Elaboration d'un modèle long terme pour optimiser le renouvellement des canalisations d'eau potable. Application au SEDIF : rapport de deuxième année de thèse (1/3)

A. Large, M. Tomasian

▶ **To cite this version:**

A. Large, M. Tomasian. Elaboration d'un modèle long terme pour optimiser le renouvellement des canalisations d'eau potable. Application au SEDIF : rapport de deuxième année de thèse (1/3). [Rapport de recherche] irstea. 2014, pp.132. hal-02603651

HAL Id: hal-02603651
https://hal.inrae.fr/hal-02603651

Submitted on 16 May 2020

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

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

Rapport dans le cadre du projet « Optimeau » à partir des données du SEDIF

Elaboration d'un modèle long terme pour optimiser le renouvellement des canalisations d'eau potable

Application au SEDIF

Rapport de deuxième année de thèse (1/3)

Par Aurore LARGE, Max TOMASIAN

IRSTEA, Groupement de Bordeaux
Unité de Recherche ETBX
Equipe GPIE

50, avenue de Verdun, Gazinet
33612 Cestas cedex

1ere version Nov 2014
révisée le 03/09/2015

Encadrement : Dr. Y. Le Gat, Dr. S.M. Elachachi, Ing. E. Renaud & Pr. D. Breysse

Remerciements

Nous tenons tout d'abord à remercier les trois autorités organisatrices qui financent le projet « optimeau », nous fournissent leur données et nous apportent un appui technique :

- le service des eaux de la ville de Lausanne (eauservice Lausanne) en Suisse,
- la communauté urbaine de Lyon (Grand Lyon) et
- le syndicat des eaux d'île de France (SEDIF).

Pour le financement indirect du projet « optimeau », merci :

- au fonds de recherche pour l'eau (Forschungsfonds Wasser = FOWA) de la Société Suisse de l'Industrie du Gaz et des Eaux (SSIGE),
- à l'agence de l'eau Rhône Méditerranée et Corse et
- à l'agence de l'eau Seine-Normandie.

Nous adressons des remerciements tout particuliers aux membres de notre comité d'encadrement : Dr. Yves LE GAT, Dr. Sidi-Mohammed ELACHACHI, Ing. Eddy RENAUD et Pr. Denis Breysse. En effet venant chacun de disciplines différentes la communication a parfois été difficile mais les complémentarités ont été d'une grande richesse.

Un grand merci à tous les participants aux comités de suivi du projet « optimeau » (cf. Tableau 1) pour leurs réponses à nos questions et leurs regards critiques tant sur le point pratique que théorique.

Prénom NOM	Structure	Pays	Statut	Présent com suivi 1	Présent com suivi 2	Présent com suivi 3
M. Yves LE GAT	IRSTEA	France	Encadrement	Oui	Oui	Oui
M. Sidi-Mohammed ELACHACHI	Université Bordeaux (Génie Civil)	France	Encadrement	Oui	Oui	Oui
M. Eddy RENAUD	IRSTEA	France	Encadrement	Oui	Oui	Oui
M. Denys BREYSSE	Université Bordeaux (Génie Civil)	France	Encadrement	Oui	Oui	Non
M. Sébastien APOTHELOZ	Eauservice Lausanne	Suisse	Terrain d'étude	Oui	Non	Non
Mme Linda VIGUET	Eauservice Lausanne	Suisse	Terrain d'étude	Oui	Oui	Non
Mme Christelle SANZ	Eauservice Lausanne	Suisse	Terrain d'étude	Non	Oui	Oui
Mme Lisa SCHOLTEN	EAWAG	Suisse	Experte	Non	Oui	Non
M. Joao LEITAO	EAWAG	Suisse	Expert	Non	Oui	Non
Mme Anne LANG	SED Ile de France	France	Terrain d'étude	Oui	Non	Oui
M. Bruno CHOUX	SED Ile de France (Hydratech)	France	Terrain d'étude	Oui	Non	Non
M. Olivier CHESNEAU	SED Ile de France	France	Terrain d'étude	Oui	Non	Oui
Mme Claire LEFORT	SED Ile de France	France	Terrain d'étude	Non	Oui	Non
M. Sylvain CHARRIERE	SED Ile de France	France	Terrain d'étude	Non	Non	Oui
M. Christophe PERROD	SED Ile de France	France	Terrain d'étude	Non	Non	Oui
M. Sébastien BUTTOUDIN	Véolia Ile de France	France	Expert	Non	Non	Oui
Mme Anne-Claire SANDRAZ	Véolia Ile de France	France	Experte	Oui	Oui	Oui
M. Didier FANGEAT	Grand Lyon	France	Terrain d'étude	Oui	Oui	Oui
Mme Elisabeth SIBEUD	Grand Lyon	France	Terrain d'étude	Non	Oui	Non
M. Pascal LE GAUFFRE	INSA Lyon	France	Expert	Non	Oui	Oui
M. David POINARD	Véolia Gd Lyon	France	Expert	Non	Oui	Oui
M. Frédéric IMMEDIATO	AERMC	France	Expert	Non	Oui	Non
Mme Caty WEREY	IRSTEA-ENGEES	France	Experte	Oui	Oui	Non
Mme Sylvie FERRARI	Université Bordeaux (Economie)	France	Experte	Oui	Non	Non
M. Vincent COUALLIER	Université Bordeaux (Statistiques)	France	Expert	Oui	Oui	Oui

Tableau 1 : Liste des membres du comité de suivi du projet « optimeau »

Merci à toutes les autres personnes que nous avons interviewées : M. Gildas Lévesque (SEDIF, responsable SIG), M. Florent Alix (SEDIF, contrôle, suivi des travaux, exécution des travaux), M. Pierre Chopard (SEDIF, directeur des travaux), M. Bruno Choux (Bureau d'étude hydratech,

assistance à maîtrise d'ouvrage au SEDIF), M. Pierre Chopard, Mme Julie Tastet (SEDIF stagiaire gestion patrimoniale) et Mme Morgane Le Quiniou (Véolia eau, direction technique).

Nous remercions toute l'équipe GPIE de l'IRSTEA, et le département GCE de l'université de Bordeaux, pour leur aide, leur soutien et leur disponibilité.

Résumé

Dans les pays développés, l'eau potable est distribuée au domicile des usagers. Ce confort requiert un long linéaire de réseau de forte valeur patrimoniale. Pour optimiser ressources et performances, il est important d'essayer de renouveler les tronçons au meilleur moment possible. Dans ce cadre, ce rapport présente les modèles « court terme » (1 à 3 ans) et « long terme » (> 30 ans) actuellement employés au SEDIF. Les processus « court terme » semblent pertinents, mais les méthodes « long terme » restent frustes. Ce rapport propose de mettre en place une nouvelle approche « long terme » pour obtenir une vision globale du patrimoine de canalisations et rationaliser sa gestion. Cette approche est testée à partir des données réelles du SEDIF, le plus grand syndicat d'eau potable français.

Dans un premier temps nous estimons la courbe de survie empirique des âges à la mise hors service des canalisations observés entre 1995 et 2012. Après analyse et traitement de cette courbe puis création de scénarios prospectifs nous déduisons des indicateurs de réalisation tels que les futurs taux de renouvellement, et des indicateurs financiers tels que les montants futurs à investir.

Sommaire

Introduction 15

I Présentation du SEDIF 17

I.1 Territoire, mode de gestion et histoire 17

I.2 Caractéristiques des canalisations 18
I.2.1 Origine des données 18
I.2.2 Caractéristiques des linéaires des tronçons 19
I.2.3 Comparaison du linéaire en service et hors service en fonction des dates de pose 21
I.2.4 Période d'observation du linéaire hors service 22
I.2.5 Les matériaux des canalisations 24
I.2.6 Les diamètres des canalisations 25

I.3 Caractéristiques des défaillances 26
I.3.1 Caractéristiques des défaillances FR (Fuites Réparées) 26
I.3.2 Les défaillances Qualitatives Réparées 29

I.4 Les données environnementales 30

II Vue globale du modèle long terme que l'on souhaite développer dans « optimeau » 32

III Présentation de la gestion du patrimoine canalisation du SEDIF 34

III.1 Présentation du renouvellement passé du SEDIF 34

III.2 Méthode décisionnelle de « court terme » (≤ 3 ans) pour renouveler les tronçons au SEDIF 35
III.2.1 Méthode décisionnelle de « court terme » employée par le SEDIF 35
III.2.2 Méthode décisionnelle de « court terme » employée par le VEDIF 37

III.3 Méthode décisionnelle long terme (> 70 ans) 39
III.3.1 Méthode « Cador » 39
III.3.2 Prévision de l'évolution de l'âge moyen du réseau 41
III.3.3 Linéaire à renouveler fixe à atteindre 42
III.3.4 Patrimoine expert 43

IV Méthode pour le calcul de la survie empirique passée 45

IV.1 Préliminaires 45

IV.2 Nettoyage des données 45
IV.2.1 Nettoyage des données tronçons 45
IV.2.2 Nettoyage des données défaillances FR 47

IV.3 Nos données tronquées à gauche, censurées à droite 47

IV.4 Survie empirique des tronçons hors service et en service 48
IV.4.1 Kaplan-Meier simple pour données censurées à droite 48
IV.4.2 Kaplan-Meier élargi (en effectif) pour données censurées et tronquées 50
IV.4.3 Kaplan-Meier élargi (en linéaire) pour données censurées et tronquées 51

IV.5 Obtention d'une fonction continue à partir de la fonction de survie empirique discontinue 52
IV.5.1 Ajustement d'un modèle paramétrique classique à notre courbe de survie empirique 52
IV.5.2 Estimation des points inexistants par interpolation linéaire 53

IV.6 Modèle de Cox 54

IV.7 Modèle de Weibull avec covariables au tronçon 55

V Calcul de la survie passée, application aux données du SEDIF 58

V.1 Courbes de survie sans partitionnement 58
V.1.1 Comparaison des courbes de survie en linéaire et en effectif 58
V.1.2 Courbe de survie tous matériaux, tous diamètres en linéaire 59

V.2 Courbes de survie avec partitionnement 61
V.2.1 Comparaison des courbes de survie pour différents diamètres 61
V.2.2 Comparaison courbes de survie pour différentes longueurs 63
V.2.3 Comparaison courbes de survie pour différents matériaux (en linéaire) 65
V.2.4 Comparaison courbes de survie pour des tronçons avec ou sans défaillances 66

V.3 Courbe de survie avec trois fenêtres d'observation différentes 68

V.4 Résultats du modèle de Cox 70

V.5 Résultats modèle de Weibull avec covariables au tronçon 71

VI Scénario statu quo, calcul d'indicateurs clés au SEDIF 75

VI.1 Avec une seule courbe de survie 75
VI.1.1 Courbe de survie prospective 75
VI.1.2 Linéaires futurs à installer 76
VI.1.3 Taux de renouvellement 78
VI.1.4 Age moyen du réseau 78
VI.1.5 Coûts d'investissement 79

VI.2 Avec plusieurs courbes de survies (partitionnement) 80
VI.2.1 Courbes de survie prospectives avec partition par diamètre 80
VI.2.2 Linéaires futurs à installer 82
VI.2.3 Taux de renouvellement 83
VI.2.4 Age moyen du réseau 84
VI.2.5 Coûts d'investissement 84

Conclusion 88

VII Bibliographie 89

Annexe 1: Fiche intervention pour une défaillance, exemple 92

Annexe 2: Nettoyage du fichier défaillances FR 93

Annexe 3: Méthode de nettoyage de la variable type joints 95

Annexe 4: Méthode de nettoyage de la variable technique de pose 96

Annexe 5: Méthode de nettoyage de la variable emplacement 97

Annexe 6: Méthode de nettoyage de la variable type appui 98

Annexe 7: Les critères de renouvellement des canalisations par le VEDIF sur le territoire du SEDIF en 2010 99

Annexe 8 : Définition de la censure et de la troncature 102

Annexe 9 : Méthode pour calculer la survie des tronçons hors service = « survie brute » 103

Annexe 10 : Description de la méthode de Turnbull 104

Annexe 11 : Description de la méthode d'optimisation de Nelder & Mead 106

Annexe 12 : Validation des méthodes de Kaplan-Meier et de Turnbull, démonstration en effectif 109

Annexe 13 : Validation des méthodes de Kaplan-Meier et de Turnbull, démonstration en Linéaire 113

Annexe 14: Courbe de survie des tronçons partitionnés par matériaux 115

Annexe 15 : Le modèle de Weibull à covariables est un modèle de Cox particulier 120

Annexe 16: Corrélation entre longueur et diamètre 121

Annexe 17: Résultats Weibull avec covariables de la survie de Weibull avec covariables 122

Annexe 18: Indice *TP10A*, canalisation, égouts, assainissement et adduction d'eau avec fourniture de tuyaux 123

Annexe 19: Courbe de survie avec d'autres stratifications 125

Annexe 20: Conseils d'un point de vue statistique pour la gestion des données 132

Table des illustrations

Figures

FIGURE 1 : L'IDEAL DE L'ARTICULATION ENTRE LES DIFFERENTS OUTILS FONCTIONNANT A DIFFERENTES ECHELLES TEMPORELLES 16
FIGURE 2: TERRITOIRE COUVERT PAR LE SEDIF EN 2013 17
FIGURE 3 : EXTRAIT DU JEU DE DONNEES DES TRONÇONS 18
FIGURE 4 : NOMBRE DE TRONÇONS EN FONCTION DE LEUR LONGUEUR 19
FIGURE 5 : FREQUENCE CUMULEE DU NOMBRE DE TRONÇONS (EN SERVICE ET HORS SERVICE) EN FONCTION DE LEUR LONGUEUR (EN METRE) (100% = 287 716 TRONÇONS) 20
FIGURE 6 : DIAGRAMME DU LINEAIRE (KM) EN SERVICE AU SEDIF AU 31/12/2012 EN FONCTION DES DATES DE POSE (8380.6 KM AU TOTAL) 21
FIGURE 7 : HISTOGRAMME DU LINEAIRE EN SERVICE (8 296 KM) ET HORS SERVICE (HS) (680 KM) AU 31/12/2012 AYANT UNE DATE DE POSE CONNUE (EN ROUGE NE SONT REPRESENTES QUE LES TRONÇONS MIS HS ENTRE [1995, 2012]). 21
FIGURE 8 : HISTOGRAMME DU LINEAIRE HORS SERVICE ARCHIVE AU 01/01/2013 23
FIGURE 9 : LES MATERIAUX DU RESEAU EN SERVICE AU 31/12/2012 24
FIGURE 10 : HISTOGRAMME DU LINEAIRE EN FONCTION DES DDP DES DIFFERENTS MATERIAUX AU 31/12/2012 .. 24
FIGURE 11 : LINEAIRE DE TRONÇONS EN FONCTION DE LEUR DIAMETRE 25
FIGURE 12 : FREQUENCE CUMULEE DU NOMBRE DE TRONÇONS (EN SERVICE ET HORS SERVICE) EN FONCTION DE LEUR DIAMETRE (EN MINI METRE) (100% = 287 716 TRONÇONS) 25
FIGURE 13 : FREQUENCE CUMULEE DU NOMBRE DE TRONÇONS EN FONCTION DE LEUR DIAMETRE 26
FIGURE 14 : SCHEMA RECAPITULATIF DU LIEN ENTRE FUITES, CASSES, TROUS ET DEFAILLANCES FR 26
FIGURE 15 : EXTRAIT DU FICHIER SIG DEFAILLANCES FR AU SEDI 28
FIGURE 16 : NOMBRE DE DEFAILLANCES ARCHIVEES PAR AN, JUSQU'A MARS 2013 28
FIGURE 17 : REPARTITION DES TYPES DE DEFAILLANCES FR SUR TRONÇON AU SEDIF AU COURS DU TEMPS 29
FIGURE 18 : SCHEMA DES TYPES DE DEFAILLANCES « FUITES REPAREES » (POINARD 2006) 29
FIGURE 19 : SCHEMA DE LA DEMARCHE GLOBALE DU MODELE « LONG TERME » OPTIMEAU 32
FIGURE 20 : TAUX DE RENOUVELLEMENT ANNUEL, CALCULE A PARTIR DU LINEAIRE HORS SERVICE DANS LE SIG 34
FIGURE 21 : TAUX MOYEN DE RENOUVELLEMENT SUR 5 ANS 35
FIGURE 22 : LES ETAPES DE LA DECISION DE LA MISE HORS SERVICE DE TRONÇONS AU SEDIF A COURT TERME (EN 2013) 37
FIGURE 23 : LES ETAPES DE LA DECISION DU VEDIF DE LA MISE HORS SERVICE DE TRONÇONS A COURT TERME (EN 2010 (LES POURCENTAGES INDIQUES SONT ISSUES DU PREVISIONNEL DE LA FIGURE SUIVANTE) 38
FIGURE 24 : LES CAUSES DE RENOUVELLEMENT DE CANALISATIONS PAR VEDIF DANS LE PROGRAMME PREVISIONNEL EN 2010 SOURCE : (BRUYERON & RE 2012A; BRUYERON & RE 2012B) 38
FIGURE 25 : LES DUREES MOYENNES DE MAINTIEN EN SERVICE SELON LES EXPERTS DU SEDIF (BRUYERON & RE 2012B) 39
FIGURE 26 : FONCTION DE SURVIE DES TRONÇONS EN FONTE GRISE SELON LES EXPERTS DU SEDIF. 40
FIGURE 27 : PREVISION DES BESOINS EN RENOUVELLEMENT SELON LA METHODE « CADOR »(BRUYERON & RE 2012B) 40
FIGURE 28 : LES ETAPES DE LA DECISION DE LA MISE HORS SERVICE DE TRONÇONS DU SEDIF SUR DU LONG TERME 41
FIGURE 29 : PREVISION DE L'EVOLUTION DE L'AGE MOYEN DU RESEAU (LARGE 2013A) 41
FIGURE 30 : COMPARAISON DE PREVISIONS A LA REALITE (LARGE 2013A) 42
FIGURE 31 : REPARTITION DU LINEAIRE A RENOUVELER ENTRE SEDIF ET VEDIF JUSQU'EN 2022 POUR LE RESEAU DE DISTRIBUTION (LARGE 2013A) 42
FIGURE 32 : EXEMPLE DE SORTIE DU LOGICIEL PATRIMOINE EXPERT (SEDIF 2013B) 44
FIGURE 33 : LINEAIRE EN FONCTION DES DATES DE POSE AVANT NETTOYAGE 45
FIGURE 34 : DISPOSITIONS POSSIBLES DES TRONÇONS VIS-A-VIS DE LA FENETRE D'OBSERVATION, ILLUSTRATION DE LA TRONCATURE A GAUCHE ET DE LA CENSURE A DROITE 48
FIGURE 35 : COURBES DE SURVIE PASSES EN LINEAIRE ET EN EFFECTIF, SEDIF, TOUS TRONÇONS CONFONDUS ENTRE [1995, 2012] 58

FIGURE 36 : COURBE DE SURVIE EN NOMBRE DE TRONÇONS ACCOMPAGNEE DE SON INTERVALLE DE CONFIANCE, TOUS TRONÇONS CONFONDUS. 59
FIGURE 37 : COURBES DE SURVIE EN LINEAIRE, TOUS TRONÇONS CONFONDUS 60
FIGURE 38 : COURBES DE SURVIES PASSEES [1995-2012] EN EFFECTIF (TRONÇONS PARTITIONNES EN 2 SOUS-ENSEMBLES RESEAU DE TRANSPORT ET RESEAU DE DISTRIBUTION) ET LEURS INTERVALLES DE CONFIANCES (EN GRIS FONCE) 62
FIGURE 39 : COURBES DE SURVIE PASSES [1995, 2012] EN LINEAIRE (TRONÇONS PARTITIONNES SELON LE DIAMETRE) ET LEURS SURVIES DE WEIBULL AJUSTEES 62
FIGURE 40 : COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 3 SOUS-ENSEMBLES : PETITES, MOYENNES ET GRANDES LONGUEURS) ET LEURS INTERVALLES DE CONFIANCE (EN GRIS FONCE) ... 63
FIGURE 41 : REPARTITION DU LINEAIRE EN SERVICE (FIN 2012) DU SEDIF EN 3 SOUS-ENSEMBLES SELON LA LONGUEUR DES TRONÇONS 64
FIGURE 42 : COURBES DE SURVIE EN LINEAIRE (TRONÇONS PARTITIONNES EN 3 SOUS-ENSEMBLES : PETITES, MOYENNES ET GRANDES LONGUEURS) 64
FIGURE 43 : COURBES DE SURVIE EN LINEAIRE POUR DIFFERENTS MATERIAUX 65
FIGURE 44: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 2 SOUS-ENSEMBLES : AVEC ET SANS DEFAILLANCE) ET LEURS INTERVALLES DE CONFIANCE (EN GRIS FONCE) 66
FIGURE 45: TAUX DE DEFAILLANCE ANNUEL SUR TRONÇONS EN SERVICE (EN ROUGE) ET TAUX DE DEFAILLANCE SUR TRONÇONS HORS SERVICE (EN BLEU) 67
FIGURE 46 : COURBES DE SURVIE AU LINEAIRE, TOUS TRONÇONS CONFONDUS, POUR TROIS FENETRES D'OBSERVATION DIFFERENTES 68
FIGURE 47 : EQUATION A POUR 4 000 TRONÇONS SELECTIONNES PAR TIRAGE ALEATOIRE SANS REMISE ET EQUATION A MOYENNE SUR TOUS LES TRONÇONS 73
FIGURE 48 : COURBES DE SURVIE DE KAPLAN-MEIER TOUS TRONÇONS CONFONDUS ET EQUATION A 73
FIGURE 49 : COURBE DE SURVIE PROSPECTIVE (ROUGE) OBTENUE A PARTIR DE LA COURBE EMPIRIQUE (ISSUE DE KM EN VERT) 75
FIGURE 50 : RECONSTITUTION DU LINEAIRE PASSE ET PREVISION DES LINEAIRES FUTURS JUSQU'EN 2050, PHOTOGRAPHIE DE LA SITUATION EN 2050, RESEAUX DE TRANSPORT ET DE DISTRIBUTION. 77
FIGURE 51 : RECONSTITUTION DU LINEAIRE PASSE ET PREVISION DES LINEAIRES FUTURS JUSQU'EN 2120, PHOTOGRAPHIE DE LA SITUATION EN 2120, RESEAUX DE TRANSPORT ET DE DISTRIBUTION. 77
FIGURE 52 : PREVISION DES TAUX DE RENOUVELLEMENT FUTURS DE 2013 JUSQU'EN 2120 AU SEDIF 78
FIGURE 53 : PREVISION DES TAUX DE RENOUVELLEMENT FUTURS DE 2013 JUSQU'EN 2120 AU SEDIF AVEC LE SCENARIO « SATU QUO » 79
FIGURE 54 : PREVISIONS DES BUDGETS NECESSAIRES POUR LE RENOUVELLEMENT ENTRE 2013 ET 2120 AVEC 2 TAUX D'ACTUALISATION DIFFERENTS, RESEAUX DE TRANSPORT ET DE DISTRIBUTION 80
FIGURE 55 : COURBES DE SURVIE PROSPECTIVES EN LINEAIRE (CF. TRAIT FIN) OBTENUE A PARTIR DES COURBES EMPIRIQUE DE KAPLAN-MEIER ET AJUSTE (EQUATION DE WEIBULL), TOUS TRONÇONS CONFONDUS, PARTITIONNES EN 3 SOUS-ENSEMBLES (PETITS, MOYENS ET GROS DIAMETRE) 81
FIGURE 56 : RECONSTITUTION DU LINEAIRE PASSE ET PREVISION DES BESOINS (EN LINEAIRE) FUTURS JUSQU'EN 2120, PHOTOGRAPHIE DE LA SITUATION EN 2120, PARTITIONNEMENT PAR LE DIAMETRE 83
FIGURE 57 : PREVISION DES TAUX DE RENOUVELLEMENT FUTURS DE 2013 JUSQU'EN 2120 POUR LES TRONÇONS DE « GROS DIAMETRES », DE « MOYEN DIAMETRES », DE « PETITS DIAMETRES » ET « TOUS DIAMETRES », AU SEDIF AVEC LE SCENARIO « SATU QUO : PLUSIEURS COURBES » 83
FIGURE 58 : PREVISIONS DE L'EVOLUTION DE L'AGE MOYEN DU RESEAU DE 2013 JUSQU'EN 2120 AU SEDIF, SCENARIO « STATU QUO, PLUSIEURS COURBES », TRONÇONS PARTITIONNES PAR DIAMETRE. 84
FIGURE 59 : PREVISIONS DES BUDGETS NECESSAIRES POUR LE RENOUVELLEMENT ENTRE 2013 ET 2120 AVEC 2 TAUX D'ACTUALISATION DIFFERENTS, PARTITIONNEMENT PAR LE DIAMETRE. 85
FIGURE 60 : EVOLUTION DU COUT D'INVESTISSEMENT SUR 108 (SOMME DES COUTS SUR 2013-2120) ANS SI R VARIE, E = 3,7% 86
FIGURE 61 : EVOLUTION DU COUT D'INVESTISSEMENT SUR 108 ANS (SOMME DES COUTS SUR 2013-2120) SI E VARIE ET R=3%. 86
FIGURE 62 : PREVISIONS DES BUDGETS NECESSAIRES POUR LE RENOUVELLEMENT ENTRE 2013 ET 2120, SCENARIO « STATU QUO, PLUSIEURS COURBES », TRONÇONS PARTITIONNES PAR DIAMETRE. 87
FIGURE 63 : FONCTION DE SURVIE NON PARAMETRIQUE 105
FIGURE 64 : EVOLUTION POSSIBLE DU SIMPLEXE DE NELDER & MEAD 107
FIGURE 65 : ORGANIGRAMME DE LA METHODE DE NELDER & MEAD CLASSIQUE 108
FIGURE 66 : COURBE DE SURVIE REELLE 110

FIGURE 67 : DATE DE POSE DES 40 000 TRONÇONS CONSTITUANT L'ECHANTILLON (EN EFFECTIF) 111
FIGURE 68 : REPARTITION DES DATES A LA MISE HORS SERVICE (EN EFFECTIF) DES 40 000 TRONÇONS 111
FIGURE 69 : COURBES DE SURVIE ESTIMEES (EN EFFECTIF) ET INTERVALLE DE CONFIANCE 112
FIGURE 70 : DATE DE POSE DES TRONÇONS CONSTITUANT L'ECHANTILLON (EN LINEAIRE) 113
FIGURE 71 : REPARTITION DES DATES A LA MISE HORS SERVICE (EN LINEAIRE) 114
FIGURE 72 : COURBES DE SURVIE ESTIMEES (EN LINEAIRE) 114
FIGURE 73 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU FONTE GRISE CONDITIONNE 115
FIGURE 74 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU FONTE GRISE DECONDITIONNEES 115
FIGURE 75 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU FONTE DUCTILE 116
FIGURE 76 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU BONNA 116
FIGURE 77 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU PEHD 117
FIGURE 78 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU ACIER 117
FIGURE 79 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU AUTRE CONDITIONNE 118
FIGURE 80 : COURBES DE SURVIE EN LINEAIRE POUR MATERIAU AUTRE DECONDITIONNE 118
FIGURE 81 : CORRELATION ENTRE LONGUEUR ET DIAMETRE EN NOMBRE DE TRONÇON (ECHELLE REDUITE) 121
FIGURE 82 : CORRELATION ENTRE LONGUEUR ET DIAMETRE EN NOMBRE DE TRONÇON 121
FIGURE 83 : COURBES DE SURVIE DE KAPLAN-MEIER TOUS TRONÇONS CONFONDUS ET EQUATION B 122
FIGURE 84 : EVOLUTION DE L'INDICE TP 10A ENTRE JANVIER 2004 ET SEPTEMBRE 2013 123
FIGURE 85: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 2 SOUS-ENSEMBLES SELON LE NIVEAU PIEZOMETRIQUE : [0 :150M [ET [150M : MAX]) ET LEURS INTERVALLES DE CONFIANCE 125
FIGURE 86: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LE NIVEAU PIEZOMETRIQUE (NOMBRE DE TRONÇONS) 125
FIGURE 87: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 3 SOUS-ENSEMBLES SELON LE NIVEAU D'ENFOUISSEMENT : [MIN : 0[, [0 :1.15M [ET [1.15M : MAX]) ET LEURS INTERVALLES DE CONFIANCE 126
FIGURE 88: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LE NIVEAU D'ENFOUISSEMENT (NOMBRE DE TRONÇONS) 126
FIGURE 89: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 6 SOUS-ENSEMBLES : JOINTS SOUDES, JOINTS VISSES A BRIDE, JOINTS A EMBOITEMENT, JOINTS ELECTRO-SOUDES, JOINTS COULES A PLOMB ET JOINTS AUTRE) 127
FIGURE 90: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LE TYPE DE JOINTS (NOMBRE DE TRONÇONS) 127
FIGURE 91: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 2 SOUS-ENSEMBLES : TECHNIQUE DE POSE TRADITIONNELLE ET TECHNIQUE DE POSE AUTRE) ET LEURS INTERVALLES DE CONFIANCE 128
FIGURE 92: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LA TECHNIQUE DE POSE TRADITIONNEL OU NON (NOMBRE DE TRONÇONS) 128
FIGURE 93: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 3 SOUS-ENSEMBLES : APPUI BETON, APPUI SOL COMPACTE ET APPUI AUTRE) ET LEURS INTERVALLES DE CONFIANCE 129
FIGURE 94: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LE TYPE D'APPUI (NOMBRE DE TRONÇONS) 130
FIGURE 95: COURBES DE SURVIE EN EFFECTIF TRONÇONS (TRONÇONS PARTITIONNES EN 3 SOUS-ENSEMBLES : EMPLACEMENT SOUS CHAUSSEE, EMPLACEMENT SOUS TROTTOIR OU EMPLACEMENT AUTRE) ET LEURS INTERVALLES DE CONFIANCE 130
FIGURE 96: DIAGRAMME DE LA STRATIFICATION DES TRONÇONS SELON LEUR EMPLACEMENT SOUS CHAUSSEE OU NON OU AUTRE (NOMBRE DE TRONÇONS) 131

Tableaux

TABLEAU 1 : LISTE DES MEMBRES DU COMITE DE SUIVI DU PROJET « OPTIMEAU » 2
TABLEAU 2 : QUARTILES ET MOYENNE DU NOMBRE DE TRONÇONS EN FONCTION DE LEUR LONGUEUR (100% = 287 716 TRONÇONS)..... 20
TABLEAU 3 : CORRESPONDANCE ENTRE L'INDICE DE RENOUVELLEMENT EN POURCENTAGE ET L'INDICE DE RENOUVELLEMENT EN CLASSE (NALDEO 2013). 43
TABLEAU 4 : MODIFICATIONS REALISEES SUR LES DATES DE POSE EN FONCTION DU MATERIAU 46
TABLEAU 5 : PRIX MOYEN POUR LE RENOUVELLEMENT D'UN METRE DE TRONÇON PAR CLASSE DE DIAMETRES (PILLOT 2011)..... 84
TABLEAU 6 : LES CRITERES DE RENOUVELLEMENT DES CANALISATIONS PAR VEOLIA SUR LE TERRITOIRE DU SEDIF EN 2010 (BRUYERON & RE, 2012) 101
TABLEAU 7 : SYNTHESE DE LA DISTRIBUTION DES AGES A LA MISE HORS SERVICE EN FONCTION DU MATERIAU . 119
TABLEAU 8 : STRUCTURE DE L'INDICE TP 10A 123
TABLEAU 9 : EVOLUTION DE L'INDICE TP 10A ENTRE JANVIER 2004 ET SEPTEMBRE 2013..... 124

Sigles, acronymes et abréviations

AEP Alimentation en Eau Potable

ARP Annual Rehabilitation Programmes

AERMC Agence de l'Eau Rhône Méditerranée et Corse

AESN Agence de l'Eau Seine Normandie

CARE-W Computer Aided REhabilitation of Water networks (GB) =
réhabilitation des réseaux d'eau potable assistée par ordinateur (F)

CEMAGREF Centre du Machinisme Agricole du Génie Rural des Eaux et des Forêts

CGCT Code Général des Collectivités Territoriales

CVM Chlorure de Vinyle Monomère

Dr. Docteur

EAWAG Swiss Federal Institute of Aquatic Science and Technology

ENGEES Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg

ES En Service

ETBX Environnement, Territoires et infrastructures, Bordeaux (ex REBX + ADBX)

FOWA Forschungsfonds Wasser

FR Fuites Réparées

GPIE Gestion Patrimoniale des Infrastructures liées à l'Eau (ex NETWATER)

HS Hors Service

i.i.d. indépendantes et identiquement distribuées

Ing. Ingénieur

IRSTEA Institut de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

KM Kaplan-Meier

NA Not Available

NETWATER NETwork WATER (UK), réseau d'eau (F)

NT Non Trouvées

PEBD Polyéthylène Basse Densité

PEHD Polyéthylène Haute Densité

Pr. Professeur d'Université

PVC Polychlorure de Vinyle

QR Qualitatives Réparées

RPQS Rapport sur le Prix et la Qualité du Service

SEDIF Syndicat des Eaux d'Ile-de-France

SIG Système d'Information Géographique

SSIGE Société Suisse de l'Industrie du Gaz et des Eaux

VEDIF Véolia Eau d'Ile de France

Notations

θ	Année, ex : 1995
DDP	Année de pose
DHS	Année à la mise hors service
θ_1	Année de démarrage de la fenêtre d'observation
θ_2	Dernière année de la fenêtre d'observation
T ou t	Age à la mise hors service, ex : 60 ans ; t = DHS-DDP
a_i	Age au début de la fenêtre d'observation de l'individu i
b_i	Age à la fin de la fenêtre d'observation de l'individu i
c_i	L'individu i est censuré si c_i=1, non censuré si c_i = 0
Lg	Longueur
ø	Diamètre
NDOA	Nombre moyen de défaillances FR observées par an sur la fenêtre d'observation
r_a	Taux d'actualisation l'année a
S(t)	Fonction de survie (dépendant de t)
F(t)	Fonction de répartition (dépendant de t)
P	Mesure de probabilité
Cθ	Cohorte (groupe de tronçons) posée l'année θ
TR	Taux de renouvellement
L	Fonction de vraisemblance
TxD5	Taux de défaillances FR sur les 5 dernières années (Somme sur 5 an par hectomètre)
W	Temps de censure
$\boldsymbol{\Theta}$	Vecteur de paramètres
$I\{...\}$	La fonction indicatrice qui prend 1 si les conditions entre accolades sont respectées, 0 sinon

Introduction

Dans les pays développés, l'eau potable est distribuée au domicile des usagers. Ce confort implique un long linéaire de réseau (plus de 900 000 km en France en 2014) avec une valeur des ouvrages élevée (entre 120 et 170 milliards d'euros en France dont 70% correspond aux canalisations) (Bouchet 2013) & (Charrière 2013).

Le renouvellement annuel d'une fraction du linéaire des canalisations est nécessaire à cause du vieillissement « naturel » de ces objets techniques, qui entraîne l'érosion progressive de leur niveau de performance. En parallèle du vieillissement, l'évolution des besoins et des contraintes (coordination avec les travaux de voirie, urbanisation, législation) peut nécessiter le remplacement de tronçons obsolètes en bon état.

Dans une optique d'optimisation des ressources et des performances, il est alors important de renouveler les tronçons en temps opportun (ni trop tôt, ni trop tard).

Dans ce cadre, l'IRSTEA, organisme public de recherche finalisée sur la gestion des eaux et des territoires, a lancé le programme de recherche « Optimeau » dont le but est d'optimiser le renouvellement des canalisations d'eau potable. Ce rapport s'inscrit dans ce programme de recherche. Ce projet est financé directement par le syndicat des eaux d'île de France (SEDIF), la communauté urbaine de Lyon (Grand Lyon) et le service des eaux de la ville de Lausanne (eauservice Lausanne). Il est financé indirectement par l'agence de l'eau Seine-Normandie, l'agence de l'eau Rhône Méditerranée et Corse, et le fonds de recherche pour l'eau (Forschungsfonds Wasser = FOWA) de la Société Suisse de l'Industrie du Gaz et des Eaux (SSIGE).

L'approche proposée pour optimiser ressources et performances est de partir d'une vision « long terme » (>30 ans) permettant de prédire le linéaire de canalisations qui doivent être renouvelées chaque année, pour respecter des objectifs stratégiques en termes de performance. Puis d'en déduire le besoin annuel en investissement à « moyen terme » (≈10 ans, échelle des budgets des services). Il faut ensuite appliquer des méthodes de « court terme » (≤ 3 ans) permettant d'obtenir pour l'année à venir une liste de tronçons hiérarchisés par niveau de besoin et opportunités en renouvellement (cf. Figure 1).

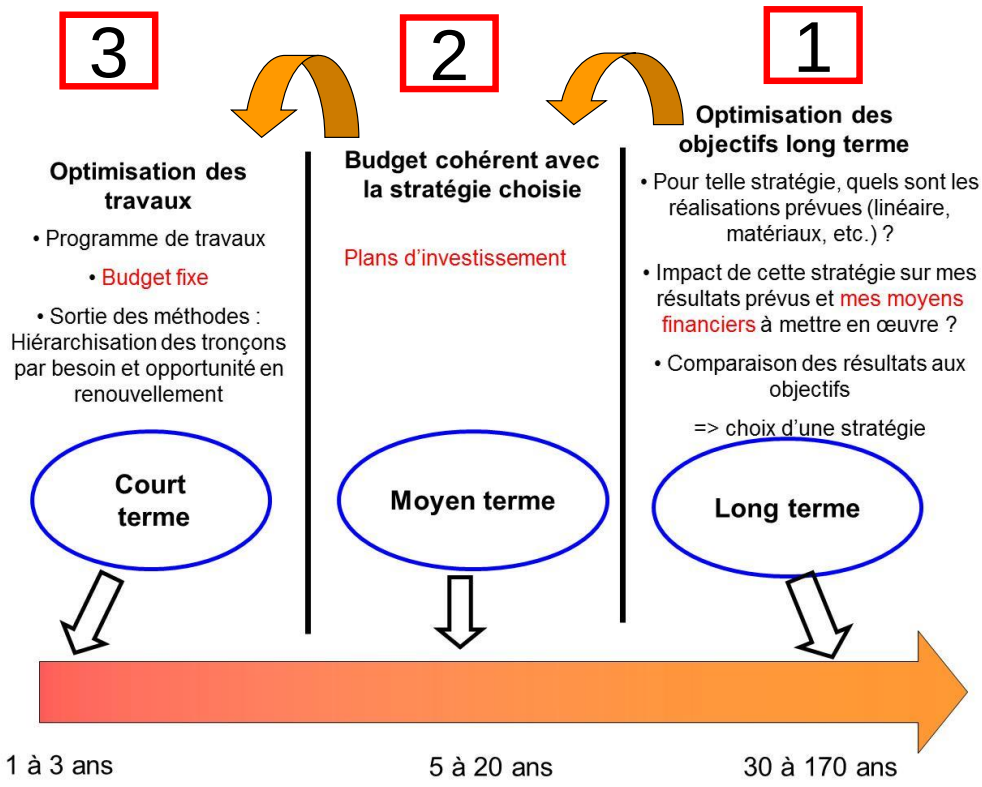


Figure 1 : L'idéal de l'articulation entre les différents outils fonctionnant à différentes échelles temporelles

Le problème est qu'actuellement les méthodes décisionnelles sur le long terme employées par les services des eaux ou dans la bibliographie sont souvent inexistantes ou trop simples et déconnectées des méthodes décisionnelles sur du court terme.

L'objectif final du projet « optimeau » est de construire une méthode à long terme (> 30 ans) permettant d'aider aux choix des réalisations (linéaire, matériaux, etc.) afin d'optimiser le renouvellement des canalisations.

Dans un **premier temps** nous présenterons le terrain d'étude (SEDIF) et le jeu de données sur lequel nous nous sommes appuyés pour obtenir tous nos résultats dans ce rapport.

Dans la **seconde partie** nous présenterons de manière globale la méthode « long terme » que nous souhaitons développer dans le projet « optimeau ».

Dans un **troisième temps** nous exposerons les différentes méthodes de gestion patrimoniale actuellement employées au SEDIF (méthodes décisionnelles long terme et court terme).

Puis via la **quatrième partie** nous exposerons différentes méthodes statistiques de construction de courbe de survie passée d'âge à la mise hors service (modèle de Kaplan-Meier, modèle de Cox, modèle de Turnbull….).

Ensuite dans un **cinquième** temps nous appliquerons ces méthodes avec les données du SEDIF.

Finalement nous construirons des indicateurs clés tel que des taux de renouvellement futurs, des linéaires futurs, l'âge moyen du réseau ou encore les coûts d'investissement toujours à partir des données du SEDIF.

I Présentation du SEDIF

I.1 Territoire, mode de gestion et histoire

Le Syndicat des eaux d'Ile-de-France (SEDIF) est un établissement public de coopération intercommunale créé en 1923 qui gère en 2013 le service public d'eau potable pour 149 communes de l'agglomération parisienne. Son réseau (transport[1] et distribution[2]) constitué de 8 300 km s'étend sur plus de 7 départements (cf. Figure 2). Plus de 4.3 millions de personnes consomment l'eau provenant pour 95 % d'eau de surface (Marne, Seine et Oise), c'est à dire d'origine fluviale. Le réseau comprend 3 usines de traitement, 45 stations de relevage, 63 réservoirs et 38 stations de chloration (SEDIF 2012b). Le SEDIF est l'autorité organisatrice la plus importante de France en nombre d'usagers et linéaire de réseau.

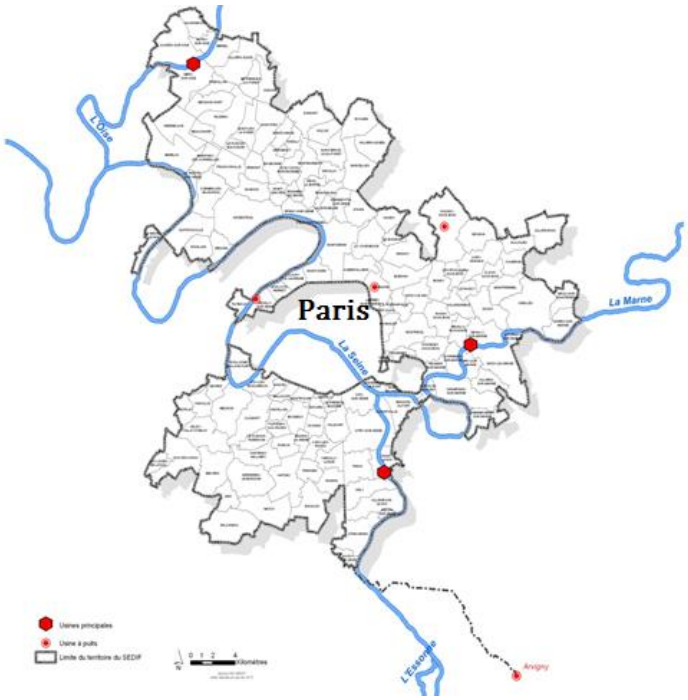


Figure 2: Territoire couvert par le SEDIF en 2013

Le SEDIF est propriétaire de toutes les installations de production de transport et de distribution (usines, réservoirs, canalisations…).

Jusqu'en décembre 2010 l'exploitation était déléguée à Véolia Eau au travers des contrats de régie intéressée d'une durée de 39 ans de 1923 à 1962, puis de 30 ans de 1962 à 1992 et enfin de 19 ans de 1992 à 2011. Depuis le premier janvier 2011, le nouveau contrat, d'une durée de 12 ans, attribue à VEDIF (Véolia Eau d'Ile de France), filiale de Véolia, toute l'exploitation de son réseau (Lanez 2010; Lion 2009), seuls les investissements sont en majorité faits par le SEDIF.

[1] Transport : Réseau de diamètre strictement supérieur à 300mm, ce qui correspond à peu près au réseau entre la rivière et les châteaux d'eau.

[2] Distribution : Réseau de diamètre inférieur ou égal à 300mm, hors branchement, ce qui correspond à peu près au réseau entre les châteaux d'eau et les branchements. Sachant qu'un branchement correspond à peu près au réseau entre le réseau de distribution et le compteur de l'abonné.

I.2 Caractéristiques des canalisations

I.2.1 Origine des données

Ce travail est effectué notamment sur les données des canalisations d'eau potable du réseau de distribution et de transport du SEDIF. Ce réseau est composé par plus de 290 000 tronçons. Un tronçon est défini comme l'ensemble des tuyaux unitaires, assemblés par des joints, homogènes en matériau, diamètre, date de pose et formant un segment de réseau généralement délimité à ses deux extrémités par des organes de sectionnement (vannes) (Le Gat 2009). Mais au SEDIF la définition n'est pas tout à fait celle-là. Au SEDIF : un tronçon est une portion élémentaire de conduite qui tient compte des changements d'environnement : par exemple, lors d'un changement de voie ou d'un passage « sous chaussée » à « sous trottoir » d'une conduite, un nouveau tronçon est créé (en général, une conduite est constituée de plusieurs tronçons) (Levesque 2014).

L'ensemble de l'étude est basé sur deux fichiers tronçons comportant chacun des informations fondamentales :

Fichier tronçons OCL[3]: Ce premier fichier illustré par la Figure 3 comporte des informations pour chaque tronçon. Nous pouvons y distinguer sa date de pose, son matériau, son identifiant, son statut (en service (ES) ou hors service (HS)), sa date de mise hors service si le tronçon a été mis hors service dans la fenêtre d'observation, etc. Toutes ces précieuses descriptions permettent d'obtenir les caractéristiques propres à chaque tronçon. Ce fichier comprend 99 variables. Ces données ont été extraites en mars 2013 du système d'information géographique (SIG) du SEDIF.

OBJ_ID	COMMUNE	DATE_POSE	DIAM	EMPLACEMEN	MATERIAU	LONG_REF
20120427102529000031ZJ	Aubervilliers	01/01/1995	50	Chaussée centra	PEHD	3,68
20120427104534000054BB	Alfortville	01/01/1975	100	Chaussée côté b	Fonte ductile	59,35
20120427110220000006V1C	Montfermeil	01/01/1900	100	Trottoir pair	Fonte grise	
20120427105547000067 8C	Bagneux	01/01/1980	200	Trottoir impair	Fonte ductile	2,04
20120427104625000057B0	Pierrelaye	01/01/1975	100	Chaussée	Fonte ductile	77,66
20120427110118000060Q8	Saint-Gratien	30/04/2008	150	Chaussée	Fonte ductile	17,14
20120427110646000007AIK	Cormeilles-en	01/01/1987	100	Chaussée	Fonte ductile	33,5
20120427104452000051OH	Bondy	01/01/1963	100	Trottoir impair	Fonte grise	25,08
20120427104241000047TU0	Châtillon	01/01/1962	100	Trottoir impair	Fonte grise	6,87
20120427110828000007GH8	Viroflay	26/02/2010	125	Trottoir	PEHD	3,12
20120427103705000049CS	Montrouge	01/01/1970	100	Trottoir impair	Fonte ductile	5,37
20120427101744000028LV	Margency	01/01/1988	50	Parking-esplana	PEHD	5,95
20120427105903000060KLO	Neuilly-sur-Se	01/01/1997	100	Trottoir pair	Fonte ductile	8,03
20120427101147000001NEX	Champigny-su	01/01/1981	100	Trottoir impair	Fonte ductile	1,55
20120427104249000047UAE	Palaiseau	01/01/1992	100	Chaussée paire	Fonte ductile	7,06
20120427105441000062MZ	Joinville-le-Po	01/01/1997	125	Chaussée paire	PEHD	8,66
20120427103730000047B39	Le-Perreux-sui	01/01/1924	80	Chaussée impair	Fonte grise	16,4
20120427101242000001R0D	Levallois-Perre	01/01/1872	100	Trottoir impair	Fonte grise	91,1
20120427110142000065CY	Argenteuil	01/01/1992	100	Chaussée impair	Fonte ductile	0,32
20120427103133000037P49	Saint-Prix	01/01/1931	80	Chaussée centra	Fonte grise	56,47

Figure 3 : Extrait du jeu de données des tronçons

Fichier tronçons FCL[4] : Ce fichier est assez similaire au premier mais il comporte en plus les coordonnées (x, y), associées à chaque identifiant, représentant le début et la fin du tronçon. Il est alors possible de comparer deux canalisations en termes de distance et d'établir des conditions de voisinage. Le système de localisation géographique "Lambert 93" est utilisé.

Ce jeu de données présente 93 % de tronçons en service, pour environ 7% hors service et 0.06 % de statut « inconnu ». Pour ces mêmes tronçons, 94% ont une date de pose connue et apparemment

[3] OCL : table attributaire des tronçons
[4] FCL : géométrie des tronçons

correcte, 1% ont des dates de pose inconnues et 5 % ont une date de pose incohérente (cf. « III.1. Nettoyage des données »).

I.2.2 Caractéristiques des linéaires des tronçons

Les graphiques de la Figure 4 illustrent la distribution de la taille des tronçons. Il en ressort que la plupart des tronçons étudiés est de petite taille (minimum 1 cm, et 90% des tronçons ont une longueur inférieure à 93 m [Cf. Figure 5 et Tableau 2]). Effectivement, la médiane se situe à 7 mètres alors que le plus long tronçon mesure 1 492 mètres et la moyenne est de 31 m.

Le nombre de très petits tronçons (<6m) apparaît comme étant très important (132 000 tronçons). Cela est dû à la définition du tronçon au SEDIF. Cette définition induit un découpage des canalisations en fonction de leurs caractéristiques et de leur environnement (types de sol, profondeur, chaussée/trottoir, aspects hydrauliques, etc.) (Levesque 2014).

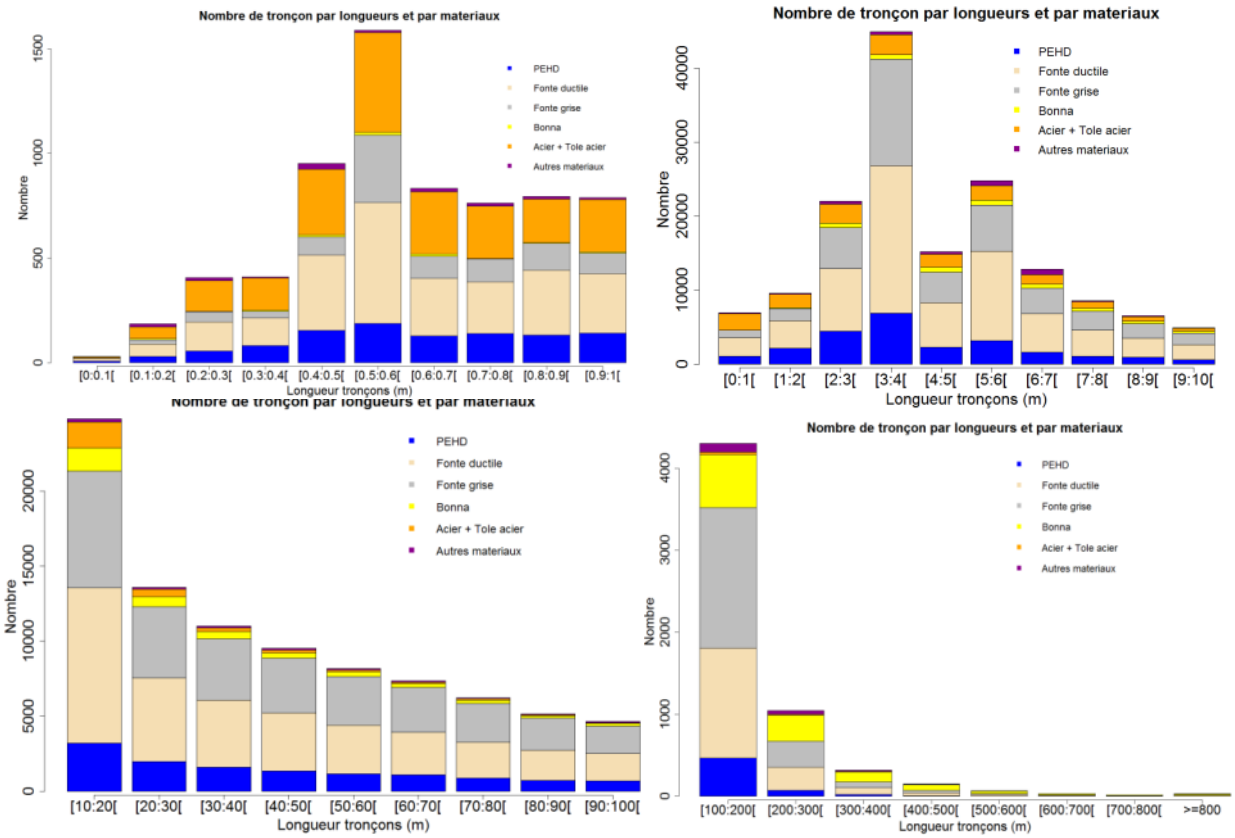


Figure 4 : Nombre de tronçons en fonction de leur longueur

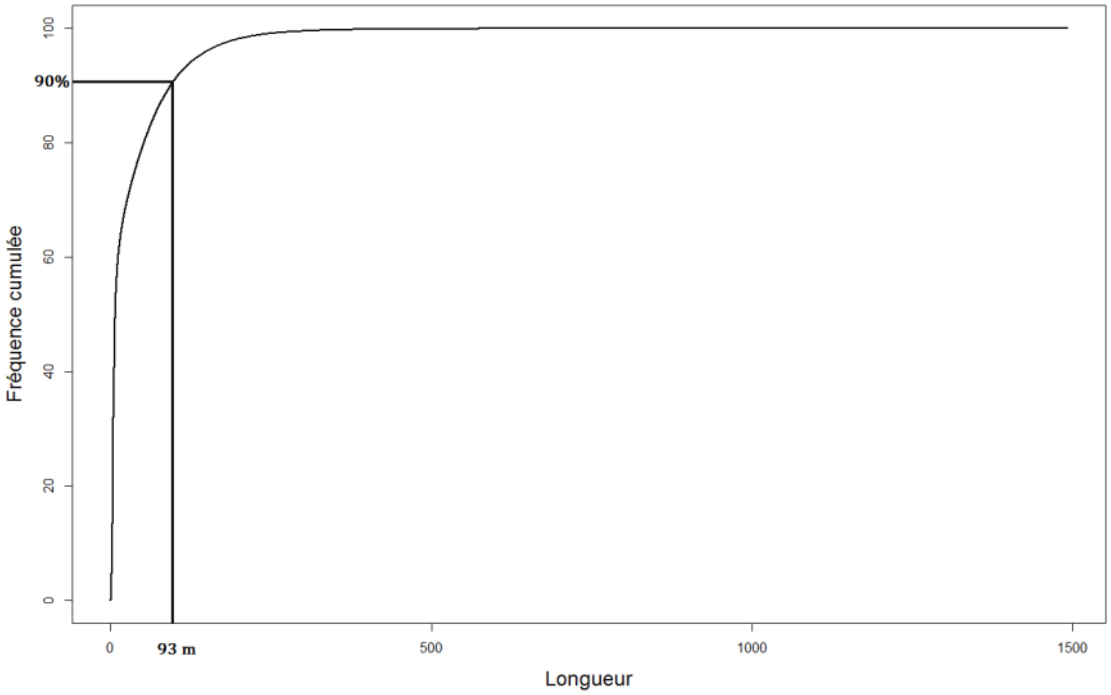


Figure 5 : Fréquence cumulée du nombre de tronçons (en service et hors service) en fonction de leur longueur (en mètre) (100% = 287 716 tronçons)

Nbre de tronçons (fréquence cumulée)	Le plus petit	1er quartile (25%)	Médiane (50%)	Moyenne	3e quartile (75%)	(90%)	Le plus grand (100%)
Valeur	0.01 m	3.6 m	7 m	31 m	38 m	93 m	1 492 m

Tableau 2 : Quartiles et moyenne du nombre de tronçons en fonction de leur longueur (100% = 287 716 tronçons)

I.2.3 Comparaison du linéaire en service et hors service en fonction des dates de pose

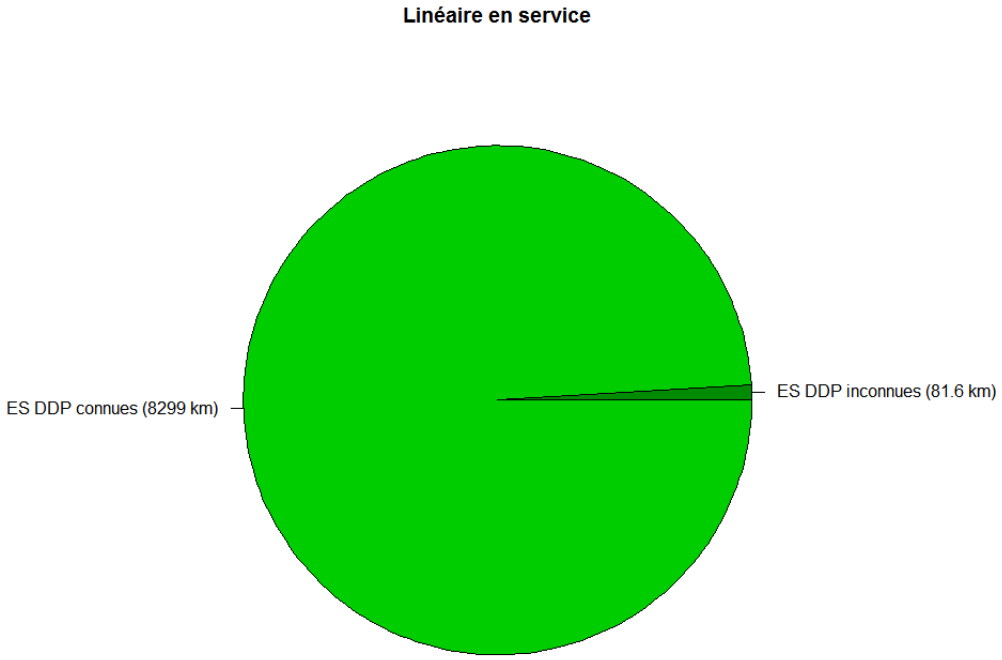


Figure 6 : Diagramme du linéaire (km) en service au SEDIF au 31/12/2012 en fonction des dates de pose (8380.6 km au total)

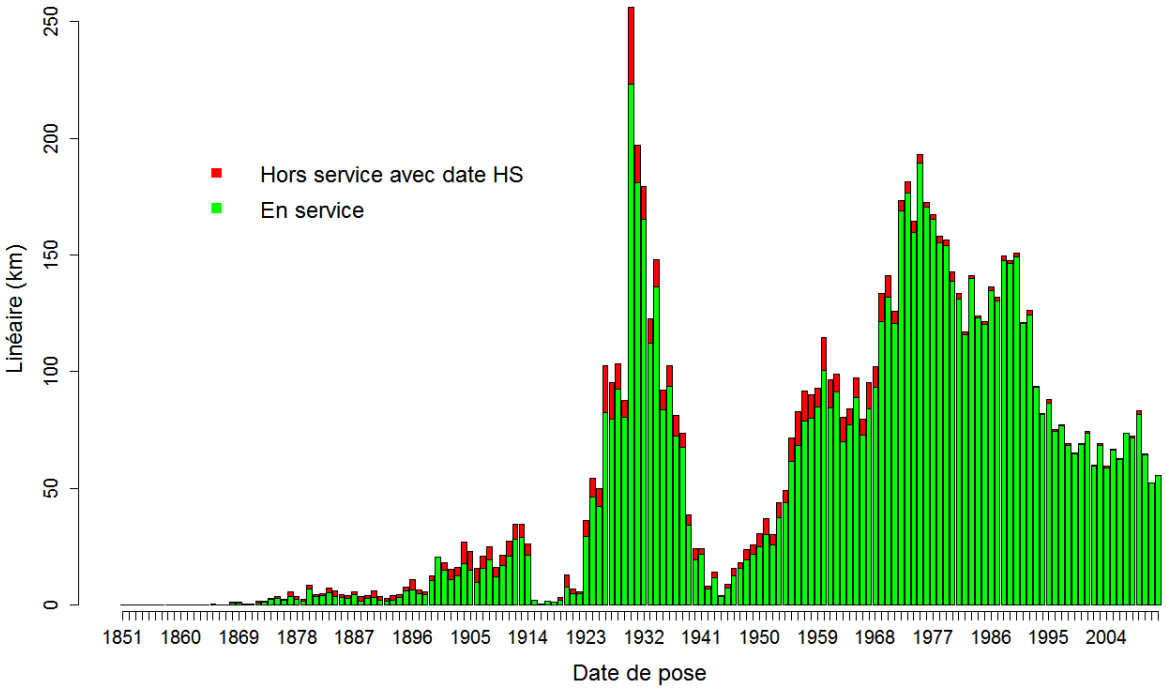


Figure 7 : Histogramme du linéaire en service (8 296 km) et hors service (HS) (680 km) au 31/12/2012 ayant une date de pose connue (en rouge ne sont représentés que les tronçons mis HS entre [1995, 2012]).

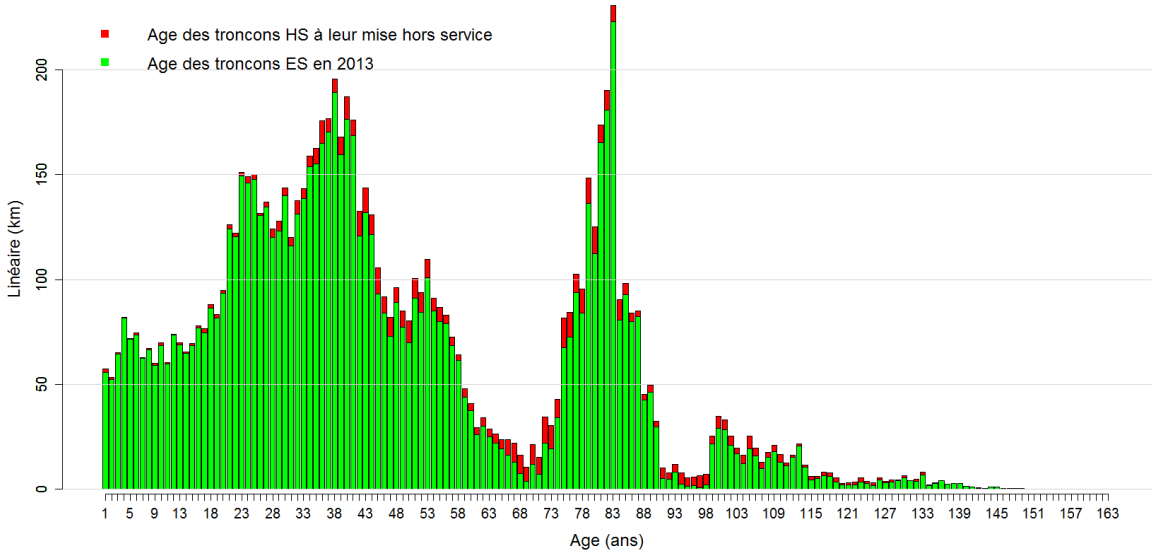


Figure 8 : Histogramme du linéaire En Service et Hors Service au 31/12/2012 en fonction de leur âges (en rouge ne sont représentés que les tronçons mis HS entre [1995 ; 2012])

L'histogramme de la Figure 7 nous donne une photographie des dates de pose des canalisations en service en décembre 2012 (cf. barres vertes). La plus vieille canalisation en service fin 2012 date de 1851. Ce graphique nous indique distinctement les deux guerres mondiales où il y a eu une nette diminution de la pose de canalisation. L'après-guerre s'accompagne d'une recrudescence des travaux qui atteignent leur apogée dans les années 1970. Dès 1995, les travaux concernaient essentiellement des réparations ou des remplacements plutôt que la mise en place de nouveaux conduits (Ajuste et al. 2004). L'âge moyen du réseau en service au SEDIF est d'environ 46 ans au 31/12/2012.

Au SEDIF, l'ensemble des dates de pose des tronçons hors service sont connues.

I.2.4 Période d'observation du linéaire hors service

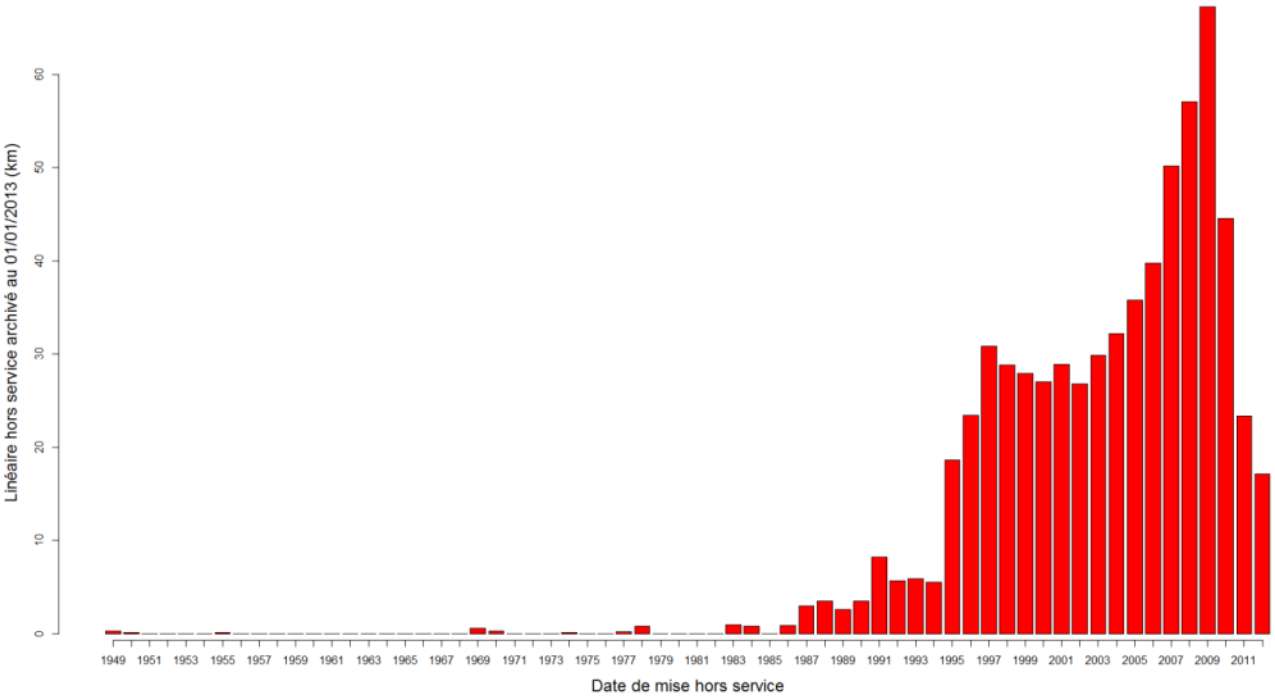


Figure 9 : Histogramme du linéaire hors service archivé au 01/01/2013

La Figure 9 permet d'observer que les dates de mise hors service n'ont commencé à être significativement archivées dans le SIG qu'à partir de 1995. De plus comme les données nous ont été fournies début 2013, nous n'avons pas de date de mise hors service pour cette année. Les courbes de survie estimées ne seront calculées qu'avec des tronçons archivés dans **l'intervalle considéré comme fiable [1995 ; 2012].** Ce sera notre « fenêtre d'observation ». Au final entre 1995 et 2012 environ 608 km de réseau ont été mis hors service.

D'après (Optimeau 2014) les données des dates de mise hors service sur la fenêtre [1995, 2012] ne sont pas exhaustives mais elles sont significatives. En effet l'archivage sous format informatique des dates à la mise hors service était inexistant avant 2005. L'archivage au format informatique s'est mis en place en 2005. Les données avant 2005 ont été collectées à partir des formulaires papiers. Elles sont alors très partielles avant 1995 et représentatives mais pas complètes entre [1995,2004]. A partir de 2005 cette donnée est exhaustive.

I.2.5 Les matériaux des canalisations

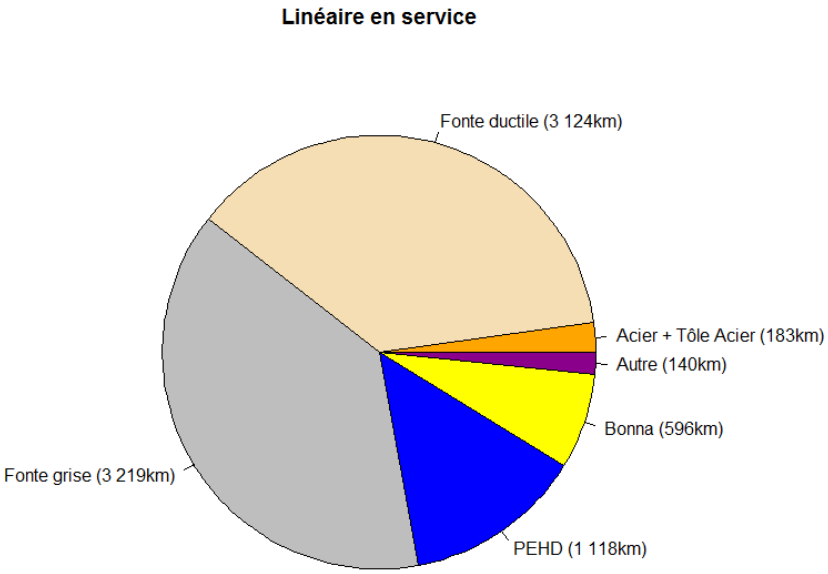


Figure 10 : Les matériaux du réseau en service au 31/12/2012

D'après le graphique Figure 10 nous pouvons observer qu'une importante partie du réseau en service fin 2012 est constituée de fonte grise qui n'est plus posée depuis le début des années soixante-dix (cf. Figure 11). Ce type de canalisation représente 38% du linéaire du réseau actuel. La fonte ductile, posée à partir de 1965 représente également à ce jour 38% du réseau du SEDIF. Les autres matériaux sont comme suit : PEHD (Polyéthylène haute densité) 13.5%, Bonna (béton âme tôle) 7%, acier + tôle acier 2% et autres 1.5%.

Les tuyaux en Bonna constituent une partie du réseau de transport : ils assurent la liaison entre les usines de production et les grands réservoirs principaux (SEDIF 2014), leur diamètre est supérieure à 300 mm.

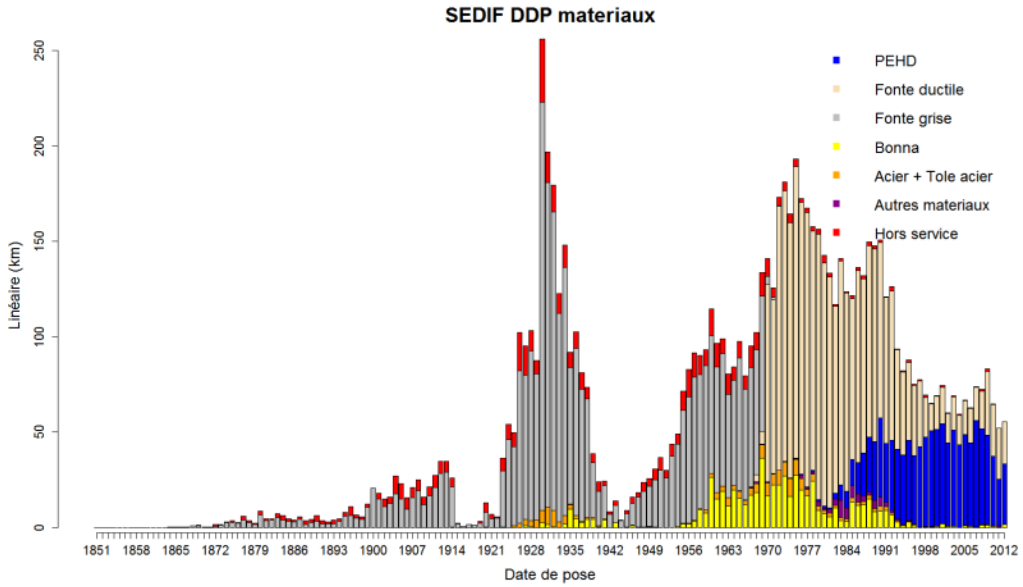


Figure 11 : Histogramme du linéaire en fonction des DDP des différents matériaux au 31/12/2012

Le graphique de la Figure 11 permet d'avoir une vue globale sur l'évolution des matériaux utilisés pour les canalisations. Il est important de noter que la fonte ductile remplace définitivement la fonte grise dès le début des années 1970 et que le PEHD, matériau relativement récent, est de plus en plus utilisé de nos jours.

I.2.6 Les diamètres des canalisations

Les conduites sont regroupées dans un premier temps en 2 catégories :

- les conduites de transport dont le diamètre intérieur est supérieur à 300 mm,
- les conduites de distribution, dont le diamètre intérieur est inférieur ou égal à 300 mm. Elles sont ensuite regroupé en 2 groupes (Bruyeron & Re 2012b):
    - les canalisations locales (ø ≤ 150 mm),
    - les canalisations intercommunales (150 mm < ø ≤ 300 mm).

Aucun branchement n'est réalisé sur les conduites de transport, sauf rares exceptions.

Le graphique suivant présente le linéaire de conduite par catégories de conduites :

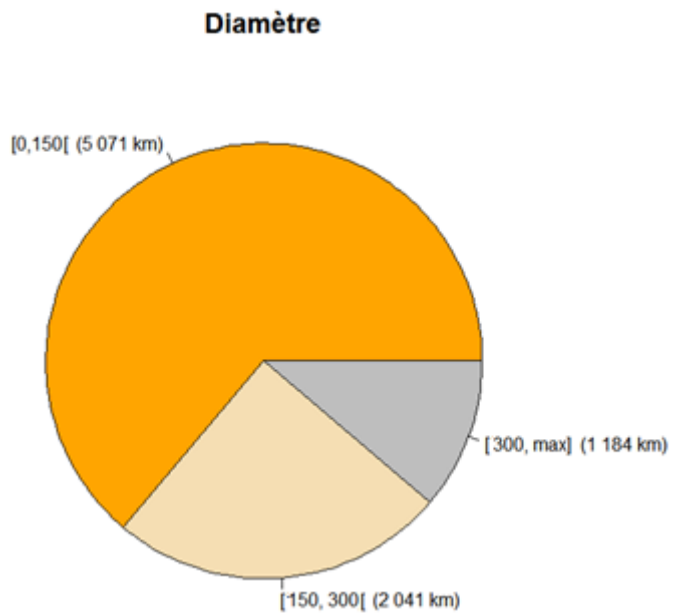


Figure 12 : Linéaire de tronçons en fonction de leur diamètre

La majorité du réseau (86 % du linéaire total) est constituée par des conduites de distribution (ø ≤ 300 mm), les canalisations intercommunales représentent 25% de ce réseau. Le réseau de transport (ø > 300 mm) ne représente que 14 % du réseau total.

Nbre de tronçons (fréquence cumulée)	Le plus petit	1er quartile (25%)	Médiane (50%)	Moyenne	3e quartile (75%)	(90%)	Le plus grand (100%)
Valeur	20 mm	100 mm	100 mm	169.1 mm	150 mm	93 mm	2 000 mm

Figure 13 : Fréquence cumulée du nombre de tronçons (en service et hors service) en fonction de leur diamètre (en mini mètre) (100% = 287 716 tronçons)

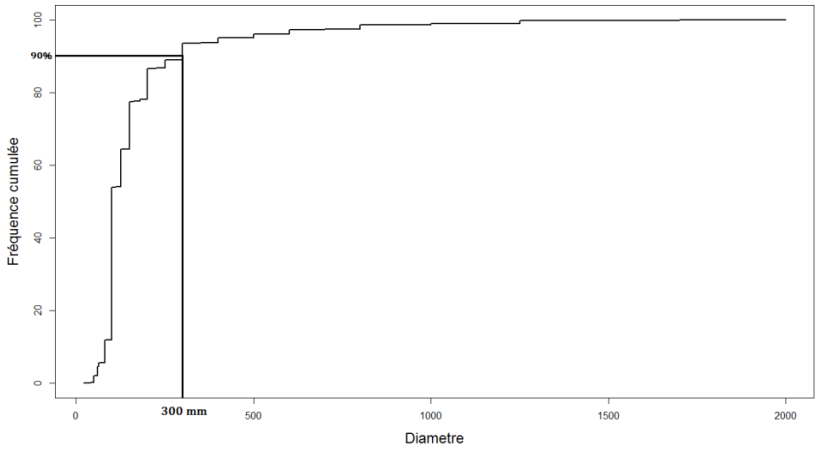


Figure 14 : Fréquence cumulée du nombre de tronçons en fonction de leur diamètre

I.3 Caractéristiques des défaillances

Une défaillance est la cessation de l'aptitude d'une unité fonctionnelle à accomplir une fonction requise (Breysse 2009).

En gestion patrimoniale des canalisations d'eau potable, il existe trois types de défaillances : les défaillances FR (Fuites Réparées), les défaillances QR (Qualitatives Réparées) impactant la qualité de l'eau, et les défaillances NT (Non Trouvées). Il n'existe pas de base de données sur ces dernières car par définition, elles nous sont inconnues.

I.3.1 Caractéristiques des défaillances FR (Fuites Réparées)

I.3.1.1 Définitions, généralités

Le schéma Figure 15 récapitule les liens notamment entre fuites, casses et défaillances FR.

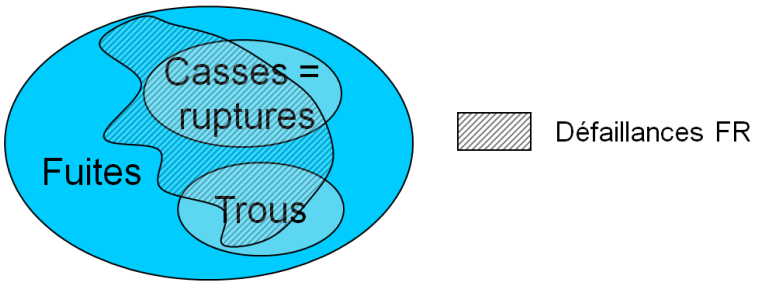


Figure 15 : schéma récapitulatif du lien entre fuites, casses, trous et défaillances FR

Une **Fuite** est un écoulement indésirable d'une canalisation, d'un branchement[5], d'un équipement ou d'un ouvrage qui nécessite une réparation (Charrière 2013).

Les casses ou les ruptures sont des sous-ensembles des fuites. Une **Casse** est une rupture d'un tronçon de canalisation ou d'un branchement (Charrière 2013) avec écoulement d'eau apparent en surface. Enfin une **Défaillance FR** est une fuite intervenant sur un tronçon de canalisation et donnant lieu à une réparation (Charrière 2013).

Ces fuites représentent chaque année en France 1.3 milliard de mètres cubes. 20% du volume traité par les usines n'arrivera jamais au robinet (France_Libertés 2014).

[5] Un **branchement** est une conduite et accessoires mis en œuvre pour amener l'eau du réseau de desserte jusqu'au point de livraison de l'eau à l'usager abonnée, à l'exception des conduites et accessoires privés des immeubles collectifs (Charrière 2013).

I.3.1.2 Présentation du fichier défaillances FR au SEDIF

Fichier défaillances FR : C'est une base de données de toutes les fuites réparées archivées. Celles-ci sont associées à l'identification du tronçon détérioré avec la date de réparation de la fuite, ses coordonnées (SIG) et sa description. Ce fichier est capital notamment pour la construction d'un modèle de prédiction des futures défaillances. Avant d'être rentré dans le SIG, les informations sur les défaillances sont collectées par les agents de terrain en remplissant la fiche en Annexe 1. Ce fichier contient 25 964 défaillances avec 23 variables.

ID_ACTIF	COP	DT_TRAVAUX	DateDef	COD_COM	COD_VOIE	LIBELLE_COMMUNE
215008	CIT EPINAY-SUR-SEINE	06-juin-12	2012	515	NA	MONTMORENCY
215630	CIT EPINAY-SUR-SEINE	14-juin-12	2012	434	NA	SOISY-SOUS-MONTMORENCY
215722	CIT PAVILLONS-SOUS-BOIS	12-juin-12	2012	220	NA	VILLEMOMBLE
215524	CIT PAVILLONS-SOUS-BOIS	11-juin-12	2012	502	NA	AULNAY-SOUS-BOIS

Figure 16 : Extrait du fichier SIG défaillances FR au SEDI

En accord avec les responsables techniques chargés de gestion patrimoniale du SEDIF, sont considérées comme défaillances FR :

1. Les fuites manifestes en surface réparées,
2. Les fuites non manifestes repérées suites à une campagne de recherche active puis réparées,

La Figure 17 montre le nombre de défaillances FR archivé par le SEDIF chaque année. Nous constatons que, comme pour les dates de mise hors service, l'enregistrement est significatif dans le SIG depuis début 1995.

I.3.1.3 Localisation des défaillances FR au SEDIF

Les défaillances FR surviennent au SEDIF (après nettoyage cf. annexe 2) cf. Figure 17 :

- Sur corps de conduite (78.6%),
- Sur joint (6.8%),
- Ou sur appareillage (vanne, robinet prise, etc.) (13.6%). (Le Gat 2009) (Large 2013b)

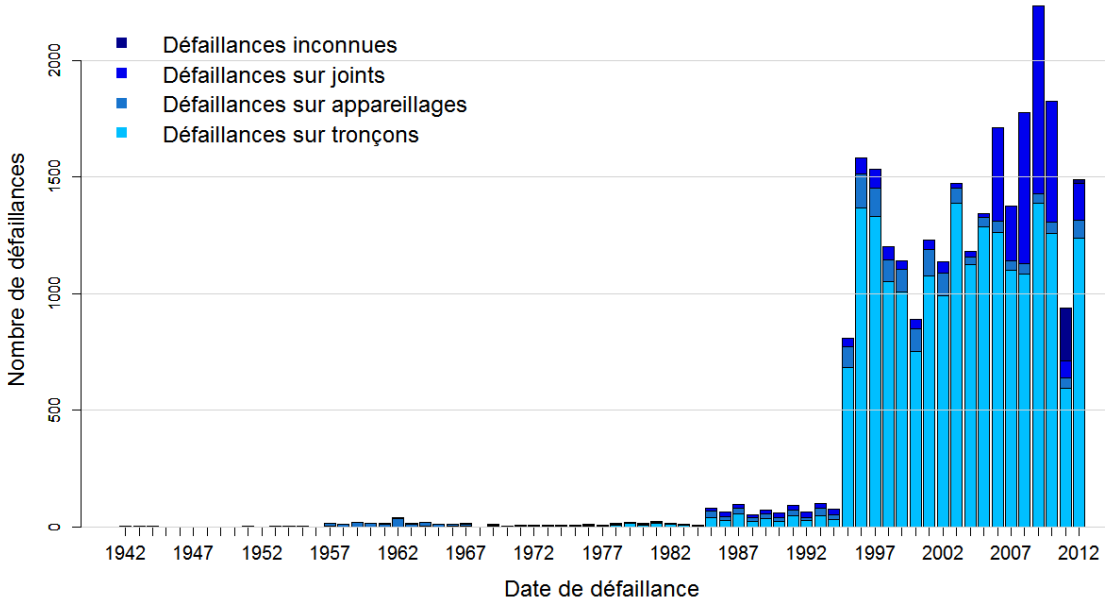


Figure 17 : Nombre de défaillances archivées par an, jusqu'à mars 2013

Dans ce rapport toutes les analyses ont été réalisés en comptabilisant les défaillances quelle que soit leur localisation. Néanmoins par la suite, dans les rapports suivant, conformément à ce qu'il a été demandé lors du comité de suivi n°3 (Optimeau 2014) nous prendrons en compte uniquement **les défaillances sur conduites, sur joints ou inconnues**.

I.3.1.4 Types de défaillances FR au SEDIF

Il existe plusieurs types de fuites : les fissures circulaires, les ruptures longitudinales , les piqûres, les ruptures d'éclatement… Représentant respectivement environ 57%, 13%, 7.5%, et 8.4% des défaillances rencontrées au SEDIF (Cf. Figure 19).

Dans notre base de données nous avons 14.1% de défaillances de type inconnu.

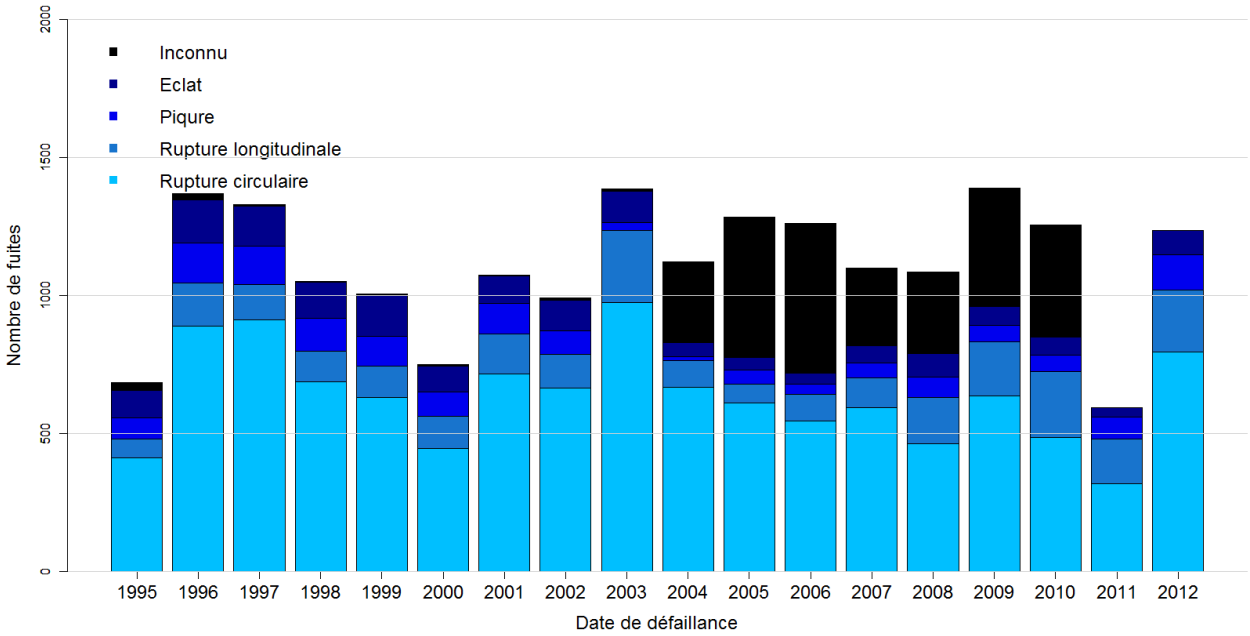


Figure 18 : Répartition des types de défaillances FR sur tronçon au SEDIF au cours du temps

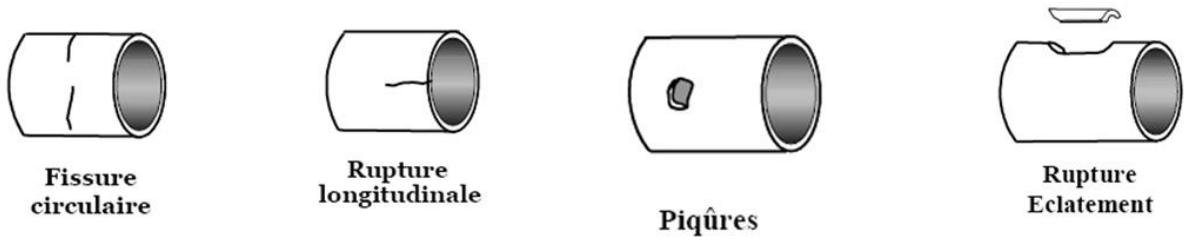


Figure 19 : Schéma des types de défaillances « Fuites Réparées » (Poinard 2006)

I.3.2 Les défaillances Qualitatives Réparées

Les défaillances liées au réseau et pouvant impacter la qualité de l'eau au SEDIF (eaux rouges, eaux noires, eaux riches en CVM [Chlorure de Vynil Monomère], eaux riches en plomb) sont très minoritaires (2 à 3 par an) et elles sont rapidement réparées. C'est pourquoi il n'existe aucune base de

données sur celles-ci. Cependant il existe une enquête annuelle de satisfaction client. Il en ressort en 2012 qu'environ 70% des consommateurs ont répondu comme étant plutôt satisfais des services de l'eau dans leur commune. Ils donnent une note de 8.1/10 pour la limpidité (couleur) de l'eau, 7.1/10 pour l'odeur et 6.9/10 pour le goût (TNS/SOFRES 2012).

I.4 Les données environnementales

Plusieurs paramètres de l'environnement extérieur aux conduites peuvent favoriser les défaillances comme l'indique (Berland et al. 2005) & (Babykina 2010) :

- le climat (variations de températures, nombre de jours de gel, nombre de jours très chaud, etc.);
- les charges du trafic routier transmises aux conduites ;
- l'agressivité naturelle ou anthropique de certains terrains qui peuvent induire des phénomènes de corrosion externe ;
- les mouvements de terrain ;
- la qualité des remblais et des travaux de compactage des sols ;
- conduite immergée dans des nappes phréatiques ;
- les mouvements de déstabilisation des sols causés par la pose, le remplacement ou les interventions d'entretien d'autres réseaux techniques ou par des travaux de voirie ;
- la présence de courants vagabonds générés par des installations électriques.

De plus les données environnementales concernant les éléments vulnérables (consommateurs, usagers de la route, biens des tiers, etc.) peuvent être très utiles pour calculer des indicateurs de risques :

- Nombre d'abonnés reliés au tuyau ;
- Localisation des transports en commun par rapport aux canalisations ;
- Localisation des cavités souterraines, etc.

Le SEDIF possède les données environnementales suivantes :

- **Les données sur les branchements.** Par exemple, il a été constaté que plus le nombre de piquages sur un tronçon augmente, plus le nombre de défaillances augmente. De plus le nombre de branchements sur un tronçon est corrélé au nombre de personnes reliés à ce tronçon.

- **Les données sur la corrosivité du sol.** En effet, le SEDIF a fait des mesures de corrosivité des sols et des canalisations corrodées et s'interroge sur la pertinence de poursuivre ces investigations. Le SEDIF mène aussi une étude plus spécifique afin de mesurer la corrosion des "Bétons Ame Tôle » (Large 2013a) et (SEDIF 2012a) .

- **Données sur la voirie.** Le SEDIF possède des cartes sur les voiries mais pas sur le niveau du trafic routier.

- **Le SEDIF a les zones des carrières et cavités sur son territoire**. Pour le SEDIF cela est une donnée pertinente car par exemple à Maisons Alfort, en 2011, il y a eu un trou assez dangereux (40 m de large) qui s'est formée suite à une fuite dans une cavité de carrière.

Le SEDIF pourrait collecter les données environnementales suivantes :

- MétéoFrance doit posséder les **données météorologiques**. Pour le SEDIF cela pourrait être une donnée intéressante dans l'analyse rétrospective des défaillances. En effet, tout comme (Babykina 2010) et (Claudio 2011), les agents du SEDIF ont constaté que les fuites apparaissent plutôt l'hiver dès que la température de l'eau passe en-dessous de 4°C (SEDIF 2013a).

Nous venons de présenter les différentes caractéristiques de notre terrain d'étude et des données disponibles. Nous pouvons retenir que pour les futures analyses statistiques que nous réaliserons la fenêtre d'observation historique fiable est de **18 ans : de 1995 à 2012** (bornes incluses).

Nous allons maintenant présenter de manière synthétique la démarche globale de notre modèle « long terme » que nous souhaitons créer dans le projet optimeau.

II Vue globale du modèle long terme que l'on souhaite développer dans « optimeau »

La démarche globale de notre modèle « long terme » peut se décomposer en 6 grandes étapes (cf. Figure 20).

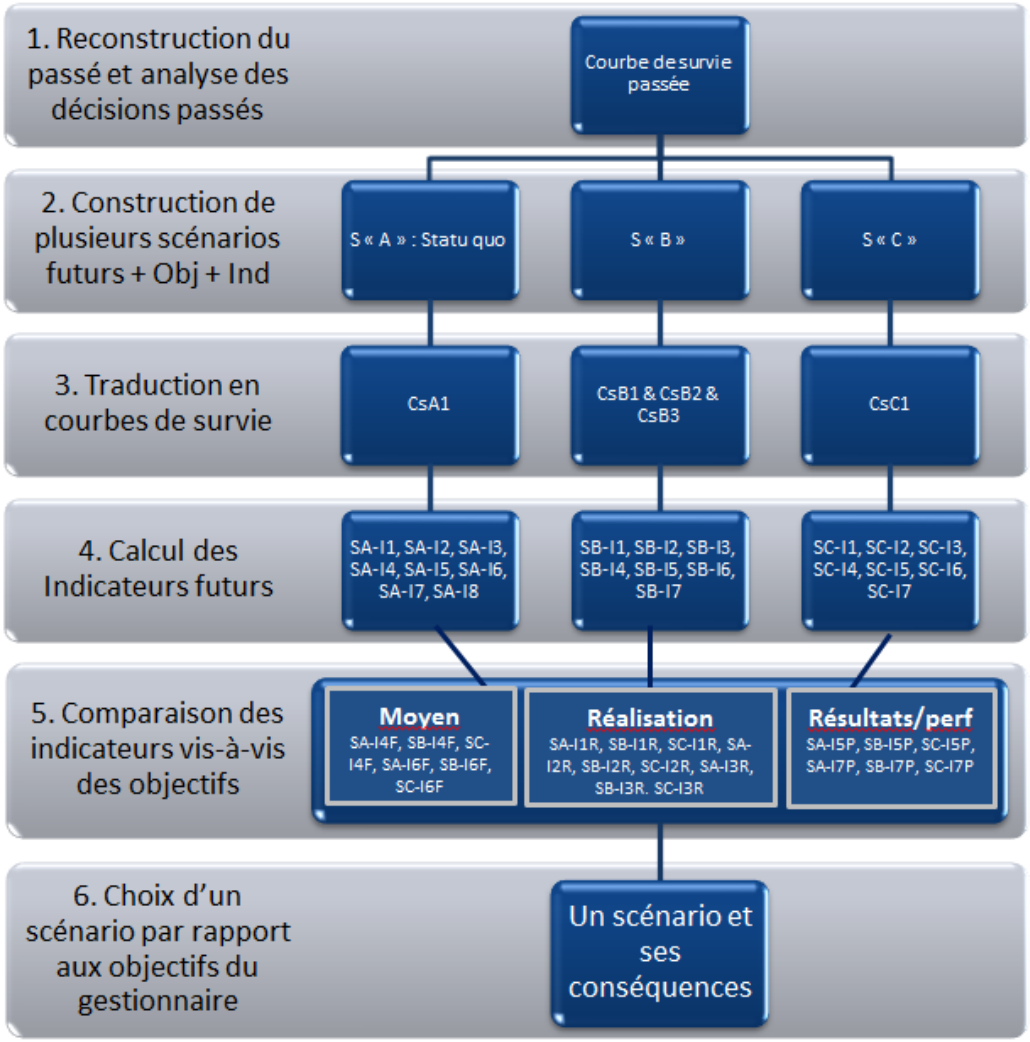


Figure 20 : Schéma de la démarche globale du modèle « long terme » optimeau

La première étape comprend l'analyse de la politique de gestion patrimoniale que le service d'eau a réalisé par le passé. C'est ce que nous allons faire pour le SEDIF dans le **chapitre III**.

Au cours de cette étape il est aussi nécessaire de modéliser la courbe de survie empirique passée. Au cours du **chapitre IV** nous allons expliquer les méthodes pour calculer la courbe de survie empirique passée et dans le **chapitre V** nous allons appliquer ces méthodes aux données du SEDIF.

L'étape deux consiste à élaborer différents scénarios prospectifs pour les années à venir.
Par exemple le scénario prospectif « A » peut être : « A l'avenir, la distribution des âges à la mise hors service des canalisations sera strictement identique à celle du SEDIF par le passé ». C'est le scénario que nous avons choisi dans le **chapitre VI**.
Néanmoins les scénarios prospectifs peuvent être plus complexes. Par exemple le scénario prospectif « B » peut être : « Les canalisations avec une forte probabilité de défaillances seront renouvelés plus tôt en âge que par le passé (décalage de la distribution des âges à la mise hors service vers la gauche) alors que les canalisations avec une faible probabilité de défaillances seront renouvelés plus tard en âge que par le passé (décalage de la distribution des âges à la mise hors service vers la droite) ».
Les scénarios prospectifs peuvent être encore plus complexes pour refléter au mieux la réalité. Par exemple le scénario prospectif « C » peut être : « Les canalisations en PVC et en amiante ciment doivent être éliminés. Les canalisations ayant une forte probabilité d'induire des dommages considérables suite à leurs défaillances seront renouvelées en priorités. »

A cours **de l'étape trois,** il est nécessaire de traduire chaque scénario prospectifs en une ou plusieurs courbes de survie prospective. Par exemple ici le scénario « statut quo simple» a été traduit en une seule courbe de survie cf. Figure 50 alors que le scénario « statu quo partitionné par diamètre » a été traduit en 3 courbes de survie cf. Figure 56**.**

Lors **de l'étape quatre,** des indicateurs futurs sont déduits pour chaque scénario. Nous avons prévu à minima de calculer les indicateurs suivants.

- I1-R le linéaire à renouveler chaque année, cf. **paragraphe VI.1.2**
- I2-R le taux de renouvellement annuel, cf. **paragraphe VI.1.3**
- I3-R l'âge moyen du réseau chaque année, **paragraphe VI.1.4**
- I1-F le coût d'investissement annuel, cf. **paragraphe VI.1.5**
- I1-P le nombre de défaillances au tronçon annuel,
- I2-F le coût de la maintenance annuelle,
- I2-P le niveau de risque au tronçon annuel[6].

Ensuite aux cours de **l'étape cinq**, il faut comparer les différents scénarios entre eux par le biais des indicateurs ci-dessus.

Lors de **l'étape 6** le gestionnaire par rapport à ses objectifs et ses contraintes (budgétaires ou autres) va choisir alors le scénario qu'il souhaite appliquer à l'avenir en ayant en mémoire ses conséquences (mise en évidence par les 7 indicateurs associés).

Nous venons de présenter de manière synthétique la démarche globale de notre modèle « long terme » que nous souhaitons créer dans le projet optimeau.

Nous allons maintenant démarrer l'étape 1 en résumant les méthodes employées par le passé par le SEDIF pour la gestion patrimoniale de ses canalisations d'eau potable.

[6] R : indicateur de réalisation dans le cycle de l'action. F : indicateur financier ou de moyen économique dans le cycle de l'action. P : indicateur de résultats (≈performance) dans le cycle de l'action

III Présentation de la gestion du patrimoine canalisation du SEDIF

III.1 Présentation du renouvellement passé du SEDIF

Le taux de renouvellement annuel de canalisation au SEDIF pour ces quinze dernières années est compris entre environ, 0.3% et 0.76% (Cf. Figure 21). Pour une année ce taux est obtenu en divisant le linéaire qui est mis hors service cette année (source SIG) par le linéaire en service de cette même année (source RPQS). Conformément avec la loi Barnier du 2 février 1995 (codifié dans l'article L2224-5 du CGCT) le SEDIF calcule des taux de renouvellement moyennés sur les 5 dernières années (Cf. Figure 22). Nous constatons une différence de résultat entre les taux moyennés calculés grâce au SIG et ceux données par le « Rapport sur le Prix et la Qualité du Service » (RPQS). D'après (Optimeau 2014) cette différence pourrait peut-être être expliqué par :

- le fait que le taux de renouvellement est calculé à partir de la ligne budgétaire " renouvellement " alors que les linéaires HS dans le SIG le sont à partir des travaux liés à ligne budgétaire " renouvellement " et " tiers " ;
- en général il y a un temps de latence entre la date effective des travaux de renouvellement et leurs remontées dans le SIG ;
- une petite partie des tronçons mis hors service ne sont pas renouvelés.

Donc dans ces 3 cas le taux de renouvellement RPQS a tendance à être inférieur à celui calculé à partir du SIG (tronçons mis HS).

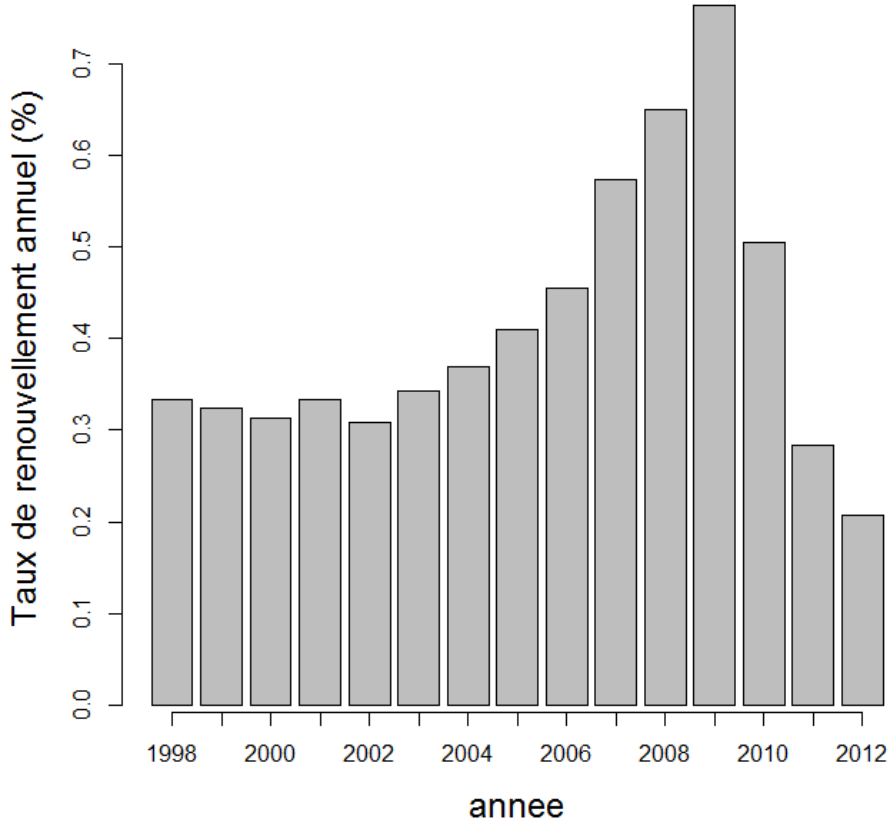


Figure 21 : Taux de renouvellement annuel, calculé à partir du linéaire hors service dans le SIG

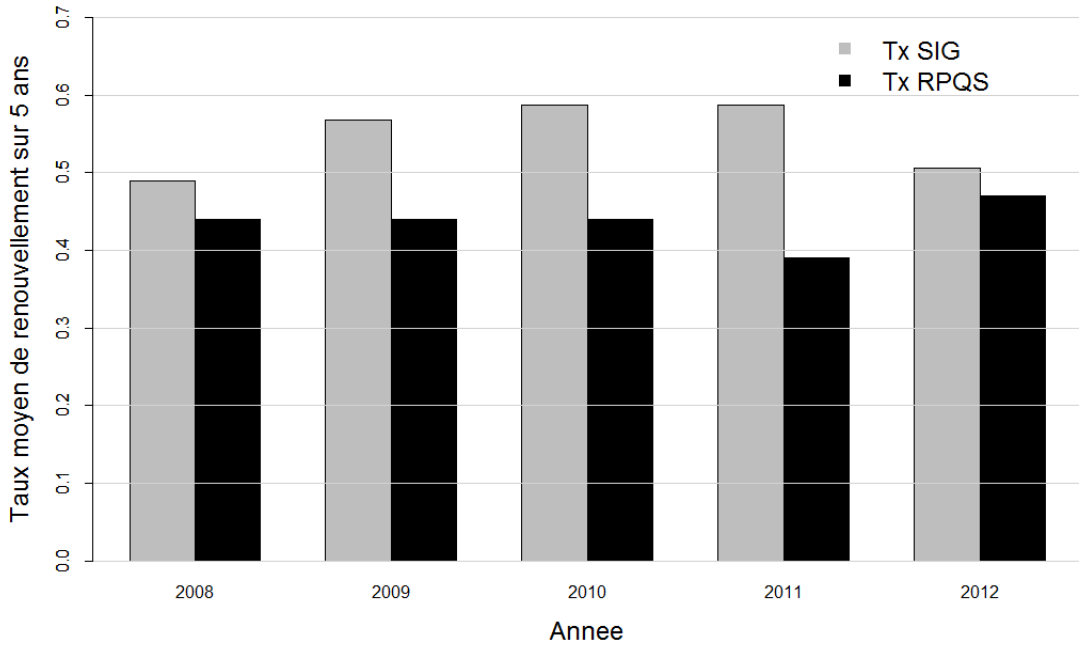


Figure 22 : Taux moyen de renouvellement sur 5 ans

III.2 Méthode décisionnelle de « court terme » (≤ 3 ans) pour renouveler les tronçons au SEDIF

Nous appelons méthode « court terme », le processus décisionnel permettant d'obtenir pour l'année à venir (ou la suivante) une liste de tronçons hiérarchisés par niveau de besoin en renouvellement ou opportunité / obligation de renouvellement (Large 2013b).

La méthode « court terme » (c'est-à-dire à l'échelle du programme des travaux, 1 à 3 ans max) appliquée au SEDIF est assez robuste. Une série d'étapes a été mise au point par les gestionnaires qui utilisent notamment plusieurs logiciels permettant d'établir une liste de tronçons hiérarchisés selon le niveau d'opportunité du renouvellement.

Avant 2011 (date du dernier contrat), le SEDIF et le VEDIF utilisaient la même méthode (cf. méthode multicritères « Vétusté » en Annexe 7: Les critères de renouvellement des canalisations par le VEDIF sur le territoire du SEDIF en 2010) pour le choix des canalisations à renouveler. Après 2011, VEDIF a gardé sa méthode et le SEDIF utilise une autre méthode utilisant notamment les résultats du logiciel « Casses ». Actuellement, sur le territoire du SEDIF, le renouvellement des canalisations est assuré en partie par le SEDIF (22 km/an en 2011-2012, 44 km/an en 2013-2014 et 66 km/an en 2015-2022) et en partie par son délégataire VEDIF (18.2 km/an en 2011-2022). Les processus décisionnels du SEDIF et de Véolia ne sont pas les mêmes. De plus, chaque année leurs méthodes décisionnelles « court terme » peuvent changer sensiblement. Par conséquent nous ne présenterons qu'une seule année.

III.2.1 Méthode décisionnelle de « court terme » employée par le SEDIF

Le processus décisionnel du SEDIF peut se diviser en 2 étapes. (Cf. Figure 23)

1. La probabilité de défaillances (fuites) futures est estimée pour chaque tronçon. Pour les défaillances impactant la quantité d'eau (fuites) l'estimation est effectuée à l'aide du logiciel « Casses » via un modèle statistico-probabiliste (Le Gat, 2009). Les données

d'entrée de ce modèle sont les données sur les défaillances FR, les caractéristiques des tronçons et de leur environnement.

Par exemple, à partir des données 2012 et de l'utilisation du logiciel « Casses » par (Poulton & Saignol 2013) le taux de défaillances prévu en 2013 (TxD) pour les tronçons en acier pouvait s'écrire selon l'équation suivante :

$$\mathrm{TxD} = [(1 + 2.9NDO) \times 1.36t^{0.36} \times \exp(-2.32 + 1.1\ln(Lg) - 0.004Ø + 0.004PZ - 0.32E)]/(Lg \times 100) \quad \textit{Eq. 1}$$

NDO : nombre de défaillances FR total sur le tronçon jusqu'en 2012
t : âge du tronçon en 2013 (en siècle) = 2013 – année de pose
Lg : longueur du tronçon (en mètre)
Ø : diamètre du tronçon (en mm)
PZ : niveau piézométrique, niveau du trop-plein du réservoir auquel est relié le tronçon selon le NGF (nivellement général de la France en mètre)
E : niveau d'enfouissement, différence entre l'altitude du sol et l'altitude du tronçon (en mètre)

2. Le gestionnaire prend ensuite en compte d'autres critères d'arbitrage tels que les opportunités et les contraintes externes à son service. Elles répondent à des besoins de coordination avec d'autres services (travaux de voiries essentiellement, etc.), ou sont induites par des demandes de tiers. Ce sont des demandes de renouvellement suite à des défaillances ayant provoquées beaucoup de dégâts, ou pour mise aux normes de défense incendie, etc. En croisant toutes ces informations, le gestionnaire établit la liste définitive des chantiers (groupes de tronçons adjacents) à mettre en œuvre.

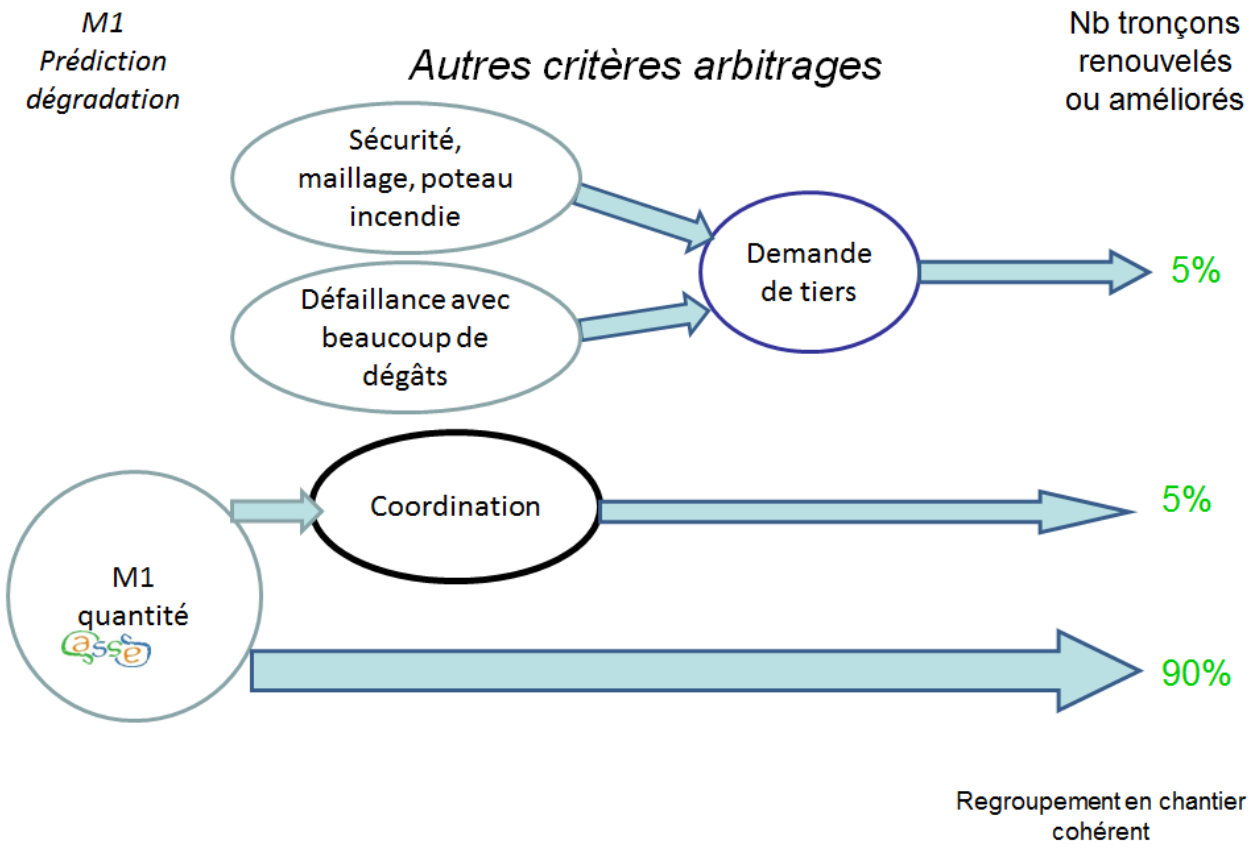


Figure 23 : Les étapes de la décision de la mise hors service de tronçons au SEDIF à court terme (en 2013)

III.2.2 Méthode décisionnelle de « court terme » employée par le VEDIF

Le processus décisionnel court terme de VEDIF est basé sur une analyse multicritères (cf. Annexe 7). Par exemple en 2010 :

- la prévision de la dégradation est principalement approchée par le nombre de défaillances FR à l'hectomètre sur les 5 dernières années (TxD5) (cf. Figure 25). Néanmoins, d'autres critères peuvent être pris en compte tels que la longueur du tronçon, son âge et son matériau. Il est important de préciser que le calcul du TxD5 ne se fait pas sur les « tronçons SIG » car certains sont trop petits, mais sur des « tronçons SIG » agrégés nommés « tronçons chantiers ». Néanmoins cette notion n'est pas formalisée. Les « tronçons chantiers » sont rarement inférieurs à 100m alors que le plus petit « tronçon SIG » fait 1 cm.
- Une autre approche de la prévision de la dégradation basée sur d'autres outils (comme les logiciels « Casses » ou « MOSARE »), complète ce travail de sélection des conduites à renouveler en priorité.
- Véolia prend ensuite en compte les opportunités et les contraintes externes à son service telles que la coordination avec les services de voirie.

Le programme de renouvellement se décompose en un programme initial, complété par un programme complémentaire si nécessaire.

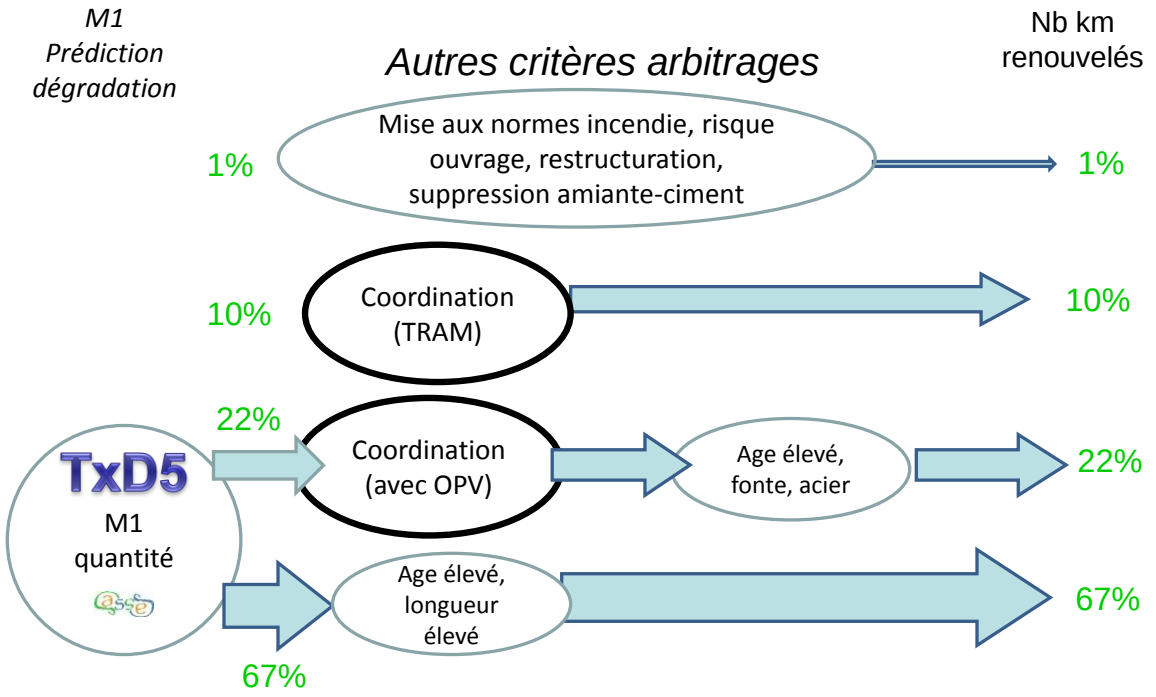


Figure 24 : Les étapes de la décision du VEDIF de la mise hors service de tronçons à court terme en 2010

Causes de renouvellement (mètre de canalisations)	Défaillances	Coordination	Défaillances et coordination	Autres	Total
TxD5 > = à 2	3 853				
Tronçons > 100 m et 1,8 < TxD5 < 2	1 441				
Tronçons > 150 m et 1,5 <TxD5 < 2	15 791				
TxD5 >=1	896				
Age élevé et TxD5 >= 0,5	230				
Taux de défaillances prévus par "Casses" élevé	4 137				
Coordination obligatoire (voirie)		6 596			
Coordination (voiries) avec incertitude		6 495			
Coordination et (âge avancé et TxD5 élevé ou fonte à risque ou acier)			7 060		
Coordination et (TxD5 très élevé ou fonte à risque ou acier)			846		
Risque pour les éléments vulnérables "ouvrages "				470	
Mise aux normes incendies + risque				200	
Restructuration du réseau (abandon de conduite, etc.)				280	
suppression amiante ciment				320	
Total (mètre)	26 348	13 091	7 906	1 270	48 615
Total (%)	54%	27%	16%	3%	100%
programme complémentaire					
Programme initial					
TxD5 : taux de défaillances FR sur les 5 dernières années					

Figure 25 : Les causes de renouvellement de canalisations par VEDIF dans le programme prévisionnel en 2010 source : (Bruyeron & Re 2012a; Bruyeron & Re 2012b)

Il est à noter qu'une partie du programme initial et du programme complémentaire a été réalisée en 2010 puisque selon le SIG seuls 44 560 mètres de canalisations ont été renouvelés en 2010 sur les 48 615 prévus.

L'année 2010 est l'année avant le changement de contrat (2011) de délégation. Or cette année-là d'après (Optimeau 2014) il y a eu un grand décalage entre le prévisionnel présenté dans la Figure 25 et la réalité. En effet selon (Optimeau 2014) ces dernières années : 70% des tronçons renouvelés l'ont été principalement sous la contrainte coordination contre 30% un peu plus choisi.

III.3 Méthode décisionnelle long terme (> 70 ans)

Le SEDIF utilise actuellement plusieurs méthodes afin d'avoir une vision sur le long terme (c'est-à-dire à l'échelle de la durée de maintien en service des canalisations).

III.3.1 Méthode « Cador »

À partir des durées de maintien en service fixées « à dires d'expert » pour chaque matériau (cf. Figure 26), les besoins futurs en renouvellement sont estimés année par année. Cette méthode se résume simplement à additionner ces durées de maintien en service « à dires d'expert » à la date de pose de chaque tronçon pour obtenir la date de mise hors service présumée de ce tronçon (cf. Figure 28).

Matériaux	Linéaire en service (km) fin 2012	Linéaire en service (%) fin 2012	Durée moyenne de maintien en service selon Expert (ans)	Durée de maintien en service cible selon Expert (ans)
Fonte grise	3 220	38%	70	35
Fonte ductile	3 125	37%	120	60
PEHD	1 118	13%	100	50
Bonna	596	7%	120	60
Acier+ tôle acier	183	2.2%	60	30
Autres	140	1.7%	60 ou 85	30 ou 43
Tous confondus	**8 382**	**100%**	**95**	**47.5**
appelation SEDIF			Durée de vie	Durée de demi-vie

Figure 26 : Les durées moyennes de maintien en service selon les experts du SEDIF (Bruyeron & Re 2012b)

Cette démarche ne convainc pas les praticiens du SEDIF car elle fait l'hypothèse que les canalisations doivent être remplacées à âge fixe, ce qui est non pertinent. Effectivement, en fonction des contraintes que subissent les canalisations, certaines très âgées peuvent persister en bon état tandis que d'autres plus jeunes, sont en mauvais état. De plus, aucun lien n'est fait entre cette valeur « à dires d'expert » et le processus décisionnel opérationnel de « court terme » (coordination avec les travaux de voirie, risques, etc.), dont résulte effectivement l'âge à la mise hors service. Ensuite, pour les matériaux récents (PEHD, fonte ductile) âgés seulement d'une quarantaine d'années, il est très difficile d'estimer leur durée de maintien en service en raison d'un manque de retour d'expérience. Enfin cette méthode fait l'hypothèse que la survie par matériau est une fonction heaviside (fonction escalier d'une marche) (cf. Figure 27) ce qui n'est pas du tout réaliste (raisonnement on/off). La survie par matériau est une fonction plutôt sigmoïde décroissante (pente de « montagne ») (cf. paragraphes suivants).

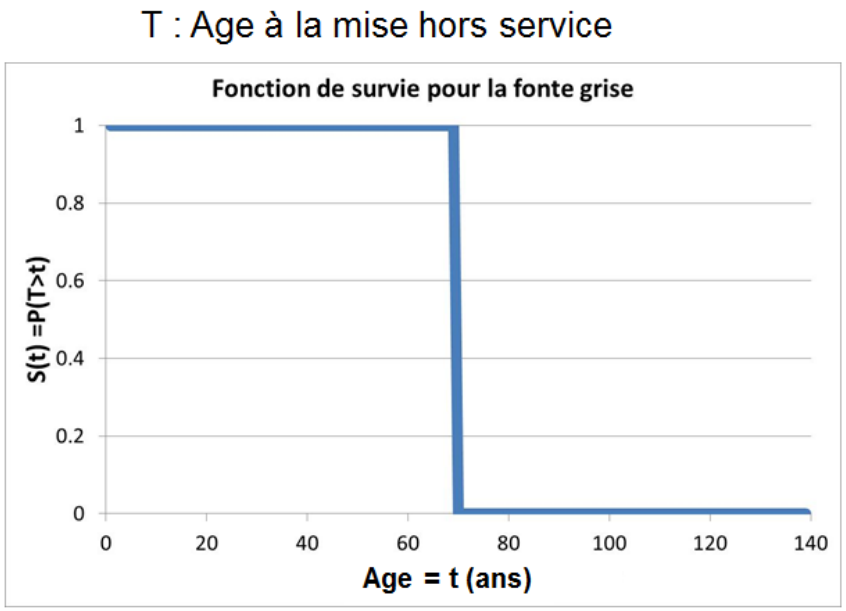


Figure 27 : Fonction de survie des tronçons en fonte grise selon les experts du SEDIF.

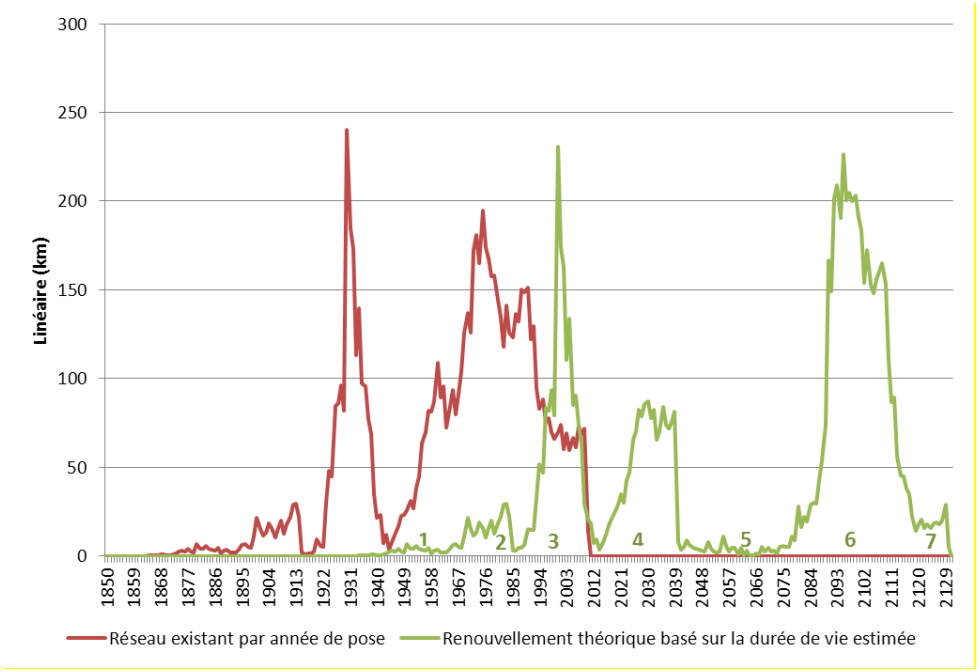


Figure 28 : Prévision des besoins en renouvellement selon la méthode « Cador »(Bruyeron & Re 2012b)

Les prévisions issues de la méthode « Cador » induisent des variations trop brusques dans les taux de renouvellement à venir. Par exemple cette méthode prévoit un pic de renouvellement en 2095 de 225 km (Cf. Figure 28). Ce qui correspond à un taux maximal de renouvellement de 225/8300 =2.7%. Cette méthode prévoit un minimum de renouvellement de 2 km en 2065. Ce qui correspond à un taux minimal de renouvellement de 0.02%.

Après avoir estimé les besoins en renouvellement (en km), la multiplication par des coûts unitaires permet d'estimer le budget nécessaire (cf. Figure 29). Les besoins techniques déterminent ici le budget. (Bruyeron & Re 2012b)

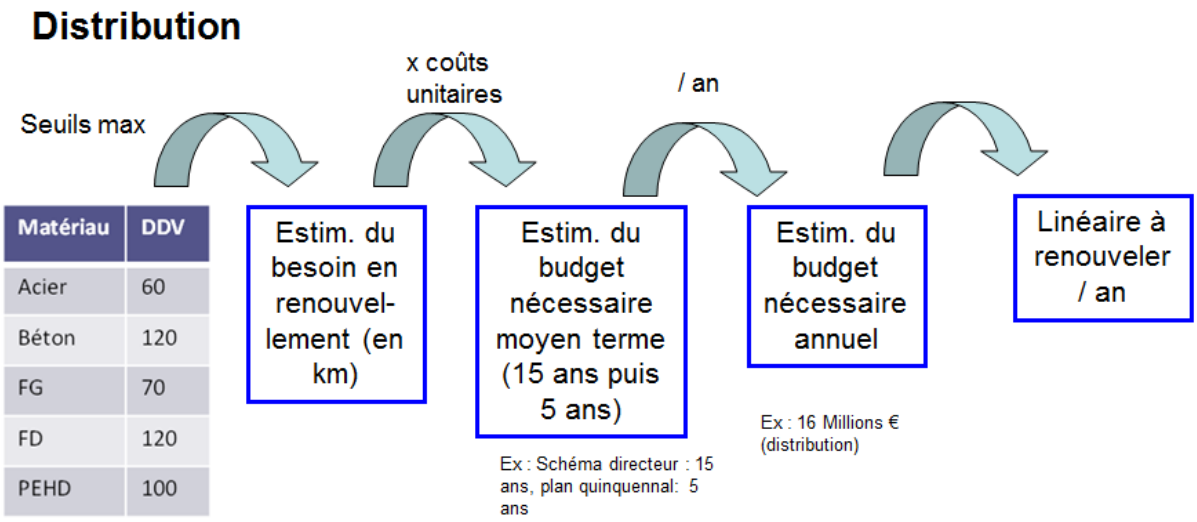


Figure 29 : Les étapes de la décision de la mise hors service de tronçons du SEDIF sur du long terme

III.3.2 Prévision de l'évolution de l'âge moyen du réseau

Afin d'obtenir une vision sur l'évolution de l'âge moyen du réseau à long terme, les experts du SEDIF font l'hypothèse que le renouvellement futur portera uniquement sur les canalisations les plus anciennes.
Puis l'évolution de l'âge moyen du réseau est calculée à partir de 2 taux de renouvellement futur :

1. Taux de renouvellement après 2013 identique aux 5 dernières années
2. Taux de renouvellement de 0,8 à 1%

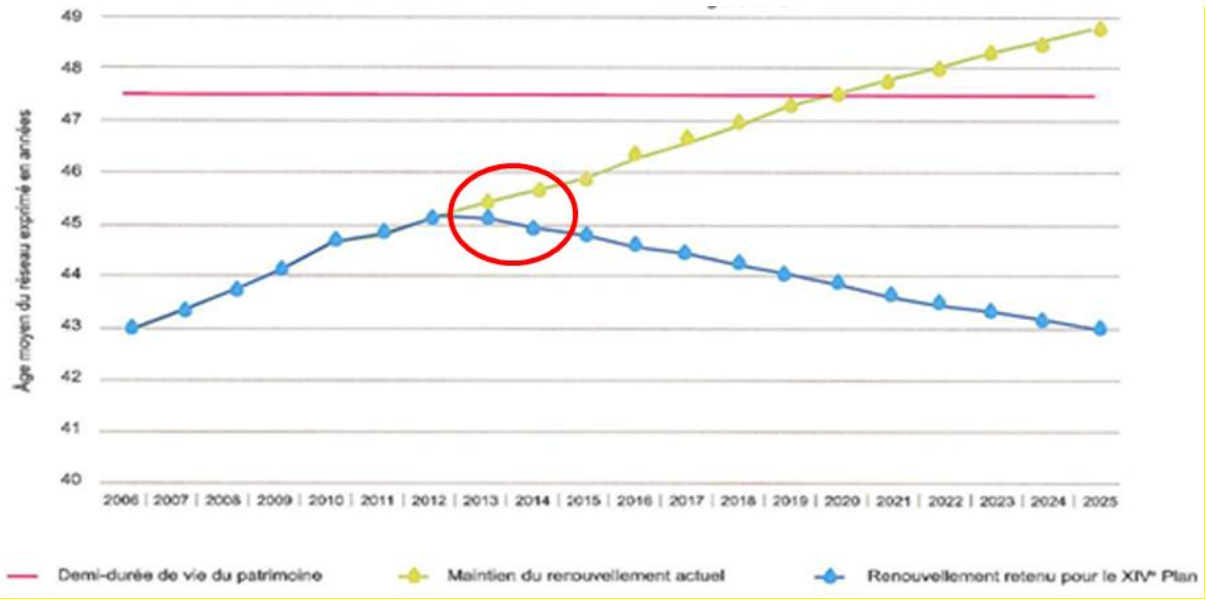


Figure 30 : Prévision de l'évolution de l'âge moyen du réseau (Large 2013a)

Malheureusement cette hypothèse va à l'encontre des pratiques réelles « court terme » du SEDIF. Ce ne sont pas les canalisations les plus anciennes qui sont remplacées mais celles ayant notamment la probabilité de défaillance la plus élevée.

Ainsi les prévisions sur l'évolution de l'âge moyen du réseau, liées à ces pratiques réelles (cf. courbe bleu foncé Figure 31) sont de 2 ans au-dessus des prévisions sur l'évolution de l'âge moyen (avec un taux de renouvellement après 2013 identique aux 5 dernières années).

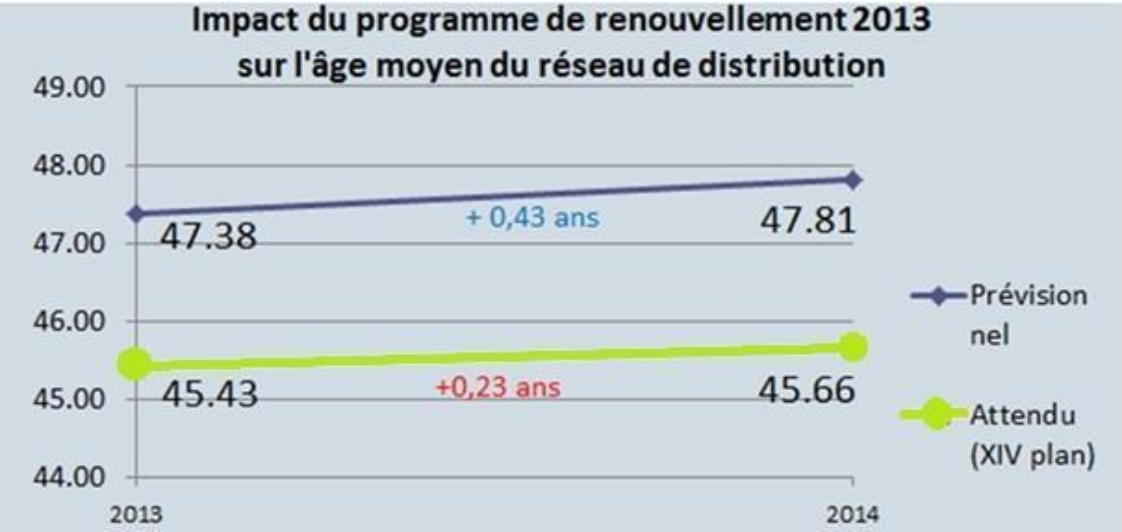


Figure 31 : Comparaison de prévisions à la réalité (Large 2013a)

III.3.3 Linéaire à renouveler fixe à atteindre

Le nouveau contrat de 2011 (cf. article 27 bis), fixe les linéaires de canalisations à renouveler par le VEDIF et par le SEDIF.

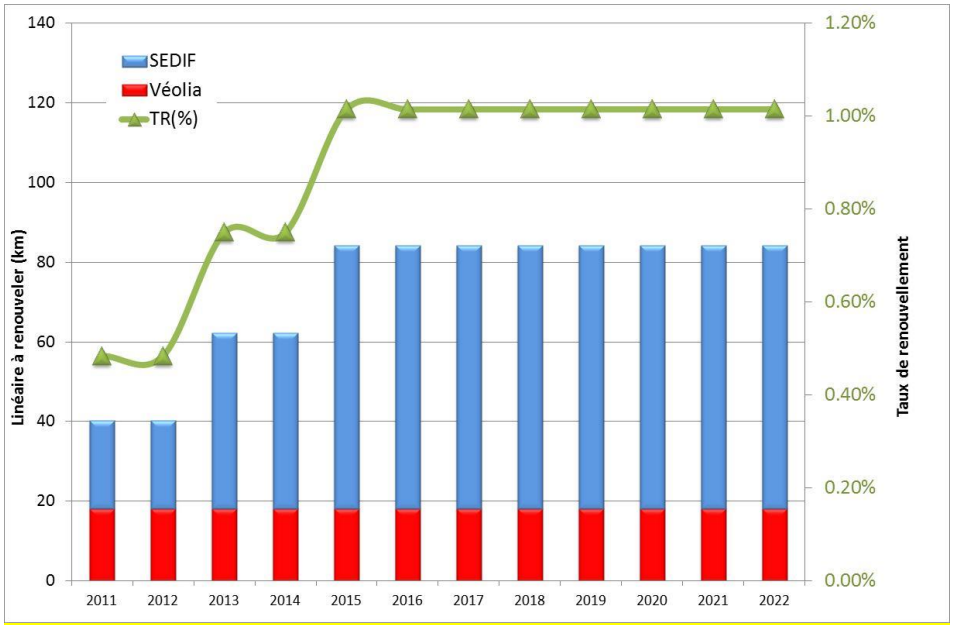


Figure 32 : Répartition du linéaire à renouveler entre SEDIF et VEDIF jusqu'en 2022 pour le réseau de distribution (Large 2013a)

Le renouvellement du réseau de transport doit être réalisé à 100% par le SEDIF pour 7km/an en moyenne sur la période 2011-2025 (Large 2013a).

III.3.4 Patrimoine expert

La méthode contenue dans le logiciel « Patrimoine expert » est très proche de la méthode « Cador ».
Cet outil effectue des simulations du renouvellement des canalisations sur une période longue à budget renouvellement annuel fixe selon l'âge des canalisations (cf. Figure 33). L'outil conseille de renouveler les canalisations qui ont l'indice de remplacement (IR) le plus élevé (Cf. Tableau 3).

$$\text{IR1}(\%) = \left(\frac{\hat{a}ge}{DS}\right) \times 100$$

$$IR2(Classe) = IR1(Classe) + C1 + C2 + C3 + C4$$

Avec :
t : âge du tronçon
DS : Durée de maintien en service par matériau fixée à dires d'expert. Elle est fonction du matériau et du diamètre.
C : 4 critères de modulation au tronçon. Pour chaque critère, le gestionnaire doit attribuer à chaque tronçon une valeur numérique parmi : {0, 1, 2, 3, 4}. Par exemple le critère de modulation peut-être : la pression de l'eau ou le niveau de corrosivité du sol ou des caractéristiques de l'environnement (trafic routier), ou l'indice linéaire de perte, etc. Mettre un 1 sur un critère de modulation signifie que l'indice de remplacement sera augmenté d'une classe pour la canalisation concernée.

IR1(%)	IR1(Classe)
[0,75%]	1
]75,100%]	2
]100%,125%]	3
]125%,150%]	4
>150%	5

Tableau 3 : Correspondance entre l'indice de renouvellement en pourcentage et l'indice de renouvellement en classe (Naldeo 2013).

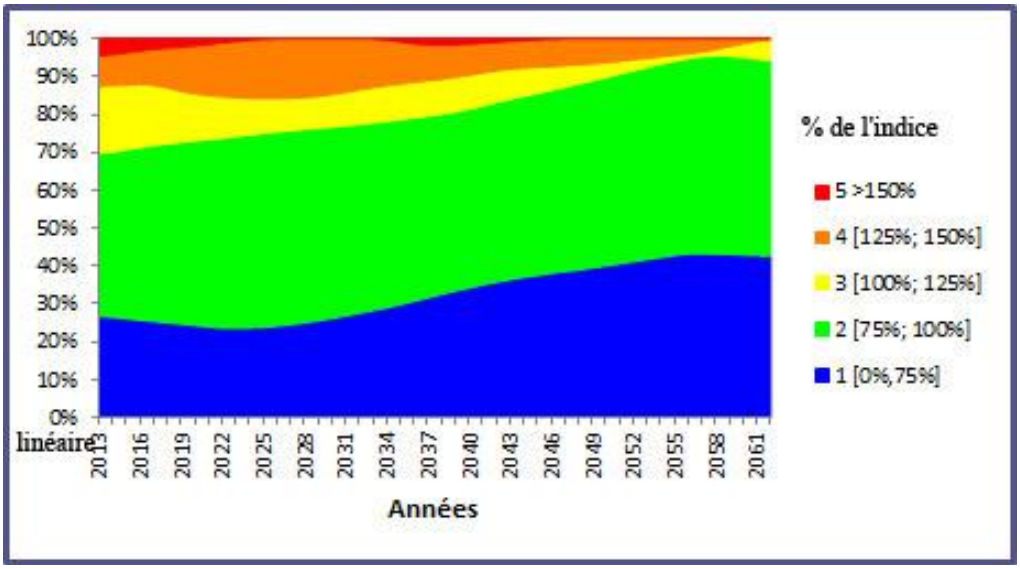


Figure 33 : Exemple de sortie du logiciel Patrimoine expert (SEDIF 2013b)

Les données d'entrée de ce modèle sont notamment :
- les caractéristiques des canalisations à l'échelle du segment (groupe de tronçons homogènes adjacents) : matériau, diamètre, linéaire et date de pose ;
- les durées de maintien en service des canalisations ;
- les prix unitaires de remplacement des canalisations ;
- le montant annuel d'investissement pour renouveler les canalisations ;
- et le nombre d'années de la simulation.
(Naldeo 2013)

Les inconvénients majeurs de ce modèle sont :
- les durées de maintien en service sont fixées à dires d'expert ;
- le budget est une donnée d'entrée alors que les gestionnaires attendent que cela soit une donnée de sortie sur du long terme ;
- les critères complémentaires (C1, C2, C3 et C4) de modulation ont un poids trop important dans le calcul de l'indice de remplacement.

Nous venons de décrire les méthodes « court terme » employées par VEDIF et par le SEDIF sur le territoire du SEDIF. Puis nous avons exposé les différents modèles « long terme » (Cador, Patrimoine expert, etc.) utilisé par le SEDIF. En résumé, les méthodes « court terme » (1 à 3 ans max) appliquée au SEDIF sont nombreuses et assez robustes. En revanche les méthodes « long terme » (> 30 ans) ont des lacunes importantes.

Nous allons maintenant présenter les méthodes employés pour calculer la survie empirique passée, faisant partie de la première étape nécessaire de notre nouveau modèle « long terme ».

IV Méthode pour le calcul de la survie empirique passée

IV.1 Préliminaires

Pour tous les travaux effectués nous avons utilisé le logiciel « R », version 3.1.0. Nous avons écrit nous-même la majorité des codes (Turnbull, Kaplan-Meier, Nelder-Mead, etc.). Cependant afin de valider nos codes, nous avons aussi utilisé certaines fonctions du package « Survival » notamment pour le modèle de Cox ou de Kaplan-Meier.

Les atouts majeurs de ce logiciel sont multiples :

- Il est un standard international de calcul statistique ;
- Il permet une traçabilité des transformations opérées sur les données brutes, et sur les modèles ;
- Il améliore grandement la reproductibilité du travail ;
- Les codes sont réutilisables notamment entre les différents acteurs du projet ;
- Les codes pourront ensuite être intégrés dans un logiciel avec une interface homme machine conviviale.

IV.2 Nettoyage des données

IV.2.1 Nettoyage des données tronçons

Avant de commencer toute étude statistique, un nettoyage des données est nécessaire pour optimiser la qualité de l'ensemble des données. En effet, les données remises par le SEDIF comportent quelques incohérences qu'il faut prendre le temps de corriger. (Cf. Figure 34)

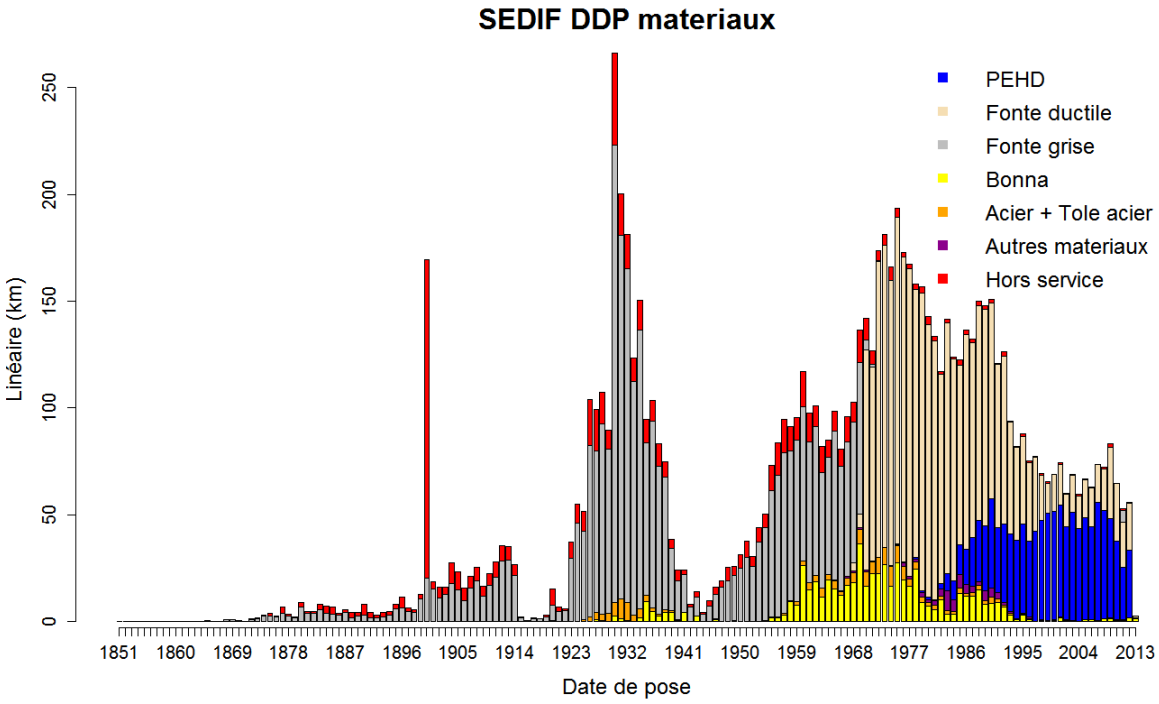


Figure 34 : Linéaire en fonction des dates de pose avant nettoyage

Des tests préliminaires sur les variables ont permis de déceler des incohérences dans les données :

1. Dans la variable des types de matériaux nous avons trouvé l'existence de canalisations en fonte ductile (FD) posées avant 1966. Or ce matériau n'était pas utilisé en ce temps. Nous avons décidé de remplacer par de la fonte grise (FG), très présente au début du 20ème siècle. Nous avons fait d'autres modifications analogues pour d'autres matériaux cf. : Tableau 4. Les dates sur lesquelles nous nous sommes appuyés pour effectuer ces changements proviennent de (Charrière 2013), (Poinard 2006) et de (SEDIF 2013a).

Si nous trouvons :	Nous remplaçons par :
Fonte < 1972	Fonte grise
Fonte ≥ 1972	Fonte ductile
Fonte grise > 1972	Fonte ductile
Fonte ductile < 1966	Fonte grise
PEHD < 1974	Autre

Tableau 4 : Modifications réalisées sur les dates de pose en fonction du matériau

2. Le pic observable à l'année de poses de 1900 (cf. Figure 34) correspond à l'attribution arbitraire par Véolia d'une série de tronçons hors service dont toutes les dates de poses étaient inconnues. Ils ont alors été supprimés de la base de données afin de ne pas fausser l'étude.

3. Pour la variable représentant la longueur[7] des tronçons, toutes les longueurs non renseignées et toutes les valeurs négatives ou nulles sont remplacées par la longueur calculée par le SIG[8] grâce aux coordonnées spatiales.

4. Pour faciliter l'analyse des données, des modifications ont été effectuées sur la variable « statuts ». Composée initialement de trois modalités : la modalité 3 pour « en fonctionnement », 7 pour « supprimé » et 9 pour « abandonné ». Elle a alors été réduite à deux modalités : « Hors service » si abandonné ou supprimé et « En service » si en fonctionnement.

5. Afin d'avoir une photographie de la situation au 31/12/2012. Les tronçons posés en 2013 sont supprimés et les tronçons mis hors service en 2013 retrouvent le statut « En service ».

6. Vu que la fenêtre fiable des dates de mise hors service des tronçons est [1995; 2012]. Les tronçons mis hors service avant 1995 sont supprimés.

7. Après un certain nombre de tests, 136 tronçons avec un statut « En service » et une date de mise hors service ont été observés. Or il est impossible qu'un tronçon soit en service avec une date de mise hors service. Pour corriger ce problème, toutes les dates hors service de ces individus ont été remplacées par « Not Available » (NA). La variable de date de mise hors service est notée « DHS ».

8. Toutes les dates de pose et dates de mise hors service non renseignées ont également été remplacées par « NA ». La variable des dates de pose est constituée des années durant lesquelles ont été posés les tuyaux, elle est notée « DDP ».

[7] Champ intitulé "Long Ref » dans le SIG
[8] Champ intitulé « shape Length » dans le SIG

9. Enfin, une variable « t » qui correspond à l'âge du tronçon à sa mise hors service, est obtenue en faisant le calcul suivant :

$$t = DHS - DDP$$

10. Ces modifications ont permis de mettre en avant l'existence de 104 tronçons dont l'âge de mise hors service était négatif. L'abréviation « NA » a été affectée à tous ces individus dans la variable des dates de mise hors service.

IV.2.2 Nettoyage des données défaillances FR

Nous avons nettoyé les défaillances FR de la façon suivante :

- la **localisation** des défaillances a été « nettoyée » selon la méthode décrite dans l'Annexe 2 afin de ne conserver uniquement que les 4 catégories suivantes : sur tronçon, sur joint, sur appareillage, et inconnue.
- Ensuite, nous avons conservé uniquement les défaillances FR observables sur notre fenêtre d'observation **[1995, 2012]**

Enfin nous avons pu rattacher aux tronçons, leurs défaillances (100 % des défaillances ont pu être rattaché à aux tronçons) et ainsi nous avons créé l'indicateur **NDOA** qui est le nombre de défaillance moyen observé par an sur la fenêtre d'observation [1995, 2012].

IV.3 Nos données tronquées à gauche, censurées à droite

Comme nous l'avons vu partie I.2.4, Figure 9, la période d'observation du linéaire hors service s'étend sur 18 ans, de 1995 à 2012. Avant 1995 les dates de mise hors service n'étaient pas enregistrées ou commençaient seulement à l'être sur le SIG. Nous ne connaissons donc pas les durées de maintien en service des tronçons mis hors service avant cette date, ils sont donc retirés de la base de données. Nos données ne sont alors complètement observables que sur la fenêtre [1995, 2012]. Ainsi l'estimation de courbe de survie sera faite sur des données tronquées à gauche.

Les données sont censurées à droite du fait que tous les tronçons ne sont pas encore hors service à l'arrêt des observations fin 2012. La fenêtre d'observation étant limitée nous observons la réelle mise hors service d'une faible proportion du réseau. Pour la plupart des tronçons observés, la valeur réelle de la durée de maintien en service est inconnue car elle est supérieure à l'âge atteint à la fin de la fenêtre d'observation.

Notons θ_1 l'année de début d'observation (1995), θ_2 l'année de fin d'observation (2012). Pour chaque tronçon i, posons : a_i l'âge du tronçon en θ_1 ($a_i = 0$ an si le tronçon a été posé après θ_1), b_i l'âge du tronçon à sa date de mise hors service ou son âge en θ_2, et T la variable aléatoire représentant la durée de maintien en service. La Figure 35 illustre les concepts de troncature à gauche et de censure à droite.

- Toute date de mise hors service inférieure à θ_1 est inobservable (troncature à gauche).
- Les dates de mise hors service entre $[\theta_1 ; \theta_2]$ sont exactement observées.
- Les dates de mise hors service postérieures à θ_2 ne sont pas observées et les âges à la mise hors service correspondants sont censurés à droite.

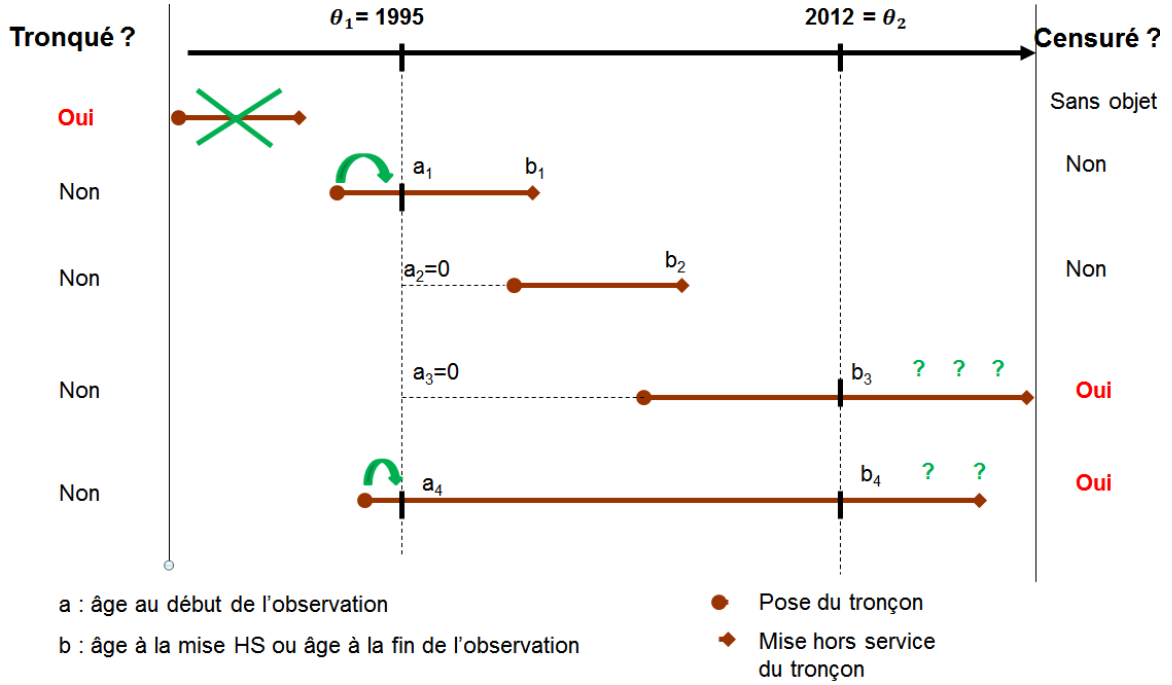


Figure 35 : Dispositions possibles des tronçons vis-à-vis de la fenêtre d'observation, illustration de la troncature à gauche et de la censure à droite

Enfin les courbes de survie sont estimées tantôt à partir du linéaire de tronçons et tantôt à partir de leur nombre. Les terrains d'étude sont plus partisans des courbes de survie basées sur le linéaire. En effet, vu que tous les tronçons ont des longueurs différentes il est plus réaliste d'exposer des résultats en termes de linéaire hors service chaque année plutôt qu'en nombre de tronçons. D'un point de vue statistique cela ne diffère pas significativement d'une estimation de courbes de survie en nombre de tronçons (cf. démonstration ci-après).

IV.4 Survie empirique des tronçons hors service et en service

IV.4.1 Kaplan-Meier simple pour données censurées à droite

L'estimateur de Kaplan-Meier (Kaplan & Meier 1958) permet de calculer une fonction de survie d'après des données de durées de maintien en service. Il repose sur **l'idée** qu'être en service après un temps *t*, c'est être en service juste avant *t* (= *t*-1 = *t* moins une année) et ne pas passer hors service au temps *t*. L'évènement d'intérêt ici est la mise hors service d'un tronçon.

Considérons les notations suivantes :

- t_j avec $(j = 1, \dots, m)$ les âges à l'évènement, des évènements (décès ou censure) rangés par ordre croissant, avec $t_{(0)} = 0$
- Y_j le nombre d'individus à risque de subir l'évènement d'intérêt juste avant le temps t_j,
- d_j le nombre de mise hors service en t_j.

Selon la définition de la fonction de survie on a :

$$\hat{S}(t_j) = P(T > t_j) \qquad \textit{Eq. 2}$$

D'après **l'idée** de Kaplan-Meier, il en découle que :

$$\hat{S}(t_j) = P(T > t_j \; et \; T > t_{j-1}) \qquad \textit{Eq. 3}$$

Or, en théorie des probabilités, la probabilité conditionnelle d'un événement A, sachant qu'un autre événement B de probabilité non nulle s'est réalisé (ou probabilité de A, sachant B) est le nombre noté P(A|B), il est défini par :

$$P(A|B) = \frac{P(A \; et \; B)}{P(B)} \qquad \textit{Eq. 4}$$

D'où

$$\hat{S}(t_j) = P(T > t_j \; et \; T > t_{j-1}) = P(T > t_j \,|T > t_{j-1}) \times P(T > t_{j-1}) \qquad \textit{Eq. 5}$$

Ainsi de suite de proche en proche on obtient :

$$\hat{S}(t_j) = P(T > t_j \,|T > t_{j-1}) \times P(T > t_{j-1} \,|T > t_{j-2}) \times \ldots \times P(T > t_2 \,|T > t_1) \times P(T > t_1 \,|T > 0) \qquad \textit{Eq. 6}$$

Or

$$P(T > t_j \,|T > t_{j-1}) + P(T \leq t_j \,|T > t_{j-1}) = 1 \qquad \textit{Eq. 7}$$

D'où

$$\hat{S}(t_j) = [\,1 - P(T \leq t_j \,|T > t_{j-1})] \times [1 - P(T \leq t_{j-1} \,|T > t_{j-2})] \times \ldots \times [1 - P(T \leq t_2 \,|T > t_1)] \times [1 - P(T \leq t_1 \,|T > 0)] \qquad \textit{Eq. 8}$$

Alors la probabilité p_j de mise hors service dans l'intervalle $]t_{(j-1)}, t_{(j)}]$ sachant que le tronçon était en service en $t_{(j-1)}$, *i.e.* $p_j = P(X \leq t_{(j)} | X > t_{(j-1)})$, peut être estimée par

$$\hat{p}_j = \frac{d_j}{Y_j} \qquad \textit{Eq. 9}$$

D'où

$$\hat{S}(t_j) = \left[1 - \frac{d_j}{Y_j}\right] \times \left[1 - \frac{d_{j-1}}{Y_{j-1}}\right] \times \ldots \times \left[1 - \frac{d_2}{Y_2}\right] \times \left[1 - \frac{d_1}{Y_1}\right] \qquad \textit{Eq. 10}$$

L'estimateur de Kaplan-Meier de la survie est alors un produit de la forme (Saint Pierre 2013):

$$\hat{S}(t) = \prod_{\substack{j=1,\ldots,m \\ t_j \leq t}} \left(1 - \frac{d_j}{Y_j}\right) \qquad \textit{Eq. 11}$$

Et sa variance est estimée par la formule de Greenwood (Aalen et al. 2008):

$$\tilde{\tau}^2(t) = \hat{S}(t)^2 \sum_{t_j \leq t} \frac{d_j}{Y(T_j)\{Y(T_j) - d_j\}} \qquad \textit{Eq. 12}$$

L'intervalle de confiance peut être estimé alors de la façon suivante :

$$IC_{95\%} = \hat{S}(t)^2 \pm 1.96 \times \tilde{\tau}(t) \qquad \textit{Eq. 13}$$

Ou selon la formule de Rothman :

$$IC_{95\%} = \frac{M}{M + (1.96)^2}\left[\hat{S}(t) + \frac{(1.96)^2}{2M} \pm 1.96\sqrt{\tilde{\tau}(t)^2 + \frac{(1.96)^2}{4M^2}}\right] \qquad \textit{Eq. 14}$$

Où

$$M = \frac{\hat{S}(t)\left[1 - \hat{S}(t)\right]}{\tilde{\tau}(t)^2} \qquad \textit{Eq. 15}$$

L'estimateur de Kaplan-Meier présenté ci-dessus permet de corriger la censure à droite.

IV.4.2 Kaplan-Meier élargi (en effectif) pour données censurées et tronquées

Nos données étant tronquées à gauche et censurées à droite nous avons dû ajuster l'estimateur de Kaplan-Meier comme l'indique (Claude & Lyon 1997). Pour cela il suffit de faire très attention dans le comptage de nos individus à risque de subir l'événement d'intérêt.

Soit :

i est le tronçon $n°i$

a_i est l'âge du tronçon en 1995 (âge en début d'observation)

b_i est l'âge du tronçon à sa mise hors service ou s'il est censuré, b_i = âge en 2012

$c_i = \begin{cases} 0 \text{ si l'âge de mise hors service est connu (= individu non censuré)} \\ 1 \text{ si l'âge de mise hors service est inconnu (= individu censuré)} \end{cases}$

t_j avec $(j = 1, \dots, m)$ l'âge aux évènements (décès et censure) rangés par ordre croissant, avec $t_j = 0$

Nous obtenons alors l'estimateur de Kaplan-Meier élargi :

$$\hat{S}(t) = \prod_{t_j \le t}\left(1 - \frac{Card\{i, b_i = t_j \ \& \ c_i = 0\}}{Card\{i, a_i \le t_j \le b_i\}}\right) \qquad \textit{Eq. 16}$$

À chaque âge t, nous calculons 1 moins la somme des tronçons (mis hors service) morts au temps t divisée par la somme de tous les tronçons qui ont un âge en début d'observation inférieur à t et un âge en fin d'observation supérieur à t (donc tous les individus vivants au temps t).

L'estimateur de Kaplan-Meier élargie présenté ci-dessus permet de corriger la censure à droite et prendre en compte la troncature à gauche.

IV.4.3 Kaplan-Meier élargi (en linéaire) pour données censurées et tronquées

Les longueurs des tronçons étant trop disparates, il a été décidé de considérer aussi comme étant un individu, un mètre de tronçon. Ainsi au lieu de compter le nombre de tronçons qui passent hors service à chaque âge, il est compté le nombre de mètres passant hors service à chaque âge. Notons alors :

i le mètre $n°i$

a_i l'âge du tronçon en 1995 (âge en début d'observation)

b_i l'âge du tronçon à sa mise hors service ou s'il est censuré, b_i = âge en 2012

$$c_i = \begin{cases} 0 \text{ si l'âge de mise hors service est connu } (= \text{individu non censuré}) \\ 1 \text{ si l'age de mise hors service est inconnu } (= \text{individu censuré}) \end{cases}$$

t_j avec $(j = 1, \dots, m)$ les temps d'évènements (décès et censure) rangés par ordre croissant, avec $t_j = 0$

Nous obtenons alors l'estimateur de Kaplan-Meier élargi en linéaire :

$$\hat{S}(t) = \prod_{t_j \le t} \left(1 - \frac{Card\{i, b_i = t_j \ \& \ c_i = 0\}}{Card\{i, a_i \le t_j \le b_i\}}\right) \qquad \text{Eq. 17}$$

À chaque âge t, nous calculons le complémentaire à 1 de la somme des linéaires mis hors service au temps t divisée par la somme de tous les linéaires qui ont un âge en début d'observation inférieur à t et un âge en fin d'observation supérieur à t (donc tous les mètres de tronçons en service au temps t).

Il est à noter que la courbe de survie brute (cf. Annexe 9) sous-estime en général très largement la survie des tronçons puisqu'elle ne prend en compte que les tronçons hors service. Il faut donc utiliser la courbe de survie empirique (calculé par exemple par la méthode de Kaplan-Meier ou de Turnbull (cf. Annexe 10) qui prend en compte les tronçons hors service mais aussi les tronçons en service (= tronçons censurés). La démonstration de l'utilité de ces méthodes est réalisée en Annexe 12 et Annexe 13.

Il est important de retenir que la méthode de Turnbull et de Kaplan-Meier élargie donne systématiquement strictement les mêmes résultats. Vu que la méthode de Turnbull est bien plus longue en temps de calcul nous n'utiliserons par la suite que la méthode de Kaplan-Meier.

La courbe de survie estimée par la méthode de Kaplan-Meier est discontinue, c'est-à-dire qu'elle ne pas possède pas de valeurs à tous les âges. En effet cela arrive lorsqu'un âge n'est pas observé dans la fenêtre d'observation. Or pour les calculs d'indicateur à l'étape 4 de notre modèle nous avons besoin de fonction de survie continue. Nous allons donc présenter dans la partie IV.5 des méthodes pour obtenir une fonction continue la plus proche de notre survie empirique discontinue.

IV.5 Obtention d'une fonction continue à partir de la fonction de survie empirique discontinue

IV.5.1 Ajustement d'un modèle paramétrique classique à notre courbe de survie empirique

Plusieurs modèles paramétriques de survie sont classiques en statistique : la survie de Weibull, de Poisson, Lognormale, de Herz, normale etc. Toutes ont des équations mathématiques associées dont les paramètres peuvent varier.

L'ajustement consiste à trouver les paramètres de ces fonctions de survie classiques afin de les faire correspondre au mieux à notre courbe de survie empirique (Kaplan-Meier). La méthode d'ajustement utilisée est la méthode des moindres carrés, elle consiste à minimiser [par la méthode de (Nelder & Mead 1965), cf. Annexe 11] la somme des écarts (élevés au carré pour s'affranchir des signes) entre les courbes empirique et théorique.

IV.5.1.1 Ajustement d'un modèle paramétrique de Weibull

Le modèle de Weibull est très utilisé dans le cadre de l'analyse des données de survie. La fonction de survie de Weibull dépend de deux paramètres, un paramètre d'échelle $\lambda \geq 0$ et un paramètre de forme $\delta \geq 1$, tous deux rassemblés dans le vecteur de paramètres $\boldsymbol{\Theta} = (\boldsymbol{\lambda}, \boldsymbol{\delta})$

$$S_\Theta(t) = e^{-t^\delta \times e^{-\lambda}} \qquad \text{Eq. 18}$$

Les paramètres $\boldsymbol{\Theta}$ optimaux sont ceux qui minimisent la quantité :

$$S(t) = \sum_t \left(\widehat{S_{KM}}(t) - \frac{S_\Theta(t)}{S_\Theta(t\min_{KM})} \right)^2 \qquad \text{Eq. 19}$$

Avec t les réalisations de la variable T âge à la mise hors service, $\widehat{S_{KM}}(t)$ l'estimateur empirique de Kaplan-Meier et $t\min_{KM}$ l'âge minimum à partir duquel Kaplan-Meier a été estimé, c'est-à-dire 0 pour la majorité des matériaux sauf pour la fonte grise. En effet, nos âges de mise hors service sont observés sur une fenêtre restreinte de 1995 à 2012 et le plus jeune tronçon en fonte grise mis hors service est âgé de 27 ans. La courbe de Kaplan-Meier commence alors avec $\widehat{S_{KM}}(27) = 1$, il faut donc le prendre en compte dans l'estimation de la courbe d'ajustement de Weibull. En effet nous ne pouvons estimer que la survie conditionnelle : $P\{T \geq t | T \geq t_{min}\}$ (cf. méthode de déconditionnement en annexe 15).

Nous faisons deux tests de rapport de vraisemblance pour valider les paramètres $\boldsymbol{\Theta} = (\boldsymbol{\lambda}, \boldsymbol{\delta})$ du modèle.

- Premier test :
  Posons l'hypothèse nulle H_0 : $\Theta = \Theta_0$ avec $\Theta_0 = (\boldsymbol{\lambda}_{NM}, \delta_0 = 1)$
  δ_0 est le paramètre de référence, si $\hat{\delta} \leq 1$ alors il n'y a pas de vieillissement.et $\boldsymbol{\lambda}_{NM}$ est la valeur issue de l'optimisation par la méthode de Nelder Mead.

Contre l'hypothèse alternative H_1 : $\Theta \neq \Theta_0$

On définit alors $\widehat{\Theta}$ l'estimateur du maximum de vraisemblance et $\widehat{\Theta_0}$ l'estimateur du maximum de vraisemblance sous H_0. On définit enfin la statistique du test du rapport de vraisemblance :

$$T = 2 \times \left| ln\mathrm{L}\left(\widehat{\Theta}\right) - ln\mathrm{L}\left(\widehat{\Theta_0}\right) \right| \qquad \text{Eq. 20}$$

On sait que sous l'hypothèse nulle, la statistique du test du rapport de vraisemblance suit une loi du χ^2 avec un nombre de degrés de liberté égal au nombre de contraintes imposées par l'hypothèse nulle. Donc sous H_0, T suit une loi du χ^2 à 1 degré de liberté ($T \sim X^2(1)$).

Nous choisissons un seuil de 5% (noté α) d'acceptation de l'erreur. C'est la probabilité que l'on a de se tromper d'hypothèse.

Par conséquent, on rejette H_0 (à α près) lorsque la statistique du test du rapport de vraisemblance est supérieure au quantile d'ordre 1- α de la loi du χ^2 à 1 degré de liberté.

Si la P-value est < 5% alors nous rejetons H_0 et acceptons H_1 et si la P-value ≥ 5% nous acceptons H_0.

- Deuxième test :

Posons l'hypothèse nulle H_0 : $\Theta = \Theta_0$ avec $\Theta_0 = (\boldsymbol{\lambda_0} = 0, \delta_{NM})$. Ensuite le test se déroule de la même façon que précédemment.

IV.5.1.2 Ajustement d'un modèle paramétrique de Herz

Le modèle de Herz est un modèle qui a été développé spécialement pour la survie de tronçons (Herz 1996). La fonction de survie de Herz a pour paramètres $\boldsymbol{\Theta} = (\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{\tau})$ et est défini par :

$$S_\Theta(t) = \begin{cases} 1 & si\ t \le \tau \\ \dfrac{(\eta + 1)}{\eta + e^{\gamma(t-\tau)}} & si\ t > \tau \end{cases} \qquad \text{Eq. 21}$$

Le paramètre τ représente un paramètre de temps de la résistance au-dessous du quelle aucun renouvellement de tuyau n'est supposé se produire. Cependant, des tronçons renouvelés dès l'année où ceux-ci sont posés sont observés dans nos données, le paramètre est alors fixé à 0.

Les paramètres $\boldsymbol{\Theta}$ optimaux sont ceux qui minimisent la quantité :

$$S(t) = \sum_t \left(\widehat{S_{KM}}(t) - \frac{S_\Theta(t)}{S_\Theta(t\min_{KM})} \right)^2 \qquad \text{Eq. 22}$$

Pour le calcul de la probabilité d'erreur sur chaque paramètre voir le paragraphe précédent.

IV.5.2 Estimation des points inexistants par interpolation linéaire

Une autre méthode pour obtenir une fonction continue à partir de notre fonction de survie discontinue empirique est d'estimer uniquement la valeur les points inexistants de proche en proche. Cette méthode est bien plus simple que la précédente. Elle consiste à tracer une droite entre deux points existants ce qui permet d'estimer les points inexistants au milieu.

Les intérêts d'ajuster une équation mathématique aux courbes de survie sont d'une part d'avoir une équation en tous points pour calculer des indicateurs clefs de l'étape 4 de notre modèle. Et d'autre part, de permettre de déconditionner (cf. Annexe 14) la survie empirique notamment pour la fonte grise.

Les méthodes suivantes permettent aussi d'avoir une fonction de survie continue mais elles nous seront aussi très utiles pour passer de l'étape 2 à l'étapes 3 de notre modèle long terme.

IV.6 Modèle de Cox

Dans cette partie, l'âge à la mise hors service est fonction de variables explicatives telles que la longueur ou le diamètre des tronçons.

Nous utilisons le modèle des risques proportionnels proposé par D. R. Cox (Cox 1972). Il permet d'établir une relation paramétrique entre les facteurs de « risque » de survenue de l'événement et la distribution des durées de survie sans postuler pour celle-ci une forme paramétrique particulière. L'analyse qui en découle présente donc un caractère semi-paramétrique.

Le modèle de Cox exprime une relation entre la fonction de risque associée à la survenue d'un événement et le vecteur des *p* variables explicatives $\boldsymbol{z} = (\boldsymbol{z_1}, \boldsymbol{z_2}, \dots, \boldsymbol{z_p})$

$$\lambda(t, z) = \lambda_0(t) r(\beta, z) \qquad \text{Eq. 23}$$

où $\boldsymbol{\beta} = (\boldsymbol{\beta_1}, \dots, \boldsymbol{\beta_p})$ est le vecteur des coefficients de régression et $\lambda_0(t)$ est la fonction de risque de base.

$$\lambda(t) = \frac{f(t)}{S(t)} \qquad \text{Eq. 24}$$

Plus précisément $\lambda_0(t)$ est la fonction de risque des sujets pour lesquels toutes les variables explicatives z_j $(j = 1, \dots, p)$ sont nulles. La fonction $r(\beta, z)$ dépend des caractéristiques z du sujet, la dépendance étant mesurée par les coefficients β. En général, on prend $r(\beta, z) = \exp(\beta_1 z_1 + \cdots + \beta_p z_p)$ de façon à obtenir une fonction de risque positive sans contrainte sur les coefficients β quelles que soient les valeurs de z.

Nous avons alors :

$$\lambda(t, z) = \lambda_0(t) \times \exp(\beta_1 z_1 + \cdots + \beta_p z_p) \qquad \text{Eq. 25}$$

La fonction de risque du modèle de Cox est alors le produit d'une fonction (non-paramétrique) $\lambda_0(t)$ qui ne dépend que du temps et d'une fonction (paramétrique) $\exp(\beta_1 z_1 + \cdots + \beta_p z_p)$ qui n'en dépend pas.

Soient $t_1 < t_2 < \cdots < t_k$ les différents temps d'évènements observés et (1)… (k) les indices des sujets ayant subi l'événement respectivement en $t_1 \dots t_k$. La probabilité conditionnelle que le sujet *i* subisse l'événement en t_i sachant qu'il est à risque au temps t_i et qu'il n'y ait qu'un seul évènement en t_i parmi les sujets à risque au temps t_i est égale à :

$$p_i = \frac{\lambda_0(t_i) \times \exp(\beta z'_{(i)})}{\sum_{l:y_l \geq t_i} \lambda_0(t_i) \times \exp(\beta z'_{(l)})} \qquad \text{Eq. 26}$$

Une solution proposée par Cox est l'élimination de $\lambda_0(t)$ en scindant la vraisemblance en deux parties et en ne conservant que la partie de la vraisemblance qui concerne uniquement les coefficients $\beta_j (j = 1, \dots p)$ que l'on cherche à estimer. La probabilité conditionnelle devient :

$$p_i = \frac{\exp(\beta z'_{(i)})}{\sum_{l:y_l \geq t_i} \exp(\beta z'_{(l)})} \qquad \text{Eq. 27}$$

Cette quantité ne dépend pas de la fonction de risque de base $\lambda_0(t)$ qui est considérée ici comme un paramètre de nuisance. La vraisemblance partielle est le produit des probabilités conditionnelles calculées à chaque temps d'évènements.

$$V(\beta) = \prod_{i=1}^{k} \frac{\exp(\beta z'_{(i)})}{\sum_{l:y_l \geq t_i} \exp(\beta z'_{(l)})} \qquad \text{Eq. 28}$$

Cette vraisemblance partielle n'est pas une vraisemblance au sens statistique, mais il a été établi qu'elle peut être utilisée comme telle pour estimer les coefficients de régression β et tester l'influence de variables explicatives sur la fonction de risque (Cox 1972)(Alioum & Leffondre 2014).

Pour que les β puissent être pris en compte il faut tester l'hypothèse de proportionnalité des risques. Une méthode possible est celle des résidus de Schoenfeld (D. Schoenfeld 1982). Pour chaque covariable, nous testons si son effet est indépendant du temps. Nous pouvons vérifier que l'hypothèse des risques proportionnels est satisfaite par lecture graphique ou par analyse de la P-value.

Les résidus de Schoenfeld se calculent de la manière suivante : pour chaque date de mis hors service t_i, nous calculons la différence entre les caractéristiques de l'individu décédé et une moyenne pondérée des caractéristiques des individus à risque d'être mis hors service au temps t_i. Ceci donne :

$$R_{ij} = z_{ij} - \bar{z}_{ij}(t_i) \qquad \text{Eq. 29}$$

Avec :

- R_{ij} résidu au temps t_i
- z_{ij} valeur de la covariable *j* pour l'individu décédé au temps t_i
- $\bar{z}_{ij}(t_i)$ moyenne pondérée de la covariable *j* chez les individus à risque au temps t_i

Si l'hypothèse des risques proportionnels est vérifiée, alors les résidus doivent être distribués de la même manière au cours du temps.

IV.7 Modèle de Weibull avec covariables au tronçon

Le modèle de Cox ne permettant pas de tracer une courbe de survie, dans cette partie nous avons choisi une équation le permettant. Il est possible de calculer une fonction de survie de Weibull en fonction de différentes covariables. Ce modèle de Weibull a les propriétés du modèle de Cox (Cf. Annexe 15). Ici les covariables sont « NDOA » le nombre de défaillances observé par an, « ø » le diamètre du tronçon, « *Lg* » la longueur du tronçon. Ainsi que cinq covariable-indicatrice pour chacun des matériaux : fonte grise (FG), fonte ductile (FD), PEHD, Bonna, et autre, prenant pour valeur 1 si le tronçon est de ce matériau, 0 sinon. La fonction de survie s'écrit alors comme suit :

$$S_\Theta(t) = e^{-t^{\delta_1} \times e^{Z^T \beta}} \quad \text{Eq. 30}$$

Avec $Z^T\beta = \beta_0 + \beta_1 ø + \beta_2 Lg + \beta_3 FG + \beta_4 FD + \beta_5 PEHD + \beta_6 Bonna + \beta_7 Autre + \beta_8 NDOA$

Avec $\boldsymbol{\Theta}$ le vecteur de paramètre $(\boldsymbol{\delta_1}, \boldsymbol{\delta_2}, \boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \boldsymbol{\beta_4}, \boldsymbol{\beta_5}, \boldsymbol{\beta_6}, \boldsymbol{\beta_7}, \boldsymbol{\beta_8})$ à estimer. Notons $\widehat{\Theta}$ l'estimateur de Θ, il se calcule en maximisant la vraisemblance.

Pour la construction d'une fonction de vraisemblance il est important de prendre en compte la censure et la troncature de l'échantillon observé. Si la date de mise hors service est observée ($c_i = 0$) alors la probabilité que l'événement se produise dans un intervalle arbitrairement petit contenant cet âge est connue. Cette probabilité est approximativement égale à la fonction de densité de l'individu pour ce moment.

Fonction de densité : $f_\Theta(b_i|a_i)$

Si la date de mise hors service n'est pas observée, l'individu est censuré ($c_i = 1$). Alors, la seule information dont nous disposons est : la probabilité que l'âge à la mise hors service d'un tronçon soit supérieur à un certain âge qui se lit donc par sa fonction de survie :

$$S_\Theta(b_i|a_i)$$

Enfin comme l'échantillon est tronqué à gauche il est nécessaire de prendre en compte sa probabilité conditionnelle: (Klein & Moeschberger 2003)

$$\frac{f_\Theta(b_i|a_i)}{S_\Theta(a_i)}$$

La fonction de vraisemblance s'écrit alors en composant ces trois éléments :

$$L(\Theta) = \prod_{i=1}^{n} \frac{f_\Theta(b_i|a_i)^{1-c_i} \times S_\Theta(b_i|a_i)^{c_i}}{S_\Theta(a_i)} \quad \text{Eq. 31}$$

Par la méthode de Nelder et Mead nous minimisons la fonction objectif correspondant à –ln(L(Θ)), ce qui revient finalement à maximiser la vraisemblance.

$$-\ln\big(L(\Theta)\big) = -[\sum_{i=1}^{n}(1-c_i)\big[\ln\big(f_\Theta(b_i|a_i)\big) - \ln\big(S_\Theta(a_i)\big)\big] + \sum_{i=1}^{n} c_i\big[\ln\big(S_\Theta(b_i|a_i)\big) - \ln\big(S_\Theta(a_i)\big)\big]] \quad \text{Eq. 32}$$

Avec :

$$\ln\big(S_\Theta(t)\big) = -\,t^{\delta_1} \times e^{Z^T\beta} \quad \text{Eq. 33}$$

Et comme :

$$F_\Theta(t) = 1 - S_\Theta(t) \quad \text{Eq. 34}$$

$$f_\Theta(t) = \frac{-dS_\Theta(t)}{dt} \quad \text{Eq. 35}$$

$$= -\left[-\delta_1 t^{\delta_1 - 1} \times e^{Z^{\mathrm{T}}\beta} \times e^{-t^{\delta_1} \times e^{Z^{\mathrm{T}}\beta}}\right] \quad \text{Eq. 36}$$

$$= \delta_1 t^{\delta_1 - 1} \times e^{Z^{\mathrm{T}}\beta} \times e^{-t^{\delta_1} \times e^{Z^{\mathrm{T}}\beta}} \quad \text{Eq. 37}$$

Finalement,

$$\ln\big(f(t)\big) = (\delta_1 - 1) + \ln(\delta_1 t) + Z^{\mathrm{T}}\beta - t^{\delta_1}[e^{Z^{\mathrm{T}}\beta}] \quad \text{Eq. 38}$$

Nous effectuons plusieurs tests de rapport de vraisemblance (cf. méthode décrite dans le paragraphe IV.5.1.1) pour valider chacun des paramètres du modèle.

Après avoir présenté des méthodes pour calculer notamment la survie empirique, nous allons les appliquer aux données du SEDIF. En effet la fonction de survie empirique passée fait partie de la première étape nécessaire à la création de notre nouveau modèle « long terme ».

V Calcul de la survie passée, application aux données du SEDIF

V.1 Courbes de survie sans partitionnement

Pour rappel dans les paragraphes suivants :

- Notre événement d'intérêt est la mise hors service d'un tronçon ou d'un mètre de tronçon.
- Notre fenêtre d'observation est [1995,2012]

V.1.1 Comparaison des courbes de survie en linéaire et en effectif

La Figure 36 représente les courbes de survie de Kaplan-Meier. La courbe « S(t) Kaplan-Meier (linéaire) » est réalisée en linéaire de tronçons (nous comptons le nombre de mètres de tronçons qui meurt [=est mis hors service] à chaque âge), alors que la courbe « S(t) Kaplan-Meier (effectif) » est calculée en comptant le nombre de tronçons qui deviennent hors service à chaque âge. Il est alors possible de voir que ces deux courbes se superposent quasiment à l'identique jusqu'à environ 90 ans puis se séparent légèrement après. Cette écart peut être expliqué, du fait qu'au-delà de 90 ans, les petits tronçons survivent mieux que les longs (cf. Figure 41), cela provient sans doute d'un artefact SIG dans le découpage en tronçons. Comme les tailles des tronçons sont très disparates, pour les gestionnaires des services d'eau, il est plus facile de parler en pourcentage du linéaire de réseau qui est en service après un certain âge.

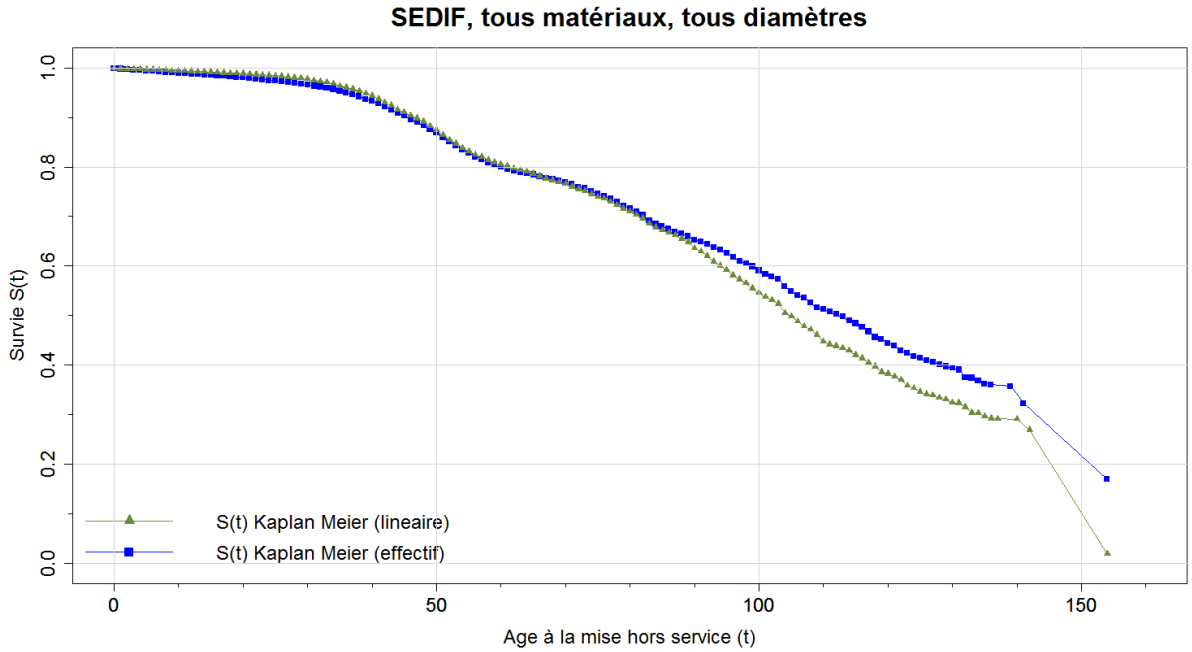


Figure 36 : Courbes de survie passés en linéaire et en effectif, SEDIF, tous tronçons confondus entre [1995, 2012]

La Figure 37 représente la survie de Kaplan-Meier en effectif ainsi que son intervalle de confiance calculé grâce à la formule de Greenwood et de Rothman (Cf. IV.4.1). Nous constatons que jusqu'à 150 ans l'intervalle de confiance est plutôt très bon.

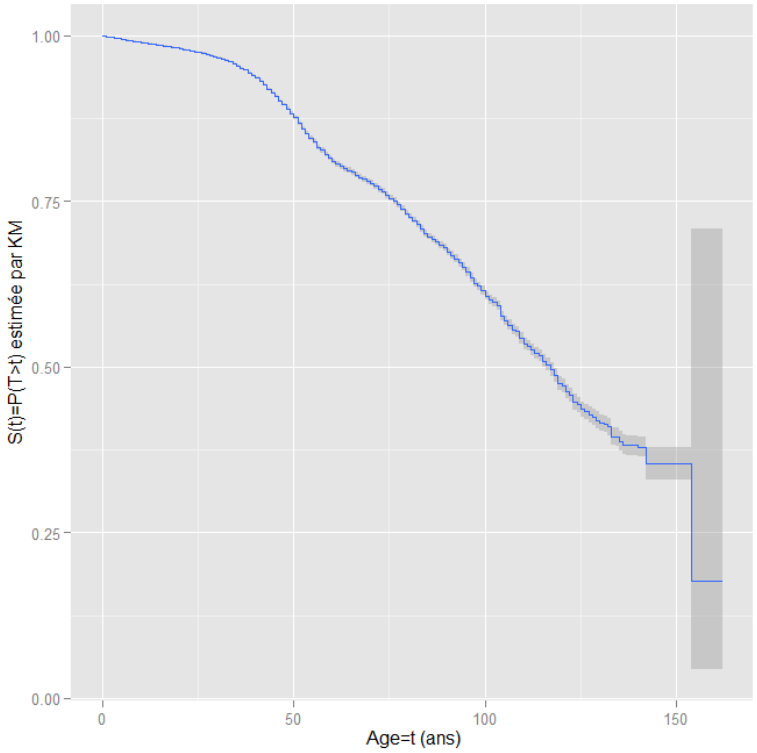


Figure 37 : Courbe de survie en nombre de tronçons accompagnée de son intervalle de confiance, tous tronçons confondus.

V.1.2 Courbe de survie tous matériaux, tous diamètres en linéaire

Sur la Figure 38 sont représentées trois courbes :

- « Sb(t) Brute » représente la survie des tronçons (en linéaire) hors service. Cette courbe n'est pas représentative de la durée de maintien en service des tronçons car elle ne tient pas compte des tronçons encore en service après 2012. C'est pourtant celle qui est utilisée à l'heure actuelle par certains bureaux d'étude. Nous ne nous appuierons donc pour la suite que sur les courbes de survie issues des méthodes de Kaplan-Meier.

- « Se(t) Kaplan Meier» décrit la survie empirique des tronçons en service et hors service estimée par la méthode de Kaplan Meier

- « S(t) Weibull » est la fonction de Weibull ajustée sur la survie de Kaplan Meier par la méthode des moindres carrées.

Toutes ces courbes sont estimées sur tous les tronçons, sans être différenciées par le diamètre ou par le matériau ou tout autre critère. La médiane de survie (courbe de Kaplan-Meier) d'un linéaire de canalisation est de 105 ans. La survie ici S(t) est la probabilité pour plusieurs mètres de tronçon de ne pas avoir été mis hors service au-delà d'un âge t.

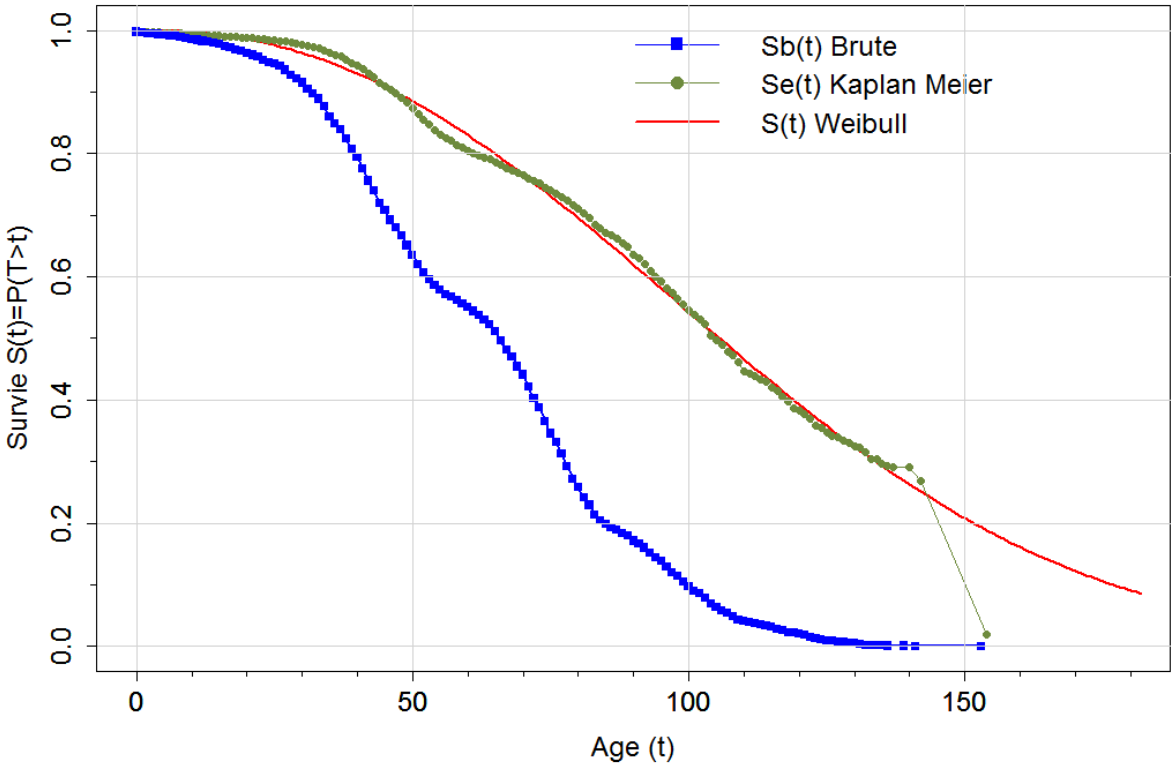


Figure 38 : Courbes de survie en linéaire, tous tronçons confondus

La fonction de survie de Weibull ajusté celle de Kaplan-Meier par la méthode de Nelder and Mead (avec pour fonction objectif : les moindres carrés) est très proche de l'optimum (Cf. ci-dessous).

Paramètres (Estimate), valeur initial pour l'optimisation de Nelder & Mead (Ref), statistique de test (Chi2), degré de liberté (df) et P-value (Pval) associé à la survie de Weibull:

Label	Estimate	Ref	Chi2	df	Pval
delta	2.3267016e+00	1	3.7834656e+01	1	0.0000000
lambda	1.1206531e+01	10	2.3714826e+01	1	0.0000011

La courbe brute est très biaisée. Elle ne prend pas en compte les tronçons encore en service. En aucun cas il ne faut utiliser cette courbe pour effectuer des calculs prospectifs.

V.2 Courbes de survie avec partitionnement

Définition du « partitionnement » :
Soit Ω un ensemble de tronçons. Ω est partitionné si :

- aucune de ces partitions n'est **vide**;
- l'**union** de ces partitions est égale à Ω;
- ses partitions sont deux à deux **disjointes**.

NB : il est à noter que le partitionnement est souvent aussi appelé « stratification ». Donc chaque partition est souvent nommée « strate ».

V.2.1 Comparaison des courbes de survie pour différents diamètres

Dans cette section nous avons partitionné l'ensemble des tronçons en 2 sous-ensembles selon le diamètre (réseau de distribution [inférieurs ou égale à 300 mm] ou réseau de transport [supérieurs à 300 mm]) (Cf. Figure 12).

L'étude des courbes de survie en différenciant les tronçons par leur diamètre indique un premier résultat très intéressant. En effet, il est possible de constater que les tronçons du réseau de transport sont mis hors service plus tard que les tronçons du réseau distribution sur une période d'environ 0 à 70 ans. Puis la survie des tronçons de transport plonge brusquement, alors que celle des tronçons de distribution diminue plus doucement. Le diamètre est donc une variable qui a un intérêt certain pour la construction d'un modèle de prévision à long terme. La chute brutale des tronçons représentant le réseau de transport peut être expliquée par une décision politique du SEDIF qui est de renouveler tous les tronçons de gros diamètres à partir un certain âge. En effet le SEDIF ne peut pas se permettre d'observer des casses sur ces tronçons de gros diamètres, qui pourraient induire des risques de dommages trop importants. (Cf. Figure 40 pour voir ces courbes en linéaire)

Les intervalles de confiance de ces deux courbes ne se chevauchent que très peu, nous pouvons donc dire qu'elles sont significativement différentes.

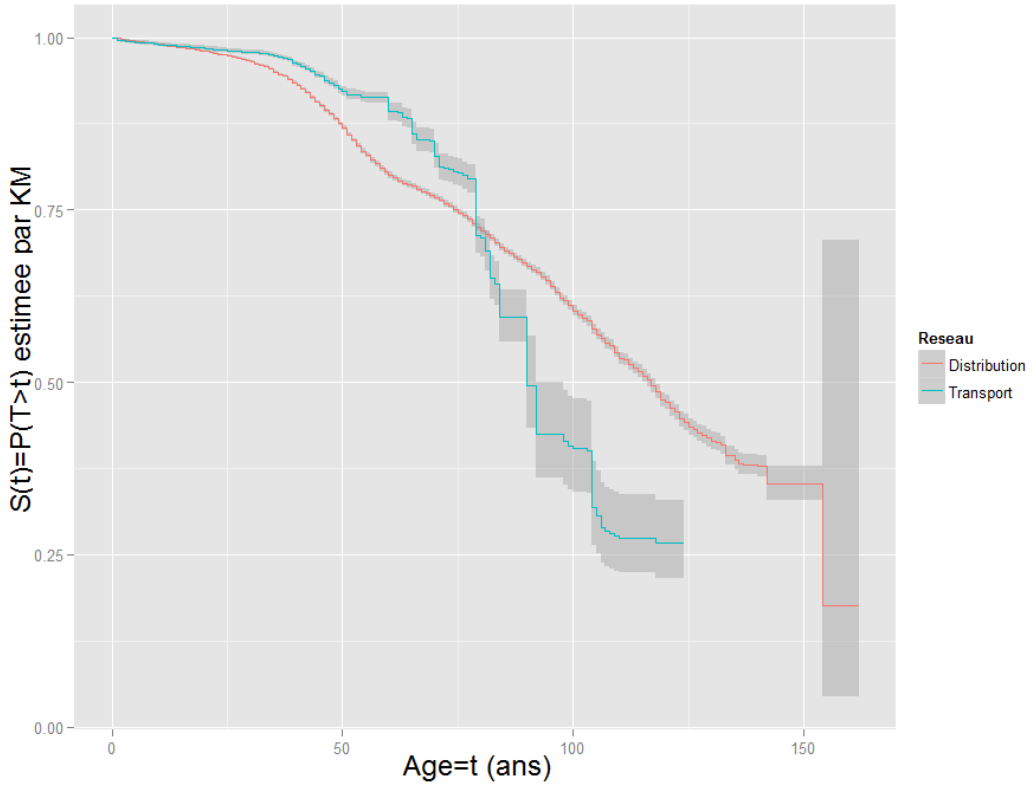


Figure 39 : Courbes de survies passées [1995-2012] en effectif (tronçons partitionnés en 2 sous-ensembles réseau de transport et réseau de distribution) et leurs intervalles de confiances (en gris foncé)

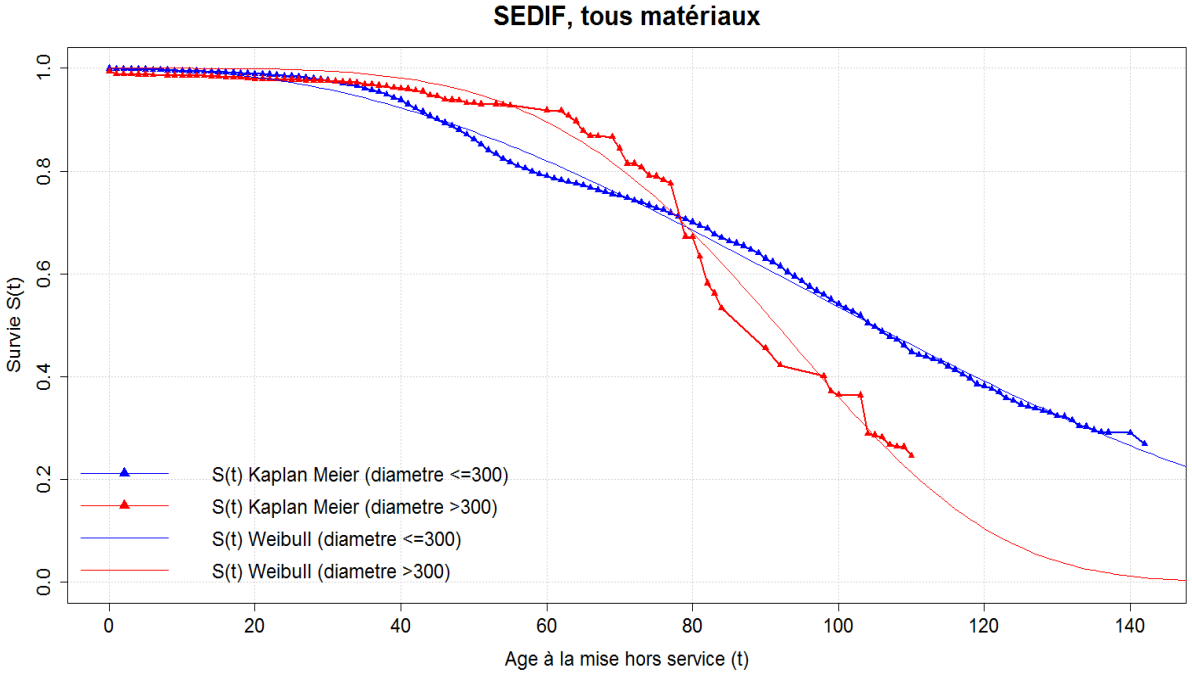


Figure 40 : Courbes de survie passés [1995, 2012] en linéaire (tronçons partitionnés selon le diamètre) et leurs survies de Weibull ajustées

V.2.2 Comparaison courbes de survie pour différentes longueurs

Nous avons partitionné, ici, l'ensemble des tronçons en 3 sous-ensembles selon leur longueur. Nous avons utilisé les seuils employés par VEDIF dans son modèle court terme (100m et 150 m) (cf. figure 45).

L'étude des courbes de survie en différenciant par la longueur, présente un deuxième résultat intéressant. En effet, il est possible de constater que les tronçons de grande (supérieure à 150m), moyenne ou petite longueur (inférieure à 100m) sont mis hors service à peu de la même façon jusqu'à environ 50 ans. Puis, après 50 ans les petits et moyens tronçons sont mis hors service à des âges plus avancés que les tronçons de longues tailles. La survie des tronçons de moyenne et petite tailles reste la même jusqu'à environ 80 ans puis la survie des moyens tronçons décroit plus vite que celle des tronçons de petites tailles. Cela peut être expliqué par le fait qu'en général, les longs tronçons sont aussi ceux qui constituent le réseau de transport (Cf. Annexe 16 corrélation longueur diamètre). La longueur est donc une variable qui peut avoir un intérêt pour la construction d'un modèle de prévision à long terme. (Cf. Figure 43 pour voir ces courbes en linéaire)

Les intervalles de confiance de ces trois courbes se chevauchent de 0 à environ 80 ans, nous ne pouvons donc pas dire qu'elles soient significativement différentes cependant, après cette dates elles sont significativement différentes.

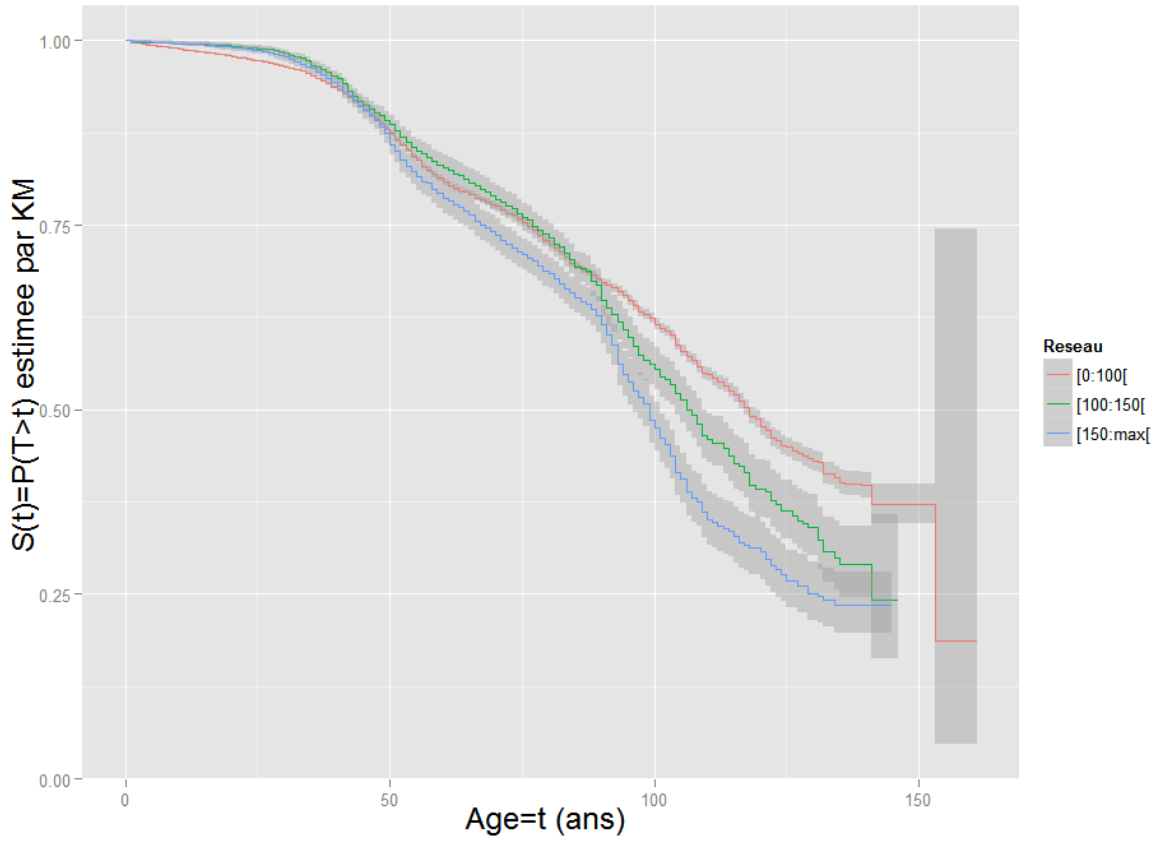


Figure 41 : Courbes de survie en effectif tronçons (tronçons partitionnés en 3 sous-ensembles : petites, moyennes et grandes longueurs) et leurs intervalles de confiance (en gris foncé)

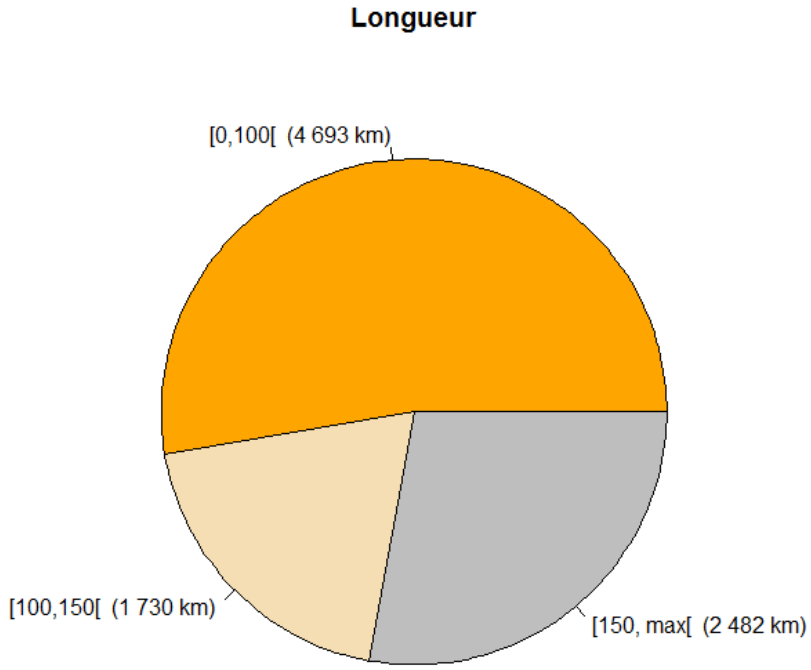


Figure 42 : Répartition du linéaire en service (fin 2012) du SEDIF en 3 sous-ensembles selon la longueur des tronçons.

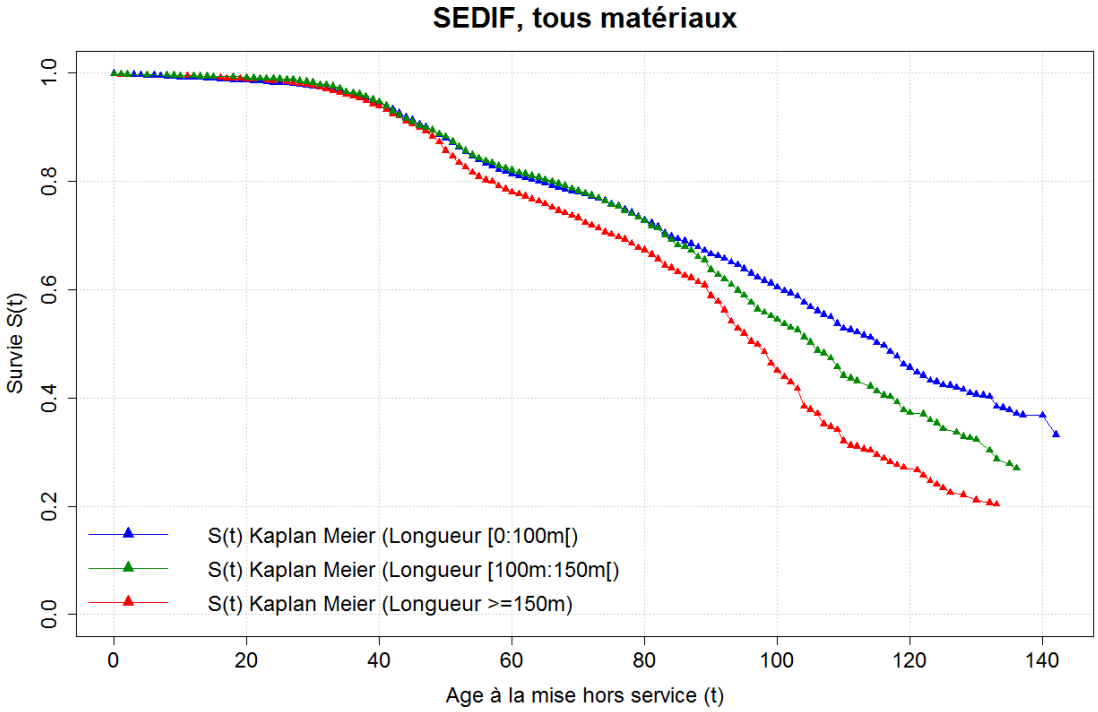


Figure 43 : Courbes de survie en linéaire (tronçons partitionnés en 3 sous-ensembles : petites, moyennes et grandes longueurs)

V.2.3 Comparaison courbes de survie pour différents matériaux (en linéaire)

Dans ce paragraphe nous avons partitionné l'ensemble des tronçons en 6 sous-ensembles selon leur matériau (Cf. Figure 10).

Sur la Figure 44 sont représentées les courbes de survie selon les différents matériaux. Notons qu'elles sont très différentes d'un matériau à l'autre. Il est impossible d'estimer la survie après 40 ans des tronçons en fonte ductile et après 30 ans pour le PEHD. Effectivement, ce sont tous deux des matériaux jeunes et nous n'avons pas suffisamment d'information dans le temps pour calculer leur survie après ces dates.

Pour les tronçons en Bonna, nous retrouvons plus ou moins la même courbe que pour les tronçons de gros diamètre. Ce qui est cohérent car les canalisations de gros diamètres sont en général en Bonna.

Enfin, pour les tronçons en fonte grise, nous ne connaissons pas le début de la courbe de survie entre 0 et 26 ans. Ceci s'explique par le fait que dans notre fenêtre d'observation, le plus jeune tronçon de fonte grise observé en 1995 a 27 ans. En effet, les derniers tronçons en fonte grise ont été posés aux alentours de 1970 tandis que la fenêtre d'observation des âges à la mise hors service commence en 1995. Pour obtenir la survie de Kaplan-Meier ci-dessous pour ce matériau nous avons déconditionné la survie d'origine par S(26) de Weibull (Cf. Annexe 14). De même, la survie de Kaplan Meier pour le type de matériau « Autre » a été déconditionnée.

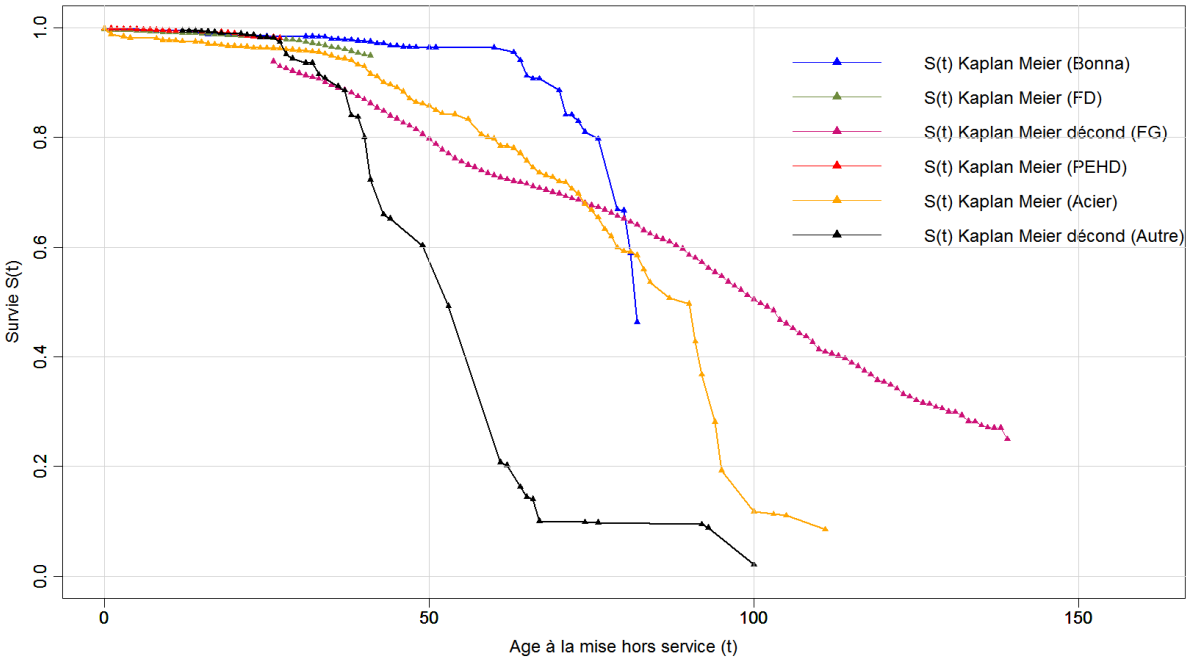


Figure 44 : Courbes de survie en linéaire pour différents matériaux

Nous constatons que les « autres matériaux » ont été mis hors service bien plus tôt en âge que les autres matériaux. D'après (Optimeau 2014) la survie des matériaux « autres » au SEDIF est très faible car le SEDIF a testé des matériaux composites, qui malheureusement en pratique se sont révélés de très mauvaise qualité. Donc ils ont dû être renouvelés rapidement.

V.2.4 Comparaison courbes de survie pour des tronçons avec ou sans défaillances

Nous avons partitionné, ici, l'ensemble des tronçons en 2 sous-ensembles selon leur nombre de défaillances. Un ensemble de tronçons sans défaillance et un autre ensemble de tronçons avec une ou plusieurs défaillances.

L'étude des courbes de survie (cf. Figure 45) en différenciant par l'existence de défaillances ou non, présente un autre résultat intéressant. En effet, il est possible de constater que les tronçons avec ou sans défaillance suivent sensiblement la même probabilité à la mise hors service entre 0 et 40 ans (leurs intervalles de confiance ce chevauchent). Puis les tronçons sans défaillance sont mis hors service à des âges plus avancés que les tronçons avec une ou plusieurs défaillances. Cette différence est significative jusqu'à environ 140 ans (les intervalles de confiance ne se chevauchent pas). Enfin cette différence n'est plus significative au-delà de 140 ; c'est un effet de l'effectif trop faible.

Ce résultat coïncide avec ce que nous attendions (cf. Figure 46). En effet il est logique que des tronçons ayant un certain nombre de défaillances soient renouvelés plus tôt que des tronçons sans défaillance.

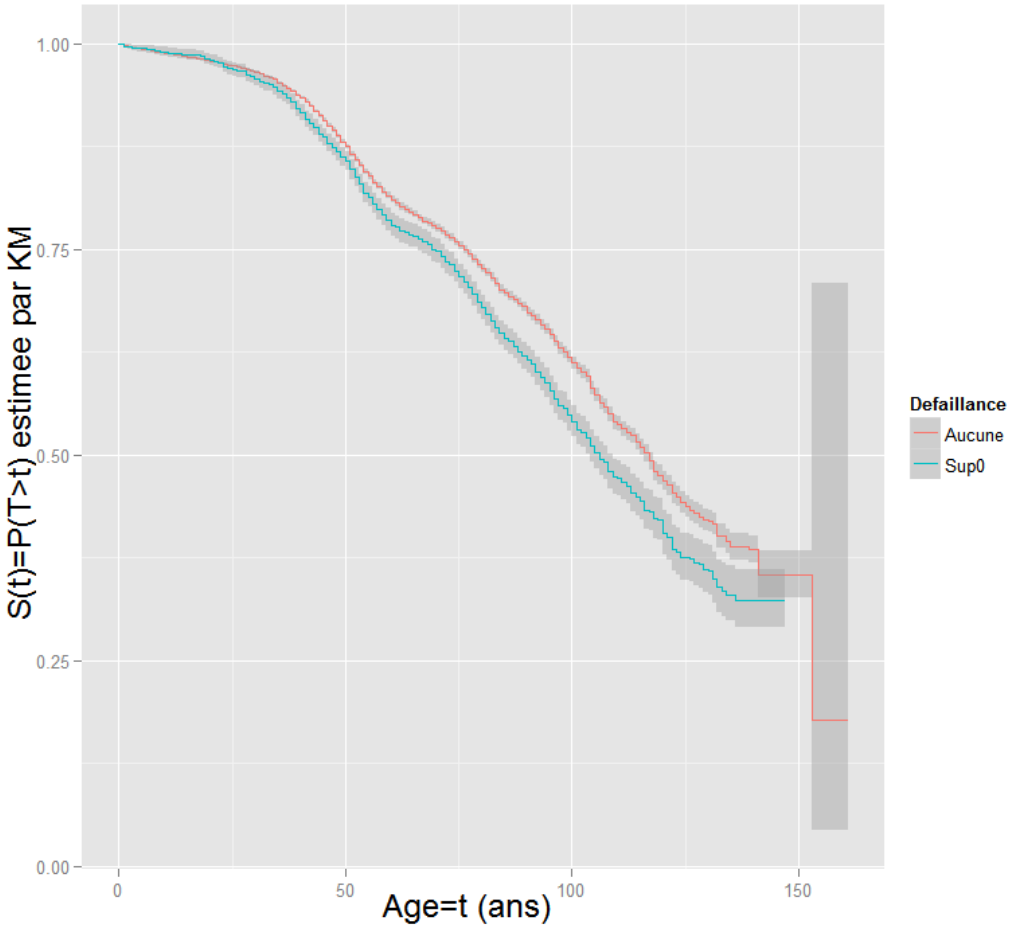


Figure 45: Courbes de survie en effectif tronçons (tronçons partitionnés en 2 sous-ensembles : avec et sans défaillance) et leurs intervalles de confiance (en gris foncé)

Ce résultat coïncide avec ce que nous attendions et les taux de défaillances passées (cf. Figure 46). En effet il est logique que des tronçons ayant un certain nombre de défaillances soient renouvelés plus tôt que des tronçons sans défaillance.

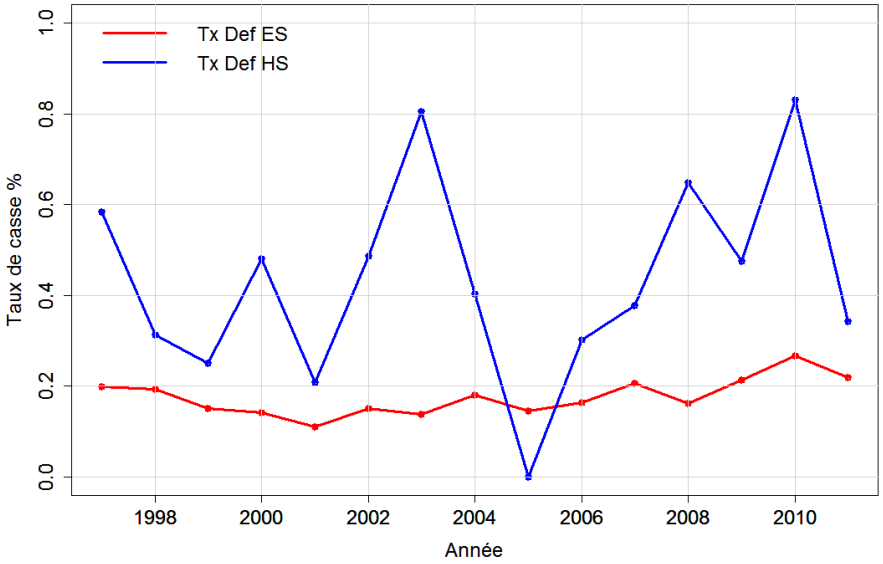


Figure 46: Taux de défaillance annuel sur tronçons en service (en rouge) et taux de défaillance sur tronçons hors service (en bleu)

Il est important de préciser que les points pour chaque année θ de la Figure 46 sont obtenus:

1. En recréant une photographie des tronçons en service à chaque date θ
2. Puis en partitionnant le réseau de chacune de ces photographies en deux strates : « tronçon en service l'année θ +1 » et « tronçons mis hors service l'année θ +1 »
3. Ensuite nous calculons pour ces deux strates pour chacune de ces photographies les taux de défaillances pour l'année θ.

Le graphique de la Figure 46 présente un taux de défaillance nul pour les tronçons en service en 2005 et mis hors service en 2006. Cette valeur nulle semble liée à un artefact dans les données fournies. D'après (Optimeau 2014) cela peut, peut-être, être expliqué par le fait que le SIG du SEDIF a été mis en place en 2005.

Nous avons effectué des courbes de survie à partir d'autres stratifications :

- Niveau piézométrique = niveau du trop-plein du réservoir
- Niveau d'enfouissement
- Type de joint
- Technique de pose
- Type d'appui
- Emplacement (chaussée/trottoir)

Elles sont disponibles dans l'Annexe 19 .

V.3 Courbe de survie avec trois fenêtres d'observation différentes

Dans cette partie nous avons choisi de tester 3 fenêtres d'observation différentes. Nous créons une fenêtre où nous observons les tronçons entre 1995 et 2003, et une autre fenêtre où on les observe entre 2004 et 2012. Pour la construction de la première fenêtre, les données ont légèrement été modifiées en associant notamment « En service » à tous les tronçons mis hors service après 2003. Toutes les dates à la mise hors service après 2003 ont été substituées par « NA ». Ainsi nous recréons une photo de notre jeu de données tel qu'il était en 2003, nous ne connaissons donc pas toutes les informations sur les tronçons après cette date.

Sur la Figure 47 sont représentées les deux courbes de survies de Kaplan-Meier associées à chacune des fenêtres d'observation citées ci-avant, ainsi que la courbe de survie pour la fenêtre d'observation [1995, 2012]. Notons que ces courbes sont quasiment identiques, excepté sur l'intervalle de 55 et 110 ans. Il s'agit sûrement d'un artéfact de donnée d'entrée pour S(t) Kaplan Meier, [1995 :2004[. Le changement de fenêtre d'observation a donc peu d'impact sur les résultats.

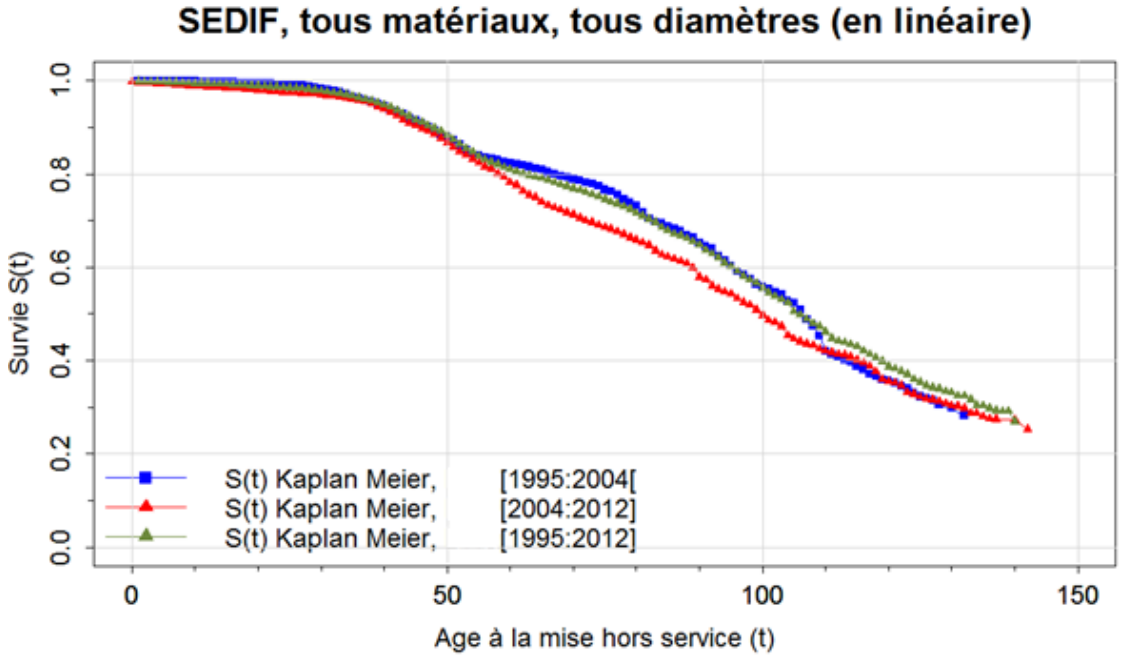


Figure 47 : Courbes de survie au linéaire, tous tronçons confondus, pour trois fenêtres d'observation différentes

Les courbes de survie avec partitionnement nous ont montré que sur la fenêtre d'observation [1995 ,2012] :

- Les tronçons de gros diamètres ont été mis hors service plus tard que les tronçons de petit diamètre avant 80 ans, puis après 80 ans ce phénomène s'inverse ;
- Les tronçons de grande longueur ont été mis hors service plus tôt que les tronçons de plus petite taille ;
- Les courbes de survie par matériaux sont très différentes les unes des autres ;
- Les tronçons sans défaillances sont mis hors service plus tard que ceux avec des défaillances.

Cependant ces variables sont corrélées entre elles. Or il serait intéressant d'observer comment toutes ces variables ont agi conjointement sur les mises hors service. C'est que permet le modèle de Cox et le modèle de Weibull avec covariables.

Nous allons maintenant présenter les résultats du modèle de Cox avec les données du SEDIF

V.4 Résultats du modèle de Cox

Nous avons utilisé le modèle de Cox en choisissant d'étudier l'influence de plusieurs covariables explicatives tels que : le diamètre, la longueur, le matériau et le nombre de défaillance par an (« NDOA ») sur la fenêtre [1995,2012]. Le modèle s'écrit alors :

$$\lambda(t,z) = \lambda_0(t) \times \exp(\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \beta_5 z_5 + \beta_6 z_6 + \beta_7 z_7 + \beta_8 z_8) \qquad \textit{Eq. 39}$$

Avec z le vecteur de 8 covariables ø, Lg, Matériau(FG), Matériau(FD), Matériau(PEHD), Matériau(Bonna), Matériau(Autre) et NDOA.

En maximisant la vraisemblance $V(\beta)$ (cf. IV.8) nous obtenons les résultats suivant :

Covariable	$\hat{\beta}$	$\exp(\hat{\beta})$	P-value
Diamètre	-0.0009	0.999	$<2\times10^{-16}$
Longueur	0. 0005	1.001	2.74×10^{-5}
Matériau(Acier)	0	1	Référence
Matériau(FG)	-0.23	0.79	1.16×10^{-11}
Matériau(FD)	-0.07	0.93	0.05
Matériau(PEHD)	0.25	1.28	8.53×10^{-05}
Matériau(Bonna)	-0.26	0.76	0.0005
Matériau(Autre)	0.50	1.65	5.44×10^{-08}
$NDOA = NDO_i/(b_i - a_i)$	2.26	9.63	$< 2\times10^{-16}$

Les P-value sont toutes inférieures à 5%, donc les coefficients $\hat{\beta}$ sont significatifs. Nous pouvons donc déduire que :

- pour la covariable « Longueur », pour un mètre de longueur en plus, le risque relatif d'être mis hors service est multiplié par 1.001, donc il augmente plus la longueur du tronçon est grande. Ce résultat semble cohérent avec les survies obtenues par la méthode de Kaplan-Meier partitionné par Longueur.
- concernant la variable : « $NDOA$ ». Pour une défaillance par an en plus, le risque relatif d'être mis hors service est multiplié par 9.63, donc il augmente, ce qui est aussi cohérent.
- Les risques relatifs pour les matériaux sont à comparer avec le matériau de référence, ici l'acier. Par exemple, pour le Bonna, le risque relatif d'être mis hors service par rapport à l'acier est multiplié par 0.76 donc il est plus faible. Ce qui est là aussi en accord avec le résultat obtenu grâce aux courbes de survies de Kaplan-Meier en stratifiant par matériaux.

Pour valider le modèle de Cox il nous faut regarder si l'hypothèse de proportionnalité des risques est vérifiée tout au long de la vie du tronçon (âge). Les résultats obtenus par la méthode des résidus de Schoenfeld sont :

Covariable	X^2	P-value
Diamètre	1.09	0.29
Longueur	109.28	0
Matériau(FG)	72.01	0
Matériau(FD)	13.17	2.84×10^{-04}
Matériau(PEHD)	2.05	0.15
Matériau(Bonna)	0.61	0.43
Matériau(Autre)	11.83	5.81×10^{-04}
$NDOA$	8.79	3.01×10^{-03}

Ici pour que l'hypothèse de proportionnalité des risques soit acceptée il faut que la P-value soit supérieure à 5%, hors ce n'est pas le cas ici. Les coefficients $\hat{\beta}$ ne peuvent donc pas être interprétés. Sauf pour les variables « Matériau(Bonna) », « Matériau(PEHD) », et « Diamètre ». Cependant ce type de test n'est valable que sur de petits échantillons, notre échantillon étant de plus de 290 000 données nous n'en tiendrons pas compte. En effet c'est un test « sensible » aux moindres écarts à l'hypothèse H_0 quand la taille de l'échantillon tend vers l'infini.

Le modèle de Cox précédent utilise la fonction de risque instantanée de la variable aléatoire « âge à la mise hors service » comme le produit d'un risque de base inconnue et d'une autre partie connue. Or pour obtenir une fonction de survie nous avons besoin d'une fonction entièrement connue. C'est ce que permet le modèle de Weibull avec covariables au tronçon dont nous allons exposer des résultats à l'aide des données du SEDIF dans la partie suivante.

V.5 Résultats modèle de Weibull avec covariables au tronçon

Le modèle de Weibull avec covariables au tronçon n'est qu'un cas particulier du modèle de Cox (cf. Annexe 15).

En utilisant le modèle de Weibull avec covariables au tronçon que nous écrivons de la façon suivante :

$$S_\Theta(t) = e^{-t^{\delta_1}\times e^{Z^T\beta}} \quad \text{(Cf. IV.7)}$$

(que nous appellerons équation A par la suite) on obtient les résultats suivant :

Variable	Coefficient	Estimation	$\exp(\hat{\beta})$	Initialisation de Nelder & Mead	P-value
Age	$\widehat{\delta_1}$	2.05	Sans objet	1	$< 2\times10^{-16}$
Sans objet	$\widehat{\beta_0}$	-10.43	Sans objet	-4	$< 2\times10^{-16}$
NDOA	$\widehat{\beta_1}$	2.53	12.55	0	$< 2\times10^{-16}$
Diamètre	$\widehat{\beta_2}$	0.0002	1.0002	0	$< 2\times10^{-16}$
Longueur	$\widehat{\beta_3}$	0.0005	1.0005	0	$< 2\times10^{-16}$
Acier	Référence	0	1	Sans objet	Sans objet
PEHD	$\widehat{\beta_4}$	0.95	2.59	0	$< 2\times10^{-16}$
FG	$\widehat{\beta_5}$	0.25	1.28	0	$< 2\times10^{-16}$
FD	$\widehat{\beta_6}$	-0.05	0.95	0	$< 2\times10^{-16}$
Autre	$\widehat{\beta_7}$	1.91	6.75	0	0.0016
Bonna	$\widehat{\beta_8}$	-1.23	0.29	0	$< 2\times10^{-16}$

Toutes les P-value sont inférieures à 5%, nous avons donc un modèle de survie avec des coefficients significatifs. Nous obtenons d'ailleurs des coefficients qui se rapprochent sensiblement des coefficients calculés avec le modèle de Cox, ce qui est rassurant.

Nous calculons la survie de chaque tronçons à chaque âge issue de l'équation A. Puis nous effectuons la moyenne à chaque âge (cf. Eq. 40 & Figure 48) afin de comparer nos résultats à la courbe de survie empirique tous tronçons confondus calculée avec la méthode de Kaplan-Meier. Et qui en plus nous permet d'analyser la mise hors service des tronçons en fonction de covariables, ici l'âge, le diamètre, la longueur, les différents matériaux et le nombre de défaillances observé moyen. Pour un tronçon de diamètre moyen, de longueur moyenne et de nombre de défaillance moyenne on obtient la courbe suivante (Cf. Equation A, Figure 49) :

$$\frac{\sum_{i=1}^{n} S_{\Theta}(t)}{n} = e^{-t^{\delta_1} \times e^{Z^{T}\beta}} \qquad \text{Eq. 40}$$

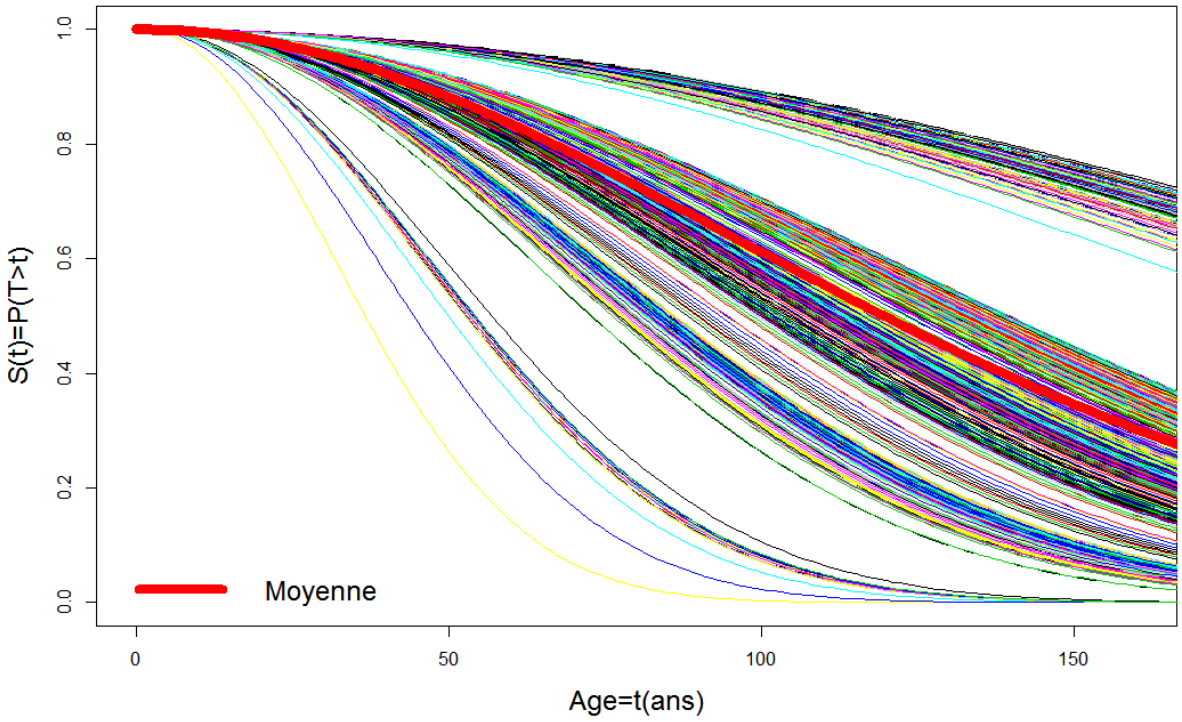


Figure 48 : Equation A pour 4 000 tronçons sélectionnés par tirage aléatoire sans remise et équation A moyenné sur tous les tronçons

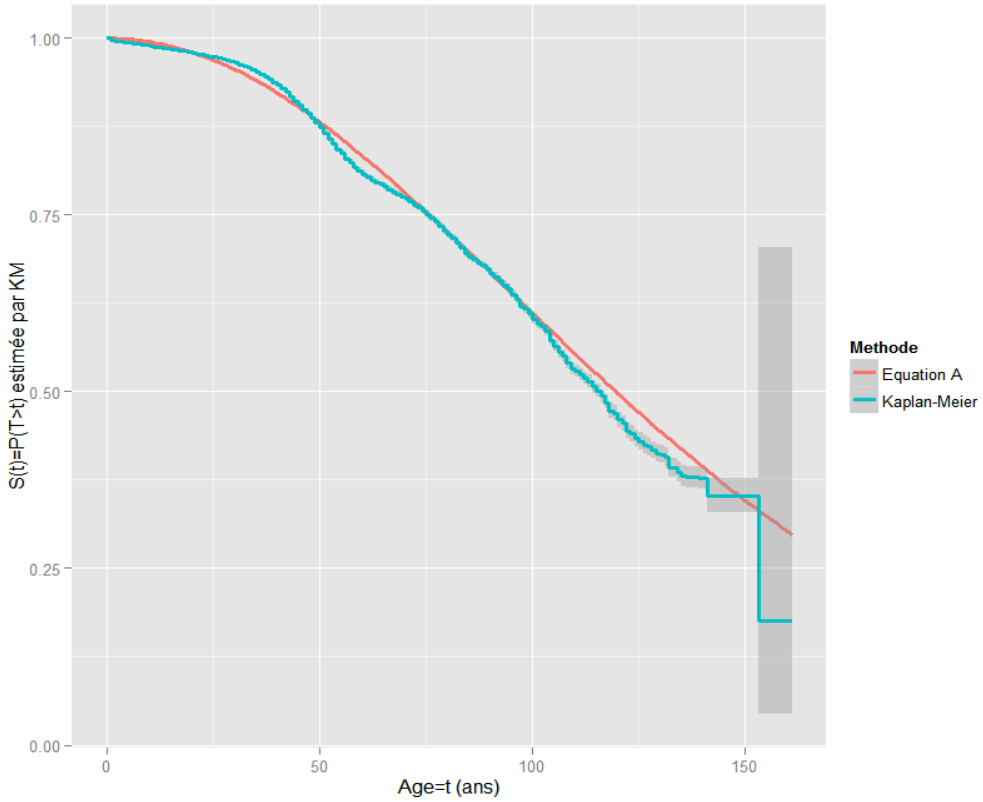


Figure 49 : Courbes de survie de Kaplan-Meier tous tronçons confondus et équation A

L'équation A s'ajuste très bien sur la courbe de Kaplan-Meier, nous obtenons donc un premier résultat très satisfaisant qui explique les mises hors service en fonction de covariables. De plus la survie peut s'écrire sous forme analytique et est donc représentable graphiquement. (Cf. Annexe 17 pour voir une équation à 4 paramètres).

Dans cette cinquième partie nous présentons deux méthodes afin d'obtenir une forme analytique de la fonction de survie des tronçons (fonction de leur âge à la mise hors service). La première méthode consiste à ajuster par la méthode des moindres carrés une fonction analytique (Weibull ou Herz, etc.) à la courbe empirique obtenue par la méthode de Kaplan-Meier. La seconde méthode consiste, à l'échelle du tronçon, à écrire une fonction analytique pour la courbe de survie fonction de diverses covariables, et de déterminer les coefficients de cette fonction.

Maintenant, nous allons choisir un scénario prospectif, **deuxième** étape de notre nouveau modèle long terme (cf. Figure 20).
Puis de ce scénario prospectif nous en déduirons une à plusieurs courbes de survie, **troisième** étape de notre modèle long terme (cf. Figure 20).
Enfin à partir des fonctions de survie prospectives et des données du SEDIF, nous allons calculer 4 indicateurs clés prévisionnels (linéaire à renouveler, taux de renouvellement, âge moyen du réseau et coûts d'investissement), **quatrième** étape du modèle long terme optimeau (cf. Figure 20).

VI Scénario statu quo, calcul d'indicateurs clés au SEDIF

Les courbes de survie empirique passée (fonction de l'âge à la mise hors service) (issue de la méthode Kaplan-Meier) sont représentatives :

- non seulement de l'évolution physique passée du patrimoine (augmentation du mauvais état i.e. des défaillances).
- mais aussi des décisions passées (coordination avec la voirie ou non, aversion plus ou moins grande aux conséquences des défaillances…).

Si le SEDIF applique à l'avenir le **scénario « *statu quo* »** (*i.e.* rien ne change ni dans les objectifs, ni dans les mécanismes impliqués, ni dans les pratiques de gestion), l**es courbes de survie prospectives** sont, dans ce cas très particulier, égales aux courbes de survie passés.
Il est possible d'utiliser ces courbes de survies présentées précédemment pour calculer des indicateurs clés futurs.

VI.1 Avec une seule courbe de survie

VI.1.1 Courbe de survie prospective

La courbe de survie prospective (cf. Figure 50) du **scénario « statu quo : 1 courbe »** est obtenue à partir de la courbe de survie empirique du paragraphe V.1.2 et de la méthode d'ajustement de Weibull décrit dans le paragraphe IV.5.1.1. L'évènement d'intérêt de cette fonction de survie est l'âge à la mise hors service.

La fonction de Weibull, *Sw*(*t*), ajustée sur cette fonction de survie empirique par la méthode des moindres carrés (Cf.IV.5), pour tous les tronçons (Cf. V.1.2) est la suivante

$$Sw(t) = e^{-t^{\delta} \times e^{-\lambda}} = e^{-t^{2.33} \times e^{-11.2}} \qquad \text{Eq. 41}$$

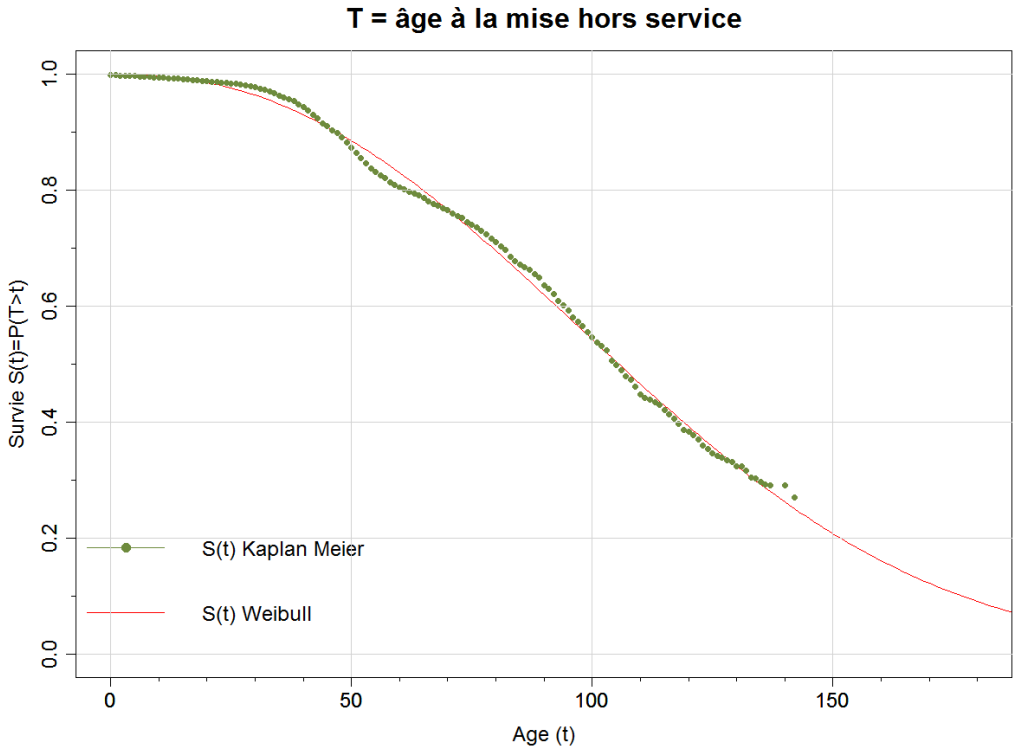


Figure 50 : Courbe de survie prospective (rouge) obtenue à partir de la courbe empirique (issue de KM en vert)

VI.1.2 Linéaires futurs à installer

Pour calculer le premier indicateur I1R nous avons besoin de la courbe de survie prospective du paragraphe précédent mais aussi du linéaire en service en 2012 par date de pose.

Il est possible alors de reconstituer le linéaire passé et de prévoir les besoins (en linéaire) futurs. La démarche suit une procédure itérative pour reconstituer les linéaires posés dans le passé à partir des linéaires observés en 2012. Posons $C\theta$, l'ensemble des tronçons (cohorte) posés la même année (θ). Au début du processus itératif, l'année de pose (θ) varie de 1851 à 2012 (cf. Figure 7). Pour chaque année N, chaque cohorte est caractérisée par une longueur en service : $LES_{c\theta}(N)$ et un âge : $t_{c\theta}(N)$. La survie de la cohorte $C\theta$, dépend de son âge à l'année N: $S[t_{c\theta}(N)]$. L'objectif est de prédire la longueur de la nouvelle cohorte en 2013. Nous pouvons dans un premier temps calculer la longueur qu'avait à l'origine $L_{iniC\theta}(2012)$ à partir de la longueur encore en service en 2012 ($LES_{C\theta}(2012)$) et de la survie attendue pour les tronçons de cet âge à la date d'observation ($S[t_{C\theta}(2012)]$). Cette étape n'est pas obligatoire.

$$L_{iniC\theta}(2012) = \frac{LES_{C\theta}(2012)}{S[t_{C\theta}(2012)]} \qquad \text{Eq. 42}$$

On souhaite par la suite prédire la longueur de la nouvelle cohorte en 2013. On calcule alors pour la cohorte la longueur en service en 2013 (Large 2013b).

$$LES_{C\theta}(2013) = L_{iniC\theta}(2012) \times S[t_{C\theta}(2013)] \qquad \text{Eq. 43}$$

Nous souhaitons conserver un linéaire en service constant. A partir de cette hypothèse nous calculons la longueur des canalisations à poser en 2013.

$$LES_{C2013}(2013) = \sum_{\theta=1851}^{2012} [LES_{C\theta}(2012) - LES_{C\theta}(2013)] \qquad \text{Eq. 44}$$

θ varie maintenant de 1851 à 2013. Pour établir des prévisions jusqu'à l'échéance souhaitée, il est nécessaire d'itérer cette méthode année par année (Large et al. 2014).

La Figure 51 présente les résultats si on itère cette méthode jusqu'en 2050 et la Figure 52 expose les résultats si on itère cette méthode jusqu'en 2120.

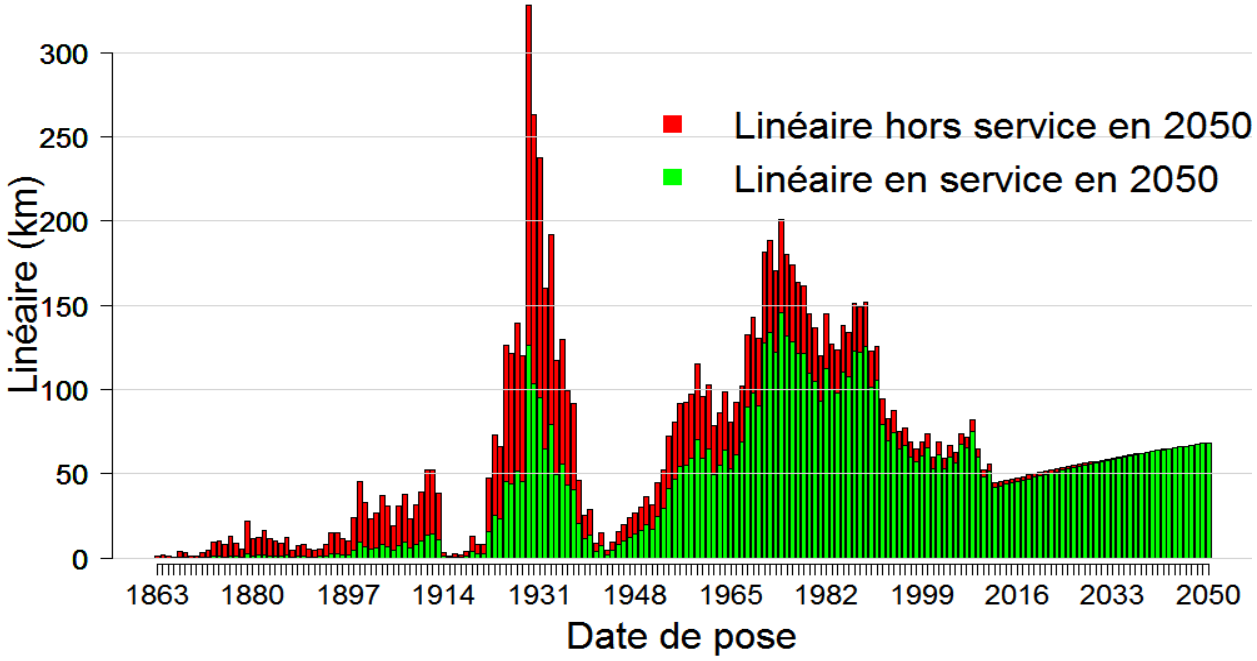


Figure 51 : Reconstitution du linéaire passé et prévision des linéaires futurs jusqu'en 2050, photographie de la situation en 2050, réseaux de transport et de distribution.

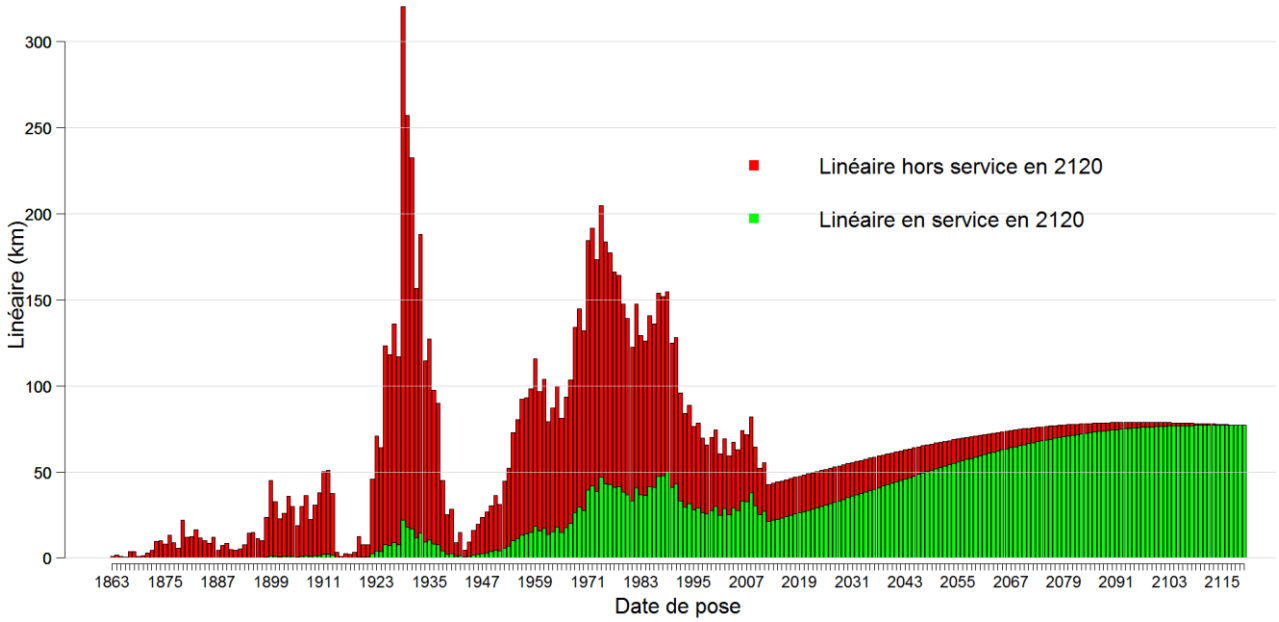


Figure 52 : Reconstitution du linéaire passé et prévision des linéaires futurs jusqu'en 2120, photographie de la situation en 2120, réseaux de transport et de distribution.

VI.1.3 Taux de renouvellement

Le second indicateur, I2R, est le taux de renouvellement annuel futur.

Le taux de renouvellement pour l'année 2013 (*TR*) est donné par :

$$TR(2013) = \frac{LES_{C2013}(2013)}{\sum_{\theta=1851}^{2012} LES_{C\theta}(2012)} \qquad \text{Eq. 45}$$

En réalisant ce calcul chaque année nous obtenons I2.

Avec le scénario « statu quo : 1 courbe» le taux de renouvellement annuel futur du SEDIF reste dans une fourchette relativement étroite et stable, entre 0.5% et 1% (cf. Figure 53). Les résultats sont légèrement supérieurs aux taux de renouvellement appliqués par le SEDIF dans le passé proche, qui étaient en moyenne entre 2005 et 2012 de 0.44% avec un pic en 2009 à 0.8%. Ces taux sont proches des objectifs du SEDIF qui souhaite effectuer 1% de renouvellement annuel à partir de 2015.

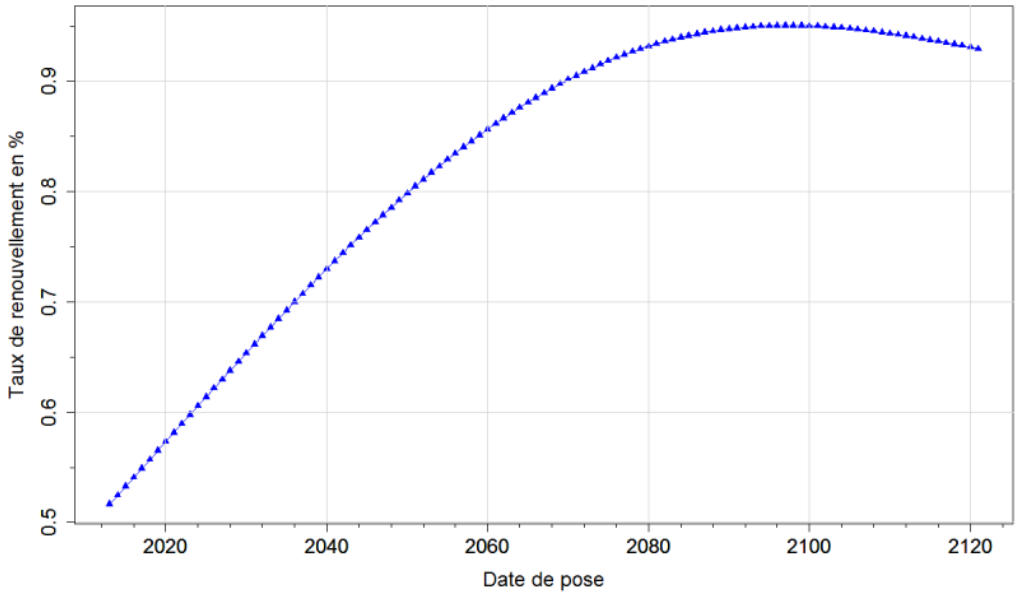


Figure 53 : Prévision des taux de renouvellement futurs de 2013 jusqu'en 2120 au SEDIF

Ces résultats sont intéressants mais il faudrait les comparer aux résultats d'un autre scénario prospectif pour juger de leur pertinence.

VI.1.4 Age moyen du réseau

Le troisième indicateur, I3R, est l'âge moyen du réseau futur. L'âge moyen du réseau pour l'année 2013 (AGE) est donnée par :

$$AGE(2013) = \frac{\sum_{\theta=1851}^{2013}(LES_{C\theta}(2013) \times t_{C\theta}(2013))}{\sum_{\theta=1851}^{2013} LES_{C\theta}(2013)} \qquad \text{Eq. 46}$$

Avec le scénario « statu quo » les âges moyens annuels futurs du SEDIF sont compris entre environ 45 et 66 ans (cf. Figure 54). Le réseau vieillit jusqu'à l'année 2092 puis rajeunit jusqu'à l'année 2120.

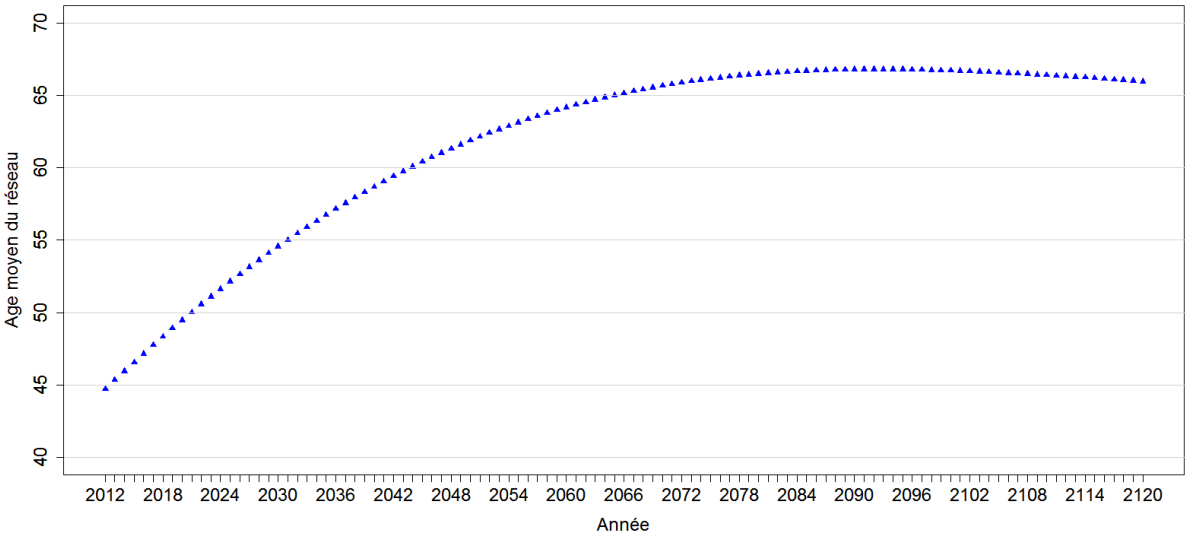


Figure 54 : Prévision des taux de renouvellement futurs de 2013 jusqu'en 2120 au SEDIF avec le scénario « satu quo »

VI.1.5 Coûts d'investissement

Le quatrième indicateur, I1F, est le coût d'investissement annuel futur. Ce coût peut être actualisé ou non actualisé.
L'actualisation est une méthode économique permettant en théorie de rendre commensurable les sommes perçues et dépensées à des époques différentes, tout en intégrant une estimation quantifiée des risques de toutes natures (économiques, financiers, industriels, technologiques, sociaux…) susceptibles de peser sur la « rentabilité » des investissements publics.

Pour les simulations budgétaires, nous faisons l'hypothèse que le renouvellement d'un mètre de canalisation coûte 530€ (500€ comprenant la maîtrise d'œuvre et les études connexes pour un diamètre de 100 mm en fonte + 6% (prise en compte de la diversité des diamètres au SEDIF)). Suite à l'analyse de (Hardelin & Marical 2011) sur les valeurs du taux d'actualisation (r), nous avons choisi d'effectuer deux simulation budgétaires : une avec un taux d'actualisation nul, et une autre avec le taux fixé par le « commissariat général au Plan » (Lebègue 2005) soit 4% de 2013 à 2043 puis décroissant après 2043 (asymptote à 2%). La formule d'actualisation des coûts est donnée par :

$$Coût_{actualise}(2013+k) = \frac{Coût(2013+k) \times (1+e)^k}{(1+r_k)^k} \qquad \text{Eq. 47}$$

Avec :

- r_k le taux d'actualisation pour l'année supplémentaire k avec :

- pour $k \leq 30$: $r_k = 0.04$

- et pour $k > 30$: $r_k = \sqrt[k]{1.04^{30} 1.02^{k-30}} - 1$

- k : année(s) supplémentaire(s) vers lesquelles nous souhaitons faire tendre la prévision, dans notre exemple k varie de 0 à 107 ans
- e : évolution annuelle de l'indice TP10a, il représente l'indice canalisations, égouts, assainissement et adduction d'eau avec fournitures de tuyaux, il est de 3.7% (Cf. Annexe 18)

Pour les conséquences budgétaires, un taux d'actualisation même faible a un très fort impact financier sur du « long terme ». Par conséquent, le scénario avec un taux d'actualisation nul, nous semble ici plus adapté. Ce scénario indique un pic budgétaire en 2096 à 41.8 millions d'euros en investissement pour renouveler les canalisations du réseau principal (hors branchement) (cf. Figure 55, courbe noire). Ce montant est à comparer aux investissements du SEDIF (sous maîtrise d'ouvrage publique uniquement) dans ce domaine en 2012 qui s'élevaient à 19.6 millions d'euros (SEDIF 2012b).

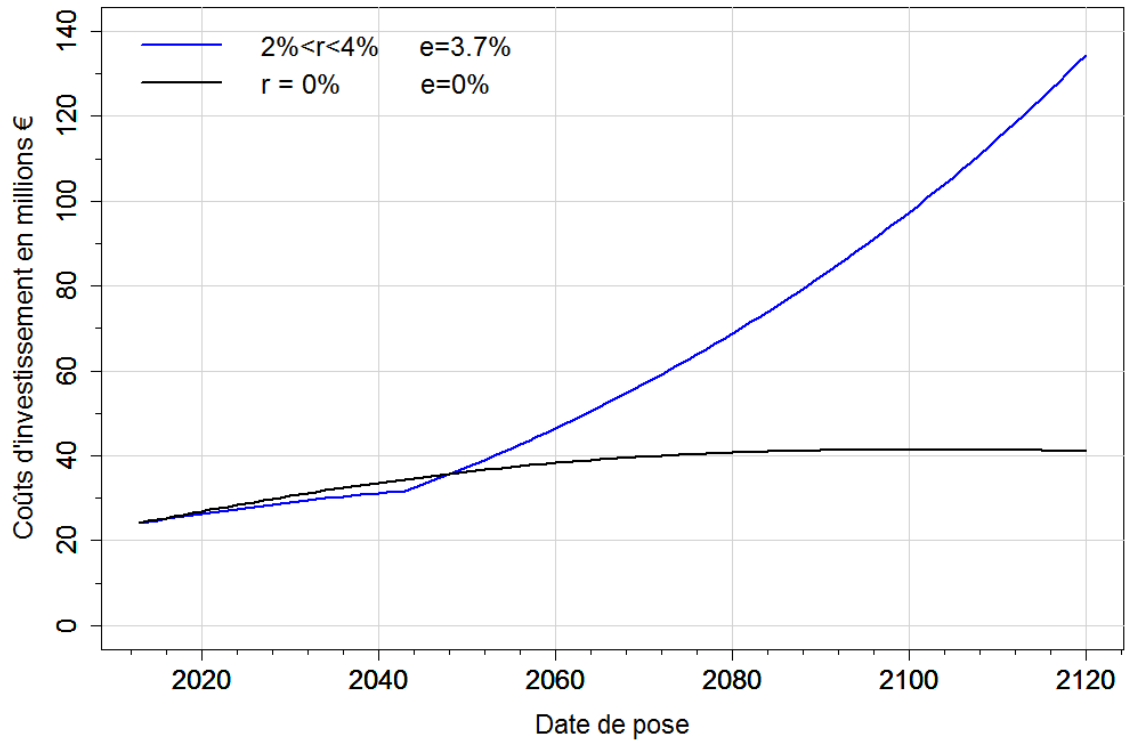


Figure 55 : Prévisions des budgets nécessaires pour le renouvellement entre 2013 et 2120 avec 2 taux d'actualisation différents, réseaux de transport et de distribution

VI.2 Avec plusieurs courbes de survies (partitionnement)

VI.2.1 Courbes de survie prospectives avec partition par diamètre

Les courbes de survie prospectives (cf. Figure 56) du **scénario « statu quo : plusieurs courbes »** sont obtenues à partir :

- des courbes de survie empirique, obtenues en partitionnant le réseau en 2 strates selon le diamètre des tronçons:
    - [min, 150 mm] : 61.1% du linéaire en service en 2012
    -] 150 ; 300 mm] : 24.6% du linéaire en service en 2012
    -] 300, max] : 14.3% du linéaire en service en 2012
- et de la méthode d'ajustement de Weibull décrite dans le paragraphe IV.5.1.1 .

Nous avons décidé ici de partitionner notre échantillon initial selon le diamètre. En effet, ce facteur de partitionnement nous semble très important car les courbes de survie passées pour différents diamètres sont très distinctes (Cf. Figure 56). Or en utilisant le scénario « statu quo : plusieurs courbes » elles continueront de l'être à l'avenir. Les besoins en renouvellements futurs seront donc également différents. De plus, le prix des canalisations est très variable selon le diamètre de celles-ci. L'indicateur des coûts d'investissement prévisionnel sera alors plus précis.

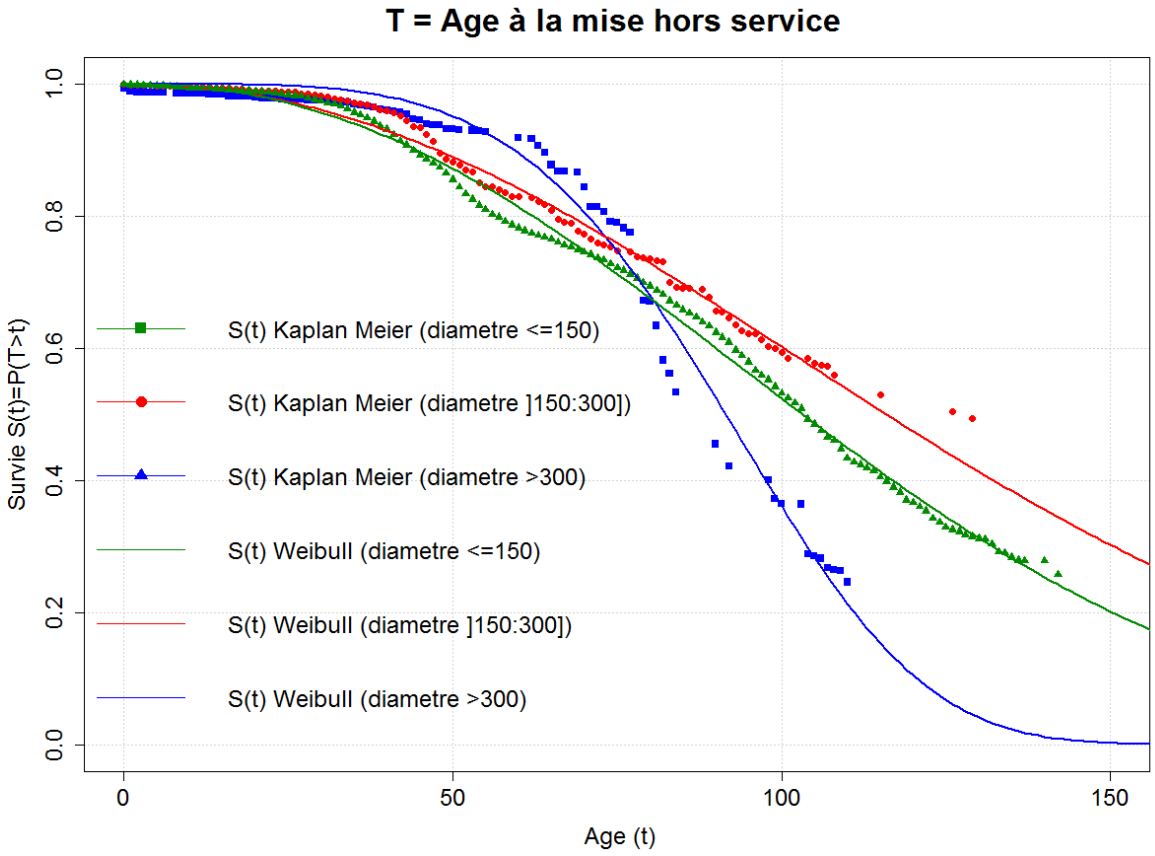


Figure 56 : Courbes de survie prospectives en linéaire (cf. trait fin) obtenue à partir des courbes empirique de Kaplan-Meier et ajusté (équation de Weibull), tous tronçons confondus, partitionnés en 3 sous-ensembles (petits, moyens et gros diamètre)

Les trois fonctions de survie prospectives de Weibull utilisées sont les suivantes :

$$S_{ø>300}(t) = e^{-t^{\delta} \times e^{-\lambda}} = e^{-t^{4.3} \times e^{-20}} \qquad \text{Eq. 48}$$

Label	Estimate	Ref	Chi2	df	Pval
delta	4.3535873e+00	1	1.3391686e+01	1	0.0002527
lambda	2.0030182e+01	12	1.0042234e+02	1	0.0000000

$$S_{150>\text{ø}\geq 300}(t) = e^{-t^{2.1}\times e^{-10.4}} \quad \text{Eq. 49}$$

Label	Estimate	Ref	Chi2	df	Pval
delta	2.1162315e+00	1	9.8934775e+00	1	0.0016587
lambda	1.0425312e+01	12	5.7299270e+00	1	0.0166782

$$S_{150\geq \text{ø}}(t) = e^{-t^{2.23}\times e^{-10.7}} \quad \text{Eq. 50}$$

Label	Estimate	Ref	Chi2	df	Pval
delta	2.2391966e+00	1	3.8999364e+01	1	0.0000000
lambda	1.0749403e+01	10	9.8497563e+00	1	0.0016985

Les « P-values » (Pval) indiquent qu'elles sont bien ajustées aux courbes de survie empiriques.

VI.2.2 Linéaires futurs à installer

Pour calculer le premier indicateur I1R nous avons besoin comme données d'entrée :

- des courbes de survie prospectives de la Figure 56.
- mais aussi des linéaires en service en 2012 par date de pose selon ces trois partitions des tronçons.

Dans le scénario « statu quo : plusieurs courbes », chaque cohorte est scindée en trois selon leurs diamètres :

- [min, 150 mm]
-] 150 ; 300 mm]
-] 300, max]

Ainsi, le linéaire en service de la cohorte 2013 l'année N est alors la somme des linéaires en service de chacune des partitions de cette cohorte soit :

$$LES_{C2013}(\text{N}) = LES_{\substack{C2013 \\ \text{ø}>300}}(\text{N}) + LES_{\substack{C2013 \\ 150>\text{ø}\geq 300}}(\text{N}) + LES_{\substack{C2013 \\ 150\geq \text{ø}}}(\text{N}) \quad \text{Eq. 51}$$

En utilisant la même méthode que celle décrite dans le paragraphe VI.1.2 nous obtenons la Figure 57. Elle illustre la prévision des linéaire à poser jusqu'en 2120.

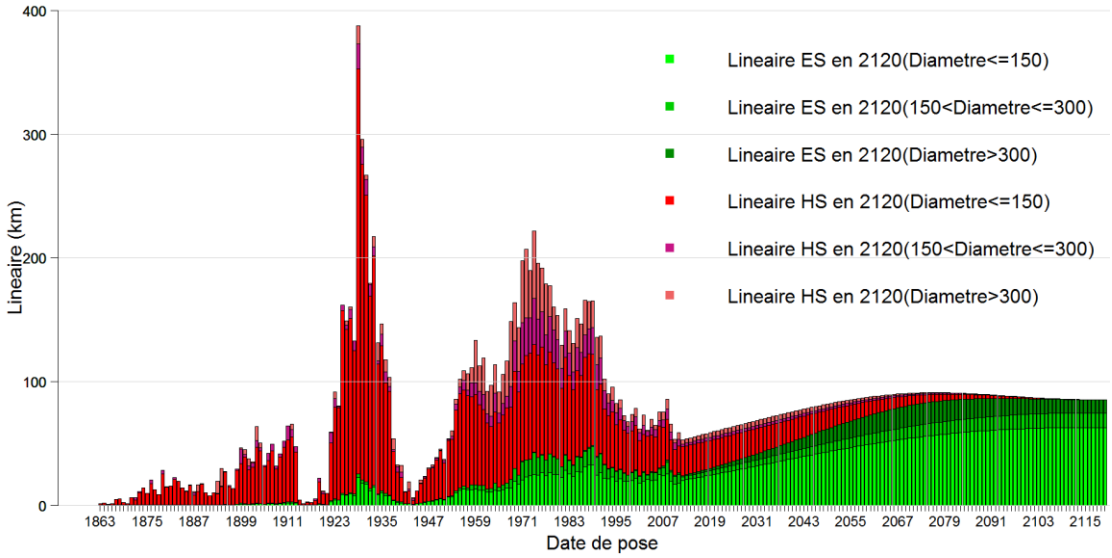


Figure 57 : Reconstitution du linéaire passé et prévision des besoins (en linéaire) futurs jusqu'en 2120, photographie de la situation en 2120, partitionnement par le diamètre.

VI.2.3 Taux de renouvellement

En utilisant le scénario « statu quo, plusieurs courbes » et la méthode décrite dans le paragraphe VI.1.3 pour calculer les taux de renouvellement (I2) nous obtenons les résultats de la Figure 58.

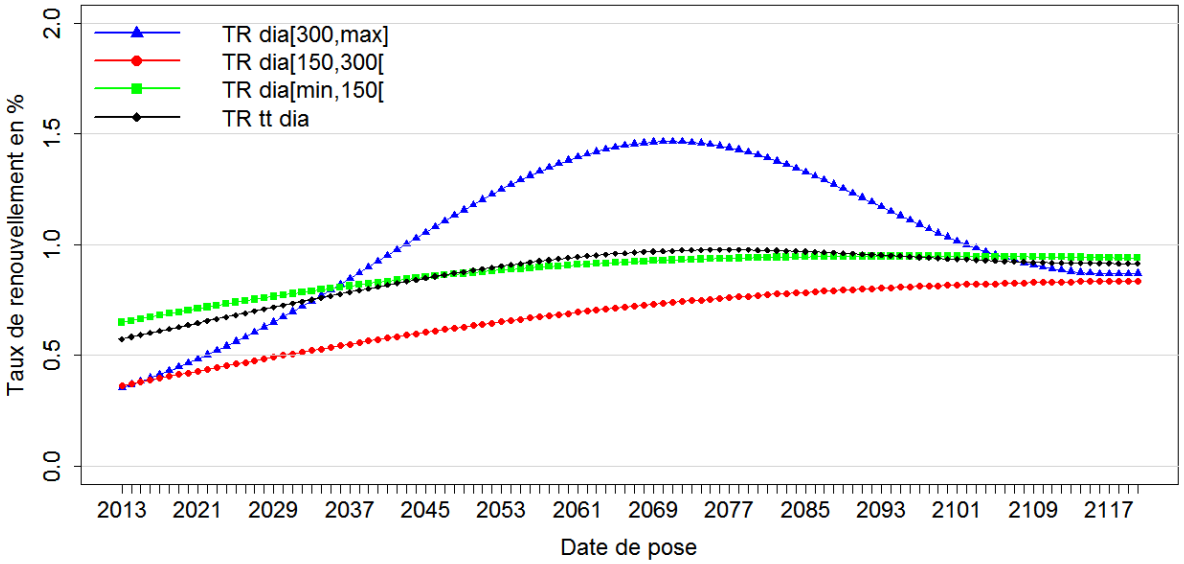


Figure 58 : Prévision des taux de renouvellement futurs de 2013 jusqu'en 2120 pour les tronçons de « gros diamètres », de « moyen diamètres », de « petits diamètres » et « tous diamètres », au SEDIF avec le scénario « statu quo : plusieurs courbes »

Avec le scénario « statu quo, plusieurs courbes », nous observons que les taux de renouvellement des « gros diamètres » sont en général supérieurs au taux de renouvellement des « petits diamètres », eux même supérieure au taux de renouvellement des « moyens diamètres ».

VI.2.4 Age moyen du réseau

En utilisant le scénario « statu quo, plusieurs courbes » et la méthode décrite dans le paragraphe VI.1.4 pour calculer l'évolution de l'âge moyen du réseau (I3) nous obtenons les résultats de la Figure 59.

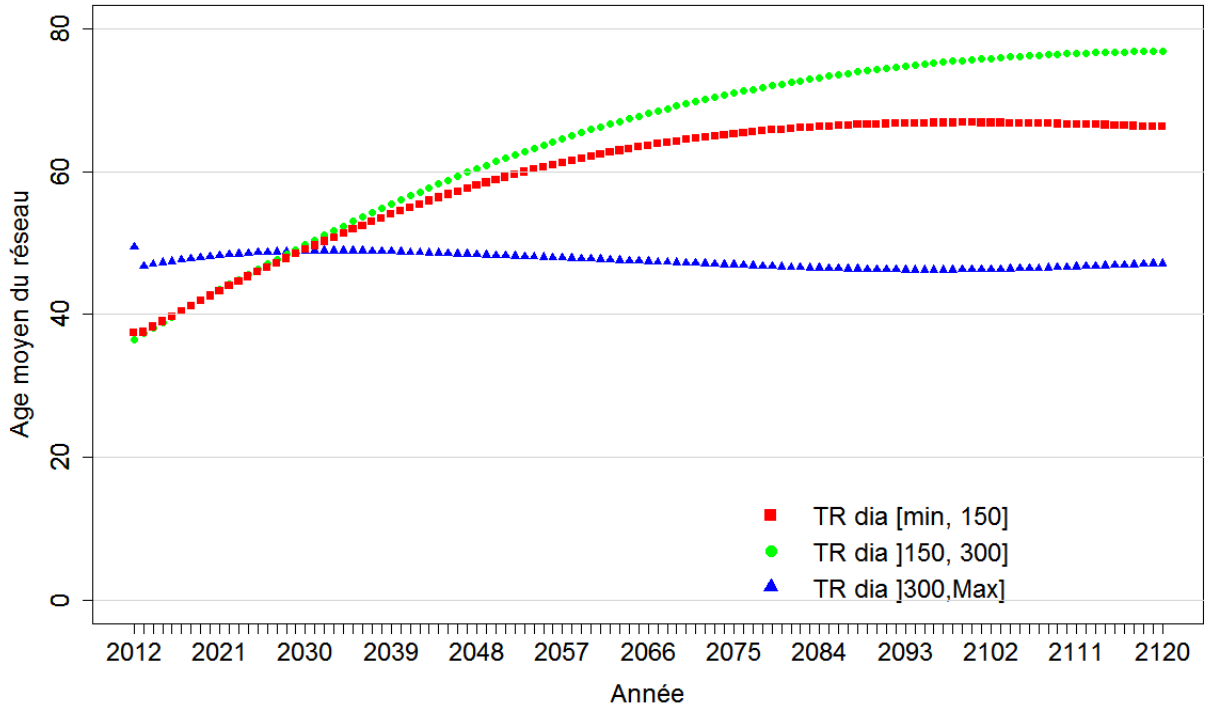


Figure 59 : Prévisions de l'évolution de l'âge moyen du réseau de 2013 jusqu'en 2120 au SEDIF, scénario « statu quo, plusieurs courbes », tronçons partitionnés par diamètre.

L'âge moyen du réseau à long terme pour les tronçons de petits et moyens diamètres est plus grand que pour les tronçons de gros diamètres, ce qui est cohérent car les tronçons de gros diamètres sont ceux qui sont le plus renouvelés.

VI.2.5 Coûts d'investissement

VI.2.5.1 Calcul de l'indicateur coûts d'investissement

Pour les simulations budgétaires, nous prenons comme valeur pour le renouvellement d'un mètre de canalisation, les valeurs ci-dessous. (cf. Tableau 5)

Diamètre	Ø ≤ 150	150 < Ø ≤ 300	Ø > 300
Prix moyen pour le renouvellement d'un mètre de tronçon	407€	567€	840€

Tableau 5 : Prix moyen pour le renouvellement d'un mètre de tronçon par classe de diamètres (Pillot 2011)

Suite à l'analyse de (Hardelin & Marical 2011) sur les valeurs du taux d'actualisation (*r*), nous avons choisi d'effectuer deux simulations budgétaires : une avec un taux d'actualisation nul, et une autre avec le taux fixé par le « commissariat général au Plan » (Lebègue 2005) soit 4% de 2013 à 2043 puis décroissant après 2043 (asymptote à 2%). La formule d'actualisation des coûts est donnée par Eq. 47.

Pour les conséquences budgétaires, un taux d'actualisation même faible a un très fort impact financier sur du « long terme ». Par conséquent, comme ci-avant, le scénario avec un taux d'actualisation nul, semble ici plus adapté. Ce scénario indique un pic budgétaire en 2076 à 46.5 millions d'euros en investissement pour renouveler les canalisations du réseau principal (hors branchement) (cf. Figure 55, courbe noire). Ce montant est à comparer aux investissements du SEDIF (sous maîtrise d'ouvrage publique uniquement) dans ce domaine en 2012 qui s'élevaient à 19.6 millions d'euros (SEDIF 2012b).

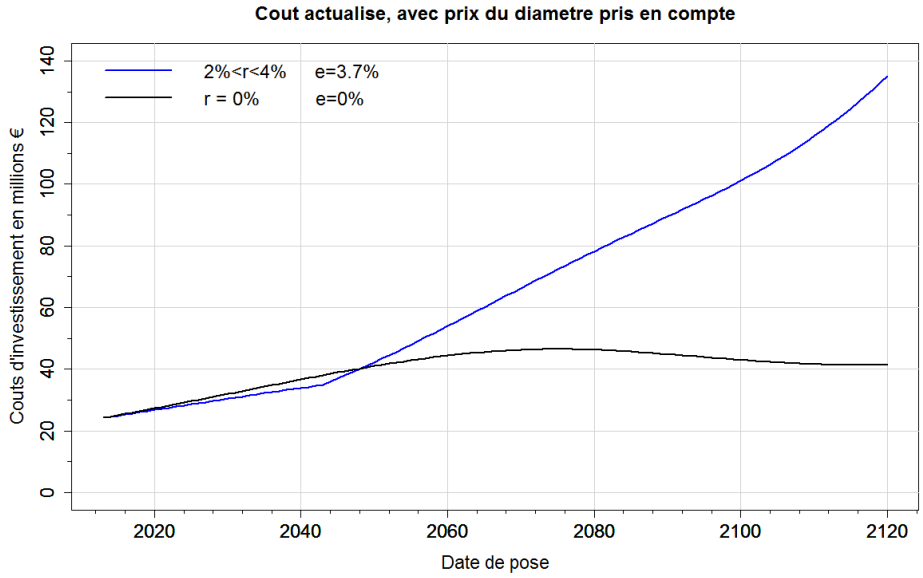


Figure 60 : Prévisions des budgets nécessaires pour le renouvellement entre 2013 et 2120 avec 2 taux d'actualisation différents, partitionnement par le diamètre.

VI.2.5.2 Analyse de sensibilité des coûts d'investissement au taux d'actualisation et à l'indice TP

Nous avons testé l'évolution du coût d'investissement sur la période 2013, 2120 en fonction du taux d'actualisation (r) et de l'indice TP 10a (e).

- Si r varie de 0 à 10% cf. Figure 61

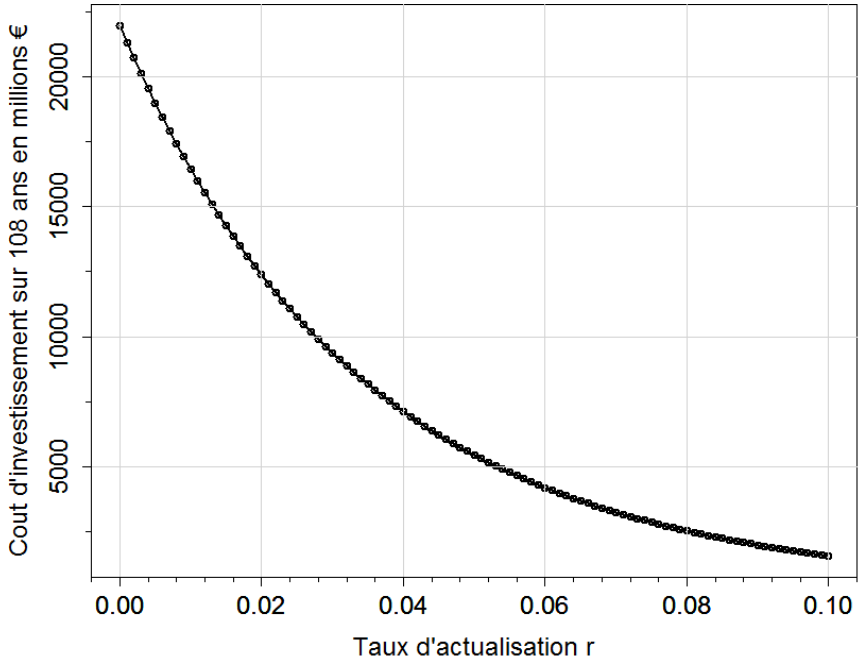


Figure 61 : Evolution du coût d'investissement sur 108 (somme des coûts sur 2013-2120) ans si r varie, e = 3,7%

Nous pouvons noter que pour un indice e fixé à 3.7%, la variation du taux d'actualisation r influence beaucoup le coût d'investissement total. Il semble alors important de bien choisir cette valeur.

- Si e varie de 0 à 6% cf. Figure 62

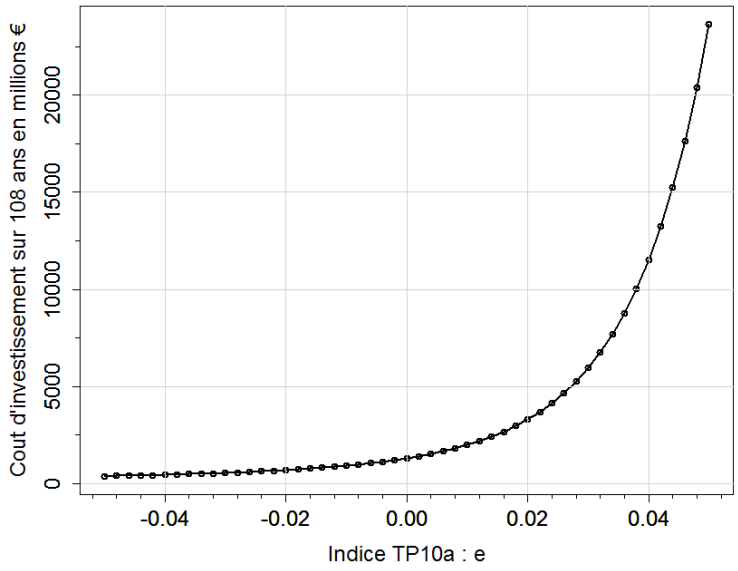


Figure 62 : Evolution du coût d'investissement sur 108 ans (somme des coûts sur 2013-2120) si e varie et r=3%.

Nous pouvons constater là aussi que pour un indice r fixé à 3%, la variation de l'indice TP 10a (e) influence beaucoup le coût d'investissement total. Il semble alors important de bien choisir cette valeur.

VI.2.5.3 Coûts d'investissement non actualisés avec différentiation des strates

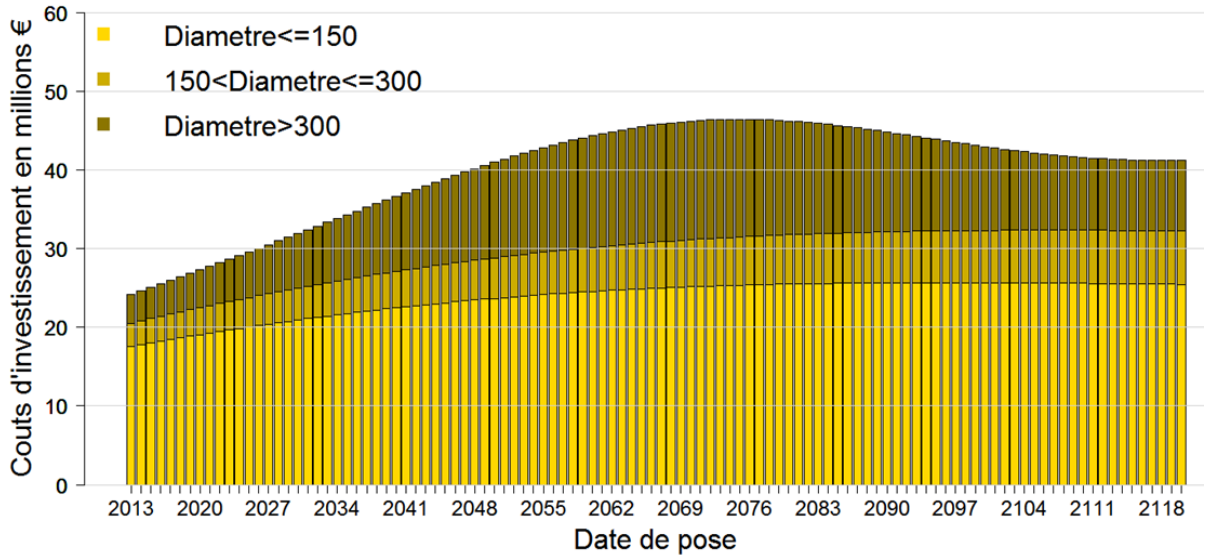


Figure 63 : Prévisions des budgets nécessaires pour le renouvellement entre 2013 et 2120, scénario « statu quo, plusieurs courbes », tronçons partitionnés par diamètre.

Le coût d'investissement à long terme pour les tronçons de petits diamètres est plus important que pour les tronçons de gros diamètres, bien que le coût d'investissement des tronçons de gros diamètres soit tout aussi important. Du faite du prix élevé de c'est tronçons de gros diamètres, donc même si ceux-ci sont moins nombreux que les tronçons de petit diamètres, l'investissement pour c'est tronçons reste relativement importante.

Dans cette sixième partie nous venons de calculer, à partir de deux scénarios prospectifs, 3 indicateurs clés très utile dans un modèle long terme (le linéaire à renouveler, le taux de renouvellement, l'âge moyen du réseau et les coûts d'investissement à venir). Cette ébauche de modèle long terme semble bien plus précis que ceux présentés au chapitre trois, qui utilisent des courbes de survie en escalier complètement déconnectés des pratiques « court terme » réelle au SEDIF.

Néanmoins ces scénarios trop « simplistes » aboutissent à des coûts d'investissement qui semblent excessifs. De plus il aurait été intéressant de calculer l'impact de ces scénarios sur l'évolution probable future des nombres de défaillances.

Conclusion

Le travail réalisé s'est fixé pour **objectif** de créer un modèle afin d'avoir une vision long terme pour optimiser le renouvellement des canalisations d'eau potable du SEDIF.

Dans un premier temps nous avons effectué un **portrait du SEDIF** ainsi qu'une description détaillée du jeu de données brutes fournies. Dans cette partie il a été identifié une fenêtre d'observation fiable des dates à la mise hors service de 18 ans de 1995 à 2012.

Ensuite nous avons **présenté** d'une manière globale **le modèle « long terme »** que l'on souhaite développer. Il en ressort 6 grandes étapes clés.

Puis, un bref état de l'art des **méthodes** de renouvellement « court terme » et « long terme » **employées** au SEDIF a montré que les méthodes « court terme » sont pertinentes. En effet elles prennent notamment en compte la probabilité de défaillance future, les risques de dommages, et les contraintes/opportunités extérieures au service (coordination, etc.). En revanche celles de « long terme » sont grossières, car notamment elles ne prennent pas en compte les dates de pose passées et les pratiques de gestion « court terme ».

Nous avons ensuite présenté des **méthodes** pour estimer **les courbes de survie passés**.

Ceci nous a conduits à construire avec les données du SEDIF plusieurs **courbes de survie passés**, avec (ou pas) différentes stratifications ou en utilisant différentes covariables. Nous avons pu notamment calculer des courbes de survie propres au SEDIF en fonction du diamètre, de la longueur, du nombre de défaillances et du matériau des tronçons.

Pour finir nous avons choisi **deux scénarios prospectifs** pour le SEDIF: « statu quo : 1 courbe » et « statu quo : plusieurs courbes ». De ces scénarios nous en avons déduit **des courbes de survie prospectives**, que nous avons ensuite utilisées pour calculer **4 indicateurs clés** (le linéaire à renouveler, les taux de renouvellement, l'âge moyen du réseau encore les coûts d'investissement).

Il en ressort d'une part que ces scénarios sont encore trop « simplistes » et ils aboutissent à des coûts excessifs. D'autre part pour ces scénarios et tous les autres à venir, il serait très utile d'en déduire les 4 autres indicateurs prévus dans le modèle long terme « optimeau » à savoir notamment l'évolution du nombre de défaillances au cours du temps. Pour information actuellement le code informatique pour calculer cet indicateur est quasiment prêt. Donc ça ne serait tarder … patience…
En effet, simuler les effets d'un scénario sur l'évolution future d'autres indicateurs de performance, permettrait aux gestionnaires de juger si ce scénario est cohérent avec ses objectifs.

Il est à mettre en avant que cette démarche est très reproductible, nous pouvons l'appliquer sans problème sur des réseaux plus modeste. Dans le cadre du projet « optimeau » cette méthode a été appliquée sur les réseaux d'eauservice Lausanne en suisse (900 km) et du Grand Lyon (3000 km).

Un autre axe de recherche crucial sera de nous écarter de ce scénario « **statu quo** » en recherchant des stratégies optimales. Ce travail devra se faire en deux étapes. Dans un premier temps, à partir d'une analyse fine des processus opérationnels de décision « court terme », nous étudierons comment les diverses causes de mise hors service des tronçons contribuent à la construction des courbes de survie observées. Les courbes de survie brutes seront donc doublement corrigées, d'abord pour tenir compte des censures et troncatures, ensuite pour en extraire les effets des décisions prises lors de la programmation des travaux. Nous pourrons alors, dans un second temps reconstruire de nouvelles courbes de survie opérationnelles, qui intégreront de manière explicite les éléments essentiels qui président au processus de décision (par exemple le pourcentage des travaux pilotés par le mauvais état *vs* le pourcentage piloté par le renouvellement de la voirie). Nous pourrons alors estimer comment des stratégies différentes (par exemple profiter plus largement des opportunités de travaux) modifient les performances de « long terme ».

VII Bibliographie

Aalen, O., Borgan, O. & Gjessing, H., 2008. *Survival and Event History Analysis*,

Ajuste, C., Berland, J.-M. & Celerier, J.-L., 2004. *Réhabilitation/remplacement des réseaux d'eau potable en zone rurale*,

Alioum, A. & Leffondre, K., 2014. *Analyse des données de survie*,

Babykina, E., 2010. *Modélisation statistique d'évènements récurrents. Exploration empirique des estimateurs, prise en compte d'une covariable temporelle et application aux défaillances des réseaux d'eau.* Ecole doctorale de Mathématiques et Informatique, Université de Bordeaux II.

Berland, J.-M., Faby, J.-A. & Juery, C., 2005. *La gestion patrimoniale des réseaux d'eau potable : enjeux et recommandations.*,

Bouchet, C., 2013. Réseaux d ' eau potable : optimiser la gestion et améliorer les rendements. *L'eau, l'industrie, les nuisances*, 360, pp.29–43.

Breysse, D., 2009. *Maîtrise des risques en génie civil 1, multiples dimensions des risques en génie civil*. J. Mazars, ed., Lavoisier.

Bruyeron, N. & Re, C., 2012a. *Gestion patrimoniale du reseau d'eau potable du SEDIF, phase 1 - Etude bibliographique - Réalisé par le bureau d'étude Naldeo-Pöyry*,

Bruyeron, N. & Re, C., 2012b. Marche subsequent n°7 Prestations techniques de conseil, de suivi et d'expertise liées aux critères de choix des conduites à renouveler dans le cadre de la gestion patrimoniale du réseau d'eau potable du SEDIF,phase 2-Réalisé par le bureau d'étude Naldeo. , p.111.

Charrière, S., 2013. *Gestion patrimoniale des réseaux d'eau potable. Elaboration du descriptif détaillé des ouvrages de transport et de distribution d'eau .*,

Claude, U. & Lyon, B., 1997. Generalisation de l'estimateur Kaplan-Meier d'une loi de durée de maintien en presence d'observations tronquées à gauche. , 1, pp.97–114.

Claudio, K., 2011. *Prise en compte de facteurs dépendants du temps dans les modèles de vieillissement des canalisations.* Université Bordeaux 2.

Cox, D.R., 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), pp.187–220.

D. Schoenfeld, 1982. Partial Residuals for The Proportionnal Hazards Regression Model. , pp.239–241.

France_Libertés, 2014. Les fuites d'eau (vous) coûtent 20% de votre facture. *La tribune*, (78703), p.2.

Le Gat, Y., 2009. *Une extension du processus de Yule pour la modélisation stochastique des événements récurrents. Application aux défaillances de canalisations d'eau sous pression.* Institut des Sciences et Industries du Vivant et de l'Environnement (AgroParisTech).

Le Gat, Y., Kropp, I. & Poulton, M., 2013. Is the service life of water distribution pipelines linked to their failure rate? *Water Science & Technology: Water Supply*, 13.2, pp.386–393.

Hardelin, J. & Marical, F., 2011. Taux d'actualisation et politiques environnementales : un point sur le débat. Commissariat général au développement durable (CGDD). , p.18.

Herz, R., 1996. Ageing processes and rehabilitation needs of drinking water distribution networks. *Journal of Water Supply: Research and Technology*, pp.221–231.

Kaplan, E.L. & Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. , pp.457–481.

Klein, J. & Moeschberger, M., 2003. *Survival Analysis, Techniques for Censored and Truncated Data, Second Edition*,

Lanez, É., 2010. Veolia empoche le contrat du siècle. *Le Point*. Available at: http://www.lepoint.fr/actualites-economie/2010-05-27/veolia-empoche-le-contrat-du-siecle/916/0/459718.

Large, A., 2013a. *Comité de suivi de thèse 1 [Powerpoint]*,

Large, A. et al., 2014. Decision support tools: Review of risk models in drinking water network asset management. In *ICVRAM International Conference on Vulnerability and Risk Analysis and Management, ISUMA International Symposium on Uncertainty Modelling and Analysis*.

Large, A., 2013b. *Optimisation du renouvellement des canalisations. Rapport de première année de thèse : revue bibliographique et cadre méthodologique.*,

Lebègue, D., 2005. Révision du taux d'actualisation des investissements publics. Rapport du groupe d'experts présidé par Daniel Lebègue. Comissariat général au Plan. , p.112.

Levesque, G., 2014. *Interview avec monsieur LEVESQUE, Chef de projet SIG au SEDIF*,

Lion, V., 2009. L'eau à Paris, une guerre entre Veolia et Suez. *l'express*. Available at: http://lexpansion.lexpress.fr/actualite-economique/l-eau-a-paris-une-guerre-entre-veolia-et-suez_833413.html.

Luersen, M.A. & Le Riche, R., 2001. Globalisation de l ' Algorithme de Nelder-Mead : Application aux Composites Globalisation de l ' Algorithme de Nelder-Mead : Application aux Composites. , p.44.

Ministère_de_L'écologie, 2010. Ministère de l'écologie, du développement durable et de l'énergie. Available at: http://www.developpement-durable.gouv.fr/Structure-des-index-TP.html.

Naldeo, 2013. Patrimoine expert, Outil de simulation du vieillissement du patrimoine canalisations AEP, Manuel de l'utilisateur, Version 3.4. , p.25.

Nelder, J.A. & Mead, R., 1965. The downhill simplex method. *Computer Journal*, 7, pp.308–3013.

Optimeau, 2014. Compte rendu du comité de suivi n°3 du projet optimeau le 15/10/2014 au SEDIF, Paris. , p.8.

Saint Pierre, P., 2013. Introduction à l'analyse des durées de survie. , p.47.

Pillot, J., 2011. Élaboration d ' une stratégie de gestion du patrimoine de l ' Eau Potable du Grand Lyon. , p.55.

Poinard, D., 2006. *Modèles pour la conception de stratégies et de programmes de réhabilitation des réseaux urbains d'eau potable.* Institut National des Sciences Appliquées (INSA) de Lyon.

Poulton, M. & Saignol, P., 2013. Etablissement d'un programme de travaux de remplacement des conduites de distribution - Etude réalisé par le cabinet Merlin et le bureau d'étude WTSim. , p.35.

SEDIF, 2013a. *Entretien avec le SEDIF (puis VEDIF), mené le 18 et 19 juillet à Paris et Nanterre.*,

SEDIF, 2014. Fonctionnement du réseau AEP. Available at: http://www.sedif.com/fonctionnement_reseau.aspx.

SEDIF, 2012a. *Gestion patrimoniale : Diagnostic des canalisations en fonte sur le réseau du SEDIF*,

SEDIF, 2013b. SEDIF Comité de suivi de la thèse OPTIMEAU [Powerpoint]. , p.17.

SEDIF, 2012b. *SEDIF-Rapport d'activité 2012*,

TNS/SOFRES, 2012. Observatoire de la qualité du service public de l'eau (SEDIF).

Turnbull, B.W., 1976. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society*, 3, pp.290–295.

Annexe 1: Fiche intervention pour une défaillance, exemple

TRAITEMENT PREALABLE pour
Intervention urgente (demande reçue par tél. / fax)
Dernière mise à jour 23/05/2006

Réf dossier : 317194 **Num Provisoire :**
Interlocuteur : VEOLIA EAU ENTRETIEN ET DOMMAGES
N° Tél./Fax : /
Adresse : 8 B AV PIERRE ET MARIE CURIE DRANCY
Nature des Travaux : Fuite sur conduite de diamètre <= 300 mm
Appel : **Date :** ☐ **différée : Date report :**
Heure : **Heure :**

Fouille tracée :	Mise en chantier par inspecteur :	Appel suite à dommage :	Arrêt d'eau :
NON	NON	NON	

Définition de l'intervention : **Fuite sur :**
Conduite **matériau :** **Ø :**
Branchement **matériau :** **Ø :**
Robinet Vanne **Ø :**
Robinet de prise **Ø :**
Collier de prise **Ø / ø :** /
Nettoyage :
Mise à niveau de bouche à clé ☐ **coffre** ☐
Autres :

Dispositions / consignes particulières
Accès mini-pelle possible ☐
Entourage de sécurité immédiat ☐
Panneaux particuliers de signalisation ☐
Déviation / barrage de rue ☐
Feux tricolores ☐
Sablage chaussée ☐ **/ cave** ☐
Vidange cave ☐
Autres :

Informations sur le terrain
Position
Voie
RDS **RDS Veolia Eau d'Ile de France**
Revêtement
0000000000

Informé :
M. .....
à .....**heures**.....
BRS établi ☐
Réfection commandée ☐

Demande d'intervention reçue par :
Visa rédacteur :

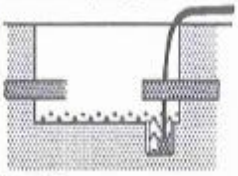
Informations Compteur : N° :
Index :
Cachetage ☐

Reste à enlever : *Terres en sac :* oui non *Boues à nettoyer :* oui non
Matériel et/ou fournitures :

Réfection provisoire de sols : Grave ciment : oui non Noir à froid : oui non
Autres : .....

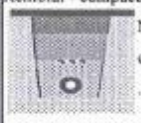
Validation Client :
Nom : Signature :
Commentaires :

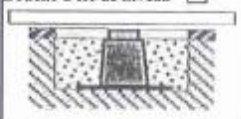
REALISATION AUTOCONTROLES

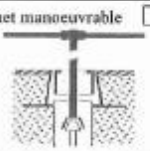
Conduite préservée hors eaux des fouilles (pompages)

Oui ☐
si non, Client informé ☐

Nettoyage, désinfection de la coupe et des pièces utilisées ☐

Pastilles introduites à la demande du Client ☐
Nombre : .....

Remblai - compactage

Nbre de couches : .....

Bouche à clé de niveau ☐


Robinet manoeuvrable ☐


Etanchéité à la remise en eau

☐

Coffre scellé de niveau et l'eau peut s'écouler ☐

Nettoyage du chantier et des abords ☐

Signature du plombier :

Annexe 2: Nettoyage du fichier défaillances FR

Si nous trouvons dans LOCALISATION	Nous remplaçons dans Type1 par
Tuyau	Tuyau
Collier	Joint
Appareillage feeder	Appareillage
Appareillage petit réseau	Appareillage
Autres appareils	Appareillage
Brt décharge ou ventouse (petit réseau)	Appareillage
Coffre	Appareillage
Nourrice	Appareillage
Pièce spéciale	Appareillage
RA ou RV	Appareillage
Robinet jauge	Appareillage
RV	Appareillage
RV ou robinet sur prise	Appareillage
Tubulure décharge ou ventouse (feeder)	Appareillage
Tubulure ventouse - décharge	Appareillage
inconnu	Inconnu
Inconnu	Inconnu
""	Inconnu

Si nous trouvons dans TYPE_FUITE	Nous remplaçons dans Type2 par
Joints	Joint
Joint (bride, vissé)	Joint
Piqûre - Corrosion	Piqure
Piqûre - cratère	Piqure
Piqûre - Cratère	Piqure
Piqûre ou cratère	Piqure
Rupture circulaire	Rupture circulaire
Rupture circulaire s/bague	Rupture circulaire
Rupture circulaire804723/1	Rupture circulaire
Rupture longitudinale	Rupture longitudinale

Fente	Rupture longitudinale
Fente PE	Rupture longitudinale
Fente sur plomb ou PE	Rupture longitudinale
Eclat	Eclat
Eclat (fonte)	Eclat
Divers sur RP et RV	Appareillage
Divers sur RV ou RPC	Appareillage
Presse-étoupe	Appareillage
Presse étoupe	Appareillage
Sur appareillage Feeder	Appareillage
Non déterminé	Inconnu
Autre	Inconnu

Si nous trouvons dans Type1 et Type2	Nous remplaçons dans Type3 par
Type1 == "Tuyau" &Type2 == "Rupture circulaire"	Tuyau
Type1 == "Tuyau" & Type2 == "Rupture longitudinale"	Tuyau
Type1 == "Tuyau" & Type2 == "Piqure"	Tuyau
Type1 == "Tuyau" & Type2 == "Eclat"	Tuyau
Type1 == "Tuyau" & Type2 == "Inconnu"	Tuyau
Type1 == "Joint" & Type2 == "Joint"	Joint
Type1 == "Tuyau" & Type2 == "Joint"	Joint
Type1 == "Joint" & Type2 == "Eclat"	Joint
Type1 == "Joint" & Type2 == "Piqure"	Joint
Type1 == "Joint" & Type2 == "Rupture circulaire"	Joint
Type1 == "Joint" & Type2 == "Rupture longitudinale"	Joint
Type1 == "Appareillage" & Type2 == "Appareillage"	Appareillage
Type1 == "Appareillage" & Type2 == "Joint"	Appareillage
Type1 == "Appareillage" & Type2 == "Eclat"	Appareillage
Type1 == "Tuyau" & Type2 == "Appareillage"	Appareillage
Type1 == "Joint" & Type2 == "Appareillage"	Appareillage
Type1 == "Appareillage" & Type2 == "Rupture longitudinale"	Appareillage
Type1 == "Appareillage" & Type2 == "Rupture circulaire"	Appareillage
Type1 == "Appareillage" & Type2 == "Piqure"	Appareillage
Type1 == "Appareillage" & Type2 == "Inconnu"	Appareillage
Type1 == "Inconnu" & Type2 == "Inconnu"	Inconnu

Annexe 3: Méthode de nettoyage de la variable type joints

Si nous trouvons	Nous remplaçons par
"Express Non Vi"	Joints_visses_a_bride
"Gibault ou assimilé"	Joints_visses_a_bride
"Vissé"	Joints_visses_a_bride
"Express Vi"	Joints_visses_a_bride
"A bride"	Joints_visses_a_bride
"Viking-Johnson"	Joints_visses_a_bride
"Automatique Non Vi (std/universal)"	Joint_a_emboitement
"Automatique Vi (std/universal)"	Joint_a_emboitement
"Automatique Non Vi "	Joint_a_emboitement
"Automatique Vi"	Joint_a_emboitement
"Emboîtement acier"	Joint_a_emboitement
"Tulipe"	Joint_a_emboitement
"Coulé (plomb)"	Coule_plomb
"Coule (plomb)"	Coule_plomb
"CoulCoulé (plomb)"	Coule_plomb
"Soudé"	Soude
"SoudSoudé"	Soude
"Electro-soudElectro-soudé"	Electro_Soude
"Electro-soudé"	Electro_Soude
"Autre"	Autre
"JK"	Autre
"Collé"	Autre
"Electro-soudé"	Autre
"Triduct"	Autre
"Emboîtement acier"	Autre
"Inconnu"	Autre
"Push-Fast"	Autre

Annexe 4: Méthode de nettoyage de la variable technique de pose

Si nous trouvons	Nous remplaçons par
"Traditionnel"	Traditionnel
"Autre"	Autre
"Caniveau"	Autre
"Eclatement"	Autre
"Extraction"	Autre
"Fonçage"	Autre
"Forage dirigé"	Autre
"Forage dirigForage dirigé"	Autre
"Fourreau"	Autre
"Microtunnel"	Autre
"Tranchée ouverte"	Autre
"Tubage"	Autre
"Galerie"	Autre

Annexe 5: Méthode de nettoyage de la variable emplacement

Si nous trouvons	Nous remplaçons par
"ChaussChaussée"	Chaussee
"ChaussChaussée impaire"	Chaussee
"ChaussChaussée paire"	Chaussee
"Chaussé"	Chaussee
"Chaussée"	Chaussee
"Chaussée centrale"	Chaussee
"Chaussée côté bâti"	Chaussee
"Chaussée côté non bâti"	Chaussee
"Chaussée impaire"	Chaussee
"Chaussée paire"	Chaussee
"Chaussée paire/impaire"	Chaussee
"Chaussée paire"	Chaussee
"Parking-esplanade"	Chaussee
"Trottoir"	Trottoir
"Trottoir côté bâti"	Trottoir
"Trottoir côté non bâti"	Trottoir
"Trottoir impair"	Trottoir
"Trottoir pair"	Trottoir
"Trottoir pair/impair"	Trottoir
"Voie piétonne"	Trottoir
"Accotement"	Autre
"Autre"	Autre
"Chemin"	Autre
"Côté place"	Autre
"Egout"	Autre
"Espace vert"	Autre
"Galerie"	Autre
"Hors-sol (ouvrage d'art)"	Autre
"Ilot central"	Autre
"Inconnu"	Autre
"Place"	Autre
"Propriété privée"	Autre
"Sous fluvial"	Autre
"Sous voie ferrée"	Autre
"Sur fluvial"	Autre
"Transversal"	Autre
"Vertical"	Autre

Annexe 6: Méthode de nettoyage de la variable type appui

Si nous trouvons	Nous remplaçons par
"Assise béton (1/4 ou 1/2 tuyau)"	Beton
"Enrobage bEnrobage béton"	Beton
"Radier béton + assise béton (1/4 ou 1/2 tuyau)"	Beton
"Enrobage bEnrobage béton"	Beton
"Sur pieux et dalle béton"	Beton
"Enrobage béton"	Beton
"Remblai béton"	Beton
"Solin"	Autre
"Ancienne conduite (tubage)"	Autre
"Collier Williamson"	Autre
"Sol naturel"	Autre
"Sur berceaux (galeries, ponts, caniveaux)"	Autre
"Sol compacté"	Sol_Compact
"Sol compactSol compacté"	Sol_Compact
"Sol compacte"	Sol_Compact

Annexe 7: Les critères de renouvellement des canalisations par le VEDIF sur le territoire du SEDIF en 2010

A. PROGRAMME INITIAL

1. Opérations obligatoires

- **Critère S :** Matériaux présentant un risque pour la santé type amiante-ciment
- **Critère OPV :** Canalisations impactées par des travaux de voirie lourds dans le cas où les conditions d'exécution des travaux imposeraient le renouvellement de la conduite

2. Critères hiérarchisés (par ordre d'importance et de priorité)

a. Ruptures constatées

- **Critère F :** Canalisations au taux de fuite supérieur à 2 sur les 5 dernières années
- **Critère F1 :** Canalisations vétustes favorisant les tronçons supérieurs à 100 mètres, avec taux de fuite >1,8 et <2 constaté sur 5 dernières années
- **Critère F2 :** Canalisations vétustes favorisant les tronçons supérieurs à 150 mètres, avec taux de fuite est supérieur à 1,5

b. Ruptures anticipées

- **Critère C :** Canalisations à risque de rupture identifiées dans le cadre de l'étude CEMAGREF

c. Travaux de voirie

- **Critère OPV1 :** Canalisations impactées par des travaux de voirie lourds dans le cas où les conditions d'exécution des travaux permettraient le maintien en l'état de la canalisation

Le remplacement peut s'effectuer sous au moins une des conditions suivantes :

- Age et taux de fuite :

Si âge de la canalisation est de :	Avec taux de fuite de :
≥ 70 ans	≥ 0,6
≥ 80 ans	≥ 0,4
≥ 90 ans	≥ 0,2
≥ 100 ans	≥ 0

- Si la canalisation correspond à une « fonte à risque » (canalisations posées de 1905 à 1919 et de 1945 à 1969)
- Si la canalisation est en acier

- **Critère OPV2 :** Canalisations impactées par des travaux de voirie légers, (exemple renouvellement d'un tapis d'enrobée)

Le remplacement peut s'effectuer sous au moins une des conditions suivantes :

- Taux de fuite supérieur à 1,3
- Conduites des années de « fonte à risque » (canalisations posées de 1905 à 1919 et de 1945 à 1969),
- Conduites en acier

d. <u>Risques identifiés pour l'environnement</u>

- **Critère R1 :** Canalisations susceptibles de représenter un risque pour un ouvrage à proximité
- **Critère R2 :** Canalisations susceptibles de représenter un risque dans une zone PPR et plus spécifiquement en zones bleues suite à la présentation et à l'analyse d'une étude d'impact sur le réseau de distribution locale
- **Critère R3 :** Incompatibilité de certains matériaux dans des terrains instables
- **Critère R4 :** Risques identifiés au tiers (mise à niveau de la sécurité incendie, risque engendré par une canalisation vétuste, etc…)

e. <u>Gestion patrimoniale :</u>

- **Critère P1 :** Canalisations en matériaux aciers devant disparaître du réseau
- **Critère P2 :** Canalisations susceptibles de représenter un risque constaté après analyses métallographiques
- **Critère P3 :** Restructuration du réseau (abandon de conduite, …)
- **Critère P4 :** Canalisations de diamètre non normalisé et de matériaux peu usités devant disparaître du réseau

<u>B. PROGRAMME COMPLEMENTAIRE</u> (inter secteur géographique)

Les critères suivants sont à retenir, dans l'ordre d'importance :

- **Critère PC1 :** Canalisations en lien avec des opérations de voirie présentant des incertitudes de planification
- **Critère PC2 :** Canalisations avec un taux de fuite supérieur ou égal à 1
- **Critère PC3 :** Canalisations posées avant 1900 avec un taux de fuite supérieur ou égal à 0,5

Critères 2010	Linéaire inscrit au programme 2010 (m)		Linéaire désaffecté en 2010 d'après SIG (m)		Ecart
S	320	320	349	349	
F	3 853	25 222	3 475	19 114	-24,2%
F1	1 441		1 081		
F2	15 791		12 155		
C	4 137		2 402		
OPV	6 596	14 502	1 215	7 066	-51,3%
OPV1	7 060		5 012		
OPV2	846		838		
R1	470	670	514	718	7,1%
R2	-		-		
R3	-		-		
R4	200		203		
P1	-	280	-	318	13,7%
P2	-		-		
P3	280		318		
P4	-		-		
Total programme initial		**40 994**		**27 565**	**-33%**

Critères 2010	Linéaire inscrit au programme 2010 (m)	Linéaire désaffecté en 2010 d'après SIG (m)	Ecart
Total programme initial	**40 994**	**27 565**	**-33%**
PC1	6 465	384	-94%
PC2	896	-	-100%
PC3	230	-	-100%
Total programme complémentaire	**7 591**	**384**	**-95%**
Non retrouvé	-	14 164	
Total général	**48 585**	**42 112**	**-13%**

Tableau 6 : Les critères de renouvellement des canalisations par Véolia sur le territoire du SEDIF en 2010 (Bruyeron & Re, 2012)

Annexe 8 : Définition de la censure et de la troncature

Définition de la Censure

La censure est le phénomène le plus couramment rencontré. Soit une population observée de n individus. Pour l'individu numéro i, $i \in \{1,2,\dots,n\}$ considérons :

- Son temps de survie X_i,
- Son temps de censure W_i,
- La durée réellement observée T_i.

L'événement aléatoire étudié est la mise hors service. La censure peut être « à droite » ou « à gauche », à droite si l'individu n'a pas subi l'événement à sa dernière observation, dans ce cas les durées de maintien en service ne sont pas toutes observées. Pour certaines d'entre elles, nous savons seulement qu'elles sont supérieures à une certaine valeur connue. Ou à gauche dans le cas où l'individu a déjà subi l'événement avant qu'il soit observé. Nous savons alors uniquement que la date de l'événement est inférieure à une certaine date connue.

- Censure à droite : Soit W une valeur fixée ; au lieu d'observer les variables $X_1,\dots,X_n$ qui nous intéressent, X_i est uniquement observé lorsque$X_i \leq W$. Sinon nous savons seulement que $X_i > W$.
  T_i se note $T_i = \min(X_i, W)$, $i = 1,\dots,n$
- Censure à gauche : Soit W une valeur fixée ; au lieu d'observer les variables $X_1,\dots,X_n$ qui nous intéressent, X_i est uniquement observé lorsque$X_i \geq W$. Sinon nous savons seulement que $X_i < W$.
  T_i se note $T_i = \max(X_i, W)$, $i = 1,\dots,n$

Définition de la Troncature

Les troncatures diffèrent des censures dans le sens où elles concernent l'échantillonnage lui-même. Ainsi, une variable X est tronquée par un sous-ensemble éventuellement aléatoire A de $\mathbb{R}^+$si au lieu de X, X est uniquement observé si $X \in A$, et suivant donc la loi de T conditionnée par l'appartenance à A. S'il y a troncature, une partie des individus n'est pas observable et l'étude porte seulement sur un sous-échantillon (Saint Pierre 2013).

Annexe 9 : Méthode pour calculer la survie des tronçons hors service = « survie brute »

Nous estimons dans ce paragraphe, la distribution des âges à la mise hors service uniquement des tronçons mis hors service dans la fenêtre d'observation.

En linéaire, l'individu est 1m de tronçon, alors qu'en effectif, l'individu est un tronçon.

- En linéaire :

Nous calculons dans un premier temps les courbes de survie « brutes » *Sb(t)* des tronçons hors service (en linéaire). Parmi les données archivées par le SEDIF nous travaillons sur la population des tronçons hors service entre [1995 ; 2012]. Chaque tronçon i a une longueur L_i, $i = 1, \dots, n$ et un âge de mise hors service b_i. Les tronçons ici sont tous non-censurés ($c_i = 0$) c'est-à-dire que leur mise HS est observable dans la fenêtre [2001 ; 2012]. Notons *t* les réalisations de la variable âge à la mise hors service. Posons $I\{\dots\}$ la fonction indicatrice qui prend 1 si les conditions entre accolades sont respectées et 0 sinon. Nous estimons d'abord la fréquence brute (en linéaire) $D_L(t)$ puis la fonction de répartition brute (en linéaire) $F_L(t)$ et enfin la survie brute (en linéaire) $Sb_L(t)$.

$$D_L(t) = \frac{\sum_i (L_i \times I\{c_i = 0 \ \& \ b_i = t\})}{\sum_i (L_i \times I\{c_i = 0\})} \qquad \text{Eq. 52}$$

$$Sb_L(t) = 1 - F_L(t) = 1 - \sum_{k \le t} D_L(k) \qquad \text{Eq. 53}$$

- En effectif :

En effectif nous ne comptons plus le linéaire mais le nombre de tronçon qui tombe hors service à un temps *t*. Nous estimons d'abord la fréquence brute *D(t)* puis la fonction de répartition brute *F(t)* et enfin la survie brute *Sb(t)*. Nous retrouvons alors :

$$D(t) = \frac{\sum_i I\{c_i = 0 \ \& \ b_i = t\}}{\sum_i I\{c_i = 0\}} \qquad \text{Eq. 54}$$

$$Sb(t) = 1 - F(t) = 1 - \sum_{k \le t} D(k) \qquad \text{Eq. 55}$$

Annexe 10 : Description de la méthode de Turnbull

L'estimateur non paramétrique de Turnbull sert à estimer une fonction de survie d'après des données de durées de maintien en service. Une fonction de survie donne la probabilité de survivre au-delà d'un certain âge, notée :

$$S(t) = P(T>t)$$

La courbe de survie a pour valeur de départ S(0)= 1, est décroissante et tend vers 0 quand t tend vers l'infini.

L'estimation non paramétrique de courbe de survie est une fonction en escalier décroissante qui saute vers le bas à chaque valeur de durée de maintien en service observée (quand un ou plusieurs individus « meurent » [ou sont mis hors service]), tout en restant constante (horizontale) entre deux valeurs successives de durée de maintien en service (Cf. Figure 64). En bref, la méthode non paramétrique de Turnbull consiste à estimer la hauteur des sauts. Une présentation plus formelle de la procédure de ce calcul nécessite d'établir des conventions de notation spécifiques.

Notons alors T la variable aléatoire durée de maintien en service, observée sur un ensemble de N tronçons du réseau, divisé en n groupes. Chaque groupe est constitué d'un nombre de tronçons e_i (avec $N = \sum_{i\in\{1,\dots,n\}} e_i$) observés dans l'intervalle d'âge $[a_i, b_i]$; la variable aléatoire c prend la valeur $c_i = 0$ si les segments du groupe i ont été mis hors service à l'âge b_i, sinon la valeur $c_i = 1$ si les segments du groupe i n'ont pas encore été mis hors service lorsque leur observation s'est arrêtée à l'âge b_i. Le premier cas $c_i = 0$ signifie l'observation exacte $T = b_i$, alors que le second cas $c_i = 1$ signifie censurés à droite et nous observons $T > b_i$. Sur l'ensemble des valeurs observées b_i, on note m le nombre de valeur distincte de b_i non censurées. Ces valeurs sont triées dans l'ordre croissant pour construire l'ensemble $\{t_j, j = 1, \dots m\}$. Posons, $a = min_{i\in\{1,\dots,n\}} a_i$ et $b = max_{i\in\{1,\dots,n\}} b_i$.

La fonction de survie non paramétrique $S_N(t|a)$ est une estimation empirique de la probabilité conditionnelle $S(t|a) = P\{T > t | T \geq a\}$ comme aucune information n'est disponible sur la durée de maintien en service des tronçons en service avant a. Aucune information n'est disponible au-delà de l'âge b non plus. Comme illustré par la Figure 64, $S(t)$ est alors définie sur l'intervalle d'âge $[a, b]$ par le vecteur de sauts $\boldsymbol{s} = (s_1, s_2, \dots, s_{m+1})$ avec

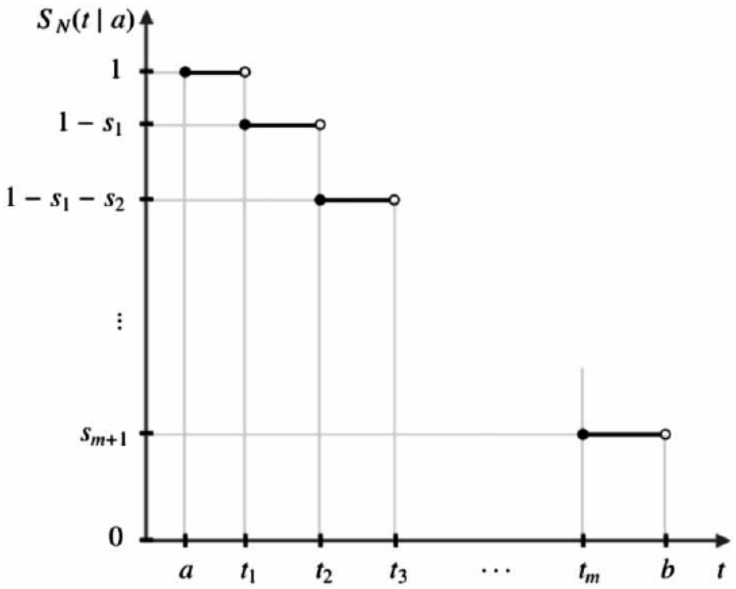


Figure 64 : Fonction de survie non paramétrique

$$\sum_{j\in\{1,\dots,m+1\}} s_j = 1 :$$

$$S_N(t|a) = 1, \text{quand } t \in [a, t_1]$$

$$S_N(t|a) = 1 - s_1, \text{quand } t \in [t_1, t_2]$$

$$S_N(t|a) = 1 - \sum_{k\in\{1,\dots,j\}} s_k \text{ , quand } t \in \left[t_j, t_{j+1}\right]$$

$$S_N(t|a) = s_{m+1}, \text{quand } t \in [t_m, b]$$

L'estimation du vecteur $\boldsymbol{s}$ repose sur les travaux de pivotement de (Turnbull 1976). Cette méthode consiste au calcul de deux matrices de $n \times (m + 1)$ termes:

$$\alpha_{ij} = c_i I\left(t_j > b_i\right) + (1 - c_i) I\left(t_j = b_i\right) \qquad \text{Eq. 56}$$

$$\beta_{ij} = I\left(t_j \geq a_i\right) \qquad \text{Eq. 57}$$

Puis de deux matrices de n × (m + 1) termes:

$$\mu_{ij}(\boldsymbol{s}) = \frac{\alpha_{ij} s_j}{\sum_{k\epsilon\{1,\dots,m+1\}} \alpha_{ik} s_k} \qquad \text{Eq. 58}$$

$$v_{ij}(\boldsymbol{s}) = \frac{\left(1 - \beta_{ij}\right) s_j}{\sum_{k\epsilon\{1,\dots,m+1\}} \beta_{ik} s_k} \qquad \text{Eq. 59}$$

Et finalement d'un vecteur de taille $m + 1$:

$$\pi_j(\boldsymbol{s}) = \sum_{i\epsilon\{1,\dots,n\}} e_i \left(\mu_{ij}(\boldsymbol{s}) + v_{ij}(\boldsymbol{s})\right) \Big/ \left(\sum_{k\epsilon\{1,\dots,m+1\}} \sum_{i\epsilon\{1,\dots,n\}} e_i\left(\mu_{ik}(\boldsymbol{s}) + v_{ik}(\boldsymbol{s})\right)\right) \qquad \text{Eq. 60}$$

Initialisons $\boldsymbol{s}^{(0)}$, tel que $s_j^{(0)} = 1/(m + 1)$ pour tout j. Calculons alors $s_j^{(1)} = \pi_j(\boldsymbol{s}^{(0)})$ pour tout j, et réitérons r fois jusqu'à ce que $\boldsymbol{s}^{(r)} \approx \boldsymbol{s}^{(r-1)}$. Il est prouvé par (Turnbull 1976) qu'une estimation auto-cohérente de $\boldsymbol{s}$ est obtenue (Le Gat et al. 2013).

Annexe 11 : Description de la méthode d'optimisation de Nelder & Mead

La méthode de Nelder-Mead est un algorithme d'optimisation non-linéaire qui a été publié par Nelder et Mead en 1965. C'est une méthode numérique heuristique qui cherche à minimiser une fonction continue dans un espace à plusieurs dimensions. L'algorithme exploite le concept de simplexe qui est un polytope de K+1 sommets dans un espace à K dimensions. Partant initialement d'un tel simplexe, celui-ci subit des transformations simples au cours des itérations : il se déforme, se déplace et se réduit progressivement jusqu'à ce que ses sommets converge sur un point où la fonction est localement minimale. C'est une méthode efficace qui n'utilise pas le gradient de la fonction mais seulement les valeurs de celle-ci.

Nous construisons dans un premier temps un simplexe à k+1 sommets dans un espace à k dimensions, par exemple en 2 dimensions, le simplexe est un triangle, en 3 dimensions, le simplexe est un tétraèdre, etc. Il y a donc k+1 vecteurs $X_1, \dots, X_{k+1}$ de $\mathbb{R}^n, k = n$, affinement indépendants, définissant chacun des k+1 sommets.

Le simplexe se modifie ensuite à travers les opérations de réflexion, d'expansion et de contraction, en changeant le sommet avec la plus grande valeur de la fonction objectif par un autre point de valeur plus petite (Nelder & Mead 1965) (Cf. Figure 65). L'algorithme s'arrête quand la longueur des arêtes du simplexe est inférieure à un seuil fixé.

La Figure 66 montre l'organigramme de la méthode classique.

Le premier sommet du simplexe initial (point initial) peut être fourni par l'utilisateur. Les autres sommets peuvent être définis à partir du premier sommet, en sommant une longueur en chaque direction des n coordonnées :

$$X_i = X_0 + \lambda E_i \qquad i = 1, \dots, n$$

Où X_0 est le point initial, E_i le vecteur de la base canonique dans la ième direction et λ le pas initial d'exploration de l'espace des paramètres (Luersen & Le Riche 2001).

L'avantage déterminant de la méthode de Nelder & Mead est qu'il peut être modifié facilement pour prendre en compte des contraintes dites « base constraints », i.e. quand un paramètre n'est pas recherché sur $\mathbb{R}$ mais sur un intervalle de $\mathbb{R}$.

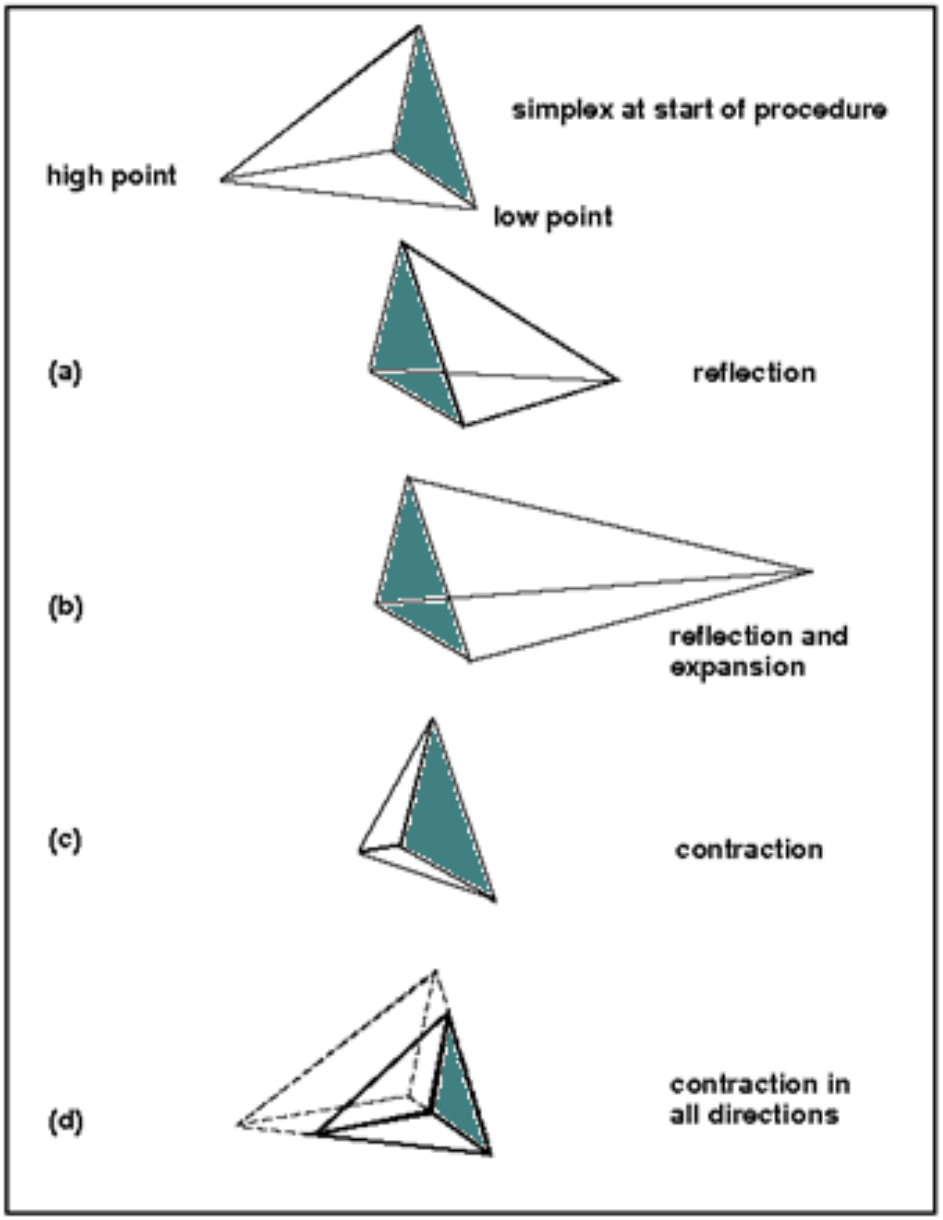


Figure 65 : Evolution possible du simplexe de Nelder & Mead

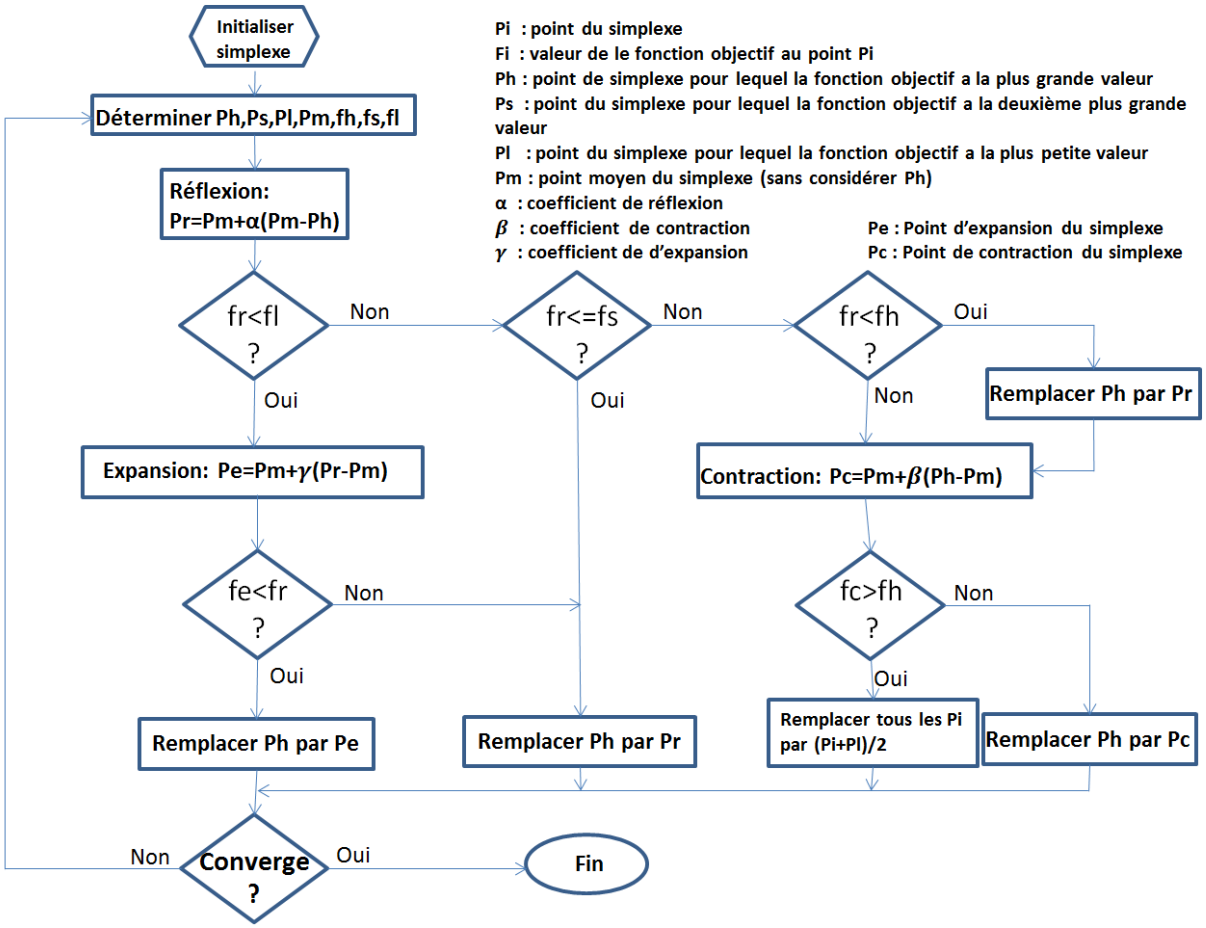


Figure 66 : Organigramme de la méthode de Nelder & Mead classique

Annexe 12 : Validation des méthodes de Kaplan-Meier et de Turnbull, démonstration en effectif

Dans cette partie nous allons démontrer l'efficacité des modèles utilisés (Kaplan-Meier et Turnbull). Pour cela nous supposons que la courbe de survie réelle d'un échantillon est connue. Ensuite, un échantillon de tronçons est créé avec des âges à la mise hors service simulés selon cette survie réelle. Puis nous tronquons à gauche et censurons à droite l'échantillon dans le but de se retrouver dans la même situation que pour nos données du SEDIF. Enfin nous évaluons l'estimateur de Kaplan-Meier et de Turnbull. Ceux-ci doivent alors s'ajuster à la courbe de survie réelle initiale.

Dans un premier temps, nous créons 40 000 âges aléatoires à la mise hors service, dérivant d'une loi uniforme entre 0 et 1. Notons U_i, avec $i = \{1, \dots, 40\ 000\}$ ces aléas uniformes.

Nous construisons ensuite $S_{Réelle}(t)$ une survie de Weibull à deux paramètres (Cf. IV.5.1.1), en fixant arbitrairement les paramètres de forme et d'échelle de cette loi, avec $\delta = 1.5$ et $\lambda = 6.2$.

$$S_{Réelle}(t) = e^{-t^{1.5} \times e^{-6.2}} \quad \text{Eq. 61}$$

Pour générer des âges à la mise hors service distribués selon la survie théorique (Weibull) il est nécessaire d'isoler t en calculant la réciproque de $S^{-1}_{Réelle}(t)$

$$U = S_{Réelle}(t) \quad \text{Eq. 62}$$

$$U = e^{-t^{1.5} \times e^{-6.2}} \quad \text{Eq. 63}$$

$$\ln(U) = -t^{1.5} \times e^{-6.2} \quad \text{Eq. 64}$$

$$\frac{\ln(U)}{e^{-6.2}} = -t^{1.5} \quad \text{Eq. 65}$$

$$\left(-\frac{\ln(U)}{e^{-6.2}}\right)^{\frac{1}{1.5}} = t = S^{-1}_{Réelle}(U) \quad \text{Eq. 66}$$

Calculons alors $S^{-1}_{Réelle}(U)$ pour obtenir 40 000 âges à la mise hors service de tronçons qui suivent la survie réelle initiale.

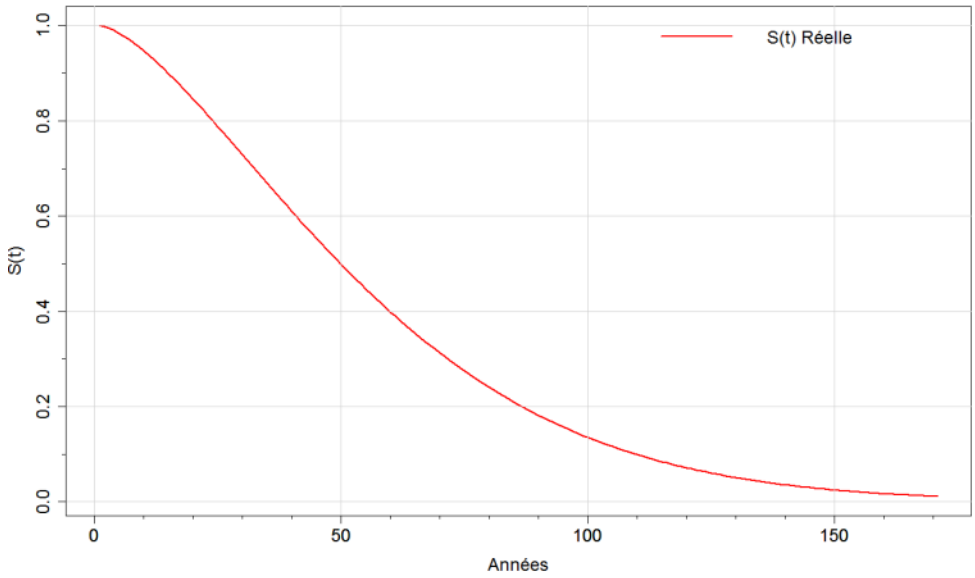


Figure 67 : Courbe de survie réelle

Dans un second temps, afin de reconstituer un échantillon type, nous tirons au sort par tirage aléatoire simple 40 000 dates de pose dans notre base de données « Caractéristiques des tronçons » (Cf. Figure 68). Il est alors possible de construire les dates à la mise hors service de chacun de ces tronçons, en ajoutant aux dates de pose, les âges à la mise hors service calculés précédemment (Cf. Figure 68, la répartition des dates à la mise hors service). Pour se retrouver dans les conditions d'expérience, nous décidons d'observer ces dates à la mise hors service sur une fenêtre qui s'étend de 1995 à 2015. Ainsi, il est possible de reproduire une troncature à gauche et une censure à droite. Nous pouvons alors construire une variable censure, avec c_i qui prend les valeurs 0 si la date de mise hors service est comprise dans la fenêtre d'observation, 1 si la date de mise hors service est supérieure à 2015. Il en résulte que 9 613 tronçons sont totalement tronqués, 22 895 tronçons sont censurés à droite et nous observons l'évènement d'intérêt (=mis hors service) pour 7 492 tronçons.

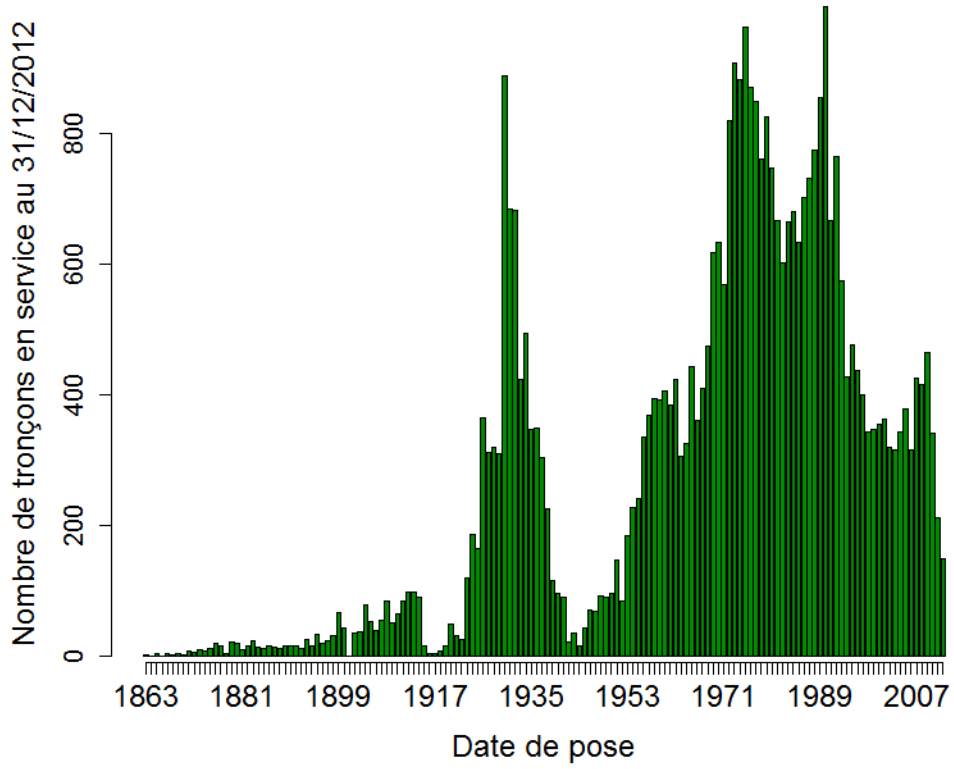


Figure 68 : Date de pose des 40 000 tronçons constituant l'échantillon (en effectif)

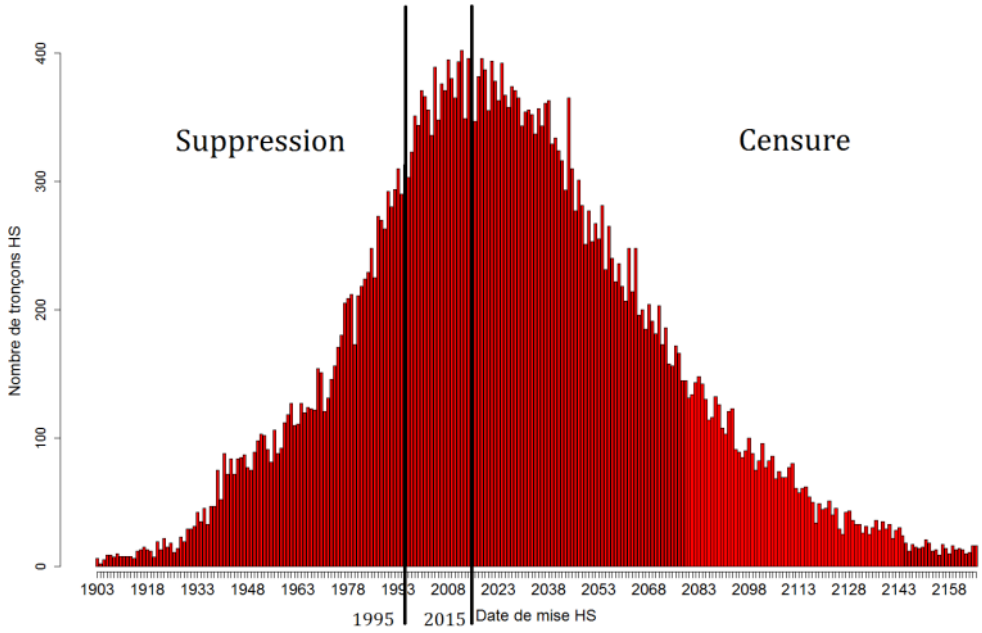


Figure 69 : Répartition des dates à la mise hors service (en effectif) des 40 000 tronçons

Il est maintenant possible d'appliquer les méthodes de Turnbull et de Kaplan-Meier élargi (en effectif) à cet échantillon tronqué, censuré et simulé.

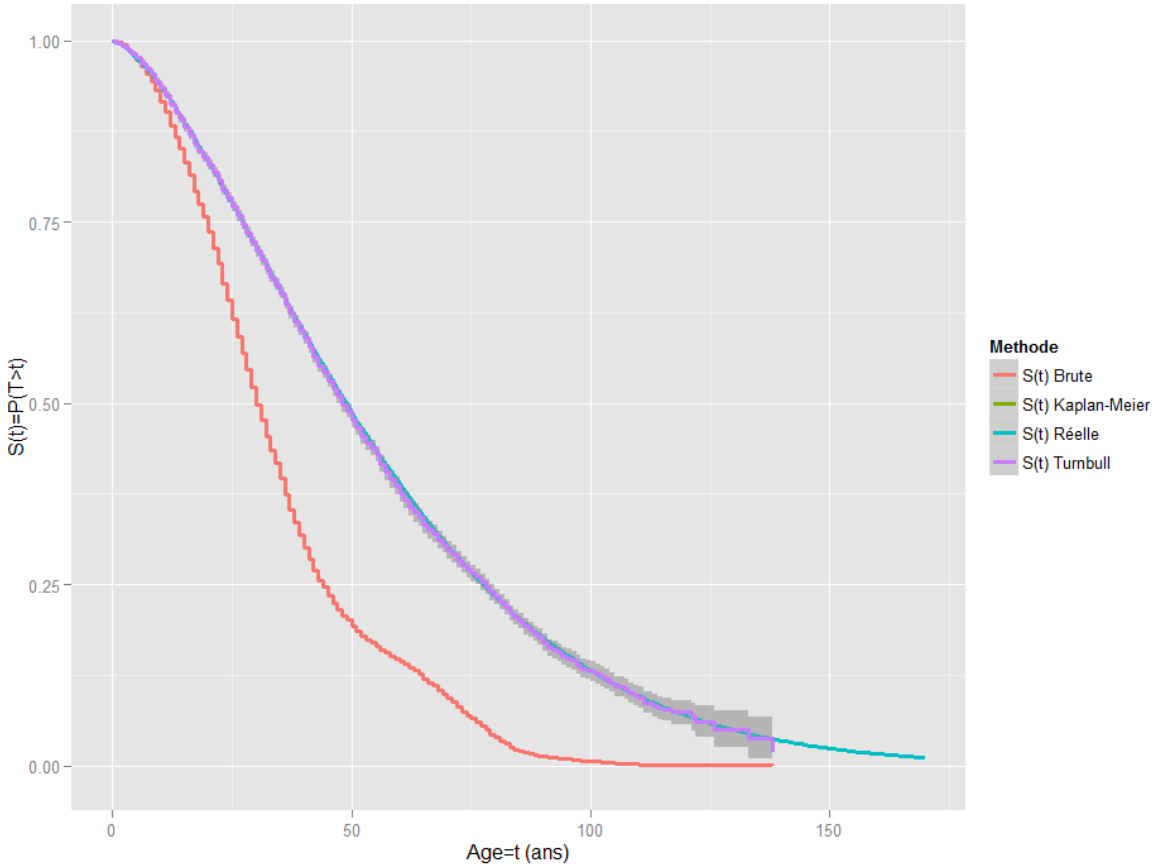


Figure 70 : Courbes de survie estimées (en effectif) et intervalle de confiance

En observant la Figure 70 : Courbes de survie estimées (en effectif) et intervalle de confiance, il est possible d'affirmer que les méthodes de Kaplan-Meier et de Turnbull sont efficaces. En effet, elles se superposent presque à l'identique avec la survie réelle initiale. Il a donc été démontré que ces deux méthodes pouvaient être utilisées quant à l'estimation de courbes de survie pour des données tronquées et censurées. Notons aussi que la survie brute n'est pas fiable car elle ne prend pas en compte la censure, elle n'est donc pas interprétable. De plus, la relation de Turnbull est satisfaite par Kaplan-Meier élargi. En effet, les méthodes de Turnbull et de Kaplan-Meier sont identiques. Pour des questions de temps de calcul il sera préférable d'utiliser la méthode de Kaplan-Meier, en effet l'algorithme de Turnbull tourne pendant environ 30 min alors que la méthode de Kaplan-Meier nous donne des résultats quasi immédiatement (< 5 sec). Pour voir la démonstration en linéaire cf. Annexe 5.

Dans la suite du rapport certains graphiques de courbe de survie de Kaplan-Meier ou de survie brute ne sont pas en escalier, même si nous savons que la représentation correcte est une courbe en escalier nous avons fait le choix dans certains cas de lisser les courbes.

Annexe 13 : Validation des méthodes de Kaplan-Meier et de Turnbull, démonstration en Linéaire

Dans un premier temps on simule 100 000 probabilités à la mise hors service, selon une loi uniforme entre 0 et 1. Notons U_i, avec $i = \{1, \dots, 100\ 000\}$ les probabilités de la mise hors service du tronçon i.

Construisons ensuite $S_{Réelle}(t)$ une survie de Weibull à deux paramètres, en fixant arbitrairement les paramètres de forme et d'échelle de cette loi, avec $\delta = 1.5$ et $\lambda = 6.2$.

Pour générer des âges à la mise hors service qui suivent la survie réelle initiale il est nécessaire d'isoler t en calculant l'inverse de $S_{Réelle}(t)$.

Calculons alors $S^{-1}(U)$ pour obtenir 100 000 âges à la mise hors service de mètres qui suivent la survie réelle initiale.

Dans un second temps, afin de reconstituer un échantillon type, nous tirons au sort par tirage aléatoire simple 100 000 dates de pose dans notre base de données « Caractéristiques des tronçons ». Attention, cette fois-ci une date de pose est associée à un mètre, et non à un tronçon, la répartition des mètres en fonction des dates de pose). Il est alors possible de construire les dates à la mise hors service de chacun de ces mètres, en ajoutant aux dates de pose, les âges à la mise hors service calculés précédemment. Pour retrouver les conditions réelles, nous décidons d'observer ces dates à la mise hors service sur une fenêtre qui s'étend de 1995 à 2015. Ainsi, il est possible de reproduire une troncature à gauche et une censure à droite. Nous pouvons alors construire une variable censure, avec c_i qui prend les valeurs 0 si la date de mise hors service est comprise dans la fenêtre d'observation, 1 si la date de mise hors service est supérieure à 2015.

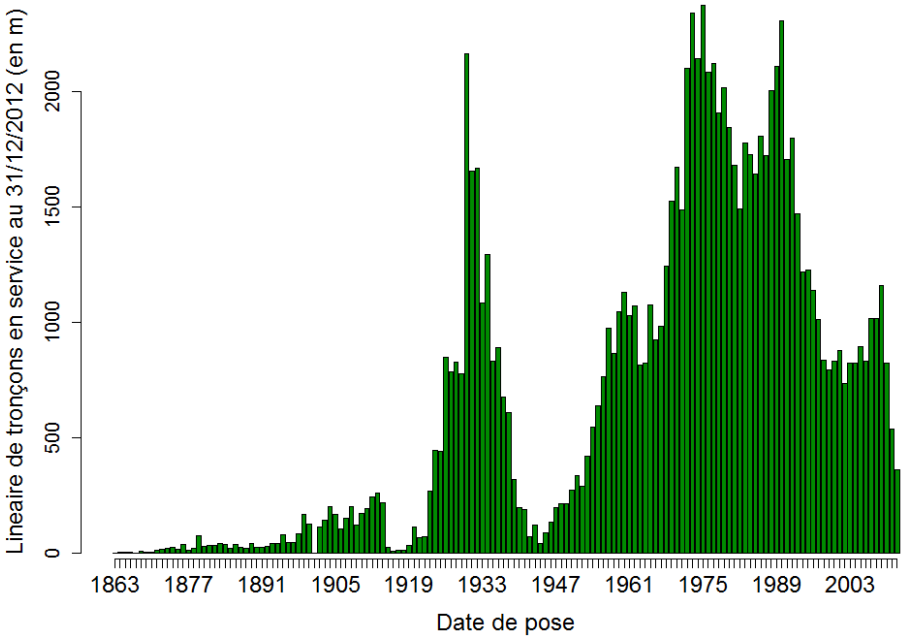


Figure 71 : Date de pose des tronçons constituant l'échantillon (en linéaire)

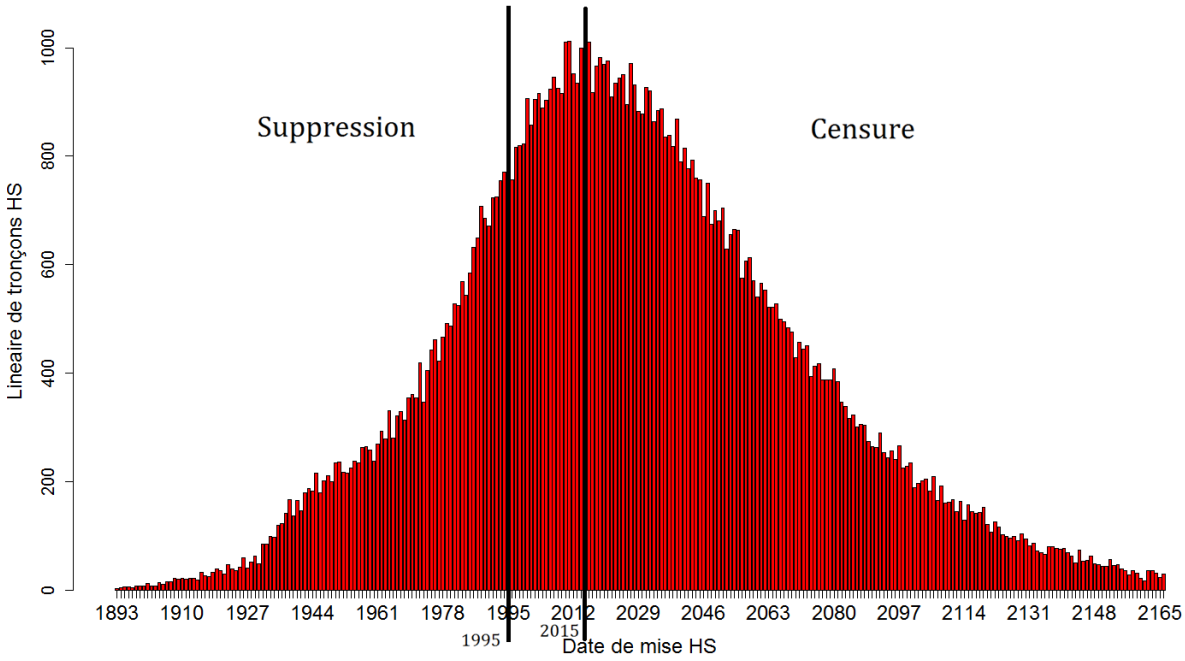


Figure 72 : Répartition des dates à la mise hors service (en linéaire)

Il est maintenant possible d'appliquer les méthodes de Turnbull et de Kaplan-Meier élargi (en linéaire) à cet échantillon tronqué, censuré simulé.

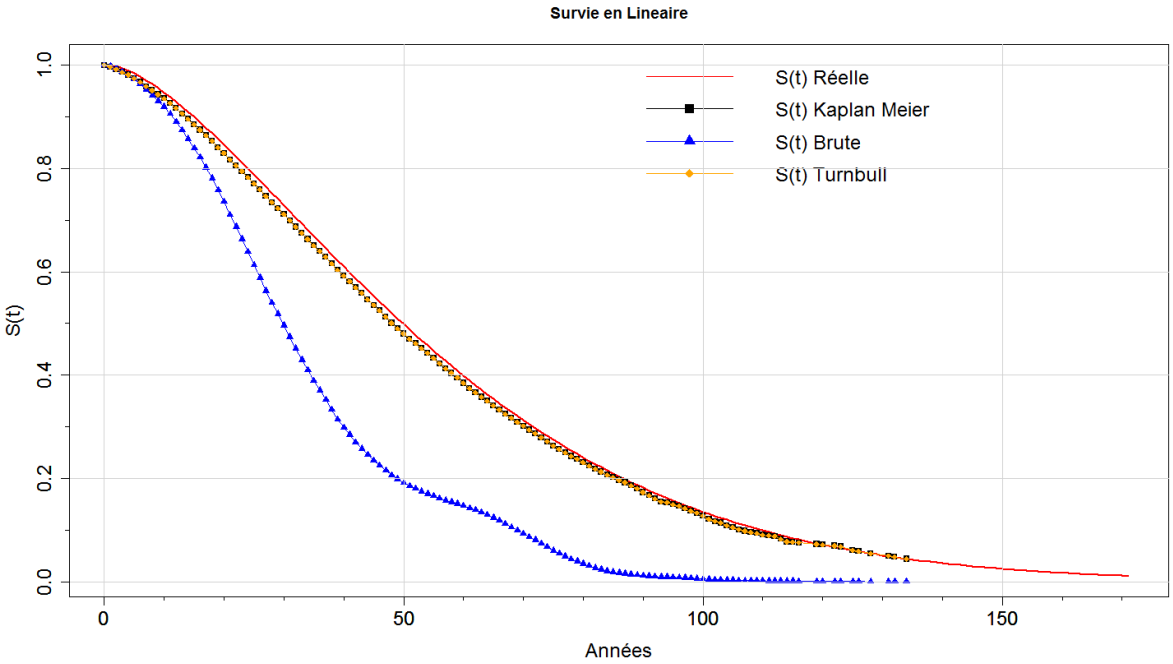


Figure 73 : Courbes de survie estimées (en linéaire)

Annexe 14: Courbe de survie des tronçons partitionnés par matériaux

- **Fonte grise :**

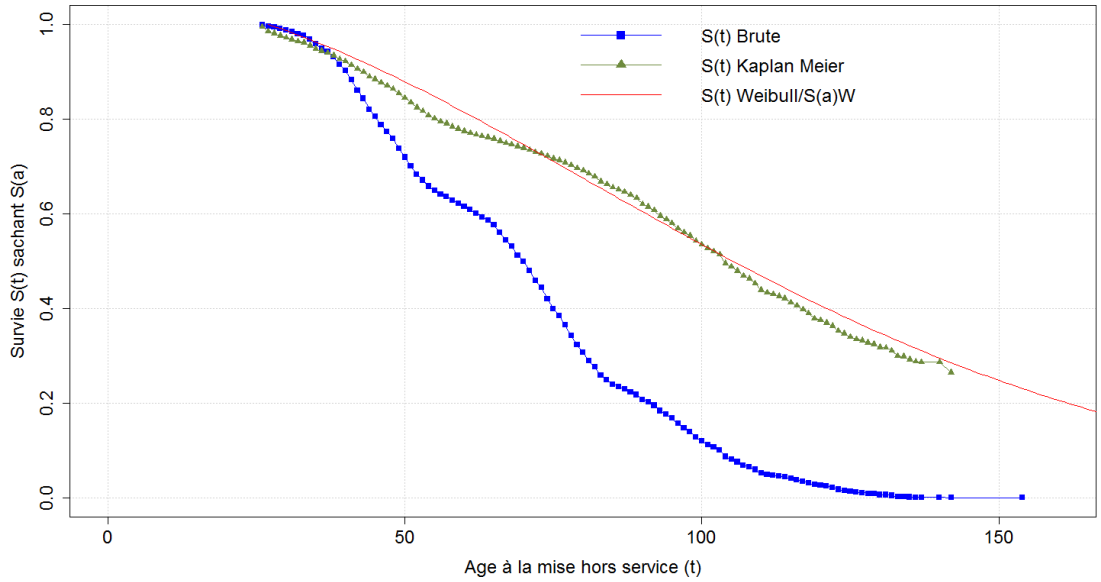


Figure 74 : Courbes de survie en linéaire pour matériau fonte grise conditionné

La survie de Kaplan-Meier de la fonte grise est conditionnée, elle commence en $S_{KM}(26)$ à 1 du fait que le plus jeune tronçon à devenir hors service a 26 ans (phénomène induit par la troncature à gauche).

Par les méthodes des moindres carrées nous ajustons à la courbe de survie empirique (Kaplan-Meier) conditionné [$S_{KM}(t)$=1 à 26 ans] de la fonte grise, une équation analytique de Weibull conditionnée [$S_W(t)$=1 à 26 ans]. Puis nous déconditionnons l'équation de Weibull ce qui donne : $S_w(t) = e^{-t^{1.8} \times e^{-8.8}}$

Ensuite, nous déconditionnons la survie empirique (Kaplan-Meier) en multipliant $S_{KM}(t)$ par $S_w(t) = e^{-26^{1.8} \times e^{-8.8}}$.

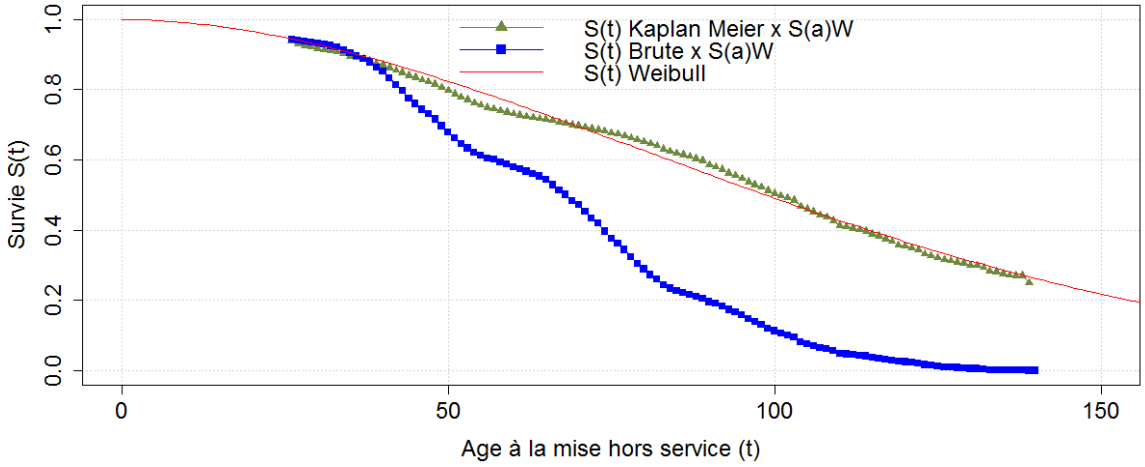


Figure 75 : Courbes de survie en linéaire pour matériau fonte grise déconditionnées

S(t)Weibull :

Label	Estimate	Ref	Chi2	df	Pval
delta	1.8761547e+00	1	3.7434251e+01	1	0.0000000
lambda	8.9810953e+00	10	1.1397618e+01	1	0.0007354

- **Fonte ductile :**

La survie de Herz associée à la courbe de Kaplan Meier pour ce matériau est fausse, en effet elle a été ajustée à la main du faite que ce matériau est trop jeune pour une extrapolation dans le temps. Sa survie ne peut donc pas être utilisée.

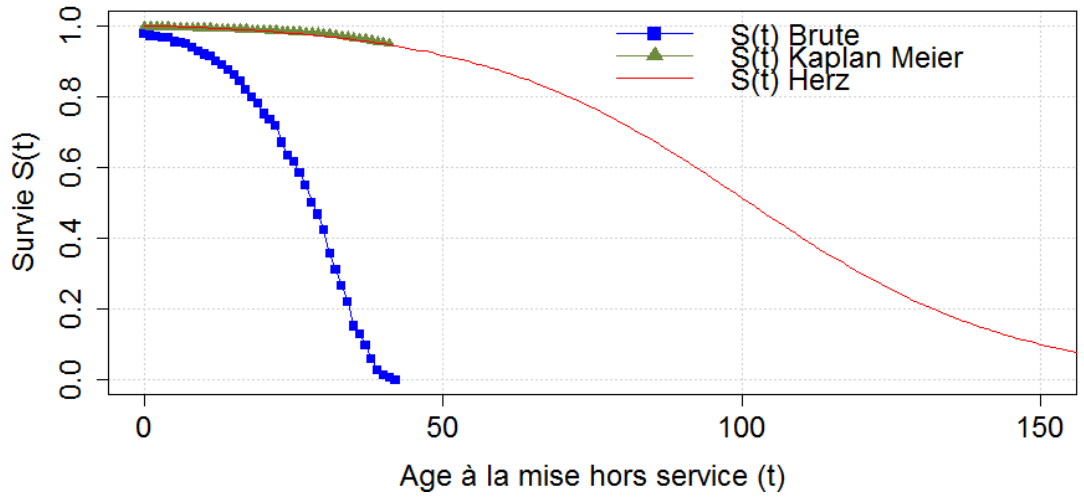


Figure 76 : Courbes de survie en linéaire pour matériau fonte ductile

- **Bonna :**

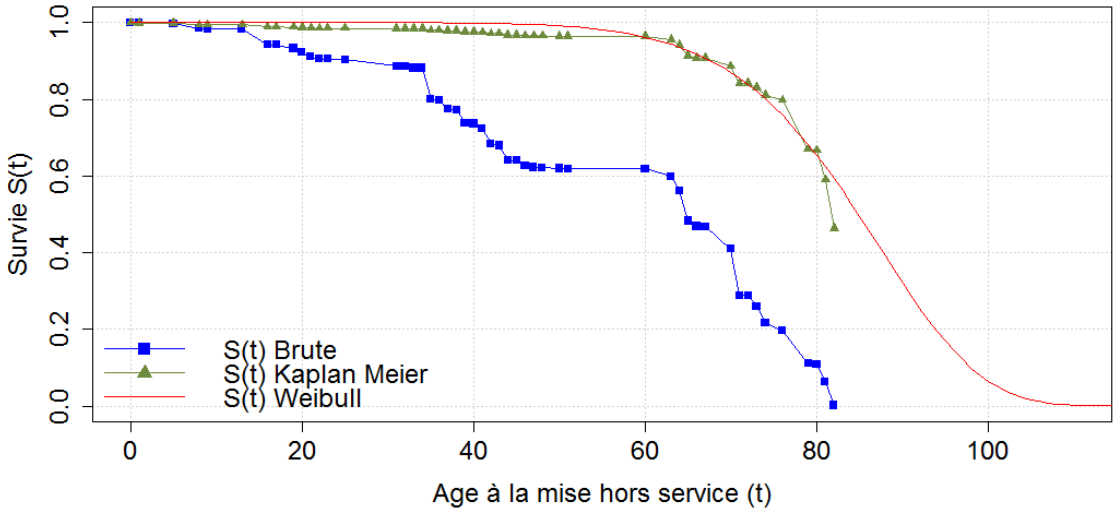


Figure 77 : Courbes de survie en linéaire pour matériau Bonna

S(t)Weibull :

Label	Estimate	Ref	Chi2	df	Pval
delta	8.3352284e+00	9	1.4203595e+01	1	0.0001641
lambda	3.7380279e+01	15	8.0738389e+01	1	0.0000000

- **PEHD :**

Comme pour la fonte ductile, la survie de Herz associée à la courbe de Kaplan Meier pour ce matériau est fausse, en effet elle a été ajustée à la main du faite que ce matériau est trop jeune pour une extrapolation dans le temps. Ca survie ne peut donc pas être utilisé.

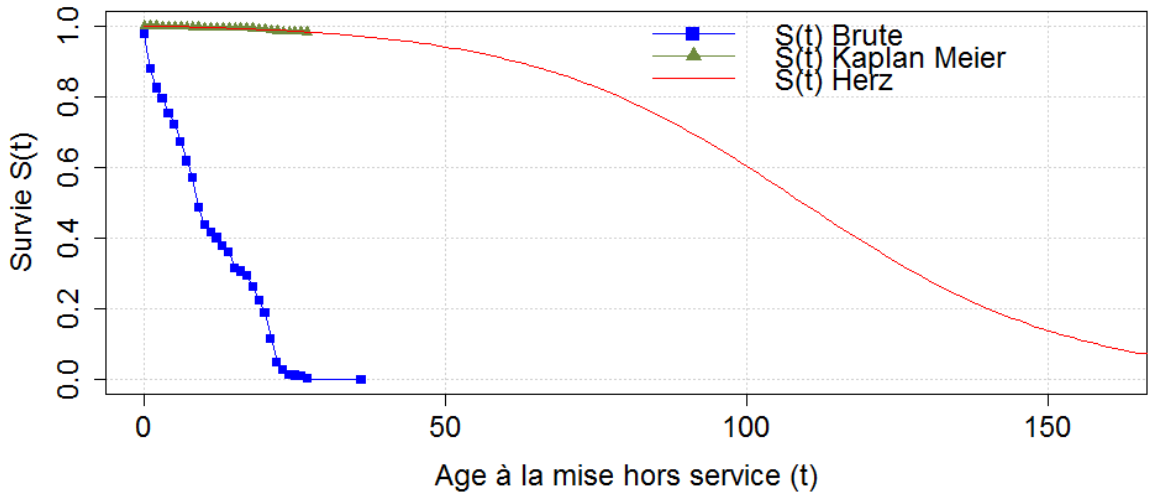


Figure 78 : Courbes de survie en linéaire pour matériau PEHD

- **Acier :**

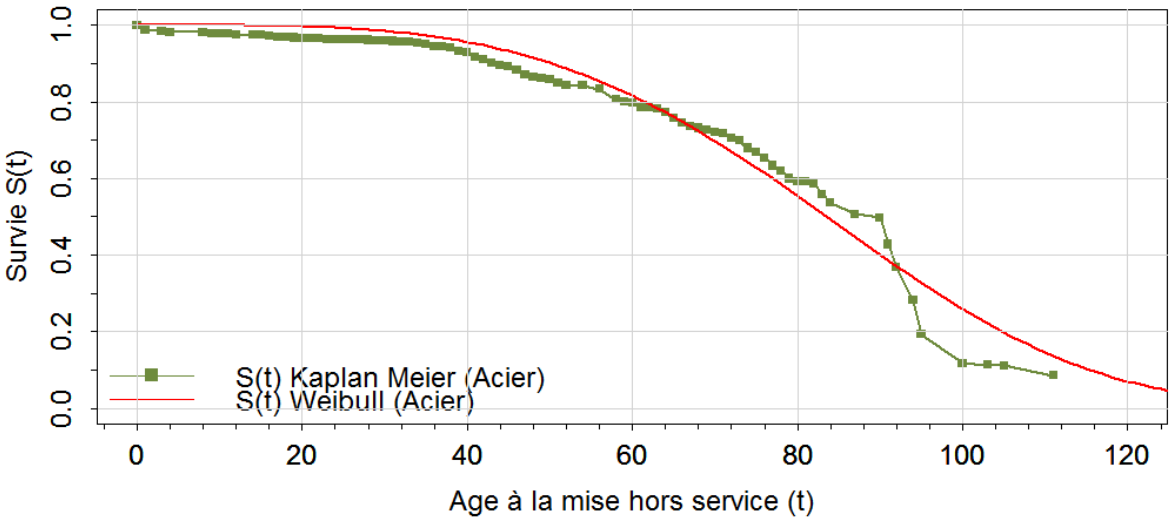


Figure 79 : Courbes de survie en linéaire pour matériau acier

S(t)Weibull :

Label	Estimate	Ref	Chi2	df	Pval
delta	3.7094970e+00	1	1.5125169e+01	1	0.0001006
lambda	1.6781839e+01	10	9.4450911e+01	1	0.0000000

- **Autre :**

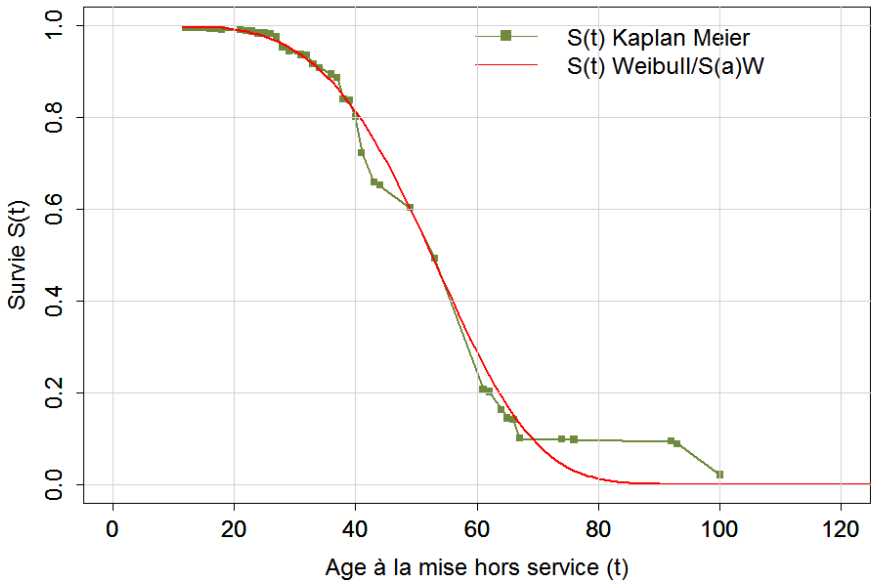


Figure 80 : Courbes de survie en linéaire pour matériau autre conditionné

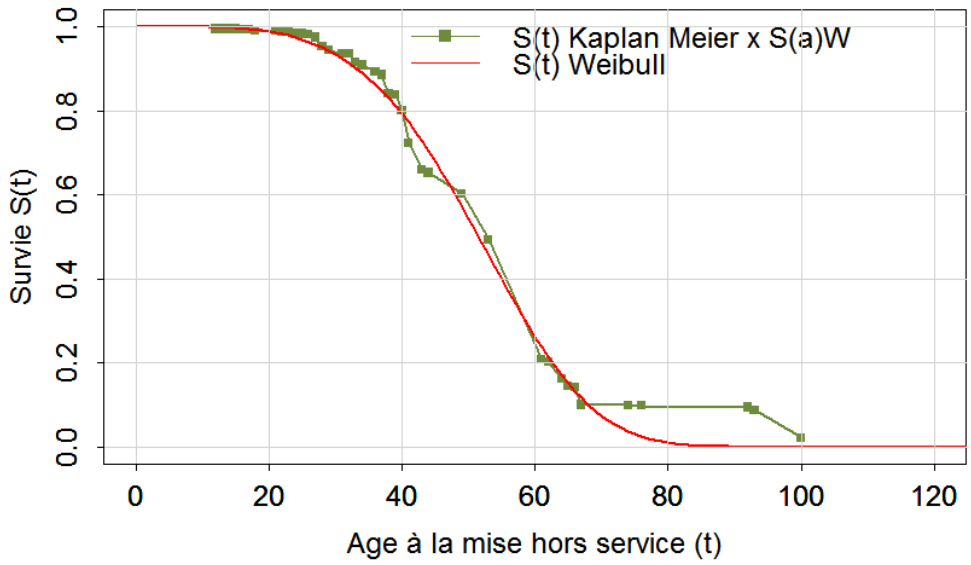


Figure 81 : Courbes de survie en linéaire pour matériau autre déconditionné

Pour déconditionner on applique la même méthode que pour la fonte grise, avec t_{min} = 12 ans.

S(t)Weibull :

Label	Estimate	Ref	Chi2	df	Pval
delta	4.3116494e+00	1	1.6718561e+01	1	0.0000434
lambda	1.7361603e+01	10	4.5306379e+01	1	0.0000000

- **Synthèse :**

Quartile	Tous matériaux	Fonte grise	Fonte ductile	PEHD	Bonna	Acier + Tole acier	Autre
S(t) = 0.75	74 ans	52 ans	77 ans	84 ans	76 ans	66 ans	41 ans
S(t) = 0.5	105 ans	95 ans	101 ans	109 ans	82 ans	90 ans	49 ans
S(t) = 0.25	142 ans	134 ans	125 ans	133 ans	92 ans	95 ans	57 ans

Tableau 7 : Synthèse de la distribution des âges à la mise hors service en fonction du matériau

Remarque : Pour la répartition du linéaire en fonction des matériaux Cf. Figure 10.

Annexe 15 : Le modèle de Weibull à covariables est un modèle de Cox particulier

On a :

$$S_\Theta(t) = e^{-t^{\delta_1} \times e^{Z^T \beta}}$$

Or :

$$h(t) = \frac{f(t)}{S(t)}$$

et :

$$f_\Theta(t) = \frac{-dS_\Theta(t)}{dt}$$

Car :

$$F_\Theta(t) = 1 - S_\Theta(t)$$

Donc nous obtenons :

$$h(t) = \frac{\delta_1 t^{\delta_1 - 1} \times e^{Z^T \beta} \times e^{-t^{\delta_1} \times e^{Z^T \beta}}}{e^{-t^{\delta_1} \times e^{Z^T \beta}}}$$

$$h(t) = \delta_1 t^{\delta_1 - 1} \times e^{Z^T \beta}$$

Avec $Z^T \beta = \beta_0 + \beta_1 ø + \beta_2 Lg + \beta_3 FG + \beta_4 FD + \beta_5 PEHD + \beta_6 Bonna + \beta_7 Autre + \beta_8 NDOA$

Et nous retrouvons la même propriété que celle du modèle de Cox:

$$\lambda(t, z) = \lambda_0(t) \times \exp\left(Z^T \beta\right)$$

Annexe 16: Corrélation entre longueur et diamètre

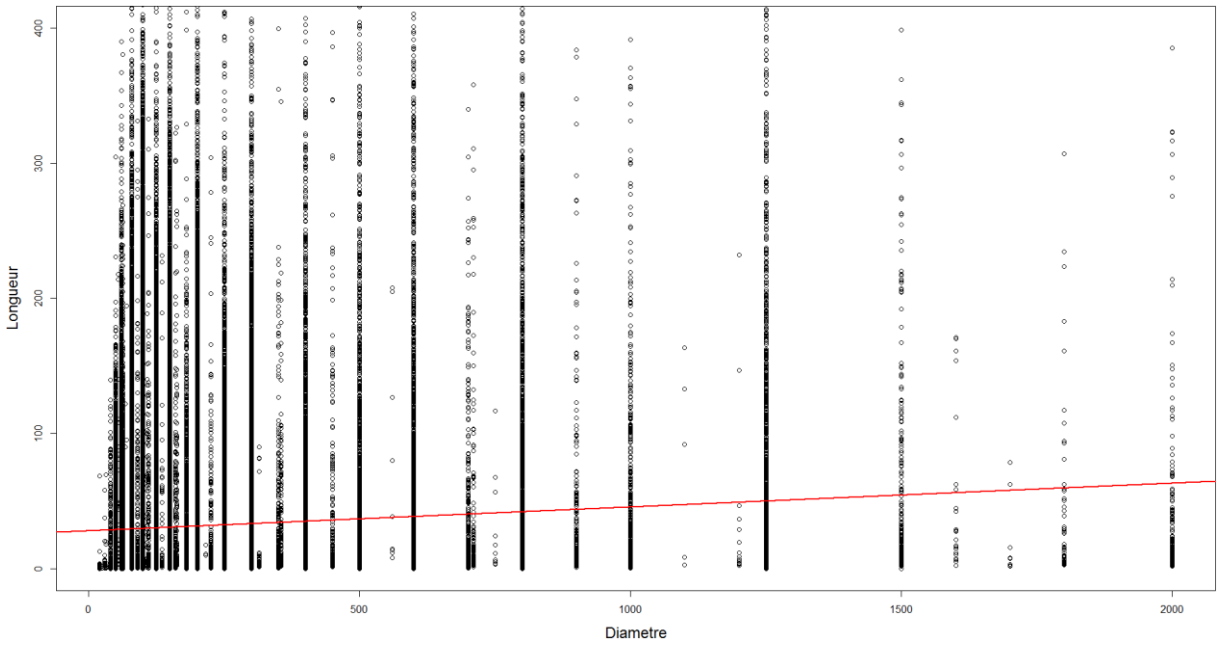


Figure 82 : Corrélation entre longueur et diamètre en nombre de tronçon (échelle réduite)

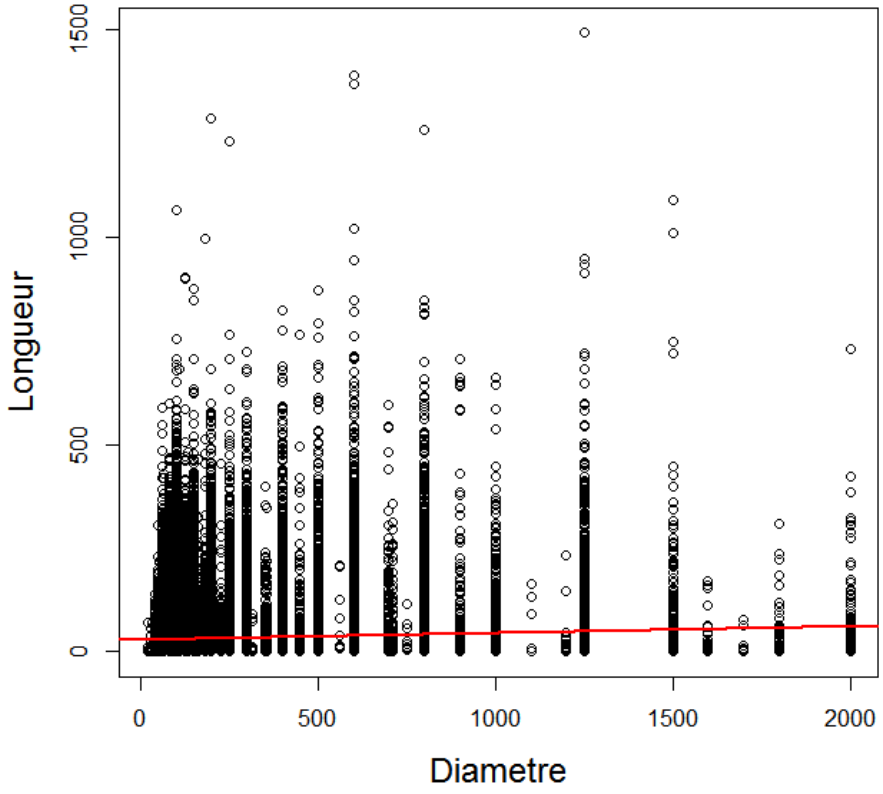


Figure 83 : Corrélation entre longueur et diamètre en nombre de tronçon

Annexe 17: Résultats Weibull avec covariables de la survie de Weibull avec covariables

Equation B :

$$S(t) = e^{-t^{\delta_1} \times e^{\beta_0 + \beta_2 Diam + \beta_3 Long}}$$

Coefficient	Estimation	Initialisation de Nelder & Mead	P-value
$\widehat{\delta_1}$	2.032	1	< 2e-16
$\widehat{\beta_0}$	-9.936	0	< 2e-16
$\widehat{\beta_2}$	-0.000886	0	< 2e-16
$\widehat{\beta_3}$	0.000924	0	< 2e-16

Courbe associé :

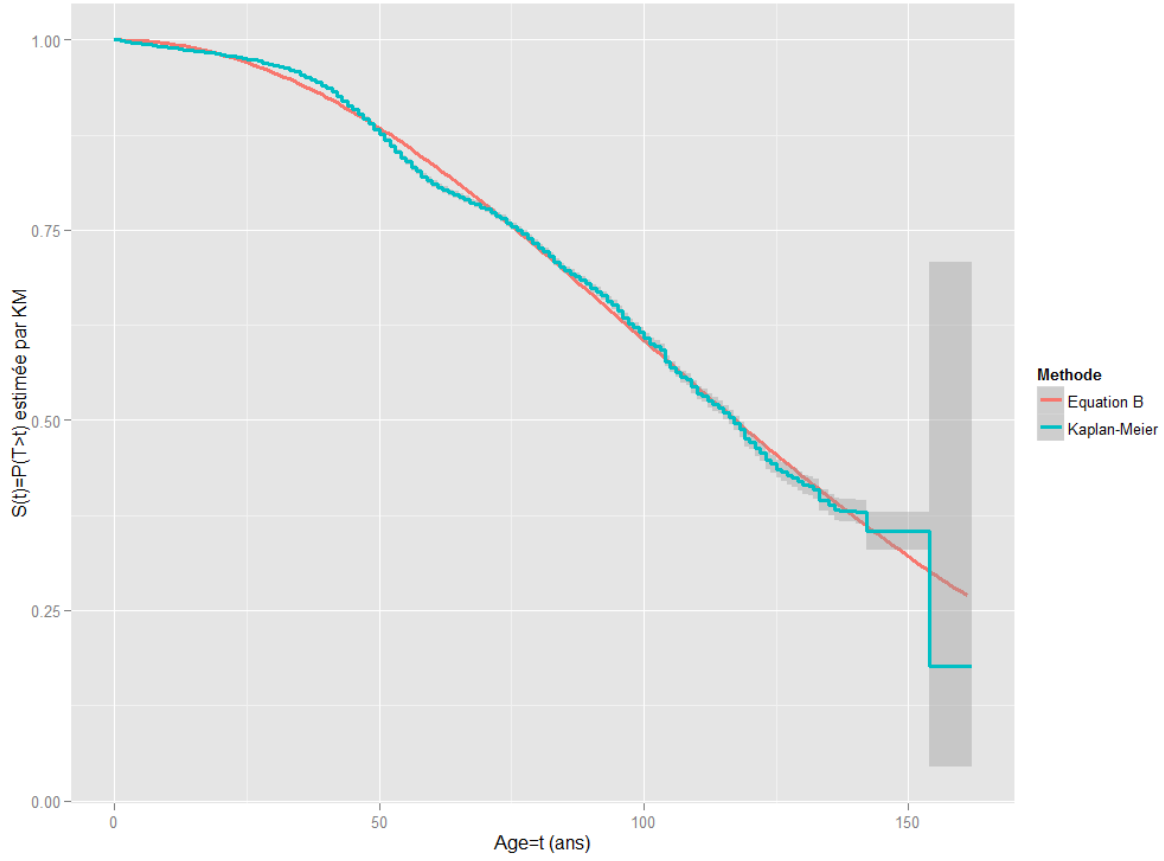


Figure 84 : Courbes de survie de Kaplan-Meier tous tronçons confondus et équation B

Annexe 18: Indice *TP10A*, canalisation, égouts, assainissement et adduction d'eau avec fourniture de tuyaux

Structure de l'indice TP10a :

Canalisations

Code-Définition	Salaires et charges	Matériaux	Matériel	Transport	Energie	Frais Divers
TP10a-Canalisations, égouts, assainissement et adduction d'eau avec fournitures de tuyaux (2)	41%	34%	15%	4%	3%	3%
TP10bis-Canalisations sans fournitures	51%	18%	18%	4%	4%	5%
TP11-Canalisations grande distance et irrigation avec fourniture de tuyaux (2)	22%	61%	8%	2%	2%	5%

Tableau 8 : Structure de l'indice TP 10a

(Ministère_de_L'écologie 2010)

Pour construire cet indice, le ministère de l'écologie recueille chaque moi des coûts sur le terrain tel que le coût des salaires, des matériaux, des matériels, etc. Ainsi il peut alors construire des taux d'augmentations ou de diminution d'un mois sur l'autre. Nous observons ci-dessous que le l'indice TP10a a une tendance à l'augmentation.

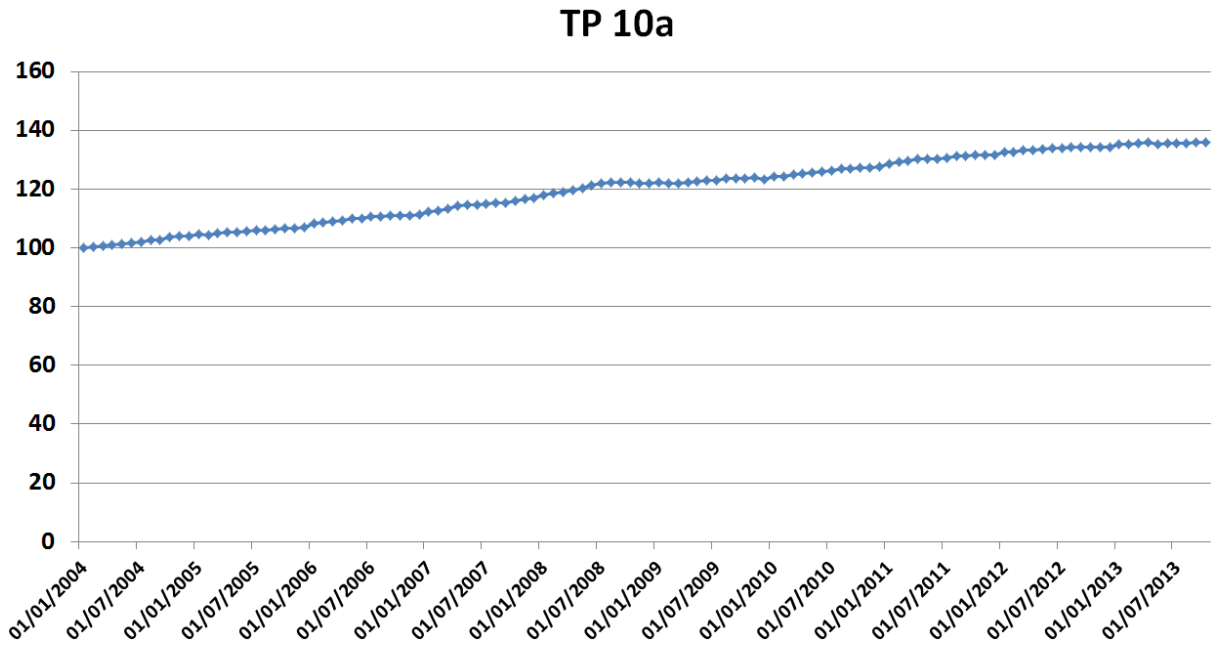


Figure 85 : Evolution de l'indice TP 10a entre janvier 2004 et septembre 2013

L'indice moyen correspond au coefficient de la pente soit : 3.7% entre janvier 2004 et septembre 2013

Date de cotation	Date de mise en ligne	Valeur	Date de cotation	Date de mise en ligne	Valeur	Date de cotation	Date de mise en ligne	Valeur
01/01/2004	29/04/2004	100	01/04/2007	29/07/2007	114,2	01/07/2010	29/10/2010	126,3
01/02/2004	28/05/2004	100,2	01/05/2007	05/09/2007	114,6	01/08/2010	28/11/2010	126,8
01/03/2004	26/06/2004	100,6	01/06/2007	29/09/2007	114,7	01/09/2010	30/12/2010	126,9
01/04/2004	30/07/2004	101	01/07/2007	28/10/2007	114,9	01/10/2010	27/01/2011	127,3
01/05/2004	03/09/2004	101,4	01/08/2007	30/11/2007	115,2	01/11/2010	25/02/2011	127,1
01/06/2004	02/10/2004	101,8	01/09/2007	29/12/2007	115,3	01/12/2010	30/03/2011	127,5
01/07/2004	30/10/2004	102	01/10/2007	31/01/2008	115,8	01/01/2011	29/04/2011	128,6
01/08/2004	27/11/2004	102,6	01/11/2007	29/02/2008	116,6	01/02/2011	29/05/2011	129,2
01/09/2004	26/12/2004	102,8	01/12/2007	30/03/2008	116,8	01/03/2011	30/06/2011	129,7
01/10/2004	30/01/2005	103,6	01/01/2008	02/05/2008	118,1	01/04/2011	30/07/2011	130,1
01/11/2004	03/03/2005	103,9	01/02/2008	30/05/2008	118,7	01/05/2011	31/08/2011	130,3
01/12/2004	27/03/2005	104	01/03/2008	04/07/2008	119,1	01/06/2011	30/09/2011	130,2
01/01/2005	29/04/2005	104,6	01/04/2008	03/08/2008	119,7	01/07/2011	30/10/2011	130,7
01/02/2005	31/05/2005	104,4	01/05/2008	03/09/2008	120,4	01/08/2011	30/11/2011	131,3
01/03/2005	30/06/2005	105	01/06/2008	30/09/2008	121,3	01/09/2011	30/12/2011	131,2
01/04/2005	28/07/2005	105,2	01/07/2008	31/10/2008	122,1	01/10/2011	31/01/2012	131,5
01/05/2005	01/09/2005	105,4	01/08/2008	30/11/2008	122,2	01/11/2011	28/02/2012	131,6
01/06/2005	01/10/2005	105,6	01/09/2008	20/12/2008	122,4	01/12/2011	31/03/2012	131,5
01/07/2005	03/11/2005	105,9	01/10/2008	05/02/2009	122,4	01/01/2012	03/05/2012	132,7
01/08/2005	30/11/2005	105,9	01/11/2008	28/02/2009	122	01/02/2012	31/05/2012	132,7
01/09/2005	27/12/2005	106,2	01/12/2008	29/03/2009	121,9	01/03/2012	29/06/2012	133,3
01/10/2005	01/02/2006	106,7	01/01/2009	30/04/2009	122,3	01/04/2012	31/07/2012	133,4
01/11/2005	01/03/2006	106,8	01/02/2009	31/05/2009	122,1	01/05/2012	01/09/2012	133,6
01/12/2005	30/03/2006	106,9	01/03/2009	28/06/2009	122	01/06/2012	02/10/2012	134
01/01/2006	28/04/2006	108,3	01/04/2009	30/07/2009	122,2	01/07/2012	31/10/2012	133,9
01/02/2006	28/05/2006	108,6	01/05/2009	28/08/2009	122,6	01/08/2012	30/11/2012	134,3
01/03/2006	01/07/2006	109	01/06/2009	01/10/2009	123	01/09/2012	28/12/2012	134,3
01/04/2006	29/07/2006	109,4	01/07/2009	30/10/2009	122,9	01/10/2012	30/01/2013	134,3
01/05/2006	31/08/2006	110	01/08/2009	29/11/2009	123,6	01/11/2012	27/02/2013	134,4
01/06/2006	04/10/2006	110	01/09/2009	27/12/2009	123,5	01/12/2012	28/03/2013	134,4
01/07/2006	28/10/2006	110,5	01/10/2009	29/01/2010	123,7	01/01/2013	08/05/2013	135,4
01/08/2006	30/11/2006	110,8	01/11/2009	26/02/2010	123,9	01/02/2013	31/05/2013	135,1
01/09/2006	28/12/2006	111	01/12/2009	28/03/2010	123,4	01/03/2013	29/06/2013	135,7
01/10/2006	31/01/2007	111	01/01/2010	30/04/2010	124,3	01/04/2013	30/07/2013	136
01/11/2006	28/02/2007	111,1	01/02/2010	28/05/2010	124,4	01/05/2013	31/08/2013	135,1
01/12/2006	31/03/2007	111,2	01/03/2010	27/06/2010	125	01/06/2013	29/09/2013	135,5
01/01/2007	29/04/2007	112,4	01/04/2010	29/07/2010	125,3	01/07/2013	31/10/2013	135,6
01/02/2007	01/06/2007	112,7	01/05/2010	31/08/2010	125,7	01/08/2013	30/11/2013	135,6
01/03/2007	30/06/2007	113,2	01/06/2010	29/09/2010	125,9	01/09/2013	31/12/2013	135,9

Tableau 9 : Evolution de l'indice TP 10a entre janvier 2004 et septembre 2013

Annexe 19: Courbe de survie avec d'autres stratifications

Niveau piézométrique : niveau du trop-plein du réservoir auquel est relié le tronçon selon le NGF (nivellement général de la France en mètre)

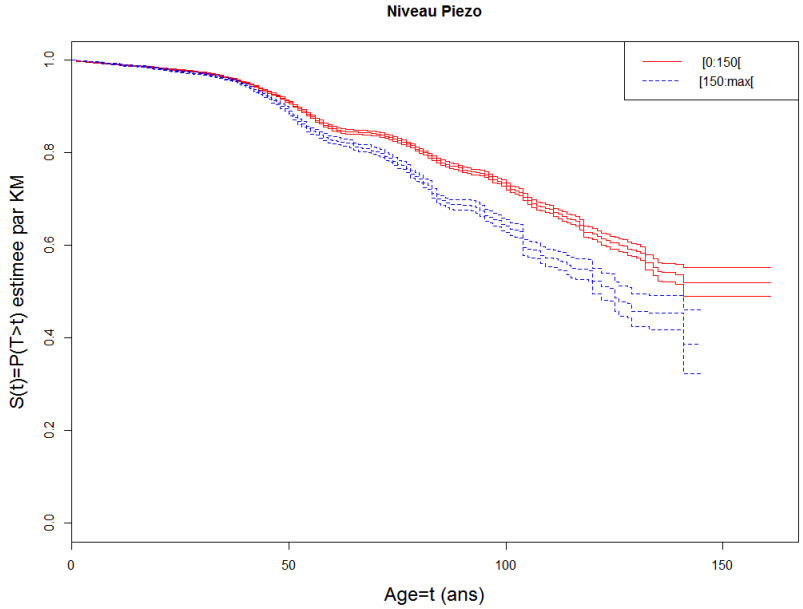


Figure 86: Courbes de survie en effectif tronçons (tronçons partitionnés en 2 sous-ensembles selon le niveau piézométrique : [0 :150m [et [150m : max]) et leurs intervalles de confiance

Le niveau piézométrique a un impact sur l'âge à la mise hors service des tronçons. En effet le SEDIF a renouvelé plus tôt les tronçons qui ont un niveau piézométrique élevé (Cf. Figure 86).

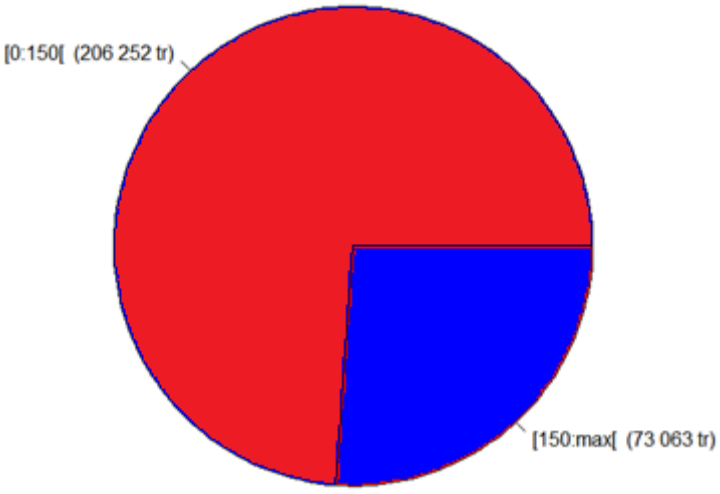


Figure 87: Diagramme de la stratification des tronçons selon le niveau piézométrique (nombre de tronçons)

Niveau d'enfouissement, différence entre l'altitude du sol et l'altitude du tronçon (en mètre)

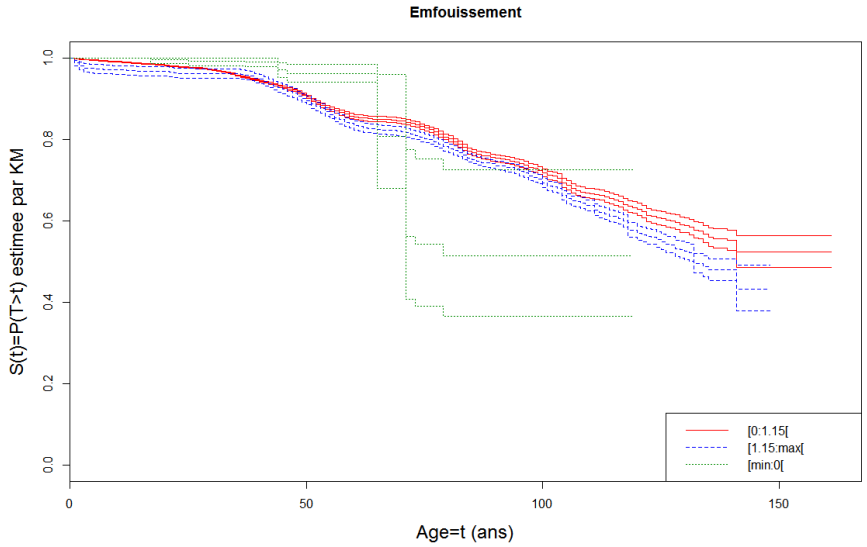


Figure 88: Courbes de survie en effectif tronçons (tronçons partitionnés en 3 sous-ensembles selon le niveau d'enfouissement : [min : 0[, [0 :1.15m [et [1.15m : max]) et leurs intervalles de confiance

L'enfouissement a un impact sur l'âge à la mise hors service des tronçons. En effet les tronçons qui ont un enfouissement élevé (Cf. Figure 88) ont été renouvelés un peu plus tôt.

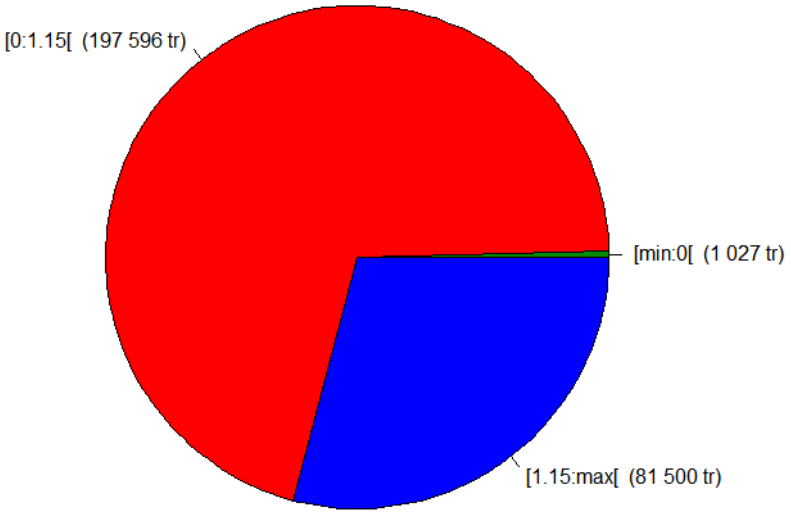


Figure 89: Diagramme de la stratification des tronçons selon le niveau d'enfouissement (nombre de tronçons)

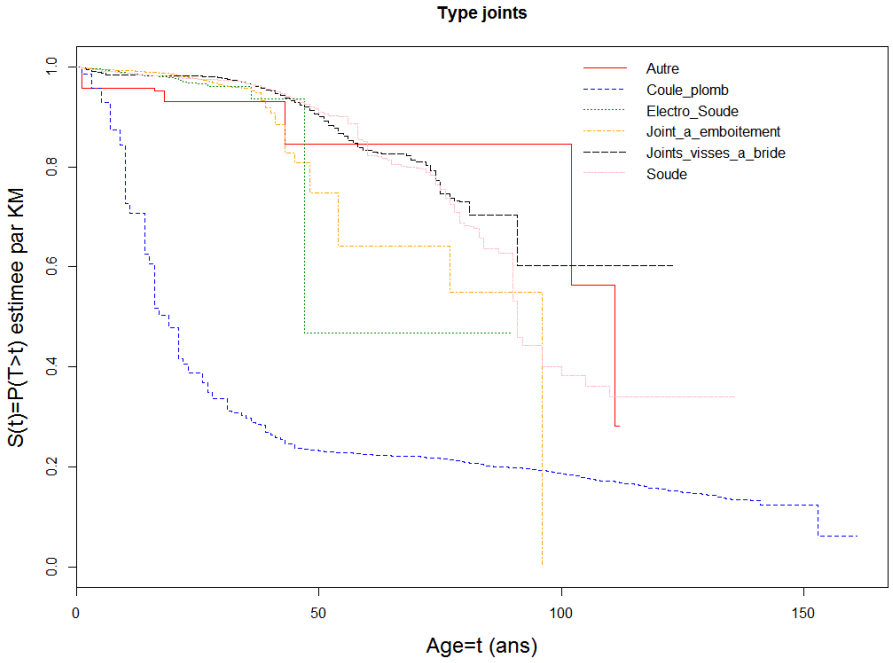


Figure 90: Courbes de survie en effectif tronçons (tronçons partitionnés en 6 sous-ensembles : joints soudés, joints visses à bride, joints à emboitement, joints électro-soudés, joints coulés à plomb et joints autre)

Le type de joints a un impact sur l'âge à la mise hors service des tronçons. En effet le SEDIF a renouvelé beaucoup plus les tronçons avec des joints coulés au plomb que ceux soudés (Cf. Figure 90 pour voir différent type de joints).

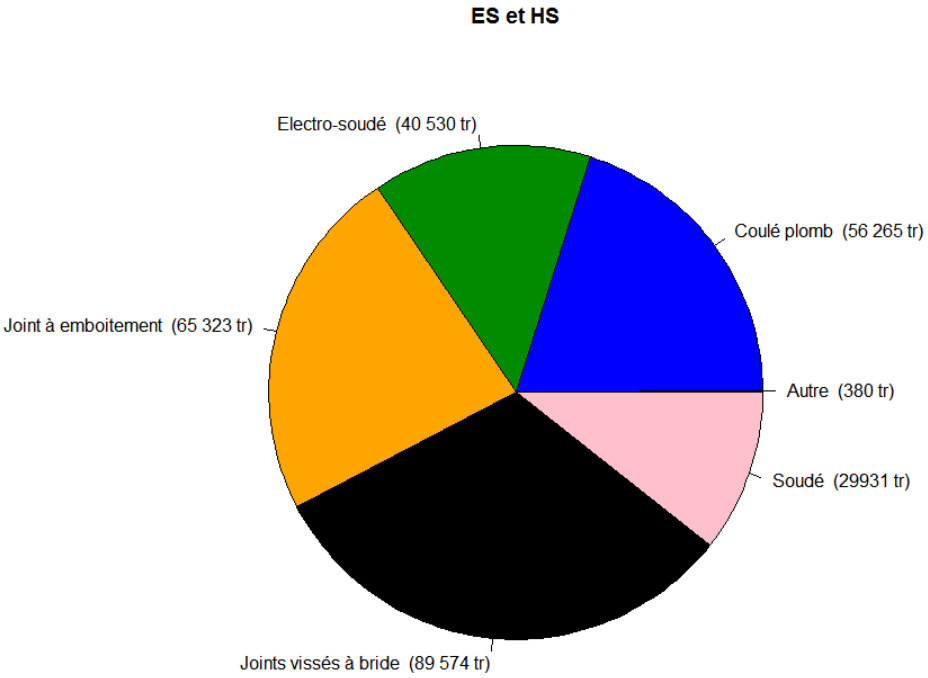


Figure 91: Diagramme de la stratification des tronçons selon le type de joints (nombre de tronçons)

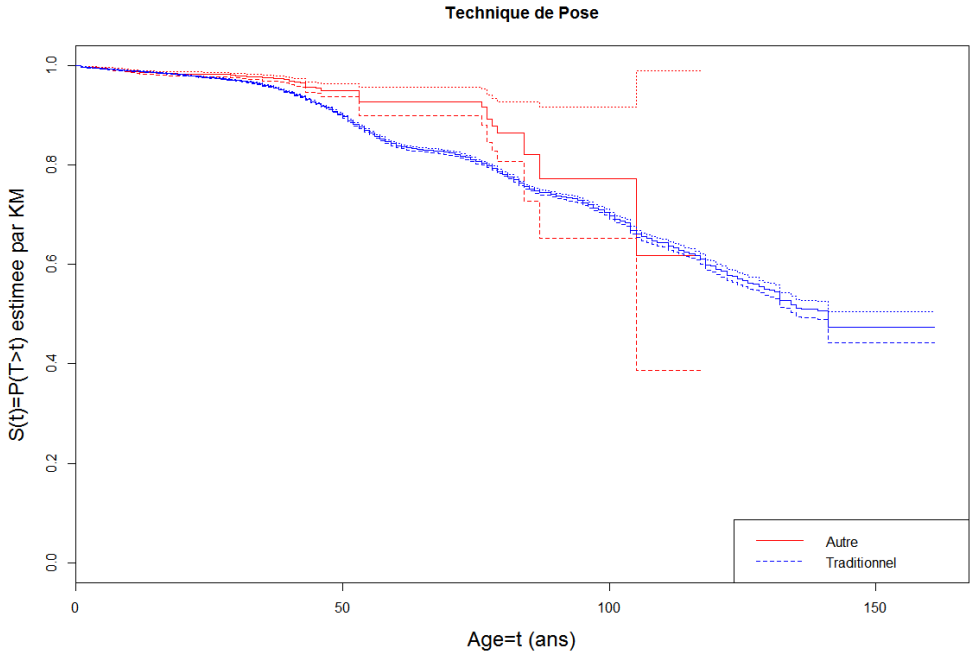


Figure 92: Courbes de survie en effectif tronçons (tronçons partitionnés en 2 sous-ensembles : technique de pose traditionnelle et technique de pose autre) et leurs intervalles de confiance

La technique de pose a un impact sur l'âge à la mise hors service des tronçons. En effet les tronçons qui ont été posé avec une technique de pose traditionnelle ont été renouvelés plus tard que ceux posés avec une technique de pose autre (Cf. Figure 92).

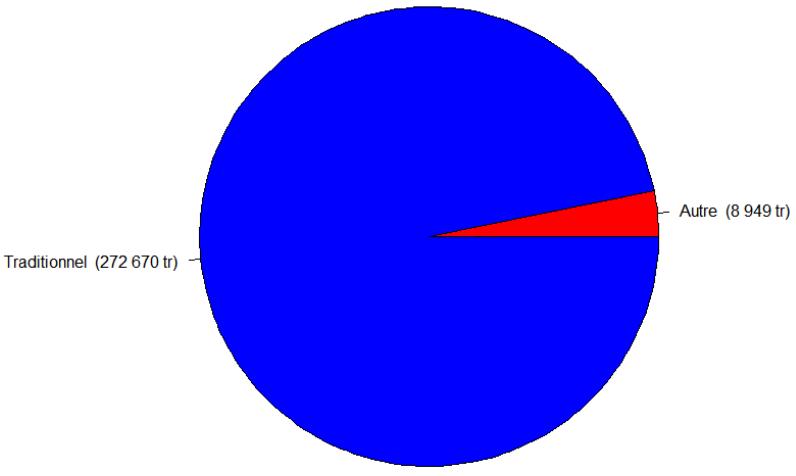


Figure 93: Diagramme de la stratification des tronçons selon la technique de pose traditionnel ou non (nombre de tronçons)

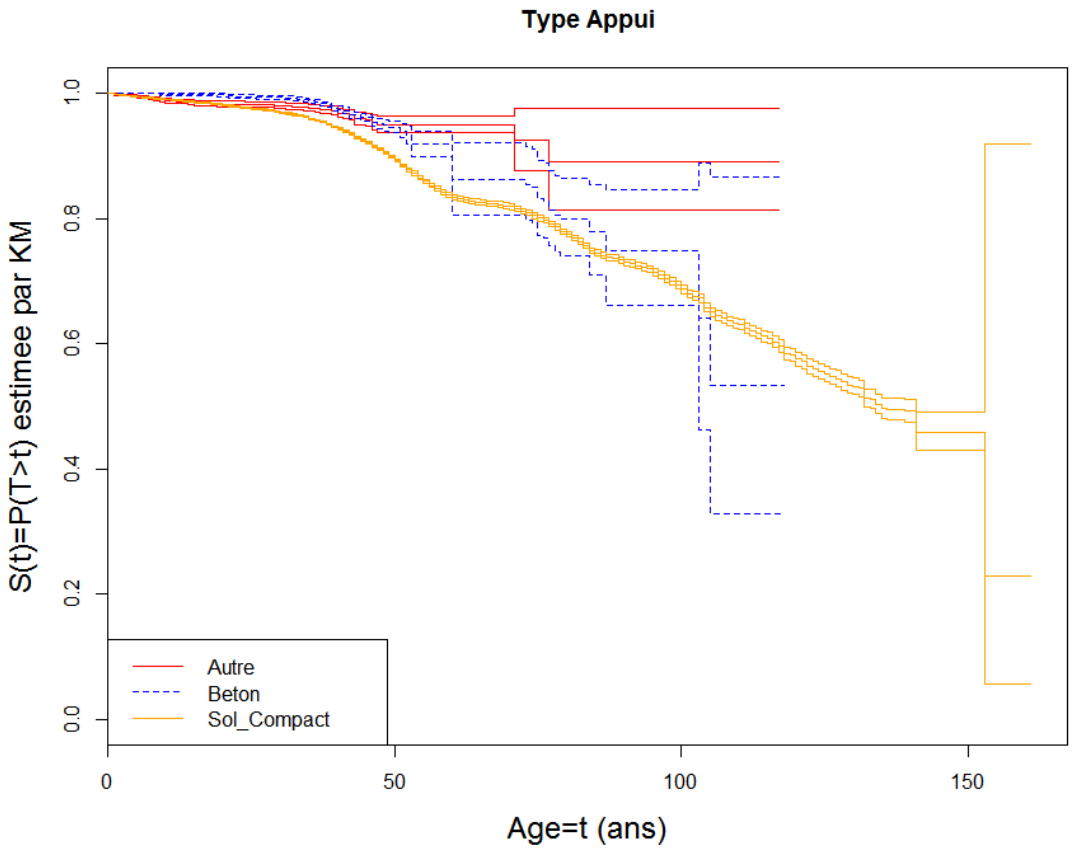


Figure 94: Courbes de survie en effectif tronçons (tronçons partitionnés en 3 sous-ensembles : appui béton, appui sol compacte et appui autre) et leurs intervalles de confiance

Le type d'appui a eu un impact sur l'âge à la mise hors service des tronçons. En effet le SEDIF a renouvelé plus les tronçons qui ont été posé sur un sol compact que les autres, cependant les intervalles de confiance se chevauchent. Cette différence est donc peu significative (Cf. Figure 94).

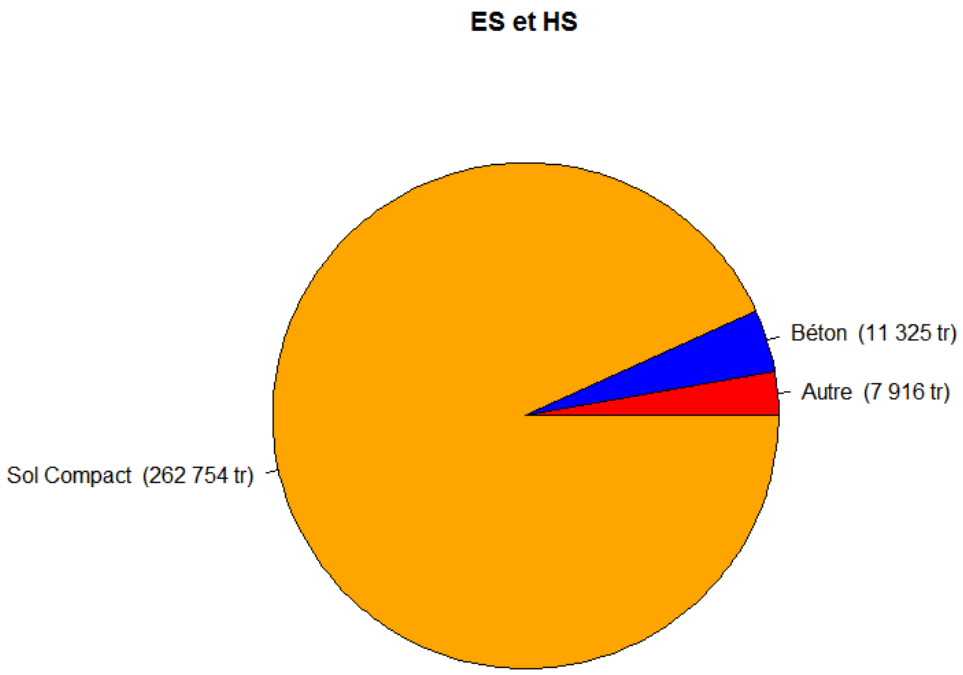


Figure 95: Diagramme de la stratification des tronçons selon le type d'appui (nombre de tronçons)

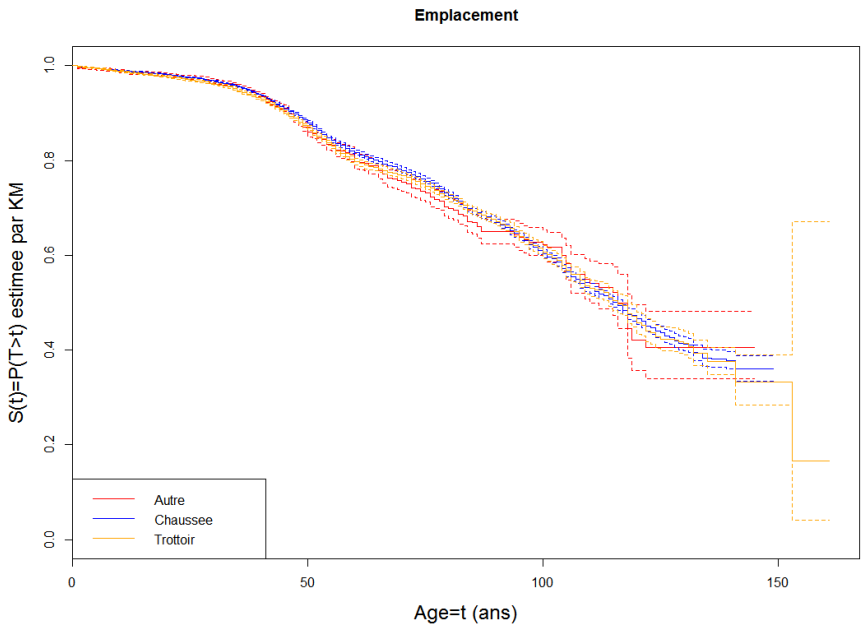


Figure 96: Courbes de survie en effectif tronçons (tronçons partitionnés en 3 sous-ensembles : emplacement sous chaussée, emplacement sous trottoir ou emplacement autre) et leurs intervalles de confiance

L'emplacement n'a pas eu d'impact sur l'âge à la mise hors service des tronçons (Cf. Figure 96).

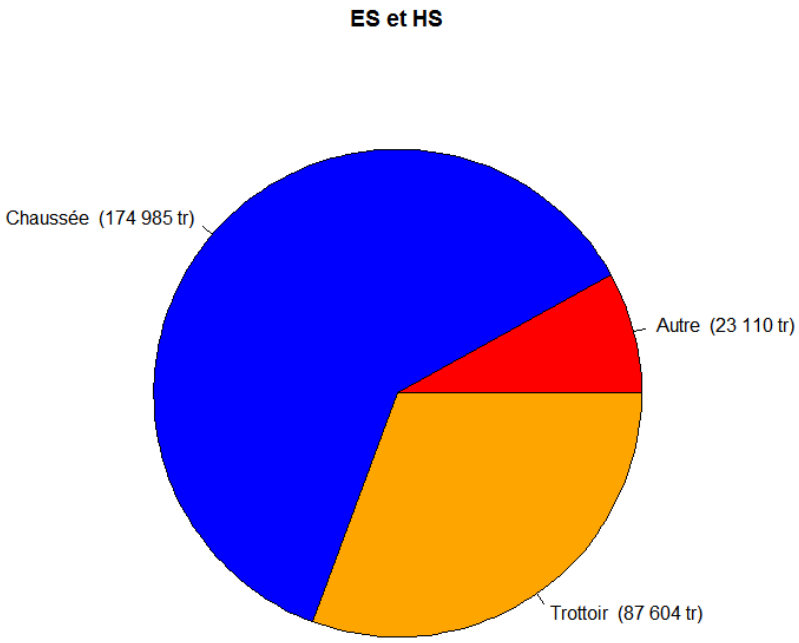


Figure 97: Diagramme de la stratification des tronçons selon leur emplacement sous chaussée ou non ou autre (nombre de tronçons)

Annexe 20: Conseils d'un point de vue statistique pour la gestion des données

Conseil n°1 d'un point de vue statistique

Nous conseillons au SEDIF, d'un point de vu statistique de modifier **sa définition du mot « tronçon ».** La nouvelle définition pourrait être : « Un tronçon est un morceau de tuyau continue, dont on connait les coordonnées (x ;y), posé à la même date par la même entreprise, avec la même technique de pose, de même matériau, de même diamètre, avec les mêmes types de joints et de longueur supérieure à 6m et inférieur à 80 m ».

En effet la définition actuelle aboutira sur du long terme à la création de tronçons de plus en plus petits. Actuellement 46% des tronçons sont des très petites tailles (<6m). 6 m correspond à la taille minimale d'un tuyau réel. Or réaliser des calculs statistiques sur des individus tronçons de moins de 6 m n'a pas de sens technique, que ce soit pour calculer des taux de défaillances au tronçon ou des courbes de survie au tronçon. Si trop de tronçons deviennent inférieurs à 6m, on sera obligé de regrouper les tronçons SIG avec leurs voisins pour dépasser au moins 6m. Ce regroupement pourra s'appeler « segments » ou « tronçons chantiers » ou « biefs »… Néanmoins ce sera un regroupement artificiel. Est-ce que les tronçons qu'on aura regroupé avaient été posés par la même entreprise ? ça on le aura plus… on aura perdu l'information…

Nous vous conseillons donc d'avoir :

- une couche SIG « tronçons stricts » basée sur leurs historiques de pose.
- et d'avoir à côté des couches SIG « données environnementales » : où sont localisés les chaussées et les trottoirs, où sont localisés les réseaux, où sont localisés les balmes, etc.

Ensuite avant toute simulation statistique, au dernier moment croiser la couche « tronçons stricts» avec la couche « environnementale » désirée. Et si par exemple un tronçon est 80% sous chaussée et 20% sous trottoir vous lui attribuez la valeur « chaussée ». Néanmoins pour nous les couches « tronçons stricts» et données « environnementales » doivent évoluer indépendamment. En effet par exemple lors d'une décision d'urbanisme une chaussée peut se transformer à une certaine date en une voie piétonne alors que le tronçon en dessous pour lui rien n'a changé il garde son historique.

Conseil n°2 d'un point de vue statistique

Nous conseillons au SEDIF de rajouter dans la base de données des défaillances FR, l'information sur le **mode de détection de la fuite** :

- fuite visible liée au mauvais état
- ou fuite visible liée à un tiers (accident de pelleteuse, etc.)
- ou fuite détectée par des méthodes de recherche active.

En effet la présence de ce champs permettrait ensuite des calculer des volumes de fuite prévisionnels.