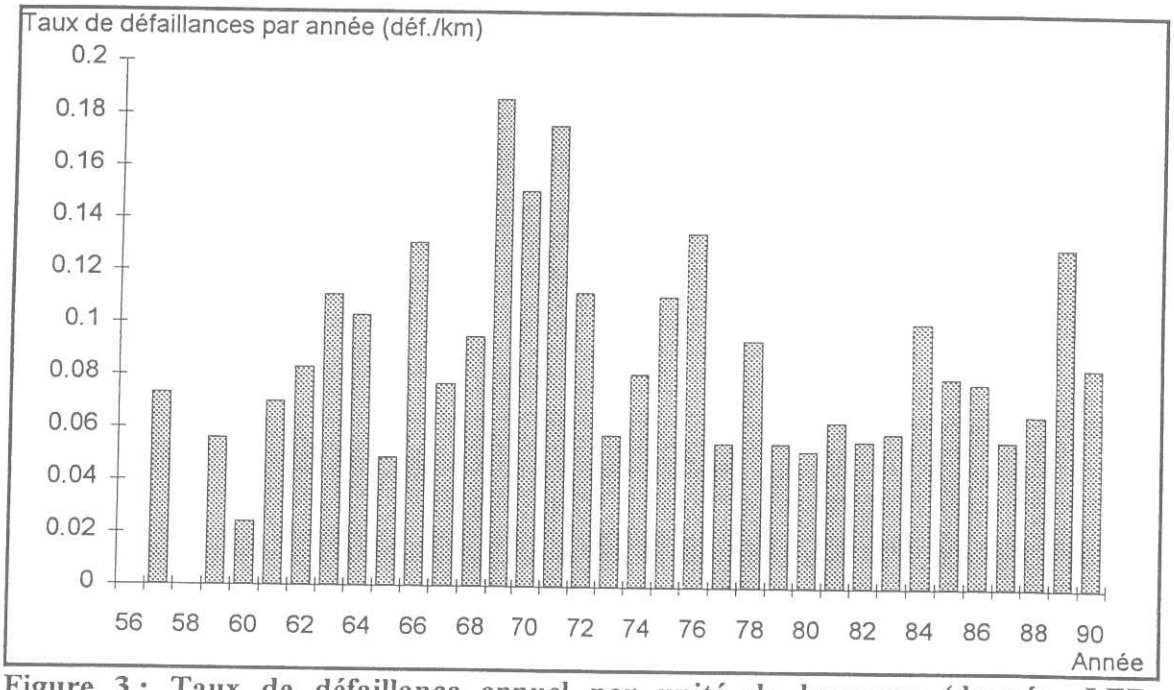


Figure 3 : Taux de défaillance annuel par unité de longueur (données LED-Bordeaux, 1991) [Eisenbeis, 1991]

Le tableau 2 présente les différentes manières d'estimer l'état de dégradation d'une conduite.

1.3. Conclusion

Dans toutes les politiques de renouvellement on essaie de définir un taux de dégradation des conduites. En général c'est le taux de ruptures qui semble prépondérant dans le choix des conduites à remplacer, même si certains services essaient de prendre en compte également la qualité de service offerte aux consommateurs. Ce taux de ruptures fait cependant partie de la qualité de service puisqu'une rupture peut entraîner une coupure d'eau chez l'abonné.

D'autre part des services, aux Etats-Unis notamment, essaient de prévoir la dégradation des conduites afin d'établir des politiques de renouvellement basées sur des analyses économiques.

Tableau 2 : Manifestations et mesures de la dégradation des conduites d'eau potable

Manifestations :	- Dégradation de la qualité de l'eau	- Diminution de la capacité de transport	- Augmentation des fuites diffuses	- Ruptures ou fuites apparentes
Origines :	- Corrosions internes (chimiques ou biologiques) - Porosité de joints ou tuyaux en matière plastique	- Entartrage - Protubérance de produits de corrosion	- Corrosions internes et externes - Détérioration des joints	- Corrosions internes et externes - Dégradation du lit de pose - Mouvements de sol
Mesure du vieillissement :	- Mesures chimiques - Plaintes des consommateurs	- Mesure de diamètre ou de coefficients de rugosité - Chutes de pression	- Nombre de fuites/km - Débit des fuites	- Taux de fuites ou de casses annuel

Parmi les méthodes d'estimation du vieillissement prises en compte, celle concernant les ruptures de conduites est la plus simple à mettre en oeuvre. En effet puisqu'elles entraînent obligatoirement une intervention sur le réseau, il suffit de décrire ces ruptures et de les caractériser (date, causes, réparation) pour pouvoir l'utiliser. Les autres méthodes le permettent plus difficilement, car elles nécessitent de faire des campagnes de mesure et ne sont pas toujours aussi précises quant à l'endroit exact de la dégradation.

Cette relative simplicité du recueil d'informations sur les ruptures permet d'avoir des données en nombre suffisant pour espérer les utiliser statistiquement et aboutir à des modèles de prévision de vieillissement. Certains services possèdent déjà ces données depuis un temps relativement important. C'est pourquoi cette étude concerne la prévision des défaillances, c'est à dire des fuites ou ruptures au niveau du tuyau ou du joint entraînant une réparation de la conduite.

Dans le chapitre suivant nous allons décrire les défaillances et leurs causes. Puis nous verrons quels résultats ont déjà été obtenus et quel modèle de prévision semble le mieux approprié.

CHAPITRE 2 :
LES DEFAILLANCES DES CONDUITES D'EAU POTABLE : CAUSES ET MODELES DE PREVISION DEJA EXISTANTS

Le chapitre précédent nous a montré que les ruptures ou fuites apparentes peuvent donner des indications précises sur la dégradation des conduites, d'autant que leur prise en compte ne nécessite rien de plus que l'observation d'interventions déjà existantes.

Après avoir défini les défaillances, nous présenterons leurs causes possibles. Puis nous verrons quels modèles de prévision existent déjà.

2.1. DESCRIPTION ET CAUSES DES DEFAILLANCES

L'EPA (Environmental Protection Agency) a regroupé et utilisé des données sur les fuites ou ruptures dans plusieurs services américains. Les données recueillies ne sont pas toujours basées sur les mêmes définitions concernant les fuites. On peut distinguer [O'day et al., 1989] :

- les fuites (leak repair) ; ce sont tous les événements entraînant une réparation sur les conduites, les hydrants, les branchements.
- les fuites sur tuyau (main leak),
- les fuites sur joint (joint leak),
- les ruptures sur conduites (main breaks) ; elles représentent les défaillances structurelles du tuyau ou de l'emboîtement dues à une surcharge excessive, une détérioration du lit de pose, un contact avec les autres structures, la corrosion ou une combinaison entre ces conditions.

Comme on le voit les définitions données différencient surtout les fuites sur conduites de celles ayant eu lieu sur les différents organes du réseau (vannes, pompes, etc.). Ensuite on distingue aussi les fuites et les ruptures.

Sachant que la limite entre rupture et fuite n'est pas toujours bien définie (une rupture est une fuite, mais une fuite n'est pas obligatoirement une rupture), nous définirons de la façon suivante la défaillance : une défaillance est une rupture ou une fuite apparente nécessitant une intervention sur le réseau. Elle peut avoir lieu soit sur le tuyau, soit sur le joint. Ceci exclut les fuites ayant lieu sur les branchements, ainsi que celles ayant lieu sur les organes du réseau, tels que les pompes, les vannes, les hydrants et autres organes régulateurs.

2.1.1. Description des défaillances

Les défaillances au niveau du tuyau peuvent être de différentes natures.(Cf. figure 4).

Elles peuvent également correspondre à un éclatement total du tuyau suite à une action violente qu'il a pu subir (Coup de bélier, mouvement de terrain...). La figure 5 montre une photographie d'un tuyau ayant subi une telle action.

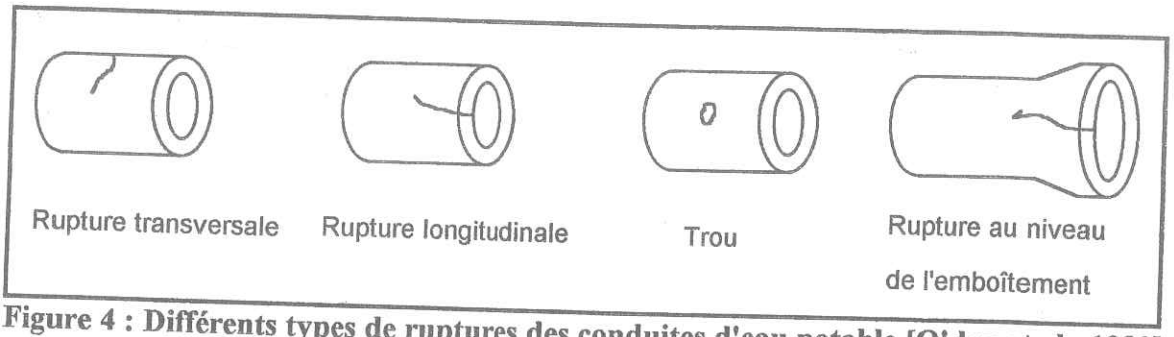


Figure 4 : Différents types de ruptures des conduites d'eau potable [O'day et al., 1989]

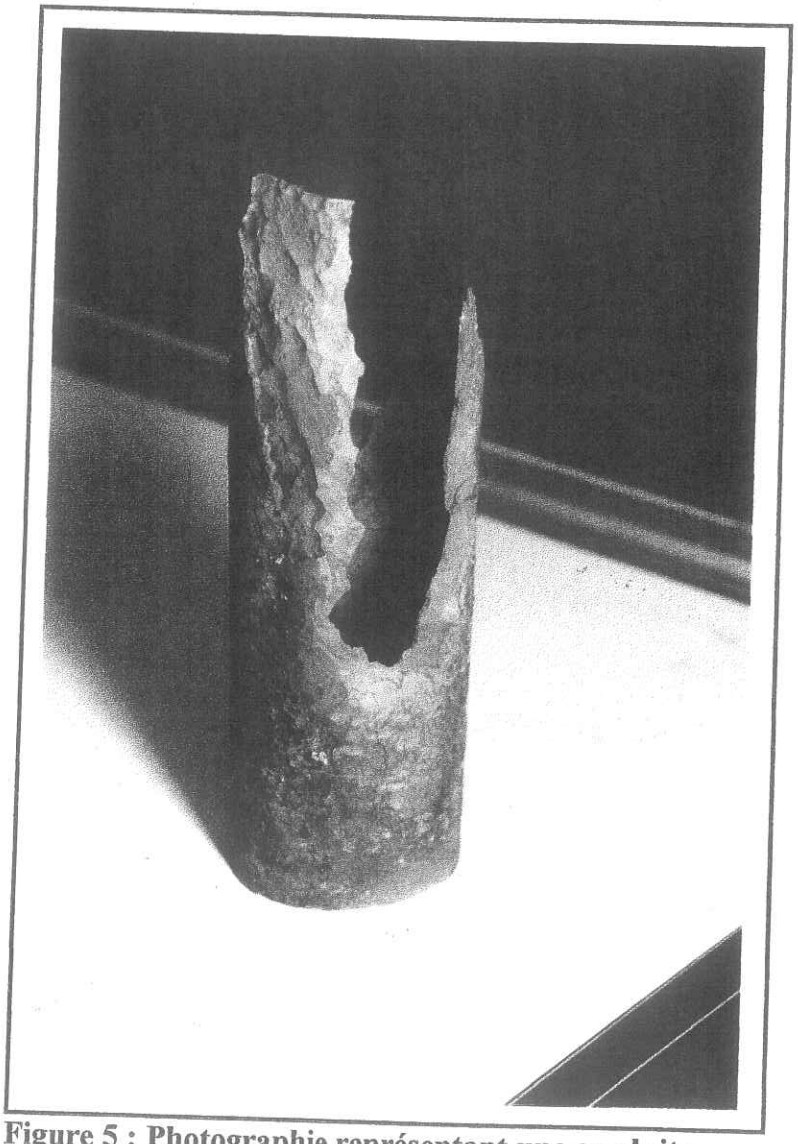


Figure 5 : Photographie représentant une conduite ayant subi une rupture.

Une rupture ou une fuite correspond à un affaiblissement de la conduite, c'est à dire soit à une diminution de la paroi, en un endroit précis (piquage) ou non (fissure) combinée à une augmentation de contraintes sur la canalisation, soit à une non-étanchéité du joint.

La figure 6 montre la fréquence des défaillances en fonction du diamètre et du type de défaillances, les ruptures transversales ont lieu surtout sur les petits diamètres alors que les gros diamètres subissent plutôt des ruptures longitudinales ou des piquages.

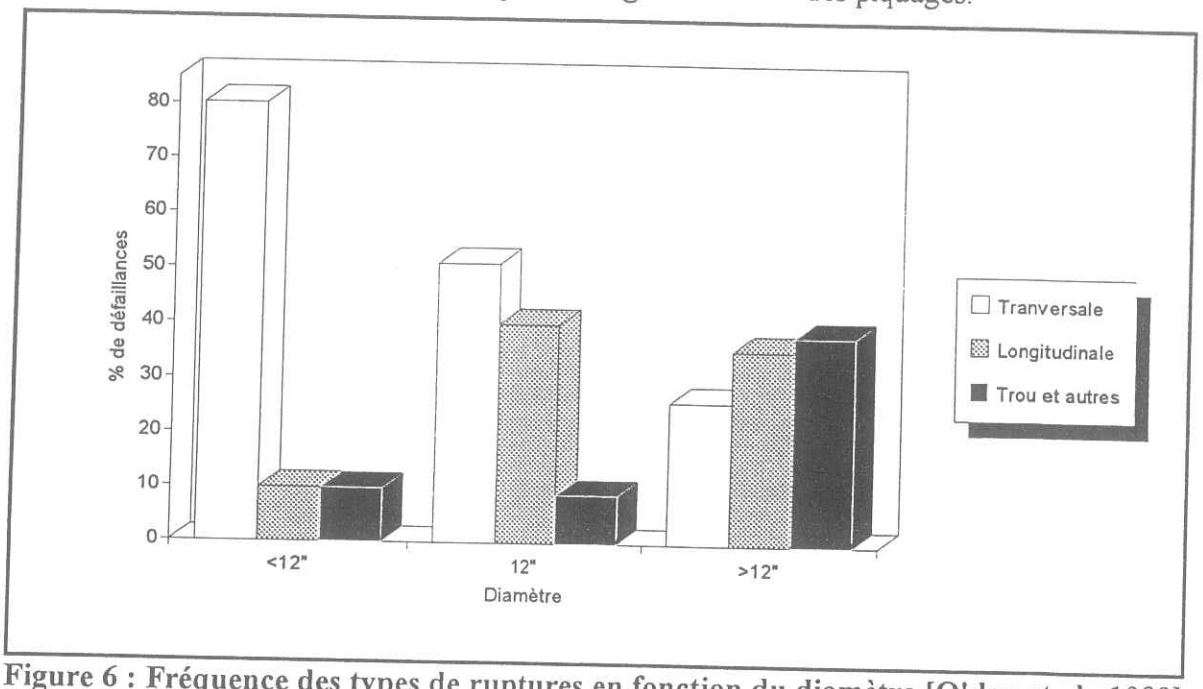


Figure 6 : Fréquence des types de ruptures en fonction du diamètre [O'day et al., 1989]

De même les fuites sur tuyaux ont surtout lieu sur les petits diamètres, alors que les diamètres plus élevés ont surtout des fuites au niveau des joints (Cf. fig. 7).

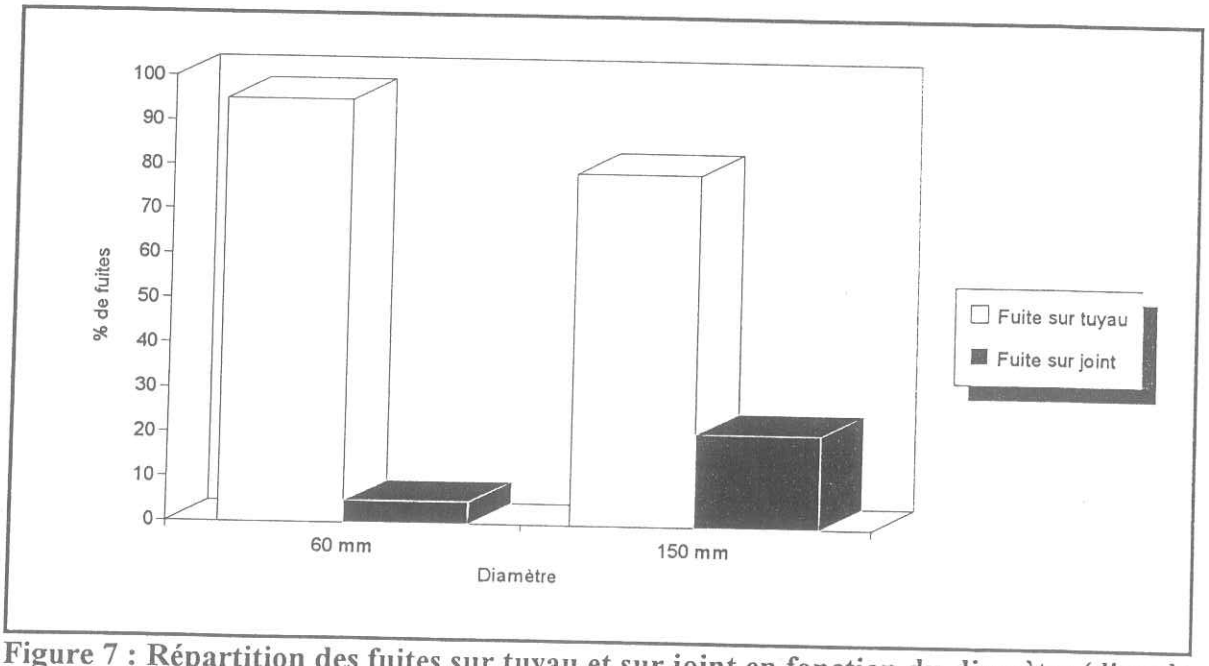


Figure 7 : Répartition des fuites sur tuyau et sur joint en fonction du diamètre (d'après les données recueillies sur la LED-Bordeaux) [Eisenbeis, 1991]

2.1.2. Causes de l'apparition des défaillances des conduites d'eau potable

2.1.2.1. Phénomènes entraînant l'apparition des défaillances

- Phénomènes mécaniques

La figure 8 récapitule les différentes contraintes mécaniques que peut subir une canalisation.

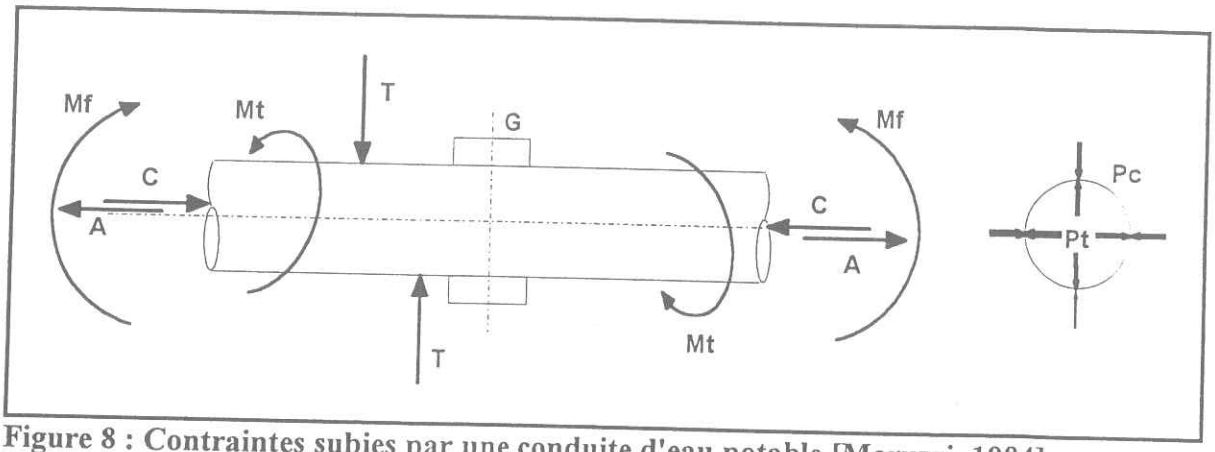


Figure 8 : Contraintes subies par une conduite d'eau potable [Moruzzi, 1984]

Ces contraintes sont les suivantes :

- moment de flexion (Mf) susceptible de provoquer une courbure de la canalisation,
- moment de torsion (Mt) susceptible de tordre la canalisation autour de son axe,
- contrainte de cisaillement (T),
- effort de traction axiale (A),
- contrainte de compression axiale (C),
- surpression interne (Pt) susceptible d'entraîner l'éclatement de la conduite,
- dépression interne (Pc) ou surpression externe pouvant provoquer un écrasement.

Les causes de variation de ces contraintes sont les suivantes :

- mouvement violent du sol (Mf, Mt, T, A, C, Pc),
- mouvement lent du sol (Mf, T, A),
- transmission directe excessive de charges accidentelles en surface (Mf, T, Pc),
- températures peu élevées (Pt, A),
- coup de bélier (Pc, Pt, A),
- travaux de voirie (Mf, Mt, T, A, Pc).

- Phénomènes chimiques ou biologiques

Ces phénomènes sont représentés principalement par la corrosion externe ou interne. Il s'établit dans ce cas un échange d'ions entre le milieu extérieur (sol ou eau) et la paroi de la conduite.

Cette corrosion est de deux types :

*** La corrosion interne**

Elle est liée à la qualité de l'eau et à ses différentes caractéristiques (pH, taux d'oxygène dissous, équilibre calco-carbonique, existences de bactéries du fer ou sulfato-réductrices). Elle entraîne une diminution de la paroi en certains points précis du tuyau, lié le plus souvent à l'existence de dépôts d'où diminution du diamètre interne.

*** La corrosion externe**

Elle correspond à un échange d'ions entre le sol et la paroi de la conduite et peut avoir diverses origines :

- Les courants vagabonds,
- Le sol,
- L'hétérogénéité de surface,
- L'hétérogénéité du sol,
- L'hétérogénéité par contact.

Tous ces facteurs de corrosion externe sont présentés plus précisément en annexe 2.

2.1.2.2. Eléments influençant ces phénomènes

La figure 9 nous montre les différentes causes qui peuvent entraîner l'affaiblissement d'une conduite d'eau potable et donc augmenter le risque d'apparition d'une défaillance. Elles sont regroupées en différentes classes [Kremer, 1990] :

- les éléments propres au type de la canalisation,
- les éléments liés à l'exploitation des réseaux,
- les éléments extérieurs aux réseaux.

Tous ces éléments sont présentés en annexe 2. Cependant une canalisation d'eau potable fait partie d'un ensemble de structures pouvant avoir des relations entre elles et pouvant subir des événements tout à fait fortuits. Il est donc impossible de lister toutes les causes agissant sur les défaillances.

Certains facteurs doivent cependant être prépondérants par rapport aux autres. C'est ce que nous allons essayer de déterminer dans la suite de l'étude (Cf. Chapitre 5). Mais auparavant il

est important de connaître les modèles de prévision des défaillances sur les conduites déjà existants.

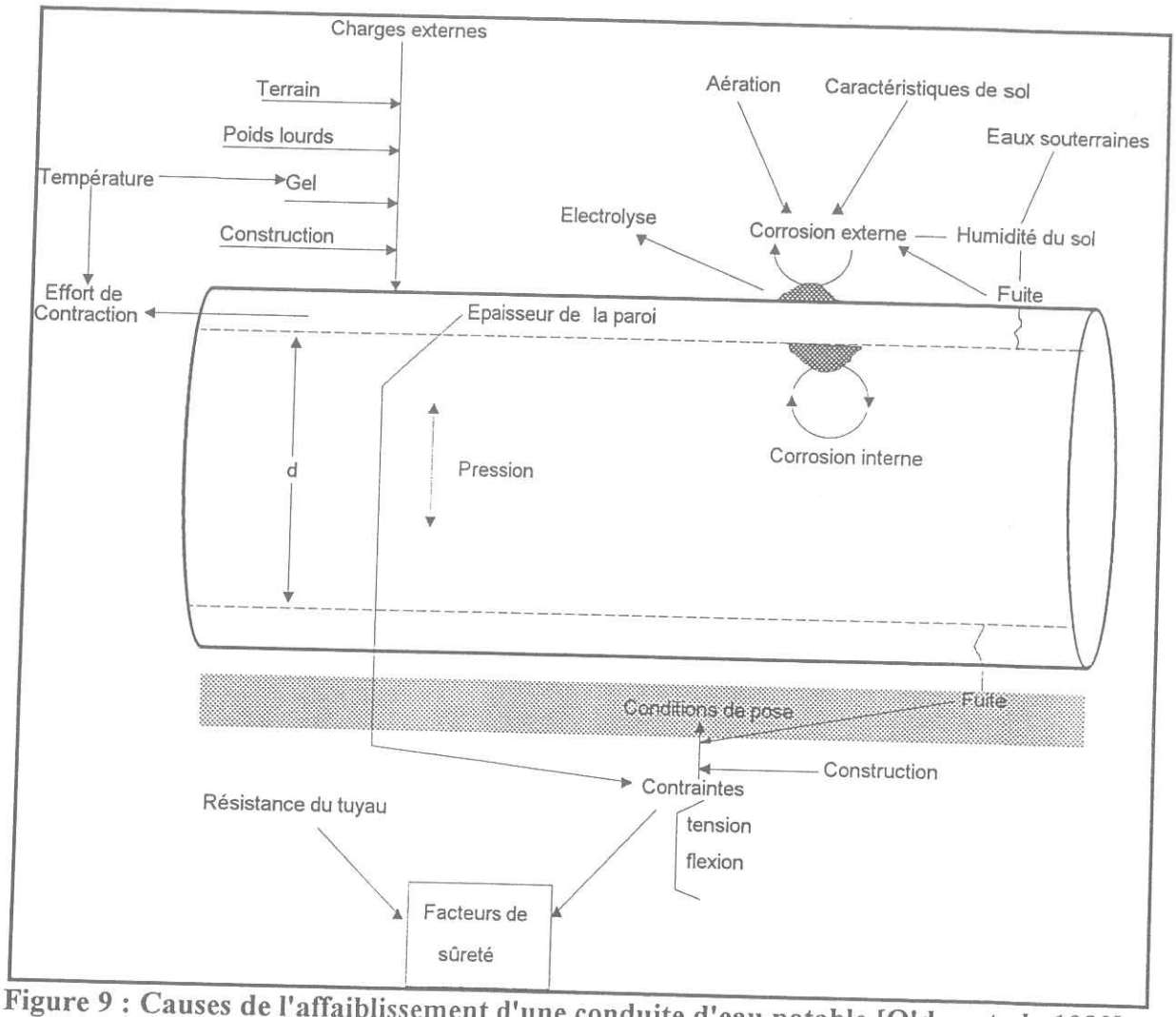


Figure 9 : Causes de l'affaiblissement d'une conduite d'eau potable [O'day et al., 1989]

2.2. LES MODELES DE PREVISION DES DEFAILLANCES DES CONDUITES D'EAU POTABLE

Des études descriptives sur les défaillances des canalisations d'eau potable ont déjà eu lieu Cependant si elles permettent de mettre en évidence des corrélations entre le nombre de défaillances et certains facteurs cités précédemment, elles n'ont pas permis de définir quels en étaient les facteurs les plus importants ni même de modéliser l'apparition des défaillances.

Pour l'instant seuls des modèles américains existent. Ils sont de deux types : des modèles physiques et des modèles statistiques.

2.2.1. Les modèles physiques de prévision des défaillances

Ces modèles permettent d'évaluer l'action des corrosions internes et externes, de la surcharge sur la canalisation et enfin les contraintes agissant sur ces canalisations. On peut alors estimer

l'état structurel dans lequel se trouve la canalisation en comparant ces contraintes à la résistance originelle du tuyau.

Plusieurs études ont été réalisées et sont présentées ci-après.

2.2.1.1. Etude de Vancouver (Canada) [O'day et al., 1989]

En 1978, Vancouver avait entrepris une étude sur les "temps de défaillance" des conduites pour déterminer la durée de mise en service des 42 miles de conduites en fonte installées à partir de 1900. Le modèle permet de prédire la profondeur maximale de piquage due à la corrosion externe.

Le service a conduit son étude de la manière suivante :

1 - Surveillance des enregistrements des réparations puis discussions pour déterminer les problèmes de la zone,

2 - Evaluation des sources de courants vagabonds,

3 - Analyse de l'eau pour déterminer ses propriétés de corrosivité,

4 - Analyse d'échantillons de sol à la profondeur de la canalisation pour déterminer le pH, l'humidité, la résistivité, le potentiel redox et la teneur en sulfure,

5 - Examen physique de vingt canalisations en ce qui concerne l'uniformité et de la profondeur de piquage de la paroi.

Les résultats de cette enquête sont les suivants :

1 : La diminution moyenne d'épaisseur de la paroi due à la corrosion est de 1,65 mm sur toutes les conduites observées,

2 : La résistivité du sol varie entre 3400 et 900000 ohm-cm, le pH entre 4.6 et 9.4, et le sol est faiblement aéré,

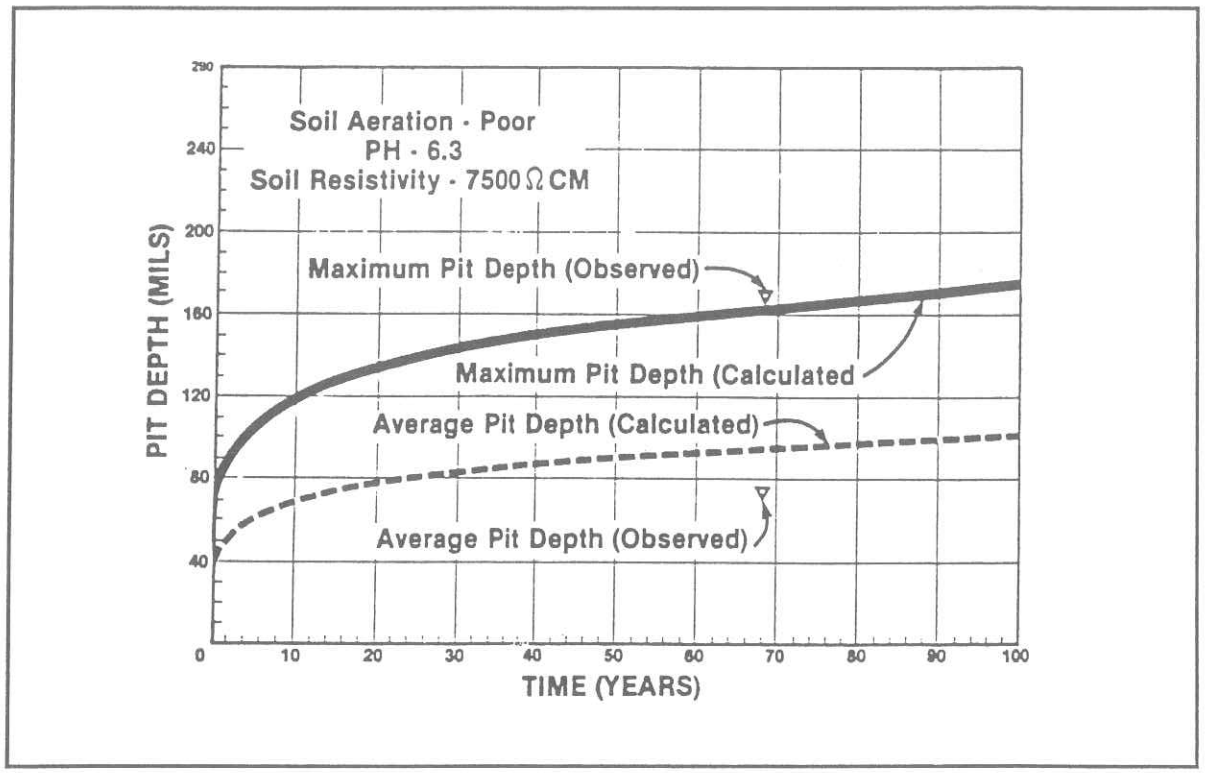
3 : Aucune trace de sulfure n'a été détectée,

4 : L'eau n'est en général pas entartrante et contient des taux d'oxygène dissous importants.

Les données de sol ont été alors incorporées au modèle pour estimer la profondeur de piquage de la paroi. A partir de là une trentaine de courbes correspondant aux différentes combinaisons des données du sol ont été établies. Une de ces courbes se trouve en figure 10.

Afin d'établir un modèle de prédiction des défaillances, il est nécessaire de connaître l'épaisseur minimale de la paroi qui provoque la casse de la canalisation. Cette donnée, non précisée dans

la publication, est déterminée à partir des forces et contraintes qui s'appliquent sur la canalisation.



"Pit depth" correspond à la profondeur de piquage de la paroi de la canalisation (1 mill = 0,0254 mm)

Figure 10 : Courbe présentant le modèle physique de Vancouver [O'day et al., 1989]

En résumé afin de pouvoir utiliser ce modèle il est nécessaire de connaître les informations suivantes concernant la canalisation et son environnement :

- La date de pose de la canalisation,
- L'épaisseur originelle de la canalisation,
- Le type de sol,
- La diminution de l'épaisseur de la paroi de la canalisation due à la corrosion interne.

Le principal avantage de ce modèle est qu'il permet d'inclure des facteurs qui ont une influence directe sur la structure de la canalisation. Cependant il est à noter que les charges externes et les conditions de lit de pose ne sont pas prises en compte dans ce modèle.

2.2.1.2. Modèle de l' "US Army Corp of Engineers Construction Engineering Research Laboratory" (CERL) [O'day et al., 1989]

Le CERL a développé un modèle physique appliqué aux canalisations de distribution de gaz. Ce modèle estime également l'influence de la corrosion externe sur l'épaisseur de la paroi de la conduite et calcule un "index de corrosion" (CSI) variant de 1 à 100, une valeur de 100

correspondant à une conduite neuve et une valeur de 0 correspondant à une canalisation complètement détériorée. Le CSI est calculé par l'équation suivante :

$$CSI = 100\left(1 - \frac{PAV}{T}\right)$$

avec : PAV = Valeur de la profondeur de piquage de la paroi
T = épaisseur de la paroi de la canalisation

Il a été observé que la casse de la canalisation a lieu lorsque le CSI est de l'ordre de 30.

Pour chaque conduite celui-ci peut être estimé à partir de différentes méthodes telles que les mesures directes sur la canalisation, les techniques de polarisation électrique ou des modèles mathématiques. Ensuite sont estimées les influences de différents facteurs déjà vus précédemment, à savoir l'effet du sol (pH, résistivité), l'effet d'un revêtement extérieur sur la corrosion, l'effet des sulfates et l'effet de l'humidité.

Cette approche peut être appliquée aux canalisations d'eau potable, en tenant toutefois compte des différences existant entre les caractéristiques des réseaux, telles qu'une pression plus élevée pour les canalisations d'eau ou la possibilité de corrosion interne due aux caractéristiques chimiques de l'eau. Le CSI correspondant à l'apparition des casses n'est pas précisé dans la publication.

2.2.1.3. Etude du Philadelphia Water Department (P.W.D.) [O'day et al., 1989]

Le P.W.D. a développé un modèle d'évaluation de l'état des conduites d'eau potable. Ce modèle est destiné à évaluer les facteurs structurels de sécurité de conduites en fonte grise exposées à des variations de conditions environnementales. Il estime les charges internes et externes, les actions des corrosions internes et externes, puis prédit l'épaisseur de la paroi et calcule un taux de sécurité donnant l'état de la canalisation.

Notons que, comme les analyses techniques concernant les deux types de fonte (grise ou ductile) sont différentes, ce modèle ne peut être appliqué qu'aux conduites en fonte grise.

2.2.2. Les modèles statistiques de prévision des défaillances des conduites d'eau potable

Ces modèles appliquent des techniques statistiques aux données de réparations afin de définir les facteurs contribuant à l'apparition des défaillances. Ces techniques peuvent être des analyses de régression multivariable (Etude de l'Environmental Protection Agency ou du Des Moines Water Works) ou des analyses plus spécifiques comme la régression de Cox (Massachusetts Institute of Technology).

2.2.2.1. Modèle de l'EPA (Environmental Protection Agency) [Clark et al., 1982]

L'EPA a effectué une analyse sur les réparations des canalisations d'eau potable dans deux services d'eau (un petit et un grand). Les données recueillies dans ces deux services renseignaient sur les facteurs suivants :

* Diamètre de la canalisation,
* Longueur de la canalisation,
* Nombre total de défaillances,
* Type de canalisation,
* Corrosivité du sol enrobant la canalisation,
* Conditions de pression,
* Age de la canalisation,
* Réparations effectuées sur la canalisation.

Le tronçon de base défini pour cet inventaire est un tronçon séparé des autres soit par des raccordements hydrauliques, soit par des changements de diamètre. Au total 307 tronçons ont été répertoriés.

Cette analyse a permis de différencier les deux phénomènes suivants que l'on peut considérer comme les variables à expliquer :

* l'intervalle de temps qui existe entre la pose de la canalisation et la première défaillance.
* après cette première défaillance, le nombre de défaillances en fonction du temps écoulé qui semble croître de manière exponentielle.

L'équipe de l'EPA a donc développé deux équations, l'une estimant la période jusqu'à la première défaillance, l'autre estimant le nombre de défaillances à la suite de la première. Ces équations ont été développées pour chacun des deux services. Ce sont les suivantes :

* Age à la première défaillance

$NY = 2,9 + 0,442\ D + 0,017\ P + 0,412\ I + 0,32\ RES$

(Coefficient de détermination, $R^2=0,34$)

pour le petit service d'eau

$NY = 11 + 0,263\ D - 0,006\ P + 0,773\ I + 0,253\ RES - 0,00006\ LH + 15,69\ T$

($R^2=0,34$)

pour le grand service d'eau

$NY = 4,13 + 0,338\ D - 0,022\ P + 0,265\ I + 0,0983\ RES - 0,0003\ LH + 13,28\ T$

($R^2=0,23$)

pour les deux services combinés

avec

NY = Nombre d'années entre la pose et première défaillance

D = Diamètre de la canalisation en pouces (= 0.3048 m)

P = Pression absolue dans la canalisation en livres par pouce carré (= 7,02 Pa)

I = Pourcentage de linéaire en zone industrielle

RES = Pourcentage de linéaire en zone résidentielle

LH = Longueur de canalisation en sol hautement corrosif

T = Type de canalisation (1=métallique, 0=béton armé)

* Nombre de défaillances à la suite de la première

$$REP = 0,386\ e^{0,139\ PRD} e^{0,0602\ A} e^{0,0208\ DEV} SL^{-0,016} SH^{0,025} \qquad (R^2=0,49)$$

pour le petit service

$$REP = 0,386\ e^{0,7198\ T} e^{0,040\ PRD} e^{0,0862\ A} e^{0,0189\ DEV} SL^{0,014} SH^{0,069} \qquad (R^2=0,47)$$

pour le grand service

$$REP = 0,386\ e^{0,7197\ T} e^{0,044\ PRD} e^{0,0865\ A} e^{0,0121\ DEV} SL^{0,014} SH^{0,069} \qquad (R^2=0,47)$$

pour les deux services combinés

avec

REP = Nombre de défaillances

PRD = Pression différentielle

A = Période depuis la première défaillance

DEV = Pourcentage de canalisations se trouvant en zone développée

SL = Surface de canalisation en contact avec un sol faiblement corrosif

SH = Surface de canalisation en contact avec un sol corrosif

T = Type de canalisation (=1 pour canalisation métallique
=0 pour béton armé)

On peut noter que, dans ces équations de nombre de défaillances, le diamètre n'est pas pris en compte.

2.2.2.2. Des Moines Water Works Study [O'day et al., 1989]

Le "Des Moines Water Works Study" a mené une étude afin d'établir les causes principales d'augmentation du nombre de casses des canalisations par année. Basés sur une analyse descriptive, les résultats ont permis de conclure que 75% de toutes les défaillances étaient dus à l'action des sols corrosifs sur les canalisations en fonte.

D'autre part plusieurs modèles de régression multiple basés sur des modèles linéaires et logarithmiques ont été étudiés. L'équation de prédiction est la suivante :

$$Age = 65,78 + 0,028\ SR - 6,338\ pH - 0,049\ r$$

avec

Age = Age de la canalisation à la première défaillance (en années)

SR = Résistivité du sol saturé (en Ohm-cm)

pH = pH du sol

r = Potentiel redox (en millivolts)

Cette équation montre bien l'influence du sol sur l'âge à la première défaillance. Ainsi une augmentation de résistivité du sol de 1000 Ohm-cm fait croître l'âge de première défaillance de 28 ans. Ici encore le diamètre n'a pas été pris en compte dans l'équation.

2.2.2.3. Massachusetts Institute of Technology (MIT) [Andreou et al., 1987][Andreou, 1986][Marks et al., 1987]

Le MIT a développé un modèle probabiliste de prédiction des défaillances à partir de réseaux situés à New Haven et Cincinnati. Pour cette étude des données informant sur les caractéristiques des canalisations et leur environnement ont été collectées. Au total 1391 tronçons de canalisations ont été répertoriés.

Le modèle utilisé est le modèle des risques proportionnels, appelé aussi modèle de Cox, qui permet de donner une "fonction de survie" de la canalisation ainsi qu'une fonction de risque, en fonction des différents facteurs recueillis.

Les relations suivantes ont pu être mises en évidence :

* Une augmentation de pression est un bon indicateur de la fonction de risque lorsque la canalisation n'a pas eu de casse antérieure,

* Un âge assez faible à la deuxième défaillance indique une forte probabilité de futures casses,

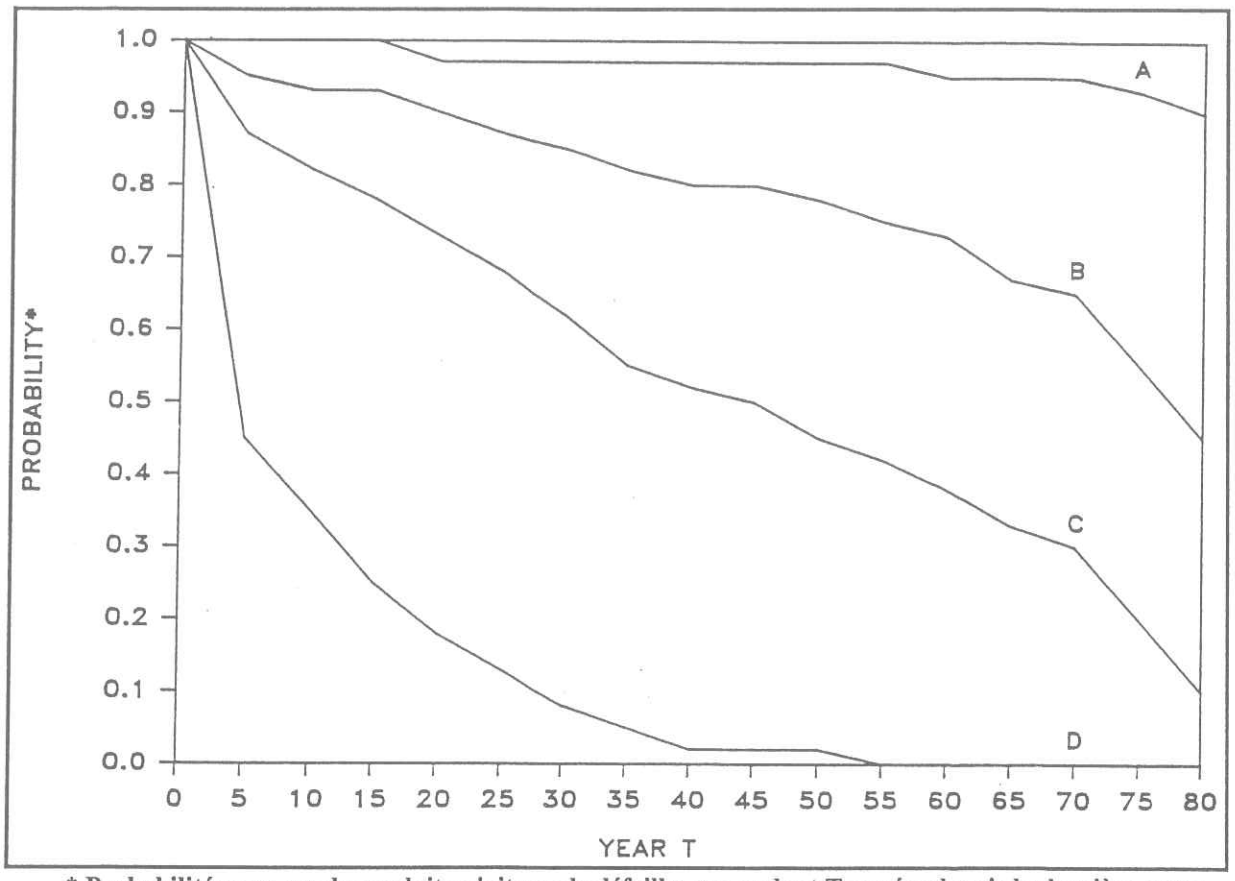
* La période au cours de laquelle a été posée la canalisation est importante dans la détermination de la probabilité de défaillance.

* Chaque défaillance qui a déjà eu lieu triple la probabilité d'apparition d'une nouvelle casse.

* Un tronçon plus long a tendance à avoir proportionnellement moins de défaillances qu'un tronçon plus court, un tronçon étant homogène par le diamètre et séparé d'un autre tronçon par des particularités hydrauliques (pompes, vannes, croisements avec d'autres canalisations, ...).

* Le risque de première casse diminue avec l'âge de la canalisation jusqu'à un âge d'environ 30 ans, puis augmente brutalement.

La figure 11 montre quatre courbes de survie différentes pour des canalisations de caractéristiques bien précises.



*** Probabilité pour que la conduite n'ait pas de défaillance pendant T années depuis la dernière casse**

Courbe A : 30m de canalisations installées entre 1930 et 1935, se trouvant dans une zone rurale et qui ont subi deux casses, la dernière ayant eu lieu au moins 77 ans après la date d'installation.

Courbe B : 30m de canalisations, sans casse antérieure, installée après 1950, se trouvant dans une zone rurale et subissant de très importantes pressions internes de l'ordre de 173 psi.

Courbe C : 30m de canalisations, se trouvant partiellement dans une zone de développement maximum, ayant subi deux casses antérieures la dernière casse ayant eu lieu 4 ans après la date de pose.

Courbe D : 4500m de canalisations, se trouvant entièrement dans une zone de développement maximal, ayant subi deux défaillances 4 ans après la date d'installation.

Figure 11 : Exemples de courbes de survie de conduites d'eau potable (New Haven) [Andreou et al., 1987]

Ce modèle de prédiction semble être bien adapté à un modèle de prévision des défaillances des canalisations, car il permet de mettre en évidence l'influence des différents facteurs pris en compte et donne en plus des probabilités de défaillance. Ces probabilités pourront ainsi être utilisées dans des modèles économiques de prévision du renouvellement.

2.3. CHOIX DU MODELE

Dans la suite de l'étude c'est un modèle statistique probabiliste qui a été choisi. Ce choix a été déterminé pour plusieurs raisons.

Tout d'abord, comme on l'a vu précédemment les modèles statistiques permettent de prendre en compte un plus grand nombre de facteurs, ce qui est important, vu la longue liste des causes des défaillances. Ces facteurs peuvent d'ailleurs être qualitatifs, comme par exemple le type de développement en surface (industriel, résidentiel,...).

Ensuite un modèle physique trop théorique ne permettrait pas de prendre en compte la spécificité de certains services, tels que les facteurs géographiques, la façon de travailler dans la région, la pluviométrie, l'existence de nappes plus ou moins profondes. Au contraire un modèle statistique, par définition, utilise des données de terrain et permet de prendre en compte de type de facteur.

Parmi les modèles statistiques c'est un modèle du même type que celui du M.I.T. qui a été choisi et adapté aux réseaux français, ceci du fait des données bien spécifiques qui ont été recueillies. Ces données sont présentées dans le chapitre suivant.

CHAPITRE 3 :
LES DONNEES DISPONIBLES

Le chapitre précédent a permis de retenir un modèle statistique, car ce type de modèle semble le mieux approprié à la prévision des défaillances sur les réseaux d'eau potable. Après avoir défini quels types de données sont nécessaires, nous présenterons de manière générale les réseaux choisis, puis nous verrons quelles données ont été recueillies. Enfin nous montrerons que ces données ont une certaine spécificité d'un point de vue statistique.

3.1. LES DONNEES RECHERCHEES

Rappelons que l'objectif de l'étude est de connaître le vieillissement des conduites d'eau potable en rapport avec l'apparition des défaillances, puis de définir quels facteurs, définis dans le chapitre précédent, sont les plus influents sur ce vieillissement.

Les données que l'on recherche doivent donc concerner des services d'eau ayant conservé trace sur une période suffisamment longue de tous les travaux de maintenance et en particulier de toutes les casses et fuites réparées sur les conduites. Afin de bien comprendre le vieillissement, cette durée doit être du même ordre que la durée de vie d'une canalisation.

Enfin afin de bien connaître l'influence des facteurs de défaillances, les données concernant ces facteurs doivent pouvoir être recueillies assez facilement et doivent, sur les réseaux étudiés, avoir une certaine variabilité.

C'est en fait la première contrainte qui a guidé notre choix, car peu de réseaux possèdent des historiques de maintenance assez longs, alors que les données concernant les caractéristiques des conduites sont assez bien connues.

Le choix s'est donc porté sur deux réseaux urbains et deux syndicats ruraux du Bas-Rhin.

Les réseaux urbains sont le réseau de la Lyonnaise des Eaux-Dumez (LED) sur la Communauté Urbaine de Bordeaux et le réseau de la Compagnie des Eaux de la Banlieue de Paris (CEB) sur les communes de la Presqu'île de Gennevilliers. Pour ces deux services, les données concernant les défaillances sont conservées depuis 40 et 54 ans respectivement.

Les deux réseaux ruraux concernent les syndicats du Kronthal et du Kochersberg et sont gérés par le Service Départemental d'Eau et d'Assainissement du Bas-Rhin (SDEA). Les données y sont conservées depuis 17 ans. Même si elle peut paraître faible cette durée est relativement importante pour un réseau rural.

3.2. PRESENTATION GENERALE DES RESEAUX

3.2.1. Le réseau de la Lyonnaise des Eaux-Dumez à Bordeaux [Eisenbeis, 1991]

Le Service des Eaux intervient dans 22 des 27 communes de la Communauté Urbaine de Bordeaux et la Lyonnaise des Eaux-Dumez en est le gestionnaire depuis 1969. La population desservie est de l'ordre de 580.000 habitants soit environ la moitié du département de la Gironde (Cf carte annexe 3).

3.2.1.1. Les Ressources

Les ressources sont disséminées en grande partie dans l'agglomération bordelaise et sont constituées de :

- ***Sources et captages*** : 12 unités dont le débit s'échelonne entre 120 et 1200 m^3/h et qui sont situées à l'émergence de nappes issues de calcaires de l'oligocène supérieur.

- ***Forages*** : plus de 100 unités, sollicitant différents niveaux aquifères appartenant à trois grands étages stratigraphiques.

Une proportion importante de pompes de forage refoule directement dans le réseau, les sources et autres forages sont reprises par des usines élévatoires qui permettent notamment de pomper l'eau d'un étage inférieur à un étage supérieur du réseau.

Grâce à ces ressources, le Service des Eaux pourrait faire face à une demande de pointe de 350.000 m^3 d'eau par jour.

3.2.1.2. Structures de la distribution

L'ensemble du réseau est naturellement divisé par la Garonne qui traverse la Communauté Urbaine de Bordeaux. On distinguera donc les deux rives de la Garonne.

- Rive gauche de la Garonne

Cette entité comprend la totalité de la ville de Bordeaux, y compris le quartier de Bordeaux-Bastide sur la rive droite, et toutes les communes situées à l'ouest du fleuve. Cette configuration nécessite la mise en place de trois étages de distribution.

** Cote 40-Ville de Bordeaux*

Deux aqueducs mis en service au 19ème siècle amènent l'eau jusqu'à des usines de refoulement situées sur la commune de Bordeaux. Ces usines permettent le refoulement de 90.000 m^3/j.

Une autre usine capte les eaux d'une galerie souterraine et refoule environ 15.000 m^3/j. D'autre part des forages profonds de Bordeaux-Nord peuvent refouler directement dans le réseau.

** Cote 60-Périphérie et Nord-Ouest de Bordeaux*

L'aqueduc de Cap-Bos, construit au début du siècle, collecte des eaux de source et de captage et les amène gravitairement au réservoir de Gajac, d'où une usine refoule dans le réseau une moyenne de 12.000 m^3/j. D'autre part certaines stations relèvent, suivant les besoins, de l'eau prélevée sur la cote 40.

Des châteaux d'eau assurent l'équilibre de l'ensemble.

** Cote 75-Grande périphérie au Sud-Ouest de Bordeaux*

L'usine de Cazeaux refoule les eaux de 4 forages tandis que l'usine de Cap-Roux peut relever l'eau de la cote 40.

- Rive droite de la Garonne

Cette partie est divisée en deux groupes de 3 communes.

** Groupe Floirac-Cenon-Lormont*

Ces trois communes sont alimentées uniquement par des forages profonds. Les fortes variations d'altitude ont nécessité la constitution de quatre étages de distribution : Cote 37, Cote 50, Cote 76, Cote 100.

** Groupe Ambès-Saint Louis de Montferrand-Saint Vincent de Paul*

Constituant la presqu'île du Bec d'Ambès, ces trois communes situées en terrain plat et bas sont alimentées par deux forages refoulant dans les trois réservoirs des trois communes.

3.2.1.3. Caractéristiques du réseau

- Longueur du réseau

La longueur totale du réseau était à la fin du mois de décembre 1990 de 2.960 km. Les tableaux 3 et 4 présentent la répartition des canalisations en fonction des diamètres et des différents secteurs.

Tableau 3 : Répartition des canalisations par diamètre (décembre 1990)

Diamètre (mm)	≤60	80	100	110	125	135	150	162	175
Longueurs (km)	528,7	199,1	787	56,9	79,8	32,9	461,1	25,7	16,1

Diamètre (mm)	200	216	225	250	300	350	400	>400	Total
Longueurs (km)	210,9	7,3	0,2	192,5	73,3	4,1	58,1	116,3	2903,9

Tableau 4 : Répartition des canalisations par secteur (décembre 1990)

	Secteur 1	Secteur 2	Secteur 3	Secteur 4	Total
Longueur (km)	438,6	1030,4	371,7	1063,2	2959

- Date de pose

L'âge moyen des différentes conduites ne peut être déterminé de manière très précise pour tout le réseau. Les données concernant l'année de pose existent :

- Depuis 1970 pour les communes entrant dans la Communauté Urbaine de Bordeaux à cette date (Parempuyre, Saint Aubin du Médoc,...);
- Depuis 1951 pour celles faisant déjà partie du Service des Eaux appelé à l'origine Syndicat Intercommunal de l'Agglomération Bordelaise pour l'Adduction et la Distribution de l'Eau et déjà géré par la Lyonnaise des Eaux-Dumez;
- Depuis environ 1930 pour certaines communes de la proche banlieue de Bordeaux.

Cependant connaissant le développement de l'agglomération bordelaise et de sa banlieue on peut estimer selon les secteurs la date de pose des canalisations d'eau potable.

Ainsi, la plus grande partie du réseau situé au centre de Bordeaux est assez âgé et date de la fin du siècle dernier ou du début du vingtième. Les communes de proche banlieue telles Mérignac, Caudéran, le Bouscat, Bruges, Pessac ont commencé à se développer vers le début du siècle. Leur développement se prolonge d'ailleurs jusqu'à aujourd'hui en raison de la construction de nombreux lotissements, qui a entraîné dans ces communes l'extension du réseau.

Puis se sont développées des communes un peu plus éloignées de Bordeaux telles Saint Aubin du Médoc ou Parempuyre, qui étaient à l'origine plus rurales et dont l'extension a eu lieu vers la fin des années soixante-début des années soixante-dix.

Il est à noter qu'en général les canalisations de diamètre plus petit, comme le 60 ou le 80 mm, sont plus âgées car elles correspondent à une époque où la densité de population était plus faible. Malgré les importants travaux de renforcement, il en existe encore beaucoup. Le diamètre quatre-vingt n'est d'ailleurs plus utilisé et le soixante n'est actuellement posé que dans les nouveaux lotissements.

- Matériau

Le matériau des canalisations n'est accessible que dans les archives répertoriant les travaux de pose. Ces données sont présentées dans le chapitre suivant mais on peut d'ores et déjà remarquer que le réseau est constitué de canalisations en fonte dans une proportion proche de 100 % pour ce qui est des diamètres supérieurs à 80mm.

Seules les canalisations de diamètre inférieur ou égal à 60 sont depuis une quinzaine d'années en PVC.

On peut également distinguer les canalisations en fonte grise et en fonte ductile. D'après le fournisseur, la fabrication de la fonte ductile a commencé vers 1960. Cependant comme le Service des Eaux possédait à l'époque un stock important de canalisations, la date de passage de la fonte grise à la fonte ductile est beaucoup plus floue et se situerait aux environs des années 64-66.

3.2.2. Le réseau de la Compagnie des Eaux de la Banlieue de Paris (CEB)

La Compagnie des Eaux de la Banlieue de Paris gère le service public de la distribution d'eau du Syndicat des communes de la Presqu'île de Gennevilliers, soit au total neuf communes : Asnières, Bois-Colombes, Colombes, Courbevoie, Gennevilliers, La Garenne-Colombes, Nanterre, Suresnes et Villeneuve la Garenne. Elle est également concessionnaire de la ville de Rueil-Malmaison. La population desservie est de 500 000 habitants, ce qui correspond à environ 50 000 branchements pour 1000 km de canalisations.

3.2.2.1. Les ressources

L'eau distribuée dans la Presqu'île de Gennevilliers, soit environ 50 millions de m^3 par an, provient de deux types de ressources :

- le captage d'eaux souterraines provenant de Aubergenville, Croissy et Villeneuve la Garenne,
- la Seine, prise au niveau de Suresnes.

3.2.2.2. Structure de la distribution

Les eaux provenant de la Seine, d'Aubergenville et de Croissy, sont traitées à l'usine du Mont-Valérien, située à Rueil-Malmaison. Elles sont ensuite distribuées :

- soit gravitairement vers Suresnes, Nanterre, Courbevoie, la Garenne et une partie de Colombes, Bois-Colombes et Asnières,
- soit par refoulement vers Rueil-Malmaison.

Les eaux provenant du captage de Villeneuve sont traitées puis refoulées vers Villeneuve, Gennevilliers et l'autre partie de Colombes, Bois-Colombes et Asnières.

3.2.2.3. Linéaire du réseau

La répartition des conduites par diamètre se trouve dans le tableau 5.

Tableau 5 : Répartition par diamètre sur le réseau de la CEB

Diamètre	≤60	63	80	100	125	150	200	250	300	350
Longueur (km)	96,3	5,7	26,3	306,2	17,6	182,5	138,4	48,7	47,7	8,6

Diamètre	400	500	600	700	800	900	1000	1200	1600	Total
Longueur (km)	18,2	36,7	17,4	2,5	22,0	0,9	5,8	3,6	1,5	986,5

Concernant la répartition par commune, on peut la retrouver dans le tableau 6.

Tableau 6 : Répartition par matériau sur le réseau de la CEB

Commune	Asnières	Bois-Colombes	Colombes	Courbevoie	Gennevilliers	La Garenne-Colombes
Longueur (km)	108,1	40,9	145,8	90,4	104,4	50,1

Commune	Nanterre	Suresnes	Villeneuve la Garenne	Puteaux	Rueil-Malmaison	Total
Longueur (km)	174,1	78,7	32,3	7,2	154,6	986,5

Là encore la distinction entre fonte grise et fonte ductile ne peut se faire qu'en fonction de la date de pose de la conduite. On peut évaluer la date limite comme se situant entre 1960 et 1966.

3.2.3. Les réseaux des syndicats d'eau du Kronthal et du Kochersberg [Zedgui, 1993]

Le syndicat du Kronthal regroupe 13 communes rurales du département du Bas-Rhin. Il se situe au pied des Vosges. La population concernée est de 9619 habitants (recensement 1990).

Le syndicat du Kochersberg est lui nettement plus grand. Il regroupe 39 communes du département, pour une population totale de 22119 habitants en 1990.

Le Syndicat du Kronthal est alimenté à partir de quatre puits implantés de part et d'autre de la rivière Mossig, qui ont produit en 1991 un peu plus de un million de m^3. Le Syndicat du Kochersberg est alimenté par 9 points d'eau (4 sources et 5 puits) qui ont fourni une production de deux millions de m^3 en 1991.

Sur le Syndicat du Kronthal comme sur celui du Kochersberg, la distribution se fait gravitairement à partir de réservoirs principaux. Le linéaire de canalisations est de 90 km pour

le Syndicat du Kronthal et de 190 km pour le Kochersberg. La répartition par diamètre se trouve dans le tableau 7. La répartition par matériau (PVC-Fonte) se trouve dans le tableau 8.

Tableau 7 : Répartition des diamètres sur les syndicats du Bas-Rhin

Diamètre (mm)	Longueur (m)		
	Kronthal	Kochersberg	Total
≤60	930	4700	5630
70	0	1425	1425
80	18066	28255	46321
90	4117	9440	13557
100	7444	30228	37672
110	18707	38786	57493
125	10863	36301	47164
140	140	0	140
150	21563	28425	49988
160	342	0	342
175	850	0	850
200	4100	4980	9080
250	2964	4980	7944
300	0	5630	5630
Total	90086	193150	283236

Tableau 8 : Répartition des matériaux sur les syndicats du Bas-Rhin

Matériau	Longueur (m)		
	Kronthal	Kochersberg	Total
PVC	22854	50151	73005
Fonte	67232	138449	205681
Total	90086	188600	278686

3.3. LE RECUEIL DES DONNEES

3.3.1. Le choix des conduites

Vu le linéaire concernant ces trois réseaux (un peu plus de 4000 km), il n'était pas possible de recueillir les données sur les réseaux complets. Un choix a donc été fait concernant les différents services.

3.3.1.1. Réseau de la LED-Bordeaux

Pour le réseau de la LED-Bordeaux, l'étude a été réalisée sur l'ancien secteur 4, qui regroupe une dizaine de communes. Ces communes, toutes situées sur la rive gauche de la Garonne au

Nord-Ouest de Bordeaux, sont les suivantes : Blanquefort, le Bouscat, Bruges, Eysines, le Haillan, Mérignac, Parempuyre, le Taillan-Médoc, Saint-Médard sur Jalles et Saint-Aubin du Médoc. La population concernée est d'environ 100.000 habitants.

Ce secteur a été choisi en raison de sa superficie (environ un tiers de celle de la Communauté Urbaine de Bordeaux), qui permet d'obtenir une certaine variabilité vis-à-vis de différents facteurs. On aura ainsi des sols marécageux près de la Garonne (Blanquefort, Parempuyre) et des sols sableux plus à l'ouest (Saint-Aubin, Saint-Médard). De même les densités de population ne sont pas les mêmes selon les communes, celles-là diminuant en s'éloignant de Bordeaux.

Le réseau concerné a une longueur totale de 1.063 km, les longueurs par commune se trouvant dans le tableau 9.

Tableau 9 : Répartition des canalisations par commune du secteur 4 (décembre 1990)

Communes	Blanquefort	Le Bouscat	Bruges	Eysines	Le Haillan	Mérignac
Longueurs (km)	115,4	92,5	68,6	110,9	53,1	293,1

Communes	Parempuyre	Saint-Aubin	Saint-Médard	Le Taillan	Total
Longueurs (km)	43,6	50,8	171,4	63,7	1063,2

Sur ce secteur deux diamètres principaux ont été choisis. Ce sont les diamètres 60 et 150mm; tout d'abord parce que leur longueur n'est pas négligeable (125 km pour le premier, 215 pour le second), ensuite parce que selon le fournisseur, ces deux diamètres réagissent de manière différente aux contraintes, enfin parce que le diamètre 60 est composé de canalisations en fonte et en PVC, ce qui permettra éventuellement d'estimer l'influence du facteur matériau.

Deux autres diamètres ont également été pris en compte sur la commune de Mérignac, qui est la plus grande du secteur. Ces diamètres sont le 80 et le 100 mm.

3.3.1.2. Réseau de la CEB

L'étude sur la LED-Bordeaux avait été réalisée préalablement à celle de la CEB et des deux syndicats du Bas-Rhin. A la suite des résultats obtenus et présentés dans le chapitre 5, il a semblé opportun d'étudier sur un plus petit secteur tous les diamètres. C'est pourquoi le choix a été fait, concernant la CEB, d'étudier moins de communes, mais sur tous les diamètres.

L'étude a été réalisée sur quatre communes du réseau : Rueil-Malmaison, Colombes, Gennevilliers, Villeneuve la Garenne. Ces quatre communes ont été choisies en raison de leur variabilité vis-à-vis de l'origine de l'eau et du type de sol (remblais du côté de Gennevilliers et Villeneuve, sable pour Colombes et Rueil-Malmaison).

Le réseau concerné a une longueur totale de 440 km, les longueurs par commune étant représentées dans le tableau 10.

Tableau 10 : Répartition des conduites selon les communes étudiées sur la CEB

COMMUNES	LONGUEURS (km)
Colombes	145,9
Gennevilliers	116,9
Rueil-Malmaison	145,4
Villeneuve-la-Garenne	32,3
TOTAL	440,50

Pour cette étude tous les diamètres de canalisations, excepté le diamètre 40 qui peut être confondu avec un branchement, ont été pris en compte. La répartition par diamètre se trouve dans le tableau 11.

Tableau 11 : Répartition des conduites selon le diamètre sur la CEB

Diamètre (mm)	Longueur (km)
40	4,06
60	53,56
63	2,07
80	15,50
100	133,47
125	7,64
150	77,20
200	67,47

Diamètre (mm)	Longueur (km)
250	21,49
300	24,30
350	1,37
400	6,49
500	11,45
600	9,14
800	3,72
1000	1,19
TOTAL	440,50

Les données par matériau sont les suivantes :

Tableau 12 : Répartition des conduites selon le matériau sur la CEB

Matériau	Longueurs (km)	%
Acier	0,63	0,14
Fonte	432,83	98,72
Polyéthylène	4,98	1,14

3.3.1.3. Syndicats du Kronthal et du Kochersberg

Etant donné le linéaire plus faible, la totalité des réseaux a été étudié, soit environ 280 km de conduites.

3.3.2. Les données concernant les conduites

La nature et l'origine des données sont différentes selon les services. On les retrouve dans les tableaux suivants :

- Le diamètre et la longueur des tronçons

LED-Bordeaux	CEB	Bas-Rhin
Plans de réseaux au 10.000ème et au 1.000ème	Plans au 200ème et base de données informatisée	Plans au 2.000ème

- La date de pose

LED-Bordeaux	CEB	Bas-Rhin
Classeurs répertoriant les travaux de pose existant depuis 1970 pour les communes de Parempuyre et de Saint Aubin et depuis 1951 pour les autres	Fichier de base de données informatisée sur les canalisations	Attachements des travaux de pose (depuis 1976)

- Le matériau de canalisation

LED-Bordeaux	CEB	Bas-Rhin
Plans ou classeurs répertoriant les travaux de pose	Fichier de base de données informatisée sur les canalisations	Plans ou attachements

- Date de passage à des nouveaux matériaux

Cette date peut être importante car elle permettra de comparer les différents matériaux : fonte grise/fonte ductile, joint au plomb/joint express, revêtement intérieur, revêtement extérieur. Les dates d'apparition de ces nouveaux matériaux, données par Pont-à-Mousson, sont les suivantes :

* Joints

1 : Joints au plomb

Ce type de joint figurait encore au catalogue en 1963 et les derniers stocks ont été liquidés en 1969.

2 : Joints mécaniques (express)

Les premiers joints mécaniques sont apparus au catalogue en 1936 tandis que le joint EXPRESS n'apparaît qu'en 1948.

3 : Joints automatiques (STANDARD 2GS, RAPID)

La mise sur le marché date de 1958.

* Revêtements

1 : Revêtement intérieur en mortier de ciment

Ce type de revêtement a été introduit entre 1957 et 1961.

2 : Revêtement extérieur en zinc

Il a été utilisé pour la première fois en 1958 pour les tuyaux en fonte grise et en 1961 pour les tuyaux en fonte ductile.

* Fonte

Les tuyaux en fonte ductile (fonte GS) apparaissent au catalogue en 1963, tandis que la fabrication de la fonte grise (fonte GL) est totalement arrêtée à l'usine en 1971.

On peut donc voir que les changements de technologie ont surtout eu lieu au début des années 60. Cependant à l'époque les problèmes de stockage, déjà cités auparavant, ne permettent pas de donner une date bien précise. On peut cependant situer ce changement aux alentours des années 64-66.

- Le sol et sa corrosivité

LED-Bordeaux	CEB	Bas-Rhin
Carte de corrosivité des sols et avis du chef du secteur	Avis des inspecteurs de réseaux	Carte géologique au 50.000ème

La carte de corrosivité des sols de la LED-Bordeaux provient d'une étude concernant "les risques de corrosion des canalisations d'eau potable sur le territoire de la Communauté Urbaine de Bordeaux". Cette étude a été réalisée en août 1989 par la Lyonnaise des Eaux-Dumez et avait pour objet de caractériser les différents sols de la Communauté Urbaine de Bordeaux selon leur corrosivité, soit par des caractéristiques évidentes de corrosivité, telles une inondabilité importante ou leur nature propre, soit à partir de campagnes de mesures de la résistivité.

Différentes classes de corrosivité de sol ont ainsi pu être établies et tracées sur un plan au 25.000ème. Malheureusement les mesures de résistivité n'ont pu être prises sur tout le terrain de la Communauté Urbaine de Bordeaux, les relevés ayant eu lieu lors d'une période de forte sécheresse.

Les différentes classes de corrosivité sont les suivantes :

* Risque de corrosion important :

Sol tourbeux hydromorphe,

Tourbes,

Palus et bords de Garonne,

Zones de remblais et de décharge d'ordures ménagères,

Argiles bleues-grises de la nappe alluviale,

* Risque moyen de corrosion :

Argilo-calcaire tertiaire,

* Risque faible de corrosion :

Remblais sablo-graveleux.

Ces données ont donc pu être utilisées pour une partie du secteur étudié sur la LED-Bordeaux.

Sur les deux autres réseaux, ces données sont beaucoup moins précises car elles proviennent soit de cartes géologiques qui ne reflètent pas forcément la corrosivité du sol,

soit des inspecteurs de réseaux, dont l'avis est forcément subjectif. Dans la suite de l'étude les sols seront caractérisés de la manière suivante :

* Sol à tendance argileuse, supposé corrosif,

* Sol à tendance sablo-graveleuse, supposé non corrosif.

- *Trafic*

LED-Bordeaux	CEB	Bas-Rhin
Comptages du service voirie de la Communauté Urbaine de Bordeaux	Carte de comptage sur les routes départementales et nationales des Hauts de Seine provenant de la DDE	Carte de comptage sur les routes départementales et nationales du Bas-Rhin provenant de la DDE

Les données sur Bordeaux n'ont pu être obtenues que pour un certain nombre de rues, qui ont fait l'objet d'études et pour lesquelles il y avait des problèmes de circulation et de signalisation. On peut supposer que les rues pour lesquelles les données sont disponibles correspondent à celles qui connaissent le plus fort trafic.

- *Pressions de service*

LED-Bordeaux	CEB	Bas-Rhin
Mesures faites par les sapeurs-pompiers et calcul résultant de la modélisation hydraulique du réseau	Mesures de pression en quelques points du réseau	Calcul provenant de la modélisation hydraulique du réseau

- *Nature de l'eau*

LED-Bordeaux	CEB	Bas-Rhin
Analyses chimiques	Analyses chimiques	**

** signifie que cette donnée n'a pu être recueillie sur le service.

Sur Bordeaux, les analyses chimiques des eaux provenant des différentes sources concernées par le secteur ont pu être obtenues, analyses contenant notamment le taux d'oxygène dissous, le pH et la conductivité de l'eau. Cependant les données entre les différents forages ne diffèrent que très peu notamment pour le taux d'oxygène dissous, qui, d'après certaines études, apparaît comme le facteur le plus influent. Il n'a donc pas été possible d'en tenir compte. D'autre part les eaux provenant des différents réservoirs se mêlent dans le réseau sans que l'on connaisse les zones limites.

Sur le secteur étudié de la CEB, les zones alimentées par les deux ressources sont plus nettes. La zone Nord (Villeneuve, Gennevilliers, une partie de Colombes) est alimentée par le captage de Villeneuve, la zone Sud (Rueil-Malmaison et l'autre partie de Colombes) par les autres captages et pour une partie par l'eau de Seine. Les eaux des deux origines se différencient surtout par leur teneur en sulfate.

- Nombre d'abonnés par unité de longueur

Cette donnée a pu être obtenue sur Bordeaux en comptant le nombre d'abonnés par rue à partir des fichiers d'abonnés du Service des Eaux et en mesurant la longueur de chaque rue, donnée qui semble bien caractériser l'environnement de la conduite, notamment le taux d'activité autour de celle-ci, ainsi que la densité de population.

- Emplacement de la conduite par rapport à la chaussée

LED-Bordeaux	CEB	Bas-Rhin
Carnets d'emplacement des vannes et autres organes hydrauliques	Plans au 200ème et Système d'Information Géographique	Visites sur le terrain

Ce facteur caractérise le fait que la conduite se trouve sous la chaussée ou le trottoir.

- Conditions de pose et protection

Des données concernant la profondeur de la canalisation et son lit de pose n'ont pu être obtenues.

Sur Bordeaux, on peut simplement dire que cette profondeur est de l'ordre de 1m et qu'un plus grand soin a été apporté au lit de pose et à sa profondeur à partir des années 70. Sur la CEB, on sait que depuis le début des années 80 toutes les conduites sont posées avec une protection extérieure en polyéthylène.

Comme on le voit les données obtenues sont différentes selon les services. De plus elles ne sont pas toujours obtenues de la même manière, notamment pour les deux services ruraux où les données sont moins précises.

3.3.3. Les données concernant les défaillances

Rappelons que les défaillances considérées ici sont les fuites ou casses ayant engendré une réparation au niveau du joint ou au niveau du tuyau, ce qui exclut les fuites diffuses.

Sur Bordeaux, la conservation des données concernant les défaillances s'est faite de deux manières différentes :

- dans des livres annuels de 1955 à 1969 dans lesquels les défaillances sont classées par commune ceci pour toutes les communes excepté Parempuyre et Saint Aubin,

- dans des fiches classées par rue et par commune depuis 1970 pour toutes les communes du secteur.

Sur la CEB, les données concernant les défaillances ont été obtenues à partir de différentes sources :

- un fichier informatisé répertoriant tous les travaux de maintenance existant depuis 1985, parmi lesquels sont différenciées les fuites sur branchement, fuites sur conduites, etc...
- des cahiers annuels de maintenance, classés par commune, depuis 1973,
- des dossiers de rue, dans lesquels sont gardés tous les événements de la vie de la conduite dans la rue (fuites, plaintes, travaux neufs,...) depuis 1938.

Sur les deux syndicats du Bas-Rhin, les défaillances sont conservées dans des récapitulatifs annuels classés par commune.

La figure 12 présente une des fiches conservées sur Bordeaux.

Dans ces livres et fiches sont conservées les défaillances à la fois sur les canalisations et sur les branchements, ainsi que les poses de compteur et tout autre travail d'entretien. Un premier tri a donc été nécessaire pour retrouver les données recherchées.

Comme on peut le voir sur la fiche, un vocabulaire spécifique et des abréviations ont été employés et les différentes défaillances ont été distinguées par les termes suivants : "Casse tuyau ∅ 60", "fuite sur canalisation", "1 manchon 2 (ou 3) pièces", "1 manchon de réparation", "1 MR", "joint rematé sur diamètre 150", "remplacement joint sur canalisation", "1 MD", "1 coupe cylindrique", "1 CC", "1m de tuyau", etc.

Il n'existe apparemment aucune norme concernant la description des différents travaux d'entretien ou de réparation. C'est pourquoi il est parfois impossible de distinguer une réparation effectuée sur un tuyau ou sur un branchement ou même, lorsque deux conduites existent sur une même rue, de distinguer la canalisation touchée, le diamètre n'étant pas toujours indiqué.

En résumé, les fiches et cahiers donnent sur les défaillances les renseignements suivants :

- la rue et la commune où a eu lieu la défaillance (et éventuellement le numéro de rue ou l'emplacement de la défaillance),
- la date de la défaillance,
- le diamètre (éventuellement),
- le type de réparation (éventuellement),
- la cause de la défaillance (lorsque celle-ci est due à une entreprise tiers).

- Présentation des types de réparation

Afin de mieux définir la notion de défaillance, il est utile de présenter les différents types de réparation, décrits dans les fiches.

Aristide Briand

Rue Baudin – Ave J. Guesde –

Nos	Objet de la Réparation	Date d'exécution	Équipes	OBSERVATIONS
45	C 28512 V 14940	9.4.70	Barrières	Chanf. Cl. 6427 G. 1 RAC 15. 1 Tête 20 (LAVILLE)
19	V 4964	14.4.70	"	Tuyau cassé 60 Eternit
19	B 4838	26.5.70	Chiquet	Fermeture – Payi –
97	V. 8612	29.5.70	Barrières	Remise Gibaudet.
104	V. 8773	11.6.70	Ménier	1 Tête 20 : DECOGNE.
21	V. 3475	15.7.70	Barrières	1 Tête 15. Bandelay
	V. 8846	21.7.70	"	1 Tête 30. Groupe Scolaire J. Jaurès.
9	V. 6001	28.7.70	"	1 RAC 15.
125	V. 6078	7.8.70	Fauceu	2 Soudures. Bt ø 40. Cité Ballarin.
	T. V 6309	18.9.70	"	1 Soudure – 01 Bt école Jean Jaurès – Filles
	SA. C 14548	19.9.70	"	Dép. 50 Rad. 17.30 – 8140 7 191 m² 2649
	B 33032	—	"	Réal p/ Mairie S. Bouzat – Deuxième Hôtel Municipal
149	C 31266 V 6412	7.10.70	Barrières	Chanf. Cl. STJ 15. H 63. 1 RAC 15. PICHARDI.
41	C 2072	13.10.70	Alfo	Bt Neuf STJ 15. G. 9994. GALAN.
"	B 33095	— " —	"	" " 3m Pk. 18 x 5 – 12 cl. 60 x 20 Eternit N.
123	V. 3283	25.11.70	Barrières	Fuite Cl. 1 cuis. Brauè.
41	B 33092 – C 6578	28.12.70	Ménier	Réal STJ. 15 – G 9994 m² – Galan –
101	V 1986	31.12.70	Barrières	Tuyau étenitcassé – refait neuf –
35	C 28438	10.2.71	Chiquet	Ch Cl STJ 15 W 2655
124	C 28439	12.2.71	—	— Arbre 12 – 5442 Tuyau
Fau 54	B 33717	23.2.71	Barrières	Tuyau cassé 3m Eternit 60 ETPR
25	C 6950	2.3.71	Lamaison	Bt neuf STJ 15 H 1832 refait
	B 33603	—	—	20/20 eternit 18/15 3m r.n
84	B 33763	3.3.71	Barrières	Bt détérioré p/ ETPR 1 Tête v. ky 20
153	B 33858	16.3.71	—	— — 1 tête 20 1 tête v. ky 20
70	V 4299	30.3.71	Barrières	Ch 1 T.A.P de 20
121	V 4323	1.4.71	—	Fuites 1/2 branch. 1 V. ky 20 + 1 tête v. ky 20
173	B 33996	2.4.71	Cagot	Tuyau cassé E.T.P.R 60 Eternit 6m
167	B 33986	—	Barrières	— — — 4m
124	V 4363	8.4.71	—	1 T.A.P 15
77	B 34087	23.4.71	Cadoi G	remise profil BE 1 T/5 1.50 C.U.B
Fau 77	B 34088	—	—	— — 0.50 C.U.B
104	B 34089	—	—	— — 0.50 C.U.B
107	B 34090	—	—	— — C.U.B
126	B 34091	—	—	remise profil BE
75	B 34092	—	—	remise profil et chaussement BE
52	V 6541	19.5.71	Ménier	Tuyau cassé Eternit 60 3m
167	V 6836	2.6.71	Barrières	" " " " 1,m

Figure 12 : Exemple de fiches de maintenance (LED-Bordeaux)

- Réparations concernant le tuyau

  * Coupe cylindrique, remplacement par une certaine longueur de tuyau, utilisation de deux joints gibault ou deux manchons deux pièces (ou droits) ; ceci correspond à une casse longitudinale du tuyau.

  * Manchon de réparation (ou manchon 3 pièces) ; ce type de manchon peut également être utilisé lorsqu' il y a eu cassure longitudinale et remplacer la réparation précédente.

  * Manchon 2 pièces ou manchon droit ou joint gibault ; ceci correspond à une cassure transversale du tuyau.

- Réparations concernant les joints

Ces réparations sont de deux types, selon le type de joint employé :

  * Matage de joints; ceci ne concerne que les joints au plomb.

  * Remplacement de joints, pour les autres types de joint.

3.3.4. Création de fichiers de base de données : fichier "Tronçons" et fichier "Défaillances"

Le recueil des différentes données s'est effectué dans deux fichiers, un fichier "tronçons" et un fichier "défaillances" établis à partir du logiciel de base de données Paradox 4.0. Les deux fichiers sont présentés dans les paragraphes suivants.

3.3.4.1. Le fichier "tronçons"

Ce fichier a permis de répertorier les différents tronçons de canalisations ainsi que leurs caractéristiques.

- *Définition du tronçon de base*

Chaque tronçon a été défini de manière homogène vis-à-vis de certains facteurs caractérisant les canalisations, à savoir le diamètre, la date de pose, le matériau, le type de sol et son environnement immédiat. Chaque tronçon est rattaché à une rue et à une commune, ceci afin de pouvoir l'identifier dans la recherche des défaillances.

Ainsi, tout tronçon est donc individualisé par le nom de la rue et de la commune, puis par un numéro d'identifiant.

- *Structure du fichier*

Une fois définis les tronçons composant les fichiers, le recueil des caractéristiques peut avoir lieu. Les caractéristiques de chaque tronçon de canalisations sont les suivantes :

- Longueur,
- Diamètre,
- Année de pose,

- Matériau,
- Type de sol,
- Trafic,
- Pression de service,
- Origine et qualité de l'eau,
- Nombre d'abonnés par unité de longueur.

Le tableau 13 représente les données disponibles dans chacun des services.

Tableau 13 : Caractéristiques des conduites prises en compte sur les différents services.

FICHIER TRONCONS			
	LED-Bordeaux	CEB	Bas-Rhin
Numéro de tronçon	X	X	X
Rue	X	X	X
Commune	X	X	X
Année de pose	X	X	X
Diamètre	X	X	X
Matériau	X	X	X
Longueur	X	X	X
Sol et corrosivité	X	X	
Trafic automobile	X	X	X
Pression de service	X	X	X
Nombre d'abonnés/km	X		
Sous trottoir ou chaussée	X	X	
Qualité de l'eau		X	

3.3.4.2. Fichier "défaillances"

Avant de commencer l'analyse il est utile de rappeler la définition du mot "*défaillance*" utilisé dans l'étude. Une défaillance correspond à une *fuite au niveau de la canalisation entraînant une réparation soit du joint de la conduite, soit du tuyau même*.

- *Structure du fichier*

Ce fichier répertorie toutes les défaillances recueillies avec chacune de leurs caractéristiques :

- Commune,
- Tronçon concerné,
- Emplacement dans la rue,
- Année de la défaillance,
- Type de réparation,

- Cause de la défaillance.

3.3.4.3. Contraintes conduisant à l'élimination de certains tronçons ou défaillances

Certains tronçons ont dû être éliminés en raison de l'absence de données. Ces contraintes sont les suivantes :

- Absence de données concernant l'année de pose de la conduite,
- Date de pose de la conduite connue, mais inférieure à la date du début d'observation des défaillances. Par exemple sur Bordeaux on sait pour certaines conduites qu'elles ont été posées avant 1930, ce qui entraîne une absence de données d'au moins 25 ans concernant les défaillances, celles-ci n'étant répertoriées que depuis 1956. Cette absence de données aurait entraîné un biais beaucoup trop important,
- Absence de données concernant les différents facteurs.

La collecte des défaillances a également entraîné l'élimination de certaines données recueillies. Les raisons sont les suivantes :

- Absence de données concernant le diamètre de la canalisation défaillante, d'où impossibilité de définir cette canalisation lorsqu'il en existe plusieurs dans la même rue,
- Impossibilité de définir si la fuite a eu lieu au niveau d'un branchement ou de la canalisation,
- Impossibilité de définir l'emplacement exact de la défaillance dans la rue, ceci posant un problème pour les rues divisées en plusieurs tronçons.

3.4. DESCRIPTION DES TRONÇONS ET DES DEFAILLANCES

Les caractéristiques des conduites étudiées après élimination des données incomplètes sont présentées dans le tableau 14.

Les conduites de la CEB sont en moyenne plus âgées, alors que celles du Bas-Rhin sont les plus jeunes. Celles de Bordeaux sont en moyenne plus longues. La figure 13 décrit la répartition des années de pose dans les échantillons.

On peut également remarquer que pour les deux réseaux urbains la fonte est le principal matériau utilisé. Les matériaux plastiques (PVC, Polyéthylène) commencent à se développer légèrement. Sur le Bas-Rhin, le PVC et la fonte sont réparties de manière à peu près égale. Notons que pour ces services, on a éliminé un grand nombre de conduites âgées en fonte, car on ne connaissait pas leur année de pose.

Tableau 14 : Description des conduites recueillies par service (données non recueillies sur le service).**

		LED-Bordeaux	CEB	Bas-Rhin
Nombre de tronçons		1334	1379	653
Longueur totale (km)		347,6	261,8	94,3
Longueur moyenne (m)		260,6	189,8	144,35
Age moyen (années)		18	24	16
Diamètre	Minimum	60	60	50
	Maximum	150	800	250
% Année de pose 66	≤66	35,2	47,1	21,3
	>66	64,8	52,9	78,7
% sol	Non corrosif	64,8	73,0	**
	Corrosif	35,2	27,0	
% Trafic	Faible	82,9	81,5	80,9
	Fort	17,1	18,5	19,1
% Emplacement	Sous trottoir	61,6	56,2	**
	Sous chaussée	38,4	43,8	
% Origine de l'eau		**	35,4 (Villeneuve) 38,7 (Rueil) 25,9(Mélange)	**
% Matériau	Acier	0	0,3	0
	Fonte	92,7	98,6	54,7
	PVC	6,5	0	45,3
	PE	0	1,1	0
	Amiante-Ciment	0,8	0	0
Pression moyenne(mce)		35,7	**	43
Nombre d'abonnés/km		48,1	**	**

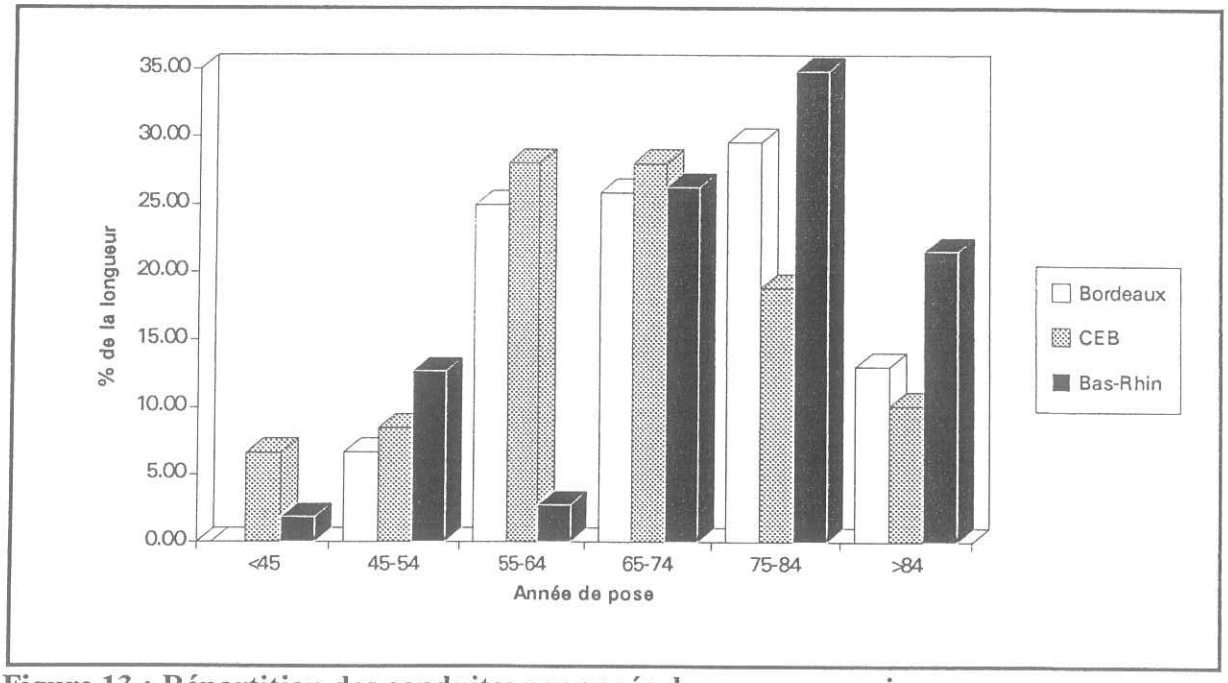


Figure 13 : Répartition des conduites par année de pose par service

La figure 14 précise la distribution des diamètres par service. Sur Bordeaux seuls quatre diamètres ont été observés. Sur la CEB les diamètres les plus fréquents sont les diamètres 100, 150 et 200 mm. Sur le Bas-Rhin, c'est le diamètre 110 qui est le plus fréquent. Sur ces deux syndicats du Bas-Rhin, certains diamètres non disponibles actuellement comme le 70, le 90 ou le 140 existent encore.

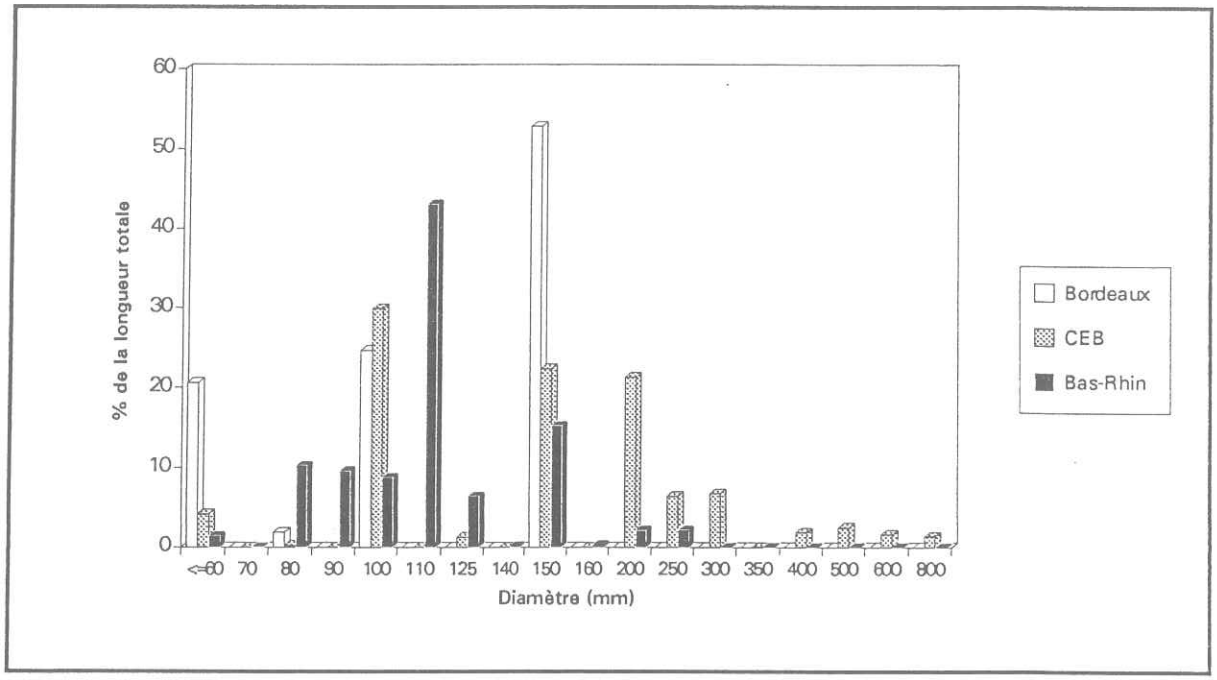


Figure 14 : Répartition des diamètres étudiés sur les trois services

Le nombre de défaillances par service se trouve dans le tableau 15. Le nombre de conduites ayant eu des défaillances est en fait très faible, entre 15 et 20 % du total (Cf. Figure 15). La proportion de celles en ayant eu au moins 6 est encore plus faible (moins de 1%).

Tableau 15 : Nombre de défaillances recueillies par service.

	Bordeaux	CEB	Bas-Rhin
Nombre de défaillances relevées	543	542	171

On peut noter également que la répartition des défaillances est différente selon la saison (Cf. Figure 16). C'est lors des mois d'hiver (décembre, janvier, février) que les défaillances sont les plus nombreuses (près de 40 % du total). L'action conjuguée de la baisse de la température et d'une humidité plus importante peut l'expliquer. Sur les deux syndicats du Bas-Rhin, on peut également se rendre compte que le nombre de défaillances augmente en juin, juillet.

3.5. LA SPECIFICITE DES DONNEES

Sur les 3366 tronçons de canalisations répertoriés sur les trois services, soit environ 700 km, 2835 n'ont pas eu de défaillance, soit environ 540 km. Près de 8 tronçons sur dix n'ont donc pas eu de casses.

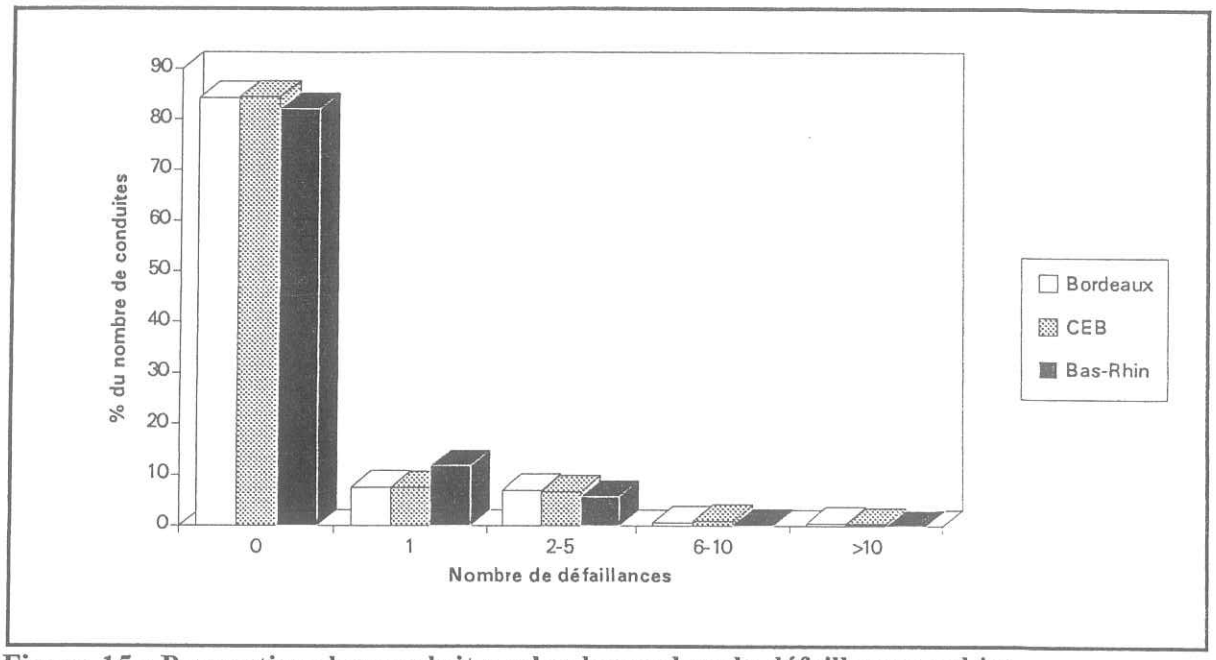


Figure 15 : Proportion des conduites selon le nombre de défaillances subies

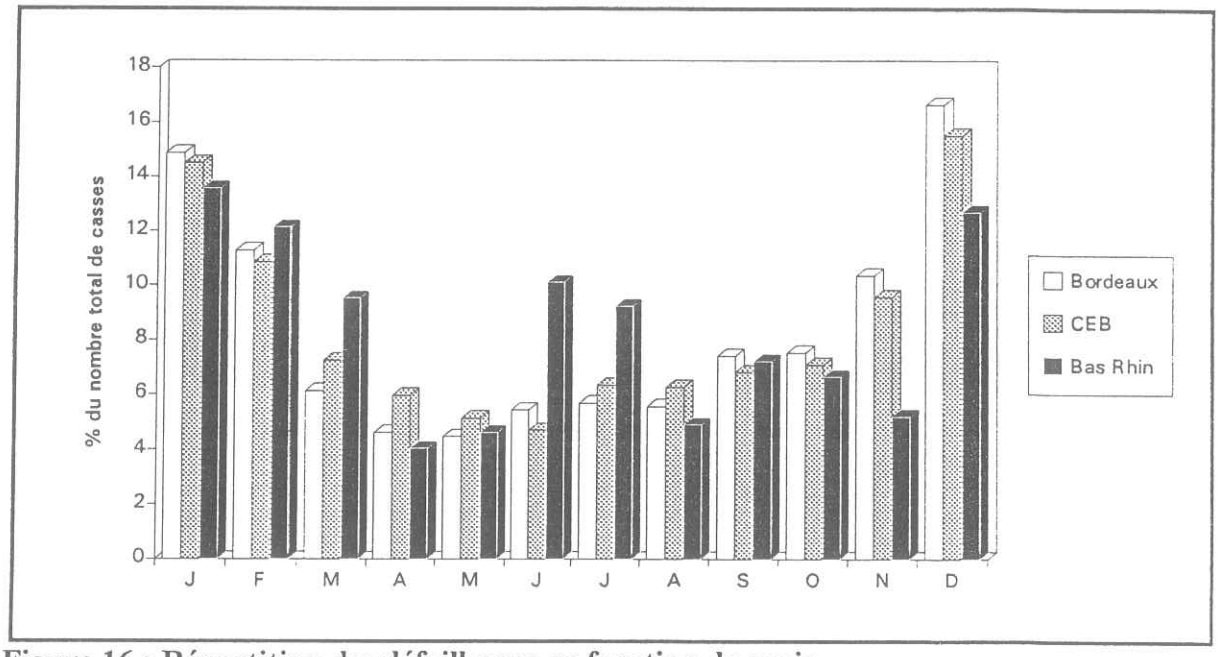


Figure 16 : Répartition des défaillances en fonction du mois

Faut-il pour autant les éliminer de l'échantillon d'observation puisque l'événement auquel on s'intéresse est l'événement "défaillance"? Ceci ferait perdre une quantité d'information trop importante puisque ces tronçons sans défaillance apportent l'information suivante : **ils n'ont pas eu de défaillance entre l'année de pose et l'année de fin d'observation.** Ces observations sont dites censurées à droite. La figure 17 représente les différents types de tronçons pouvant exister (en faisant l'hypothèse qu'un tronçon ne peut avoir qu'une seule défaillance).

TYPE DE TRONCONS	INFORMATION APPORTEE
Tronçon non censuré à droite : t0 Date de pose — td < tf Défaillance ayant lieu avant la date de fin d'observation	* Aucune défaillance entre t0 et td * Une défaillance à td
Tronçon censuré à droite : t0 Date de pose — tf Aucune défaillance	* Aucune défaillance entre t0 et tf

Figure 17 : Description de la spécificité des données

Il est donc important d'utiliser des méthodes statistiques spécifiques adaptées à ce type de données. Ce sont les méthodes d'analyse de données de survie qui sont utilisées et qui vont être décrites dans le chapitre suivant.

CHAPITRE 4 :
APPLICATION DE L'ANALYSE STATISTIQUE DES DONNEES DE SURVIE A L'ETUDE DES DEFAILLANCES

4.1. INTRODUCTION A L'ANALYSE DES DONNEES DE SURVIE

Ces méthodes d'analyse statistique s'appliquent à des types de données bien particulières. Elles sont utilisées dans des secteurs tels que l'épidémiologie (mesure de la survie d'individus malades dans des essais cliniques), l'industrie (durée de vie de composants) ou les sciences socio-économiques (mesure de la durée de grève ou de la période de chômage selon les individus par exemple) les utilisent.

4.1.1. Définitions et caractéristiques des données de survie

4.1.1.1. Les données de survie [Hill et al., 1990]

Les données de survie mesurent la *durée de vie* ou le *date d'apparition* d'un événement critique pour chaque individu étudié. Cet événement critique peut être soit la mort de l'individu, soit une défaillance, soit une guérison, soit une rechute...

Cependant la principale caractéristique de ces données est que cet événement peut ne pas apparaître chez certains individus parce que sa durée de vie (ou durée de "non-défaillance") est trop longue pour l'étude : ce sont les individus censurés à droite.

La notion d'individu "perdu de vue", courante en épidémiologie,n'intervient pas dans la présente étude.

4.1.1.2. Définitions

Dans une étude de survie il faut pour chaque sujet connaître les renseignements suivants :

* La date d'origine

La durée de survie que l'on désire étudier ne correspond pas toujours au temps écoulé entre la naissance et la mort. En général on prend plutôt le temps écoulé depuis un diagnostic ou la mise en oeuvre d'un traitement.

Chaque sujet a donc une date d'origine différente sur le calendrier, mais qui définit le temps 0.

* La date de dernières nouvelles

Cette date correspond à la date la plus récente où l'on a recueilli les renseignements concernant l'état d'un sujet. Si pour un sujet l'événement attendu (mort ou défaillance) a été observé, la date de dernières nouvelles correspond à la date d'apparition de cet événement.

Connaissant ces dates il est alors possible de calculer le *temps de participation* t_i défini pour chaque sujet i. Ce temps correspond au délai entre la date d'origine et la date de dernières nouvelles.

D'autre part un sujet est défini comme "exclu-vivant" s'il vit encore (ou s'il n'a pas eu de défaillance) au moment de la date de dernières nouvelles. Ce type d'individu correspond aux données censurées à droite (Cf. fig. 17).

Les censures présentées dans cette figure sont des *censures "à droite"* car elles ne donnent plus de nouvelles de l'individu après une certaine date, sans que cet individu soit mort ou ait eu une défaillance. Un autre type de censure existe :

* *Les censures à gauche* : ceci correspond aux données dont la date du début d'observation est postérieure à la date d'origine.

Pour ces données censurées on ne peut donc pas dire qu'il n'y a pas d'information sur la durée de vie et l'événement étudié et il est difficile de les ignorer même si l'information apportée n'est que partielle. Les bonnes nouvelles comme les mauvaises doivent être prises en compte!

En général on affecte à chaque individu son temps de participation ainsi qu'une variable de censure égale à 0 pour les individus censurés et à 1 pour les individus défaillants.

En résumé l'analyse des données de survie présente les spécificités suivantes :

* elle porte sur des variables aléatoires positives,
* elle se fait en présence de données censurées,
* elle peut prendre en compte des covariables (lorsque la population étudiée est hétérogène).

Ses objectifs sont de modéliser des lois de durée de vie (fiabilité), de comparer les durées de vie dans différents groupes (essais thérapeutiques) ou d'évaluer la valeur pronostique de facteurs à risques (études épidémiologiques).

4.1.1.3. Application au cas des conduites d'eau potable

Plusieurs défaillances peuvent survenir sur une même canalisation. Il faut alors distinguer l'événement "$k^{ième}$ défaillance".

Pour l'événement première défaillance, la date d'origine est la date de pose.

Pour l'événement $k^{ième}$ défaillance (k>1), la date d'origine est la date d'apparition de la $k-1^{ème}$ défaillance.

Chacun de ces événements peut être étudié individuellement. Il a donc été choisi de différencier une même canalisation suivant le nombre de défaillances qu'elle a subies. Dans ce cas un individu correspond au couple "canalisation, nombre de défaillances subies" (Cf figure 18).

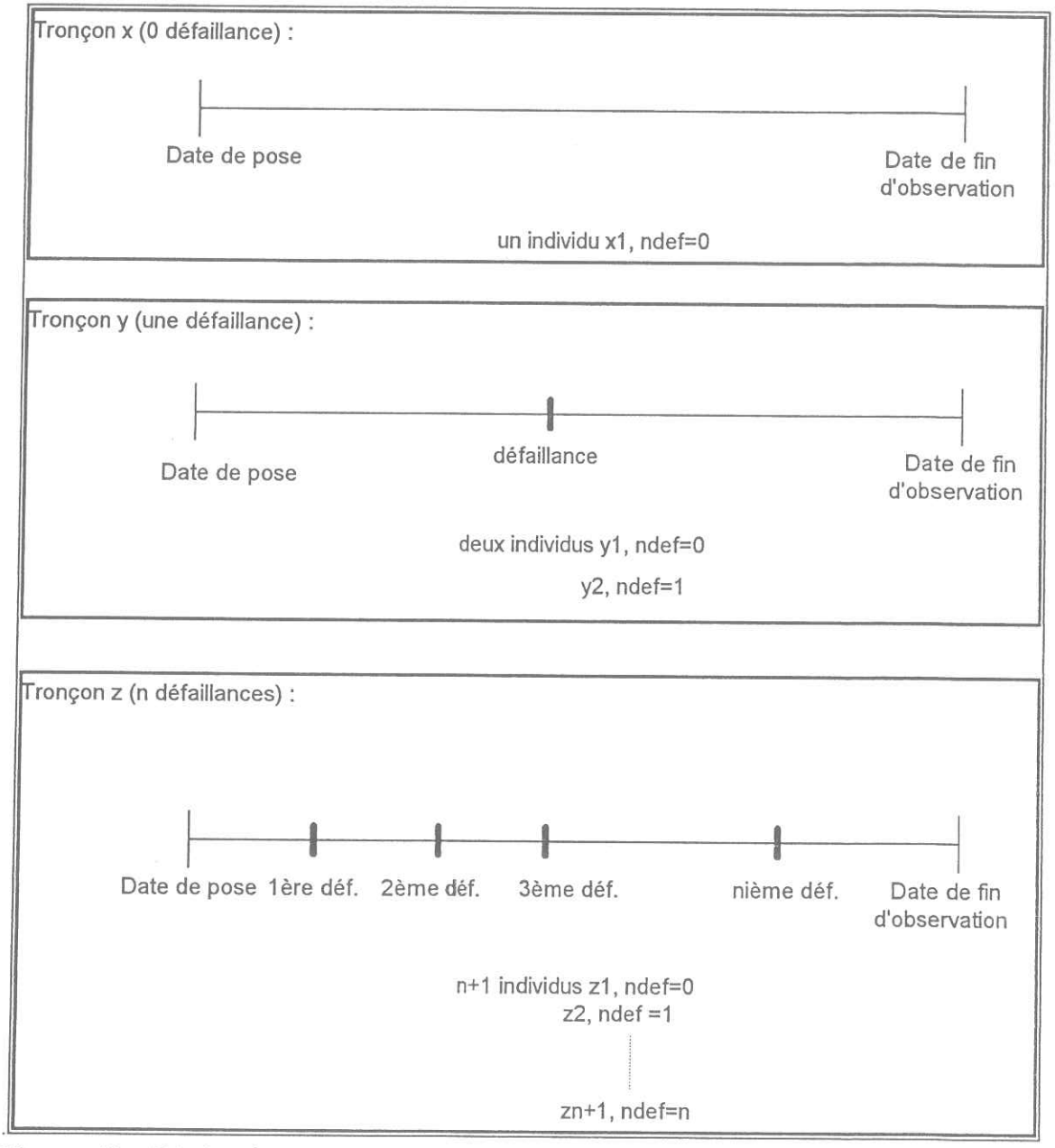


Figure 18 : Distinction des canalisations en individus (ndef correspond au nombre de défaillance).

4.1.2. Les fonctions étudiées en analyse de survie [Kalbfleisch et Prentice, 1980]

Soit T une variable aléatoire positive représentant le temps de défaillance d'un individu d'une population homogène. La distribution de probabilité de T peut être représentée de différentes manières. Trois types de fonction sont le plus utilisés : la fonction de survie, la fonction de risque instantanée, la densité de probabilité.

* La fonction de survie

Cette fonction *S(t)* correspond à la probabilité pour un individu d'avoir une durée de vie supérieure à un certain temps *t*.

On a donc :

$$S(t) = P(t \leq T), \text{ avec } t \in [0, +\infty]$$

S est une fonction décroissante et varie entre *S(0)*=1 et $\lim_{t \to +\infty} S(t) = 0$.

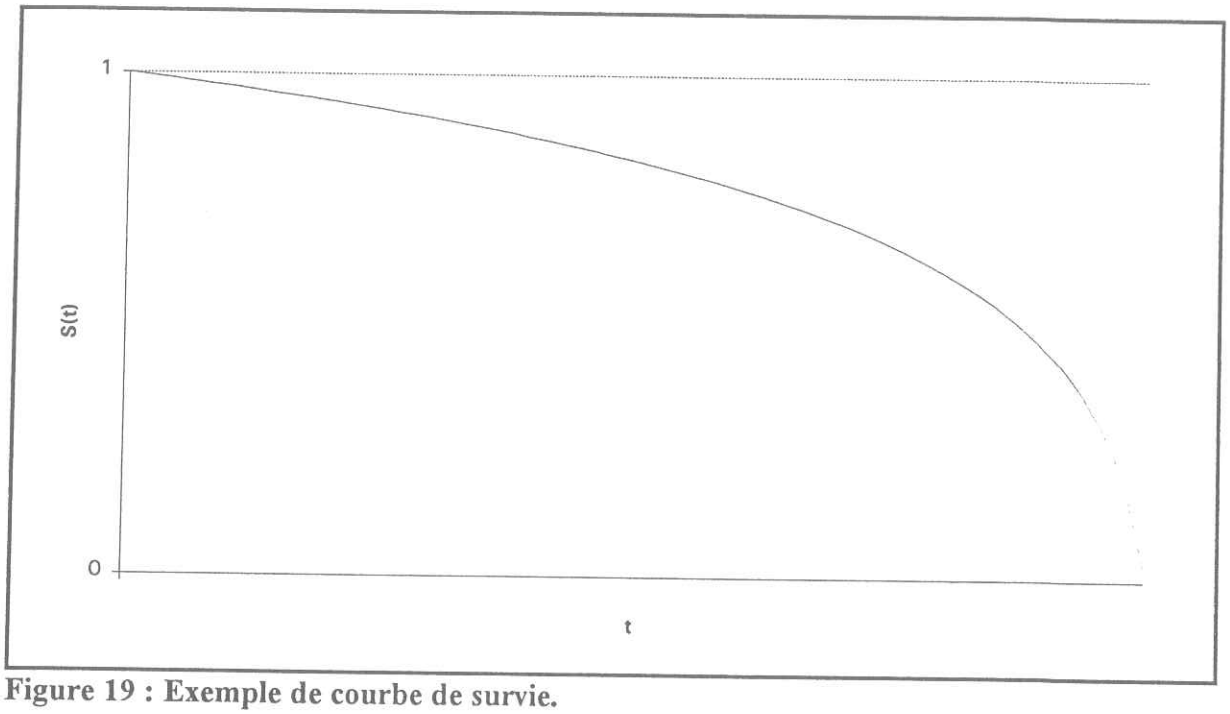


Figure 19 : Exemple de courbe de survie.

* La densité de probabilité de *T*

Cette fonction *f* est telle que :

$$f(t) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = -\frac{dS(t)}{dt}$$

* La fonction de risque (ou risque instantané)

Cette fonction *h* correspond au taux instantané de défaillance des individus qui, n'ayant pas eu de défaillance à un instant *t*, en ont une entre *t* et *t*+*Δt*. Elle est égale à :

$$h(t) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T < t + \Delta t / T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

On peut remarquer que la connaissance de h est équivalente à la connaissance de *S* ou de *f*. En effet d'après les définitions, on a :

$$h(t) = -\frac{d \log S(t)}{dt}$$

d'où, sachant que *S(0)*=1 :

$$S(t) = \exp(-\int_0^t h(u)du)$$

et

$$f(t) = h(t)\exp(-\int_0^t h(u)du)$$

Les figures 20 et 21 représentent deux exemples de courbes de risque.

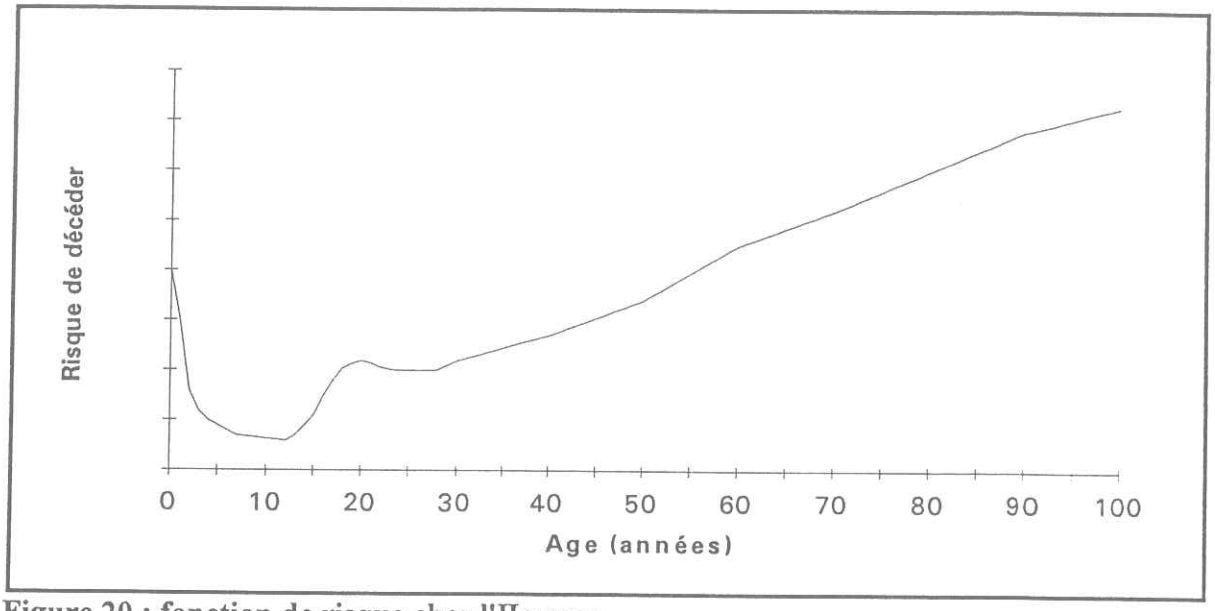


Figure 20 : fonction de risque chez l'Homme

La figure 20 correspond au taux de mortalité chez l'Homme. Ce taux est d'abord élevé juste après la naissance (mortalité infantile), puis décroît jusqu'à un certain âge, avant de croître plus lentement (vieillesse).

La figure 21 représente le "taux de sortie de l'état de chômage". Celui-ci est d'abord croissant (effet de la recherche d'emploi) puis décroît (chômeurs de longue durée qui ont de plus en plus de mal à trouver un emploi).

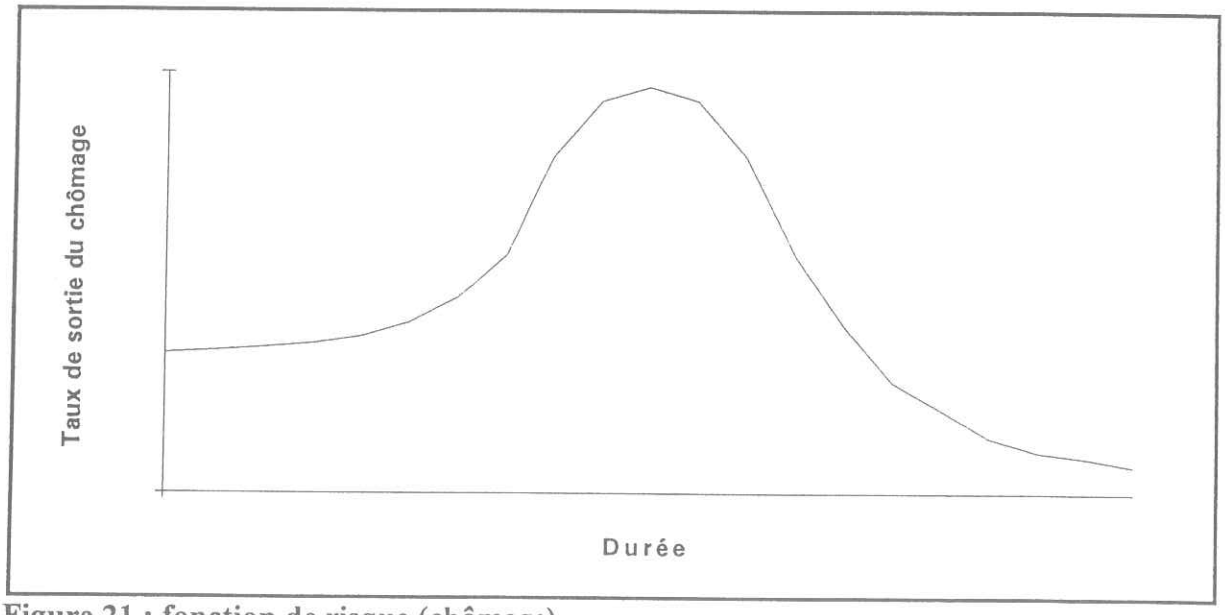


Figure 21 : fonction de risque (chômage)

Une dernière fonction est également utilisée, c'est la fonction de risque cumulée :

$$H(t) = \int_0^t h(u)du$$

On a alors :

$$S(t) = \exp(-H(t))$$

4.2. LES MODELES UTILISES EN ANALYSE DE SURVIE

4.2.1. Quelques modèles paramétriques

Il peut être utile d'examiner la fonction de risque dans l'étude des modèles probabilistes. Il est donc important de savoir comment elle évolue en fonction du temps au niveau de l'interprétation.

Voici quelques lois de probabilité souvent utilisées pour des populations homogènes.

*** Le modèle exponentiel**

Ce modèle est caractérisé par une fonction de risque constante (indépendance temporelle). On a donc :

$$\forall t > 0, h(t) = \lambda$$

$$\text{d'où } S(t) = \exp(-\lambda t)$$

Les défaillances surviennent donc de manière aléatoire au cours du temps. D'autre part on peut remarquer que le tracé de la fonction *log[S(t)]* est une droite passant par l'origine.

Ce modèle est un cas particulier du modèle de Weibull, décrit ci-après.

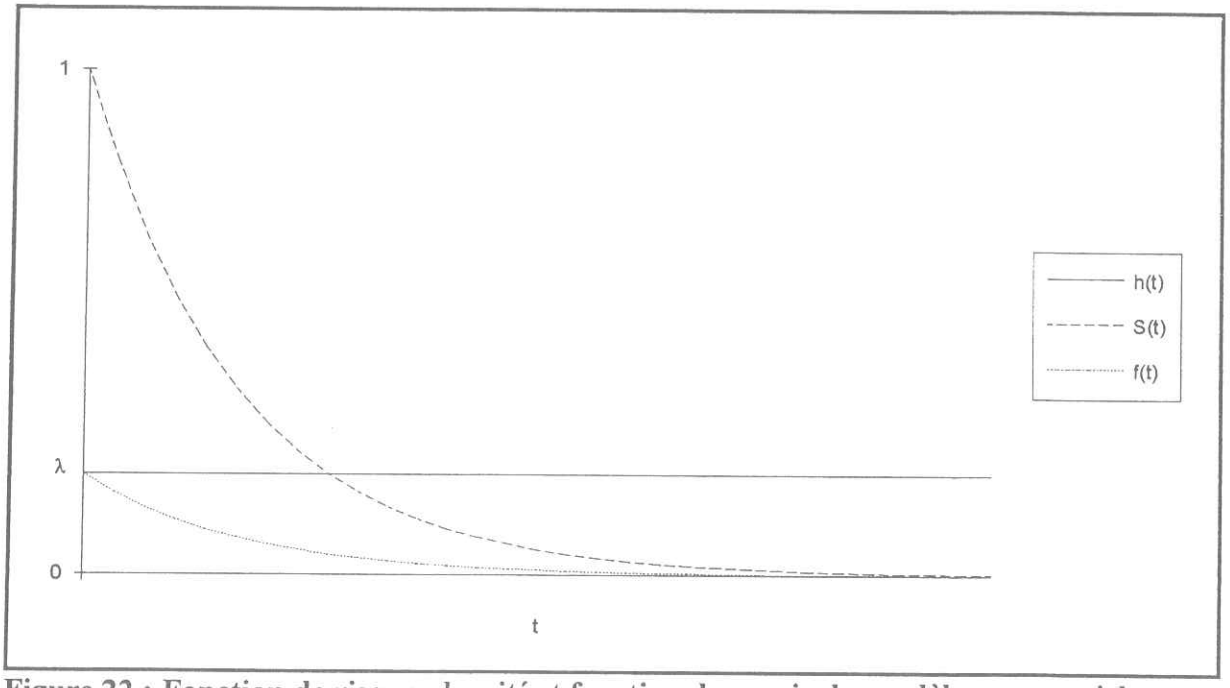


Figure 22 : Fonction de risque, densité et fonction de survie du modèle exponentiel

*** Le modèle de Weibull**

C'est une généralisation du modèle exponentiel, la fonction de risque étant une fonction puissance du temps.

$$h(t) = \lambda p(\lambda t)^{p-1}$$

et

$$S(t) = \exp\left[-(\lambda t)^{p}\right]$$

λ correspond au paramètre d'échelle (durée de vie caractéristique). p est le paramètre de forme.

Si $p<1$, h est monotone décroissante.

Si $p=1$, h est constante (modèle exponentiel).

Si $p>1$, h est monotone croissante.

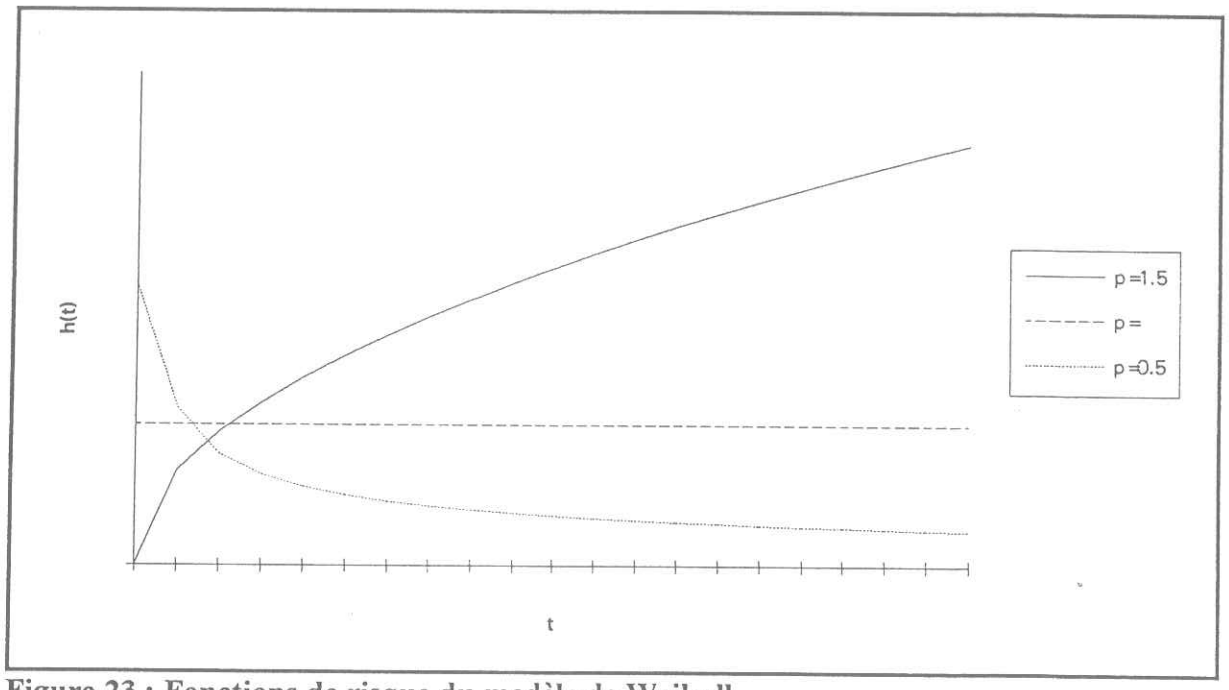


Figure 23 : Fonctions de risque du modèle de Weibull.

Le tracé de la fonction log[-log(S(t))] en fonction de log(t) est une droite.

4.2.2. Modèles de régression sur données de survie [Kalbfleisch et Prentice, 1980]

*** Introduction**

Ces types de modèle permettent de généraliser les modèles univariés au cas de populations hétérogènes pour tenir compte de l'influence de variables explicatives.

La fonction de risque h(t,z) dépend du temps et de covariables : on définit donc le vecteur z des variables explicatives, qui peuvent être soit qualitatives, soit quantitatives (constantes ou dépendant du temps).

L'objectif est de déterminer les *meilleures variables pronostiques*, de *tester et quantifier* l'effet d'une variable en tenant compte de certaines covariables ou de créer un *modèle prédictif*.

Les modèles sont basés sur la décomposition de la fonction de risque λ en :

* une fonction de risque de base $h_0(t)$ exprimant la façon dont la fonction de risque dépend du temps et commune à tous les individus.

* un *terme de régression* $\beta z = \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k$ qui inclut l'effet des variables explicatives sur la survie par l'intermédiaire des paramètres de régression β_i du modèle.

L'effet des covariables peut agir multiplicativement soit sur la fonction de risque (Modèle des Risques Proportionnels ou, en anglais, Proportional Hazards Model) soit sur le temps de défaillance (modèles des temps accélérés ou, en anglais, Accelerated Failure Time Model).

*** Le Modèle des risques proportionnels**

Dans ce modèle, le risque instantané $h(t,z)$ est le produit d'une fonction de risque de base $h_0(t)$ et d'un terme exponentiel résumant l'effet des variables explicatives :

$$h(t, z_1, z_2, \ldots, z_k) = h_0(t) \exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k)$$

où les β_i sont les *paramètres de régression* inconnus du modèle.

Les covariables agissent donc multiplicativement sur le risque instantané de défaillance. D'autre part h_0 est une fonction arbitraire du temps, qui représente la fonction de risque d'un individu pour lequel toutes les covariables z_i seraient nulles.

Les β_i sont estimés suivant les méthodes que l'on verra au paragraphe suivant. Par contre la fonction h_0 peut être calculée de deux façons :

- soit elle est choisie parmi les modèles paramétriques déjà définis auparavant (Weibull, exponentiel,...) et est mise sous forme analytique. C'est alors un modèle complètement paramétrique.

- soit elle est laissée "libre" et on ne fait qu'estimer ses valeurs pour les différents temps de défaillance observés. C'est le modèle de Cox, qui est semi-paramétrique. Celui-ci n'impose pas de contraintes sur h_0 et semble, a priori, mieux adapté pour déterminer l'influence de facteurs.

La figure 24 résume les différentes approches en analyse de données de survie.

4.3. UTILISATION DU MODELE DES RISQUES PROPORTIONNELS - ESTIMATION DES PARAMETRES

Parmi les modèles vus précédemment, c'est le modèle des risques proportionnels qui a été choisi. En effet il permet tout d'abord de mesurer l'influence des différents facteurs grâce à

l'estimation des β_i. Ensuite sa forme est simple, grâce à la séparation en deux fonctions h_0 et la fonction exponentielle d'influence des covariables.

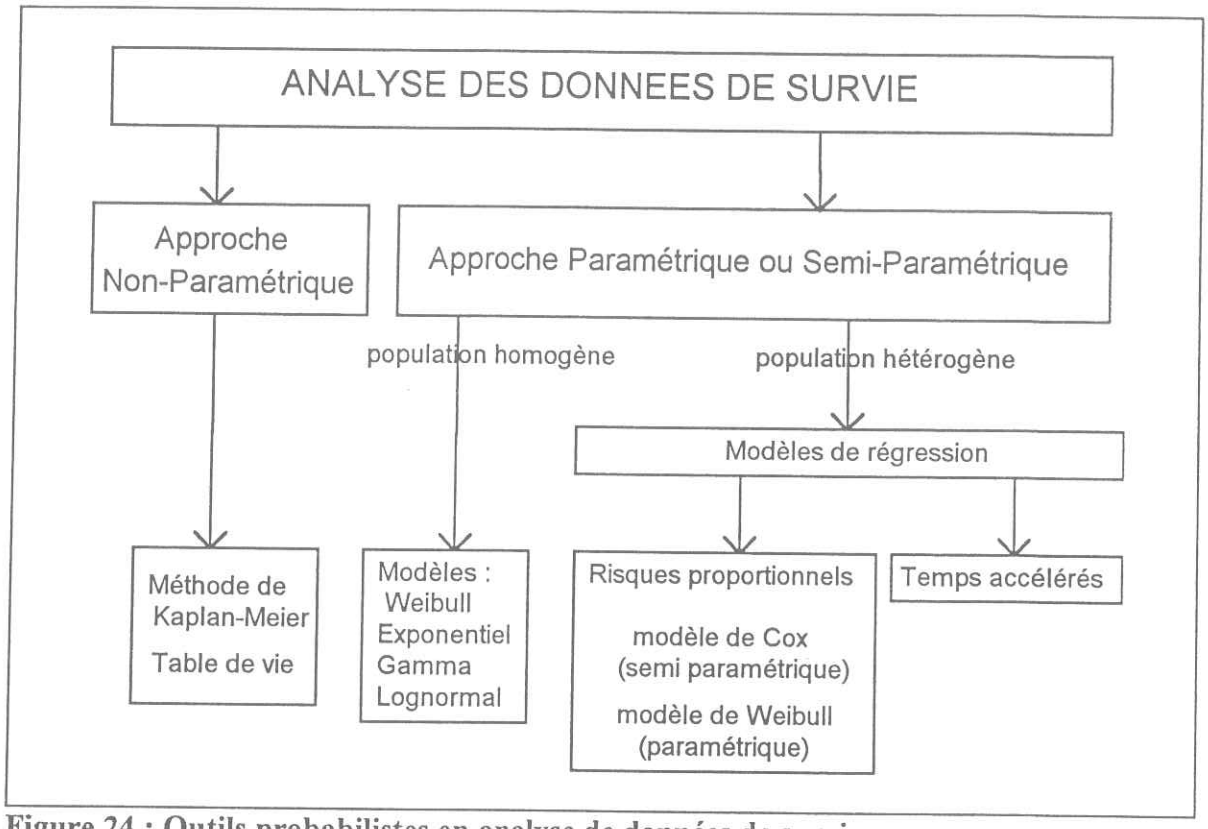


Figure 24 : Outils probabilistes en analyse de données de survie

L'estimation des paramètres β_i est réalisée par la méthode du maximum de vraisemblance. Cependant la vraisemblance utilisée ici est une vraisemblance partielle et permet de prendre en compte les données censurées.

4.3.1. Formulation de la vraisemblance

Les étapes de la construction de la vraisemblance sont données en annexe 4.La vraisemblance partielle de Cox est formulée de la manière suivante lorsqu'il n'y a pas d'ex-aequo en chacun des temps t_i :

$$V = \prod_{i=1}^{k} \frac{\exp[\beta' Z_i]}{\sum_{j \in R_i} \exp[\beta' Z_j]},$$

β' étant le vecteur des paramètres à estimer et R_i l'ensemble des individus restant au temps t_i.

Dans le cas contraire sa formulation est plus complexe, c'est à dire :

$$V = \prod_{i=1}^{k} \frac{\exp(s_i \beta)}{\sum_{l \in R_{d_i(t_i)}} \exp(s_l \beta)}$$

avec :

$s_i = \sum_{p=1}^{d_i} Z_{l_p}$, s_i la somme des covariables associées aux d_i individus défaillants au temps t_i, $R_{di(ti)}$ l'ensemble des sous-ensembles de d_i éléments choisis parmi l'ensemble des individus à risque au temps t_i R_{ti}. Cette vraisemblance partielle est d'autant plus difficile à mettre en forme que le nombre d'individus ex-aequo au temps t_i est important. Cependant une bonne approximation correspond à la formule de Breslow à savoir :

$$V = \prod_{i=1}^{k} \frac{\exp(s_i\beta)}{\left[\sum_{l \in R_{t_i}} \exp(Z_l\beta)\right]^{d_i}}.$$

Le calcul de l'estimation des paramètres se trouve en annexe 4.

4.3.2. Estimation des paramètres

La méthode du maximum de vraisemblance consiste à rechercher la valeur de β qui maximise la fonction de vraisemblance (ou de façon équivalente son logarithme). On cherche donc β tel que :

$$\frac{\partial V(t_i, \beta)}{\partial \beta} = 0.$$

Les paramètres β_i sont estimés en résolvant le système d'équation :

$$U_j(\beta) = \frac{\partial \ln V(\beta)}{\partial \beta_j} = 0.$$

On trouvera en annexe 4 la méthode d'estimation des paramètres, de leur variance et le test utilisé pour mesurer leur significativité.

Les valeurs estimées de β sont obtenues à partir de la méthode de Newton-Raphson. Elle conduit au schéma itératif suivant :

$$\beta_{k+1} = \beta_k - U'(\beta_k)^{-1} U(\beta_k) \text{ (Cf. annexe 4).}$$

La méthode utilisée pour estimer les coefficients du modèle de Weibull est présentée au chapitre 5 (§ 5.4.2.).

4.3.3. Propriétés des estimateurs du maximum de vraisemblance [Hill et al., 1990]

Les estimateurs du maximum de vraisemblance ont les propriétés suivantes :

- Ils sont asymptotiquement sans biais.

- Ils sont asymptotiquement efficaces.

- Ils ont asymptotiquement une distribution normale.

4.3.4. Le test de Wald

Pour mesurer l'intérêt d'introduire β_i dans le modèle, on pourra tester l'hypothèse $H_0 : \beta_i = \beta_{i0}$. Plusieurs tests existent pour tester H_0 contre $H_1 : \beta_i \neq \beta_{i0}$. C'est le test de Wald (ou du maximum de vraisemblance) qui est utilisé. Il mesure l'écart entre $\hat{\beta}$ *et* β_0.

La statistique de ce test est la suivante :

$$(\hat{\beta}-\beta_0)^t I(\hat{\beta})(\hat{\beta}-\beta_0),$$

$I(\hat{\beta})$ étant la matrice d'information qui contient pxp éléments (p étant le nombre de composants du vecteur β) formulés de la manière suivante :

$$I_{hj}(\beta) = -\frac{\partial^2 \ln V}{\partial\beta_h \partial\beta j} \text{ (calculé en } \beta).$$

Lorsque le vecteur β ne comprend qu'un seul paramètre, cette statistique se ramène à :

$$\frac{\left(\hat{\beta}-\beta_0\right)^2}{V\left(\hat{\beta}\right)}.$$

Cette statistique est distribuée selon une loi de χ^2 avec un nombre de degrés de liberté correspondant aux paramètres fixés par H_0. Le calcul de cette statistique permet alors d'estimer une probabilité qui déterminera l'influence réelle ou non du facteur étudié. On prendra dans l'étude une valeur égale à 0,95, valeur correspodant à une probabilité d'erreur de 95% si l'on ne considère pas quel facteur étudié est influent.

4.3.5. Estimation de la fonction de risque de base h_0 (Hill et al., 1990)

Cette fonction est estimée à partir de la fonction de risque de base cumulée H_0. Pour un modèle non stratifié cette fonction est estimée par Breslow en chaque pas de temps t_i :

$$\hat{H}_0(t_i) = \sum_{j=1}^{i}\left[\frac{m_j}{\sum_{s\in R_j}\exp\left(\hat{b}z_s\right)}\right],$$

m_j étant le nombre de défaillances au temps t_j, R_j l'ensemble des individus restant au temps t_j.

Pour un modèle stratifié et pour le groupe de stratification g, cette fonction sera égale à :

$$\hat{H}_0(t_{gi}) = \sum_{j=1}^{i} \left[\frac{m_{gj}}{\sum_{s \in R_{gj}} \exp(\hat{b} z_s)} \right],$$

m_{gj} étant le nombre de défaillances au temps t_{gj}, R_{gj} l'ensemble des individus restant au temps t_{gj}.

Les outils présentés dans ce chapitre permettent donc de tenir compte de la spécificité des données utilisées. De plus des modèles paramétriques peuvent être élaborés à partir de ces outils. La présentation de leur application aux données de défaillance se trouve dans le chapitre 5.

CHAPITRE 5 :
APPLICATION DES MODELES DE COX ET DE WEIBULL A LA PREVISION DES DEFAILLANCES DES CONDUITES D'EAU POTABLE

C'est tout d'abord le modèle de Cox qui a été utilisé afin de déterminer l'influence des facteurs, ainsi que la forme des différentes courbes de probabilité d'apparition des défaillances et de risque. Les différents calculs ont été faits à partir du logiciel EGRET, qui est présenté dans ce chapitre. Avant de présenter les résultats, nous exposerons comment les données ont été transformées, puis nous définirons la méthode de stratification. Enfin nous étudierons le modèle de Weibull.

5.1. PRESENTATION DU LOGICIEL EGRET

Le logiciel utilisé pour l'analyse de ces données est le logiciel EGRET élaboré par le SERC (Statistics and Epidemiology Research Corporation à Seattle). Il est spécialisé dans les analyses de survie et les analyses épidémiologiques. Il permet donc, entre autres, de déterminer les paramètres des modèles de Cox, de Weibull et exponentiel à partir des méthodes présentées dans le chapitre précédent. Il utilise également des tests permettant de définir si un facteur est significatif ou non.

Les variables prises en compte sont, outre les caractéristiques des individus, les variables de censure d et de temps de défaillance ou de censure t.

Le tableau 16 présente les résultats donnés pour le modèle de Cox.

Tableau 16 : Exemple de résultats donnés par le logiciel EGRET

TERM	COEFFICIENT	STDERROR	P-VALUE	RATERATIO
Longueur	0,4822	0,135	<001	1,62
Trafic-Chaussée	0,331	0,198	0,013	1,39
Sol	0,273	0,272	<001	1,31
Diamètre	-0,5957	0,112	0,001	0,55
Nature cana	-0,4446	0,2	0,009	0,64

Les différentes valeurs correspondent aux définitions suivantes :

* TERM : nom de la variable,

* COEFFICIENT : valeur du paramètre β_i calculé,

* STDERROR : écart type du paramètre estimé,

* P-VALUE : valeur correspondant à la probabilité d'une distribution de χ^2, comparable à la statistique du Test de Wald, calculé ici. Cette statistique est égale au carré du rapport entre la valeur du coefficient calculé et l'écart type.

* RATE RATIO : valeur correspondant au rapport des risques lorsque la variable change d'une unité.

Le logiciel donne également les valeurs des fonctions de risques de base h_0 et de fonctions de risques cumulées de base H_0 calculées aux différents temps de défaillances. Ces fonctions sont calculées à partir de la formule de Breslow citée dans le chapitre précédent.

D'autre part il est possible de stratifier sur le modèle de Cox suivant les valeurs prises par les variables. Le paragraphe 5.3.1. présente cette technique.

5.2. TRANSFORMATION DES DONNEES POUR L'ANALYSE DES DONNEES DE SURVIE

Afin de pouvoir différencier les variables censurées des non censurées, le logiciel EGRET nécessite de créer :

- une variable de censure appelée d

- une variable t déterminant la durée d'observation entre la dernière défaillance (ou la date de pose) et la date de censure ou la date de défaillance.

Nous allons voir comment ces données ont été transformées sachant que l'on recherche également l'influence du nombre de défaillances déjà subies par la conduite.

5.2.1. Transformation selon l'ordre d'apparition des défaillances

Les tronçons de canalisation déterminés sont de plusieurs types (Cf. fig 25).

-soit ils n'ont subi aucune défaillance. Dans ce cas on les caractérise par les valeurs d=0 et t=date de dernière observation-date de pose

Par exemple un tronçon a été posé en 1957 et la date de fin d'observation est 1991. Les variables de censures sont alors : d=0 et t=34.

- soit ils ont eu une ou plusieurs défaillances. On les caractérise alors par plusieurs individus avec *d=1* et *t=date de la défaillance-date de la défaillance précédente (ou date de pose)* et pour le dernier individu, *d=0* et *t=date de dernière observation-date de dernière défaillance)*.

Par exemple, un tronçon a été posé en 1964 et sa date de fin d'observation est 1991. De plus il a subi 3 défaillances en 1975, 1985 et 1988. Ce tronçon est transformé en quatre individus t1, t2, t3, t4, tels que :

pour t1, d=1 et t=(1975-1964)=11,

pour t2, d=1 et t=(1985-1975)=10,

pour t3, d=1 et t=(1988-1985)=3,

pour t4, d=0 et t=(1991-1988)=3.

De plus ces individus se différencient par une variable supplémentaire "Nombre de défaillances antérieures" abrégée en "Ndef". Par exemple on a :

pour t1, Ndef=0,

pour t2, Ndef=1,

pour t3, Ndef=2,

pour t4, Ndef=3,

Le nombre d'individus étudiés dans EGRET est donc différent du nombre de tronçons.

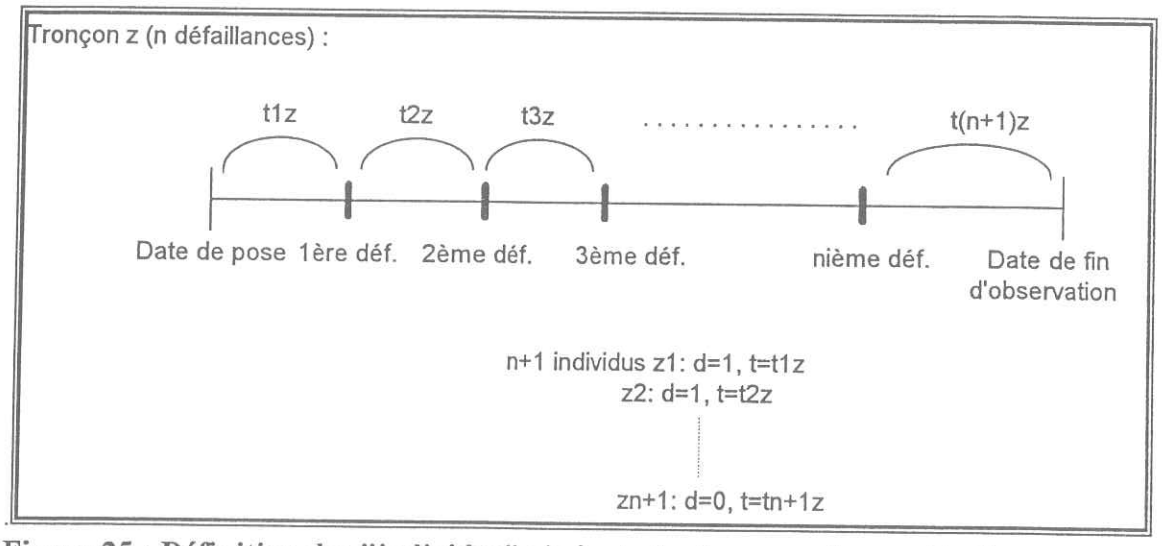


Figure 25 : Définition des "individus" vis-à-vis des tronçons de canalisations

Nous allons voir maintenant comment les variables ont été codées.

5.2.2. Le codage des variables

Les variables et leur codage sont les suivants :

* **Variable de censure d**
  =0 : Individu censuré (défaillance non observée)
  =1 : Individu défaillant

* **Temps de défaillance ou de censure T**
  si d=0, date (en années) de fin d'observation de l'individu depuis, soit la date de pose si l'individu n'a pas eu de défaillance, soit la date de la dernière défaillance, si l'individu en a eu,
  si d=1, date de la défaillance depuis, soit la date de pose si l'individu n'a pas eu de défaillance, soit la date de la dernière défaillance, si l'individu en a eu.

* **Nombre de défaillance(s) déjà subie(s) par l'individu Ndef**

*** Emplacement de la conduite sous chaussée ou sous trottoir**
=0, si la conduite se trouve sous trottoir,
=0.5, si une partie se trouve sous trottoir et une partie sous chaussée,
=1, si elle se trouve sous chaussée.

*** Sol et sa corrosivité**
=0 : Sol non corrosif
=1 : Sol corrosif
La corrosivité des sols a été déterminée à partir soit de cartes de résistivité des sols (Bordeaux), soit de renseignements pris sur le terrain auprès d'inspecteur de réseaux (CEB).

*** Trafic automobile :**
= 0 : Trafic automobile faible
= 1 : Trafic automobile élevé
Le trafic automobile est considéré élevé dans une rue soit lorsque la mesure du trafic dans cette rue a été évaluée par les services de voirie, soit

*** Origine de l'eau (CEB)**
=0 : Origine de l'eau 1
=1 : Origine de l'eau 2
=0.5 : Eau provenant des deux ressources

*** Diamètre (en mm)**

*** Date de pose de la conduite (en années)**

*** Apnn**
=0 : Conduite posée avant 19nn
=1 : Conduite posée après 19nn
Dans la suite de l'étude, le facteur Ap66 sera appelé "Nature cana." car il concerne le passage technologique de la fonte grise à la fonte ductile, ainsi que l'apparition de nouveaux joints.

*** Longueur de la conduite (en m)**

*** Matériau :**
=0 Acier
=1 fonte
=2 PVC ou Polyéthylène
=3 Amiante-ciment

*** Facteur croisé Trafic-Chaussée (Trachau) :**
=0 : Conduite se trouvant, soit sous trottoir, soit dans une rue à trafic faible
=1 : Conduite se trouvant sous chaussée dans une rue à trafic élevé

5.3. LE MODELE DE COX

5.3.1. La stratification

Le modèle des risques proportionnels est de la forme suivante :

$$h(t, z_1, z_2, \ldots, z_k) = h_0(t) \exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k)$$

Il fait donc l'hypothèse que les risques sont proportionnels lorsque la valeur d'un facteur de risque change. Par exemple, supposons que l'on ne tienne compte que du facteur z_t avec

$z_t = 0$ pour un trafic faible

$= 1$ pour un trafic élevé

Le risque pour un tronçon placé sous une route à trafic faible est de la forme :

$$h(t,0) = h_0(t) \times \exp(\beta_t \times 0) = h_0(t).$$

Pour un tronçon placé sous une route à trafic élevée, il est de :

$$h(t,1) = h_0(t) \times \exp(\beta_t \times 1) = h_0(t) \times \exp(\beta_t)$$

donc on a :

$$\forall t, \frac{h(t,1)}{h(t,0)} = \exp(\beta_t).$$

Le rapport entre les risques est donc le même quel que soit la durée t.

Il est donc important de vérifier l'hypothèse suivante pour chacun des facteurs z_i :

$$\forall t, \frac{h(t,z_i')}{h(t,z_i)} = \exp\left[\beta_i\left(z_i' - z_i\right)\right] = \text{constante}.$$

Ceci peut se vérifier graphiquement à partir de la fonction $\ln(H(t)) = \ln(-\ln(S(t)))$, H étant la fonction de risque cumulée. En effet la formule précédente donne :

$$\forall t, \frac{H(t,z_i')}{H(t,z_i)} = \exp\left[\beta_i\left(z_i' - z_i\right)\right] = \mathrm{K}.$$

D'où, en prenant le logarithme :

$$\forall t, \ln\left[H(t,z_i')\right] - \ln\left[H(t,z_i)\right] = \left[\beta_i\left(z_i' - z_i\right)\right] = \ln K = K'.$$

Le logiciel EGRET permet de vérifier graphiquement cette hypothèse. La courbe de la figure 28 nous montre un exemple pour le facteur Trafic-Chaussée.

Si cette hypothèse n'est pas vérifiée, il est alors nécessaire de "stratifier" l'échantillon en fonction des valeurs prises par le facteur i. Par exemple si la variable j prend les valeurs binaires 0 ou 1, l'échantillon est divisé en deux groupes pour lesquels sont calculées deux fonctions de risque de base h_0 :

Le modèle pour les individus dont la valeur du facteur j $z_j=0$ est :

$$h^0(t,z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_k)=h_0^0(t)\exp(\beta_1 z_1+\ldots+\beta_{j-1}z_{j-1}+\beta_{j+1}z_{j+1}+\ldots+\beta_k z_k).$$

Lorsque $z_j=1$:

$$h^1(t,z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_k)=h_0^1(t)\exp(\beta_1 z_1+\ldots+\beta_{j-1}z_{j-1}+\beta_{j+1}z_{j+1}+\ldots+\beta_k z_k).$$

Les valeurs des paramètres β_i ($i\neq j$) sont communes aux deux modèles. Le logiciel EGRET permet de calculer les β_i pour le modèle de Cox. pour le modèle de Weibull, le logiciel ne calcule pas de modèle stratifié. Une méthode, expliquée au paragraphe 5.4.2.,a été élaborée.

5.3.2. Résultats

Les résultats sont présentés suivant la chronologie des services étudiés. C'est à dire tout d'abord le service de la LED-Bordeaux, puis la CEB, enfin les deux syndicats du Bas-Rhin.

Concernant ces deux syndicats, les données de défaillances n'existaient que depuis 1976, alors qu'un certain nombre de conduites ont été posées bien avant cette date. Ne prendre que les conduites posées après 1976 aurait donné une taille d'échantillon trop faible. On a donc étudié les conduites posées avant 1976, conduites appelées "censures à gauche", en y appliquant les modèles de Weibull de Bordeaux et de la CEB sur la période comprise entre la date de pose et 1976. La création d'un échantillon de défaillances à partir d'un modèle connu est présentée au chapitre 7.

5.3.2.1. Estimation des facteurs influents

Le test de Wald donne une statistique suivant une statistique du χ^2. La probabilité (p-value dans EGRET) maximale définissant un facteur comme significatif a été prise égale à 0,05. Ceci signifie que l'on a plus de 95 chances sur 100 de faire en rejettant ce facteur.

Le χ^2 calculé a moins de 5 chances / 100 d'apparaître sous l'hypothèse modèle $\beta_i = 0$.

Afin de définir les facteurs influents, on a tout d'abord établi le test d'influence en prenant en compte tous les facteurs, puis ont été retranché, un à un, les facteurs ne satisfaisant pas le test, c'est à dire ayant une "p-value" supérieure à 0,05.

Les résultats des tests sur l'influence des facteurs se trouvent dans le tableau 17.

Les facteurs significatifs pour chacun des trois services sont donc :

Pour la LED Bordeaux :
- la longueur de tronçon,
- le sol et sa corrosivité,
- la pose de la conduite avant ou après 1966 (Nature cana),
- la pose de la conduite avant ou après 1980,

Tableau 17 : Estimation des facteurs influents (test de Wald)

	"p-value"		
Facteurs	LED-Bordeaux	CEB	Bas-Rhin
Longueur	<0,001	<0,001	<0,001
Diamètre	<0,001	0,001	0,050
Sol	<0,001	0,013	
Trafic-Chaussée	0,048	0,001	0,043 (Trafic seul)
Nature cana.	0,040	<0,001	
Ap75	0,049	0,008	

- le diamètre,
- le nombre de défaillances subies par la conduite,
- le facteur croisé Trafic-Chaussée.

Pour la CEB :

- la longueur de tronçon,
- le sol et sa corrosivité,
- la pose de la conduite avant ou après 1966 (Nature cana),
- la pose de la conduite avant ou après 1975,
- le diamètre,
- le nombre de défaillances subies par la conduite
- le facteur croisé Trafic-Chaussée.

Pour les syndicats ruraux :

- la longueur,
- le diamètre,
- le nombre de défaillances subies par la conduite,
- le trafic automobile,
- le facteur croisé Trafic-Chaussée.

Entre les deux services urbains, les facteurs significatifs sont pratiquement les mêmes, excepté le trafic et les dates de pose seuil. Le facteur croisé trafic-chaussée (Trachau) est significatif sur les deux services.

Les dates de pose seuil correspondent à deux phénomènes différents. Pour la date se situant autour des années soixante, il s'agit de différencier les conduites en fonte grise des conduites en fonte ductile. Cette date n'est pas la même pour les deux services à cause des stocks de canalisations en fonte grise constitués par ces services à cette époque.

Pour les dates se situant à la fin des années soixante-dix, il pourrait s'agir d'un meilleur soin apporté à la pose et à l'entretien des conduites. Ceci est d'autant plus net pour la CEB, car depuis 1980 toutes les conduites sont posées avec une protection extérieure en polyéthylène.

Pour les syndicats ruraux le nombre de facteurs significatifs est plus faible, ceci pour plusieurs raisons. Tout d'abord l'échantillon observé est plus petit (935 tronçons). Ensuite certains facteurs n'ont pu être recueillis avec précision. Par exemple pour le sol, on a du se baser sur les cartes géologiques au 50 000ème.

5.3.2.2. Vérification de l'hypothèse de proportionnalité des risques

Une fois les facteurs influents déterminés, la vérification de la proportionnalité a permis de définir sur quels facteurs il était nécessaire d'introduire une fonction h_0 stratifiée (Cf § 5.3.). Cette vérification a eu lieu graphiquement à partir des courbes de la fonction $\ln(H(t))$.

**** Stratification suivant le groupe de défaillances***

La figure 26 représente l'écart moyen entre les différentes défaillances suivant leur ordre. Cette figure montre que le phénomène "première défaillance" se distingue des autres défaillances. En effet la première défaillance apparaît en moyenne 14 à 16 ans après la date de pose, pour les conduites qui en ont eu une.

Concernant les défaillances suivantes, deux phénomènes apparaissent. De la deuxième à la quatrième défaillance, l'écart entre deux défaillances diminue de 6 à 4 ans. Pour les défaillances suivantes à partir de la cinquième, cet écart reste constant autour de 3 à 4 ans.

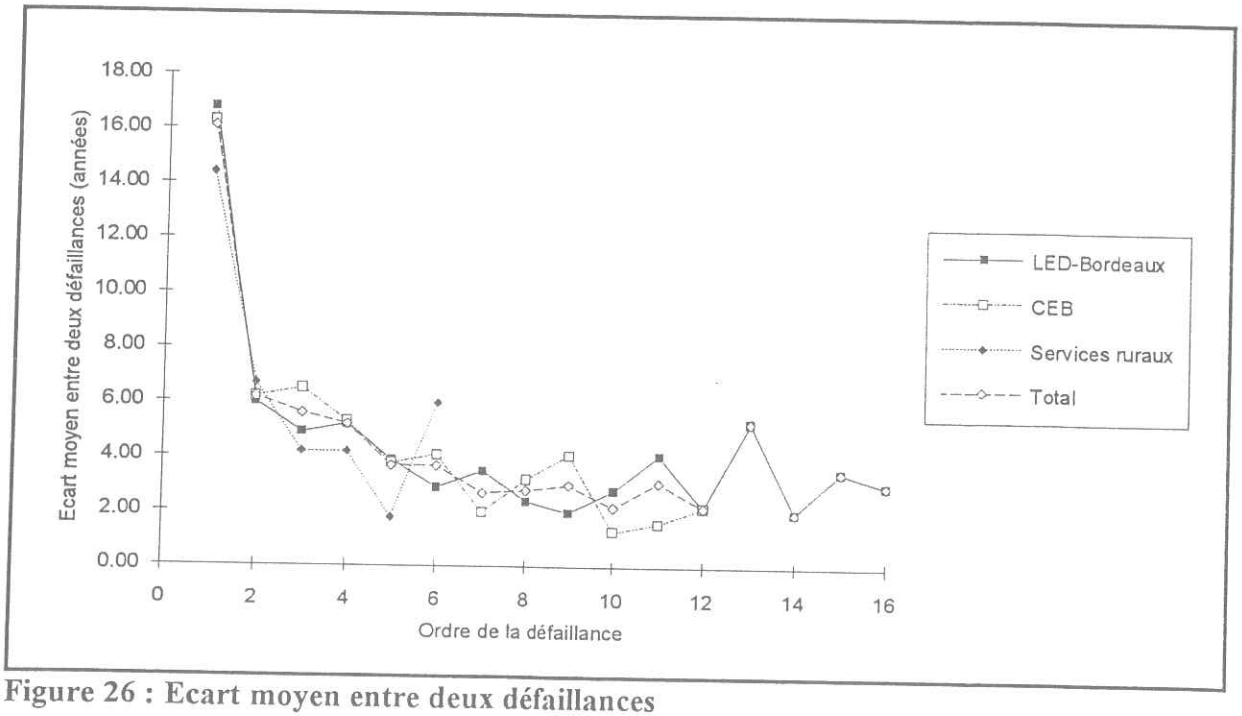


Figure 26 : Ecart moyen entre deux défaillances

Trois groupes de défaillances ont donc été distingués suivant l'ordre de la défaillance, présentés dans le tableau 18.

Tableau 18 : Groupes de défaillances

Groupe 1	1ère défaillance
Groupe 2	2 à 4ème défaillance
Groupe 3	5ème défaillance et plus

La figure 27 présente la courbe de la fonction ln(H(t)) pour chacun des trois groupes de défaillances sur la CEB.

Cette réflexion est pertinente pour les services urbains et permet donc de conclure :

- la première défaillance est un phénomène de vieillissement, car sa courbe de risque augmente avec le temps (action de l'âge) ;
- les trois défaillances suivantes sont moins dépendantes du temps;
- à partir de la 5ème, les défaillances sont indépendantes les unes des autres. Ce n'est plus un phénomène de vieillissement car le risque est le même quelque soit la défaillance.

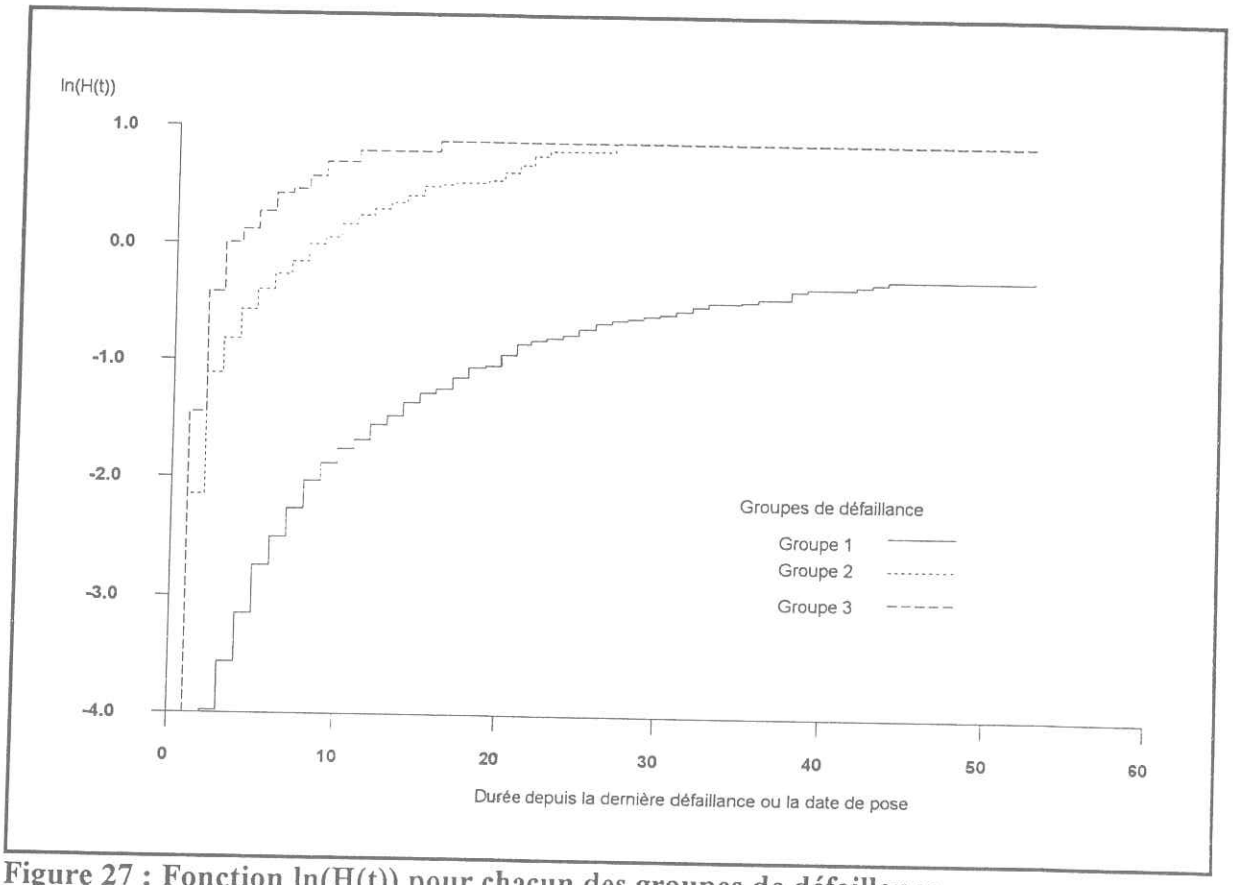


Figure 27 : Fonction ln(H(t)) pour chacun des groupes de défaillance

Pour ce qui concerne le service rural, le phénomène est un peu différent, car on ne peut voir de différences entre les défaillances à partir de la deuxième. On a alors deux groupes de défaillances : celui représentant la première et celui représentant les suivantes. Ceci est dû à la

faiblesse de l'échantillon pour les défaillances supérieures à la 4ème (uniquement deux individus).

Dans la suite de l'étude on stratifie sur le modèle de Cox, comme sur le modèle de Weibull, sur ces groupes de défaillance, pour ce qui concerne les deux services urbains. Les trois groupes ont donc trois fonctions h_0 différentes. Par contre les paramètres des facteurs influençant les défaillances sont les mêmes quelque soit le groupe de défaillance.

** Les autres facteurs*

Concernant les autres facteurs, le positionnement des courbes de la fonction ln((H(t)) est satisfaisant et vérifie l'hypothèse. La figure 28 montre les courbes correspondant aux deux valeurs possibles du facteur croisé trafic-chaussée pour la CEB.

La méthode de stratification est ainsi utilisée pour les groupes de défaillance. Pour les autres facteurs on estime les paramètres β_i comme cela avait été présenté dans le chapitre 4.

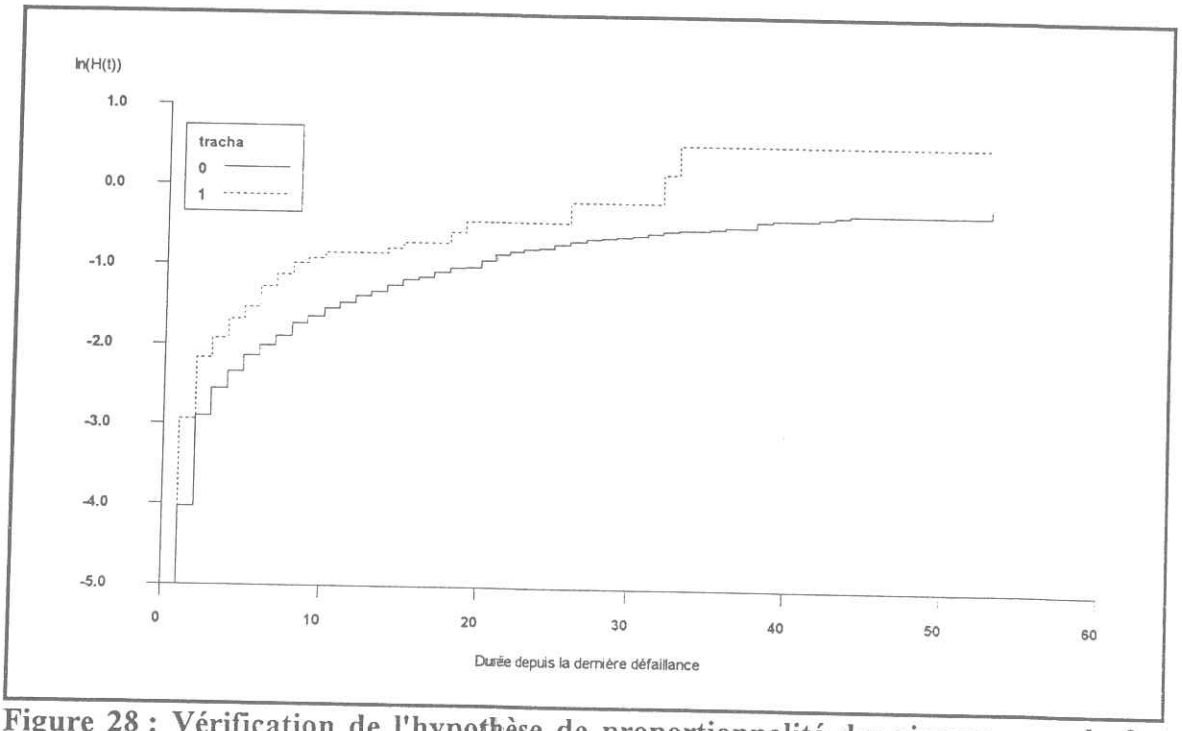


Figure 28 : Vérification de l'hypothèse de proportionnalité des risques pour le facteur Trafic-Chaussée.

5.3.2.3. Estimation des paramètres β_i

Les résultats concernant l'estimation des β_i se trouvent dans le tableau 19.

Certains paramètres sont sensiblement identiques. Ce sont :

- la longueur entre 0,37 et 0,52. Une réflexion concernant ce paramètre est présentée au paragraphe 5.5.1..
- le sol autour de 0,30.

Tableau 19 : Résultats de l'estimation des β_i du modèle de Cox (le chiffre entre parenthèse correspond à l'écart type)

	β_i					
Facteurs	LED-Bordeaux		CEB		Bas-Rhin	
Longueur	**0,52**	(0,06)	**0,37**	(0,05)	**0,48**	(0,10)
Diamètre	**-0,80**	(0,14)	**-0,34**	(0,10)	**-0,37**	(0,30)
Sol	**0,32**	(0,09)	**0,24**	(0,10)		
Trafic-Chaussée	**0,23**	(0,12)	**0,58**	(0,18)	**-0,45** (Trafic seul)	(0,22)
Nature cana.	**-0,31**	(0,15)	**-0,64**	(0,15)		
Ap75	**-0,36**	(0,27)	**-1,39**	(0,52)		

Le paramètre diamètre est du même ordre entre la CEB et les syndicats ruraux, mais est très différent pour ce qui concerne le réseau de la LED-Bordeaux. Cette différence est due à l'échantillonnage. En effet sur la CEB et les syndicats ruraux, tous les diamètres avaient été utilisés de 60 à 300 mm. Sur Bordeaux ce sont les diamètres 60 et 150 mm qui avaient été d'abord choisis sur tout le secteur, puis les diamètres 80 et 100 mm sur la seule commune de Mérignac.

Si l'on utilise uniquement les données recueillies sur cette commune, le coefficient correspondant au diamètre est de *-0,4*, ce qui rapproche des valeurs établies sur la CEB et les syndicats ruraux.

Pour les autres paramètres les valeurs sont différentes. Pour le facteur "Ap75", cela peut s'expliquer par le fait que sur la CEB les conduites sont posées avec des manchons en polyéthylène depuis 1975, et non sur Bordeaux. Ceci explique également la différence concernant le facteur "pose avant ou après 1966" ou "Nature cana", qui correspond au passage entre la fonte grise et la fonte ductile. L'incertitude liée au stockage des conduites à cette époque doit également jouer un rôle important.

Concernant le facteur croisé trafic-chaussée, la différence ne peut être expliquée que par le trafic qui est plus important sur la CEB (secteur entièrement urbain) ou par la technique de pose. Sur les deux syndicats ruraux, c'est le trafic seul qui a été pris en compte. La valeur négative montre que les conduites cassent moins lorsque le trafic est élevé, ce qui est contraire au résultat auquel on aurait pu s'attendre. Ceci peut s'expliquer par un meilleur soin appliqué à la pose des conduites (profondeur plus importante, meilleur lit de pose) dans les rues à trafic élevé et notamment sur les routes nationales et départementales.

Le tableau 20 représente les bornes des intervalles de confiance, lorsque l'on prend une valeur de 95 % (les β_i étant considéré comme suivant une loi normale de moyenne la valeur $\hat{\beta}_i$ et d'écart type l'écart type estimé $\hat{\sigma}_i$). Ces valeurs sont donc calculées pour chacun des facteurs pris séparément. Cependant on peut également estimer un intervalle de confiance sur la valeur globale $\left(\sum_i \beta_i z_i\right)$ particulière à chaque conduite.

Tableau 20 : Bornes des intervalles de confiance (95 %)

BORNES DES INTERVALLES DE CONFIANCE						
	LED-Bordeaux		CEB		Bas-Rhin	
Longueur	0.40	0.64	0.27	0.47	0.27	0.68
Diamètre	-1.08	-0.52	-0.55	-0.14	-0.98	0.24
sol	0.14	0.50	0.05	0.44	**	**
trafic-chaussée	0.00	0.46	0.22	0.93	-0.89	-0.01
Nature cana	-0.62	-0.01	-0.94	-0.34	**	**
Ap75	-0.89	0.16	-2.42	-0.36	**	**

Cette estimation se fait de la façon suivante. Comme l'on considère que les β_i suivent une loi normale $N(\hat{\beta}_i, \hat{\sigma}_i)$, la variable aléatoire $(\beta_i z_i)$ suit une loi normale $N(\hat{\beta}_i z_i, \hat{\sigma}_i z_i)$. En faisant l'hypothèse que les variables aléatoires $\hat{\beta}_i$ sont indépendantes les unes des autres, on aura la variable aléatoire $\left(\sum_i \beta_i z_i\right)$ qui suivra une loi normale $N(\sum_i \hat{\beta}_i z_i, \sqrt{\sum_i (\hat{\sigma}_i z_i)^2})$.

On peut alors estimer l'intervalle de confiance à 95 % pour chacune des conduites. La courbe de la figure 29 présente le cas de la conduite ayant l'intervalle de confiance le plus défavorable. On a dans ce cas un intervalle de 0,45 au bout de 40 ans. Pour les défaillances suivantes l'intervalle diminue en général de moitié.

La figure 30 représente une courbe dont l'intervalle de confiance est plus petit.

Dans le chapitre 6, nous verrons quelle influence a cet intervalle sur l'estimation du nombre total de défaillances.

5.3.2.4. Quelques courbes de probabilité de non-défaillance et de risque

Une fois les paramètres déterminés, il est possible de tracer les courbes de risque et de non-défaillance calculées par le logiciel EGRET.

** Courbes de fonctions de risque*

La figure 31 représente le nuage de points de la fonction de risque de base h_0 de première défaillance sur le réseau de la LED-Bordeaux tous diamètres confondus. On peut se rendre compte que pour certaines durées supérieures à 35 ans depuis la date de pose, ce risque est nul. Ceci est dû à la faiblesse de l'échantillon pour ces durées.

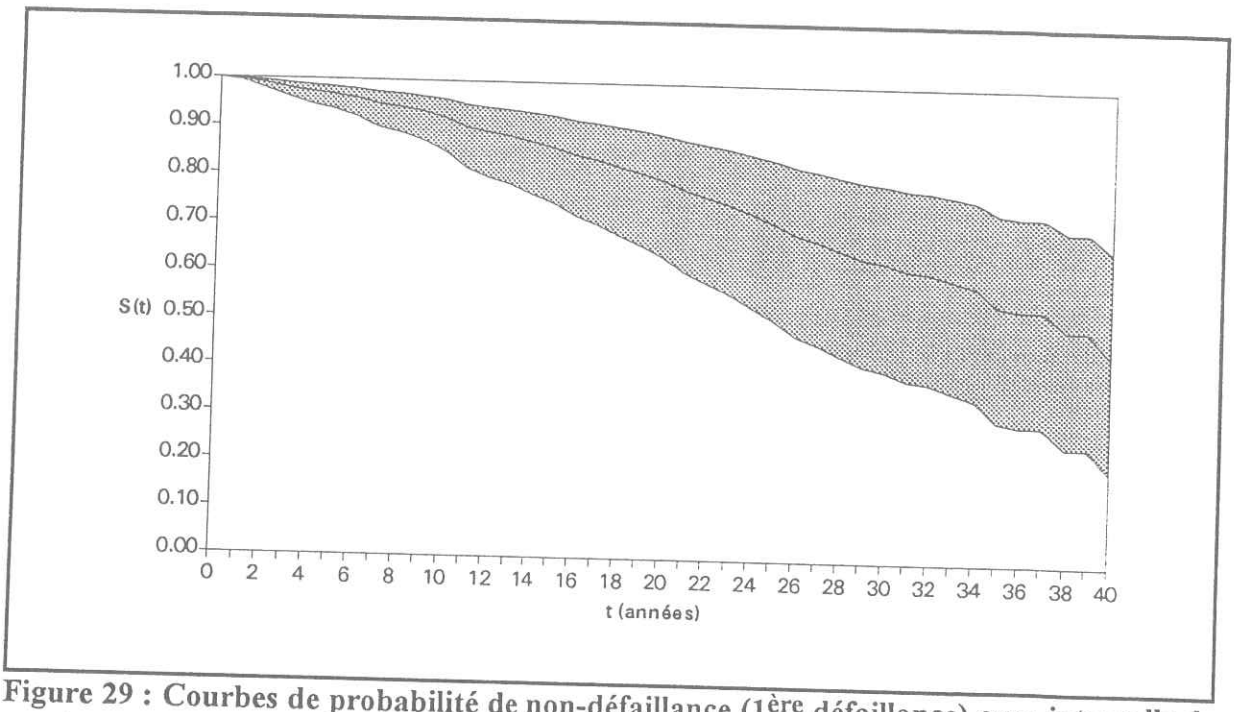


Figure 29 : Courbes de probabilité de non-défaillance (1ère défaillance) avec intervalle de confiance à 95 % pour une conduite de 500 m de long, de diamètre 100, posée après 1975, dans un sol corrosif, dans une rue à trafic fort et sous chaussée.

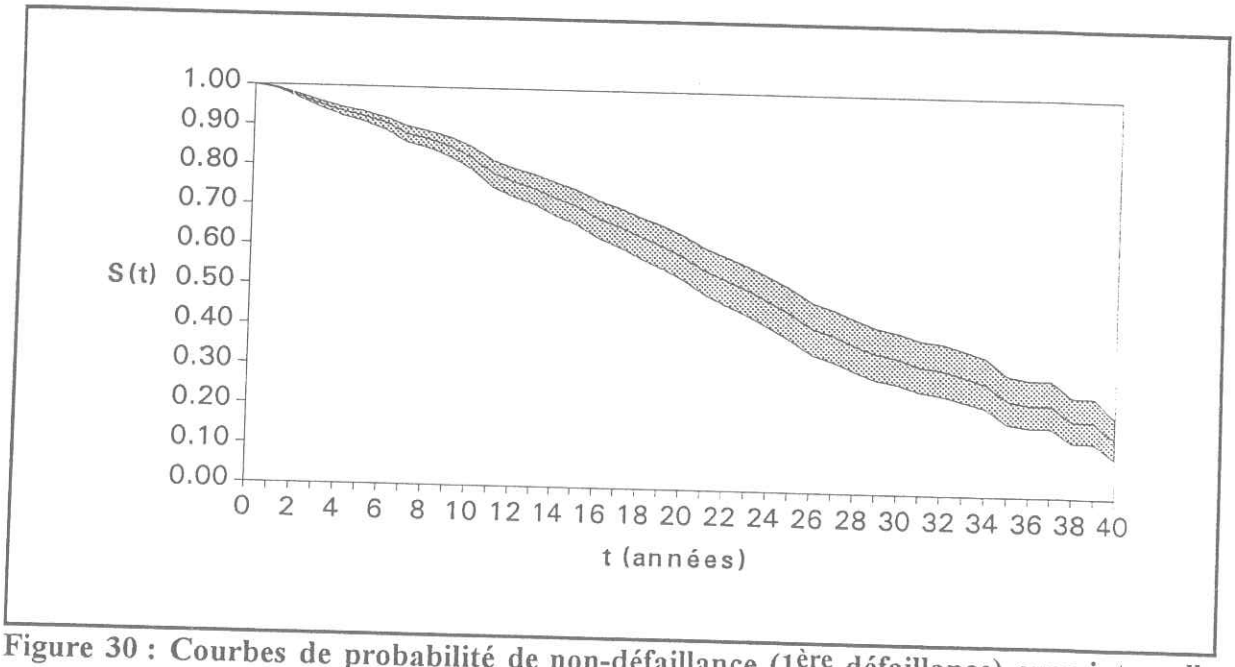


Figure 30 : Courbes de probabilité de non-défaillance (1ère défaillance) avec intervalle de confiance à 95 % pour une conduite de 500 m de long, de diamètre 100, posée avant 1966 (fonte grise), dans un sol corrosif et dans une rue à trafic faible ou sous trottoir.

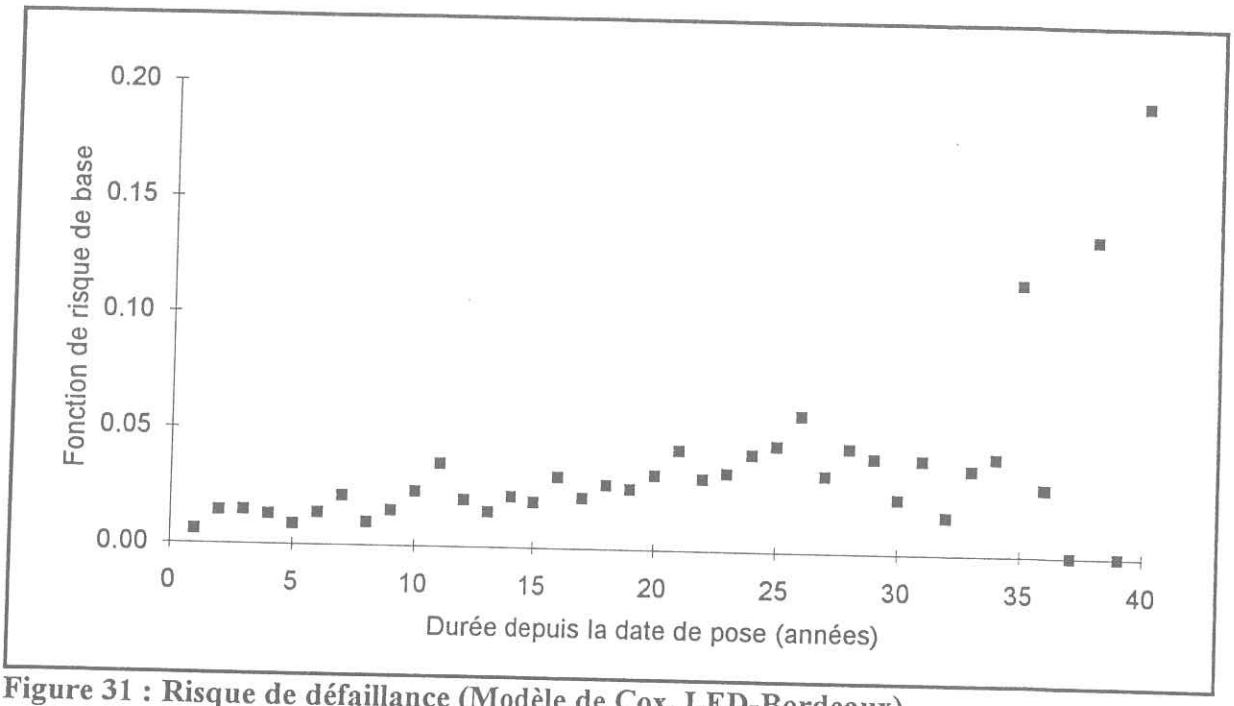


Figure 31 : Risque de défaillance (Modèle de Cox, LED-Bordeaux)

Plutôt que de calculer le risque de défaillance tous les ans, on peut aussi l'estimer tous les trois ans afin de mieux évaluer le phénomène de vieillissement. La formule utilisée est la suivante :

$$h_{moy}(t,Z) = \frac{\sum_{i \in \text{intervalle}} h(t_i,Z) * S(t_{i-1},Z)}{S(t_j,Z)}$$

t_j étant le temps au début de l'intervalle.

Par exemple si l'on veut estimer le risque d'avoir une défaillance d'ici trois ans à partir de la date de pose, on calcule :

$$h_{moy}(t,Z) = \frac{h(1,Z) * S(0,Z) + h(2,Z) * S(1,Z) + h(3,Z) * S(2,Z)}{S(0,Z)}$$

La figure 32 présente les trois courbes de fonctions de risque sur la LED-Bordeaux. Pour la première défaillance, le risque est d'abord faible et croît avec l'âge. Il y a phénomène de vieillissement. Pour les 2 à 4ème défaillances, le risque est tout de suite élevé, puis reste à peu près constant. Pour la 5ème défaillance et plus, il est encore plus élevé dès le début. Pour ces deux dernières courbes, les valeurs se trouvant aux temps les plus élevés sont moins fiables que celles données aux temps les plus faibles car l'échantillon est beaucoup plus petit.

** Courbes de probabilités de non-défaillances*

Les figures 33 et 34 présentent quelques courbes de probabilités de non-défaillance pour différents types de conduites.

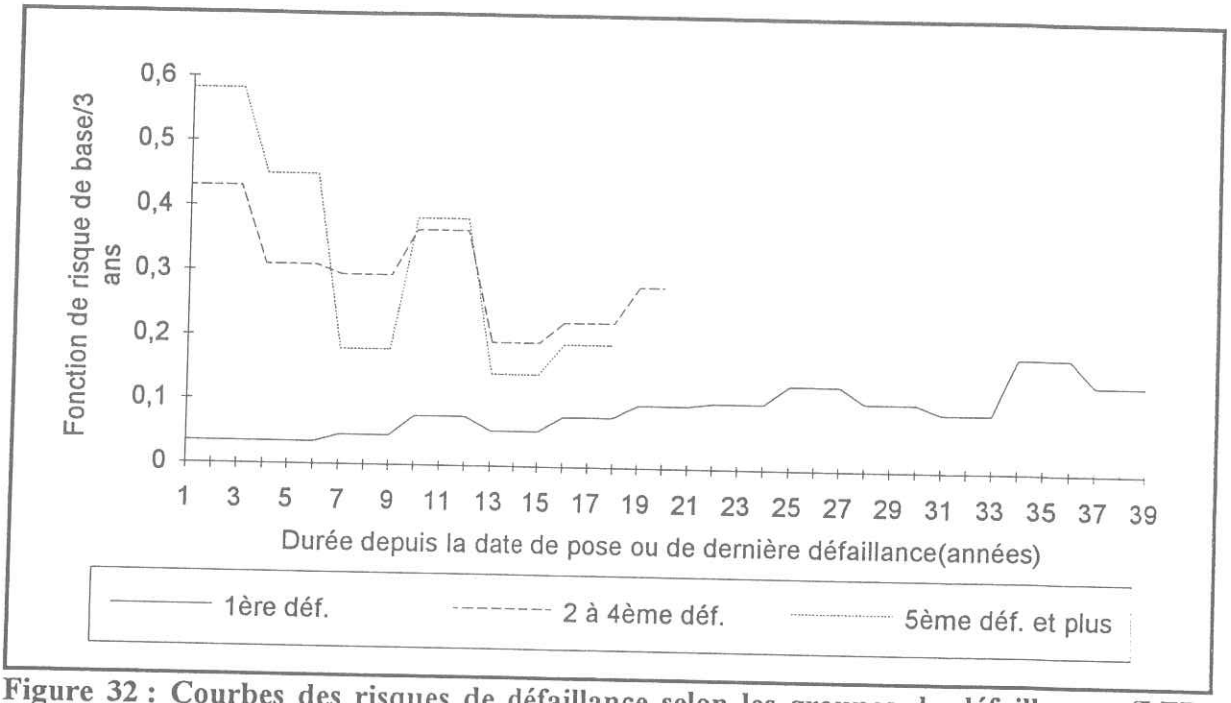


Figure 32 : Courbes des risques de défaillance selon les groupes de défaillances (LED-Bordeaux)

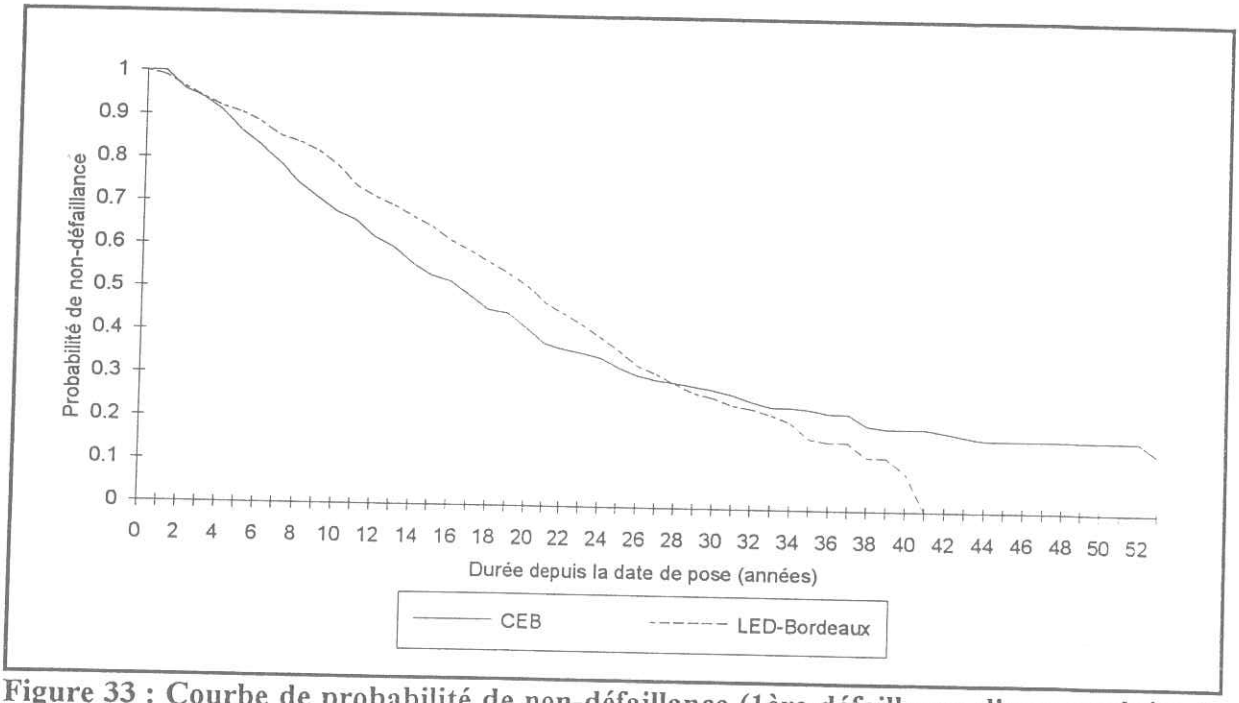


Figure 33 : Courbe de probabilité de non-défaillance (1ère défaillance d'une conduite de 150 mm, de 500 m de long, posée avant 1975, dans un sol corrosif, dans une rue à trafic élevé et sous chaussée)

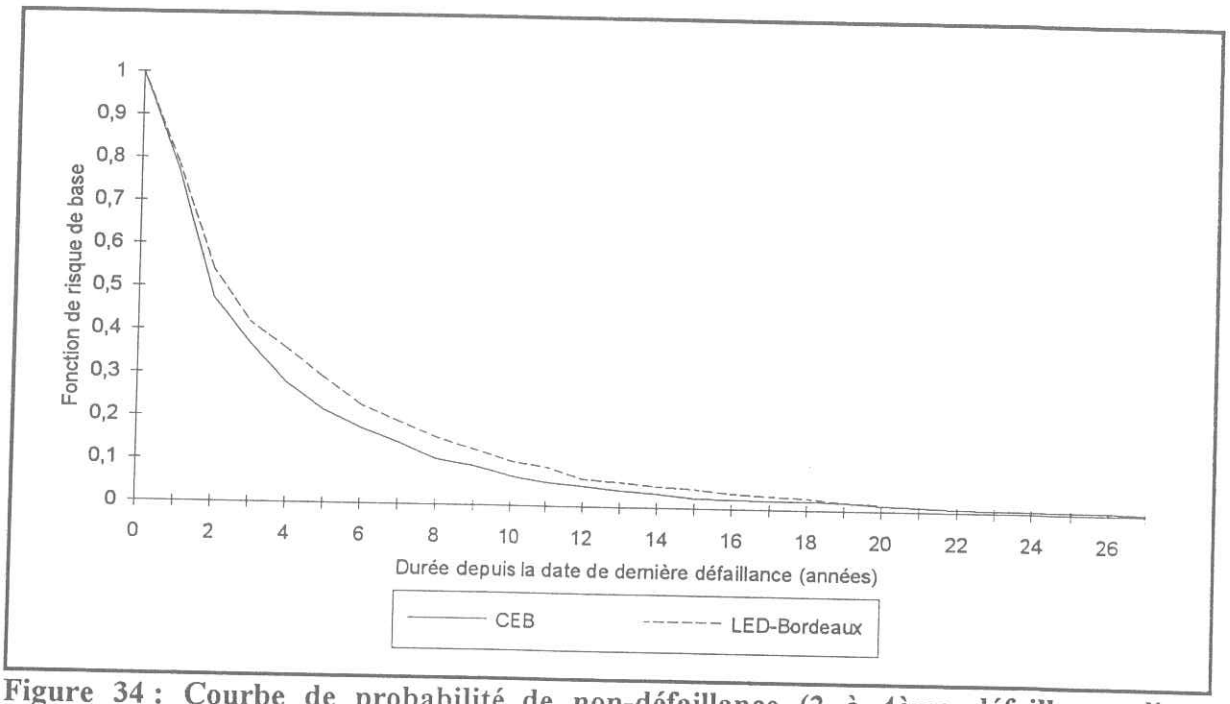


Figure 34 : Courbe de probabilité de non-défaillance (2 à 4ème défaillance d'une conduite de 60 mm, de 500 m de long, posée avant 1966 (fonte grise), dans un sol corrosif, dans une rue à trafic fort, sous chaussée)

Les autres courbes se trouvent en annexe 5.

Comme on le voit, les valeurs des différentes fonctions ne sont calculées que sur la période d'observation du fait que la fonction de risque de base h_0 n'est pas mise sous forme analytique. Ceci n'est pas gênant pour les premières défaillances car l'observation a lieu, pour la CEB et Bordeaux sur 55 et 40 ans. Par contre pour les deuxièmes défaillances et plus, cette durée d'observation n'est au maximum que de vingt ans. C'est l'inconvénient des modèles semi-paramétriques qui, s'ils estiment avec le minimum de contraintes les paramètres des facteurs influents, ne permettent pas de faire des prévisions plus longues que l'observation.

5.3.3. Un paramètre particulier : la longueur des tronçons

Les tronçons de canalisations échantillonnés pour l'étude n'ont pas tous la même longueur (entre 20 et 2000 m). Ceci est dû aux données disponibles qui ne permettent pas toujours d'identifier avec précision l'emplacement de la défaillance le long d'un tronçon.

Il est donc utile de prendre en compte la longueur comme un facteur influent. A priori il semblerait logique de considérer que le risque de défaillances varie proportionnellement à la longueur. Sachant que la fonction de risque h est de la forme :

$$h(t, z_1, z_2, \ldots, z_k) = h_0(t)\exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k),$$

nous avons pris comme facteur z_1 caractérisant la longueur, le logarithme népérien de la longueur. Ainsi le résultat attendu pour le paramètre β_1 doit être de l'ordre de 1 dans l'hypothèse

de proportionnalité entre le nombre de défaillances et la longueur. En effet dans cette hypothèse on a :

$$h(t,Z) = h_0(t) \times \exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k) \times l$$

$$h(t,Z) = h_0(t) \times \exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k) \times \exp[\ln(l)]$$

$$h(t,Z) = h_0(t) \times \exp[1 \times \ln(l) + \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k]$$

Or le paramètre β_l est de l'ordre de 0,52 pour la LED-Bordeaux, de 0,37 pour la CEB et de 0,48 pour les deux syndicats du Bas-Rhin. Le risque de défaillance serait donc plutôt proportionnel à la racine carrée de la longueur. L'étude du M.I.T. [Andreou, 1986] avait déjà permis de faire cette constatation.

Deux raisons principales pourraient expliquer cette valeur de 0,5 :

- les tronçons étudiés ne sont pas homogènes vis-à-vis des facteurs retenus, notamment pour les tronçons les plus longs,
- les défaillances ne sont pas indépendantes les unes des autres et ont souvent lieu en des endroits bien précis des tronçons.

Concernant la première explication, cela signifierait que s'il l'on prend uniquement les tronçons les plus courts, sur lesquels l'erreur est plus faible, ce coefficient β_l serait plus grand. C'est le cas sur la CEB, mais non sur Bordeaux (Cf. figure 35). Même s'il est évident que l'on fait des erreurs lorsque l'on affirme qu'un tronçon de 2500 m est, par exemple, placé dans un sol corrosif sur toute sa longueur, l'hétérogénéité des tronçons ne constitue pas l'explication principale.

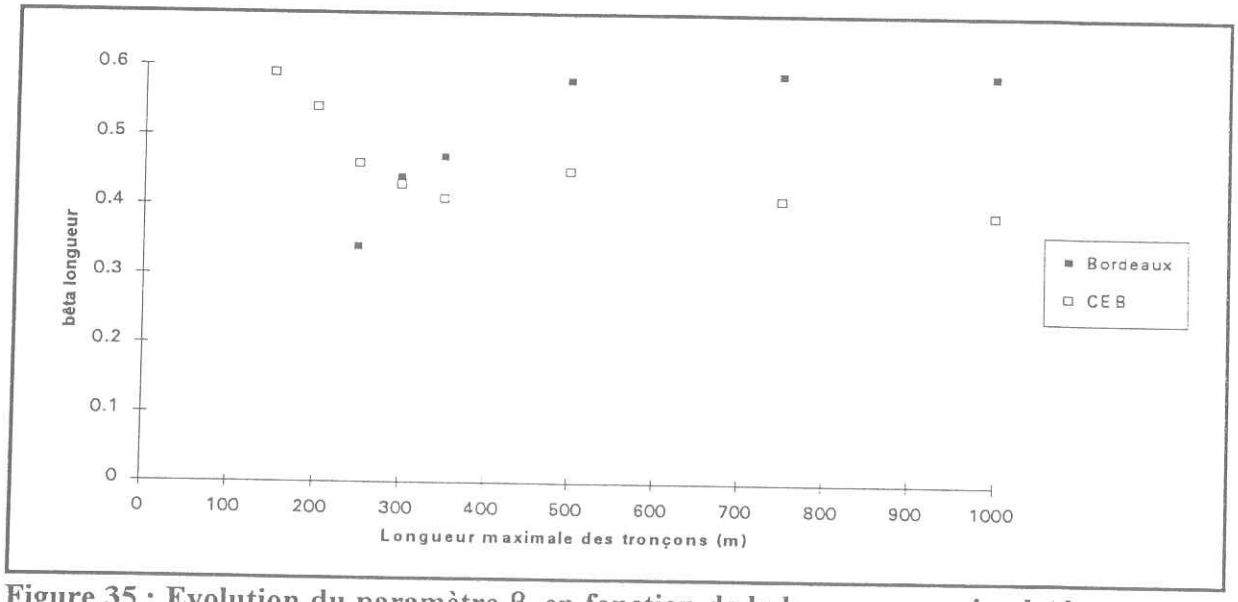


Figure 35 : Evolution du paramètre β_l en fonction de la longueur maximale des tronçons de l'échantillon initial.

L'hétérogénéité peut également concerner des facteurs dont on ne tient pas compte, comme, par exemple, le passage de nombreux poids lourds à un certain endroit de la conduite du fait de la proximité d'une usine. Ceci affaiblit alors la canalisation à cet endroit et favorise l'apparition de défaillances. Un phénomène local a alors lieu. Là encore ce phénomène est difficile à observer, et même matériellement impossible au niveau d'un service.

Concernant la deuxième explication, plusieurs raisons permettent de l'avancer. Tout d'abord on a pu se rendre compte lors du recueil des données que, lorsqu'une conduite cassait en un endroit précis, la casse suivante, lorsqu'elle a lieu, se situe souvent dans la proximité de la précédente, et même quelquefois au même endroit. La longueur de la conduite importe de moins en moins au fur et à mesure de l'apparition des défaillances.

Ceci peut se constater lorsque l'on examine les trois groupes de défaillances séparément. Le tableau 21 montre que le coefficient β_1 diminue lorsque l'ordre de la défaillance augmente. Le facteur "nombre de défaillances subies par la conduite" devient donc le facteur prépondérant et la longueur n'est plus un facteur essentiel. Ce n'est donc plus le tronçon entier qui est malade mais seulement un endroit précis.

Tableau 21 : Coefficient β_1 en fonction du groupe de défaillances.

Groupes de défaillances	β_1		
	LED-Bordeaux	CEB	Bordeaux + CEB
1ère défaillance	0,62	0,49	0,54
2 à 4ème défaillance	0,37	0,21	0,30
5ème déf. et plus	0,07	0,12	0,11

Ceci vient en fait de la difficulté à déterminer ce que peut être un individu dans ce type d'étude. Pour une étude épidémiologique, l'individu existe comme tel. Par contre pour cette étude un individu a été défini comme un tronçon de canalisations homogène suivants différents facteurs. En réalité un tronçon est ici un groupe d'individus, de tuyaux de 6m de longueur.

L'idéal serait d'utiliser les tuyaux de 6m. Là encore ceci est matériellement impossible, car on ne peut connaître mètre par mètre à quel endroit se trouvent les conduites et donc rattacher une défaillance à une conduite de 6 m.

Pour vérifier l'hypothèse du phénomène local, nous avons constitué artificiellement un échantillon de tronçons plus courts en divisant chaque tronçon en 11 parties égales. Cette façon de procéder a été adopté essentiellement pour des raisons pratiques. Puis l'hypothèse suivante a

été établie : chaque nouveau tronçon ne peut avoir plus d'une défaillance, ce qui ramène, cette fois, à l'étude de toutes les défaillances. Cette hypothèse peut se justifier du fait que la longueur moyenne des tronçons ainsi constitués est de 17 m (189 m pour les tronçons du fichier d'origine).

L'échantillon observé comporte cette fois 12551 individus, alors que dans l'échantillon de base il y en avait 1141 (échantillon sur les diamètres 100, 150, 200, 250 et 300 mm). Dans les deux cas, les modèles établis prennent en compte les mêmes facteurs significatifs. On trouve les valeurs concernant β_l dans le tableau 22 :

Tableau 22 : Valeurs du paramètre β_l de la longueur selon l'échantillon.

	Echantillon de base	Echantillon expérimental (longueur divisée par 11)
β_l	0,48	0,60

La valeur de β_l dans l'échantillon expérimental est supérieure à celle de l'échantillon de base. La longueur moyenne de l'échantillon est donc importante dans la détermination de ce facteur. De plus ceci montre que ce n'est pas la probabilité d'apparition de première défaillance qui est proportionnelle à la longueur du tronçon, mais plutôt l'apparition d'une défaillance quelle qu'elle soit.

Cependant la longueur moyenne des tronçons étudiés semble encore trop importante. Aussi pour mesurer l'effet de la longueur moyenne de l'échantillon sur l'estimation de β_l, les tronçons les plus longs ont été éliminés. Plusieurs modèles ont été établis pour des échantillons pour lesquels a été fixée une longueur maximale de tronçon. La figure 36 présente ces résultats.

On constate que l'hypothèse de proportionnalité de la longueur dans l'expression de la fonction de risque se vérifie pour des tronçons de l'ordre de 10 m maximum, ce qui est proche de la longueur d'un tuyau standard (6 m). L'hypothèse d'un phénomène local est la plus vraisemblable pour expliquer la valeur de β_l voisin de 0,6.

Les valeurs de β_l sont nettement supérieures pour les échantillons de longueur maximale 100 m, ce qui confirme bien la nécessité d'avoir, autant que possible, des tronçons de faible longueur.

En conclusion, si on veut obtenir une valeur de β_l proche de 1, il est nécessaire de faire les mesures sur des tronçons de 10 m maximum. Il faut également homogénéiser au mieux la longueur des tronçons afin que dans la prévision future, le paramètre β_l influe moins.

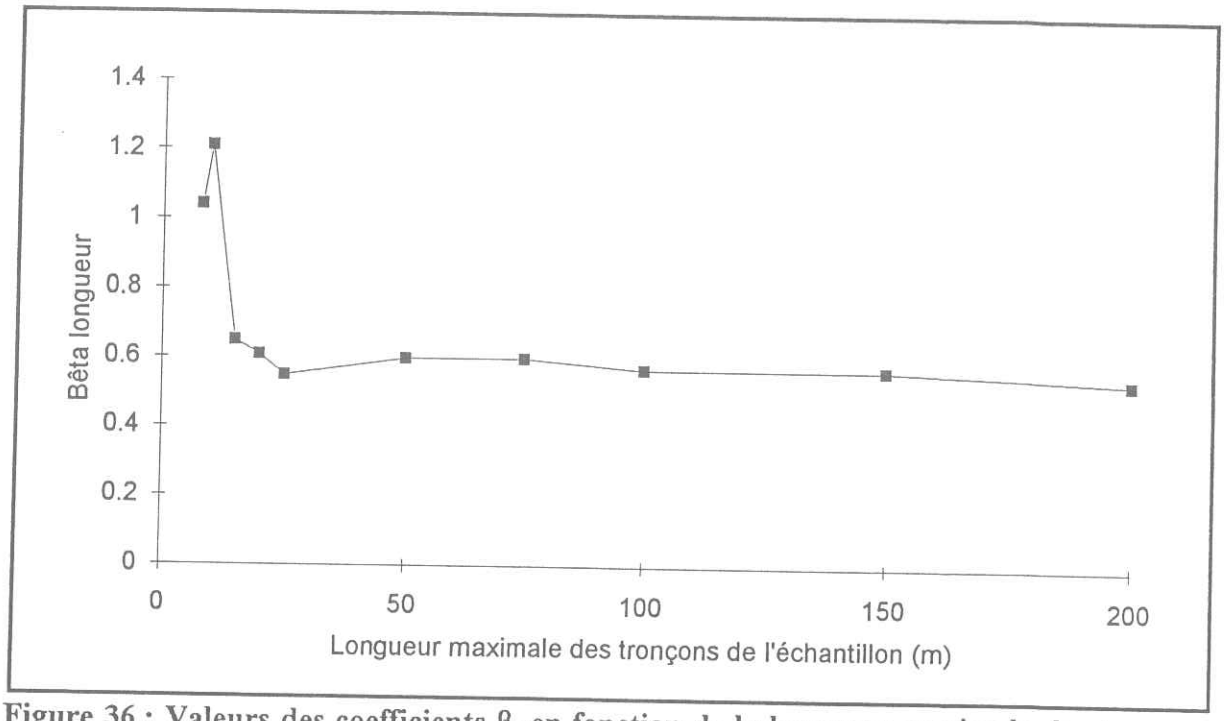


Figure 36 : Valeurs des coefficients β_l en fonction de la longueur maximale des tronçons de l'échantillon élargi.

Cependant une valeur de 0,5 n'est pas gênante et il suffira d'en tenir compte dans le calcul du risque de défaillance. Par exemple si on a un risque de 0,01 pour qu'il y ait une casse sur une conduite de 500m, sur une conduite de 1000 m de mêmes caractéristiques ce risque n'est pas de 0,02 mais de $0,01 \times \exp\left[0,5 * \ln\left(\frac{1000}{500}\right)\right] = 0,01 \times \sqrt{2} = 0,014$.

5.3.5. Conclusion

En conclusion de cette partie, l'étude du modèle de Cox a permis de tirer divers enseignements :

- les facteurs influents sur les casses des conduites sont :
  * le nombre de défaillances déjà subies par l'individu,
  * sa longueur,
  * son diamètre,

  et dans une moindre mesure,
  * le sol et sa corrosivité,
  * le facteur croisé trafic-chaussée,
  * le passage technologique de la fonte grise à la fonte ductile, ainsi que les conditions de pose (Ap75).
- ces facteurs ont une influence sensiblement identique dans les deux services urbains, ainsi que pour les deux syndicats ruraux, lorsque ces facteurs sont significatifs ;

- le risque de défaillance est proportionnel à la racine carrée de la longueur, ceci étant principalement dû au fait que les défaillances successives sont dépendantes les unes des autres.

- le modèle de Cox a l'avantage d'établir un minimum de contraintes dans l'estimation des paramètres (modèle semi-paramétrique), mais ne permet pas de faire des prévisions sur une durée supérieure à la durée d'observation des défaillances. Il renseigne cependant bien sur la forme des courbes de risque et de probabilités de non-défaillances.

Ce dernier point nous a conduit à élaborer un modèle complètement paramétrique, que nous présentons dans les paragraphes suivants.

5.4. APPLICATION DU MODELE DE WEIBULL

5.4.1. Choix et présentation du modèle

L'étude du modèle de Cox montre que les courbes de risque sont différentes par leur forme et par leur niveau de risque selon les défaillances. La première défaillance voit son risque augmenter avec le temps, tandis que pour les suivantes le risque de défaillances est beaucoup plus fort et semble moins dépendant du temps.

Parmi les modèles utilisables, c'est le modèle de Weibull qui a été choisi. Les raisons sont les suivantes :

- selon les valeurs de ces paramètres, la fonction de risque peut être soit croissante, soit décroissante, soit constante.

- sa formulation est simple. La fonction de risque peut se calculer le plus facilement à partir des paramètres λ et p, ainsi que la probabilité de non-défaillance.

- la forme de ces fonctions s'adapte bien à la forme des fonctions de risque du modèle de Cox comme on le verra par la suite.

La fonction de risque est de la forme suivante :

$$h(t) = \lambda p(\lambda t)^{p-1} \times \exp\left(\sum_i \beta_i Z_i\right),$$

λ et p étant les deux paramètres de la fonction et β_i et Z_i le paramètre et la valeur des variables étudiées, déjà utilisés dans le modèle Cox.

p est le paramètre de forme de la fonction.

H, la fonction de risque cumulée est égale à :

$$H(t) = (\lambda t)^p \exp\left(\sum_i \beta_i Z_i\right)$$

La fonction de non-défaillance est la suivante :

$$S(t) = \exp\left[-(\lambda t)^p \exp\left(\sum_i \beta_i Z_i\right)\right]$$

La fonction densité de la variable T est :

$$f(t) = \exp\left[-(\lambda t)^p \exp\left(\sum_i \beta_i Z_i\right)\right] \lambda p (\lambda t)^{p-1} \times \exp\left(\sum_i \beta_i Z_i\right)$$

Si l'on fait le changement de variable suivant : Y=logT, on a alors un modèle linéaire de la forme :

$$Y = \alpha + Z_i \beta_i^* + \sigma W,$$

où $\alpha = -\log \lambda$, $\sigma = p^{-1}$, $\beta_i^* = -\sigma \beta_i$ et W étant la variable d'erreur de densité : $f(w) = e^{(w - e^w)}$ [Kalbfleisch et Prentice, 1980].

C'est sous cette forme que le logiciel EGRET estime les différents paramètres en utilisant également la méthode du maximum de vraisemblance.

5.4.2. La stratification avec le modèle de Weibull

Comme pour le modèle de Cox, il est nécessaire de stratifier en distinguant différentes fonctions h(t) en fonction du groupe de défaillances. Cependant le logiciel EGRET ne permettant pas de pratiquer la méthode de stratification avec ce modèle, nous avons procédé de la manière suivante.

Dans un premier temps, les paramètres λ_j, p_j et β_{ij} sont estimés pour chacun des groupes j de défaillances. A l'issue de ce calcul, les vecteurs $\overline{\beta}_j$, vecteurs des paramètres estimés des facteurs influents, sont connus ainsi que les matrices de variance-covariance associées $S_j = T_j^2$. On recherche alors le vecteur β des paramètres β_i tel que $\sum_j \left\langle T_j^{-1}(\beta - \overline{\beta}_j), T_j^{-1}(\beta - \overline{\beta}_j) \right\rangle$ soit minimum, "$\langle\ \rangle$" correspondant au produit scalaire.

Cette méthode est utilisée car elle prend en compte la variance des différents échantillons, et indirectement leur taille. Cela revient donc à chercher le vecteur β au sens de la méthode des moindres carrés pondérés par T_j. Si l'on dérive par rapport à β on a alors :

$$\sum_j S_j^{-1}(\beta - \overline{\beta}_j) = 0$$

d'où :

$$\left(\sum_j S_j^{-1}\right)\beta = \sum_j \left(S_j^{-1}\overline{\beta}_j\right).$$

On a donc :

$$\beta = \left[\sum_j S_j^{-1}\right]^{-1} \sum_j \left(S_j^{-1}\overline{\beta}_j\right)$$

Cette méthode a été appliquée à l'aide du tableur Excel 4.

Dans le cas présent, comme on a trois groupes de défaillances on a :

$$\beta = \left[S_1^{-1} + S_2^{-1} + S_3^{-1}\right]^{-1}\left[S_1^{-1}\overline{\beta}_1 + S_2^{-1}\overline{\beta}_2 + S_3^{-1}\overline{\beta}_3\right].$$

Les résultats sont donnés dans le paragraphe suivant.

5.4.3. Résultats

5.4.3.1. Estimation des paramètres β_i

Les paramètres β_i du vecteur β estimés selon la méthode du paragraphe précédent sont les paramètres définis comme significatifs sur le modèle de Cox. Là encore, comme il y a stratification, on fait l'hypothèse que les facteurs influent de la même façon quelle que soit la défaillance. Le tableau 23 représente les paramètres du modèle de Weibull calculés selon la méthode du paragraphe précédent.

Tableau 23 : Paramètres β_i du modèle de Weibull avec stratification (le chiffre entre parenthèse correspond à l'écart type)

	β_i					
Facteurs	LED-Bordeaux		CEB		Bas-Rhin	
Longueur	**0,57**	(0,05)	**0,37**	(0,05)	**0,49**	(0,09)
Diamètre	**-0,85**	(0,11)	**-0,34**	(0,10)	**-0,39**	(0,26)
Sol	**0,34**	(0,07)	**0,29**	(0,09)		
Trafic-Chaussée	**0,47**	(0,09)	**0,57**	(0,17)	**-0,45**	(0,20)
					(Trafic seul)	
Nature cana	**-0,13**	(0,09)	**-0,64**	(0,14)		

Le paramètre "Ap75" ne peut être pris en compte dans ce cas car les conduites ayant eu plus de cinq défaillances ont toutes été posées avant 1975. Il n'était donc pas possible de mesurer son influence.

Comparées au modèle de Cox, les valeurs des paramètres sont proches d'un modèle à l'autre, excepté sur Bordeaux où les facteurs "Trafic-Chaussée" et "Nature cana." sont différents. C'est ici dû à l'élimination du facteur "Ap75".

5.4.3.2. Estimation des paramètres de la fonction de risque de base

Les paramètres de la fonction de risque de base sont dans le tableau 24.

Tableau 24 : Paramètres des fonctions de Weibull et exponentiel selon les services et les groupes de défaillances

Groupe de défaillance		LED-Bordeaux	CEB	Bas-Rhin
1ère défaillance	λ_1	0,03	0,02	0,03
	p_1	1,67	1,21	1,27
2 à 4 ème déf.	λ_2	0,15	0,10	0,07
	p_2	1,06	0,99	1,08
5ème déf. et +	λ_3	0,15	0,20	
(exponentiel)	p_3	1	1	

Concernant le troisième groupe de défaillance, c'est le modèle exponentiel qui a été choisi (p=1). Ceci signifie que, à partir de la 5ème défaillance, elles apparaissent à un taux constant. Notons que pour le deuxième groupe de défaillance, le paramètre p est également proche de 1. Il sera intéressant à l'avenir de regrouper les groupes 2 et 3 de défaillances.

Les différents paramètres λ et p sont assez différents d'un service à l'autre. Le paramètre d'intensité λ croît avec le groupe de défaillance. Le risque de base augmente donc avec le nombre de défaillances. Sur Bordeaux ceci est moins évident lorsque l'on compare les groupes 2 et 3 de défaillances. Le taux de casse est sensiblement le même.

Les courbes des figures 37, 38 et 39 montrent les fonctions de risque de base pour les différents services.

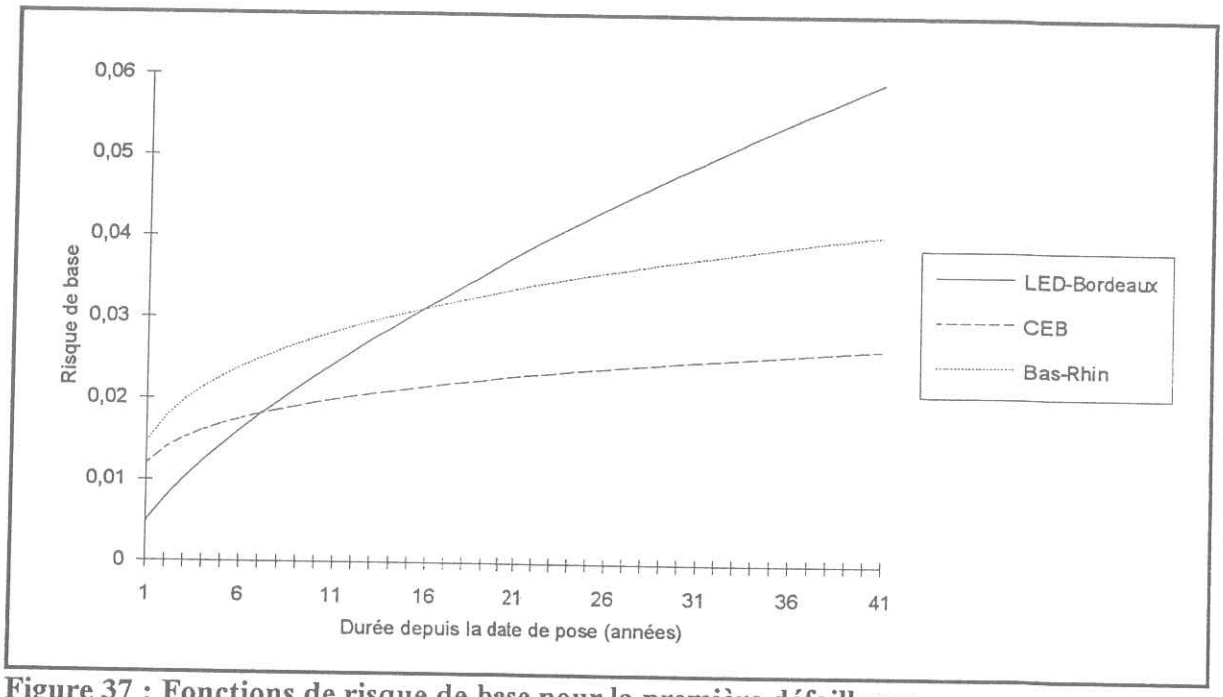


Figure 37 : Fonctions de risque de base pour la première défaillance.

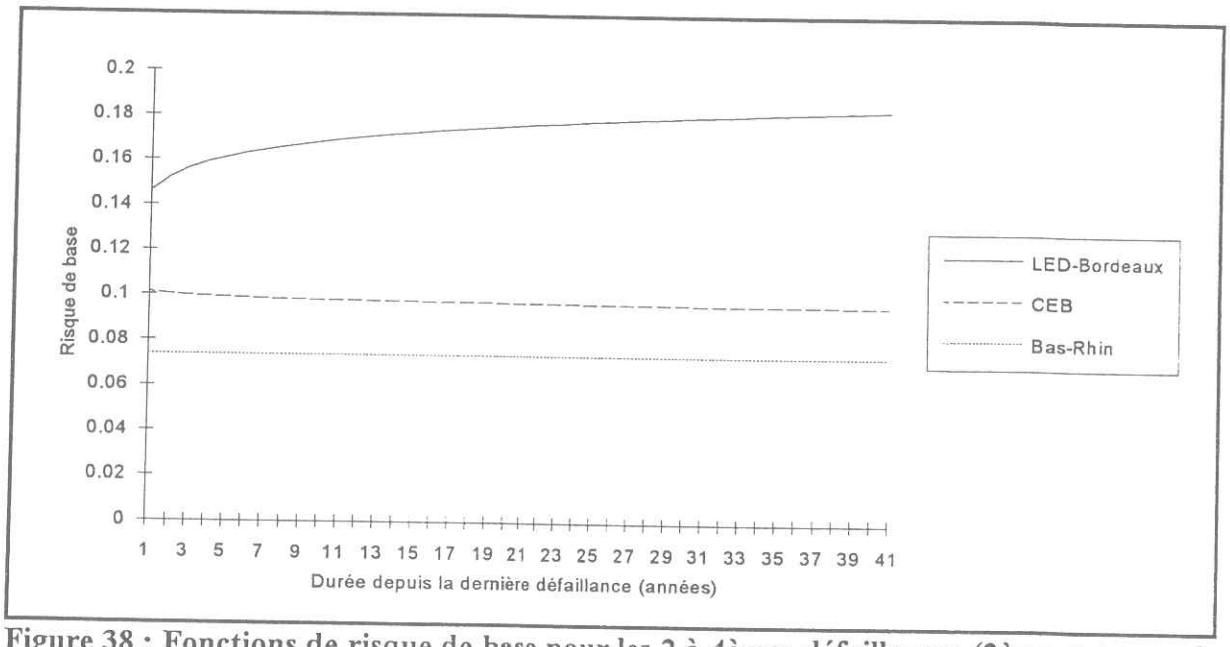


Figure 38 : Fonctions de risque de base pour les 2 à 4èmes défaillances (2ème et + pour le Bas-Rhin).

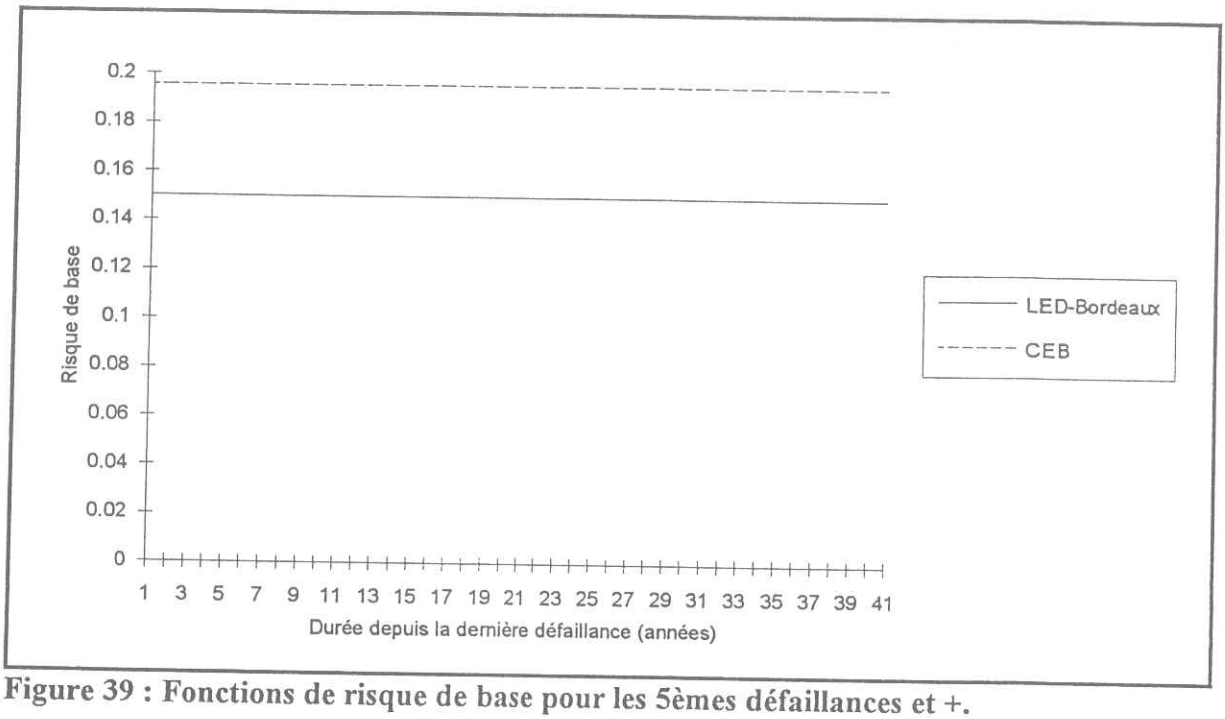


Figure 39 : Fonctions de risque de base pour les 5èmes défaillances et +.

Les fonctions de risque du modèle de Weibull sont donc les suivantes :

Sur Bordeaux :

Première défaillance :

$$h(t) = 0{,}0048\, t^{0{,}67} \times \exp\left[0{,}57\ln\left(\frac{longueur}{500}\right) - 0{,}85\ln\left(\frac{diamètre}{60}\right) + 0{,}34 \times (sol) + 0{,}47 \times (trafic - chaussée) - 0{,}13 \times (année\,de\,pose66)\right]$$

Deuxième à quatrième défaillance :

$$h(t) = 0{,}1419\, t^{0{,}06} \times \exp\left[0{,}57\ln\left(\frac{longueur}{500}\right) - 0{,}85\ln\left(\frac{diamètre}{60}\right) + 0{,}34 \times (sol) + 0{,}47 \times (trafic - chaussée) - 0{,}13 \times (année\,de\,pose66)\right]$$

Cinquième défaillance et + :

$$h(t) = 0{,}15 \times \exp\left[0{,}57\ln\left(\frac{longueur}{500}\right) - 0{,}85\ln\left(\frac{diamètre}{60}\right) + 0{,}34 \times (sol) + 0{,}47 \times (trafic - chaussée) - 0{,}13 \times (année\,de\,pose66)\right]$$

Sur la CEB :

Première défaillance :

$$h(t) = 0{,}0106\, t^{0{,}21} \times \exp\left[0{,}37\ln\left(\frac{longueur}{500}\right) - 0{,}34\ln\left(\frac{diamètre}{60}\right) + 0{,}29 \times (sol) + 0{,}57 \times (trafic - chaussée) - 0{,}64 \times (année\,de\,pose66)\right]$$

Deux à quatrième défaillance :

$$h(t) = 0{,}1013\, t^{-0{,}01} \times \exp\left[0{,}37\ln\left(\frac{longueur}{500}\right) - 0{,}34\ln\left(\frac{diamètre}{60}\right) + 0{,}29 \times (sol) + 0{,}57 \times (trafic - chaussée) - 0{,}64 \times (année\,de\,pose66)\right]$$

Cinquième défaillance et + :

$$h(t) = 0{,}20 \times \exp\left[0{,}37\ln\left(\frac{longueur}{500}\right) - 0{,}34\ln\left(\frac{diamètre}{60}\right) + 0{,}29 \times (sol) + 0{,}57 \times (trafic - chaussée) - 0{,}64 \times (année\,de\,pose66)\right]$$

Sur le Kochersberg et le Kronthal :

Première défaillance :

$$h(t) = 0{,}0148\ t^{0.27} \times \exp\left[0{,}49 \ln\left(\frac{longueur}{500}\right) - 0{,}39 \ln\left(\frac{diamètre}{60}\right) - 0{,}45 \times (trafic)\right]$$

Deuxième défaillance et + :

$$h(t) = 0{,}07 \times \exp\left[0{,}49 \ln\left(\frac{longueur}{500}\right) - 0{,}39 \ln\left(\frac{diamètre}{60}\right) - 0{,}45 \times (trafic)\right]$$

5.5. CONCLUSIONS SUR L'APPLICATION DES MODELES

5.5.1. Avantages et inconvénients des modèles de Cox et de Weibull

Comme on l'a vu, le modèle de Cox dépend de la durée d'observation des données. Il ne permet donc pas de faire des prévisions à plus long terme que la période d'observation. Pour la première défaillance ceci n'est pas gênant, car sur celle-ci la durée d'observation est suffisamment longue. Par contre sur les défaillances suivantes, la durée d'observation est de 10 à 15 ans en moyenne. Au contraire le modèle de Weibull permet de faire des prévisions à plus long terme, car il est sous forme analytique.

Le fait qu'il soit sous cette forme facilite également son utilisation. Le modèle de Cox est lui plus difficile à utiliser car la fonction de risque de base est sous forme de valeurs discrètes données à chacun des pas de temps.

Le principal avantage du modèle de Cox est qu'il n'établit qu'un minimum de contraintes dans l'estimation des paramètres β_i, qui sont alors plus précis. Cependant les paramètres estimés à partir du modèle de Weibull sont assez proches de ceux estimés par Cox.

Pour ces raisons, le modèle de Weibull semble plus opérationnel dans la prévision des défaillances sur un réseau. Cependant le modèle de Cox sera utilisé pour mettre en évidence les facteurs significatifs estimés pour le modèle de Weibull.

5.5.2. Pistes pour l'utilisation future des modèles

La comparaison des résultats selon les services permet de donner plusieurs tendances sur l'utilisation future de ces modèles dans des services d'eau potable(Cf figure 40).

Tout d'abord les facteurs influents sont les mêmes, exceptés sur les deux syndicats du Bas-Rhin, pour lesquels l'historique est plus petit. Les valeurs des paramètres β_i sont également assez proches les uns des autres. On pourra donc à l'avenir, dans un service sur lequel on a peu de données environnementales, affecter ces paramètres lors de la création d'un modèle de prévision des défaillances.

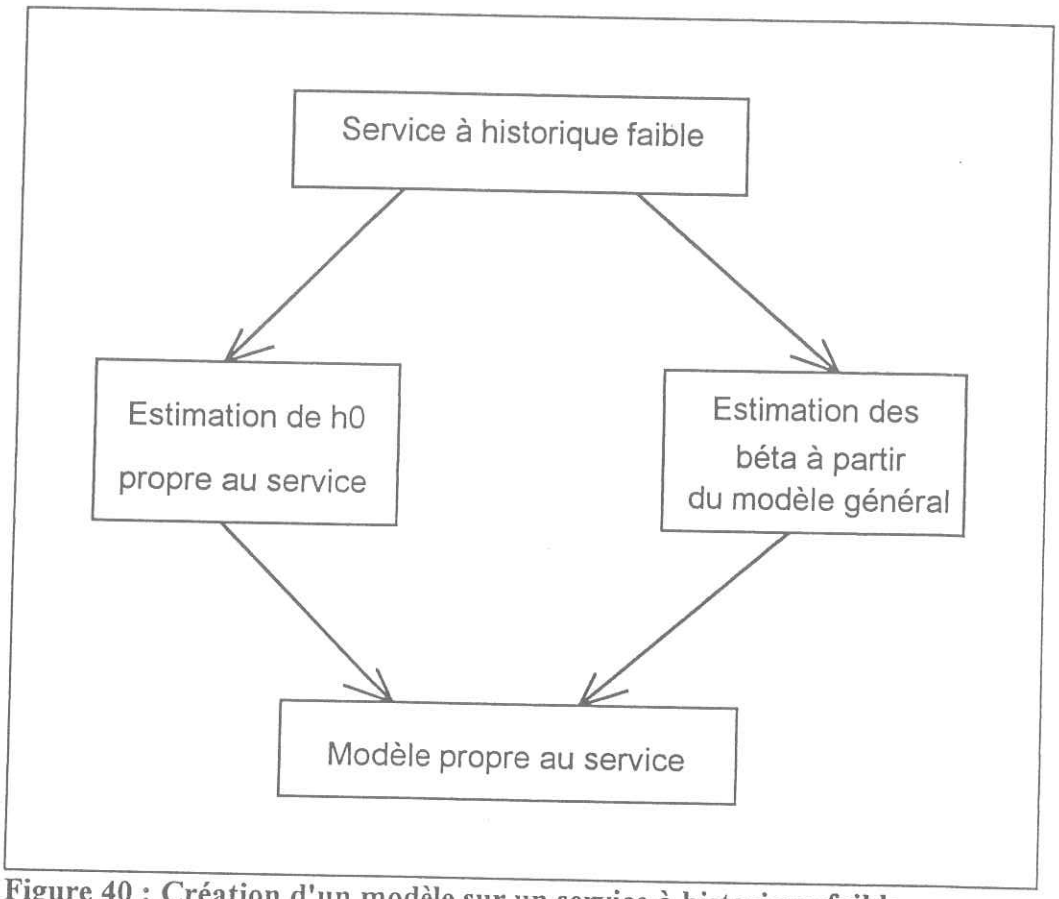


Figure 40 : Création d'un modèle sur un service à historique faible.

Par contre on a vu que les formes des courbes de risques sont bien différentes les unes des autres. La fonction h_0 exprime donc la spécificité du service (climat, service urbain-service rural, réseau âgé ou non) dans le modèle. Il est donc important de pouvoir estimer cette fonction.

La méthode utilisée pour ceci est la méthode appelée méthode d'apprentissage présentée dans le chapitre 7.

CHAPITRE 6 :
ADEQUATION DU MODELE

L'estimation du modèle dans le chapitre précédent ne donne aucune information sur la prévision du nombre de casses sur chacune des conduites ou sur un réseau complet. Il est donc important de comparer le nombre de défaillances réelles obtenu au nombre théorique de défaillances estimé à partir du modèle. Nous verrons dans les paragraphes suivants comment le nombre de défaillances théorique est estimé, puis les résultats nous montreront quelle correction peut être appliquée au modèle afin de mieux le caler sur les données réelles. Enfin nous verrons dans quelle mesure l'oubli d'un facteur estimé comme influent peut faire varier l'espérance du nombre de défaillances.

6.1. CALCUL DE L'ESPERANCE DU NOMBRE DE DEFAILLANCES

Le nombre théorique de défaillances est estimé à partir de l'espérance de défaillances. La formule générale de l'espérance mathématique est :

$$E(N(t)) = \sum_{k=0}^{\infty} k\, P(N(t) = k),$$

P(N(t)=k) étant la probabilité pour que k défaillances aient eu lieu au temps t.

Cette probabilité est égale à :

$$P(N(t) = k) = \int_0^t f_1(t_1) \int_{t_1}^t f_2(t_2 - t_1) \ldots\ldots \int_{t_{k-2}}^t f_{k-1}(t_{k-1} - t_{k-2}) \int_{t_{k-1}}^t f_k(t_k - t_{k-1}) S_{k+1}(t - t_k) dt_k dt_{k-1} dt_2 dt_1 .$$

où f_i est la fonction de densité de l'apparition de la i-ième défaillance et S_i la probabilité de non-défaillance concernant la i-ième défaillance.

La démonstration se trouve en annexe 6. Les t_i correspondent au temps d'apparition des différentes défaillances (Cf. fig. 41).

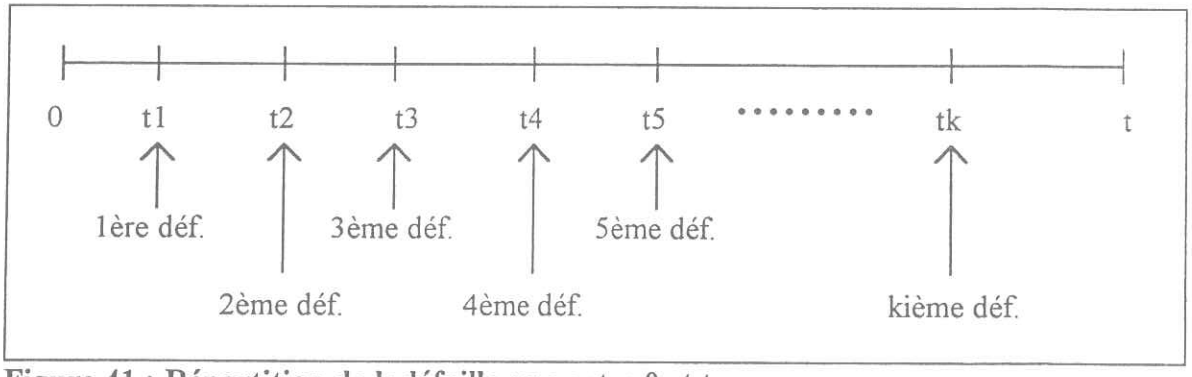


Figure 41 : Répartition de k défaillances entre 0 et t.

L'espérance sera donc égale à :

$$E(N(t)) = \sum_{k=0}^{\infty} k \left[\int_0^t f_1(t_1) \int_{t_1}^t f_2(t_2 - t_1) \ldots\ldots \int_{t_{k-2}}^t f_{k-1}(t_{k-1} - t_{k-2}) \int_{t_{k-1}}^t f_k(t_k - t_{k-1}) S_k(t - t_k) dt_k dt_{k-1} \ldots\ldots dt_2 dt_1 \right]$$

Si le modèle utilisé est un modèle de Cox (pas de temps discrets), cette espérance devient :

$$E(N(t)) = \sum_{k=0}^{\infty} k \left[\sum_{t_1=0}^{t} f_1(t_1) \sum_{t_2=t_1}^{t} f_2(t_2 - t_1) \ldots\ldots \sum_{t_{k-1}=t_{k-2}}^{t} f_{k-1}(t_{k-1} - t_{k-2}) \sum_{t_k=t_{k-1}}^{t} f_k(t_k - t_{k-1}) S_{k+1}(t - t_k) \right]$$

Si l'on utilise un modèle de Weibull pour la 1ère défaillance, puis pour les 2 à 4ème et ensuite un modèle exponentiel à partir de la 5ème, la partie de l'intégrale concernant le modèle exponentiel pourra être simplifiée. Elle sera égale à :

$$A_{\exp} = \lambda_3^{*(k-4)} \exp\left[-\lambda_3^*(t - t_4)\right] \frac{(t - t_4)^{k-4}}{(k-4)!}$$

λ_3^* étant le risque de défaillances du modèle exponentiel.

Si l'on partage l'espérance de la manière suivante :

$$E(N(t)) = \sum_{k=1}^{4} kP(N(t) = k) + \sum_{k=5}^{\infty} kP(N(t) = k),$$

on obtient alors :

$$E(N(t)) = \sum_{k=1}^{4} kP(N(t) = k)$$
$$+ \int_0^t \int_{t_1}^t \ldots \int_{t_3}^t f_1^*(t_1) f_2^*(t_2 - t_1) f_2^*(t_3 - t_2) f_2^*(t_4 - t_3) \left[4 \times \left(1 - \exp\left[-\lambda_3^*(t - t_4)\right]\right) + \lambda_3^*(t - t_4) \right] dt_4 dt_3 dt_2 dt_1$$

f_i^* étant la fonction de densité du i-ième groupe de défaillances (Cf. annexe 6).

C'est cette formule qui a été utilisée dans le calcul de l'espérance. Cette espérance correspond à une conduite étudiée depuis sa date de pose, n'ayant donc jamais eu de défaillance.

Supposons maintenant que l'on connaisse de la conduite une partie de son passé, à savoir le nombre de défaillances et leur date d'apparition jusqu'à un temps t_d, date de début d'observation (Cf. figure 42). L'espérance à calculer sera différente de la précédente, puisqu'elle tiendra compte de la date de dernière défaillance ainsi que de la durée depuis celle ci.

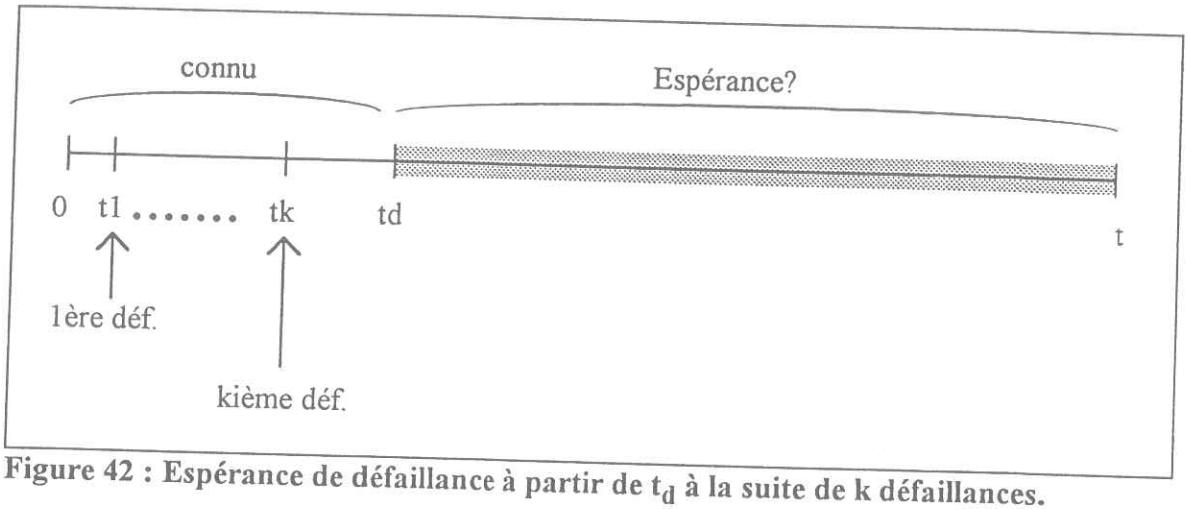


Figure 42 : Espérance de défaillance à partir de t_d à la suite de k défaillances.

On aura alors 5 espérances possibles selon le nombre de défaillances subies(Cf. annexe 6). Si aucune défaillance n'a eu lieu, l'espérance sera, par exemple, de :

$$E(N(t)) = \sum_{k=1}^{4} kP(N(t)=k)$$

$$+\int_{t_d}^{t}\int_{t_1}^{t}\cdots\int_{t_3}^{t}\frac{f_1^*(t_1)}{S_1^*(t_d)}f_2^*(t_2-t_1)f_2^*(t_3-t_2)f_2^*(t_4-t_3)\left[4\times\left(1-\exp\left[-\lambda_3^*(t-t_4)\right]\right)+\lambda_3^*(t-t_4)\right]dt_4dt_3dt_2dt_1$$

Les autres espérances se trouvent en annexe 6. Dans les deux cas l'espérance a été calculée à partir d'un programme élaboré en langage Turbo-Pascal.

Nous allons voir maintenant comment ces espérances ont été utilisées.

6.2. COMPARAISON AVEC LES DONNEES REELLES

La validation a eu lieu sur le réseau de la LED-Bordeaux en deux étapes. Tout d'abord ont été comparées sur la période 1951-1990 les défaillances réelles et les défaillances que l'on aurait pu prévoir si l'on avait eu le modèle au début de l'observation en 1951.

Ensuite ont été comparées les défaillances réelles ayant eu lieu entre 1983 et 1992 aux défaillances prévues à partir d'un modèle établi sur les données entre le début de l'observation (1951) et 1982. Une autre comparaison a eu lieu sur la période 1991-1992.

6.2.1. Comparaison des défaillances prévues et réelles sur 40 ans

Pour chacun des tronçons posés avant 1990 sur Bordeaux, ont été comparés le nombre de défaillances apparues réellement depuis leur date de pose et le nombre de défaillances prévues sur la même période calculé à partir de l'espérance (Cf. Figure 43). Cette espérance a été calculée à partir du modèle de Weibull stratifié provenant des données de Bordeaux de 1951 à 1990.

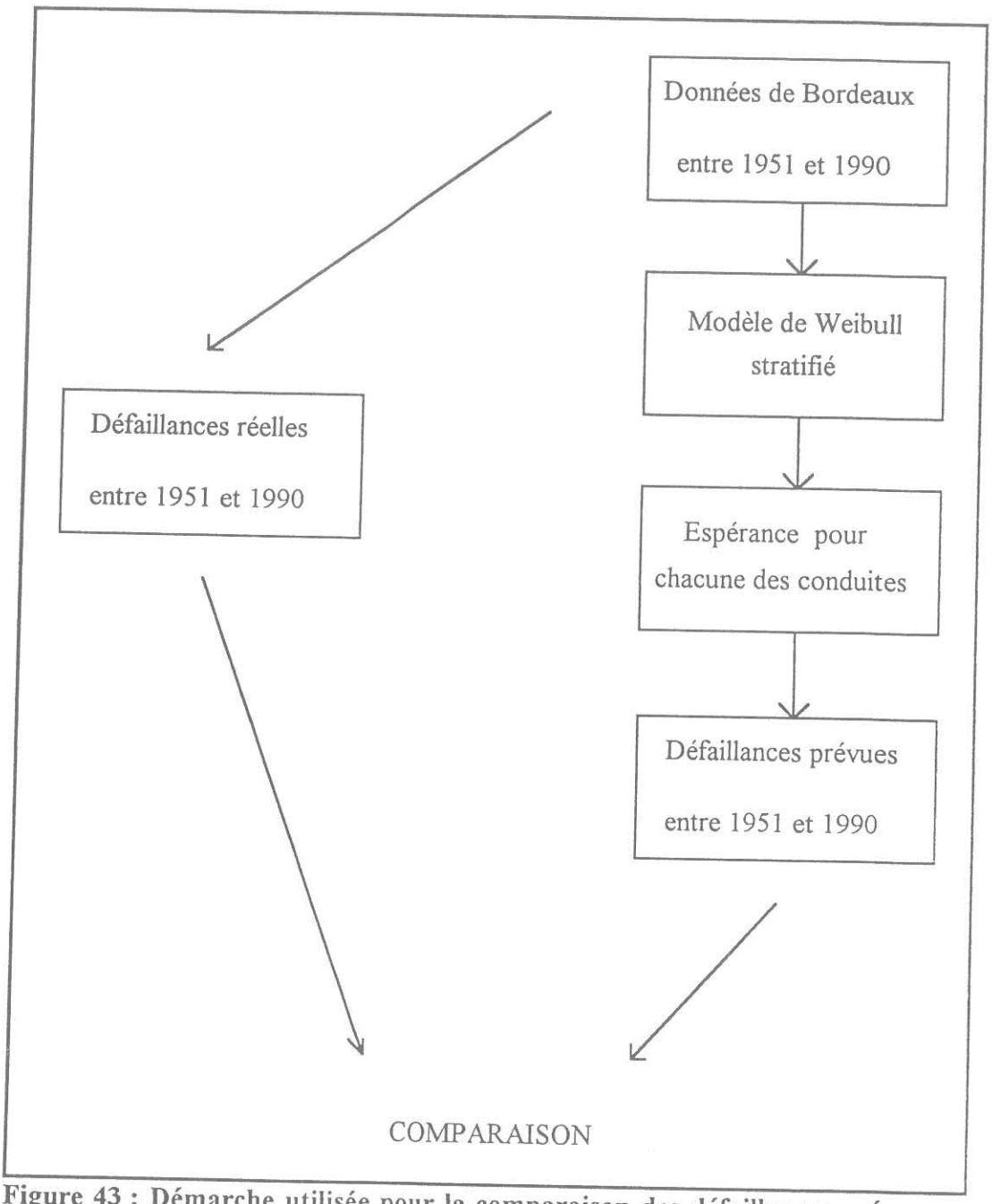


Figure 43 : Démarche utilisée pour la comparaison des défaillances prévues et réelles

Le nombre de défaillances prévues est alors de 434, pour 522 réelles sur la totalité de l'échantillon. Si l'on prend la méthode du chapitre 5 pour estimer l'intervalle de confiance à 95 % du nombre de défaillances prévues sur l'échantillon, les bornes de cet intervalle sont 346 et 546 défaillances.

Une autre comparaison peut se faire en classant les conduites par groupe de 30 par ordre décroissant des défaillances prévues (premier groupe : les trente plus défaillantes, deuxième groupe : les trente suivantes, etc...). Pour chacun des groupes est calculé le nombre moyen de défaillances prévues comparé au nombre moyen de défaillances réelles.

L'histogramme de la figure 44 représente cette comparaison.

Pour les premières classes, le nombre de défaillances est sous-estimé. C'est dans ces classes que l'erreur est la plus importante. Cependant les classes de défaillances prévues décroissent de la même manière que les classes de défaillances réelles. Cette erreur peut provenir des hypothèses que l'on a pu faire dans la construction du modèle :

- soit l'hypothèse de même influence quelle que soit la défaillance est trop forte,
- soit il existe un facteur que l'on a pu oublier.

6.2.1.1. Retour sur l'hypothèse de même influence quelle que soit la défaillance

Pour tester la première hypothèse, on calcule l'espérance à partir de trois modèles propres à chacun des groupes de défaillance. Pour chacun des groupes de défaillances on estime donc les facteurs influents et leurs paramètres β_i, ainsi que les paramètres λ et p du modèle de Weibull. Ces valeurs sont représentées dans le tableau 25.

Tableau 25 : Paramètres β_i estimés propres à chaque groupe de défaillance (signifie que le facteur n'est pas significatif).**

	1ère défaillance	2 à 4ème défaillance	5ème défaillance et +
λ	0,03	0,15	0,14
p	1,65	1,06	1
Longueur	0,74	0,59	**
Diamètre	-1,01	-1,01	-0,51
Sol	0,51	0,21	0,56
Trafic-chaussée	0,84	**	**
Nature cana	-0,34	**	0,79

Le nombre de défaillances prévues est alors de 378, ce qui est plus faible que précédemment. De plus l'histogramme par classes de défaillances décroissantes n'est pas amélioré (Cf. Fig. 45).

Dans ce cas, ce n'est donc pas l'hypothèse de même influence quelle que soit le groupe de défaillance qui entraîne un biais aussi important.

6.2.1.2. Prise en compte des défaillances juvéniles

Si l'on examine les conduites qui ont eu des défaillances alors qu'elles ont été classées comme faiblement défaillantes, on peut constater que ce sont surtout les casses qui ont eu lieu après un certain âge qui ont une réelle influence sur les casses suivantes. Des casses ("défaillances juvéniles") ayant eu lieu en début de vie de la conduite ne seraient donc qu'un "péché de jeunesse" et ne traduiraient pas un mauvais état réel.

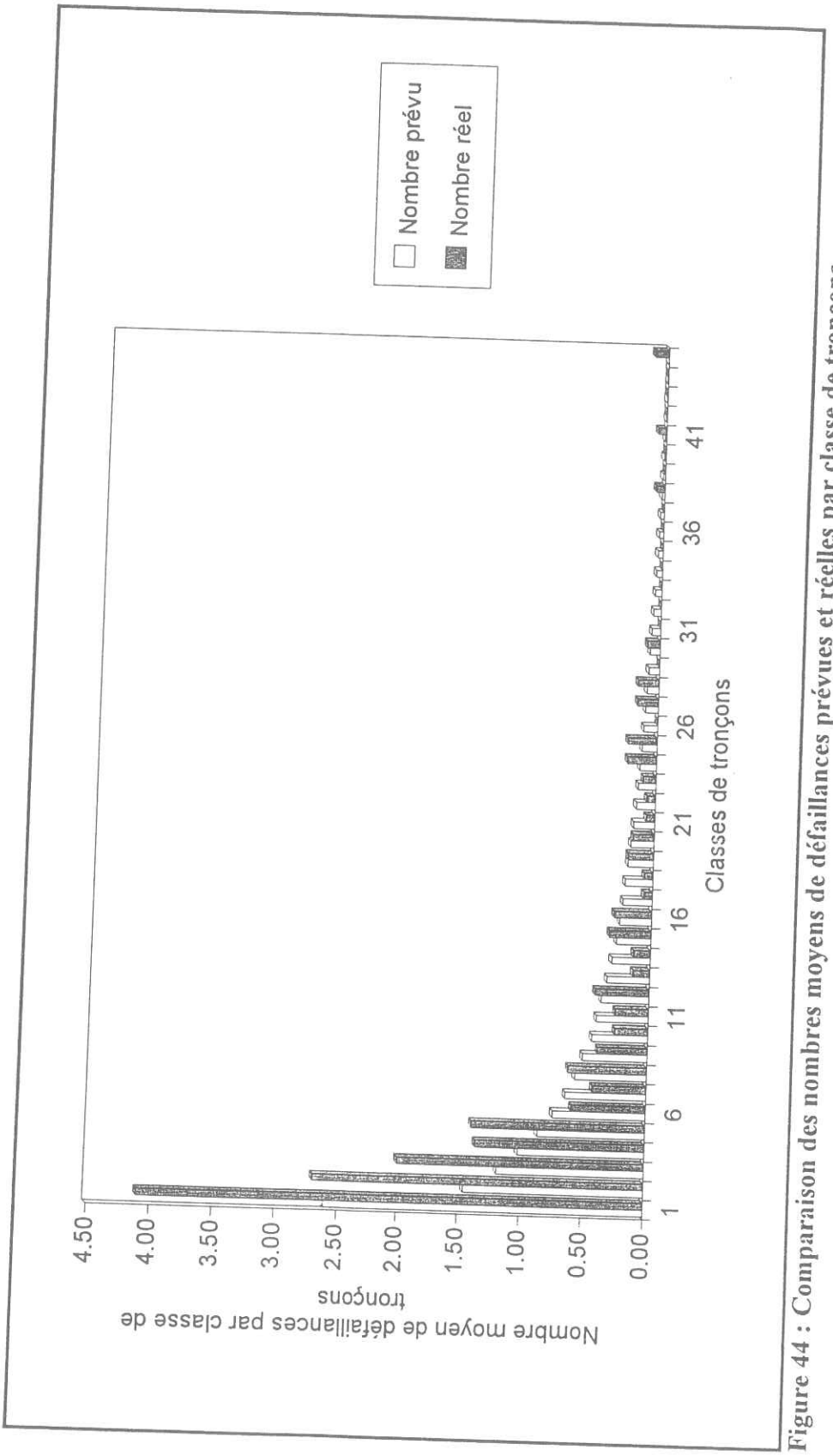


Figure 44 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons.

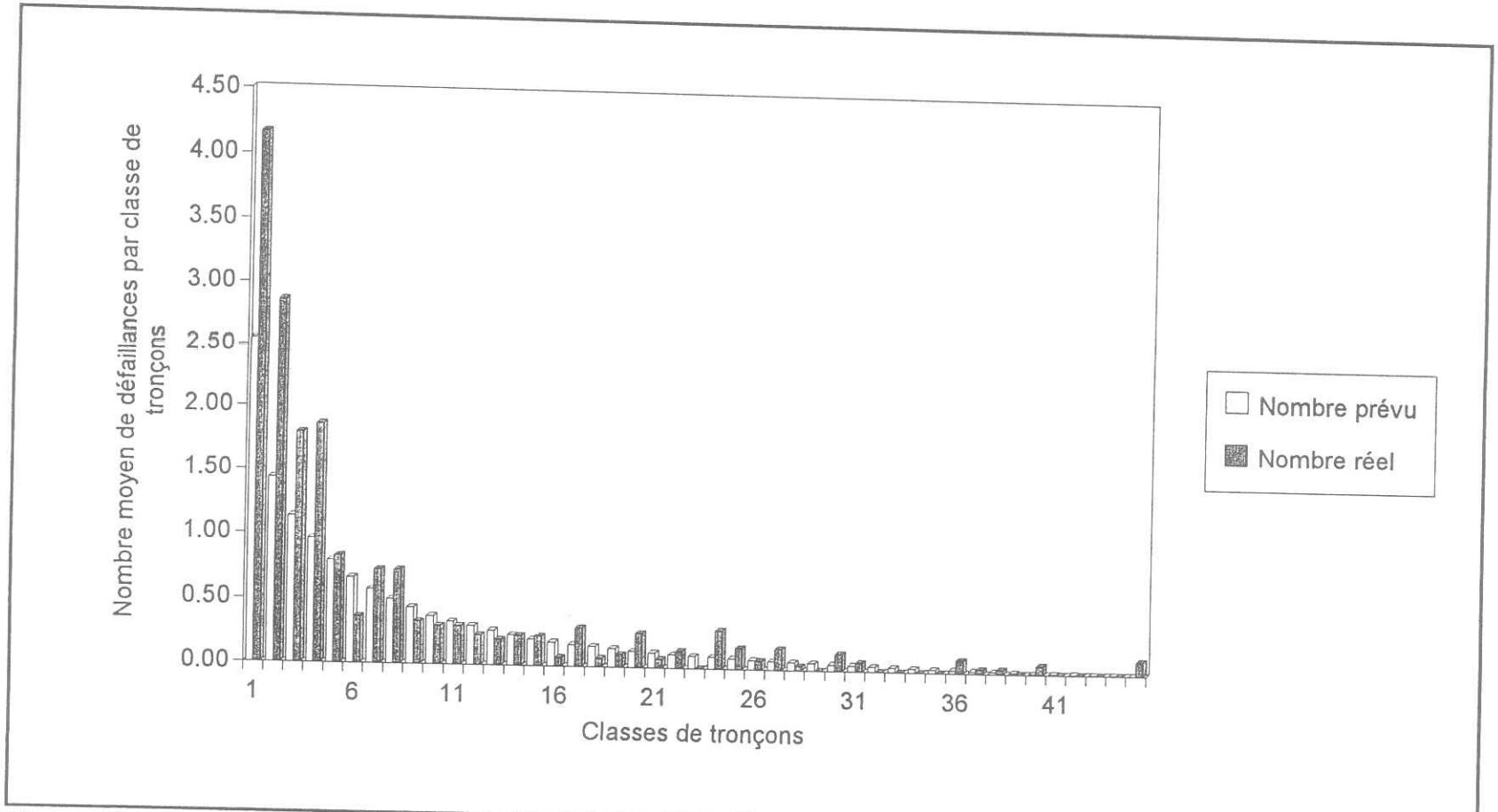


Figure 45 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (Influence différente des facteurs selon le groupe de défaillance).

L'introduction d'un paramètre supplémentaire t_0, à caler, dans la fonction de Weibull permettrait de ne considérer que les défaillances ayant eu lieu à partir de cette date t_0. L'étude ne prendrait en compte que les défaillances correspondant à une réelle dégradation dans le temps de la conduite. Le modèle de Weibull serait alors de la forme suivante :

$$h(t) = \lambda p\left[\lambda(t - t_0)\right]^{p-1} \times \exp\left(\sum_i \beta_i Z_i\right)$$

Il serait alors nécessaire de prendre construire une nouvelle vraisemblance permettant de définir les paramètres des facteurs influents. Le facteur t_0 correspond en fait au paramètre c du modèle de Herz présenté au chapitre 8 (§ 8.3.2.2.).

Une deuxième méthode serait d'introduire a posteriori un nouveau facteur permettant de définir si la défaillance qui a eu lieu est une défaillance juvénile. Ce facteur exprimerait pour la conduite le risque de rentrer dans un état de vieillissement accéléré.

C'est cette deuxième méthode qui a été choisie. Le facteur "vieillissement" établi est défini de la manière suivante. Il est égal à 0 lorsque la conduite n'a pas eu de défaillance après 15 ans, et il est égal à 1 si la conduite a eu une défaillance après 15 ans.

Le tableau 26 représente les paramètres β_i calculés en incluant ce nouveau facteur. Le paramètre du facteur "vieillissement" est le paramètre le plus fort (2,42) ce qui montre sa grande influence. On peut remarquer que le facteur "Trafic-Chaussée" n'est alors plus influent. De plus le paramètre du facteur "Nature Cana" a un signe opposé au paramètre estimé au chapitre 5. Ceci s'explique par la forte influence du facteur vieillissement qui concerne principalement les conduites posées avant 1966. Enfin les paramètres du modèle de Weibull (λ et p) sont complètement différents, encore une fois du fait de la forte influence du facteur "vieillissement".

Tableau 26 : Paramètres du modèle de Weibull stratifié avec le facteur correctif "Vieillissement".

1ère défaillance	λ	0,009
	p	1,856
2 à 4ème défaillance	λ	0,037
	p	1,161
5ème défaillance et +	λ	0,061
	p	1
β	Longueur	0,36
	Diamètre	-0,52
	Sol	0,29
	Nature cana	0,61
	"Vieillissement"	2,42

Le nombre de défaillances prévues est alors de 447, ce qui est meilleur. D'autre part l'histogramme des défaillances prévues est beaucoup plus proche des défaillances réelles.(Cf. Figure 46)

Si l'on regarde également les 90 premières conduites prévues comme ayant le plus de défaillances, on peut remarquer que le facteur "Vieillissement" apporte également de meilleurs résultats. En effet si l'on n'en tient pas compte, beaucoup de conduites prévues comme cassantes n'ont eu aucune défaillance (Cf. Fig. 47). Par contre si l'on en tient compte, uniquement deux tronçons n'ont pas de défaillance sur les 90 premiers (Cf. Fig. 48).

Le facteur "vieillissement" tel qu'on l'a pris ici a donc une influence importante sur l'apparition des défaillances. Il correspond en fait à une dégradation accélérée provoquée par un facteur de risque non-pris en compte dans le modèle(coups de bélier, choc au moment de la pose, etc.)

Dans cette partie ce facteur a été pris en compte a posteriori et on ne peut bien sûr pas prévoir sa valeur au moment de la pose. Cependant si l'on veut prévoir les défaillances sur une durée déterminée pour des conduites dont on connaît l'historique, il est possible d'en tenir compte. C'est ce que nous allons voir dans le paragraphe 6.2.2..

6.2.2. Comparaison des défaillances réelles à une prévision sur 10 et 2 ans

6.2.2.1. Prévision sur dix ans

Un modèle a été élaboré à partir des données de 1951 à 1982. Puis ont été comparées les défaillances prévues entre 1983 et 1992 à celles qui ont eu lieu réellement pendant cette période (Cf. fig. 49). La comparaison a eu lieu de la même manière que dans le paragraphe précédent.

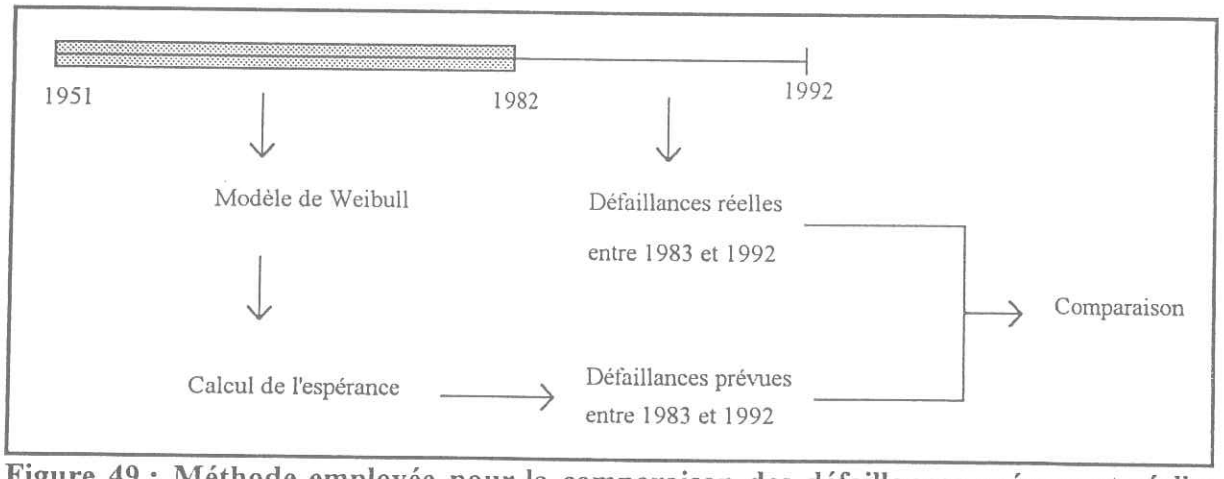


Figure 49 : Méthode employée pour la comparaison des défaillances prévues et réelles entre 1983 et 1992

Les coefficients estimés du modèle de Weibull sur la période 1951-1982 sont un peu différents de ceux estimés sur 1951-1990. Ils se trouvent dans le tableau 27.

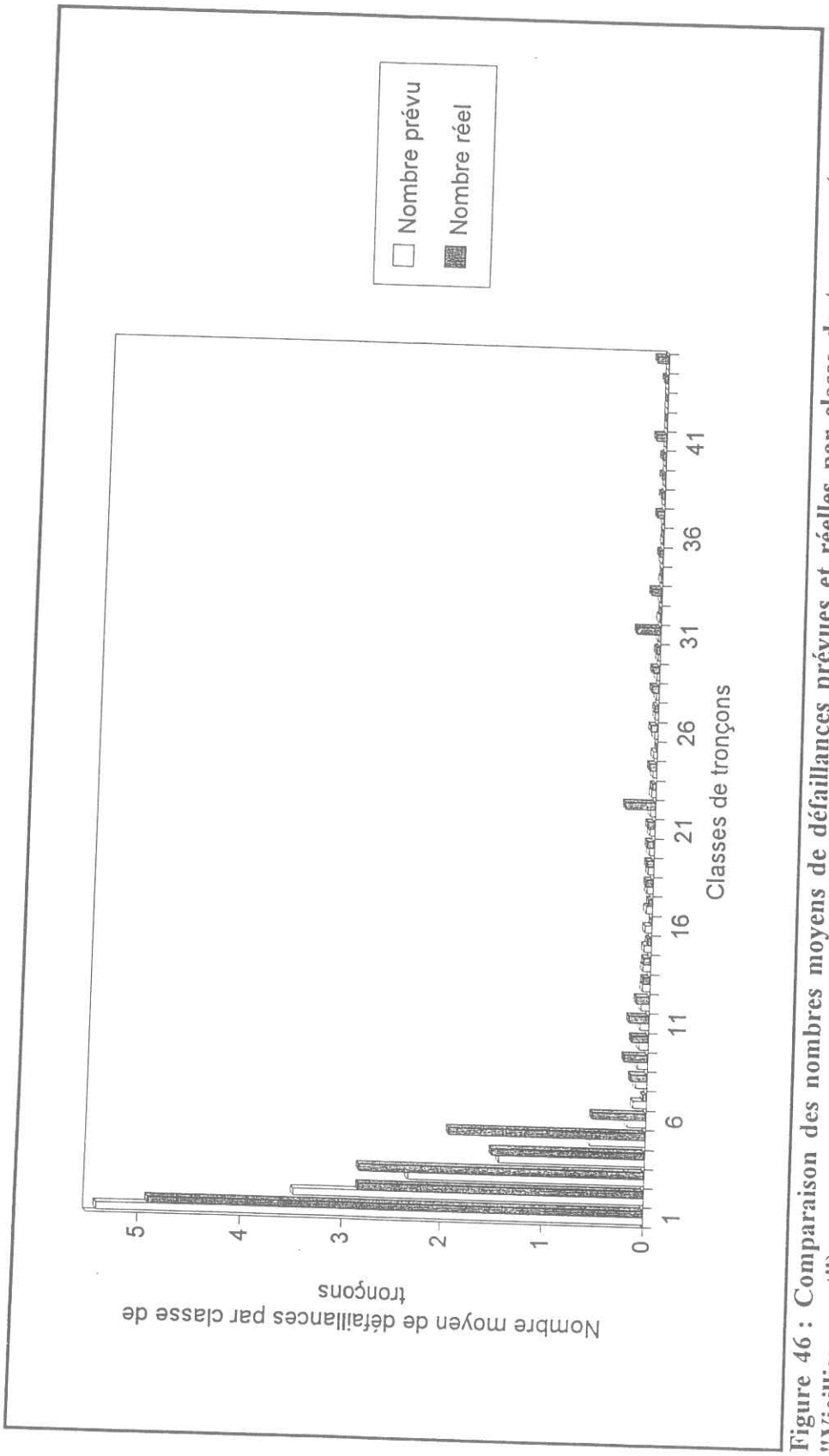


Figure 46 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (avec facteur "Vieillissement").

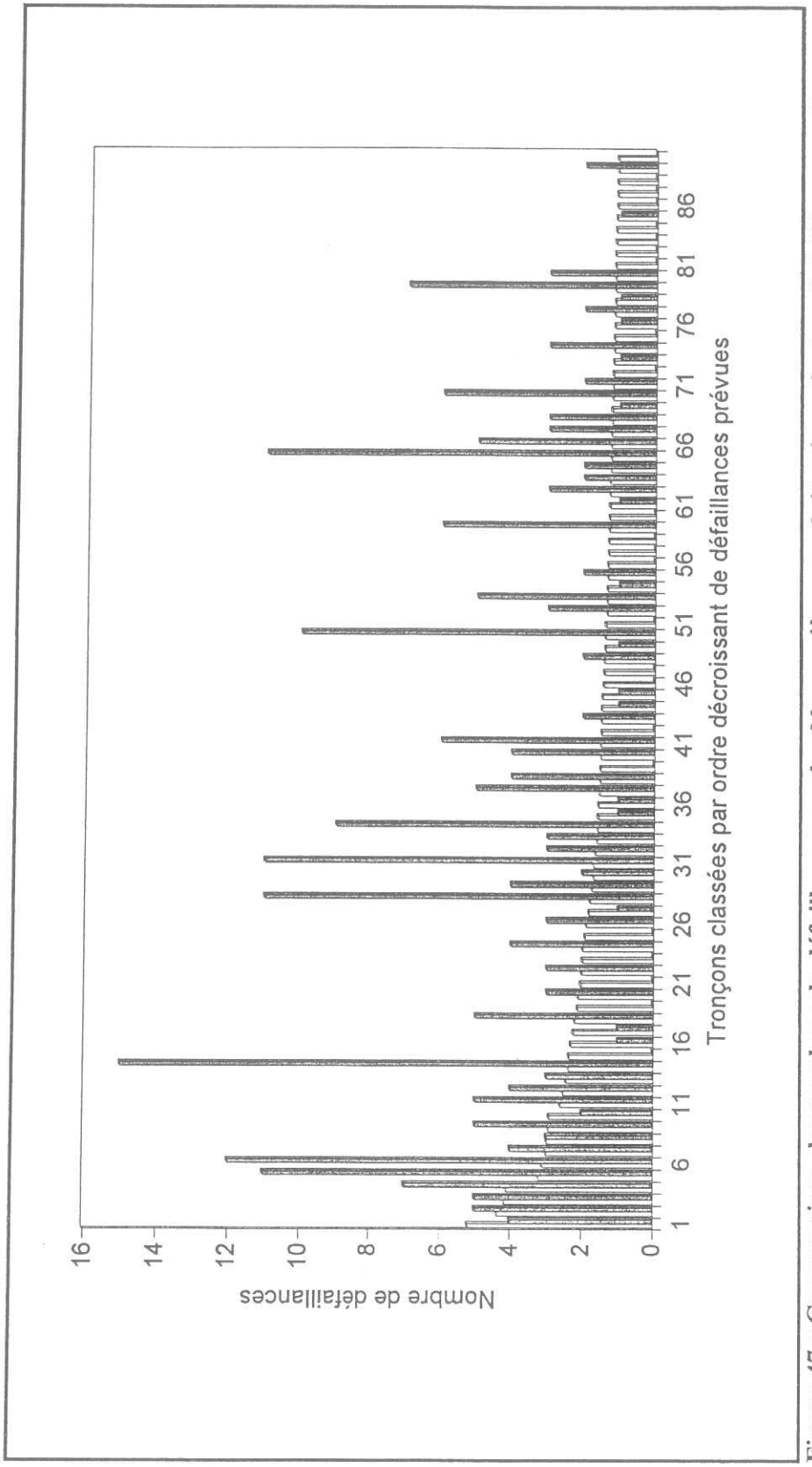


Figure 47 : Comparaison des nombres de défaillances sur les 90 premières conduites (sans facteur "vieillissement")

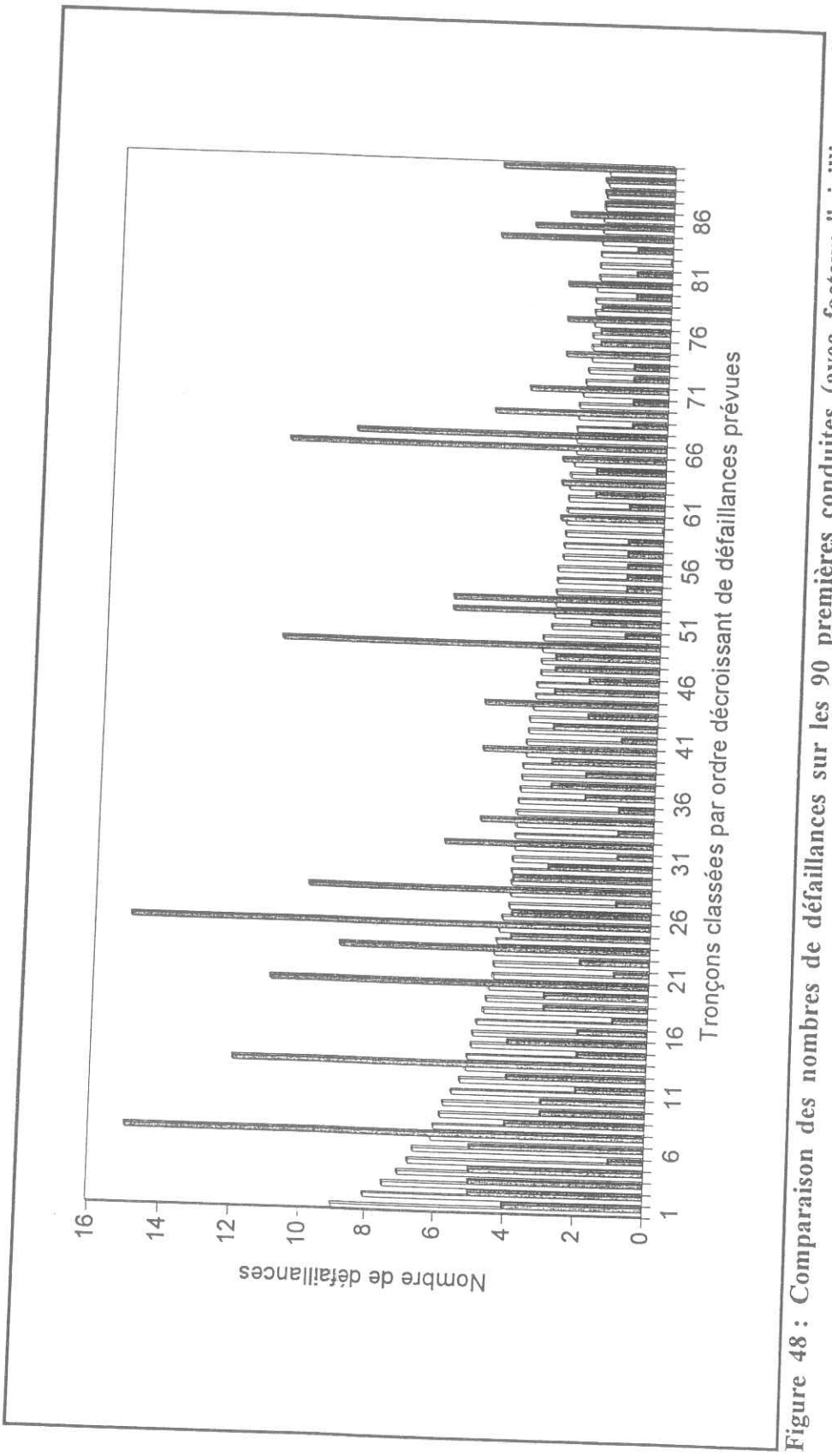


Figure 48 : Comparaison des nombres de défaillances sur les 90 premières conduites (avec facteur "vieillissement")

Tableau 27 : Paramètres du modèle de Weibull stratifié sans facteur vieillissement (données 1951-1982)

1ère défaillance	λ	0,033
	p	1,664
2 à 4ème défaillance	λ	0,178
	p	1,163
5ème défaillance et +	λ	0,15
	p	1
β	Longueur	0,79
	Diamètre	-1,17
	Sol	0,42
	Nature cana	0,13
	Trafic-Chaussée	0,32

Le nombre de défaillances prévues est de 339 alors qu'il y en a eu 267 sur cette période. La surestimation de cette valeur peut être due au fait que depuis une quinzaine d'années les travaux d'entretien ont lieu avec un plus grand soin.

Si l'on classe comme dans le paragraphe précédent les conduites par ordre décroissant des défaillances prévues, on obtient l'histogramme de la figure 50. La prévision est plus forte pour la première classe. Pour les autres classes les valeurs sont du même ordre.

Si l'on examine les 90 premières conduites prévues comme les plus défaillantes (Cf. Fig. 51), on se rend compte que parmi celles ci une vingtaine n'ont pas eu de défaillances.

Le modèle a alors été modifié en changeant les mêmes hypothèses qu'au paragraphe 6.2.1.. Le tableau 28 donne les valeurs des paramètres β_i, λ, et p des différents modèles et le tableau 29 donnent le nombre de défaillances prévues lorsque l'on crée des modèles soit non-stratifié (trois modèles par groupe de défaillances), soit tenant compte du facteur vieillissement.

Tableau 29 : Nombre de défaillances prévues selon les modèles choisis

	Défaillances prévues	Défaillances réelles
Modèle stratifié sans facteur vieillissement	339	267
Modèle non stratifié sans facteur vieillissement	329	
Modèle stratifié avec facteur vieillissement	757	
Modèle non stratifié avec facteur vieillissement	292	

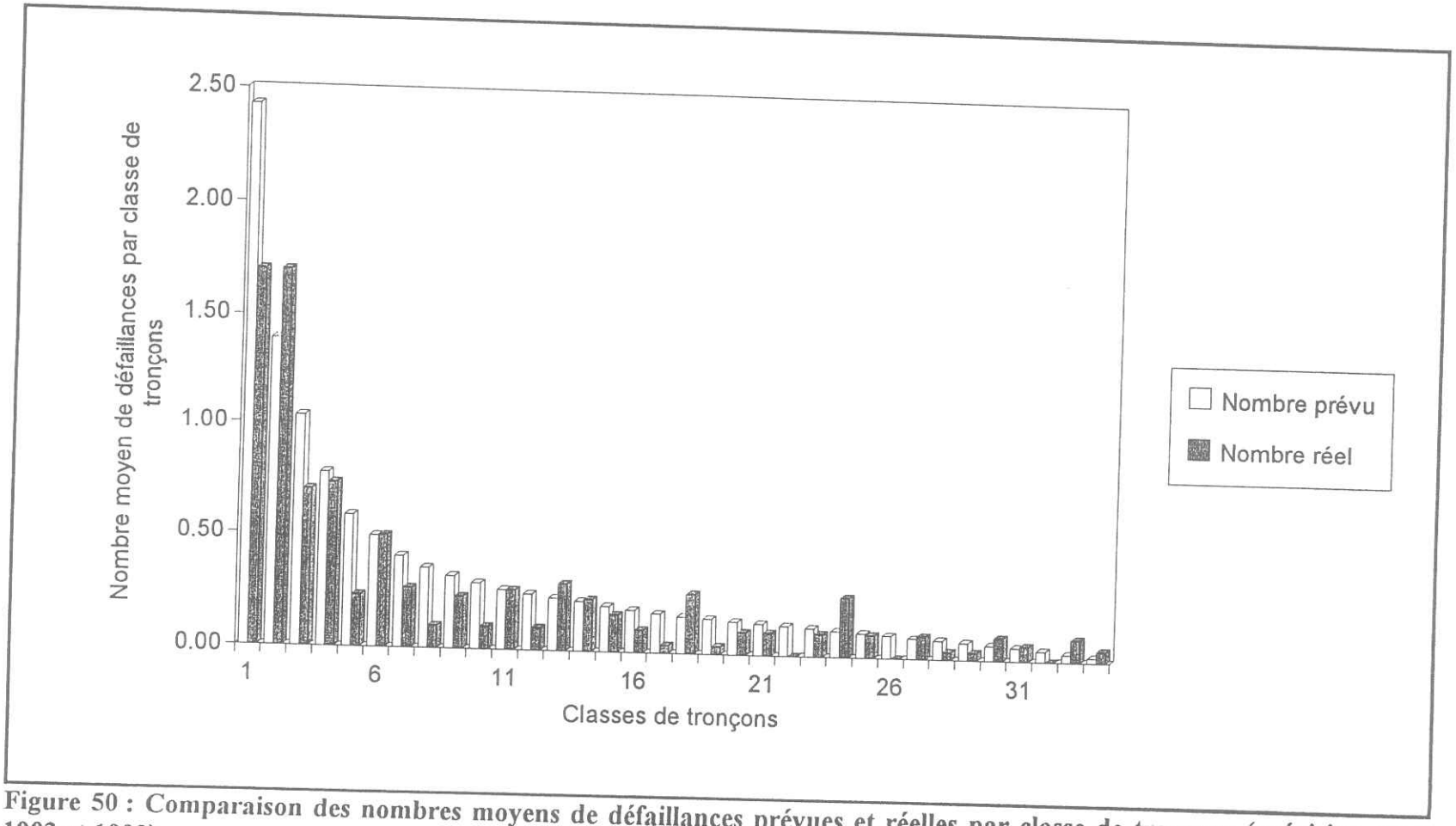


Figure 50 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (prévision entre 1983 et 1992).

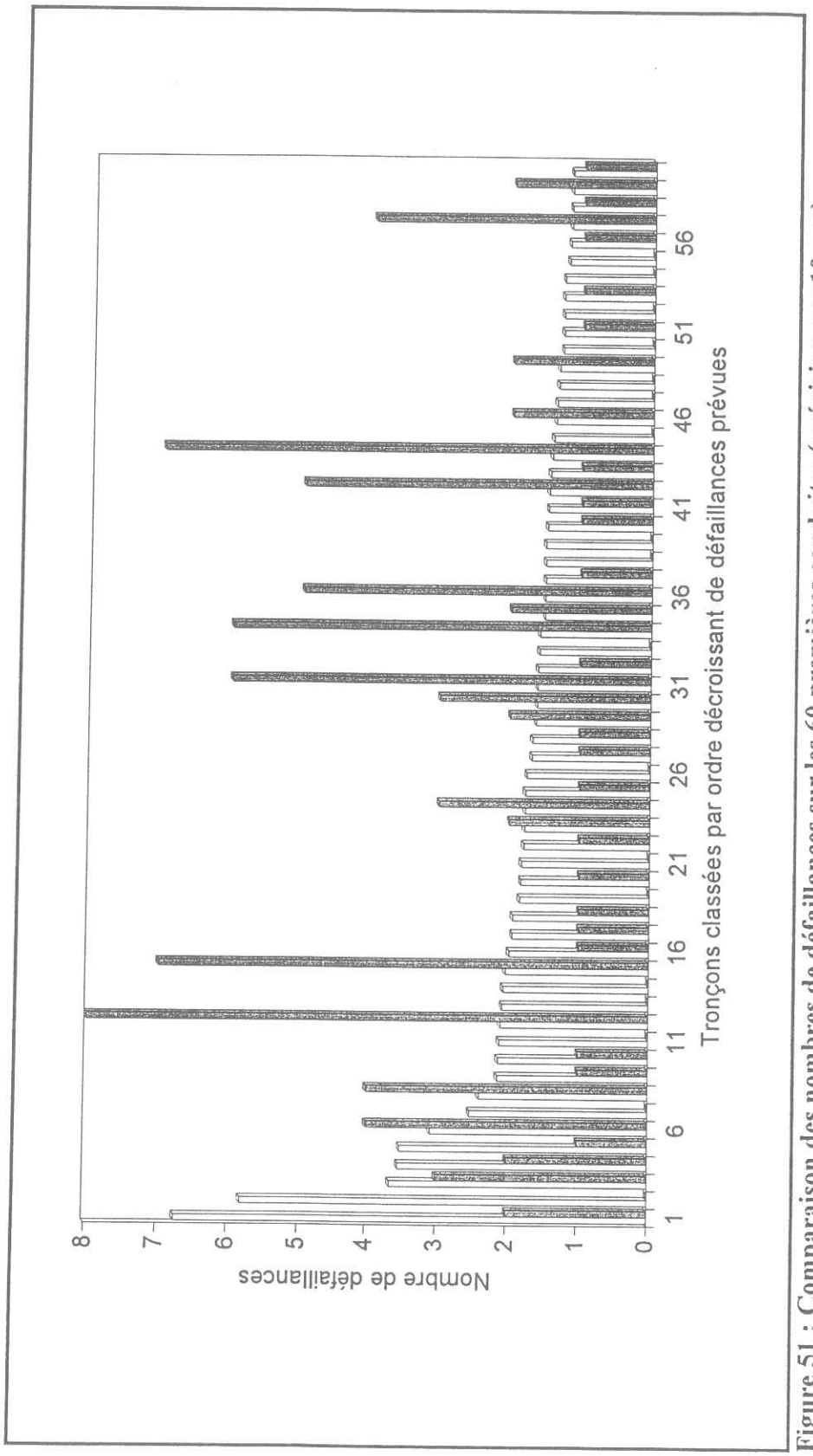


Figure 51 : Comparaison des nombres de défaillances sur les 60 premières conduites (prévision sur 10 ans)

Tableau 28 : Coefficients estimés pour les différents modèles concernant la prévision sur 10 ans.

	Modèle stratifié sans facteur vieillissement			Modèle stratifié avec facteur vieillissement			Modèle non stratifié sans facteur vieillissement			Modèle non stratifié avec facteur vieillissement		
Groupe de défaillance	1	2 à 4	5 et +	1	2 à 4	5 et +	1	2 à 4	5 et +	1	2 à 4	5 et +
λ	0,03	0,18	0,15	0,01	0,09	0,30	0,03	0,20	0,22	0,01	0,10	0,22
p	1,66	1,16	1	1,76	1,23	1	1,68	1,15	1	1,76	1,23	1
β_i Longueur		0,79			0,51		0,93	0,65	**	0,38	0,67	**
Diamètre		-1,17			-0,86		-1,15	-1,3	**	-0,45	-1,43	**
Sol		0,42			0,30		0,61	**	**	0,44	**	**
Trafic-chaussée		0,32			**		0,63	**	**	**	**	**
Nature cana		0,13			0,86		**	**	**	0,82	1,06	**
Vieillissement					1,62					2,41	1,08	**

Concernant le modèle stratifié avec le facteur vieillissement, on se rend compte que le nombre de défaillances prévues est beaucoup trop élevé. Par contre pour les deux modèles non-stratifiés, les deux valeurs sont meilleures. L'examen des histogrammes de classes de conduites montre de plus que c'est le modèle non stratifié avec le facteur vieillissement qui est le meilleur modèle pour la prévision. Les contraintes appliquées sur les paramètres β_i des facteurs entraînent donc un biais trop important, du fait de l'hypothèse de même influence quelque soit le groupe de défaillance. L'histogramme le représentant se trouve dans la figure 52. Les autres histogrammes se trouvent en annexe 6.

La comparaison des défaillances des 90 premières conduites est également meilleure puisque celles qui cassent beaucoup se rapprochent du début de la liste. Cependant encore beaucoup de conduites prévues comme défaillantes ne le sont pas.

6.2.2.2. Prévision sur deux ans

La même étude a été menée en estimant un modèle à partir des données de 1951 à 1990 et en comparant les défaillances prévues entre 1991 et 1992 à celle ayant eu lieu sur la même période. Pour le modèle non-stratifié, on prévoit 101 défaillances au lieu de 56 réelles, ce qui est proche du double. En prenant comme dans le cas précédent, différents modèles stratifiés ou non, avec ou sans facteur "vieillissement", on a les résultats du tableau 30.

Tableau 30 : Nombre de défaillances prévues selon les modèles choisis (prévision sur deux ans.

	Défaillances prévues	Défaillances réelles
Modèle stratifié sans facteur vieillissement	101	56
Modèle non stratifié sans facteur vieillissement	87	
Modèle stratifié avec facteur vieillissement	173	
Modèle non stratifié avec facteur vieillissement	72	

C'est là encore le modèle non stratifié tenant compte du facteur "vieillissement" qui donne les meilleurs résultats. L'histogramme par classe de défaillances décroissantes se trouve dans la figure 53 et celui correspondant aux 90 premières conduites en figure 54.

On a ici des prévisions beaucoup moins précises que sur 10 ans. La durée de prévision de deux ans est en fait trop petite par rapport au pas de temps utilisé dans le modèle. Il est également possible que les deux années étudiées soient des années dont le climat ait été défavorable. Cependant près de la moitié des casses réelles se trouve parmi les 90 premières conduites classées à risque.

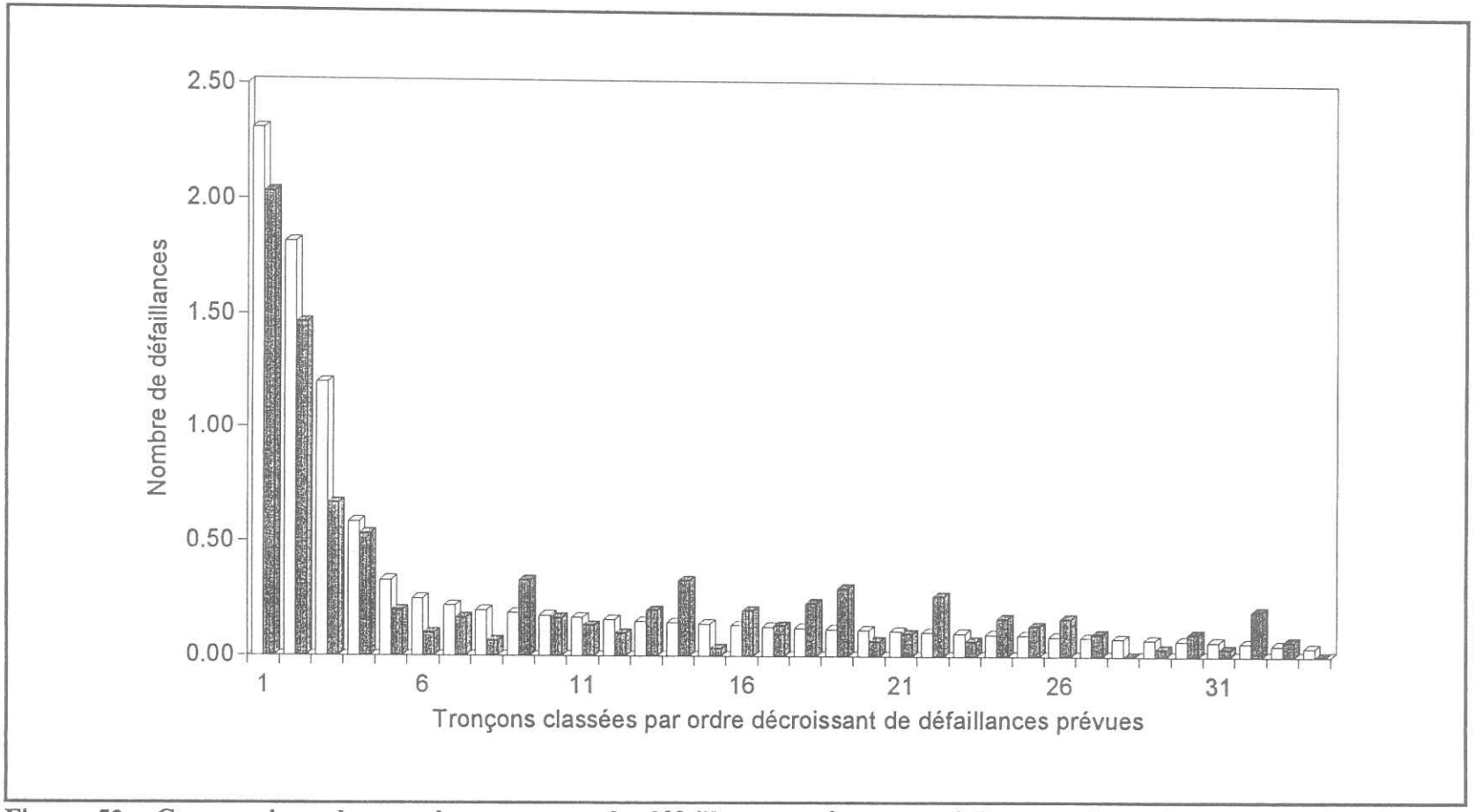


Figure 52 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (modèle non stratifié avec facteur "vieillissement", prévision entre 1983 et 1992).

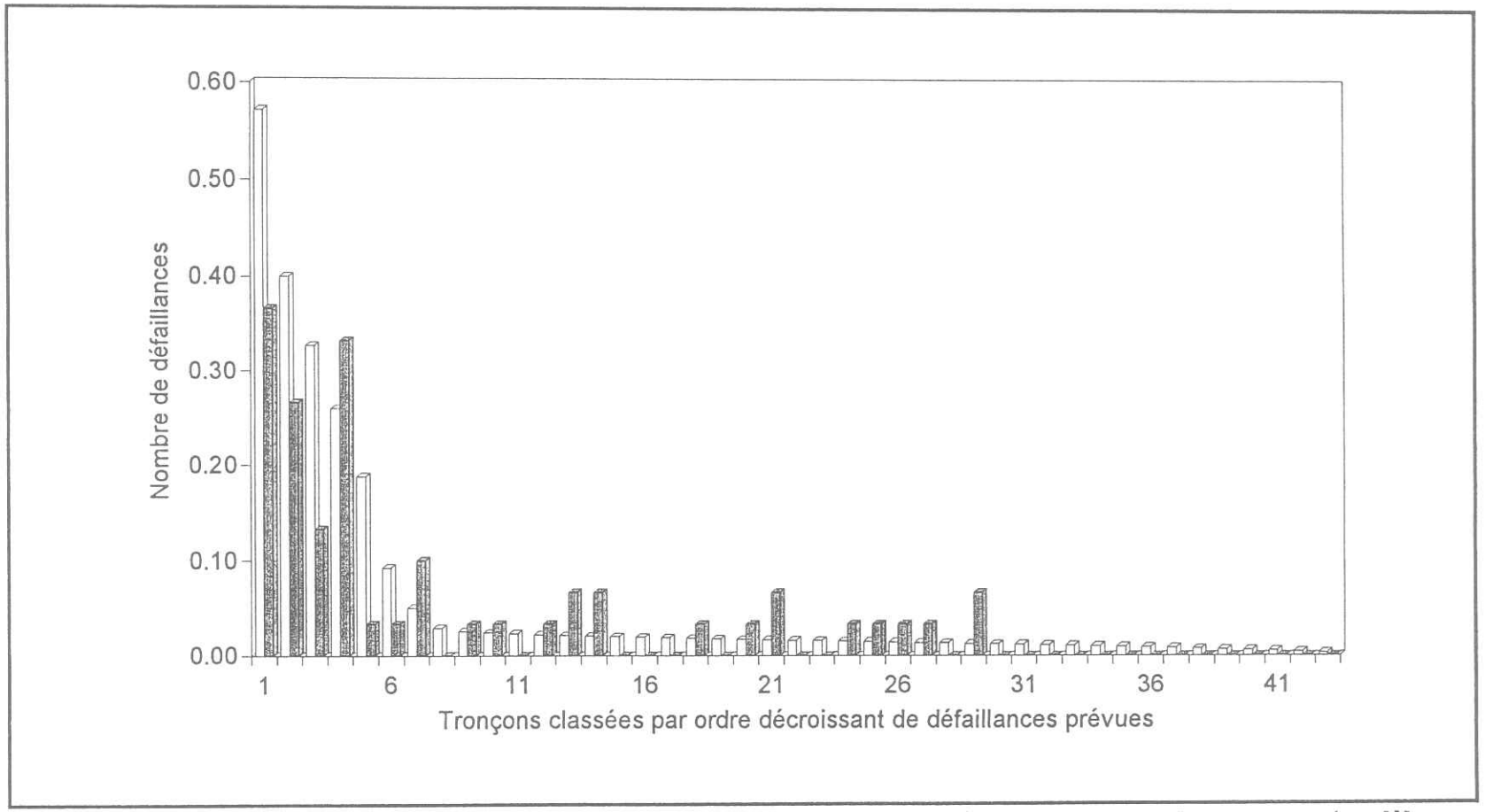


Figure 53 : Comparaison des nombres moyens de défaillances prévues et réelles par classe de tronçons (modèle non stratifié avec facteur "vieillissement", prévision entre 1991 et 1992).

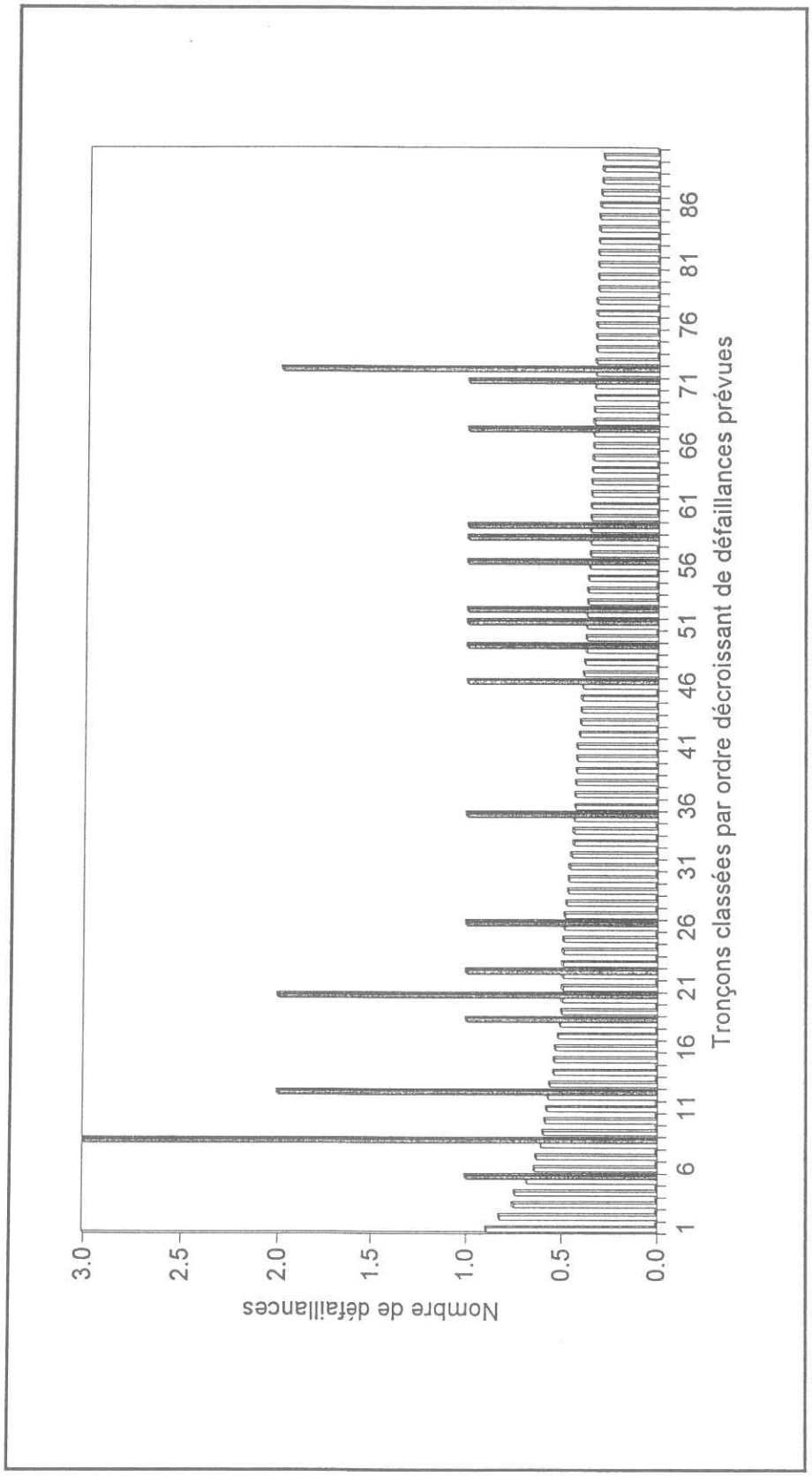


Figure 54 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision sur 2 ans)

6.3. MESURE DE L'INFLUENCE DE L'OUBLI D'UN FACTEUR

Dans certains services il est possible que les données concernant un facteur estimé comme influent ne puissent être mesurées ou ne puissent être retrouvées. D'autre part les paragraphes précédents nous ont montré que l'introduction d'un nouveau facteur, le facteur vieillissement, améliore la prévision. Quelles sont les conséquences de la non prise en compte d'un facteur sur les valeurs des paramètres β_i du modèle de Weibull, ainsi que sur l'espérance du nombre de défaillances pour un échantillon donné?

Le tableau 31 nous donne les coefficients de modèles de Weibull estimés sur Bordeaux en oubliant pour chacun de ces modèles un facteur différent et pour le dernier modèle en rajoutant le facteur vieillissement des paragraphes précédents.

Tableau 31 : Coefficients des modèles de Weibull selon les facteurs pris en compte dans le modèle.

	Modèle complet	Modèle sans le facteur : Longueur	Diamètre	Sol	Traf.-chaus.	Nature cana	Modèle avec le facteur Vieillisse-ment
λ_1	0,03	0,02	0,02	0,03	0,03	0,03	0,01
p_1	1,65	1,59	1,62	1,63	1,63	1,74	1,86
λ_2	0,15	0,09	0,09	0,16	0,15	0,15	0,04
p_2	1,06	1,01	1,02	1,06	1,06	1,06	1,16
λ_3	0,13	0,14	0,12	0,22	0,14	0,13	0,06
p_3	1	1	1	1	1	1	1
β Longueur	0,58		0,42	0,57	0,61	0,60	0,36
Diamètre	-0,87	-0,31		-0,84	-0,81	-0,92	-0,52
Sol	0,37	0,39	0,34		0,38	0,41	0,29
Tr.-ch.	0,49	0,64	0,39	0,49		0,53	
Nat. cana	-0,10	-0,23	-0,24	-0,13	-0,16		0,61
Vieillis.							2,42

Les facteurs longueurs et diamètre ont une influence plus importante que les autres facteurs. En effet, lorsqu'on les oublie, les paramètres β_i des autres facteurs varient beaucoup plus. Cependant l'oubli d'un de ces facteurs est corrigé par le paramètre λ (paramètre d'intensité) qui est beaucoup plus faible, notamment pour le deuxième groupe de défaillance.

Concernant le facteur "vieillissement", toutes les valeurs des autres paramètres sont nettement différentes. Ceci peut se comprendre du fait de la définition même de ce facteur qui dépend du nombre et de la date des défaillances des conduites. Il est ainsi évident qu'il englobe en partie le

fait que la conduite soit dans des conditions défavorables ou non, comme par exemple dans un sol corrosif ou non. Il est donc normal qu'il y ait une certaine corrélation.

L'introduction de ce facteur fait également diminuer dans des proportions plus importantes les coefficients λ et p, ceci du fait de son poids important (β=2,42).

Le calcul de l'espérance du nombre de défaillances en fonction des différents modèles donne les résultats du tableau 32.

C'est le modèle tenant compte du facteur "vieillissement" qui apparaît comme le meilleur du point de vue du nombre de défaillances total (522 défaillances réelles). Paradoxalement l'oubli des facteurs "sol" et "trafic-chaussée" voit le nombre de défaillances prévues se rapprocher du nombre de défaillances réelles.

Tableau 32 : Espérances estimées sur la période de 1951 à 1990 selon les facteurs pris en compte dans le modèle.

	Modèle complet	Modèle sans le facteur :					Modèle avec le facteur Vieillisse-ment
		Longueur	Diamètre	Sol	Traf.-chaus.	Nature cana	
Espérance	434	388	396	451	452	408	447
Intervalle de confiance (95%)	346-546	313-485	331-476	365-559	363-562	334-498	316-623

Si l'on examine les histogrammes comparant les quatre vingt-dix premières conduites prévues comme cassant le plus et leurs défaillances réelles, c'est encore le modèle avec le facteur vieillissement qui donne le meilleur résultat, puisque uniquement deux conduites parmi celles ci ne cassent pas (Cf. figure 48).

L'oubli des facteurs longueur, diamètre et trafic-chaussée augmente au contraire le nombre de conduites ne cassant pas alors qu'elles étaient prévues comme cassantes (Cf. figure 55). Par contre on peut se rendre compte que l'oubli des facteurs sol et Nature cana ne change pas fondamentalement les résultats par rapport au modèle complet même si certaines conduites qui ont eu beaucoup de casses n'apparaissent plus parmi les quatre-vingt-dix premières(Cf. figure 56).

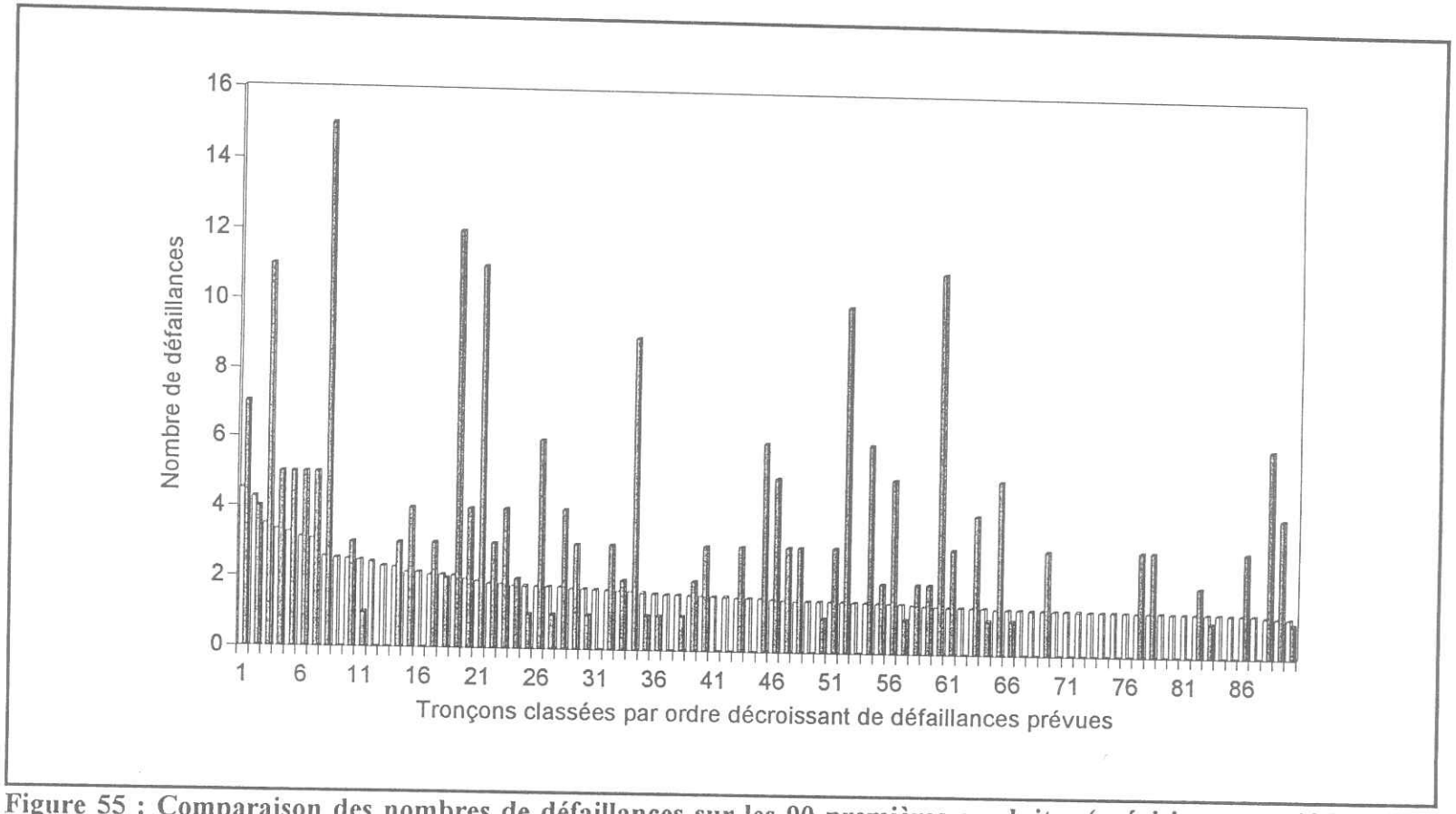


Figure 55 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision entre 1951 et 1990; oubli du facteur trafic-chaussée).

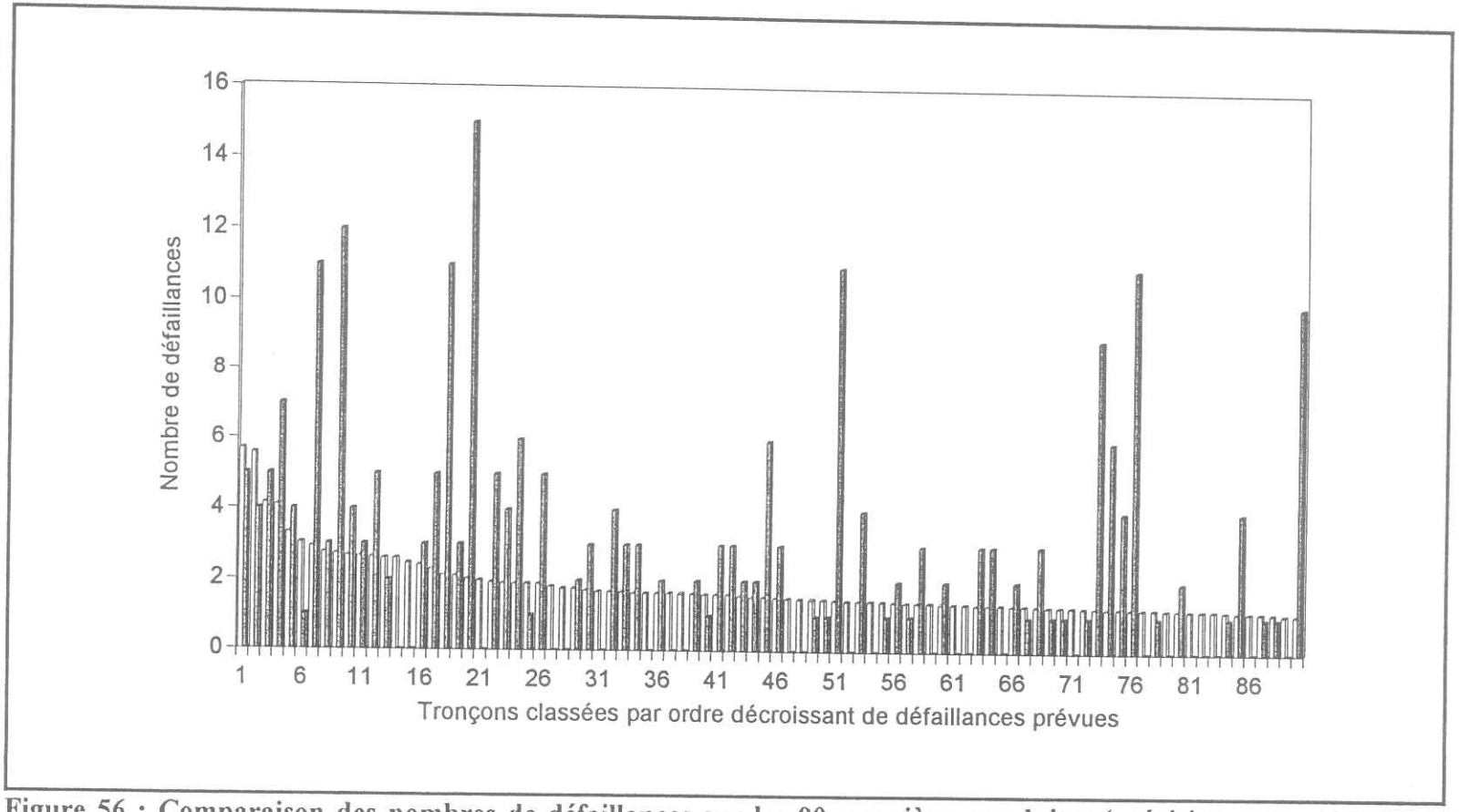


Figure 56 : Comparaison des nombres de défaillances sur les 90 premières conduites (prévision entre 1951 et 1990; oubli du facteur sol).

6.4. CONCLUSION

La comparaison des défaillances prévues et des défaillances réelles permet de faire les conclusions suivantes :

- Concernant le calage du modèle sur 40 ans, on sous-estime légèrement les défaillances. L'introduction du facteur "Vieillissement" donne un résultat plus proche et près des 3/4 des défaillances réelles se trouvent parmi les 100 premières conduites prévues comme les plus défaillantes.

- Le modèle stratifié n'entraîne pas de biais pour ce qui concerne le calage sur 40 ans. Par contre, pour la prévision le modèle non-stratifié donne des meilleurs résultats.

- Le fait de modéliser ne permet cependant pas de définir les conduites ayant un nombre de casses important. Par contre ces conduites se trouvent parmi celles prévues comme les plus défaillantes.

- Certaines conduites prévues comme non-défaillantes ont tout de même quelques casses. Cependant si l'on fait les prévisions année après année, ces conduites changeront de groupe de défaillances et leur espérance de casse sera plus grande. Elles se rapprocheront alors de la "tête" de liste.

- La prévision sur dix ans donne des résultats satisfaisants (pour le modèle non-stratifié incluant le facteur "vieillissement"). Ceci n'est cependant pas le cas de la prévision sur deux ans, même si près de la moitié des conduites qui ont eu des casses se trouvent parmi les 100 premières.

- L'oubli des facteurs longueurs, diamètre et trafic chaussée entraîne un biais important dans l'estimation des espérances. Pour les deux premiers, les informations les concernant sont souvent les mieux connues dans les services, alors que pour le facteur "Trafic-chaussée" ce n'est pas toujours le cas. Par contre l'oubli des facteurs sol et "nature cana." entraîne un biais plus faible.

CHAPITRE 7 :
L'APPRENTISSAGE

L'estimation des modèles conçus dans le chapitre 5 a nécessité de collecter un volume de données assez important. Elle a pu se faire grâce à leur existence sur une assez longue durée dans les services étudiés. Malheureusement le cas de ces services est assez exceptionnel.

En effet les données concernant les historiques de maintenance sont rares ou n'existent que sur une durée trop courte. Dans certains services ce sont les personnes travaillant sur le réseau qui connaissent les casses ou les fuites qui ont eu lieu depuis leur arrivée. L'historique est alors perdu lorsqu'elles quittent le service. Pour d'autres services les données ont pu être perdues lors de la mise en concession du réseau, lors d'un regroupement en syndicat ou bien même lors d'un déménagement.

Certains services commencent à se doter de bases de données informatisées depuis quelques années. On peut donc espérer à l'avenir avoir des historiques suffisamment conséquents.

Afin de pouvoir créer dès à présent ce type de modèles dans des services à faible historique, il est tout de même important de concevoir une méthode permettant d'utiliser des données en faible quantité. C'est cette méthode qui est présentée dans le présent chapitre.

7.1. LA METHODE

Tout d'abord il est évident que, même si les données de maintenance sont en quantité plus faible, il est indispensable que le réseau soit connu, à savoir au minimum le diamètre des conduites, leur âge, et éventuellement certaines caractéristiques comme le sol, le trafic, l'emplacement sous chaussée ou sous trottoir.

Supposons que l'on veuille créer un modèle sur un service A ayant 5 ans (de 1988 à 1993) de données et dont la date de pose la plus ancienne est 1945.(Cf. Figure 57) Supposons de plus que l'on connaisse les données d'un modèle général.

Les données du service A seules ne permettraient pas d'estimer un modèle convenable. Cependant elles apportent une information non négligeable sur le réseau et ses défaillances. On va donc tenter d'élaborer un modèle à partir des données du service A et de celles du modèle général.

Rappelons que les résultats du chapitre 5 ont montré que l'influence des paramètres était sensiblement identique d'un service à l'autre alors que les fonctions de risques de base sont différentes et représenteraient le caractère local des services. Calculer la fonction de risque de

base à partir des seules données du service A donnerait là encore des résultats trop imprécis compte tenu de la faible longueur de l'historique.

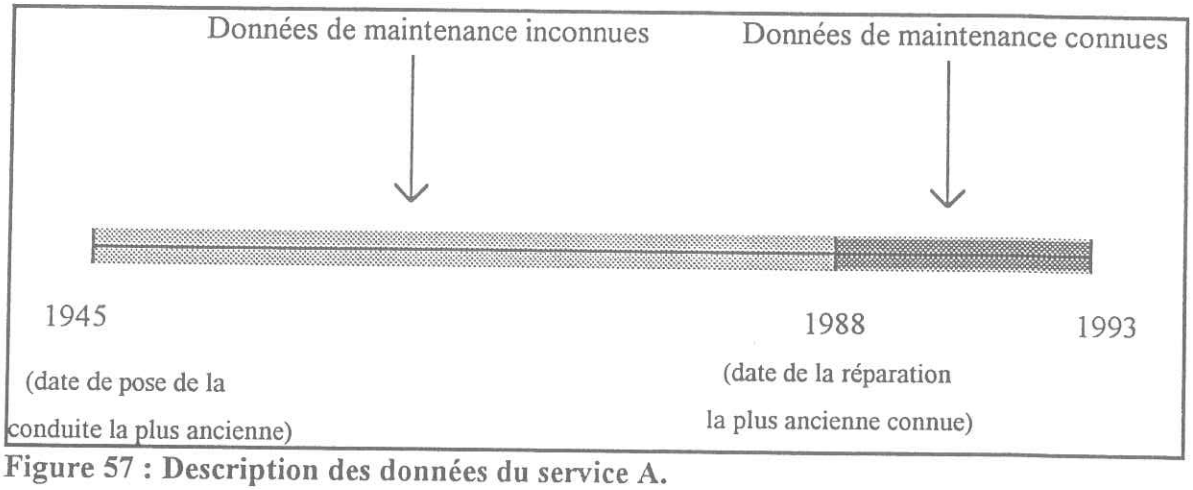


Figure 57 : Description des données du service A.

Deux méthodes sont alors possibles. On peut tout d'abord estimer un modèle à partir des données du modèle général et de celles du service (Cf. Figure 58). Ceci entraînerait cependant l'élimination de tous les tronçons du service A posés avant 1988, car sur une certaine période on ne connaît pas les travaux de maintenance les concernant. Les données du modèle général auraient alors une part trop importante.

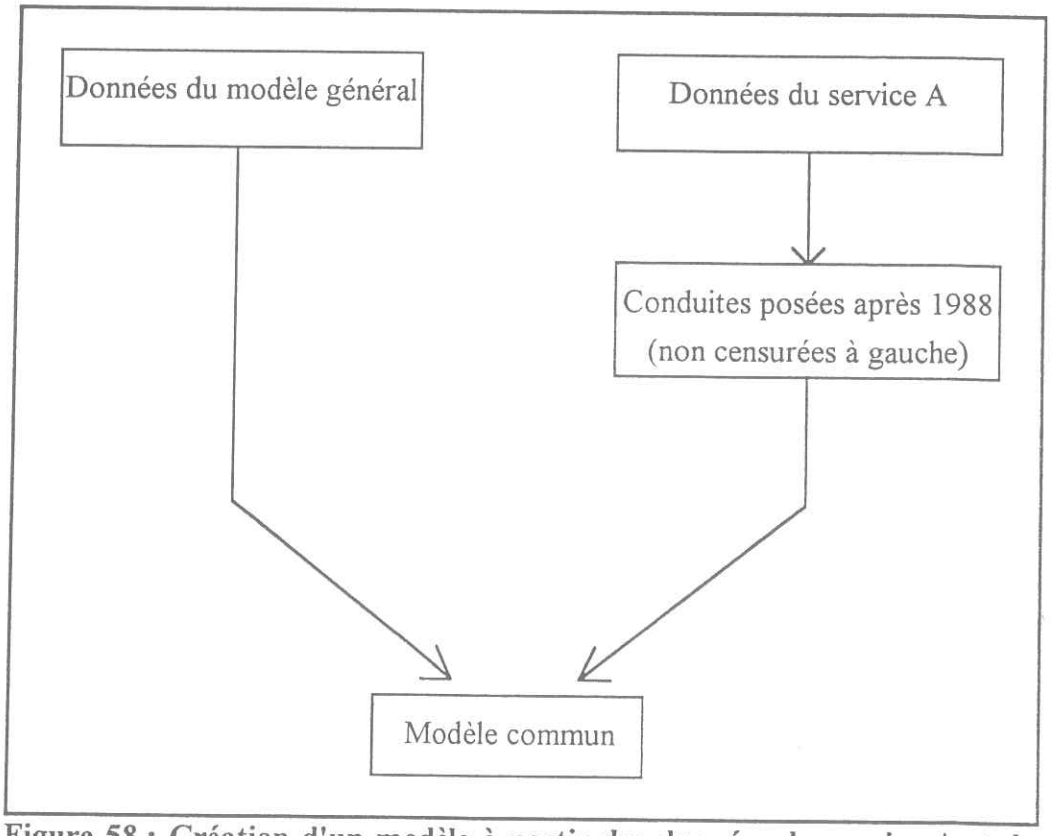


Figure 58 : Création d'un modèle à partir des données du service A et du modèle général.

Une autre méthode consiste à créer à partir du modèle général un échantillon de défaillances concernant les conduites censurées à gauche c'est à dire les conduites posées avant 1988. Cet échantillon de défaillances ajouté aux défaillances réelles apparues après 1988 permettrait alors de calculer un modèle local propre au service A. (Cf. Figure 59)

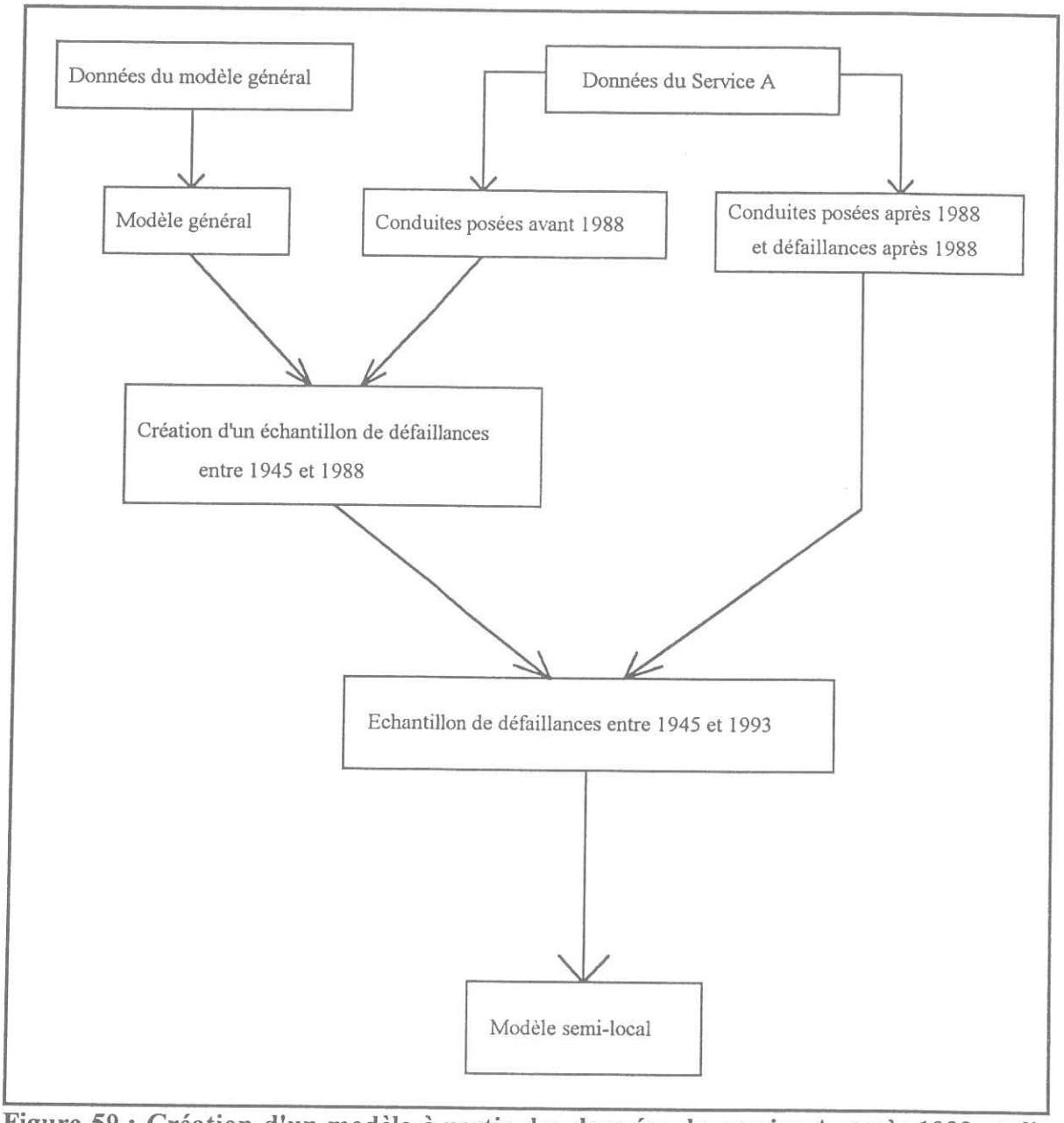


Figure 59 : Création d'un modèle à partir des données du service A après 1988 et d'un échantillon de défaillances construit à partir du modèle général pour les conduites posées avant 1988.

La principale difficulté dans cette méthode est de créer l'échantillon de défaillances pour les conduites posées avant 1988. On peut la résoudre de deux manières :

- Soit on calcule pour chacune des conduites l'espérance de défaillances en chaque année de la date de pose à 1988 et on retrouve alors la date d'apparition des défaillances. L'espérance est calculée de la même manière que dans le chapitre précédent.

- Soit on tire au hasard les défaillances de façon à obtenir la même fonction de non-défaillance que celle du modèle général. Cette méthode est appelée "anamorphose" [Saporta, 1990].

Ce sont ces deux méthodes qui ont été expérimentées. Les paragraphes suivants en présentent les résultats.

7.2. RESULTATS

L'expérimentation a eu lieu à partir des données de la LED-Bordeaux et de la CEB. Le modèle de la CEB a été considéré comme étant le modèle général. Les données de la LED-Bordeaux ont été volontairement étudiées à partir de l'année 1989. On a donc deux ans de données de défaillances entre 1990 et 1991 et il faut créer un échantillon de défaillances pour les conduites posées avant 1989.

7.2.1. Echantillonnage à partir de l'espérance tirée du modèle général

C'est la première formule de l'espérance du chapitre 6 qui a été utilisée (calcul depuis la date de pose). Sur les 1296 tronçons posés avant 1989, 211 défaillances ont été prévues.

Une fois ces défaillances ajoutées à celles qui ont eu lieu après 1989, on peut calculer le modèle de Weibull. Les paramètres de ce modèle se trouvent dans le tableau 33.

Ces résultats sont peu satisfaisants. En effet à chaque fois les valeurs absolues des β_i sont surestimées par rapport à celles obtenues avec l'ensemble des données. Ceci signifie que cette méthode accentue l'influence des facteurs. Les paramètres du modèle de Weibull sont également très différents.

Le calcul de l'espérance lisse beaucoup trop l'apparition des défaillances. Par exemple pour toutes les conduites, aucune défaillance n'aura lieu pendant une certaine période après la date de pose. Par contre les défaillances apparaîtront sur une courte période, ce qui donne une courbe de probabilité de non-défaillances de forme sigmoïde (Cf. fig. 60).

Utiliser le calcul de l'espérance n'apparaît donc pas comme une solution satisfaisante pour compléter les données.

Tableau 33 : Coefficients du modèle semi-local en échantillonnant à partir de l'espérance

		Bordeaux	Apprentissage à partir de l'espérance sur 38 ans	CEB
β	Longueur	0.57	1.90	0.37
	Nature cana	-0.13	**	-0.64
	Diamètre	-0.85	-2.11	-0.34
	Sol	0.34	1.06	0.29
	Trafic - chaussée	0.47	2.97	0.57
1ère défaillance	λ1	0.03	0.04	0.02
	p1	1.67	5.89	1.21
2 à 4ème défaillance	λ2	0.15	0.08	0.10
	p2	1.06	2.17	0.99
5ème défaillance et +	λ3	0.15	0.05	0.20
	p3	1	1	1

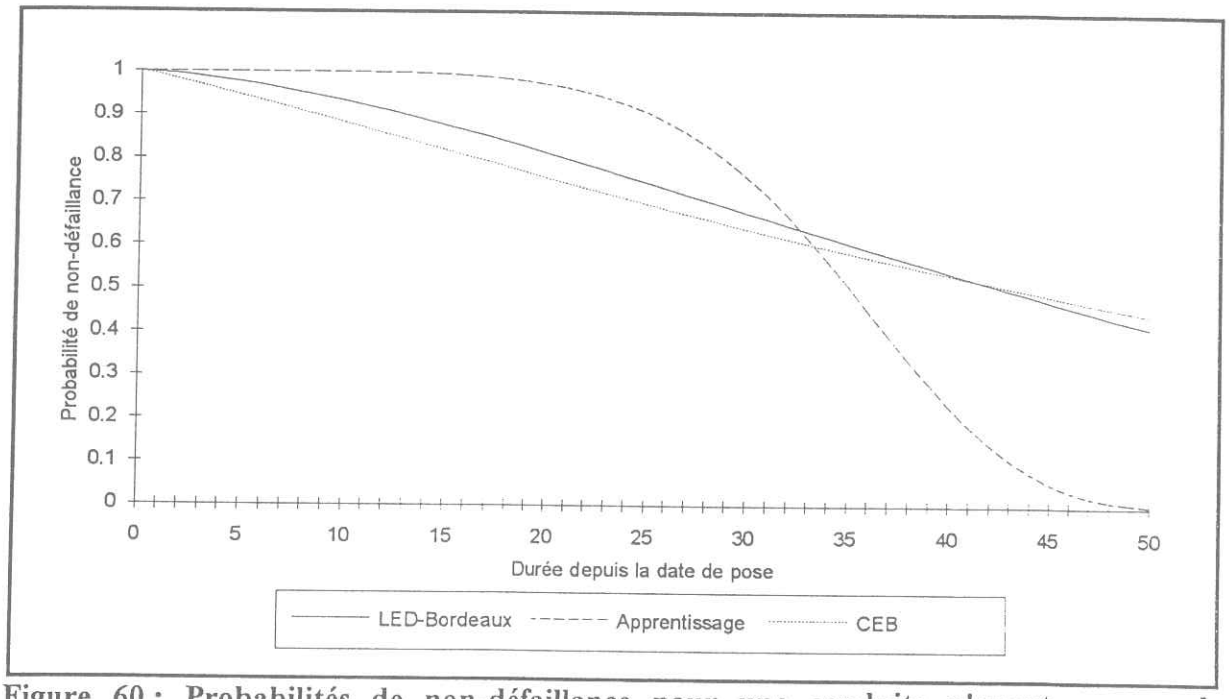


Figure 60 : Probabilités de non-défaillance pour une conduite n'ayant pas eu de défaillance, de diamètre 150, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage à partir de l'espérance).

7.2.2. Echantillonnage à partir de la méthode dite d'anamorphose [Saporta, 1990]

La méthode dite d'anamorphose est une méthode tirant au hasard les individus de façon à obtenir une même fonction de répartition (Cf. Fig. 61).

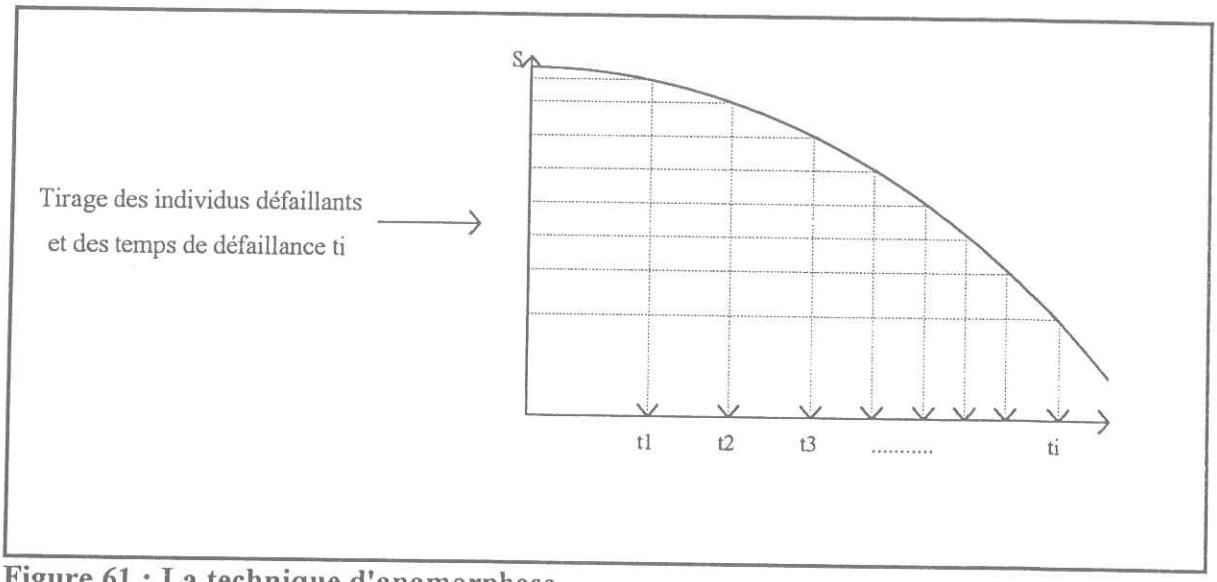


Figure 61 : La technique d'anamorphose.

En pratique la démarche est la suivante :

1) Regroupement des conduites posées entre 1951 et 1989 par classe suivant les valeurs des différents facteurs influents (sol, trafic-chaussée, diamètre, Nature cana).

2) Estimation par classe d'une longueur équivalente l_{eq} calculée de telle façon que :

N_k étant le nombre de tronçons de la classe k et l_i la longueur du tronçon i.

$$\sum_{i=1}^{N_k} h_{CEB}(t,l_i) = N_k \times h_{CEB}(t,l_{eq}).$$

Cette longueur équivalente est approximée sur l'ensemble de l'échantillon quelle que soit la valeur de t. (Cf annexe 7)

3) Estimation par classe du nombre de premières défaillances n_k. Ce nombre est égal à :

$$n_k = \sum_{i=1}^{N_k} \left(1 - S_{CEB}(t_i, l_i)\right)$$

t_i étant la durée d'observation entre la date de pose et 1989.

4) Détermination du nombre probable de défaillances à l'âge *a* depuis la date de pose pour chacune des classes. On considère qu'une défaillance a lieu au temps t_j lorsque la partie entière au temps t_j de (N*h) est égale à la partie entière au temps t_{j-1} de (N*h) auquel on a ajouté 1. On aura donc :

$$n_{ka} = \text{int}\left(\sum_{j=1}^{a} N_{kj} \times h_{CEB}(j, l_{eq})\right) - \text{int}\left(\sum_{j=1}^{a-1} N_{k(j-1)} \times h_{CEB}(j-1, l_{eq})\right)$$

int étant la partie entière, n_{ka} étant le nombre probable de casses pendant l'année a pour la classe k, N_{kj} étant le nombre d'individus de la classe k restant au temps j (non censurés).

5) Tirage au hasard, dans chacune des classes, des n_k conduites défaillantes.

6) Détermination pour chacune de ces conduites de la date d'apparition de la défaillance t_i' telle que :

$$h_{CEB}(t_i', l_i) = h_{CEB}(t_i, l_{eq}).$$

Ceci concerne les premières défaillances. Pour les deuxièmes défaillances la même méthode est employée en prenant comme échantillon de base les conduites tirées pour la première défaillance, de même pour les défaillances suivantes.

Une fois cette opération réalisée, il est ensuite possible de regrouper ces défaillances tirées au hasard avec les défaillances réelles ayant eu lieu entre 1989 et 1992. Puis on calcule le modèle de Weibull stratifié sur le groupe de défaillances comme précédemment. Les coefficients et paramètres se trouvent dans le tableau 34.

Tableau 34 : Coefficients du modèle semi-local en échantillonnant à partir de la technique d'anamorphose.

		Bordeaux	Apprentissage à partir de l'anamorphose	CEB
β	Longueur	0.57	0.57	0.37
	Nature cana	-0.13	-0.63	-0.64
	Diamètre	-0.85	-0.36	-0.34
	Sol	0.34	0.28	0.29
	Trafic - chaussée	0.47	0.40	0.57
1ère défaillance	λ1	0.03	0.03	0.02
	p1	1.67	1.59	1.21
2 à 4ème défaillance	λ2	0.15	0.07	0.10
	p2	1.06	1.12	0.99
5ème défaillance et +	λ3	0.15	0.14	0.20
	p3	1	1	1

Les coefficients sont, cette fois, du même ordre que les coefficients établis sur les modèles de la CEB et de la LED-Bordeaux. Ils sont plus proches de ceux de la CEB, mais non éloignés de ceux de Bordeaux. De ce fait les courbes sont beaucoup plus proches de celles établies sur Bordeaux et la CEB (Cf. Fig. 62, 63 et 64).

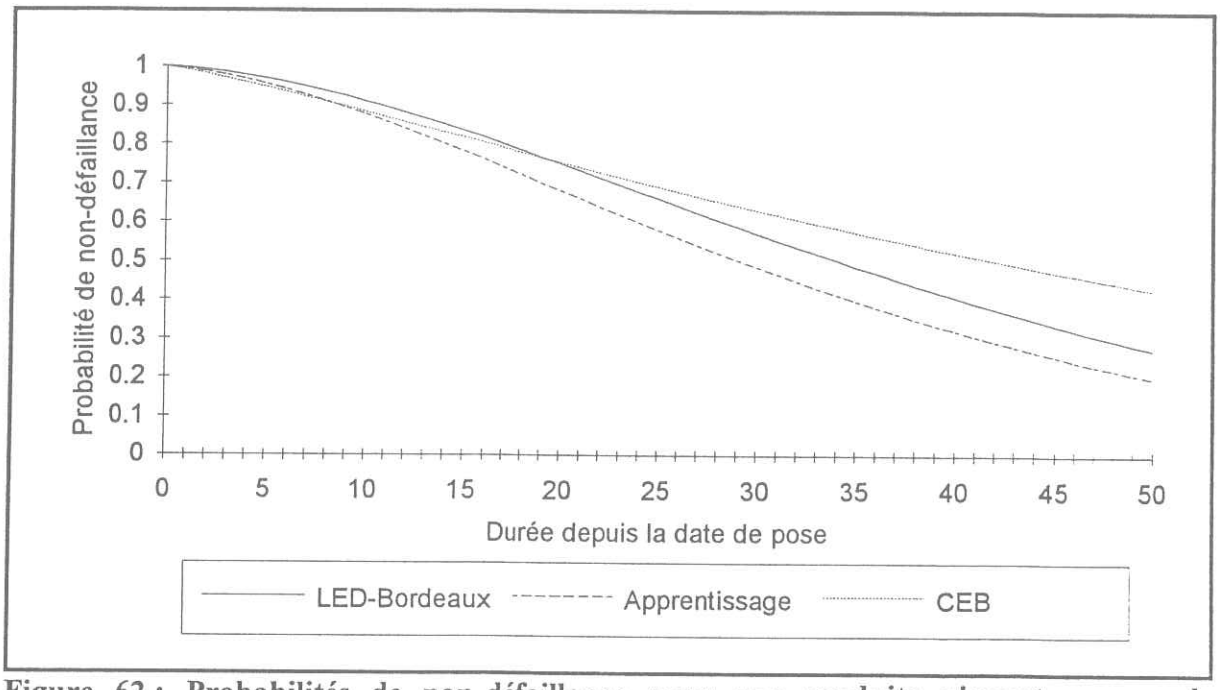


Figure 62 : Probabilités de non-défaillance pour une conduite n'ayant pas eu de défaillance, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose).

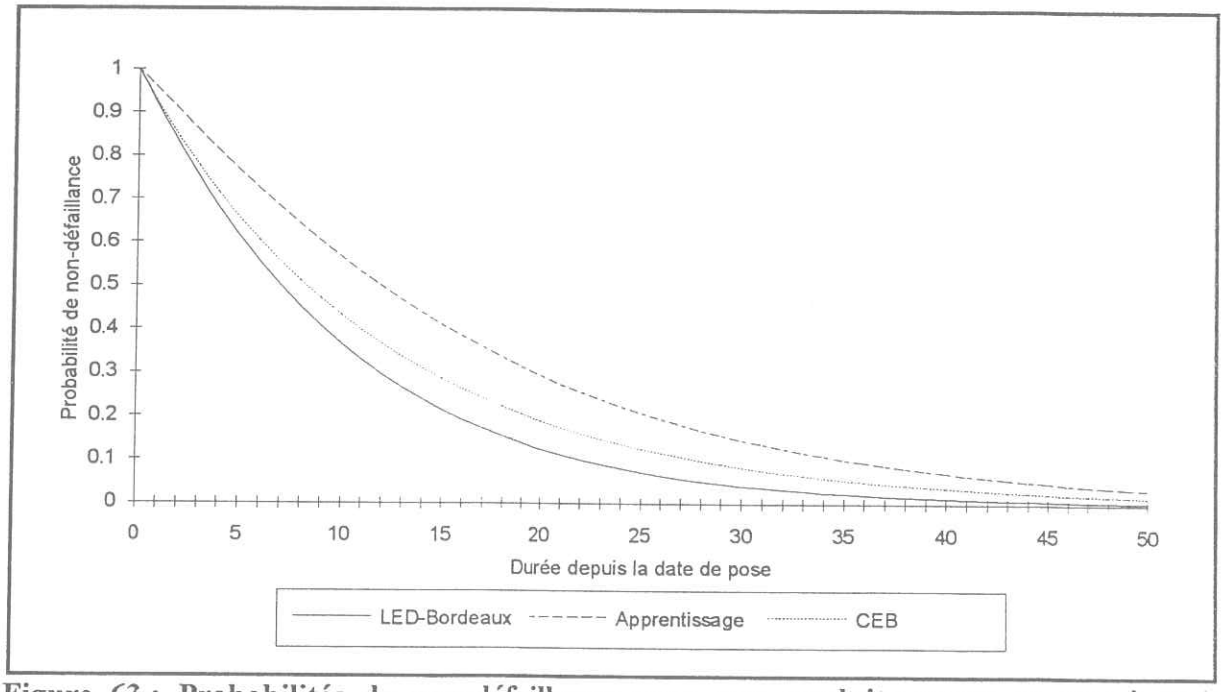


Figure 63 : Probabilités de non-défaillance pour une conduite ayant eu une à trois défaillances, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose).

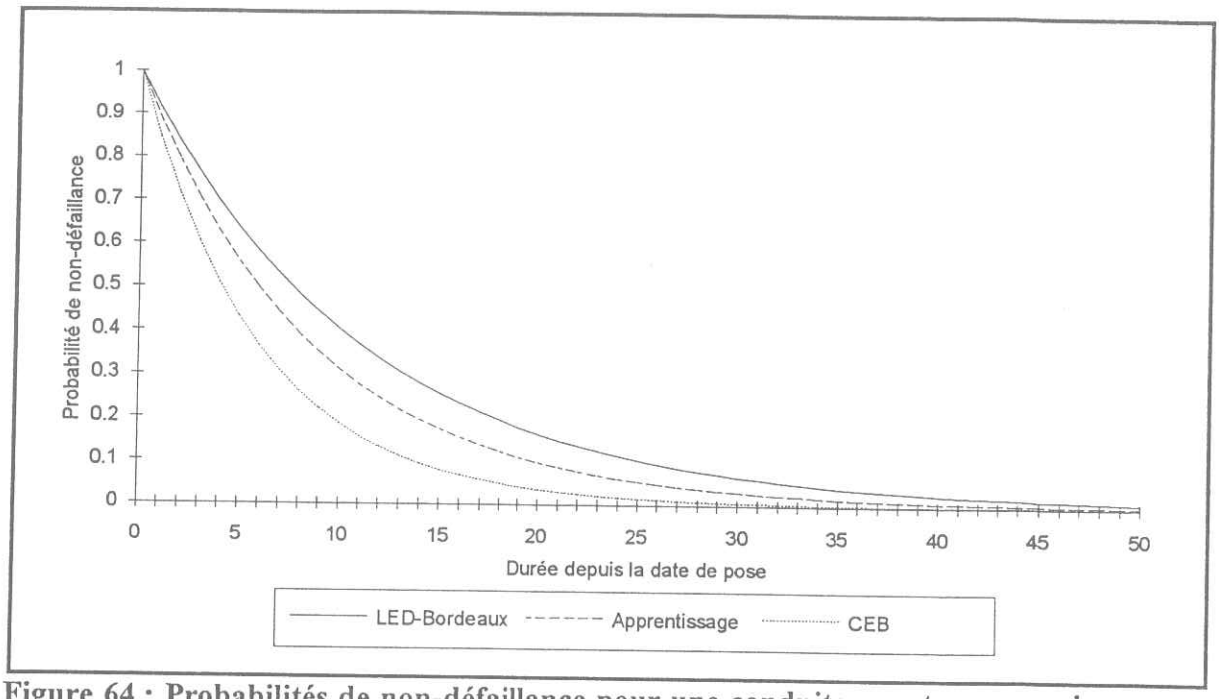


Figure 64 : Probabilités de non-défaillance pour une conduite ayant eu au moins quatre défaillances, de diamètre 100, dans un sol non-corrosif, sous trottoir ou dans une rue à trafic faible (Apprentissage par anamorphose).

Cette méthode apparaît comme donnant de meilleurs résultats que la méthode d'apprentissage à partir de l'espérance. C'est elle que l'on utilisera à l'avenir.

7.3. Conclusion

Même avec des données de défaillances en faible quantité, la méthode d'apprentissage peut être utilisée pour le modèle de Weibull sans faire une erreur trop importante. L'anamorphose est donc une méthode très satisfaisante.

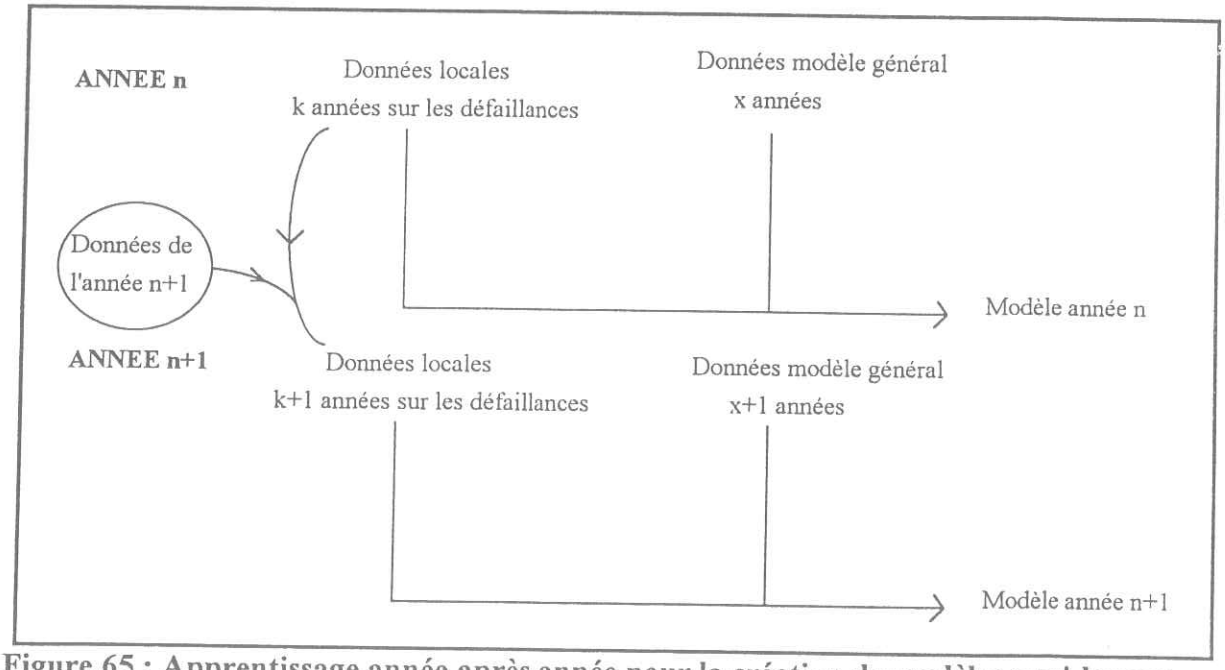


Figure 65 : Apprentissage année après année pour la création de modèles semi-locaux.

De plus d'année en année, le modèle pourra évoluer avec les données nouvelles recueillies. Les données locales auront donc un poids plus important au fil des années (Cf. Figure 65).

Il est cependant important de posséder les données sur les conduites sur la plus longue durée possible afin de probabiliser plus précisément les défaillances. Une gestion informatisée de ces données faciliterait bien entendu leur utilisation et leur manipulation en vue du calcul de l'espérance de défaillances.

CHAPITRE 8 :
UTILISATION DU MODELE PAR LES GESTIONNAIRES DE SERVICE

Cette étude a permis d'élaborer un outil statistique de prévision des défaillances des réseaux d'eau potable. Cet outil doit pouvoir être utilisable par les gestionnaires de service. Nous rappellerons donc tout d'abord la méthodologie à suivre pour son application dans un service d'eau potable. Puis nous verrons de quelles manières les résultats de ce modèle peuvent être utilisés dans le service.

8.1. METHODOLOGIE POUR LA CREATION D'UN MODELE

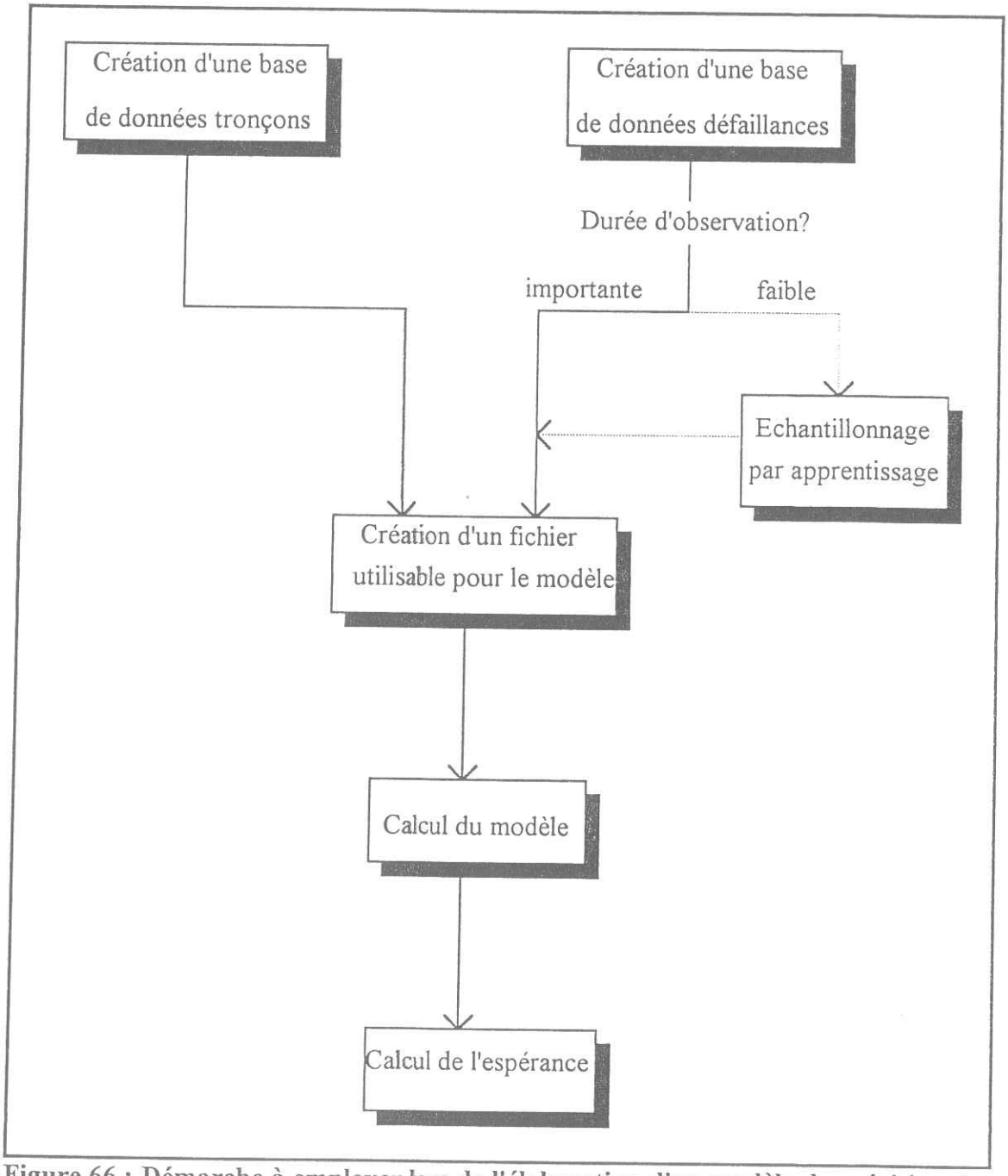


Figure 66 : Démarche à employer lors de l'élaboration d'un modèle de prévision des défaillances.

Pour la création des modèles sur les services, une démarche décrite en figure 66 a été suivie. C'est cette même démarche qui pourra être utilisée avec certains changements selon la durée d'observation des données de défaillances.

Les différentes étapes sont décrites dans les paragraphes suivants.

8.1.1. Création de bases de données

Les bases de données sont de deux types : celle concernant les tronçons et celle concernant les défaillances.

8.1.1.1 La base de données "Tronçons"

Cette base de données doit contenir les caractéristiques des tronçons de canalisations du service. Un tronçon décrit ici est un ensemble de canalisations se trouvant dans la même rue, de même diamètre, posées la même année et homogènes vis-à-vis des facteurs étudiés. Les caractéristiques des tronçons seront donc :

- la rue,
- la longueur,
- le diamètre,
- le matériau,
- l'année de pose,
- l'année de remplacement ou de réhabilitation,
- le sol et sa corrosivité,
- le trafic,
- l'emplacement de la conduite sous chaussée ou sous trottoir.

A ces caractéristiques pourront être ajoutés d'autres facteurs qui semblent avoir une influence sur les défaillances sur ce réseau (après enquête, par exemple). Concernant la longueur, il sera important de déterminer des tronçons qui soient le plus court possible, afin d'atténuer son effet.

Cette mise en place de bases de données nécessite donc que les facteurs environnementaux cités soient connus. Ceci pourra entraîner la mise en place d'études les concernant, notamment pour le sol. Si un facteur manquait, un modèle pourrait également être élaboré, mais celui ci serait bien entendu moins précis.

Il est également important de noter les numéros de rues du début et de la fin du tronçon afin, par la suite, de bien déterminer quel tronçon est concerné par une défaillance.

Ce type de base de données est différent de bases de données existant déjà et qui déterminent les tronçons de canalisations en fonction des organes de régulation existant sur les réseaux.

8.1.1.2. La base de données "Défaillances"

Cette base de données peut être rattachée à une base de données concernant tous les travaux de maintenance. Rappelons qu'une défaillance est ici une réparation due à une fuite ou une casse sur le tuyau ou sur le joint de la canalisation.

La structure de cette base de données pourra être de la forme suivante :
- Nom de la rue,
- Diamètre,
- N° dans la rue,
- Date de la défaillance,
- Nature de la défaillance,
- Cause de défaillance,
- Type de réparation.

Il sera nécessaire de bien connaître la nature de la défaillance afin de bien différencier les fuites sur joints des fuites sur tuyaux, qui n'ont pas obligatoirement une même origine. Ceci n'avait pu être déterminé lors de l'étude, compte tenu de l'imprécision existant sur les fiches de maintenance.

Une fiche type remplie sur place faciliterait la saisie de cette base de données. Cette fiche permettrait également l'obtention de renseignements sur les tronçons pour lesquels certaines caractéristiques ne sont pas connues. Un exemple de ce type de fiche est représenté en figure 67.

8.1.2. Application des données à l'étude statistique

La durée d'observation des défaillances est importante pour la suite de l'opération. Si elle correspond à l'âge de la canalisation la plus ancienne, on n'aura alors aucun tronçon censuré à gauche. L'apprentissage ne sera alors pas nécessaire. Si par contre elle est nettement inférieure à l'âge de la canalisation la plus ancienne, on aura alors beaucoup de censures à gauche et il sera nécessaire de récréer un échantillon de défaillances sur ces conduites. Enfin si aucune donnée de défaillance existe, on devra alors utiliser un modèle général.

Une fois l'échantillon des défaillances complet, il est nécessaire de réunir les deux bases de données afin de créer un fichier utilisable pour le modèle. Ce fichier est constitué d'individus décrits dans la figure 25 du chapitre 5, un tronçon pouvant constituer plusieurs individus se différenciant par le nombre de défaillances (Cf. Figure 68).

Comme on l'a vu au chapitre 6, un facteur supplémentaire peut être introduit : le facteur vieillissement. Ce facteur correspond à l'apparition ou non de défaillances après 15 ans d'âge.

NATURE

☒ Fuite
☐ Casse entreprise
☐ Reprise de réparation
☐ .....
☐ Signalée par tiers
☐ Détectée suite à recherche

Date : 4 01 90

SI COCUMONT

Commune : MARCELLUS

Lieu-dit : Au Cérac

Abonné : .....

RENSEIGNEMENTS COMPLÉMENTAIRES SUR LA FUITE (à compléter par l'intervenant)

☐ Refoulement
☒ Distribution
☐ Refoulement - Distribution

* Conduite Ø : 33 4/40
Profondeur conduite : 1 ml
Année présumée de pose : .....

☐ Fonte ductile
☒ PVC
☐ Acier
☐ Autre fonte
☐ Autre .....

Fuite constatée sur : ☒ Alignement droit ☐ Coude ☐ té ☐ Piquage
☐ Autre : .....

Pression de service : ..... bars
Débit estimé de la fuite : ..... m³
Origine de l'eau : ☐ Source, Nom..... ☒ Forage, Nom MUSCAT ☐ Prise rivière, Nom.....

FUITE PAR	CAUSE PRÉSUMÉE	NATURE DU SOL		
☐ Casse franche	☐ Corrosion intérieure	☐ Rocher	☐ Sec	☐ Sous chaussée
☐ Trou	☐ Corrosion extérieure	☐ Sable	☐ Humide	☐ Sous accotement
☐ Eclatement	☐ Poinçonnement	☒ Terre arable	☐ Détrempé	☐ Terrain cultivé
☒ Joint	☒ Décollage	☐ Terre arable (avec cailloux)	☐ Inondé	☐ Friche
☐ Pièce de raccord	☐ Pièce défectueuse			
☐	☐			

Conditions météo particulières : .....

Remblaiement tranchée : ☒ Terrain naturel ☐ Sable tout venant ☐ Grave, ciment, émulsion

Autres installations dans la tranchée : ☐ Assainissement ☐ Téléphone ☐ Electricité ☐ gaz

Ligne électrique haute tension ☐ oui ☒ non Distance en mètres .....

Présence liquide : ☐ Nappe phréatique ☐ Eaux usées ☐ Purin ☐ Autres .....

Canalisation : État du revêtement - Interne : Bon
- Externe : Bon
- Autres : .....
- Nature : .....

Dépôt interne - Epaisseur : .....

Commentaires : .....

* Conduite Ø : Intérieur pour fonte acier
Intérieur et extérieur pour PVC ET Polyéthylène.

Figure 67 : Exemple de fiche de réparation.

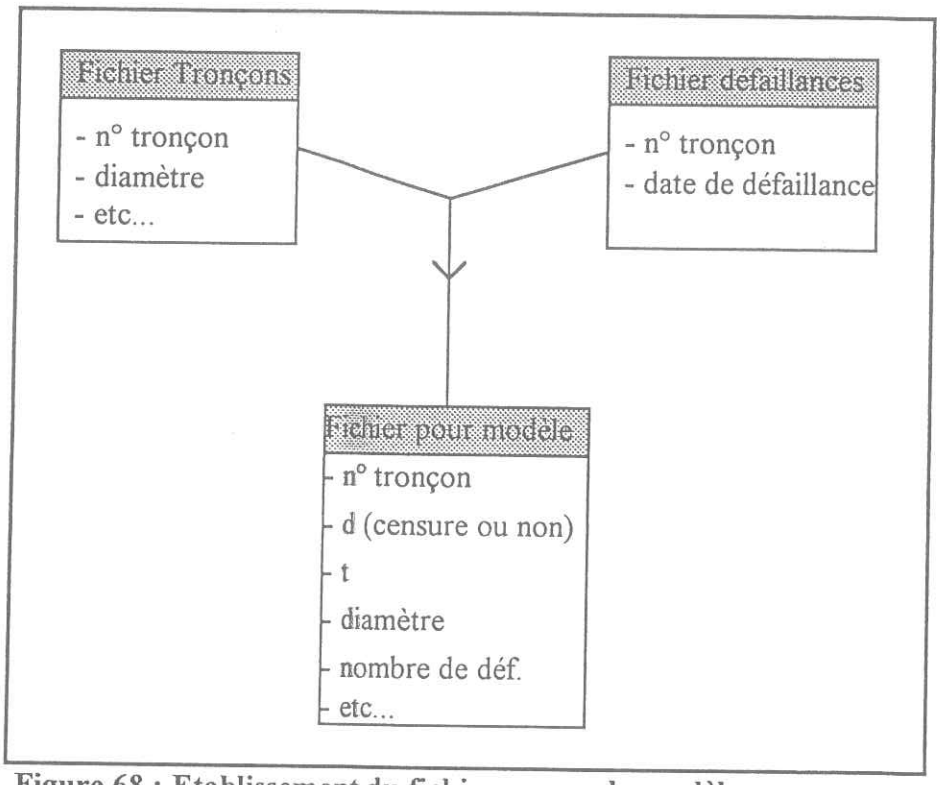


Figure 68 : Etablissement du fichier en vue du modèle.

8.1.3. Création du modèle

Le modèle peut alors être estimé. Le choix peut se faire entre un modèle stratifié ou un modèle non-stratifié. La démarche à suivre est alors la suivante :

1) Estimation des facteurs influents (test de Wald) par le modèle de Cox

2) Estimation des paramètres β_i de ces facteurs (vraisemblance de Cox) avec stratification ou non et des paramètres de Weibull λ et p (modèle de Weibull).

Le logiciel utilisé pour l'estimation du modèle est le logiciel EGRET. D'autres logiciels statistiques (SAS, BMDP, NCSS par exemple) permettent également l'analyse de données de survie.

8.1.4. Estimation de la probabilité ou de l'espérance

Une fois le modèle établi, on peut alors définir pour chaque canalisation soit la probabilité d'apparition de la prochaine défaillance , soit l'espérance à une certaine échéance.

Pour une conduite ayant eu une troisième défaillance il y a 3 ans et si l'on veut faire une prévision sur dix ans, la probabilité d'apparition de la prochaine défaillance est de :

$$P_{10} = 1 - \frac{S_2^*(3+10)}{S_2^*(3)}.$$

Pour l'espérance la formule utilisée est la même que celle décrite au chapitre 6. Il est alors possible de classer les conduites par risque de défaillance. Le listing du programme se trouve en annexe 6.

8.2. UTILISATION POUR CONNAITRE L'ETAT DU RESEAU

La démarche décrite précédemment permet de définir les conduites à risque. Le gestionnaire de service peut alors mieux connaître le réseau et définir les zones prioritaires à surveiller. Des campagnes de mesures de fuites pourront alors être axées plutôt sur ces conduites. Puis il pourra être décidé de remplacer, de réhabiliter ou de laisser telles quelles les conduites selon les résultats de ces campagnes.

Connaissant le nombre de défaillances prévues, le gestionnaire peut alors budgétiser sur cette période les travaux de maintenance futurs.

8.3. UTILISATION POUR LA PLANIFICATION DU RENOUVELLEMENT

Comme on l'a vu au chapitre 1, le taux de rupture sur les canalisations est un critère important dans le choix du remplacement d'une conduite. Cet outil est donc utilisable dans le cadre de la planification du renouvellement. Le chapitre 1 a montré les deux principales politiques de renouvellement existant : la politique curative et la politique prédictive. Ces deux politiques peuvent intégrer le modèle.

8.3.1. Intégration dans le cadre d'une politique curative

Dans ce type de politique, le taux de rupture entre en compte lorsque d'autres critères de renouvellement ont déjà été déterminés. Ce sont, par exemple, l'opportunité de travaux de voirie ou bien l'enveloppe financière prévue pour le renouvellement (Cf. Figure 69).

8.3.2. Intégration dans le cadre d'une politique prédictive

Dans une politique prédictive de renouvellement des conduites d'eau potable, seul rentre en compte dans le choix du renouvellement l'état futur de la conduite. L'estimation de cet état nécessite l'existence d'un outil d'analyse prédictive.

Cette prévision peut se faire à plus ou moins long terme. Elle peut concerner différentes représentations de l'état de la conduite, comme ceux déjà cités au chapitre 1 : le nombre de ruptures, les fuites diffuses, la qualité de l'eau, la diminution de la capacité de transport. Une politique complète devrait prendre en compte tous ces paramètres. Le modèle de Weibull mis en place pour l'étude est un outil d'analyse prédictive des réparations sur les conduites, à savoir leurs ruptures.

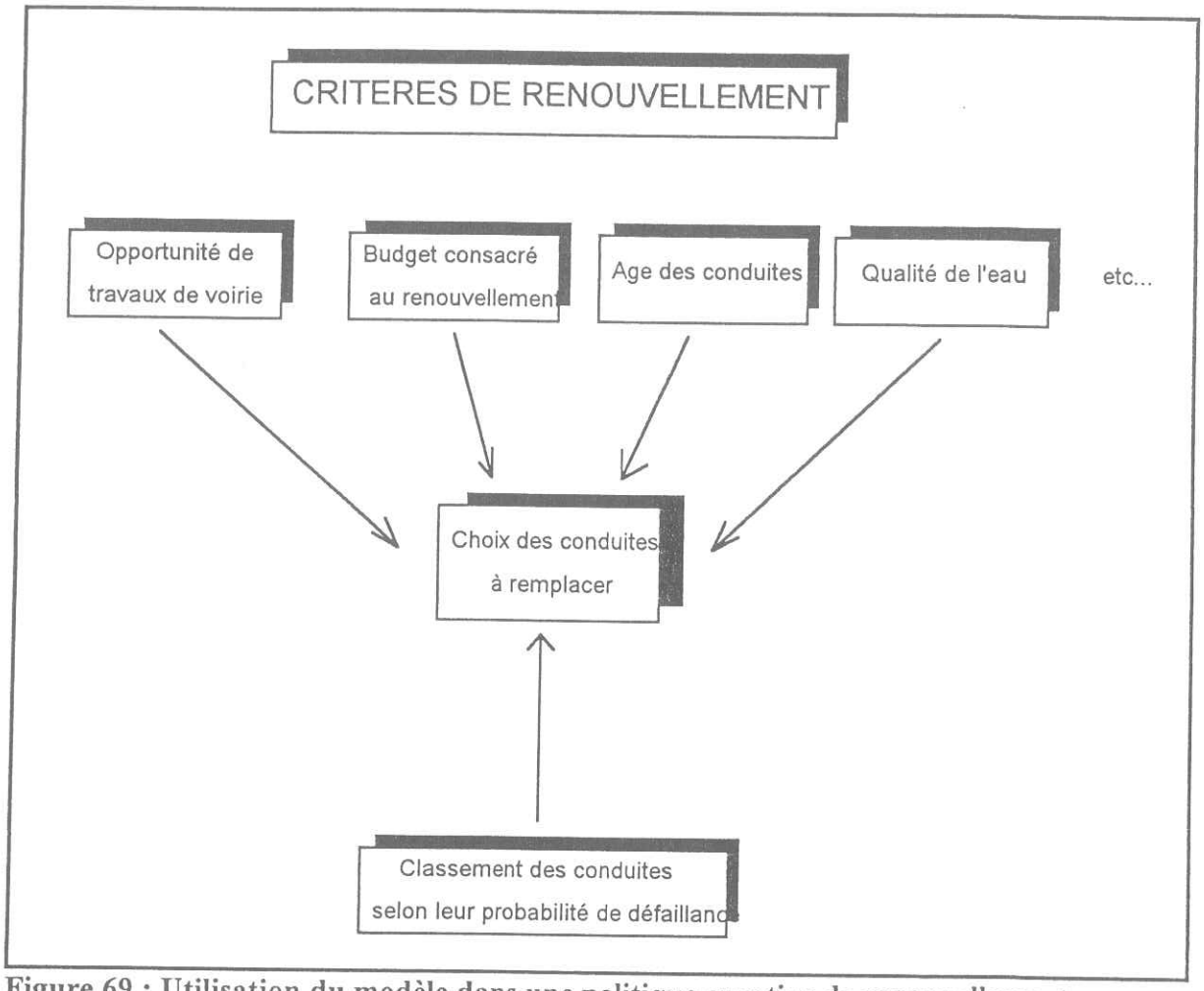


Figure 69 : Utilisation du modèle dans une politique curative de renouvellement

Il peut être utilisé dans le cadre d'une étude économique afin de définir le coût de maintien en service futur de la canalisation.

8.3.2.1. Méthode de Shamir et Howard [Shamir et al., 1979]

La méthode de Shamir et Howard a pour but de déterminer la date optimale de renouvellement en essayant de prendre en compte tous les coûts pouvant entrer en compte dans le maintien de service. Deux hypothèses sont faites pour cette analyse :

- on suppose que les nouvelles canalisations ne présentent pas de défaillances.
- on prend en considération la vie future des nouvelles canalisations avec un taux de ruptures qui augmente exponentiellement.

Deux coûts sont déterminés :

- $C_r(t)$, qui est le coût de remplacement du tronçon,
- $C_m(t)$, qui est le coût de maintien en service du même tronçon accumulé de la pose au temps t.

La date optimale de renouvellement t_r est telle que :

$$C_r(t_r)+C_m(t_r) \text{ soit minimum,}$$

C'est à dire telle que :

$$\frac{d(C_r+C_m)}{dt}(t_r)=0.$$

Dans les valeurs des coûts rentre en jeu l'actualisation. On aura donc :

$$C_r(t)=\frac{C_r^*(t)}{(1+R)^t}$$

et

$$C_m(t)=\frac{C_m^*(t)}{(1+R)^t}$$

où R correspond au taux d'actualisation.

La courbe de la figure 70 représente la courbe de Shamir et Howard.

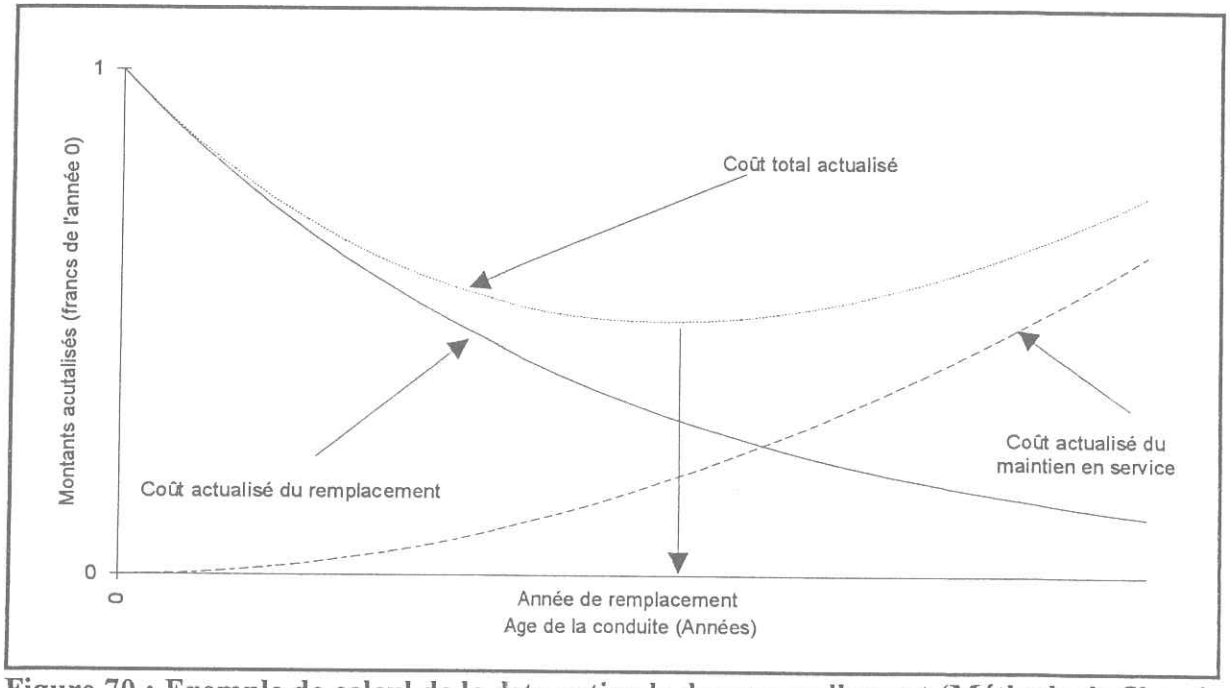


Figure 70 : Exemple de calcul de la date optimale de renouvellement (Méthode de Shamir et Howard) [Shamir et al., 1979]

Cette méthode a été utilisée dans le cadre d'une étude établie par J. El Naboulsi, en utilisant des coûts de réparation directs et indirects provenant de la Communauté Urbaine de Strasbourg. Le calcul du coût de maintien en service a été établi à partie du modèle de prévision tiré des données de la CEB et de la LED-Bordeaux.

Dans ce cas le coût de maintien en service a été calculé de la façon suivante :

$$C_m(t) = C_a \left[\begin{array}{l} \sum_{t_1=0}^{t} f_1(t_1) S_2(t-t_1) \left(\frac{1}{(1+r)^{t_1}} \right) \\ + \sum_{t_1=0}^{t} \sum_{t_2=t_1}^{t} f_1(t_1) f_2(t_2-t_1) S_3(t-t_2) \left(\frac{1}{(1+r)^{t_1}} + \frac{1}{(1+r)^{t_2}} \right) \\ + \sum_{t_1=0}^{t} \sum_{t_2=t_1}^{t} \sum_{t_3=t_2}^{t} f_1(t_1) f_2(t_2-t_1) f_3(t_3-t_2) S_4(t-t_3) \left(\frac{1}{(1+r)^{t_1}} + \frac{1}{(1+r)^{t_2}} + \frac{1}{(1+r)^{t_3}} \right) \\ + \ldots\ldots \end{array} \right]$$

C_a étant le coût d'une défaillance à l'année 0, r le taux d'actualisation.

Les courbes 71 et 72 représentent le calcul de la date optimale de renouvellement dans deux cas différents. Le premier inclut uniquement les coûts de la réparation (7000 F). Le second inclut les coûts indirects (174 000F), dont tous les coûts de dommages dus par exemple à la présence proche de cabinets médicaux.

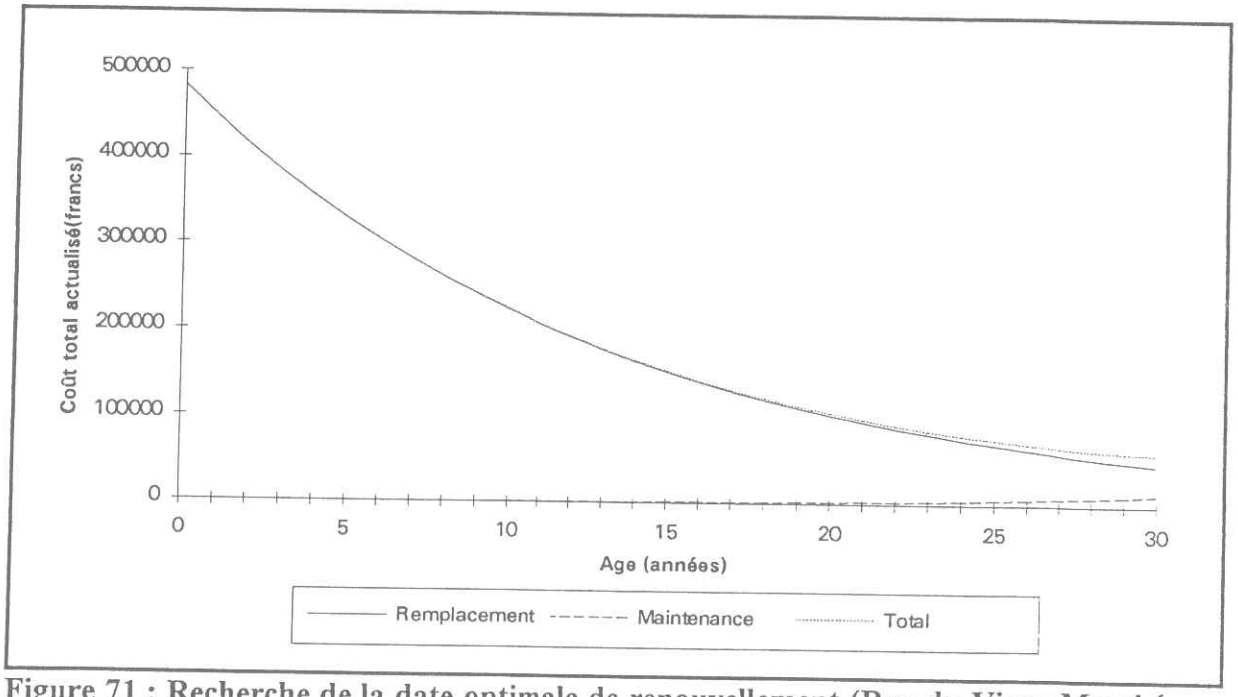


Figure 71 : Recherche de la date optimale de renouvellement (Rue du Vieux Marché aux Poissons, Hypothèse non favorable, sans coût indirect) [El Naboulsi, 1993]

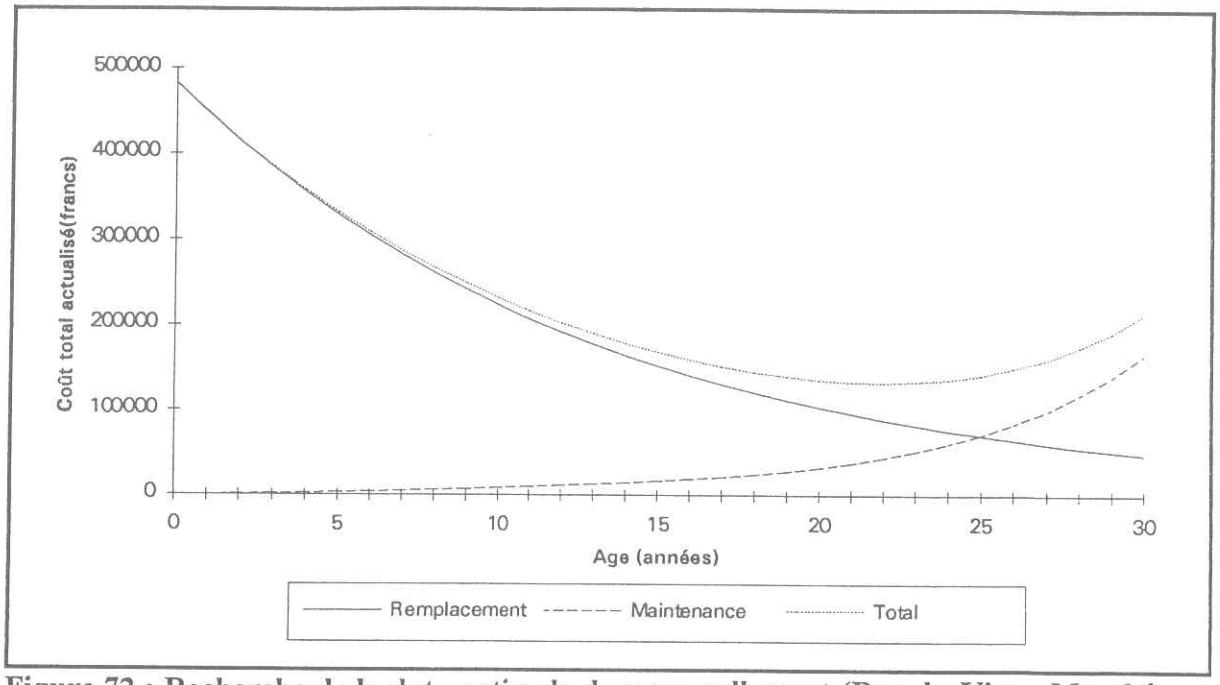


Figure 72 : Recherche de la date optimale de renouvellement (Rue du Vieux Marché aux Poissons, Hypothèse non favorable, avec coût indirect) [El Naboulsi, 1993]

Dans le premier cas la date optimale de remplacement se situe après 30 ans. Par contre dans le deuxième cas elle se situe autour de la 22ème année de la conduite. Ces cas concernent des conduites se trouvant dans les cas de défaillance les plus défavorables.

Cette méthode permet de calculer la date optimale de remplacement (Cf. Figure 73). D'autres coûts pourraient encore être pris en compte, notamment ceux concernant les autres facteurs de vieillissement cités au chapitre 1 (qualité de l'eau, fuites diffuses, diminution de la capacité de transport).

8.3.2.2. Utilisation dans la méthode de Herz [Herz, 1992]

Cette méthode vise à établir un modèle prévisionnel de besoin de renouvellement des conduites servant de base à une programmation budgétaire à long terme. Sur la base de données techniques provenant du service d'eau de la ville de Stuttgart, une espérance de vie est calculée selon le type de matériau utilisé.

Une fonction de survie est élaborée. Elle est donnée par la formule suivante :

$$y(x) = \frac{a+1}{a+e^{b(x-c)}} \text{ avec } x \geq c$$

$$y(x) = 1 \text{ avec } x < c$$

y étant le pourcentage de canalisations encore en vie, x l'âge de la conduite, a le paramètre de vieillissement, b le facteur de défaillance et c le temps de résistance.

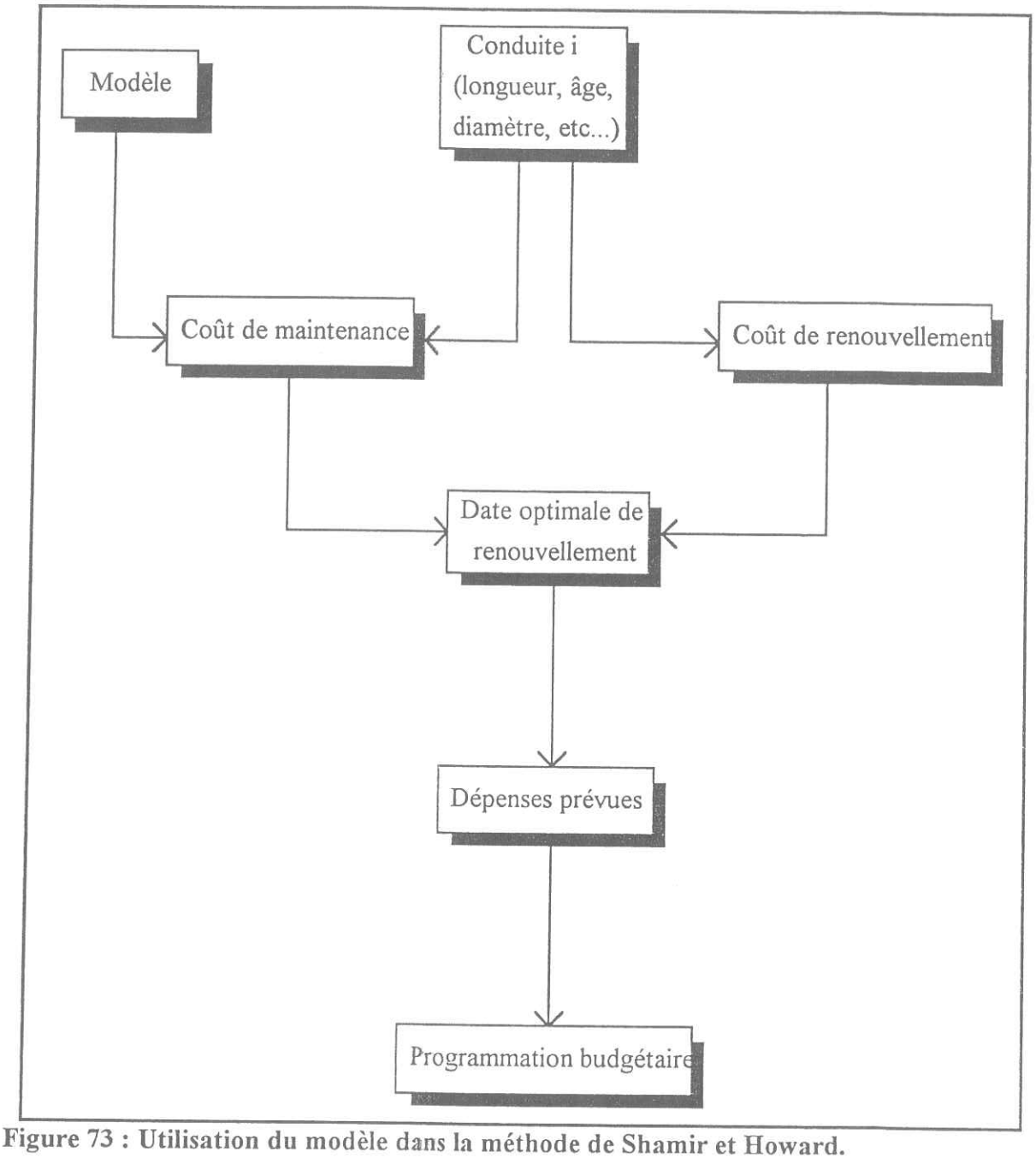


Figure 73 : Utilisation du modèle dans la méthode de Shamir et Howard.

La loi de répartition des durées de vie a pour durée de vie :

$$f(x) = \frac{(a+1)be^{b(x-c)}}{(a+e^{b(x-c)})^2} \text{ pour } x \geq c$$

$$f(x) = 0 \text{ pour } x < c$$

Le taux de défaillance correspondant au rapport entre la fonction de densité et la fonction de survie, on a :

$$z(x) = \frac{be^{b(x-c)}}{a+e^{b(x-c)}} \text{ pour } x \geq c.$$

Les courbes de la figure 74 représentent ces fonctions.

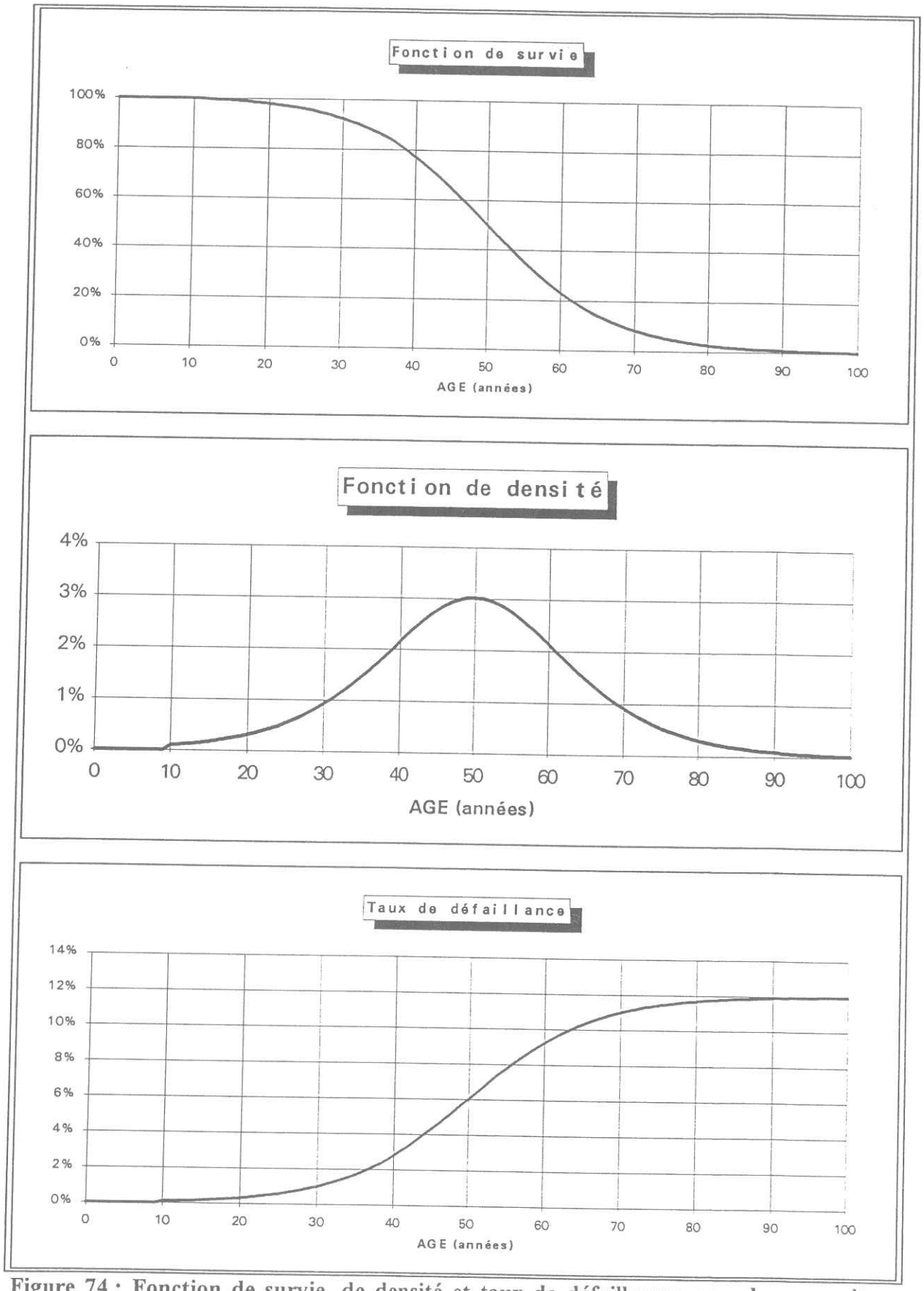


Figure 74 : Fonction de survie, de densité et taux de défaillances pour les paramètres a=120, b=0,12, c=10. [Herz, 1992]

Chaque canalisation est alors représentée selon ces trois paramètres a, b et c qui caractérisent sa durée de vie. Elle peut d'ailleurs voir varier ces paramètres le long de sa vie, lorsqu'elle passe d'une classe de vieillissement à une autre.

L'estimation de ces paramètres doit se faire en quatre temps :

- Mise en avant des connaissances empiriques,
- Utilisation de l'information stockée dans les banques de données,
- Utilisation d'un appareillage statistique,
- Correction des estimations par un apprentissage dynamique.

L'estimation de ces paramètres est encore en cours et pour l'instant une distinction n'a pu être réalisée que sur les matériaux.

Le modèle de Weibull pourrait fournir des indications intéressantes sur les paramètres a, b et c (Cf. Figure 75).

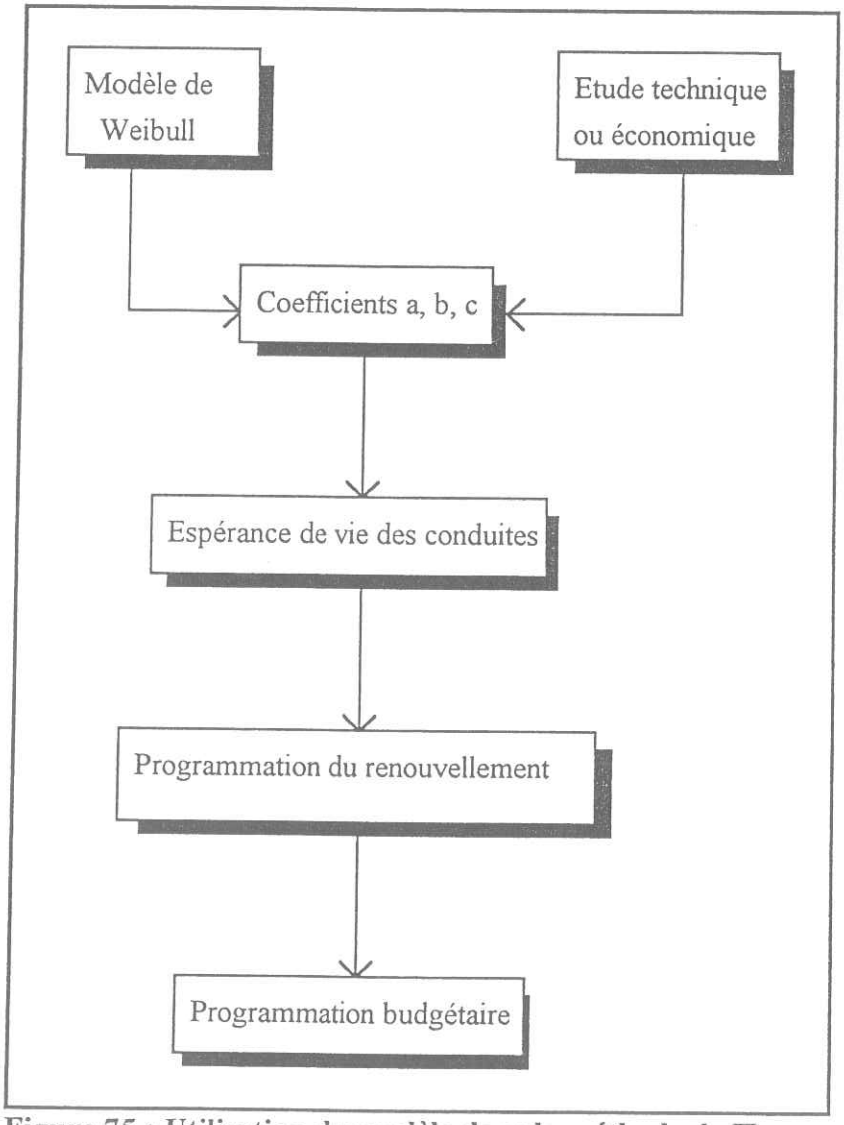


Figure 75 : Utilisation du modèle dans la méthode de Herz

Les services peuvent donc utiliser le modèle dans les différents types de politique de renouvellement, même si une politique prédictive semble préférable à une politique curative. Cependant il est impératif pour ces services de suivre auparavant la démarche proposée au paragraphe 8.1..

8.4. UTILISATION DANS L'ETUDE DE FIABILITE D'UN RESEAU

A travers cette étude probabiliste, il sera possible de fixer un risque d'interruption de service sur une conduite. En estimant pour chacune des conduites un risque de ce type, rattaché à son importance dans le bon fonctionnement du réseau (adduction ou distribution? nombre d'habitants desservis? présence de services médicaux? débit? etc...), il sera possible de déterminer un indice de fiabilité d'un réseau du point de vue de la sécurité.

Bien entendu, d'autres critères de fiabilité liés aux autres organes du réseau (réservoirs, pompes, etc.) devront rentrer en compte dans l'estimation de cet indice de fiabilité.

CONCLUSION

Les casses ou défaillances des conduites d'eau potable apparaissent comme le critère caractérisant leur vieillissement le plus simple à prendre en compte. Le modèle choisi pour prévoir leur apparition est un modèle statistique car il permet de prendre en compte des variables explicatives quantitatives, comme qualitatives.

Les données recueillies sur trois services d'eau (LED-Bordeaux, CEB et Bas-Rhin) et leur spécificité ont déterminé le choix d'un modèle basé sur l'analyse des données de survie utilisée en épidémiologie. Ce sont les modèles de Cox et de Weibull qui ont été utilisés.

Le test qui les accompagne a permis de définir les facteurs les plus influents. Ce sont pour la CEB et la LED-Bordeaux, le nombre de défaillances déjà subies, le diamètre et dans une moindre mesure le sol, l'année de pose (différenciation fonte grise-fonte ductile) et le facteur croisé trafic-chaussée. Pour les deux syndicats ruraux, où l'historique de défaillance est plus faible, seuls le nombre de défaillances déjà subies, le diamètre et le trafic sont influents.

Parmi les deux modèles, le modèle de Weibull, complètement paramétrique, semble mieux adapté, car il permet de prévoir à plus long terme les défaillances. La comparaison des résultats de ce modèle avec les données réelles sur Bordeaux a permis de faire deux constatations :

- l'hypothèse d'une même influence des facteurs quelque soit l'ordre de la défaillance est trop forte ;
- les casses ayant lieu après 15 ans ont une plus grande influence que celles ayant lieu avant 15 ans ; un nouveau facteur caractérisant le vieillissement de la conduite peut alors être pris en compte.

Le modèle pourrait être amélioré sur deux points. Tout d'abord avec des données plus précises concernant les défaillances, il serait possible de différencier les fuites sur tuyaux de celles sur joints, ce qui n'était pas le cas dans cette étude. On pourrait alors créer deux modèles spécifiques aux deux types de fuites, qui dans l'objectif du renouvellement n'ont pas la même signification. Ensuite le facteur vieillissement pourrait être affiné en prenant en compte, par exemple, l'écart entre les deux défaillances précedentes. Il caractériserait alors de manière plus précise le processus de vieillissement.

La longueur de l'historique des défaillances des services étudiés est assez considérable (35 ans pour la LED-Bordeaux, 44 pour la CEB et 17 pour le Bas-Rhin), ce qui n'est pas le cas généralement. Cependant la méthode d'apprentissage permet avec un minimum de données et à partir d'un modèle général de construire un modèle semi-local. Les services ayant de faibles historiques de maintenance pourront ainsi créer leur propre modèle.

L'élaboration d'un tel modèle nécessite dès à présent la création de bases de données sur les conduites, leurs défaillances et sur les facteurs influents. L'étude a d'ailleurs permis de préciser quels sont les facteurs essentiels à retenir dans ce cas. Elle a également montré la nécessité d'exploiter des tronçons de canalisations de longueur minimale. Ceci serait facilité par l'utilisation d'un Système d'Information Géographique, qui permettrait une meilleure localisation des conduites et de leurs défaillances, à condition que les relevés sur le terrain soient suffisamment précis.

Dans le cadre d'une politique de renouvellement, ce modèle peut être associé à une étude économique dans laquelle rentreraient en compte d'autres critères tels que l'opportunité de travaux de voirie ou le mode de financement possible. D'un point de vue technique, il sera également intéressant de prendre en compte les autres critères de vieillissement que sont la diminution de la capacité de transport de la conduite, la dégradation de la qualité de l'eau ou les fuites diffuses.

REFERENCES BIBLIOGRAPHIQUES

AGHTM-Commission "Distribution de l'eau", 1993 : "Etude sur le renouvellement des réseaux d'eau potable", TSM, n°11 bis, Novembre, pp. 1-37.

ANDREOU, S.A., MARKS, D.H., CLARK, R.M., 1987 : "A new methodology for modelling break failure patterns in deteriorating water distribution systems: theory and applications", Adv. Water Resources, Mars, vol.10, pp. 2-20.

ANDREOU, S.A., 1986 : "Predictive models for pipe break failures and their implications on maintenance planning strategies for deteriorating water distribution systems", Massachusetts Institute of Technology, 291 p..

BOUDOURESQUE, P., 1985 : "La dégradation des réseaux d'eau potable. Préventions et remèdes.", T.S.M.- L'EAU, Octobre, n°10 , pp. 427-437

CLARK, R.M., STAFFORD, C.L., GOODRICH, J.A., 1982 "Water distribution systems : a spatial and cost evaluation", Journal of the water resources planing and management division, Octobre, vol.108, pp. 243-256.

DEUTSCHER VEREIN DES GAS- UND WASSERFACHES, 1989 : "Entwurf DVGW-Hinweis: Entscheidungshilfen für die planung der Erneuerung oder der gründlichen Instandsetzung von Gas- und Wasserleitungen".

EISENBEIS, P., 1991 : "Méthodologie de prévision des défaillances des conduites d'eau potable. Etude de faisabilité.", Mémoire de DEA "Sciences et techniques de l'eau" ENGEES-ULP Strasbourg, 112 p..

ELNABOULSI, J., 1993 : "Le renouvellement des réseaux d'eau potable - Approche économique d'optimisation", Mémoire de DESS "Economie industrielle et décentralisation", Laboratoire Gestion des Services Publics ENGEES-CEMAGREF, 171 p.

HALPERN, O., PASCAL, O., 1988 : "Le renouvellement des réseaux urbains de distribution d'eau", Colloque international sur le coût et le prix de l'eau en ville, Décembre, pp.108-116.

HERZ, R., 1992 : "Alterung und Erneuerung von Infrastrukturbeständen - ein Kohortenüberlebensmodell", Veröffentlichung in Jahrbuch für Regionalwissenschaft, 22 p.

HILL, C., COM-NOUGUE, C., KRAMAR, A., MOREAU, T., O'QUIGLEY, J., SENOUSSI, R., CHASTANG, C., 1990 : "Analyse statistique des données de survie", INSERM, édition Médecine-Sciences Flammarion, 204 p..

KALBFLEISCH, J.D., PRENTICE, R.L., 1980 : "The statistical analysis of failure time data", John Wiley & Sons, 321 p..

KOTTMANN, A., 1988 : "Ist das Alter von Graugußrohren ein Kriterium für die Erneuerung einer Leitung?", 3R International, N°1, pp.55-58.

KOTTMANN, A., 1978 : "Über dir Ursachen von Rohrbrüchen in Versorgungsleitungen", Dissertation, Stuttgart.

KREMER, M., 1990 : "Vieillissement des canalisations. Exemples de politiques de renouvellement en France et en RFA.", mémoire de DEA "Sciences et techniques de l'eau" ENGEES-ULP Strasbourg, 130 p..

MARKS, D.H., ANDREOU, S.A., JEFFREY, L., PARK, C., ZASLAVSKY A., 1987 : "Statistical models for water main failures", Water Engineering Research Laboratory, USEPA, 108 p..

MORUZZI, L., 1984 : "Reasons for pipe damage", Congrès de l'AIDE (Monastir,1984) Water supply, N°3/4, pp.SS16-1-SS16-5.

O'DAY, D.K., WEISS, R., CHIAVARI, S., BLAIR, D., 1989 : "Water main evaluation for rehabilitation/replacement", Guidance manual of AWWA research foundation, 182 p..

SAPORTA, G., 1990 : "Probabilités, analyse des données et statistiques", Editions Technip, 493 p..

SHAMIR, U., HOWARD, C.D.D., 1979 : "An analytic approach to scheduling pipe replacement", Journal of AWWA, Mai, vol.71, pp 248-258.

STALDER, F., HUNKELER, F., 1986 : "Korrosion durch Streuströme", Gaz-Eaux-Eaux Usées, n° 66, pp. 821-826.

UNION DES INDUSTRIES ET ENTREPRISES DE L'EAU ET DE L'ENVIRONNEMENT, 1991 : "Les équipements de l'eau des collectivités locales", brochure, Mars, 4 p..

ZEDGUI, L., 1993 : "Diagnostic et étude des défaillances des réseaux d'eau potable des Syndicats de Kronthal et de Kochersberg", Mémoire de Mastère - ENGEES, 72 p..

BIBLIOGRAPHIE

AWWA, 1969 : "Replacement of water distribution mains-Committee report", Journal of AWWA, Septembre, vol.61, N°9, pp. 417-422.

BREMOND, B., EISENBEIS, P., COMMENGES, D., 1992 : "Outil de prévision des défaillances des conduites d'eau potable, aide à la décision du renouvellement", T.S.M.-L'EAU, Octobre, N°10, pp. 511-516.

BUTLER, D., WEST, J.W. 1987 : "Leakage prevention and system renewal", Symposium papers of Pipeline Management 87, pp. 2.1-2.17.

CARDER D.R., TAYLOR, M.E., 1983 : "Modelling the effects of traffic loading on small-diameter rigid pipelines", Pipe and pipelines international, Mars-Avril, pp. 12-14.

CLARK, R.M., STAFFORD, C.L., GOODRICH, J.A., 1984 : "Repair, replacement, and renovation of drinking water systems", ASCE water ressources planing and management annual conference, Mai, pp 1-5.

CORLESS, J.P., 1982 : "Managing a comprehensive water main reconstruction program", Journal of AWWA, Novembre, pp.560-567.

DRESSLER, H., "Wasserverlustbekämpfung in einem Wasserwerk mittlerer Größe", Gas/Wasser/Wärme, pp. 83-88.

DROESBEKE, J.J., FICHET, B., TASSI, P., "Analyse statistique des données de survie", Economica, 282 p..

GOODRICH, J.A., ADAMS, J.Q., CLARK, R.M., 1985 : "Data base development and analysis for water distribution systems", Hydraulics and Hydrology in the small computer age, Août, vol.1, pp. 336-340.

GOURIEROUX, C., 1989 : "Econométrie des variables qualitatives", Economica collection "Economie et statistiques avancées", 430 p..

GUMMOW, R.A., 1984 : "The corrosion of municipal iron watermains", National association of corrosion enginers, Mars, N°3, pp.39-42.

HERSCH, P., 1985 : "Deciding to rehabilite, repair, or replace", Water engineering and management, Mai, vol.132, N°50-53.

HERZ, R., HOCHSTRATE, K., "Long term requirements and policies for water supply network renewal", Water supply, vol.6, pp. 63-70.

HOLTSCHULTE, H., 1983 : "The acquisition and evaluation of technical data and services, as decision aids for maintenance measures in pipe networks", Water supply, vol.2, pp. B157-169.

HUDSON, W.D., "Increasing water system efficiency trough control of unaccounted-for water, Journal of AWWA, vol.70, N°7, pp.362-365.

KOTTMANN, A., GOCKEL, B., 1988 : "Planungsregeln für Wasserleitungen und Wasserrohrnetze", Wasser-Abwasser, N°129, pp. 701-702.

MOYER, E.E., MALE, J.W., MOORE, I.C., HOCK, J.G., 1983 : "The economics of leak detection and repair- a case study", Janvier, vol.75, N°1, pp.29-34.

NEWPORT, R., 1981 : "Factors influencing the occurence of bursts in iron water mains", Aqua, N°3, pp. 274-278.

O'DAY, D.K., 1987 : "Water distribution record keeping and planning approaches: current utility practices", Journal NEWWA, Juin, vol.101, N°2, pp. 145-168.

O'DAY, D.K., FOX, C.M., HUGUET, G.M., 1980 : "Aging urban water systems: a computerized case study", Public Works, Août, pp. 61-64 + p.111.

O'DAY, D.K., 1982 : "Organizing and analyzing leak and break data for making main replacement decisions", Journal of AWWA, Novembre, vol.74, N°11, pp.588-594.

PONT-A-MOUSSON, 1989 : "Modélisation du comportement des réseaux d'eau dans le temps", Centre de recherche Pont-à-Mousson,Mars.

REED, E.C., "Report on water losses", Aqua, pp. 178-191.

REY, A., 1989 : "Alimentation en eau des communes rurales", Courants, Septembre-Octobre,N°5, pp. 37-40.

STACHA, J.H., 1978 : "Criteria for pipeline replacement", Journal of AWWA, Mai, pp. 256-258.

SULLIVAN, J.P., 1982 : "Maintaining aging systems - Boston's approach", Journal of AWWA, Novembre, vol.74, N°11, pp. 555-559.

VAN DEN HOVEN, T.J.J., "Methods to determine quantitative and qualitative reliability of water distribution systems", Water Supply, vol.6, pp.13-16.

WALSKI, T.M., PELLICCIA, A., 1982 : "Economic analysis of water main breaks", Journal of AWWA, Mars, vol. 74, N°3, pp. 140-147.

WALSKI, T.M., 1982 : "Economic analysis of rehabilitaion of water mains", Journal of the water resources planing and management division, Octobre, vol.108, pp. 296-308.

WALSKI, T.M., 1984 : "Selecting an economical water main rehab", Civil engineering/ASCE, Septembre, pp.68-70.

WALSKI, T.M., 1987 : "Replacement rules for water mains", Journal of AWWA, Novembre, vol.79, N°11, pp.33-37.