



HAL
open science

The use of deterioration modelling to simulate sewer asset management strategies

Nicolas Caradot

► **To cite this version:**

Nicolas Caradot. The use of deterioration modelling to simulate sewer asset management strategies. Environmental Engineering. Université de Lyon, 2019. English. NNT : 2019LYSEI034 . tel-02484828

HAL Id: tel-02484828

<https://theses.hal.science/tel-02484828>

Submitted on 19 Feb 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.



N°d'ordre NNT : 2019LYSEI034

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
INSA Lyon

Ecole Doctorale N° 162
Mécanique, Energétique, Génie civil, Acoustique

Spécialité de doctorat : Génie Civil

Soutenue publiquement le 27/06/2019, par :
Nicolas Caradot

L'utilisation de modèles de détérioration pour l'élaboration de stratégies de gestion patrimoniale des réseaux d'assainissement

The use of deterioration modelling to simulate sewer asset management strategies

Jean-Luc Bertrand-Krajewski, Professeur (INSA Lyon, France)
Frédéric Cherqui, Maître de Conférences HDR (Université Lyon 1)
Francois Clemens, Professeur (TU Delft, Pays-Bas)
Sophie Duchesne, Professeur (INRS, Québec, Canada)
Andreas Hartmann, Directeur (SE|BS, Allemagne)
Yves Le Gat, Ingénieur en Chef des Ponts, des Eaux et des Forêts
(IRSTEA, France)
João P. Leitão, Chargé de recherche (EAWAG, Suisse)
Pascale Rouault, Responsable de département (KWB, Allemagne)

Président du jury
Co-directeur de thèse
Co-directeur de thèse
Rapporteur
Invité
Examineur
Rapporteur
Invitée

Laboratoire de recherche Déchets Eaux Environnement Pollutions (DEEP) de l'INSA de Lyon

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON http://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr INSA : R. GOURDON	M. Stéphane DANIELE Institut de recherches sur la catalyse et l'environnement de Lyon, IRCELYON-UMR 5256, Équipe CDFA, 2 Avenue Albert EINSTEIN 69 626 Villeurbanne CEDEX directeur@edchimie-lyon.fr
E.E.A.	ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr Sec. : M.C. HAVGOUDOUKIAN ecole-doctorale.eea@ec-lyon.fr	M. Gérard SCORLETTI École Centrale de Lyon 36 Avenue Guy DE COLLONGUE 69 134 Écully Tél : 04.72.18.60.97 Fax 04.78.43.37.17 gerard.scorletti@ec-lyon.fr
E2M2	ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION http://e2m2.universite-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : H. CHARLES secretariat.e2m2@univ-lyon1.fr	M. Philippe NORMAND UMR 5557 Lab. d'Ecologie Microbienne Université Claude Bernard Lyon 1 Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69 622 Villeurbanne CEDEX philippe.normand@univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTÉ http://www.ediss-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : M. LAGARDE secretariat.ediss@univ-lyon1.fr	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 Avenue Jean CAPELLE INSA de Lyon 69 621 Villeurbanne Tél : 04.72.68.49.09 Fax : 04.72.68.49.16 emmanuelle.canet@univ-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHÉMATIQUES http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 Fax : 04.72.43.16.87 infomaths@univ-lyon1.fr	M. Luca ZAMBONI Bât. Braconnier 43 Boulevard du 11 novembre 1918 69 622 Villeurbanne CEDEX Tél : 04.26.23.45.52 zamboni@maths.univ-lyon1.fr
Matériaux	MATÉRIAUX DE LYON http://ed34.universite-lyon.fr Sec. : Marion COMBE Tél : 04.72.43.71.70 Fax : 04.72.43.87.12 Bât. Direction ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIÈRE INSA de Lyon MATEIS - Bât. Saint-Exupéry 7 Avenue Jean CAPELLE 69 621 Villeurbanne CEDEX Tél : 04.72.43.71.70 Fax : 04.72.43.85.28 jean-yves.buffiere@insa-lyon.fr
MEGA	MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE http://edmega.universite-lyon.fr Sec. : Marion COMBE Tél : 04.72.43.71.70 Fax : 04.72.43.87.12 Bât. Direction mega@insa-lyon.fr	M. Jocelyn BONJOUR INSA de Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69 621 Villeurbanne CEDEX jocelyn.bonjour@insa-lyon.fr
ScSo	ScSo* http://ed483.univ-lyon2.fr Sec. : Viviane POLSINELLI Brigitte DUBOIS INSA : J.Y. TOUSSAINT Tél : 04.78.69.72.76 viviane.polsinelli@univ-lyon2.fr	M. Christian MONTES Université Lyon 2 86 Rue Pasteur 69 365 Lyon CEDEX 07 christian.montes@univ-lyon2.fr

Acknowledgments

First, I would like first to express my sincere and deep gratitude to my supervisors Frederic Cherqui from INSA Lyon and Francois Clemens from TU Delft. I know Frederic since more than 10 years, having the chance to meet him as a Master student first, as a colleague in the scientific community then and as PhD student during the last three years. He literally introduced me to the exciting world of research and water management. I always admired your patience, meticulousness and curiosity for our research topic and your dedication to guide me through my researches until the achievement of this PhD. I value a lot the numerous inspiring and passionate conversations we had on water related topics (and more!), in Lyon, Berlin, Delft and all over the world, by attending conferences such as LESAM in Sydney. I met Francois several years ago during a young scientist workshop in Graz. I remember admiring his ability to continuously question our research methodologies in urban water management as well as his interest in formulating new research questions to bring forward the current state of the art. Thanks Francois for supporting me since the beginning in this PhD journey and for the countless hours of debate, discussion, reflection. And of course I want to thank both of you for accompanying my work with this much appreciated balance between scientific rigour and friendliness. It was tough but definitely fun!

Secondly, I would like to express my deepest gratitude to Pascale Rouault for giving me the opportunity to prepare my PhD beside my involvement in other projects at Kompetenzzentrum Wasser and by being an excellent and dedicated chief since almost 10 years. I appreciate your friendship, motivation, trust and your continual support in the development of my scientific and management skills. Long story short, I simply but strongly enjoy working together since many years and probably for more years ahead.

I would like also to thank my KWB colleagues for creating such a pleasant and stimulating

working environment and in particular, Edith Roßbach for giving me the opportunity to conduct my PhD research at KWB, Mathias Riechel for being such a dedicated, meticulous and understanding colleague and Hauke Sonnenberg for introducing and guiding me into the world of programming and clean code.

I would like also to express my gratitude to Nic Lengemann, Elke Eckert, Alexander Ringe and Jan Waschnewski and the other colleagues of the Berliner Wasserbetriebe for the fruitful exchanges we had on the topic of sewer asset management in Berlin and providing the necessary data and material used in this thesis. Thanks to Andreas Hartmann and Michael Timm from Stadtentwässerung Braunschweig (SE|BS) for our friendly collaboration since many years and for being so patient in introducing me asset management from the point of view of the utility. Thanks to Stephane Denhez, Vincent Perez, Emmanuelle Gaudichet and other colleagues from Veolia for our long term collaboration on sewer asset management and for allowing me to discover the diversity of asset management issues since the SEMA project.

I would also like to express my gratitude to all colleagues and friends of the scientific community for the inspiring discussions and insights into the topic: Franz Tscheikner-Gratl, Yves Le Gat, Caty Werey, Ingo Kropp, Rita Ugarelli, Mehdi Ahmadi, Nathalie Hernandez, Andres Torres, Jeroen Langeveld, Gislain Lipeme Kouyi, Jean-Luc Bertrand-Krajewski and many more. Thanks to the members of the jury for accepting to be part of the committee and in particular a very warm thanks to Sophie Duchesne and Joao Leitao for reviewing my work and providing valuable comments for the finalization of the manuscript. I also wish to thank my dear friends in Berlin, Lyon, Oslo, Milano, Zurich, Napoli, Palma, Paris, Madrid and all over the world: thanks for being there and for sharing my life since so many years. Finally, I want to thank my family; my brother (and colleague!) Antoine for being so persevering and achieving his dreams and my parents for their unlimited love and their continuous and unconditional support.

Published papers during the PhD

Journals

- Caradot N., Riechel M., Rouault P., Caradot A., Lengemann N., Eckert E., Ringe A., Clemens F., Cherqui F. 2019. The influence of condition assessment uncertainties on sewer deterioration modelling. *Structure and Infrastructure Engineering*, accepted 05/2019
- Hernández N., Caradot N., Sonnenberg H., Rouault P. and Torres A. 2018. Support tools to predict the critical structural condition of uninspected pipes for case studies of Germany and Colombia. *Water Practice & Technology*, 13(4), 794-802
- Caradot N., Rouault P., Clemens F., Cherqui F. 2018a. Evaluation of uncertainties in sewer condition assessment. *Structure and Infrastructure Engineering*, 14(2), 264-273
- Caradot N., Riechel M., Fesneau M., Hernandez N., Torres A., Sonnenberg H., Eckert E., Lengemann N., Waschnewski J., Rouault P. 2018b. Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany. *Journal of Hydroinformatics*, 20(5), 1131-1147
- Caradot N., Sonnenberg H., Kropp I., Ringe A., Denhez S., Hartmann A., Rouault P. 2017. The relevance of sewer deterioration modelling to support asset management strategies. *Urban Water Journal*, 14(10), 1007-1015
- Riechel, M., Caradot, N., Lengemann, N., Eckert, E., Sonnerberg, H. und Rouault, P. 2018. Bewertung verschiedener Modellansätze zur Vorhersage des Zustands von Abwasserkanälen am Beispiel von Berlin. *Korrespondenz Abwasser*, 65(12), 1098-1106.

Conferences

- Hernandez N., Caradot N., Sonnenberg H., Rouault P., Torres A. 2018. Optimizing SVM Model as Predicting Model for Sewer Pipes in the Two Main Cities in Colombia. In book: *New Trends in Urban Drainage Modelling, Green Energy and Technology*, UDM 2018, Editor: Giorgio Mannina, **Palermo, Italy**
- Caradot N., Hernandez N., Sonnenberg H., Torres A., and Rouault P. 2018. From CCTV data to strategic planning: deterioration modelling for large sewer networks in Germany and Colombia. In: *Proceedings of 13th International Conference on Hydroinformatics HIC 2018*, 02-06 July 2018, **Palermo, Italy**

- Caradot N., Hernandez N., Sonnenberg H., Torres A., Hartmann A. and Rouault P. 2017. An evaluation of the performance of sewer deterioration models in Germany and Colombia. In: Proceedings of the 7th IWA Leading Edge Strategic Asset Management Conference, 20-22 June 2017, **Trondheim, Norway**
- Hernández, N., Caradot, N., Sonnerberg, H., Rouault, P. and Torres, A. 2017a. Support tools to predict the critical structural condition of uninspected sewer pipes in Bogota D.C. In: Proceedings of the 7th IWA Leading Edge Strategic Asset Management Conference, 20-22 June 2017, **Trondheim, Norway**
- Hernández, N., Caradot, N., Sonnerberg, H., Rouault, P and Torres, A. 2017b. Support Vector Machines used for the prediction of the structural conditions of pipes in Bogota's sewer system. In: Proceedings of 14th International Conference on Urban Drainage, 10-15 September 2017, **Prague, Czech Republic**
- Caradot N., Sonnenberg H., Kropp I., Ringe A., Denhez S., Hartmann A., Rouault P. 2016. The benefits of deterioration modelling to support sewer asset management strategies. In: Proceedings of 8th International Conference on Sewer Processes and Networks, 31 August – 2 September 2016, **Rotterdam, The Netherlands**

Published papers before the PhD

Journals

- Caradot N., Sonnenberg H., Rouault P., Hartmann A., Kästner A., Kropp I., Ringe A., Timm M., Mesek C. 2015. Wie Zuverlässig sind Kanalalterungsmodelle? *KA Korrespondenz Abwasser, Abfall*, 62(10), 884-892
- Caradot N., Sonnenberg H., Kropp I., Schmidt T., Ringe A., Denhez S., Hartmann A., Rouault P. 2014. Sewer deterioration modelling for asset management strategies. *IWA Water Asset Management International*, 10(3), 03-08

Conferences

- Caradot N., Sonnenberg H., Hartmann A., Kropp I., Ringe A., Denhez S., Timm M., Rouault P. 2015. The potential of deterioration modelling to support sewer asset management. In: Proceedings of 6th IWA Leading Edge Strategic Asset Management Conference, 17-19 November 2015, **Yokohama, Japan**

- Caradot N., Sonnenberg H., Hartmann A., Kropp I., Ringe A., Denhez S., Timm M., Rouault P. 2015. The influence of data availability on the performance of sewer deterioration modelling. In: Proceedings of 10th International Urban Drainage Modelling Conference, September 20-23 2015, **Mont-Saint-Anne, Quebec, Canada**
- Caradot N., Sonnenberg H., Rouault P., Hartmann A., Kästner A., Kropp I., Ringe A., Timm M., Mesek C. 2014. Wie Zuverlässig sind Kanalalterungsmodelle? In: Proceedings of Inspektions- und Sanierungstage - DWA, 12-13 November 2014, **Dortmund, Germany**
- Caradot N., Sonnenberg H., Kropp I., Schmidt T., Ringe A., Denhez S., Timm M., Hartmann A., Rouault P. 2014. What is the reliability of sewer deterioration models? In: Proceedings of IWA World Water Congress, 21-26 September 2014, **Lisbon, Portugal**
- Caradot N., Sonnenberg H., Kropp I., Schmidt T., Ringe A., Denhez S., Hartmann A., Rouault P. 2013. Sewer deterioration modeling for asset management strategies – state-of-the-art and perspectives. In: Proceedings of 5th IWA Leading Edge Strategic Asset Management Conference, 9-12 September 2013, **Sydney, Australia**

Abstract

Insufficient public and municipal investment represent a major challenge for the long term management of urban drainage systems. Utilities are challenged to develop efficient rehabilitation strategies in order to maintain the level of service. Closed-circuit television (CCTV) inspection is used since the 1980's as industry standard for sewer investigation system and structural performance evaluation. Due to budget restrictions, inspection rates are generally low and municipalities tend to inspect only a small part of their network (e.g. in France, less than 5% according to Ahmadi *et al.*, 2014c). Since the definition of rehabilitation strategies is limited by the lack of information about sewer condition and remaining life, deterioration models have been developed to forecast the evolution of the system according to its current and past condition.

One of the main factors hampering the uptake of deterioration modelling by utilities is the lack of real scale evidence of the tangible benefits provided. In particular, most utilities are concerned by the minimum amount of CCTV data required and the relevance of using such models on their networks with limited data availability. Finally, most utilities acknowledge the uncertainties in the procedure of sewer condition assessment, mainly due to the subjectivity of the coding operator. There is a strong need to quantify precisely the uncertainty of the sewer condition assessment procedure and its influence on the outcomes of deterioration modelling.

The thesis aims at addressing these gaps by assessing the performance of sewer deterioration modelling using a case study with high CCTV data availability and by identifying the influence of CCTV data quality and availability on modelling performance.

The study has been performed with a statistical (GompitZ) and a machine learning (Random Forest) deterioration models using the extensive CCTV database of the cities of Braunschweig and Berlin in Germany. Our results show, that at network level, both machine learning and

statistical models can simulate with sufficient accuracy the condition distribution of the network, even in case of low data availability. At the pipe level, the machine learning model outperforms the statistical model. Regarding CCTV data uncertainty, our results highlight that the probability to inspect correctly a pipe in poor condition is close to 80-85% and thus the probability to overestimate the (good) condition of the pipe is close to 15-20% (False Negative). The impact of the uncertainties on the prediction of a deterioration model is not negligible. The analysis shows that the required replacement rate to maintain a constant proportion of segments in poor condition is underestimated if the uncertainties are not included in the analysis.

Résumé

Les infrastructures de collecte et de traitement des eaux usées représentent un investissement considérable pour les municipalités. La plupart des villes sont aujourd'hui confrontées au problème de gestion sur le long terme de leur patrimoine vieillissant avec des besoins urgents de réparation, rénovation ou remplacement des infrastructures. Elles se doivent donc d'élaborer des stratégies de gestion patrimoniale efficaces pour maintenir le niveau de service rendu en respectant l'enveloppe budgétaire allouée.

L'inspection télévisée (ITV) reste la méthode quasi exclusivement utilisée depuis des décennies pour la surveillance du fonctionnement des réseaux d'assainissement et l'évaluation de leur condition structurelle. Dû à l'absence de réglementation et aux coûts élevés liés au linéaire important à inspecter, la plupart des collectivités n'ont qu'une connaissance partielle de la condition de leur réseau et manquent d'information pour planifier efficacement les actions de maintenance et renouvellement. Pour pallier cette situation, des modèles numériques de détérioration ont été développés pour simuler la condition des conduites non inspectées et prévoir l'évolution future de la condition du réseau.

Un des obstacles majeurs à l'appropriation de ces outils par les collectivités est le manque de preuves à grande échelle de leur bonne performance et l'absence de cas d'étude pour illustrer leurs bénéfices. D'autre part, l'influence de la quantité de données disponibles pour étalonner les modèles sur la fiabilité des prédictions est mal connue. Les taux d'inspection étant relativement faibles, les collectivités sont rarement assurées d'avoir les données nécessaires pour développer des modèles fiables. Enfin, de nombreuses collectivités reconnaissent les incertitudes liées aux données d'inspection mais aucune étude n'a pu constater l'influence de ces incertitudes sur les résultats de modélisation.

Ce travail de thèse vise à évaluer la performance des modèles de détérioration dans un cas idéal

ou l'intégralité d'un réseau d'assainissement a déjà été inspectée ainsi qu'à évaluer l'influence de la quantité et qualité des données d'inspections disponibles sur la qualité de prédiction.

Cette étude a été réalisée en utilisant deux modèles de référence : un modèle statistique (GompitZ) et un modèle d'apprentissage automatique (Random Forest). Ces modèles ont été étalonnés et testés à l'aide des bases de données d'inspection des villes de Braunschweig et Berlin en Allemagne. Nos résultats montrent que les deux modèles sont capables de reproduire finement la condition du réseau, c.à.d. la proportion de conduites dans chaque état de santé. Ils peuvent d'ailleurs faire preuve d'une précision satisfaisante même en cas de faible disponibilité des données. Le modèle d'apprentissage donne des résultats beaucoup plus satisfaisants à l'échelle de la conduite. En ce qui concerne l'incertitude des données d'inspection, nos résultats soulignent que la probabilité d'évaluer correctement l'état de santé d'une conduite en mauvais état est proche de 80-85%. La probabilité d'être trop optimiste et de surévaluer son (bon) état est proche de 15-20% (faux négatif). L'impact des incertitudes des données d'inspection sur les prédictions des modèles prévision d'un modèle de détérioration n'est pas négligeable. L'analyse montre également que le taux de remplacement requis pour maintenir la condition générale du réseau est sous-estimé si les incertitudes ne sont pas prises en compte.

Table of content

Chapter 1 Synthèse du travail de thèse14

1.1 Introduction.....	14
1.2 R1: quelle est la performance des modèles de détérioration?.....	18
1.3 R2 : comment évolue la qualité de prédiction des modèles de détérioration en fonction de la quantité de données d'inspection disponible pour l'étalonnage ?	21
1.4 R3: quelle est la fiabilité des données d'inspection et comment quantifier leurs incertitudes?	23
1.5 R4: quelle est l'influence de l'incertitude des données d'inspections sur la qualité de prédiction?.....	25
1.6 Perspectives de recherche.....	27

Chapter 2 Introduction30

2.1 Background.....	30
2.2 Research needs.....	46
2.3 Aim and structure of the thesis.....	49
2.4 Case studies.....	53

Chapter 3 The influence of data availability on modelling performance58

3.1 Overview and context of paper I.....	58
3.2 Research method.....	58
3.3 Data description.....	64
3.4 Results and discussion	64
3.5 Conclusion.....	71

Chapter 4 Practical benchmarking of the performance of statistical and machine learning models73

4.1 Overview and context of paper II.....	73
4.2 Material and Method	74
4.3 Results and discussion	87
4.4 Conclusion.....	96

Chapter 5 Evaluation of uncertainties in sewer condition assessment98

5.1 Overview and context of paper III.....	98
--	----

5.2 Research method.....	99
5.3 Evaluation of uncertainties in Braunschweig	105
5.4 Evaluation of uncertainties in Berlin	113
5.5 Conclusion.....	118
Chapter 6 The influence of condition assessment uncertainties in sewer deterioration modelling.....	120
6.1 Overview and context of paper IV	120
6.2 Material and method	121
6.3 Results and discussion	125
6.4 Conclusion.....	131
Chapter 7 Conclusion and perspectives	133
7.1 Conclusion.....	133
7.2 Perspectives	139
Chapter 8 References	144

Chapter 1 Synthèse du travail de thèse

1.1 Introduction

Les infrastructures de collecte et de traitement des eaux usées représentent un investissement considérable pour les municipalités. Une ville moyenne de 100 000 à 500 000 habitants peut avoir un patrimoine de 1 000 à 5 000 km de réseau d'assainissement souterrain. Le coût de maintenance et de réhabilitation de ces infrastructures s'élève à plusieurs millions d'euros non seulement pour les métropoles, mais aussi pour les petites municipalités (IPK, 2014). Au cours des 30 dernières années, la plupart des municipalités ont investi massivement dans la modernisation des stations d'épuration et l'extension des réseaux d'assainissement pour faire face à la croissance urbaine alors qu'une part relativement réduite du budget a été affectée à la réhabilitation des infrastructures existantes (AWWA, 2012). En conséquence, la plupart des villes sont aujourd'hui confrontées au problème de gestion sur le long terme de leur patrimoine vieillissant avec des besoins urgents de réparation, rénovation ou remplacement des infrastructures. Une enquête récente menée auprès de 397 collectivités aux États-Unis et au Canada a montré que le vieillissement des infrastructures et la gestion durable du patrimoine sont les deux principaux enjeux du secteur (Black et Veatch, 2013). Aux États-Unis, l'association nationale de génie civil évalue le besoin d'investissement pour la gestion du patrimoine à 91 milliards de dollars (ASCE, 2011). Sur ces 91 milliards, seulement 36 milliards seraient actuellement financés, laissant un déficit de financement de près de 55 milliards de dollars. En Allemagne, une étude nationale a montré que 20% des réseaux d'assainissement présentent de graves dysfonctionnements nécessitant une réhabilitation à court ou moyen terme (Berger *et al.*, 2015). Toujours en Allemagne, les investissements annuels pour la réhabilitation des réseaux d'assainissement s'élèvent à quatre milliards d'euros tandis que les besoins d'investissement sont estimés à plus de sept milliards d'euros (IPK, 2014; KfW, 2016). Les

déficits de financement et les investissements nécessaires pour faire face à la détérioration des infrastructures entraîneront nécessairement de futures augmentations des tarifs de l'eau (Oelmann *et al.*, 2017) ou une dégradation de la qualité de service. Les municipalités se doivent donc d'élaborer des stratégies de gestion patrimoniale efficaces pour maintenir le niveau de service rendu et limiter l'augmentation des redevances.

L'inspection télévisée (ITV) reste la méthode quasiment exclusivement utilisée depuis des décennies pour la surveillance du fonctionnement des réseaux d'assainissement et l'évaluation de leur condition structurelle. Une caméra embarquée sur un robot est introduite dans la conduite par un regard. L'inspection fournit des données visuelles (images ou vidéos) de la surface interne de la conduite et l'analyse de l'image permet d'identifier le type et l'emplacement des défauts tels que les fissures, effondrements partiels, fuites, sédiments, intrusion de racines, etc. La principale limitation de l'inspection télévisée est qu'elle ne fournit qu'une représentation visuelle de la surface interne des canalisations. Elle ne permet d'évaluer ni les problèmes externes de remblais ni l'intégrité ou la résistance des parois des canalisations.

Dans un premier temps, les défauts et dysfonctionnement visualisés pendant les inspections sont encodés manuellement par un technicien en suivant un système de codage normalisé (par exemple en Europe la norme EN 13508-2). Dans un second temps, l'état de santé des conduites et l'urgence de réhabilitation sont évalués en analysant le type et l'étendue des défauts. Plusieurs méthodologies existent pour classier automatiquement l'état de santé : par exemple RERAU en France (Le Gauffre *et al.*, 2004) ; SRM au Royaume-Uni (WRc, 2013) ; PACP aux États-Unis (NASSCO, 2007) ; DWA M149-3 en Allemagne (DWA, 2011). Ces méthodes permettent de traduire la quantité considérable de données d'inspection visuelle en une note d'état de santé.

Les données d'inspection sont cruciales pour la planification des programmes de réhabilitation des réseaux. Cependant, leur qualité et précision sont rarement mises en doute. Dirksen *et al.* (2013) ont analysé leur fiabilité et mis en évidence la subjectivité de la procédure d'inspection à trois niveaux : lors de la reconnaissance des défauts, lors de la description des défauts et lors de l'évaluation de l'état de santé. La probabilité qu'un inspecteur manque un défaut (faux négatif) est nettement plus élevée que la probabilité qu'un défaut soit signalé sans être présent (faux positif). La probabilité de faux positif est de l'ordre de quelques pourcents alors que la probabilité de faux négatif s'élève à 25%. En outre, Dirksen *et al.* (2013) ont démontré que l'évaluation de l'état de santé d'une même conduite par plusieurs inspecteurs peut mener à des interprétations radicalement différentes.

Dans la plupart des pays, les taux d'inspection des réseaux sont relativement faibles dû à l'absence de réglementation nationale et aux coûts d'inspections jugés relativement élevés. Ainsi la plupart des collectivités n'ont qu'une connaissance partielle de l'état de leur réseau et manquent d'information pour planifier efficacement les actions de maintenance et de renouvellement. Pour pallier cette situation, de nombreux chercheurs ont développés des modèles de détérioration pour simuler la condition des conduites non-inspectées et prévoir l'évolution future de l'état du réseau. Ces modèles permettent aux municipalités de planifier les programmes d'inspection (priorisation des inspections avec un budget limité) et les besoins d'investissement à moyen ou long terme. La plupart des outils disponibles aujourd'hui sont des modèles statistiques ou d'apprentissage automatique (intelligence artificielle). Par exemple, les fonctions de survie permettent de simuler la durée de vie restante des conduites à partir de leurs caractéristiques techniques et environnementales : âge, matériau, type d'effluent, diamètre, intensité du trafic en surface, etc. Les notes d'état de santé sont corrélées avec l'âge et les caractéristiques des conduites, en suivant un modèle mathématique. Les méthodes d'apprentissage automatique suivent le même principe mais ne requièrent aucune hypothèse sur

la structure du modèle. Le modèle identifie les variables explicatives pour décrire la détérioration des conduites et peut ensuite prédire la condition de conduites non inspectées.

Bien que de nombreux outils de modélisation existent, peu de collectivités utilisent leur potentiel pour planifier leur stratégie de gestion patrimoniale. Un des freins majeurs est le manque de preuves à grande échelle de leur bonne performance et l'absence de cas d'étude pour illustrer leurs bénéfices. De nombreuses études ont visé à évaluer la qualité de prédiction des modèles mais la plupart des tests ont été réalisés sur des jeux de données synthétiques ou relativement réduits (< 2 000 conduites). D'autre part, l'influence de la quantité de données disponibles pour étalonner les modèles sur la fiabilité des prédictions est mal connue. Les taux d'inspection étant relativement faibles, les collectivités sont rarement assurées d'avoir les données nécessaires pour développer des modèles fiables. Enfin, de nombreuses collectivités reconnaissent les incertitudes liées aux données d'inspection, en lien notamment avec la subjectivité de l'évaluation de l'état de santé des conduites. Ces données constituant la base des modèles, il est crucial de quantifier leurs incertitudes et d'évaluer les conséquences de ces incertitudes sur les résultats de modélisation.

Pour répondre à ces préoccupations, ce travail de thèse est structuré autour de quatre problématiques principales :

- R1 : quelle est la performance des modèles de détérioration ? En particulier, quelle est la qualité de prédiction des modèles de détérioration dans un cas idéal où l'intégralité du réseau d'assainissement a déjà été inspectée ?
- R2 : comment évolue la qualité de prédiction des modèles de détérioration en fonction de la quantité de données d'inspection disponible pour l'étalonnage ? Quelle est la taille minimale requise du jeu de données d'inspection pour obtenir un modèle fiable ?

- R3 : quelle est la fiabilité des données d’inspection et comment quantifier leurs incertitudes ?
- R4 : quelle est l’influence de l’incertitude des données d’inspection sur la qualité de prédiction des modèles de détérioration ?

La figure 1.1 présente l’articulation de ces quatre problématiques dans le contexte de la gestion patrimoniale, de la collecte de données jusqu’à la définition des stratégies d’inspection et de réhabilitation. Les paragraphes suivants détaillent les questions de recherche.

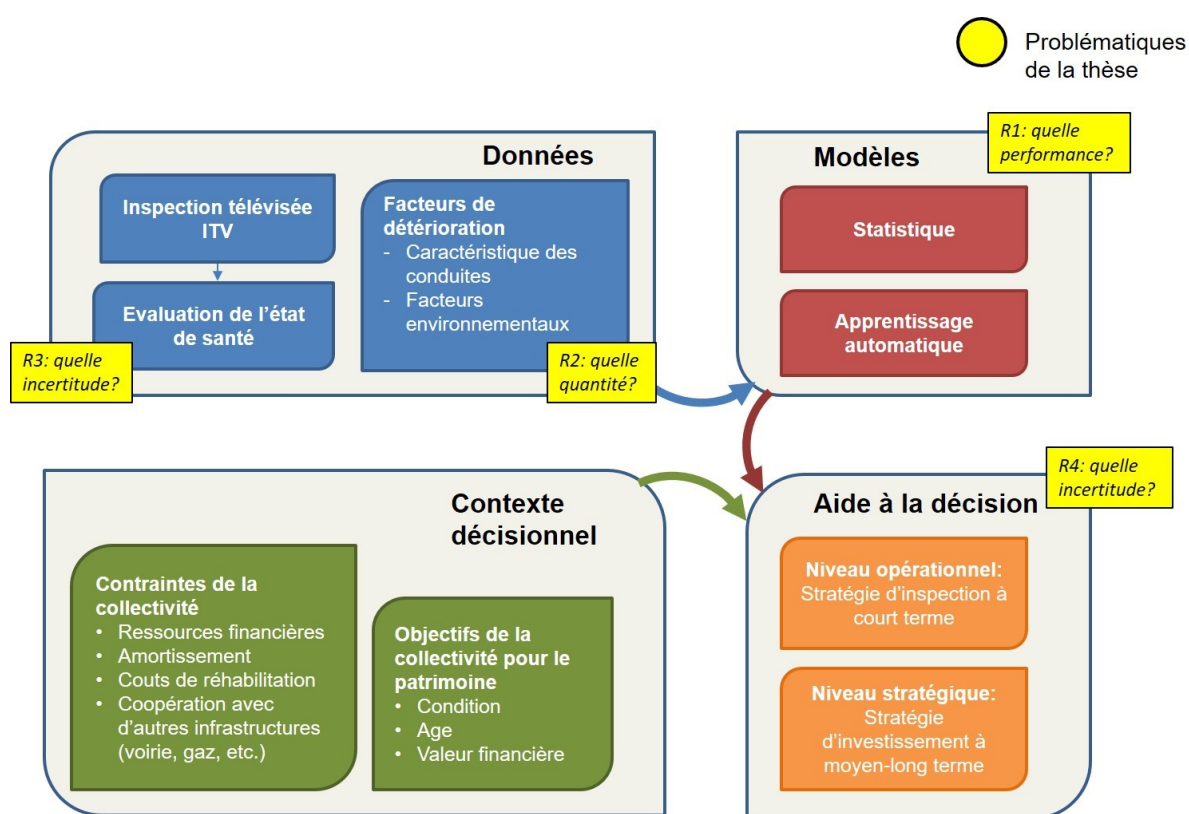


Figure 1.1. Articulation des problématiques de la thèse dans le contexte de la gestion patrimoniale

1.2 R1: quelle est la performance des modèles de détérioration?

La question de la performance des modèles de détérioration est cruciale pour justifier de leur implémentation par les collectivités. Il s’agit de vérifier leur qualité de prédiction à grande échelle en utilisant des données réelles d’inspections pour étalonner les modèles et valider leur application.

1.2 R1: quelle est la performance des modèles de détérioration?

La performance de deux modèles de détérioration (un modèle statistique et un modèle d'apprentissage automatique) a été évaluée en utilisant les données du réseau d'assainissement de la ville de Berlin en Allemagne. La collectivité dispose d'un jeu de données d'inspection unique avec l'évaluation de l'état de santé de plus de 100 000 conduites. Presque toutes les conduites ont été inspectées au moins une fois, ce qui fait de Berlin un cas idéal pour analyser la qualité de prédiction des modèles à grande échelle.

Un jeu d'indicateurs a été développé en concertation avec la collectivité pour pouvoir quantifier la performance des modèles du point de vue de l'utilisateur final. Les indicateurs doivent être intuitifs, explicites et clairement compréhensibles par la collectivité pour pouvoir statuer sur l'utilité d'utiliser ou non la modélisation pour planifier les stratégies de gestion patrimoniale à venir.

L'approche statistique retenue est le modèle GompitZ (Le Gat, 2008), basé sur les théories de l'analyse de survie et des chaînes de Markov. La méthode d'apprentissage automatique choisie est une forêt d'arbres décisionnels (Random Forest), une méthode d'apprentissage d'ensemble pour la classification ou la régression, déjà appliquée avec succès au Canada, en Norvège et au Portugal (Harvey & McBean, 2014; Rokstad *et al.*, 2015; Vitorino *et al.*, 2014). Les résultats principaux sont les suivants :

- **Les deux modèles sont capables de reproduire finement la condition du réseau,** c.à.d. la proportion de conduites dans chaque état de santé (Figure 1.2). Les déviations entre les proportions prédites et observées sont très faibles (<5% pour le modèle d'apprentissage et <1% pour le modèle statistique). Ces premiers résultats soulignent le potentiel de la modélisation pour estimer la condition du réseau.

- A l'échelle de la conduite, le modèle statistique n'est pas plus fiable qu'un modèle aléatoire** qui attribuerait une note d'état de santé aléatoire à chaque conduite. L'approche statistique, utile à l'échelle du réseau, n'est pas prévue ou adaptée pour simuler précisément l'état de chaque conduite.
- Le modèle d'apprentissage donne des résultats beaucoup plus satisfaisants à l'échelle de la conduite** : 67% des conduites inspectées en mauvais état ont été correctement prédites par le modèle (taux de vrais positifs) et seulement 9% des conduites inspectées en mauvais état ont été prédites en bon état (taux de faux négatifs). Le taux de vrais positifs de 67% est à rapprocher du taux de vrais positifs d'une inspection télévisée, de l'ordre de 80% (voir chapitre 5). Le modèle ne peut évidemment pas être plus précis que les données utilisées pour son étalonnage.

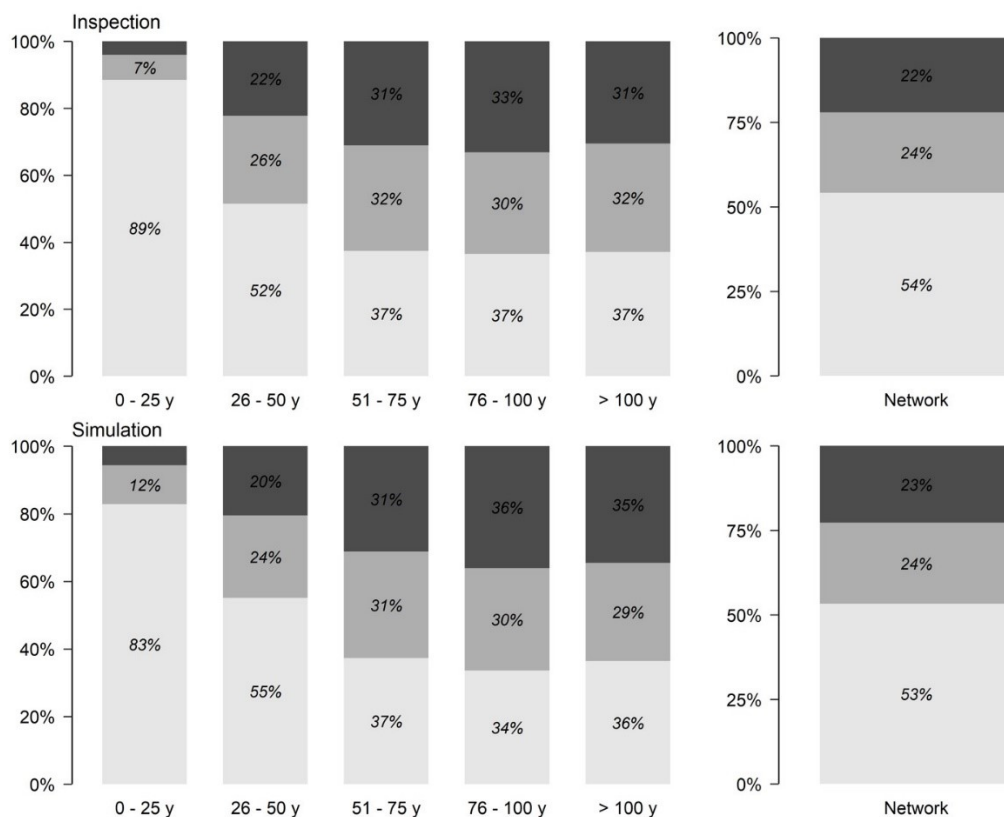


Figure 1.2. Proportions de conduites dans chaque état de santé pour l'intégralité du réseau (droite) et pour différentes classes d'âge (gauche). Les graphiques du haut montrent les proportions dérivées des résultats d'inspection. Les graphiques du bas montrent les proportions simulées par GompitZ

1.2 R1: quelle est la performance des modèles de détérioration?

Un des désavantages du modèle d'apprentissage est qu'il peut conduire à des prévisions erronées telles que l'amélioration de l'état d'une conduite dans le temps. On peut en conclure que les modèles d'apprentissage ne doivent être utilisés que pour la classification de l'état de santé des conduites, mais pas pour la prévision à long terme. Ce problème ne concerne pas le modèle statistique puisque le processus de détérioration suit une distribution statistique donnée qui empêche toute amélioration de l'état de santé dans le temps.

1.3 R2 : comment évolue la qualité de prédiction des modèles de détérioration en fonction de la quantité de données d'inspection disponible pour l'étalonnage ?

Les premières conclusions de R1 permettent de connaître la performance des modèles dans le cas particulièrement avantageux où l'intégralité du réseau a été inspectée et donc un jeu de données conséquent est disponible pour l'étalonnage des modèles. Comme de nombreuses collectivités ne disposent pas aujourd'hui d'un tel jeu de données, notre deuxième question de recherche vise à évaluer l'influence de la disponibilité des données sur la performance des modèles et la taille requise du jeu de données pour pouvoir obtenir des prédictions relativement fiables.

Le modèle statistique GompitZ a été étalonné en utilisant les données d'inspection de la ville de Braunschweig en Allemagne. Cette ville est également un cas d'étude idéal, car toutes les conduites du réseau d'assainissement (environ 40 000 conduites, 1 300 km) ont déjà été inspectées au moins une fois. La performance du modèle a été étudiée en réduisant la taille du jeu de données utilisé pour l'étalonnage (Figure 1.3). En utilisant 50% des données d'inspection (sélectionnée aléatoirement sur le réseau), la qualité de prédiction est sensiblement la même qu'en utilisant 100% des données. Même avec 3% des données (environ 1 000 conduites), le modèle statistique est encore capable de simuler la proportion de conduites en mauvais état avec une déviation inférieure à 10%.

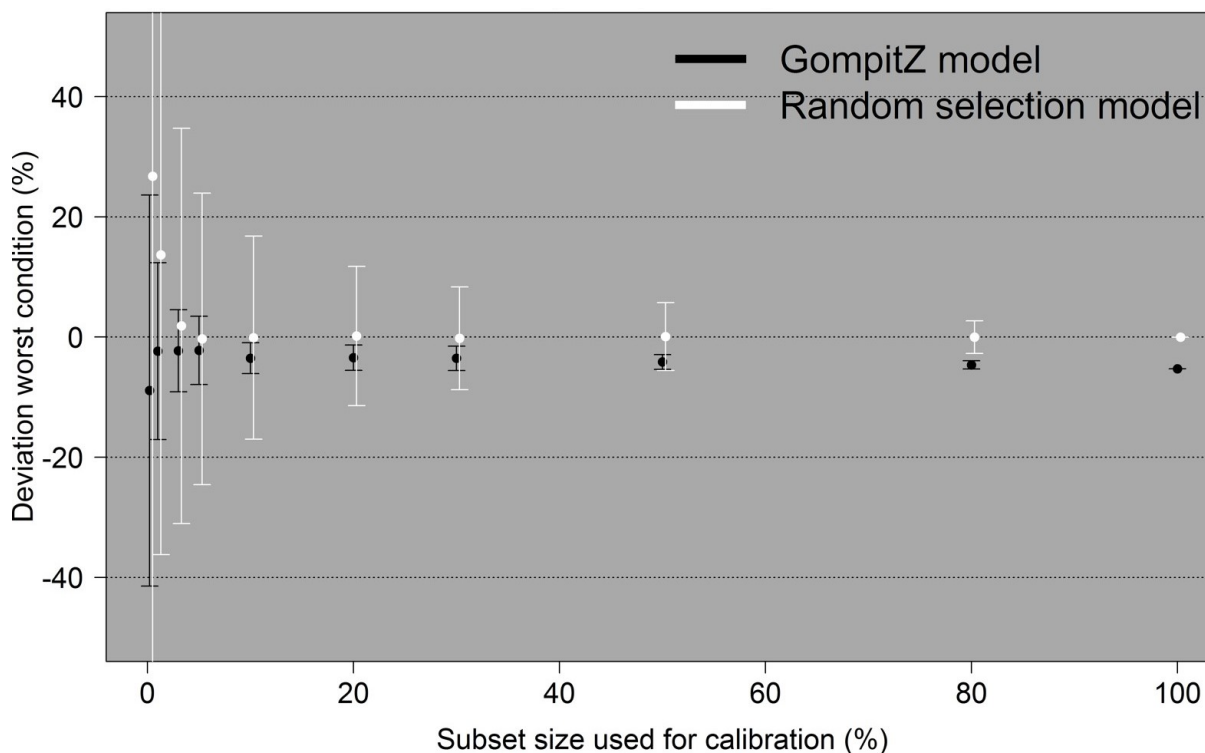


Figure 1.3. Analyse des déviations entre prédiction et observation pour le modèle statistique et pour un modèle aléatoire, pour différentes tailles de sous-ensembles d'étalonnage. Les déviations représentent l'écart entre la proportion de conduites en mauvais état prédites et observées pour les anciennes conduites uniquement (conduites > 70 ans)

Sur la base de cette étude, il semble raisonnable de constituer **un jeu de données d'au moins 1 000 conduites (environ 50 km) pour obtenir des résultats de simulation fiables**. Ces résultats obtenus à Braunschweig confirment les résultats obtenus par Tran (2016) en Australie et Ahmadi *et al.* (2016) qui recommandent respectivement un jeu minimal de 700 ou 1 000 conduites. Toutefois cette conclusion n'est valable que si les inspections utilisées pour l'étalonnage du modèle ont été sélectionnées aléatoirement sur l'intégralité du réseau. Le nombre de conduites nécessaires pourra donc dépendre de la méthode d'échantillonnage utilisée notamment si les conduites en mauvais états sont sous-représentées dans la base de données (Ahmadi *et al.*, 2016).

1.4 R3: quelle est la fiabilité des données d'inspection et comment quantifier leurs incertitudes?

La fiabilité des données d'inspection a été fortement mise en doute ces dernières années dans la littérature scientifique (Dirksen *et al.*, 2013 ; Korving et Clemens, 2004; van Riel *et al.*, 2017; Roghani *et al.*, 2019 ; van der Steen *et al.*, 2014). Selon ces études, l'évaluation de l'état de santé des conduites tend à sous-estimer leur niveau de détérioration. Les erreurs d'évaluation de l'état de santé proviennent notamment (Cherqui *et al.*, 2017) : (1) de l'environnement de la conduite (par exemple, les obstacles qui entravent la visualisation précise) ; (2) du processus d'évaluation de l'état santé (par exemple, le système de codage utilisé, l'expérience et la subjectivité de l'opérateur) ; (3) des caractéristiques de la conduite, par exemple le diamètre ou le matériau ; (4) et des caractéristiques des défauts (par exemple, taille, nombre, répartition spatiale). La plupart des collectivités reconnaissent les incertitudes liées à la procédure d'évaluation de l'état de santé des conduites même si elles n'ont pas encore été quantifiées précisément.

Une méthodologie basée sur l'analyse des inspections répétées des mêmes conduites a été développée pour déterminer les incertitudes liées à l'évaluation de l'état de santé des conduites. Cette méthodologie est basée sur une procédure d'optimisation couplée à une simulation de Monte-Carlo. Elle a été utilisée pour déterminer la probabilité d'évaluer correctement l'état de santé d'une conduite suite à une inspection télévisée (vrai positif pour une conduite correctement évaluée en mauvais état et vrai négatif pour une conduite correctement jugée en bon état) ainsi que les probabilités de juger l'état de santé de manière trop optimiste (faux négatif) ou trop pessimiste (faux positif). Cette étude a été réalisée en utilisant conjointement les bases de données des villes de Berlin et Braunschweig. Les principaux résultats sont les suivants :

- **La probabilité d'évaluer correctement l'état de santé d'une conduite en mauvais état est proche de 80-85%.** La probabilité d'être trop optimiste et de surévaluer son

(bon) état est proche de 15-20% (faux négatif).

- La probabilité d'évaluer correctement l'état de santé d'une conduite en bon état est légèrement supérieure à la probabilité d'évaluer correctement l'état de santé d'une conduite en mauvais état. Cela signifie que **les incertitudes sont réduites lors de l'évaluation de l'état de santé des conduites avec peu de défauts** et que les inspecteurs sont plus enclins à commettre des erreurs lorsque de nombreux défauts sont présents.
- **Pour les conduites en bon état, la probabilité de faux positif est plus élevée que la probabilité de faux négatif.** Il y a donc une plus grande probabilité d'être trop pessimiste, de considérer les conduites comme plus détériorées qu'elles ne sont réellement. Au contraire, **l'évaluation des conduites en mauvais état est plus sujette au faux négatif qu'aux faux positifs.** Il y a une plus grande probabilité d'être trop optimiste.

Les incertitudes liées à chaque évaluation de l'état de santé ne sont pas uniquement dues à la procédure d'inspection et à la subjectivité de l'inspecteur lors de l'encodage des défauts. L'analyse des écarts lors d'inspections répétées met en évidence d'autres sources d'incertitudes, telles que l'absence de documentation de certaines réhabilitations (par exemple, la conduite a été réhabilitée mais cette information n'a pas été saisie dans la base de donnée) ou la présence dans la base de données d'inspections effectuées dans un autre but que l'analyse de l'état de santé de la conduite (par exemple, vérifier l'emplacement des branchements sur la conduite sans encoder les défauts structurels). La réduction des incertitudes pourrait commencer par l'amélioration des procédures de gestion des données afin de pouvoir filtrer les inspections non conformes avant d'étalonner les modèles de détérioration.

1.5 R4: quelle est l'influence de l'incertitude des données d'inspections sur la qualité de prédiction?

Les incertitudes liées à l'évaluation de l'état de santé des conduites ne sont pas négligeables. Il s'agit de savoir si elles peuvent avoir une influence sur les prédictions des modèles de détérioration et éventuellement de corriger les prédictions pour considérer les biais inhérents aux données d'étalonnage.

Une méthode a été proposée pour propager les incertitudes des données d'inspection dans les prédictions d'un modèle de détérioration statistique. Le modèle a été utilisé à Berlin pour simuler des stratégies de réhabilitation simples et évaluer l'impact des incertitudes sur les prédictions du modèle. Les résultats suivants peuvent être soulignés.

- La propagation des incertitudes dans les courbes de survie du modèle statistique produit un intervalle de confiance autour des courbes de survie d'origine. A titre d'exemple, **pour une conduite âgée de 100 ans, l'incertitude sur la probabilité d'être en mauvais état est de $\pm 12\%$.**
- L'analyse de cet intervalle de confiance met en évidence **la présence d'une erreur systématique** : la courbe de survie corrigée moyenne ne chevauche pas la courbe de survie originale (Figure 1.4). Comme indiqué dans la section précédente, les erreurs les plus probables sont les faux positifs pour les conduites en bon état et les faux négatifs pour les conduites en mauvais état. Ce résultat peut paraître trivial mais il est crucial pour l'analyse des incertitudes à l'échelle du réseau. La plupart des jeunes conduites (<30 ans) sont en bon état et sont donc plus enclines aux faux positifs qu'aux faux négatifs. La probabilité d'être trop pessimiste étant plus grande, la courbe de survie corrigée est plus optimiste que la courbe de survie initiale. Au contraire, la plupart des conduites les plus âgées (> 75 ans) sont en mauvais état et sont donc plus sujettes aux

faux négatifs qu'aux faux positifs. La probabilité d'être trop optimiste étant plus grande, la courbe de survie corrigée est plus pessimiste que la courbe de survie initiale.

- L'incertitude systématique augmente avec le nombre d'années simulées. L'analyse montre également que **le taux de remplacement requis pour maintenir la condition générale du réseau est sous-estimé** si les incertitudes ne sont pas prises en compte dans le modèle.

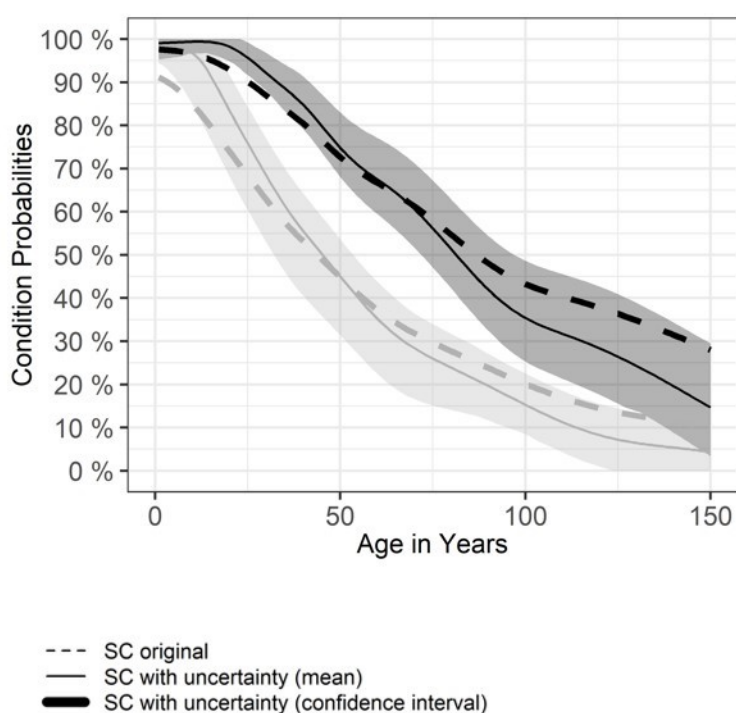


Figure 1.4. Courbes de survie (SC = Survival Curve) originales et corrigées avec intervalle de confiance pour la ville de Berlin avec trois classes d'état de santé. Les courbes de survie représentent la probabilité d'être dans chaque état de santé en fonction de l'âge de la conduite. Les courbes gris clair montrent la transition du bon état à l'état intermédiaire ; les courbes gris foncé montrent la transition de l'état intermédiaire au mauvais état.

Même influencés par les incertitudes, les modèles restent un outil fiable pour prévoir l'évolution future de la condition du réseau pour différents scénarios de réhabilitation. Cependant, il est crucial d'intégrer et de communiquer ces incertitudes aux décideurs pour éviter les interprétations erronées lors de la prise de décision.

1.5 R4: quelle est l'influence de l'incertitude des données d'inspections sur la qualité de prédiction?

1.6 Perspectives de recherche

Au cours des dernières décennies, de nombreux efforts ont été déployés pour développer des modèles de détérioration afin de soutenir les collectivités dans la planification des stratégies de gestion patrimoniale. Les modèles de détérioration se sont avérés être des outils fiables pour simuler la détérioration du réseau et identifier les conduites en état critique. L'analyse de survie et les modèles de Markov se révèlent les approches les plus fiables pour simuler la détérioration des conduites au niveau du réseau. Ils peuvent faire preuve d'une précision satisfaisante même en cas de faible disponibilité des données. Les méthodes d'apprentissages démontrent également une bonne performance à l'échelle du réseau mais leur utilisation n'est pas judicieuse pour des prédictions. De nouveaux travaux seront nécessaires pour considérer un processus de détérioration continu dans les modèles d'apprentissage et adapter leurs prédictions à une utilisation à long terme. En revanche, les méthodes d'apprentissage se sont avérées être les outils les plus fiables au niveau de la conduite pour identifier les conduites dans un état critique. Leur taux de vrais positifs de 67% est à rapprocher du taux de vrais positifs d'une inspection télévisée, de l'ordre de 80%. Ils peuvent donc être pertinents pour planifier les programmes d'inspections visant à identifier les conduites en mauvais état.

Ce travail de thèse s'est concentré sur l'influence de la qualité et disponibilité des données d'inspections sur la performance des modèles de détérioration. Des investigations supplémentaires seront encore nécessaires pour quantifier soigneusement toutes les sources d'incertitude, évaluer leurs propagations cumulées dans les modèles de détérioration et atténuer leur impact sur les décisions de gestion patrimoniale. En particulier, de nouvelles études seront nécessaires pour communiquer aux décideurs les résultats de la modélisation ainsi que leurs incertitudes associées. La plupart des approches de modélisation élaborées par la recherche sont des outils d'experts : l'appropriation des modèles par les collectivités nécessitera l'élaboration d'interfaces conviviales et communicatives pour faciliter les discussions avec les décideurs. Ces

interfaces devront intégrer pleinement les incertitudes associées aux prédictions pour accompagner la prise de décision.

Concernant les sources d'incertitudes, le biais de survie semble être un problème crucial pour le développement futur des modèles de détérioration (Ouellet & Duchesne, 2018). La plupart des modèles surestiment la condition réelle des réseaux car les jeux de données d'inspection sont souvent biaisés : les conduites utilisées pour l'étalonnage sont uniquement celles qui ont survécu jusqu'à la date de l'inspection, c.à.d. les conduites qui n'ont pas été remplacées avant d'atteindre leur état de dégradation actuel (Le Gat, 2008). Les modèles sous-estiment inévitablement la probabilité d'être en mauvais état et, par conséquent, surestiment la condition et la durée de vie résiduelle utile des conduites. Plusieurs travaux ont mis en évidence l'existence de ce biais sur des jeux de données synthétiques (Ouellet & Duchesne, 2018) ou en associant les modèles de détérioration avec des modèles théoriques de réhabilitation (Egger *et al.*, 2013). De nouveaux travaux seront nécessaires pour quantifier finement ce biais sur des jeux de données réels en intégrant les données des conduites déjà réhabilitées dans l'étalonnage des modèles. Il s'agira également de proposer des solutions pratiques pour corriger ce biais quand les données ne sont pas disponibles, pour éviter la présence d'erreurs systématiques dans les prédictions à long terme. Plus généralement, la prise en compte du biais de survie pourrait amener les municipalités à améliorer leurs pratiques de gestion des données. Un point de départ serait la documentation complète de chaque action de réhabilitation dans le réseau et la collecte de données sur les conduites réhabilitées (condition avant réhabilitation et durée de vie).

Les efforts de recherche se sont principalement portés sur le développement d'approches fiables pour modéliser la détérioration des réseaux. À l'avenir, la recherche devra être axée sur le développement et l'implémentation d'outils prédictifs d'aide à la décision pour soutenir les stratégies de gestion des patrimoniales. Cela implique de soutenir (i) la planification des

programmes d'inspection et des programmes de réhabilitation à court terme en se basant sur l'analyse des risques, (ii) la sélection des techniques de réhabilitation en se basant sur l'analyse du cycle de vie et (iii) la définition des stratégies d'investissement à moyen-long terme. En particulier, le développement d'outils pour la planification à long terme nécessitera de nouveaux travaux pour considérer la détérioration des conduites rénovées (chemisage) ou réparées. Les techniques de rénovation sont de plus en plus populaires pour réduire les coûts de réhabilitation des conduites ; cependant, l'influence de ces pratiques sur l'évolution de la condition future des réseaux est mal connue et ne peut être simulée finement en l'absence d'études.

Enfin, les réseaux d'assainissement s'intègrent dans un complexe d'infrastructures urbaines qui inclut entre autres (i) les réseaux d'eau potable, de chaleur et d'électricité en souterrain et (ii) la voirie en surface. Les interactions entre ces infrastructures et notamment le potentiel de synergie pour la réhabilitation intégrée des infrastructures sont mal comprises. Les pratiques varient beaucoup d'une collectivité à l'autre ; il est souvent admis que la réhabilitation intégrée peut réduire les coûts globaux de réhabilitation et diminuer les nuisances aux riverains en mutualisant les travaux de plusieurs infrastructures. Plusieurs travaux ont abordé le sujet sous un angle théorique en utilisant des modèles de détérioration (Carey & Lueke, 2013; Nafi & Kleiner, 2010; Marzouk & Osama, 2015; Tscheikner-Gratl *et al.*, 2015; Tscheikner-Gratl *et al.*, 2016). Même si les premières études sont prometteuses, de nouveaux travaux seront nécessaires pour quantifier les avantages et inconvénients de cette coopération et proposer des outils pratiques à destination des municipalités pour améliorer la gestion patrimoniale intégrée des infrastructures urbaines. De prochains travaux de recherche pourront notamment viser à identifier le juste niveau de coopération pour maximiser les objectifs de la collectivité ou municipalité.

Chapter 2 Introduction

2.1 Background

2.1.1 The funding gap in sewer infrastructure

Urban drainage systems including collection pipes and treatment facilities represent an important investment in physical assets for a city. A typical city of 100,000 to 500,000 inhabitants may have between 1,000 km and 5,000 km of underground network for wastewater and storm water collection. The replacement cost of these assets could amount to millions of euros not only for major cities, but for small municipalities as well (IPK, 2014). Over the last 30 years, most municipalities have invested in sewer system expansion to meet growth and treatment plant upgrades, but a relatively small proportion of the budget has been allocated to sewer rehabilitation (AWWA, 2012). As a result, most cities face the problem of an aging infrastructure in need of extensive and ongoing repair, renovation or replacement.

Insufficient public and municipal investment represent a major challenge for the long-term management of ageing urban drainage systems in Europe, Australia and North America but also for growing networks in Asia, South America and Africa. More than a decade ago, the American Water Works Association estimated that a new era was dawning: the replacement era in which the country will need to rehabilitate massively the water and sewer networks built by the previous generations (AWWA, 2012). In many cities worldwide, the underground infrastructure is nearing the end of its technical lifetime and will reach soon the age of renewal. A recent survey among 397 water and wastewater industry participants in the USA and Canada highlighted that the aging of infrastructures and the management of capital and operational costs are the two main industry issues (Black and Veatch, 2013). The ASCE estimated the required capital investment to maintain and upgrade water infrastructure in the USA at \$91 billion (ASCE, 2011). However, only \$36 billion of this \$91 billion needed was funded, leaving a

capital funding gap of nearly \$55 billion. Water infrastructure is clearly aging, and the investments planned are not able to keep up with needs (ASCE, 2011). In Germany, a recent national study highlighted that 20% of the sewer network has severe defects that require short or mid-term rehabilitation (Berger *et al.*, 2015). Over the last years, the annual investment for sewer rehabilitation was about 4 billion € whereas capital need is estimated to be more than 7 billion €, indicating a capital deficit of at least 3 billion € (IPK, 2014; KfW, 2016). Delaying further the investment will result in degrading water and drainage services, escalating flood risk, increasing environmental impacts and raising expenditures for emergency repairs.

The funding gap and the investments needed to cope with sewer deterioration will necessarily lead to future increases of the water tariff (Oelmann *et al.*, 2017). Utilities are challenged to develop efficient rehabilitation strategies in order to keep the same level of service. Traditionally it has been economically feasible to apply reactive management strategies, repairing mainly when failures occur; however, this strategy will become less viable as the systems age and the funding gap increases (Rokstad & Ugarelli, 2015). In this context, a promising leverage of utilities is the improvement of technical asset management and, in particular, the use of modelling solutions to improve the efficiency of inspection and rehabilitation strategies. Utilities often lack appropriate tools to plan and manage long-term investment needs (Black and Veatch, 2013) and rely on reactive strategies, rehabilitating mainly when failures occur.

2.1.2 Sewer condition assessment

Sewer asset management can be defined as managing infrastructure capital assets to minimize the total cost of owning and operating them, while delivering the service levels customers desire (EPA, 2002). A key element of asset management programs is an efficient rehabilitation and replacement strategy. Technical needs for sewer replacement are evaluated mainly based on the

structural and the hydraulic performance of the network, the structural performance being the most dominant aspect for budget allocation by asset managers (van Riel *et al.*, 2014). van Riel *et al.* (2016) analysed the information sources of 150 sewer replacement projects in the Netherlands; the main information used by operators was found to be Closed Circuit Television Inspection (CCTV) in 60% of the cases, followed by pipe age (30% of the cases), planning of urban development (25% of the cases) and road works (20% of the cases). Several sources of information are often considered together by the operators to plan rehabilitation actions (e.g. CCTV and opportunity of road works).

CCTV is applied since the 1980's as industry standard for sewer system inspection and the main source of information for structural performance evaluation. It provides visual data (images or videos) of the internal surface of the inspected pipe (the term "pipe" indicates here a pipe segment from manhole to manhole). The manual or automatic analysis of the images enables to identify the type and location of defects like offset joints, cracks, leaks, sediment, debris and root intrusion. Generally, the camera is mounted on a tractor or crawler, which enables the camera system to drive through the sewer pipe and record the entire pipe section. The main limitation of CCTV inspection is that it only provides a visual representation of the interior pipe surfaces; it cannot assess external voids, deteriorated bedding conditions, pipe wall integrity and mechanical strength. Despite these drawbacks and because of the lack of widely accepted and cost-effective alternative, CCTV is still the state-of-the-art technology commonly used by sewer operators to assess sewer structural condition and plan rehabilitation programs. Inspection is also the only technology with national or international standard providing guidance (such as for example EN 13508-2, 2011).

Only few legal requirements exist regarding sewer inspection frequency. In France, each utility decides the annual inspection needed for its asset. In Germany, local regulations commit sewer operators to inspect their network regularly. Inspection frequency is defined in the self-

monitoring ordinance of each region. For example, operators in North Rhine-Westphalia have to inspect their entire network within 15 years and at least 5% of the network each year (SüwVO Abw, 2013).

In most countries, pipe defects recorded during CCTV inspections are manually coded according to standard coding systems and the overall sewer condition is assessed using an automatic classification methodology (Figure 2.1).

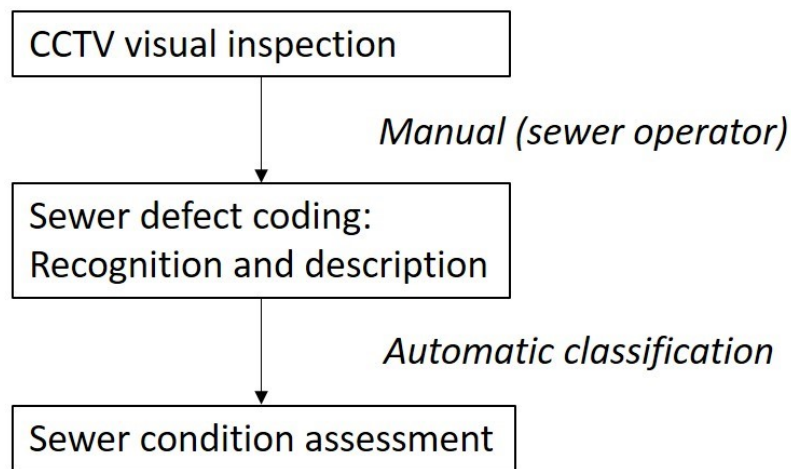


Figure 2.1. Workflow describing the sewer condition assessment procedure as basis for the prioritization of sewer rehabilitation

Sewer defect coding

The defect codification is the documentation of the CCTV inspected sewer. It is often performed manually by the inspection staff. It describes the inspected sewer defects with standard codes and asset information. In Europe, the current codification system is the normative EN 13508-2 (2011) for visual inspection. Observed defects are coded with letters on three positions and a numerical value is added to quantify the defect. The first position indicates the main code that describes the observed defect (e.g. BAB for a fissure). The second and third positions can be used to indicate the defect characterization (e.g. B-A for a crack in longitudinal direction or B-B for a crack which is around the circumference of the pipe). Other information are given such as the circumferential location or the longitudinal location.

2.1 Background

Due to the labour-intensive and error-prone manual detection and interpretation of pipe defects, recent research projects have intended to automate this procedure (Halfawy *et al.*, 2014; Kumar *et al.*, 2018; Müller *et al.*, 2007; Li *et al.*, 2019). The proposed approaches are not commonly used by municipalities because they were not successful enough or not fully validated using visual data acquired from actual sewer inspections (Guo *et al.*, 2009).

Sewer condition classification

The primary interest of condition assessment methodologies is to transfer the extensive amount of visual inspection data into an easily manageable condition class, useful to support asset management. Many methodologies exist for sewer condition classification: e.g. RERAU in France; SRM in the UK; PACP in the US; DWA M149-3 in Germany (see Kley *et al.*, 2013a, for a review of sewer classification methods). These methodologies typically use algorithms to assess the importance of each sewer defect and aggregate them in order to obtain an assessment of the overall condition of each inspected pipe for different requirements (e.g. structural and operational condition). Such condition should be then combined with complementary performance indicators (Le Gauffre *et al.*, 2007) to prioritize rehabilitation needs and support the definition of rehabilitation programs.

Uncertainties in sewer condition assessment

In current practice, CCTV data are crucial to support asset management decisions. However, the quality and uncertainty of sewer condition assessment is rarely questioned. Dirksen *et al.* (2013) published a comprehensive analysis of the accuracy and reliability of data obtained from visual sewer inspection. The authors highlighted the high subjectivity of the inspection procedure at three main steps: (1) the recognition of defects, (2) the description of defects and (3) the evaluation of sewer condition. It was found that the probability that an inspector fails to recognize the presence of a defect (False Negative FN) is significantly higher than the

probability that a defect is reported although it is not present (False Positive FP). The probability of a FP is in the order of a few percent whereas the probability of a FN is in the order of 25%. The probability of an incorrect observation using the norm EN 13508-2 can be very high when considering all defects because the level of details required by the coding system cannot be accomplished by visual inspection. Further it was shown that individual inspectors arrive at different results when evaluating a given set of CCTV reports, thereby highlighting the subjectivity of interpreting images.

Hüben (2002) analysed condition classes from repeated CCTV sewer inspections of a German city. The results showed that over 50% of the sewers changed of condition classes between the repeated inspections. He concluded that the uncertainties in the defect recognition and description are propagated in the assessment of the sewer condition class and significantly influence the results. Sousa *et al.* (2014a) also quantified CCTV uncertainties by comparing periodic inspection reports from three trunk sewers of a Portuguese sewer system. Over the 25 km of inspected pipes, 25% of the sewer pipes had different structural condition ratings between the repeated inspections. The authors highlighted a high degree of consistency of the inspectors regarding the most severe defects.

2.1.3 Sewer condition prediction

Due to budget restrictions, inspection rates are generally low and municipalities tend to inspect only a small part of their network (Harvey & Mc Bean, 2014; ONEMA, 2012). Few data exist on the national inspection rates of countries worldwide, but many publications highlight the low availability of CCTV data, e.g. in Colombia (Hernández *et al.*, 2018), in Canada (Harvey & Mc Bean, 2014), in France (ONEMA, 2012), in Portugal (Sousa *et al.*, 2014b) and in Australia (Tran, 2016). In France, the annual inspection rate of most municipalities does not exceed 5% (Ahmadi *et al.*, 2014c). In Germany, inspection rates are usually higher due to local regulations which commit sewer operators to inspect their network regularly (StwVO Abw,

2013). Furthermore, CCTV data as such (without further analysis) are insufficient to determine long term asset management strategies since they provide only a snapshot of the sewer condition at the date of inspection and no information on the expected remaining lifetime of the sewers.

Over the last decades, modelling has gained an increasing importance to assist proactive management due to a better data availability, the possibility to relate several data sources, the increase of computational power and the development of operational software. Modelling tools support utilities in addressing a broad range of issues such as sewer deterioration (Ana & Bauwens, 2010; Egger *et al.*, 2013; Kleiner *et al.*, 2006; Salman, 2010), selection of rehabilitation technique (Das *et al.*, 2018), infiltration and exfiltration (Bertrand-Krajewski *et al.*, 2005), flood risk (Dey & Kamioka, 2007; Yazdi *et al.*, 2015), sewer blockage (Jin & Mukherjee, 2010), sediment deposition (Ashley *et al.*, 2000; Rodríguez *et al.*, 2012) or combined sewer overflows (Morales *et al.*, 2017).

Since the definition of rehabilitation strategies is generally limited by the lack of information about sewer condition and remaining life, deterioration models have been developed to forecast the evolution of the system given its current and past condition. Deterioration models can be used (i) to presume the condition class of non-inspected pipes and (ii) to forecast the evolution of the system's condition. Such model outputs provide key information to operators and municipalities for the scheduling of inspection programs (i.e. the detection of sewers in critical condition) and the planning of rehabilitation budgets (i.e. the comparison of different sewer rehabilitation scenarios and the evaluation of necessary investment rates). Deterioration models can support operators and municipalities in defining mid to long-term asset management strategies with only limited availability of CCTV data. They may also help validating and showing the viability of current strategies or providing information to justify the relevance of additional investments and expenditures. Several modelling approaches are available; for a

detailed review of deterioration modelling approaches, you may refer to Kley & Caradot, 2013b; Ana & Bauwens, 2010; Marlow *et al.*, 2009; Rokstad & Ugarelli, 2015; Santos *et al.*, 2017. Models are generally classified in three groups: deterministic, statistical and machine learning models.

Deterministic models

Deterministic models aim at understanding the physical mechanisms that drive sewer deterioration. However, even if a single dysfunction such as corrosion could be modelled physically, sewer deterioration as a whole remains a very complex process that is not completely understood and depends on a large number of factors (Davies *et al.*, 2001; El-Housni *et al.*, 2018; Nariné Torres *et al.*, 2017; Schmidt, 2009). Even sophisticated deterministic models are often too simplistic to reflect the complexity of the deterioration process and the scarcity of available data needed to simulate deterioration mechanisms (initial condition, process parameters, etc.) decreases the applicability of such models (Ana *et al.* 2009; Kleiner & Rajani, 2001a).

Statistical models

Statistical models use mathematical relationships to relate the history of sewer conditions with sewer deterioration factors. Outcomes of statistical models are expressed as probability values of pipes condition; uncertainties can be assessed in the form of probability density functions. Survival analysis and Markov-chain are the most common types of statistical deterioration models (Baik *et al.*, 2006, Baur & Herz, 2002, Caradot *et al.*, 2017, Duchesne *et al.*, 2013; Duchesne *et al.*, 2014, Kleiner & Rajani, 2001a, Micevski *et al.*, 2002, Rokstad *et al.*, 2014, Tran *et al.*, 2007, Werey *et al.*, 2012, Wirahadikusumah *et al.*, 2001). Survival analysis and Markov-chain are often used in combination to quantify pipe condition probabilities given a set of explanatory factors. Prior to model calibration, pipes are generally grouped in cohorts, i.e.,

homogenous groups of sewer pipes sharing similar features, e.g., same material and type of effluent, and assumed to have a similar deterioration behaviour. During the calibration procedure, survival functions are estimated for each cohort. Survival curves have the mathematical form of a given statistical distribution (e.g. GompertZ, Weibull, and Exponential) and are calibrated by a regression procedure using e.g. the method of Maximum Likelihood Estimation (Le Gat, 2008). They represent the mean deterioration of pipes over time: they define the proportion of pipes that have survived at a given age. Additionally, the shape of the survival curves can eventually be modulated by further numerical or categorical variables. Markov-chains are used to simulate the probability for a pipe to be in a new condition at year $t+dt$ given its condition at year t and a transition matrix Q .

$$P(t + dt) = Q(t) P(t)$$

The transition matrix can be mathematically derived from the slope of the survival curves (Le Gat, 2008). Depending on the model formulation, the elements of the matrix can be time-independent (homogeneous Markov chain) or time-dependent (non-homogeneous Markov chain).

For example, the GompitZ model developed by (Le Gat, 2008) is based on a Non-Homogenous Markov Chain approach and can consider explanatory factors to modulate the transition matrix. This model has been successfully deployed in Germany (Baur & Herz, 2002; Caradot *et al.*, 2017; Le Gat, 2008), Norway (Rokstad & Ugarelli, 2015) and France (Werey *et al.*, 2012).

Regression methods also have been successfully used to determine the probability of failure of individual pipes (Ahmadi *et al.*, 2014a, 2015; Ariaratnam *et al.*, 2001, Chughtai & Zayed, 2008, Fuchs-Hanusch *et al.*, 2015, Salman & Salem, 2012, Tscheikner-Gratl *et al.*, 2016). In Canada, Ariaratnam *et al.* (2001) developed a binary logistic regression for the prediction of the likelihood that a sewer is in a deficient state for the sewer network of Edmonton. Chughtai &

Zayed (2008) used a multiple regression model to simulate the condition state of sewers using data from two Canadian municipalities, Pierrefonds and Niagara Falls. In Austria, Fuchs-Hanusch *et al.* (2015) developed a logistic regression model to predict sewer condition and each sub-type of defects such as collapse or blockage. According to the authors, the performance of the defect-based model is better than the condition-based model, suggesting that model performance can be improved if the prediction is based on the CCTV reports rather than on the assessed conditions.

Machine learning models

Unlike statistical models, machine learning models do not require assumptions about the model structure, they are purely data or information-driven. Model outputs are classified from a set of input variables by learning from the available data. Their advantage is that they could mimic the complex and non-linear relationships between explanatory variables and sewer condition states by “learning” the deterioration behaviour of pipes from inspection data (Scheidegger *et al.*, 2011; Ana & Bauwens, 2010). Therefore, the information gained on the available inspection data is generalized to non-inspected pipes. Main machine learning methods applied in deterioration models are Random Forest (Harvey & McBean, 2014, Rokstad & Ugarelli, 2015, Syachrani *et al.*, 2013, Vitorino *et al.*, 2014), Support Vector Machines (Mashford *et al.*, 2010, Hernández *et al.*, 2017, Sousa *et al.*, 2014b), and Neuronal Networks (Jiang *et al.*, 2016, Najafi *et al.*, 2005, Sousa *et al.*, 2014b, Tran *et al.* 2006).

Harvey & McBean (2014) used Random Forest to predict sewer pipes in poor condition in Ontario, Canada. They found that the predictive performance can be improved by including in the modelling dataset upstream and downstream condition information of the adjacent manholes. Vitorino *et al.* (2014) found that Random Forest provides an efficient balance between calculation time and predictive performance and that the most useful feature is the

ability to explore an open range of explanatory factors. In Finland, Laakso *et al.* (2018) outlined that the analysis of partial dependencies in Random Forest model is a valuable means of visualizing the connection between predictor variables and poor condition. Rokstad & Ugarelli (2015) used Random Forest with data of the city of Oslo, Norway. They concluded that deterioration model application may be beneficial for prioritizing inspection programs and that the performance is limited by the adequacy of the explanatory variables available.

Tran *et al.* (2007) tested an artificial neural network to simulate the condition of storm water pipes in Victoria. They found that Bayesian weight estimation is better than the conventional backpropagation weight estimation. In Australia, Jian *et al.* (2016) developed a neural network for modelling the concrete corrosion processes in sewers. The model showed excellent performance with an R^2 of 0.89 on an independent test dataset and outperformed a multiple regression model. Sousa *et al.* (2014b) tested several types of neural networks in Portugal. They found that beside the type of neural network used, the initial weights of the neuron connections introduce a significant variability into the results.

Mashford *et al.* (2010) applied Support Vector Machines (SVM) to predict the structural condition of uninspected sewer pipes in the sewer system of Adelaide. They argue that SVM have advantages over conventional machine learning classifiers such as neural networks: (i) lack of necessity for specification of internal architecture and (ii) the ability to be trained successfully on smaller training sets than those required by neural networks. Sousa *et al.* (2014b) highlight that SVM is sensible to the values of the parameters of the kernel, the type of kernel used and the capacity.

Modelling performance

Recent research initiatives have investigated the potential of deterioration modelling to support sewer inspection strategies. In France, Ahmadi *et al.* (2014a, 2015) analyzed the benefits of

using regression-based modelling to plan inspection strategies. Inspection programs based on modelling results were at least twice more efficient than random-based programs: they were able to identify at least twice the number of pipes in poor condition. Rokstad & Ugarelli (2015) in Norway, Fuchs-Hanusch *et al.* (2015) in Austria and Harvey & McBean (2014) in Canada also highlighted the relevance of modelling approaches to detect sewers in critical condition and schedule inspection planning.

Few studies evaluated the performance of deterioration models **at the network level**, i.e. the ability of the model to simulate the condition distribution of the network (Duchesne *et al.*, 2013, Hernández *et al.*, 2018, Ugarelli *et al.*, 2013). They compared the proportions of observed and predicted pipes in a given condition and highlighted relatively good modelling performance with deviations lower than 10%.

Most studies assessed model performance **at the pipe level**, i.e. the ability of the model to identify pipes in poor condition. Main metrics found in the above presented studies are statistical metrics (chi-square statistic, Goodness-of-fit, Root Mean Square Error, Coefficient of determination, etc.) and a list of indicators derived from the confusion matrix, including ROC curves. Modelling performance varies a lot between the case studies. Statistical indicators have the advantage that they are standardized and scientifically recognized but they have the main drawback that they do not allow outlining clear and understandable conclusions adapted to utilities objectives (Caradot *et al.*, 2018b).

Indicators from the confusion matrix and ROC curves (Brown *et al.*, 2006) are generally used to assess model performance at the pipe level (Harvey & Mc Bean, 2014; Sousa *et al.*, 2014b; Fuchs-Hanusch *et al.*, 2015). The confusion matrix is a simple way to summarize the number of correctly and incorrectly predicted observations (Table 2.1).

Table 2.1. Structure of a confusion matrix

		Predicted condition	
		0 (good)	1 (poor)
Observed condition	0 (good)	TN	FP
	1 (poor)	FN	TP

Several indicators can be derived from the matrix such as the

- TPR, True Positive Rate, i.e. the percentage of pipes observed in poor condition and correctly predicted in poor condition: $TP / (FN + TP)$
- PPV, Positive Predictive Value, i.e. the percentage of pipes predicted in poor condition, which have been actually observed in poor condition: $TP / (FP + TP)$
- FPR, False Positive Rate, i.e. the percentage of pipes observed in good condition and wrongly predicted in poor condition: $FP / (FP + TN)$
- Accuracy, i.e. the percentage of correct predictions: $(TP + TN) / (TN + FN + FP + TP)$

ROC curves are built by plotting TPR against the FPR for different cut-off values. The accuracy can be a misleading indicator: it assesses the overall performance of the models without exploring the single performances for the prediction of each condition class. Since the condition of sewer networks is mostly imbalanced (many pipes in good condition and few pipes in poor condition), the accuracy can be excellent even if the True Positive Rate is very low. On the other hand, assessment of only one indicator, such as the True Positive Rate, can also be misleading: a model that predict all pipes in poor condition would have 100% True Positive Rate but would not be a satisfying model. In conclusion, the combined analysis of TPR, PPV and FPR on an independent test dataset seem essential to fully describe the predictive performance of a model. Table 2.2 shows the value of these indicators from a list of previous studies that publish the confusion matrix of the outcomes or the values of the indicators. Model performance varies a lot between the case studies (Table 2.2).

Table 2.2. Predictive performance of deterioration models

Reference	Model	Condition assessment	City	Country	Number of data Training/Test	PPV Positive predictive value	TPR True Positive Rate	FPR False Positive Rate
Salman & Salem, 2012	Multinomial log. regression	PACP	Cincinnati	USA	11,373 -80/20	53%	73%	29%
	Binary logistic regression	PACP	Cincinnati	USA	11,373 -80/20	55%	45%	22%
Hernández <i>et al.</i> , 2017, 2018	Random Forest	NS-058	Bogota	Colombia	4,633 - 70/30	53%	57%	17%
	Logistic Regression	NS-058	Bogota	Colombia	4,633 - 70/30	60%	38%	7%
	Multinomial logistic regression	NS-058	Bogota	Colombia	4,633 - 70/30	-	71%	21%
	Linear Discriminant Analysis	NS-058	Bogota	Colombia	4,633 - 70/30	-	70%	20%
	Support Vector Machine	NS-058	Bogota	Colombia	4,633 - 70/30	52%	67%	22%
Laakso <i>et al.</i> , 2018	Random Forest	Finnish guidelines	-	Finland	6,700 -70/30	-	80%	53%
Harvey & Mac Bean, 2014	Random Forest	WRC	Guelph	Canada	1,825 - 80/20	30%	89%	25%
Sousa <i>et al.</i> , 2014b	Artificial Neural Network	WRC	Costa do Estoril	Portugal	745 - 80/20	67%	71%	18%
	Support Vector Machine	WRC	Costa do Estoril	Portugal	745 - 80/20	69%	60%	19%
	Logistic Regression	WRC	Costa do Estoril	Portugal	745 - 80/20	62%	39%	16%
Mashford <i>et al.</i> , 2010	Support Vector Machine	-	Adelaide	Australia	1,441 - 75/25	88%	74%	1%
Fuchs-Hanusch <i>et al.</i> , 2015	Logistic Regression	ISYBAU	-	Austria	4,577 - 62/38	-	60%	35%

- The average PPV is 57% meaning that in average 57% of the pipes predicted in poor condition are actually in poor condition. PPV ranges between 30% and 88%.
- On the other hand, the average TPR is 64% meaning that in average 64% of the pipes in poor condition are correctly identified by the model. TPR ranges between 38% and 89%.
- Finally, the average FPR is 22% meaning that in average 22% of pipes in good condition are wrongly predicted in poor condition. FPR ranges between 1% and 53%.

From these outcomes, it is not possible to outline clear conclusions regarding the best modelling approach at the pipe level, especially because modelling performance is a trade-off between several indicators and data availability differs between the case studies. However, the benchmark of several models obtained on the same cities showed that machine learning models seem to outperform statistical models to identify pipes in critical conditions (Hernández *et al.*, 2018; Sousa *et al.*, 2014b).

Uncertainties in sewer condition prediction

Recent studies focused on quantifying the sources of uncertainties which might affect model predictions. They can be summarized as follows:

- **The availability of CCTV data for model calibration** (Ahmadi *et al.*, 2016; Duchesne *et al.*, 2013; Rokstad & Ugarelli, 2015; Tran, 2016).

Duchesne *et al.* (2013) performed cross-validation tests using various sample sizes (from 250 to 4,000 pipes) for model calibration. They showed that statistical models remain robust even when a small proportion of pipes is used for model calibration, provided that the inspected pipes are representative of the pipe characteristics for the whole network. Tran (2016) highlighted that a minimum sampling size between 600 and 700 pipes is recommended for utilities to collect condition data in order to provide

a starting view on deterioration patterns with less than 10% error rate. Ahmadi *et al.* (2016) showed that a sample whose size of at least 1,000 random pipes could be representative of the asset stock.

- **The availability of historical records of rehabilitated segments for model calibration** (Egger *et al.*, 2013; Duchesne *et al.*, 2013; Ouellet & Duchesne, 2018; Scheidegger *et al.*, 2013)

The lack of availability of historical record generally leads to the development of a survival bias in modelling prediction (Le Gat, 2008). Current models are expected to overestimate the real condition of the network (i.e. give a too optimistic image) because the observed pipes used for model calibration are only those that “survived” until the date of inspection, i.e. pipes that were not replaced before they reach their current degradation state (Le Gat, 2008; Ouellet & Duchesne, 2018). Since the models are calibrated using available data concerning only pipes that were in place during the inspection period, they will inevitably underestimate the probability to be in a poor state, and consequently, overestimate the duration of useful life of pipes. Egger *et al.* (2013) propose to combine the deterioration model with a probabilistic replacement model that characterizes the probability that a pipe was not replaced, i.e. the chance that a pipe is still in service. This type of approach is, reportedly, already successfully applied for water networks (Large, 2016; Le Gat *et al.*, 2013). Ouellet & Duchesne (2018) proposed a method to recreate missing information about replaced pipes by using the characteristics of pipes still in operation.

- **The availability and quality of appropriate explanatory factors for model calibration** (Ahmadi *et al.*, 2015; Mashdorf *et al.*, 2010)

Ahmadi *et al.* (2015) showed that having data influenced by high uncertainties is preferable to having incompleteness within the utility database. For example, having

age in an imprecise form of four age groups was nearly as informative as having age in a scale form. The consideration of material or road class along with pipe age improves the efficiency of inspection programs by 15% (i.e. the ability of the model to correctly identify pipes in poor condition).

2.2 Research needs

2.2.1 The assessment of modelling performance

As the benefits of deterioration modelling to support inspection programs are now better understood by scientists and practitioners, research efforts are still needed to investigate the performance of deterioration models and the benefits of modelling for planning mid to long-term rehabilitation budgets (Alegre & Matos, 2009; WERF, 2012). Deterioration models are still not commonly used by sewer operators and municipalities to support strategies and there is still a major gap between utilities' day-to-day management practices and the available modelling approaches. No data is available for the European market, but a national survey in the USA highlights that less than 10% of municipalities and utilities are currently using deterioration models to support their asset management strategies (Black and Veatch, 2013). One of the main factors hampering the uptake of deterioration modelling by utilities is the lack of real scale evidence of the tangible benefits provided (Scheidegger *et al.*, 2011; WERF, 2007).

Many studies cited above (section 2.1.3) intended to evaluate the performance of statistical and machine learning deterioration models. The outcomes of these studies underline the relevance of using deterioration models to support asset management strategies but suffer from two main shortcomings.

- The lack of data for model calibration: models are often calibrated with less than 2,000 pipes which represent only a small part of the networks (except for the studies of Salman (2010) and Rokstad & Ugarelli (2015) which developed deterioration models in the

cities of Cincinnati in the USA and Oslo in Norway using respectively 11,373 and 12,003 pipes, respectively). Further very recent studies in Austria (Fuchs-Hanusch *et al.*, 2015), Colombia (Hernández *et al.*, 2018) and Finland (Laakso *et al.*, 2018) also provided evidences using larger inspection datasets.

- The lack of clear metrics adapted to utilities' issues for the assessment of model performance: most metrics are based on statistical tests (e.g. Mean Square Error, goodness-of-fit and coefficient of determination) and do not provide a full understanding of the potential of deterioration models for municipalities and sewer operators. Furthermore, the metrics often assess the overall performance of the models without exploring the single performance for the prediction of each condition class. Since the condition of sewer networks is mostly imbalanced (many pipes in good condition and few pipes in poor condition), this assessment can lead to biased conclusions.

In order to cope with this limitation, we propose to investigate the following research question:

Research question

- R1: what is the performance of sewer deterioration modelling in a case of high CCTV data availability?
 - How to assess the quality of prediction of a sewer deterioration model?
 - How to define a relevant and understandable set of metrics to communicate the performance of sewer deterioration models with utilities?

2.2.2 The influence of CCTV availability on modelling performance

As presented above, inspection rates are generally low and municipalities tend to inspect only a small part of their network due to budget restrictions. Most utilities are concerned by the

minimum amount of CCTV data required and the relevance of using such models on their networks with limited data availability. The assessment of the influence of CCTV data availability on the reliability of deterioration modelling is still a key step to build the trust of utilities regarding modelling outcomes. This corresponds to our second research question:

Research question

- R2: what is the influence of CCTV data availability on the performance of sewer deterioration modelling?
 - What is the minimum required amount of data to achieve accurate modelling outcomes?

2.2.3 The influence of CCTV uncertainty on modelling performance

The reliability of CCTV has been strongly questioned these last years in the scientific literature (Caradot *et al.*, 2018a; Dirksen *et al.*, 2013; Korving and Clemens, 2004; van Riel *et al.*, 2017; Roghani *et al.*, 2019; Sousa *et al.*, 2014a; van der Steen *et al.*, 2014). According to these studies, condition assessment based on CCTV tends to underestimate the level of deterioration of segments. Condition assessment errors originate from (Cherqui *et al.*, 2017): (1) the environment of the pipe (e.g. obstacles which hinder the accurate visualisation); (2) the condition assessment process (e.g. coding system used, experience and subjectivity of the operator); (3) the pipe characteristics (e.g. diameter, material); (4) and the defect characteristics (e.g. size, number, spatial distribution). Most utilities acknowledge the uncertainties in the procedure of sewer condition assessment, mainly due to the subjectivity of the operator. There is still a need to quantify the uncertainty of the sewer condition assessment procedure and its influence on the outcomes of deterioration modelling. These are our third and fourth research questions:

Research questions

- R3: what is the uncertainty of the sewer condition assessment procedure?
- R4: what is the influence of CCTV data uncertainty on the performance of sewer deterioration modelling?

2.3 Aim and structure of the thesis

This thesis aims at filling the gap presented above in chapter 2.2 by addressing the series of research questions. Figure 2.2 visualizes the links between the four research questions in the context of sewer asset management.

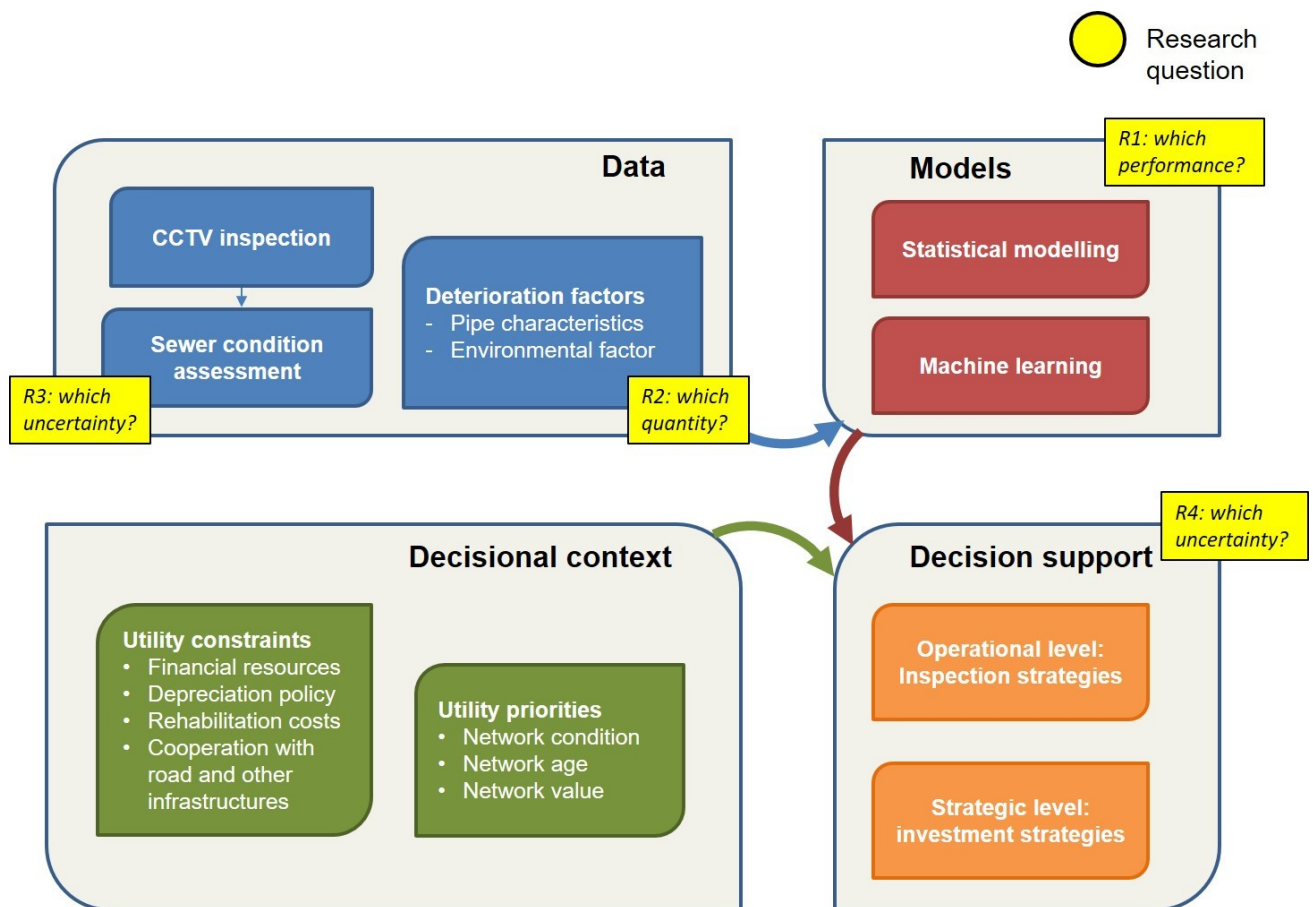


Figure 2.2. Structure of the thesis and articulation of the research questions

It presents the four main components of the development of an asset management strategy: data

are collected to gather knowledge on the pipes condition and characteristics; these data are used to calibrate deterioration models which simulate the condition of the pipes (including non-inspected pipes); such models are combined with the decisional context (mainly translated as utility constraints and priorities) to build the decision support information for the planning of inspection and rehabilitation strategies.

The thesis is structured around four research papers.

Paper I addresses the research questions R1 and R2 (in chapter 3). It aims at

- Assessing the ability of a statistical deterioration model to simulate the condition of a sewer network and
- Evaluating the influence of CCTV data availability on modelling performance.

A statistical model has been applied using the extensive GIS and CCTV database of the city of Braunschweig in Germany. This city is an ideal case study, since the entire sewer system has already been inspected once and over 50% of the sewers have been inspected at least twice. The performance of the model has been assessed by analysing the deviations between predicted and inspected condition distributions. Then the performance of the model has been evaluated with different random subset sizes n (number of pipes) for model calibration ranging from $n = 0.2\%$ (about 70 pipes) to $n = 100\%$ (35,826 pipes).

Paper II addresses the research question R1 (in chapter 4). It aims at

- Developing a set of metrics adapted to the local utility needs to assess the performance of sewer deterioration models from an end-user perspective. The metrics must be intuitive, self-explanatory and thus clearly understandable by the sewer operator. Thus, they should be able to convince the utility about the relevance or uselessness of using

deterioration models to support asset management strategies. The latter point is regarded as being crucial for facilitating the communication of the outcomes and ensuring the acceptance of the results.

- Applying the set of developed metrics to benchmark the performance of a statistical and a machine learning deterioration models trained with the extensive CCTV dataset of the sewer of Berlin, Germany. In Berlin, each pipe of the almost 10,000 km sewer network has been inspected at least once by the end of 2016. The inspection database is a highly valuable knowledge that can be exploited to assess model performance in the special case of full data availability.

The statistical approach selected is the model GompitZ (Le Gat, 2008), based on the theories of survival analysis and Markov-chains. The machine learning approach selected is Random Forest, an ensemble learning method for classification or regression, already successfully implemented for the prediction of sewer pipes condition (Harvey & McBean, 2014; Rokstad & Ugarelli, 2015; Vitorino *et al.*, 2014). Outcomes of the benchmark analysis will be used to conclude on the strengths and weaknesses of both approaches regarding asset management objectives.

Paper III addresses the research question R3 (in chapter 5). It aims at

- Quantifying the uncertainty of the whole inspection procedure from CCTV inspection to condition assessment
- Determining the probability to underestimate, overestimate or accurately estimate the real condition of a pipe using CCTV inspection

This study introduces and demonstrates a methodology based on double inspections of the same pipes to determine the uncertainties of the structural condition assessment, i.e. the probability to estimate correctly the structural condition of a pipe from a CCTV inspection. The

methodology is based on a non-linear optimization procedure coupled with a Monte-Carlo simulation and has been used to determine the probabilities of False Positive and False Negative when inspecting a pipe in a given condition.

Paper IV addresses the research question R4 (in chapter 6). It aims at

- Assessing the influence of sewer condition uncertainty on the shape of the survival curves and the prediction outcomes of a deterioration model.

First, the methodology proposed in paper III has been applied to quantify uncertainties in sewer condition assessment from the analysis of a set of repeated inspections. Then, a method has been proposed to propagate uncertainties in the survival curves and predictions of a statistical deterioration model. The deterioration model has been used to simulate long-term strategies and evaluate the impact of uncertainties over model prediction. The method has been demonstrated using the unique inspection dataset of the city of Berlin, Germany, where 13,753 segments have been inspected at least twice.

Table 2.3 illustrates the link between the papers, chapters and research questions addressed.

Table 2.3. Link between papers, chapters and research questions addressed

Paper	Chapter	Research question addressed
1	3	R1 (model performance) and R2 (influence of data quantity on model performance)
2	4	R1 (model performance)
3	5	R3 (CCTV uncertainty)
4	6	R4 (influence of CCTV uncertainty on model performance)

2.4 Case studies

This research had been conducted using the extensive dataset of the cities of Braunschweig and Berlin in Germany. The following sub-sections present briefly the case studies and the available GIS and CCTV data. More details are given in the following chapter of the thesis.

2.4.1 Braunschweig

The city of Braunschweig is in Lower Saxony, Germany (Figure 2.3). It has a population of around 250,000 inhabitants.

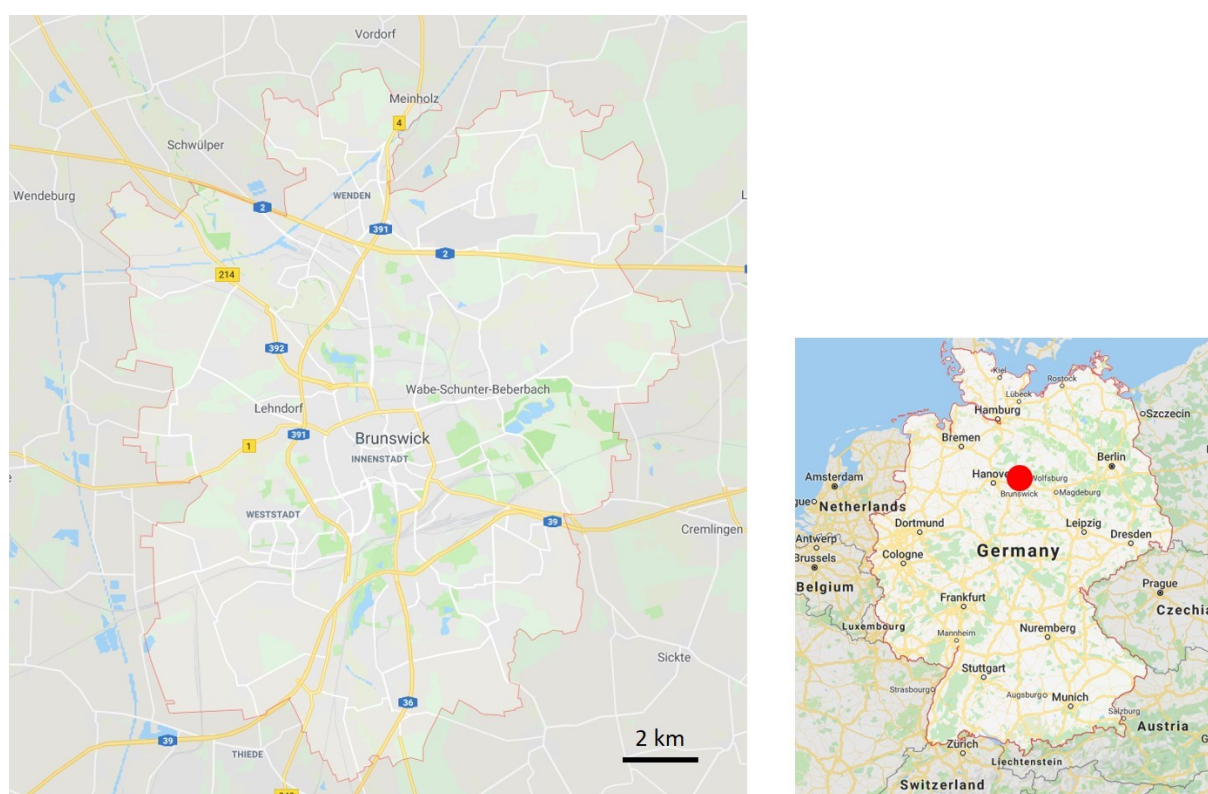


Figure 2.3. Map of the city of Braunschweig (source: google.com/maps)

The sewer network of Braunschweig is operated by SE|BS (Stadtentwässerung Braunschweig). It has a length of about 1,300 km (SE|BS, 2019). The sewer system is mostly separated: 53% of the pipes are storm water sewers whereas 44% are sanitary sewers. Clay and concrete (incl. reinforced concrete) are the two dominating materials with a share of more than 90% of the network length. Almost every sanitary pipe is made out of clay (99%) whereas most storm

water pipes are made of concrete and reinforced concrete (99%). Pipes with rarely used materials (e.g. brick, plastic, asbestos) have been excluded from the dataset, as well as non-circular pipes (3%) and pipes with missing characteristics. In order to remove (or at least reduce) the influence of the survival bias in model calibration, old segments (age > 80 years old) have been removed from the dataset. The final dataset contains 25,787 sewer pipes with a total length of 1,027 km. The distribution of pipe characteristics and environmental features for the sewer network of Braunschweig can be seen in Figure 2.4.

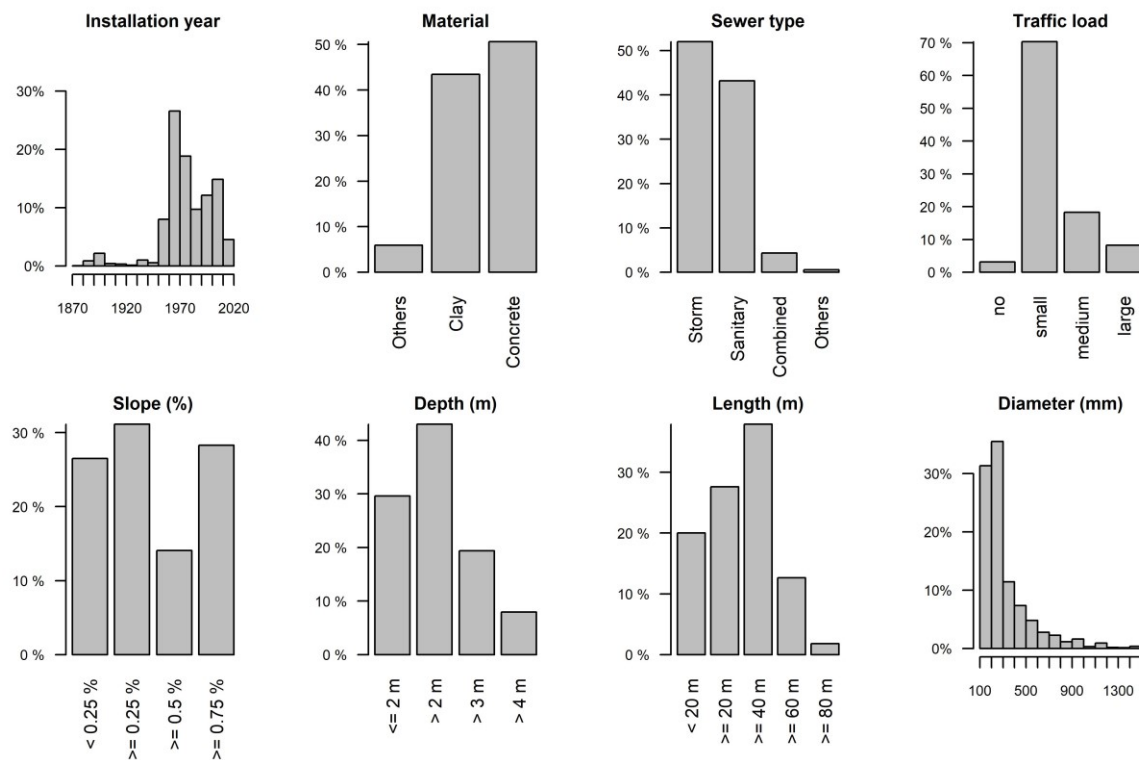


Figure 2.4. Distribution of pipe characteristics and environmental features for the sewer network of Braunschweig

CCTV inspections have been carried out for decades and have been stored systematically in a database since 1998. Inspections reports are encoded using an adaptation of the German standard ATV-M 143-2 (1999). In the frame of this research, structural condition assessment is performed using the French RERAU methodology; the scores range from 1 to 4, with 4 being the worst condition. All pipes have been inspected at least once.

2.4 Case studies

In 2016, the total number of available CCTV inspections was 55,319. Inspections with inconsistent defect coding, without age (inspection year or construction year is missing) or that could not be linked to a specific sewer pipe have been discarded from the database. This filtering step should not influence the analysis since it can be assumed that the characteristics of the filtered pipes are random: there is no evidence that the discarded pipes have specific characteristics so we assume that the filtered inspection dataset is representative of the network.

After the data clean-up, the database contained 35,826 inspections with an inspected length of 1,027 km. This database has been used for the analysis in chapter 3. The database has been updated in 2018 with additional inspections. In 2018, the total number of available CCTV inspections was 69,384. After the data clean-up, the database contained 45,049 inspections with an inspected length of 1,784 km. This database has been used for the analysis in chapter 5.

2.4.2 Berlin

The city of Berlin is the capital and largest city in Germany with a population of around 3,700,000 inhabitants (Figure 2.5).

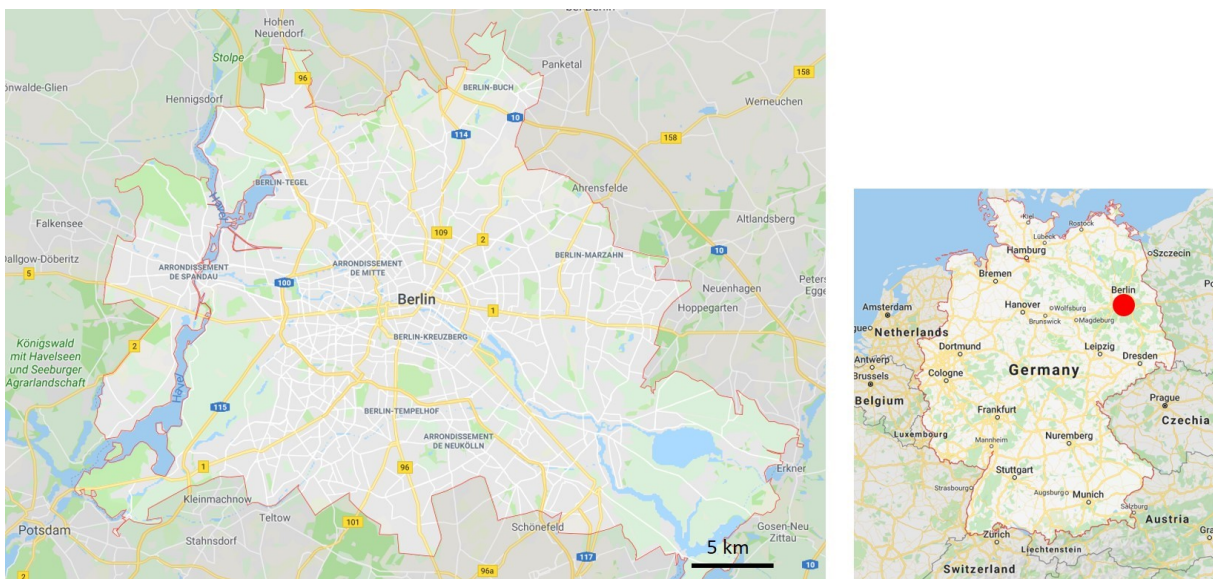


Figure 2.5. Map of the city of Berlin (source: google.com/maps)

2.4 Case studies

The sewer network of Berlin is operated by Berliner Wasserbetriebe (BWB). It has a length of about 9,710 km and is composed of 235,988 pipes. Most pipes are sanitary sewers (45%) whereas 35% are storm water drainage pipes and 20% are combined. Clay and concrete are the two dominating materials with portions of 54% and 25%, respectively. The GIS database contains the main pipes characteristics (construction year, material, type of effluent, shape, diameter, length, depth, slope, city district) and has been extended with environmental features expected to influence sewer deterioration (tree density, i.e. the number of trees in a radius of 10 m, proximity of tramway or subway, groundwater level, type of soil) (Figure 2.6).

The Berlin water company conducts extensive CCTV inspection programs since the 80s. Defects observed during inspections are systematically coded in a local coding system similar to the German guideline ATV M 143-2 (1999). Sewer structural condition is evaluated using an internal company classification system with six grades. The six grades have been aggregated in three grades indicating the emergency of rehabilitation (i.e. good condition 1; intermediate condition 2; poor condition with urgent rehabilitation need 3).

Inspections with inconsistent defect coding, without age (inspection year or construction year is missing), with negative age or that could not be linked to a specific sewer pipe have been discarded from the database. Repaired pipes and liners have also been excluded from the analysis. As in Braunschweig, there is no evidence that the discarded pipes have specific characteristics so we assume that the filtered inspection dataset is representative of the network. After the data clean-up, the database contains 115,258 inspections with an inspected length of 4,825 km over 102,258 pipes. This database has been used for the analysis in chapter 4.

In 2018, the database has been updated with additional inspections. After data clean-up, the database contained 124,450 inspections with a length of 5,222 km over 107,788 pipes. This database had been used for the analysis in chapter 5 and 6.

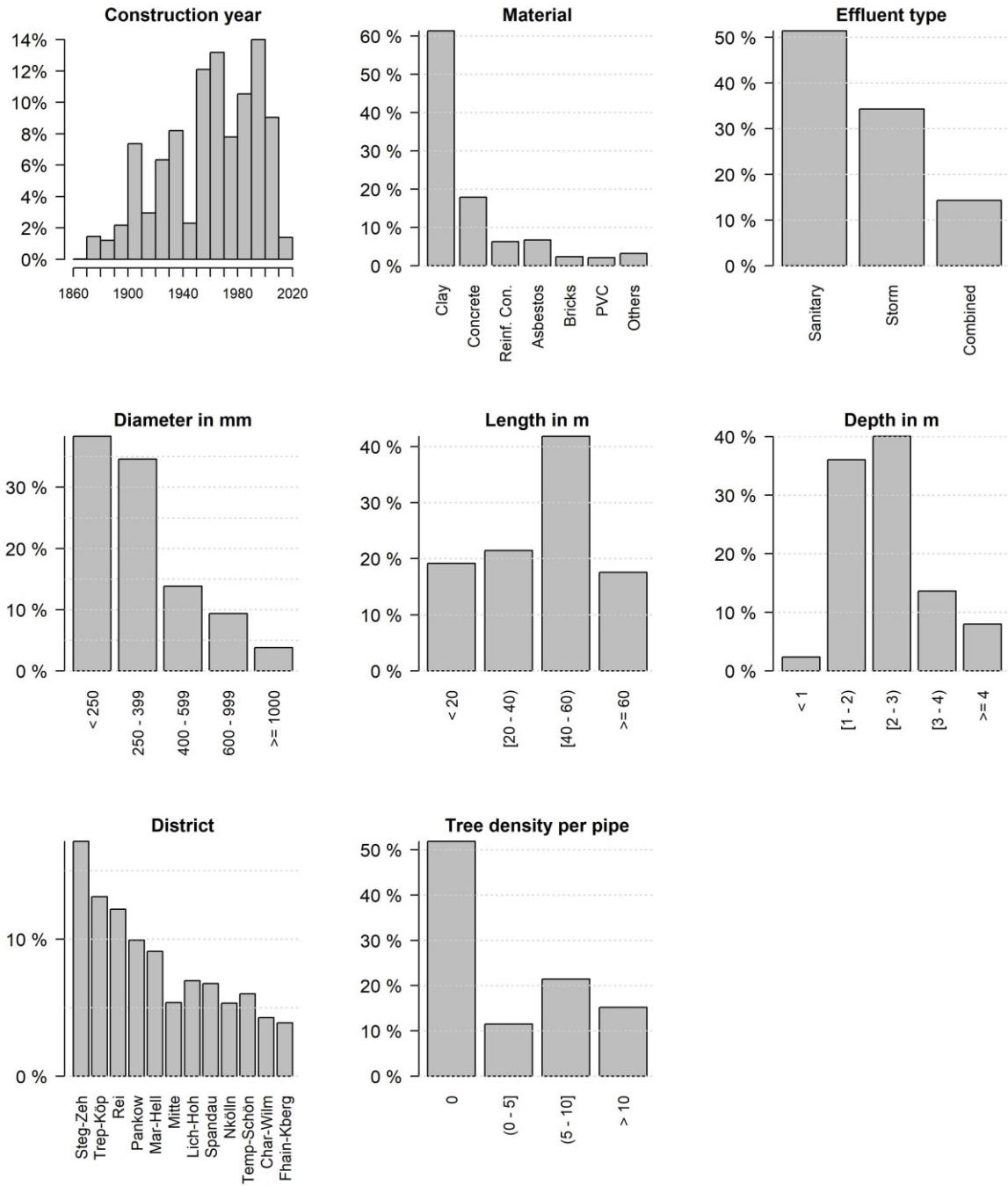


Figure 2.6. Distribution of pipes characteristics in the sewer network of Berlin – only inspected pipes are considered

2.4 Case studies

Chapter 3 The influence of data availability on modelling performance

This chapter is an adapted version of the original paper:

Caradot N., Sonnenberg H., Kropp I., Ringe A., Denhez S., Hartmann A., Rouault P. 2017.

The relevance of sewer deterioration modelling to support asset management strategies. *Urban Water Journal*, 14(10), 1007-1015.

<http://doi.org/10.1080/1573062X.2017.1325497>

3.1 Overview and context of paper I

This paper has been published in the frame of the research project SEMA (2012-2016), funded by Veolia. The project SEMA aimed at investigating the suitability of sewer deterioration models to predict sewer condition state and at identifying the relevant specifications of sewer deterioration models and input data for reaching accurate model predictions. Sewer deterioration models have been applied in the cities of Braunschweig in Germany, in close collaboration with the German engineering company 3S Consult GmbH (3SC).

This paper aims at assessing the ability of a statistical deterioration model to simulate the condition of a sewer network and evaluating the influence of CCTV data availability on modelling performance.

3.2 Research method

The statistical model GompitZ (Le Gat, 2008) has been selected since many authors have already highlighted the advantages of using Markov chain models to simulate sewer degradation (Le Gauffre *et al.*, 2014; Rokstad & Ugarelli, 2015) and several consulting offices use similar statistical approaches to plan investment strategies (e.g. in Germany: 3S Consult, Stein und Partner).

GompitZ has been calibrated and validated using real data of the city of Braunschweig, Germany, where all pipes have been inspected at least once. For each sewer, the predicted condition at the year of inspection was compared with the inspected condition. The model performance was then assessed by analyzing the deviation between model predictions and observations. Monte-Carlo simulation has been used to investigate the influence of sample size on model performance. Finally, the performance obtained was compared to the results of a simplified deterioration model that estimates network condition, based only on the condition distribution of random subsets of inspected pipes.

3.2.1 Deterioration modelling

The GompitZ model is based on the theory of Non-Homogeneous Markov Chains (NHMC), with transition probabilities derived from the Gompertz distribution. The transition probabilities are conditional on the values of a set of covariates which are assumed to contribute to sewer deterioration. The Markov process describes the behavior of a system that passes through a finite number of condition states. At each time step, the system may change its state from the current to a worse one, or remain in the same condition, according to a given probability.

Input data

Input data to the model are the entire set of inspections S .

$$S = \{O_i, i = 1, 2, \dots, N\}$$

O_i are the N pipes observations, each composed of

- ID_i the identification number of the inspected pipe
- c_i the structural condition class obtained from the CCTV inspection
- $t_{const, i}$ the construction year of the pipe

- t_i the year of the inspection
- L_i the length of the inspected pipe
- Z_i the set of covariates values for the inspected pipe

$$Z_i = \{Z_{ij}, j = 1, 2, \dots, z\}$$

With Z_{ij} the values of each covariate j among the number of covariates z . Covariates are the factors that are suspected to drive the deterioration process. It includes a list of pipe features (e.g. material, diameter, type of sewage) as well as environmental factors (e.g. traffic load, water level, presence of trees) (Carvalho *et al.*, 2018; Davies *et al.* 2001; Nariné Torres *et al.*, 2017).

The structural condition class of the inspected pipes is evaluated using the French classification methodology RERAU (Ahmadi *et al.* 2014b; Le Gauffre *et al.* 2004). The aim of this methodology is to rank inspected sewer pipes based on the urgency of their technical rehabilitation needs. In this classification, a condition class is assigned to each sewer segment (from manhole to manhole) on a four-grade scale (1 to 4, 4 being the worst condition: “in need of immediate rehabilitation”). The condition class is calculated using the characterization and quantification of sewer defects such as fissures, corrosion and surface damages that may lead to structural failure such as a pipe collapse. Using the methodology RERAU, the condition distribution of the entire set of inspections S is defined as

$$I = \{I_1, I_2, I_3, I_4\}$$

With I_k being the percentage of pipes observed in condition class $c = k$ in the set of inspections S .

Model calibration and prediction

The model is calibrated with a set of inspection data using the maximum marginal likelihood estimation. Calibration parameters are then used to compute a vector of condition probabilities for each sewer and each year, over a given prediction period. In this study the probability vector $\{p_1, p_2, p_3, p_4\}$ indicates the probability of the pipe to be in a given condition class ranging from 1 to 4, with 4 being the worst condition, for each year.

For more details on the GompitZ model and the calibration methodology, you may refer to Le Gat (2008).

3.2.2 Assessment of model's performance

Deviation between model predictions and observations

Monte-Carlo simulations are performed to assess the prediction quality of the model GompitZ. The model is run M times; for each Monte-Carlo run m , a random subset R_m of size n is selected among the entire set of inspections S of size N

$$S = \{O_i, i = 1, 2, \dots, N\}$$

$$R_m = \{O_j, j = 1, 2, \dots, n\} \quad R_m \subset S$$

For each set m , the following steps are performed

- GompitZ is calibrated using the subset of inspections R_m
- GompitZ predicts the condition of each pipe of the entire set of inspections S , at their year of inspection, i.e. GompitZ computes the condition probabilities $\{p_1, p_2, p_3, p_4\}$ of each observation O_i at the year of observation t_i
- For the entire set of inspections S , the proportion of pipes in each condition class k can be calculated as the average of all condition probabilities p_k weighted by pipe length L

$$P_k = \frac{\sum_{i=1}^N p_k \cdot L_i}{L}$$

The predicted condition distribution of the entire set of inspections S is defined as

$$P = \{P_1, P_2, P_3, P_4\}$$

- The performance of the model is assessed by the deviation between the inspected and predicted condition distributions

$$d_{GompitZ} = \{d_1, d_2, d_3, d_4\} = \{I_1 - P_1, I_2 - P_2, I_3 - P_3, I_4 - P_4\}$$

The value $d_{GompitZ}$ is computed for each Monte-Carlo simulation. Finally, the mean values and variances of the deviations are calculated over the M Monte-Carlo simulations and plotted for graphical analysis.

Comparison with a simple model based on random selection of pipes

Rokstad & Ugarelli (2015) have already proposed to compare the variance of a given model's prediction with the variance of a simple model, based only on a random selection of sewer pipes. This approach enables one to assess the added value of using a more complex approach (GompitZ with additional covariates) when compared to a simplified model based on the available inspection data only.

With the random selection model, the whole network is assumed to have the same condition distribution as a random sample of inspection data. If the entire network has already been inspected, there will be no uncertainty on the condition distribution of the network (except the uncertainties of the condition assessment procedure itself). On the other hand, if only a subset of the network has been inspected, it is possible to calculate the proportion of pipes in each condition class and extrapolate this condition distribution to the entire network. The

extrapolated condition distribution would be uncertain as only a limited subset of the network was inspected. The random selection model does not consider any influence of additional covariates on sewer condition. Predictions are based only on the observed condition classes of inspected sewers.

The random selection model estimates the condition distribution of the entire set of inspections S as being equal to the condition distribution of the subset of inspections R_m .

- For a subset R_m of size n , selected randomly among the entire set of inspections S of size N , the proportion of pipes in each condition class k is expressed as

$$R = \{R_1, R_2, R_3, R_4\}$$

With R_k being the percentage of pipes observed in condition class $c = k$ in the set of inspections R_m .

- For the entire set of inspections S of size N , the condition distribution is assumed to be equal to R
- The performance of the simplified model is assessed by the deviation between the inspected and predicted condition distributions

$$d_{random} = \{d_1, d_2, d_3, d_4\} = \{I_1 - R_1, I_2 - R_2, I_3 - R_3, I_4 - R_4\}$$

The operation is repeated for M Monte-Carlo simulations: d_{random} is computed for each Monte-Carlo simulation and finally the mean values and variances of the deviations are calculated over the M Monte-Carlo simulations and plotted for graphical analysis.

It is of interest to compare the variance and mean values of d_{random} and $d_{GompitZ}$. For a subset of size n , if the variance and mean values of d_{random} are similar to the variance and mean values of $d_{GompitZ}$, it means that the complexity of the deterioration model and the consideration of additional covariates are not justified since similar results can be obtained with the simple

random selection model. On the other hand, if the variance of $d_{GompitZ}$ is smaller than the variance of d_{random} , it can be concluded that GompitZ performs better than the simple random selection model.

3.3 Data description

This study has been performed using the GIS and CCTV database of the city of Braunschweig in Germany (Refer to chapter 2.4 for a presentation of the case study and available data). After the data clean-up, the database contains 35,826 inspections with an inspected length of 1,027 km linked to the 25,787 sewer pipes. The condition is assessed with the RERAU methodology; the scores range from 1 to 4, with 4 being the worst condition. The condition distribution of the entire set of inspection data is defined as

$$I = \{I_1, I_2, I_3, I_4\} = \{53\%, 10\%, 24\%, 13\%\}$$

3.4 Results and discussion

The performance of the deterioration model GompitZ has been assessed using a sample of 35,826 inspections. In order to evaluate the influence of the size of the calibration dataset, Monte-Carlo simulations have been performed with different subset sizes n (number of pipes) ranging from $n=0.2\%$ (about 70 pipes) to $n=100\%$ (35,826 pipes). For each subset size GompitZ and the random selection model were run 150 times, and the mean values and variance of deviations between model predictions and inspections were then computed.

GompitZ has been calibrated considering sewer material, sewage type (sanitary, storm, combined), diameter and depth as covariates. They were found to be the only relevant covariates and thus included in the model (analysis based on significance of model parameters, p-values). The material and sewage type are highly correlated variables since most stormwater pipes are made of concrete and most sanitary sewers are made of clay. However, the integration of both

variables into the model allow to consider a specific degradation behavior for combined sewers, which are mainly made of clay. Several authors have already highlighted the influence of these variables on the deterioration processes (Baur & Herz, 2002; Carvalho *et al.*, 2018; Chughtai & Zayed 2008; Davies *et al.* 2001). Other potentially relevant covariates such as the type of bedding soil and the presence of trees were not available and thus not included in the model.

3.4.1 Model assessment using all inspections for model calibration

Figure 3.1 shows model results with a calibration subset size of $n=100\%$ (35,826 pipes):

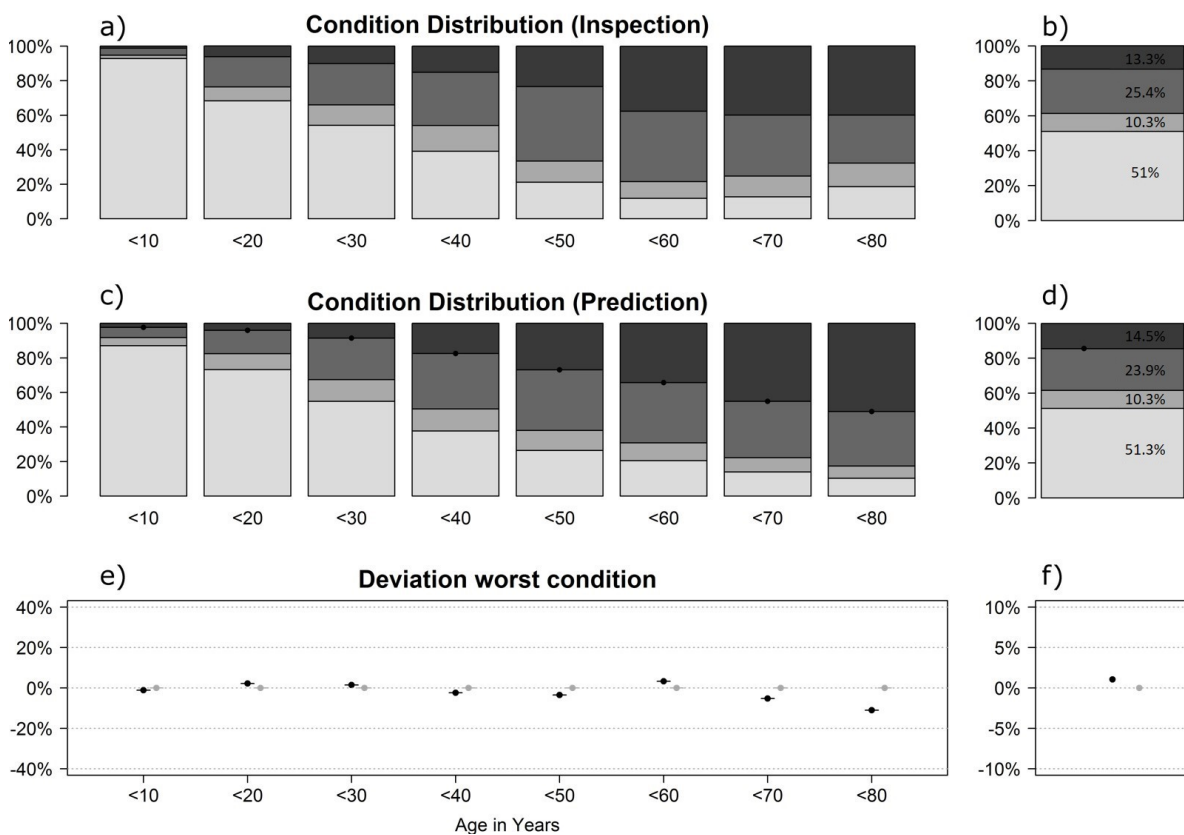


Figure 3.1. Analysis of model performance using a subset size $n=100\%$ for model calibration. Inspected (a and b) and predicted (c and d) condition distributions of the 35,826 inspected sewers in the city of Braunschweig. Deviations between the predicted and inspected proportions of sewers in poor condition 4 (e and f) - black for GompitZ and grey for the random selection model. For convenience, the labels of the age groups have been simplified: “<x” represents the interval $[x-10 ; x[$

- Figure 3.1 [b] shows the inspected condition distribution of the entire dataset (35,826 pipes). Figure 3.1 [a] shows the inspected condition distribution grouped in age periods of 10 years.

- Figure 3.1 [c and d] shows the predicted condition distribution (with GompitZ) of the entire dataset (35,826 pipes).
- Figure 3.1 [e and f] shows
 - the deviation between the inspected and predicted (with GompitZ) proportions of sewers in the worst condition (4): $I_4 - P_4$ (black points)
 - the deviation between the inspected and predicted (with random selection model) proportions of sewers in the worst condition (4): $I_4 - R_4$ (grey points)

Using the random selection model, deviation is always 0, indicating that the model reproduces perfectly the condition distribution of the network. This is an expected result: since the entire network has already been inspected, there is no uncertainty on the condition distribution of the network.

Using the GompitZ model, very low deviations are achieved both for the entire network (Figure 3.1 [f]) and for each age period (Figure 3.1 [e]). However, and obviously, the model cannot provide better results than the knowledge of the entire network.

Additionally, the predicted condition distribution of old pipes appears to be slightly biased. The model accounts for continuous network degradation and does not manage to reproduce the rather stable inspected condition of old pipes. Inspection data tend to be biased as the observations are carried out in a restricted time window (in Braunschweig from 1998 to now). Most old or deteriorated pipes have already been replaced, thus are not fully represented in the sample of inspection data. Several methods have been proposed to correct this "survival selection bias" (Le Gat, 2008; Scheidegger & Maurer, 2012).

3.4.2 Model assessment using a calibration subset size n=3%

Figure 3.2 shows the model outputs with a calibration subset size of n=3% (about 1,000 pipes). The whiskers indicate the uncertainties of the prediction for the worst condition class 4. This is the most critical condition as rehabilitation decisions would be based on this value. Uncertainties come from the random selection of calibration subsets and are computed as twice the standard deviation of the predicted proportion of sewers in poor condition P₄.

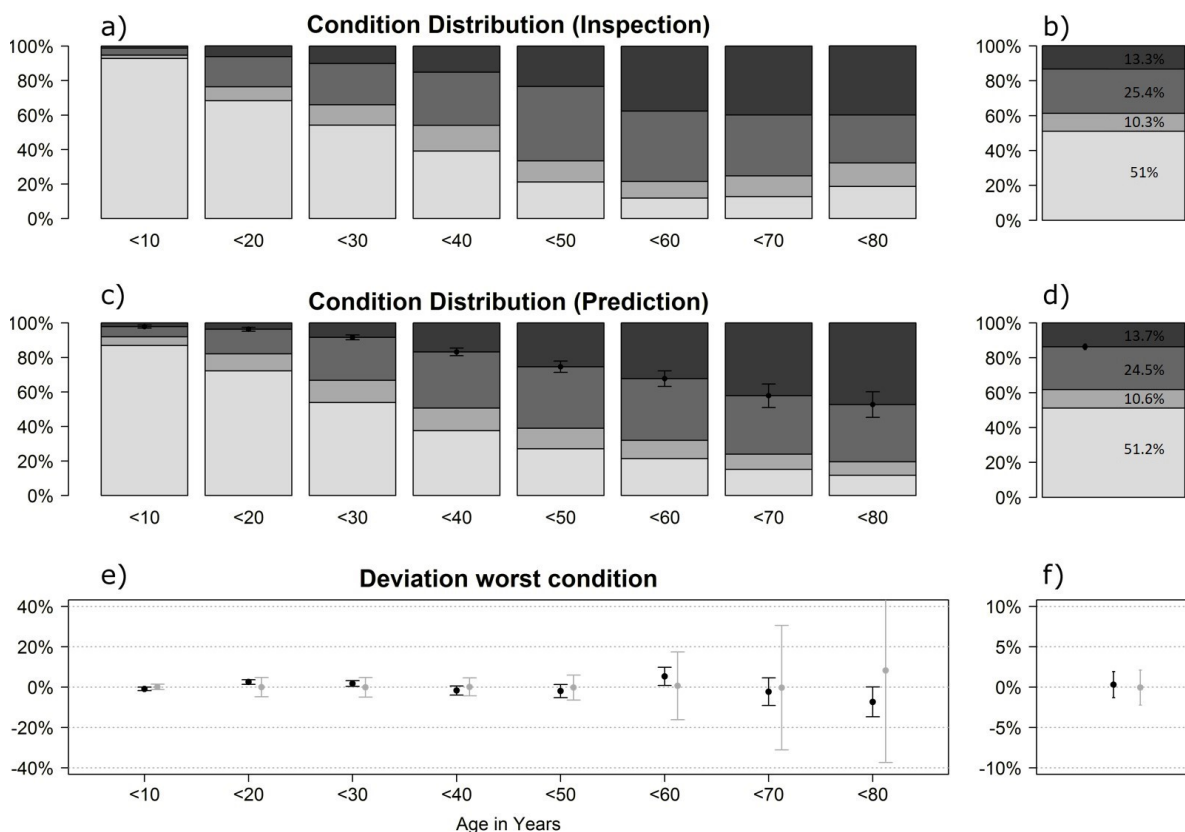


Figure 3.2. Analysis of model performance using a subset size n=3% for model calibration. Inspected (a and b) and predicted (c and d) condition distributions of the 35,826 inspected sewers in the city of Braunschweig. Deviations between the predicted and inspected proportions of sewers in poor condition 4 (e and f) - black for GompitZ and grey for the random selection model. For convenience, the labels of the age groups have been simplified: “<x” represents the interval [x-10 ; x]

At the network scale, the variances of the deviations obtained with random selection model and GompitZ are similar (Figure 3.2 [f]). The deviation obtained with GompitZ ranges between [-1.3%; 1.9%] whereas the deviation obtained with the random selection model ranges between [-2.1%; 2.2%] (Figure 3.2 [f]). If only 1,000 pipes were to be randomly inspected, GompitZ would perform a little better than the random selection model; the model complexity of

GompitZ is not justified since similar results can be obtained with a very simple approach.

However, looking at the deviations for old pipes only (older than 50 years), GompitZ performs better than the random selection model (Figure 3.2 [e]). For pipes older than 70 years, the deviation obtained with GompitZ ranges between [-9.1%; 4.5%], whereas the deviation obtained with the random selection model ranges between [-29.7%; 30.3%] (Figure 3.2 [e]). Even with this small amount of data for model calibration, GompitZ is able to simulate the condition of the network with very low deviation. In this case, GompitZ performs much better than the random selection model, indicating that of these two models only GompitZ is appropriate to simulate sewer deterioration (i.e. the condition of old pipes).

3.4.3 Influence of the calibration subset size on model performance

The means and variabilities of the deviations d_{random} and $d_{GompitZ}$ have been computed for several subset sizes $n=[0.2, 1, 3, 5, 10, 20, 30, 50, 80, 100]\%$. Figure 3.3 [a] shows the means and confidence intervals (twice the standard deviation) of d_{random} and $d_{GompitZ}$ for each subset size. It represents the deviations between model predictions and inspections at the network scale. For the subset sizes 3% and 100%, the values can also be read from Figures 3.1 and 3.2 [f]. The results show that the random selection model consistently performs as well or better than the GompitZ model when considering the entire data set. For the purpose of estimating the condition of the network, the simple model based on the random selection of sewer pipes performs similar or even better than the GompitZ approach. This is especially true when more than 5% (1,800 pipes) of data are available to calibrate the models. This result confirms the findings of Rokstad & Ugarelli (2015). With an important subset (more than 50%) used for calibration, the random selection model performs better, that is to say has a lower (and almost null) deviation compared to the GompitZ model. Such cases does not generally correspond to the reality and the utility has often a smaller subset available for calibration.

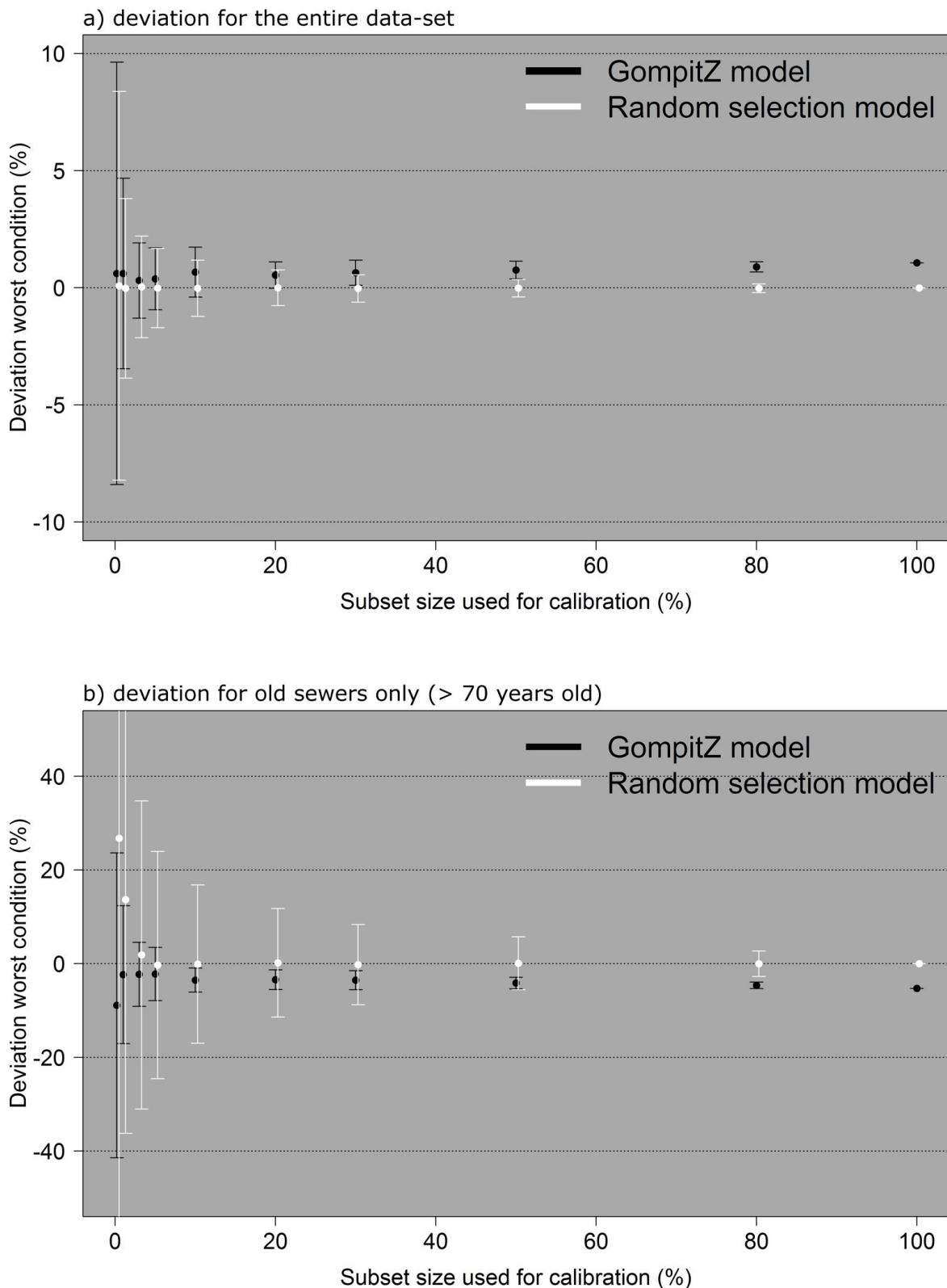


Figure 3.3. Mean and confidence intervals (twice the standard deviation) of both deviations d_{random} and d_{GompitZ} for different calibration subset sizes. It represents the deviation between model predictions and inspections for the entire inspection dataset (a) and for old sewers only, > 70 years old (b)

3.4 Results and discussion

Figure 3.3 [b] also shows the models deviations (mean and confidence intervals of d_{random} and $d_{GompitZ}$) for each subset size but only for old sewers, i.e. sewers that were older than 70 years at the time of the inspections. For the subset sizes 3% and 100%, the values can also be found on Figures 3.1 and 3.2 [e, $x = <80$].

This graph is of interest as it assesses the ability of the models to simulate sewer deterioration, i.e. the condition of old sewers, and not only the general condition distribution of the network. For a subset size of less than 50% (about 18,000 pipes), the GompitZ model performs much better than the random selection model. Indeed, the confidence intervals of the deviations from the GompitZ model are included in the confidence intervals of the random selection model. In the case of low data availability, the GompitZ approach provides much better results:

- With 3% of data used for model calibration (about 1,000 pipes), GompitZ is able to simulate the number of pipes in poor condition with a mean deviation of $-2.3\% \pm 6.8\%$ (confidence interval $[-9.1\%; 4.5\%]$). The random selection gives a mean deviation of $0.3\% \pm 30.0\%$ indicating much higher uncertainties (confidence interval $[-29.7\%; 30.3\%]$).
- With 1% of data for model calibration (about 350 pipes), GompitZ is able to simulate the number of pipes in poor condition with a mean deviation of $-2.3\% \pm 14.7\%$ (confidence interval $[-17.0\%; 12.4\%]$). The random selection gives a mean deviation of $13.6\% \pm 50.9\%$ (confidence interval $[-37.3\%; 64.5\%]$). The uncertainty is much higher using the random selection model.

With 80% of data used for model calibration (about 29,000 pipes), the deviation obtained with GompitZ ranges between $[-5.3\%; -3.9\%]$ whereas the deviation obtained with the random selection model is smaller ranging between $[-2.8\%; 2.8\%]$. This confirms also an intuitive idea:

if a very extensive inspection dataset is available, the best estimation of the condition distribution of old pipes is achieved by extrapolating the condition distribution of the already inspected pipes; the model is no longer of use.

3.5 Conclusion

This study investigated the ability of a statistical deterioration model to simulate the condition distribution of a given network. The statistical model GompitZ has been applied using the extensive GIS and CCTV database of the city of Braunschweig in Germany. Regarding data availability, this city is an ideal case study, since the entire sewer system has already been inspected at least once. The performance of GompitZ has been evaluated by calculating the deviation between the predicted and inspected condition distributions. The obtained performance has been compared to the performance of a simplified deterioration model that estimates network condition, based only on the condition distribution of random subsets of inspected pipes. Key results can be summarized as follows:

- For the purpose of estimating the current condition of the network, the simple model based on the random selection of sewer pipes generally performs as well or better than the GompitZ approach (based on the deviations between the predicted and inspected condition distributions). This is especially true when more than 5% (1,800 pipes) of data are available to calibrate the models. It indicates that a complex model such as GompitZ is not useful since similar results can be obtained with a very simple approach.
- For the purpose of simulating sewer deterioration (the condition distribution of old pipes), GompitZ outperforms the simple approach, especially in the case of low data availability. Even with 3% of data used for model calibration (about 1,000 pipes) GompitZ is able to simulate the amount of pipes in poor condition with a deviation smaller than 10%. The deviation obtained with GompitZ ranges between [-9.1%; 4.5%]

whereas the deviation obtained with the random selection model ranges between [-29.7%; 30.3%].

- Based on this outcome, a subset of at least 1,000 pipes (around 50 km) selected randomly in the network seems to be a usable starting point to simulate network condition with relatively good accuracy. This result confirms the conclusions of Ahmadi *et al.* (2016) and Tran (2016) who respectively recommended a subset of at least 1,000 and 700 pipes for model calibration.
- GompitZ has been calibrated considering sewer material, sewage type (sanitary, storm, combined), diameter and depth as covariates. They were found to be the only relevant covariates and thus included in the model. Other potentially relevant covariates such as the type of bedding soil and the presence of trees were not available and thus not included in the model.

This finding highlights the relevance of using statistical modelling tools to simulate sewer deterioration and support strategic asset management. The simple model is useful to estimate the current overall condition of the network but fails at predicting the condition of old pipes. On the other hand, GompitZ is successful at simulating the condition of old sewers, thus the deterioration of the network. With at least 1,000 inspected pipes for model calibration, municipalities get a reliable tool for the scheduling of inspection programs (by identifying sewers in critical condition) and for the planning of rehabilitation budgets (by simulating several sewer rehabilitation scenarios).

These results underline the advantages of deterioration modelling for the mid to long-term management of sewer systems but also show the high uncertainties related to model outputs. Perspectives for future works include the careful assessment of each source of uncertainty, especially errors induced by the subjectivity of sewer coding and condition classification.

3.5 Conclusion

Chapter 4 Practical benchmarking of the performance of statistical and machine learning models

This chapter is an adapted version of the original paper:

Caradot N., Riechel M., Fesneau M., Hernández N., Torres A., Sonnenberg H., Eckert E., Lengemann N., Waschnewski J., Rouault P. 2018. Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany. *Journal of Hydroinformatics*, 20(5), 1131–1147.

<https://doi.org/10.2166/hydro.2018.217>

4.1 Overview and context of paper II

This paper has been published in the frame of the project SEMA-Berlin (2016-2017) funded by the Berliner Wasserbetriebe. This project aimed at comparing the performance of several statistical and machine learning deterioration models for the sewer network of the city of Berlin, Germany. This project also benefited from a collaboration with the Javeriana University in Bogotá (Colombia) supported by DAAD with funds from the German Federal Ministry of Education and Research (BMBF). The researchers Nathalie Hernández and Andres Torres spend several months in Berlin contributing to the research.

This paper aims at

- Developing a set of metrics adapted to the local utility needs to assess the performance of sewer deterioration models from an end-user perspective. The metrics must be intuitive, self-explanatory and thus clearly understandable by the sewer operator. Thus, they should be able to convince the utility about the relevance or uselessness of using

deterioration models to support asset management strategies. This point is crucial for facilitating the communication of the outcomes and ensuring the acceptance of the results.

- Applying the set of developed metrics to benchmark the performance of a statistical and a machine learning deterioration models trained with the extensive CCTV dataset of the sewer of Berlin, Germany. In Berlin, each pipe of the almost 10,000 km sewer network has been inspected at least once by the end of 2016. The inspection database is a valuable knowledge that can be exploited to assess model performance in the special case of full data availability.

The statistical approach selected is the model GompitZ (Le Gat, 2008), based on the theories of survival analysis and Markov-chains. The machine learning approach selected is Random Forest, an ensemble learning method for classification or regression, already successfully implemented for the prediction of sewer pipes condition (Harvey & McBean, 2014; Rokstad & Ugarelli, 2015; Vitorino *et al.*, 2014). Outcomes of the benchmark analysis will be used to conclude on the strengths and weaknesses of both approaches regarding asset management objectives of the case study.

4.2 Material and Method

4.2.1 Data preparation

The study has been performed using the extensive GIS and CCTV database of the city of Berlin in Germany (Refer to chapter 2.4 for a presentation of the case study and available data). After the data clean-up, the database contains 115,258 inspections with an inspected length of 4,825 km over 102,258 pipes. The number of inspections is higher than the number of pipes because several pipes have been inspected more than once.

22% of the inspected pipes are in poor or very poor condition (condition class 3) and require

immediate or short-term rehabilitation measures; 24% of the pipes are in a medium condition (condition class 2) and must be rehabilitated in the medium term (time horizon: 10 years) whereas 54% of the pipes are in good or perfect condition (condition class 1). Figure 4.1 shows correlations between the main sewer characteristics and the sewer structural condition. The condition is clearly correlated with the pipe age; old pipes are generally in a worse condition than new pipes. However, the condition of very old pipes (> 100 years old) seems to improve slightly. This phenomenon is known as survival selection bias. Inspection data tend to be biased as the observations are carried out in a restricted time window (for this study from 2001 to 2016). Most old and deteriorated pipes have already been replaced, thus are not fully represented in the sample of inspection data.

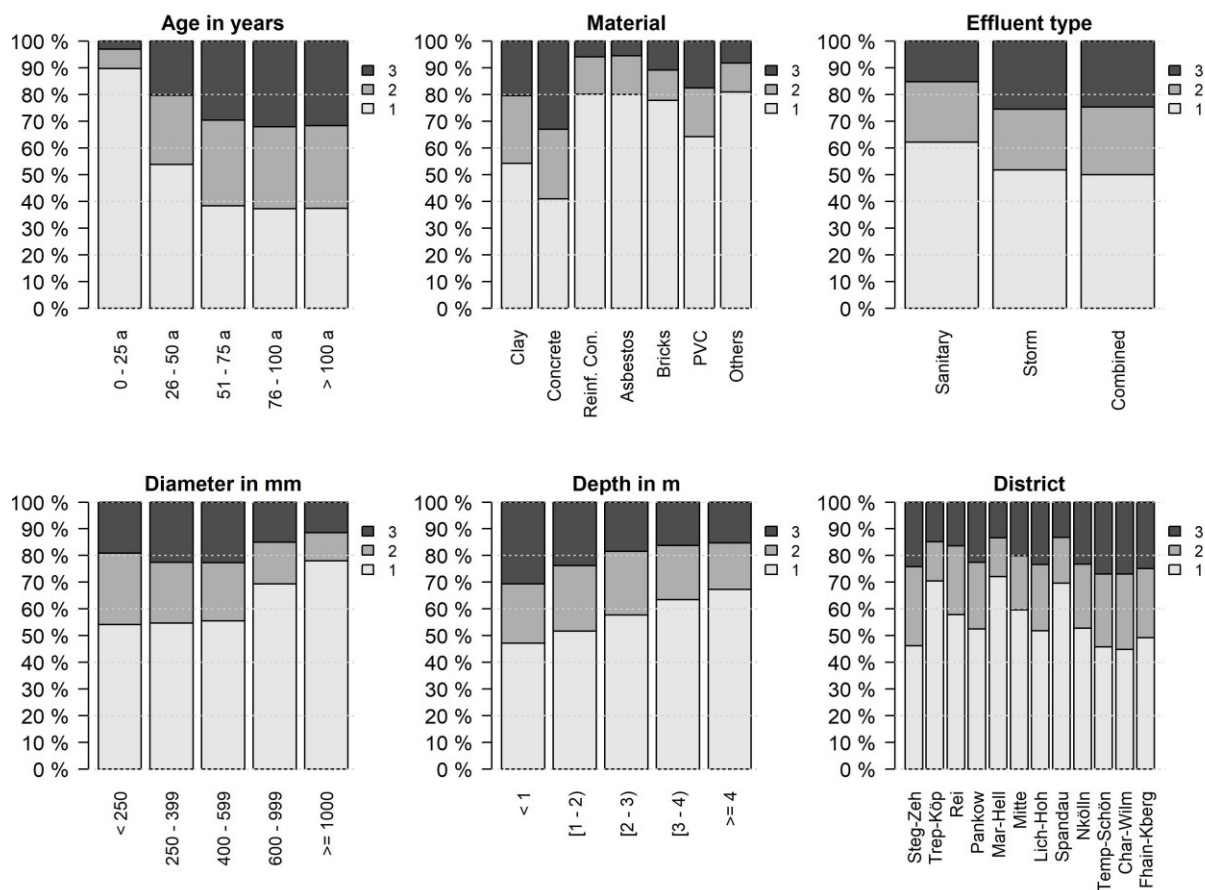


Figure 4.1. Correlation between sewer characteristics and condition. Class 1 (light grey), 2 (medium grey) and 3 (dark grey) represent good, medium and poor condition, respectively

4.2 Material and Method

The condition is also correlated with the pipe material; sewers made of concrete are generally in a worse condition than clay pipes or PVC pipes. Since storm water pipes are mainly constructed with concrete, they appear to be in worse condition than sanitary pipes. The width and the depth seem also to play a relevant role: small-diameter and shallow pipes are in worse condition than big pipes and deeply laid pipes. The condition distribution varies strongly between the districts probably due to cross correlations with other pipes' external characteristics (e.g. type of soil, type of effluent, age of the network, etc.).

4.2.2 Modelling approaches

Random Forest

Random Forest (RF) is an ensemble learning method for classification or regression. It is based upon growing hundreds of decision tree classifiers - in our case of type "CART Classification and regression tree" (Breiman *et al.*, 1984) - and combining them in a single ensemble of models (Breiman, 2001). For classification tasks, the goal is to predict a class output (e.g. condition class) from a set of numerical or categorical variables. The algorithm builds individual unpruned trees using bootstrap aggregation with the following procedure:

- Sample n instances randomly (with replacement) from the original training dataset
- Start the construction of the tree from the root with the n instances
- Search through $mtry$ random variables among M variables to find the best binary split into two children nodes. The best split is determined by minimizing the Gini criterion (Breiman, 2002). The criterion evaluates the performance of the split to classify the output: the maximum value of the Gini criterion is obtained if the distribution of the output is the same in both nodes (poor classification); the Gini criterion is 0 if the values of the outputs are perfectly separated between the two nodes (excellent classification).

- Repeat the previous step to grow the tree until the number of instances in the children nodes reach the critical size defined by the hyperparameter *nodesize*. The *nodesize* determines the size of the trees.

A total of *ntree* trees are grown with the same procedure. The resulting ensemble of trees composes the RF. For a given set of variables, each tree delivers a class output; the prediction of the RF is the mode of the *ntree* class outputs. Class probabilities can also be estimated as the percentage of each class among the *ntree* class outputs.

RF can also deal with imbalanced data in which one of the output classes constitutes a small minority of the data. In such cases, the interest usually leans towards the correct classification of the minority class (e.g. fraud detection, disease diagnostic, etc.). A classical RF might fail because it will seek to minimize the overall error rate, rather than paying special attention to the minority class. The main approach to tackle the problem of imbalanced data is to incorporate class weights into the RF to penalize the misclassification of the minority class. Class weights are used to weight the Gini criterion and to determine the class output at the terminal node of each tree using a weighted majority vote (Chen & Breiman, 2004).

The analysis of the trees structure highlights the relative importance of each variable in the model. The minimal depth is a dimensionless statistic measuring the productiveness of a variable in a decision tree (Ishwaran *et al.*, 2010). The minimal depth of a given variable in a tree is the highest level of the nodes in which the variable has been selected to classify the pipes in the training process. Since the algorithm selects at each node the variable that lead to the best classification, important variables have small minimal depths.

Harvey & McBean (2014) used RF to predict sewer pipes in poor condition in Ontario, Canada. Results were satisfying with false negative rate of 11%, false positive rate to 25% and an area

under the ROC curve >0.80 (with 1.0 indicating a perfect model). Only 1,255 pipes were available for training from which around 10% in poor condition. Rokstad & Ugarelli (2015) used RF with data of the city of Oslo, Norway. They conclude that deterioration model application may be beneficial for prioritising inspection programs and that the performance is limited by the adequacy of the explanatory variables available. Vitorino *et al.* (2014) also demonstrated an application of a RF implemented in the software platform Baseform.org.

Random forests for this study were developed using the randomForest package in the software R (Liaw & Wiener, 2002). Using three output classes, the following hyperparameters need to be set up by the user:

- *mtry* number of variables randomly sampled as candidates at each split
- *nodesize* minimum size of terminal nodes
- *ntree* number of trees in the forest
- w_1, w_2, w_3 priors (i.e. weights) of the condition classes 1, 2 and 3 (Chen & Breiman, 2004)

Markov chains and survival analysis

The model GompitZ (Le Gat, 2008) is based on the theories of survival analysis and Markov-chains. The goal of the model is to predict a probability class output (probability for a pipe to be in a given condition class) from the pipe age and a set of numerical or categorical variables.

Prior to model calibration, pipes are generally grouped in cohorts, i.e. homogenous groups of sewer pipes sharing similar features, e.g. same material and type of effluent. During the calibration procedure, survival functions are estimated for each cohort. Survival curves have the mathematical form of a Gompertz distribution. They are calibrated during a regression procedure using the method of Maximum Likelihood Estimation (MLE) and represent the mean

deterioration of pipes over time: they define the proportion of pipes that have survived at a given age. Additionally, the shape of the survival curves can be modulated by further numerical or categorical variables (also called covariates).

The calibrated survival curves are used to calculate the probability for a pipe to be in a given condition at a given age. If the pipe has never been inspected, the class output is estimated directly from the survival curves. For example for a given pipe and three possible condition classes, a probability vector P is estimated at year T from the survival curves $SC1$ and $SC2$:

$$P(T) = (P_1(T), P_2(T), P_3(T)) = (SC_1(T), SC_2(T) - SC_1(T), 100 - SC_2(T)) \quad (4.1)$$

If the pipe has been inspected at least once, the condition is known at the year of the inspection and the survival curve cannot be used straightforwardly to estimate the future condition. For example if the pipe has been inspected in condition 1 at year T :

$$P(T) = (1, 0, 0) \quad (4.2)$$

At year T , the pipe is in condition 1 with 100% probability. The probability vector has been initialised at year T of the inspection. In this case, a Markov-chain is used to simulate the future evolution of the pipe condition. The probability vector at year $T+dt$ depends on the probability vector at year T and on the transition matrix Q at year $T+dt$.

$$P(T + dt) = P(T) Q(T + dt) \quad (4.3)$$

The transition matrix can be mathematically derived from the slope of the survival curves (Le Gat, 2008). The elements of the matrix are time-dependent and indicate the probability for a pipe to stay in a given condition i (probability $q_i(T)$) or to transit to the next condition $i+1$ (probability $1 - q_i(T)$)

$$Q(T) = \begin{pmatrix} q_1(T) & 1 - q_1(T) & 0 \\ 0 & q_2(T) & 1 - q_2(T) \\ 0 & 0 & 1 \end{pmatrix} \quad (4.4)$$

The Markov-chain for a prediction of the pipe condition at year n can be written as followed.

$$P(T + n) = P(T) Q(T + 1) Q(T + 2) \dots Q(T + n) \quad (4.5)$$

$$P(T + n) = P(T) \prod_1^n Q(T + i) \quad (4.6)$$

4.2.3 Performance metrics

A set of performance metrics has been defined in consultation with the utility manager, Berliner Wasserbetriebe (BWB), in order to benchmark and evaluate model performance. The metrics assess model performance at two main levels: the network and the pipe levels. At network level, the metrics indicate to which extent the model can predict the condition distribution of the entire network, i.e. the number of pipes in each condition. At pipe level, the metrics verify to which extent the model can predict correctly the inspected condition class of each single pipe. Both information are needed for different purposes: network level metrics show the model's relevance for supporting strategic rehabilitation planning; pipe level metrics illustrate the potential for supporting inspection strategies by identifying pipes in critical condition.

Metrics at network level

The performance metrics at network level describe the deviation between the predicted and inspected condition distributions, for the entire network and for different age groups. Six metrics have been defined with the sewer operator.

Deviation of the condition distribution – all pipes; K1, K2, K3 are the absolute deviations between the percentages of sewers predicted and inspected in each condition - K1 for good condition, K2 for medium condition and K3 for poor condition.

Deviation of the condition distribution – only pipes category 51-75 years; K4, K5, K6 are the absolute deviations - for the age category 51-75 years only - between the percentages of sewers predicted and inspected in each condition – K4 for good condition, K5 for medium condition and K6 for poor condition.

K1, K2 and K3 assess the ability of the model to simulate the condition distribution of the entire network whereas K4, K5 and K6 evaluate the ability of the model to consider the deterioration process. The age category of 51-75 years has been selected instead of the oldest category because it corresponds to the depreciation period of concrete and clay pipes. Indeed, older age categories might be biased since many pipes have already been rehabilitated introducing a survival selection bias.

Metrics at pipe level

A model can provide excellent results at network level and nevertheless fail to simulate the right condition of each pipe by simulating the right proportions of each condition but the wrong pipes in each condition. An exhaustive model assessment requires the analysis of the confusion matrix of the outcomes (Table 4.1). The confusion matrix compares the predicted and observed class of each pipe and counts the number of agreements and disagreements.

Table 4.1. Example of confusion matrix with fictive numbers

		Prediction			Sum observations
		Good	Medium	Poor	
Observation	Good	420	56	32	508
	Medium	64	140	25	229
	Poor	36	28	123	187
Sum predictions		520	224	180	

Several metrics have been derived from the matrix and validated with the sewer operator.

The **True rate**; also called sensitivity, indicates the percentage of sewers inspected in condition “i” that have been correctly predicted in the same condition “i”.

$$K7 = \frac{\text{number of correct predictions in good condition}}{\text{number of observations in good condition}} = \frac{420}{508} = 83\%$$

$$K8 = \frac{\text{number of correct predictions in medium condition}}{\text{number of observations in medium condition}} = \frac{140}{229} = 61\%$$

$$K9 = \frac{\text{number of correct predictions in poor condition}}{\text{number of observations in poor condition}} = \frac{123}{187} = 66\%$$

The **False Negative rate**; also called miss rate, indicates the percentage of sewers inspected in condition “i” that have been wrongly predicted in a better condition “j”. False Negative predictions overestimate the inspected condition of the pipes (i.e. give a too optimistic image).

$$K10 = \frac{\text{number of pipes observed in medium condition but predicted in good condition}}{\text{number of observations in medium condition}} = \frac{64}{229} = 28\%$$

$$K11 = \frac{\text{number of pipes observed in poor condition but predicted in good condition}}{\text{number of observations in poor condition}} = \frac{36}{187} = 19\%$$

There is no False Negative rate for the good condition since it cannot be overestimated. The number of pipes observed in medium condition are not considered in the calculation of K11 even if they could be considered as False Negative. It was a decision of the utility to focus on pipes in poor condition in the definition of this indicator.

The **False Positive rate**; also called false alarm probability, indicates the percentage of sewers inspected in condition “i” that have been wrongly predicted in a worse condition “j”. False Positive predictions underestimate the inspected condition of the pipes.

$$K12 = \frac{\text{number of pipes observed in good condition but predicted in poor condition}}{\text{number of predictions in poor condition}} = \frac{32}{180} = 18\%$$

Summary metrics

The metrics defined above are intuitive and clearly understandable by the sewer operator and will be used to convince the utility about the relevance or uselessness of using deterioration models. For the purpose of simplifying the search of the best combination of hyperparameters in the step of model training, metrics at network and pipe levels have also been summarized in one unique single metric each. The summary metrics are defined as the root mean square of the six indicators on both network and pipe level. The square root is used to give more weight to large errors. K7, K8 and K9 are normalized, i.e. subtracted from 100, to have an optimum value of 0. It is worth mentioning that these metrics have no physical meaning: they are not used to communicate the performance of the deterioration model with the utility. However, they are useful to identify the best combination of hyperparameters which maximize the performance of the model at pipe or network level (section 4.2.4).

$$K_{Network} = \sqrt{\frac{K1^2 + K2^2 + K3^2 + K4^2 + K5^2 + K6^2}{6}}$$

$$K_{Pipe} = \sqrt{\frac{(100 - K7)^2 + (100 - K8)^2 + (100 - K9)^2 + K10^2 + K11^2 + K12^2}{6}}$$

$K_{network}$ tends toward 0 when the deviations between model predictions and observations tend toward 0 (metrics K1 to K6). $K_{network}$ is equal to 0 for a perfect model prediction at the network level: the deviations between model predictions and observations are equal to 0. K_{pipe} tends toward 0 when the True rates tend toward 100% (K7, K8 and K9) and when the False Negative and False Positive rates tend toward 0 (K10, K11, K12). K_{pipe} is equal to 0 for a perfect model prediction at the pipe level with 100% True rates and without False Negative and False Positive. Table 4.3 summarizes the defined performance metrics.

4.2.4 Model training and testing

After the withdrawal from the analysis of pipes with missing age, length or depth information or highly underrepresented profiles, material or soil types, the dataset has been separated in two random subsets: training (60%, 58,528 pipes) and test (40%, 39,019 pipes) subsets. The partition 60/40 is commonly used in statistical and machine learning studies, among ratios usually varying between 50/50 and 90/10. In this study, considering the large size of the dataset, the partition size has little influence and does not influence significantly the values of the parameter estimates and performance metrics obtained.

A Chi-squared test of independence (χ^2) has been performed for each pipe feature presented in Figure 2.4 to compare the training and test subsets. It tests the equality of proportions between the two subsets with the null-hypothesis that the distributions of the categorical variables are the same in the two subsets. Reported p-values were higher than the significance level of 0.05 indicating that the null hypothesis cannot be rejected and that the distributions are the same for each categorical variable.

For the Random Forest model, the best combination of hyperparameters have been analysed in two steps in order to reduce the computation time. A first coarse grid search is performed to find the optimal values for the most sensitive parameters, in our case the weighting factors, and then a second fine grid search with fixed weight values is used to identify the optimal values of the remaining two hyperparameters. With this two-step procedure, the computation time for the parameter search can be reduced compared to a full grid search covering all hyperparameters (Bergstra & Bengio, 2012).

- Step 1: random search: a list of 1,000 random hyperparameters combinations has been prepared based on the reasonable range of variation of the four hyperparameters ($w1$, $w2$, $mtry$ and $nodesize$). The weight $w3$ for pipes in poor condition has been set to a

fixed value since the value of the two other weights $w1$ and $w2$ are tested (Chen & Breiman, 2004). For each combination of hyperparameters, a 5-fold cross-validation procedure has been performed on the training dataset using the performance metrics defined in the previous section.

- (1) The training dataset is divided into five random equal sized subsets;
- (2) Of the five subsets, a single subset is retained as the validation data to calculate the performance metrics and the remaining four subsets are used to train the model.

The procedure (1 and 2) is repeated five times with each subset used once as validation data. Finally, the mean of the five sets of performance metrics is calculated. The values of the metrics K_{Network} and K_{Pipe} are plotted against the weight values $w1$ and $w2$ in order to identify the values that maximise the performance.

- Step 2: grid search: step 1 is repeated with fixed values of weights and varying the values of the remaining hyperparameters *mtry* and *nodesize*. The values of the metrics K_{Network} and K_{Pipe} are plotted against the hyperparameters *nodesize* and *mtry* in order to identify the combination of hyperparameters that maximise the performance. Finally, the best combination of hyperparameters is implemented to train the Random Forest model.

For the GompitZ model, the training consists in identifying the relevant cohorts and variables for the calibration of the survival curves. Similar to the Random Forest, a 5-fold cross-validation procedure has been performed on the training dataset using the performance metrics defined in the previous section.

- In a first step, a cross-validation has been run for all combinations of cohorts built with one to five categorical variables.

- If only one variable is used to build the cohorts, pipes groups are composed based on the categorical values of the variable (e.g. for the variable Material: Clay, Concrete, etc.).
- If several variables are used to build the cohorts, the combination of the variables values composes the cohorts (e.g. for variables Material and Sewerage: Concrete-Sanitary, Concrete-Storm, Clay-Sanitary, etc.).
- For the best combination of cohorts, a second cross-validation has been applied by considering additional numerical variables as variables for model calibration.

Finally, the trained models have been tested on the independent test subset to assess model performance. The sewer condition has been predicted at the year of inspection with each sewer and the performance metrics have been evaluated.

4.2.5 Assessment of model performance

The metrics proposed in the previous section enables the evaluation of model performance. In order to understand the variation's range of the metrics, the model metrics have been compared with the performance of a simple random model and an ideal model. The goal is to assess the meaning and interpretation of the model metrics within their range of variations from a poor performing model to an ideal model.

The simple random model attributes randomly a condition class to each inspected pipe. The values of the metrics obtained represent a poor performing model. The ideal model is the achievable ideal knowledge of the condition of all pipes in the network including the uncertainties associated to the sewer condition assessment procedure. The ideal model has the same accuracy/uncertainty as a sewer inspection; it is based on the assumption that it is impossible to know the sewer condition better than with a CCTV inspection. A methodology to assess CCTV uncertainties has been proposed by (Caradot *et al.*, 2018a; see chapter 5). The

approach is based on the analysis of repeated inspections of sewer pipes; it considers only repeated inspections that occur within a short time period (< 3 or 5 years) in order to neglect sewer deterioration. Thus, variations between the condition classes reflect the uncertainties of the procedure of sewer condition assessment. The methodology has been applied using repeated CCTV inspections of 13,753 pipes in Berlin in order to assess the True Positive, False Negative and False Positive rates of sewer inspection. The methodology is presented in detail in chapter 5.

4.3 Results and discussion

4.3.1 Random Forest

The random search analysis (step 1) has been performed on the training dataset using the four hyperparameters. The tested ranges of the hyperparameters are; *nodesize*: 4–1808; *mtry*: 1–12; *w1*: 0.2–3; *w2*: 0.2–3. Figure 4.2 shows the sensitivity of the summary indicators at network and pipe levels depending on the weight factors *w1* and *w2*. It indicates optimal weight values at network level ($w1 = 2$; $w2 = 1$) and at pipe level ($w1 = 1$; $w2 = 0.8$) and suggests the training of one model for each level. The optimal weight *w1* is higher at network level to compensate the imbalanced distribution of condition classes and account for the higher number of pipes in good condition 1.

The grid search analysis (step 2) has been performed with fixed weight values (defined in step 1) to identify the optimal values for *nodesize* and *mtry* at both network and pipe levels. Figure 4.3 shows the sensitivity of the pipe level metrics to the variation of the two hyperparameters: *nodesize* (minimum size of terminal nodes for each tree) and *mtry* (number of variables randomly sampled as candidates at each split). At pipe level, best results are obtained with high values of *mtry* (11) and values of *nodesize* between 20 and 90. The number of trees grown to build the forest (*ntree*) has no influence on model performance (results not shown here). Table

4.2 summarizes the best combination of hyperparameters at both network and pipe levels.

Finally, the trained models have been tested on the independent test subset to assess model performance. Figure 4.4 compares the inspected and predicted condition distributions of the network for the test dataset.

Table 4.2. Best combinations of hyperparameters at both network and pipe levels

<i>hyperparameter</i>	<i>Best model at network level</i>	<i>Best model at pipe level</i>
<i>n_{tree}</i>	100	100
<i>n_{odesize}</i>	7	55
<i>m_{try}</i>	10	11
<i>w₁</i>	2.0	1.0
<i>w₂</i>	1.0	0.8
<i>w₃</i>	1.0	1.0

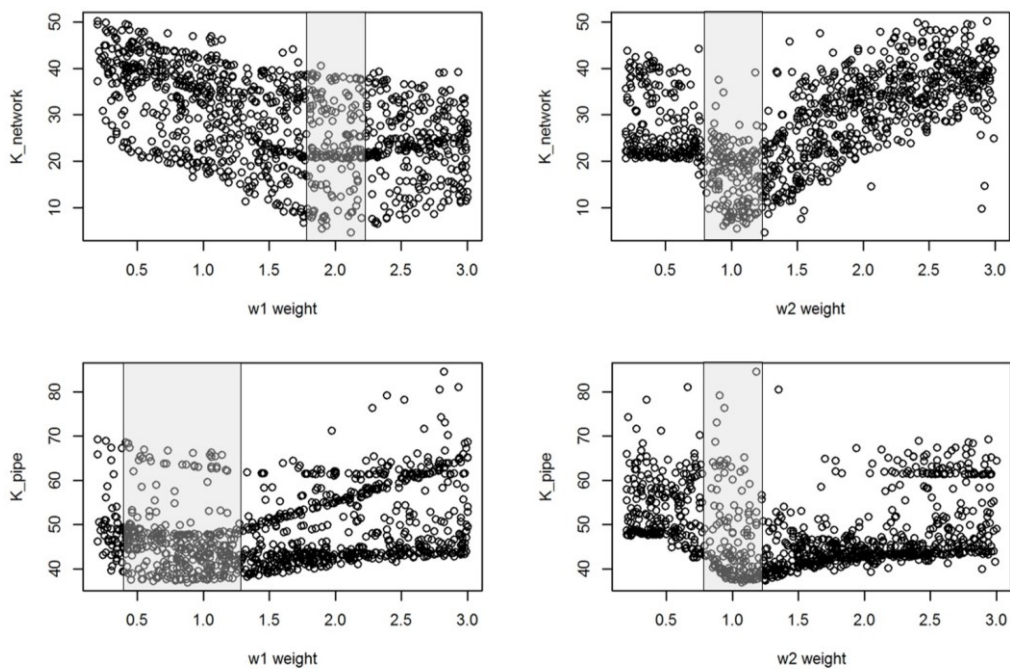


Figure 4.2 Influence of the weights factors w_1 (left) and w_2 (right) over the summary indicators $K_{network}$ (top) and K_{pipe} (bottom); the optimal parameter's window is shaded in grey

4.3 Results and discussion

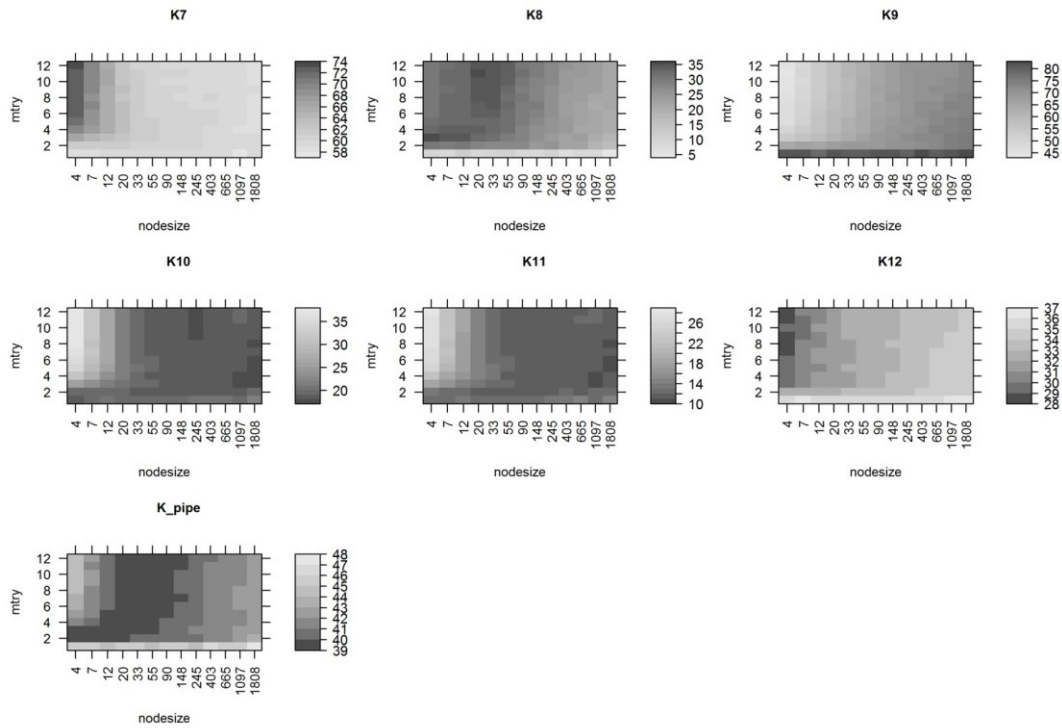


Figure 4.3. Sensitivity of the pipe level metrics to the variation of the hyperparameters *nodesize* (minimum size of terminal nodes) and *mtry* (number of variables randomly sampled as candidates at each split)

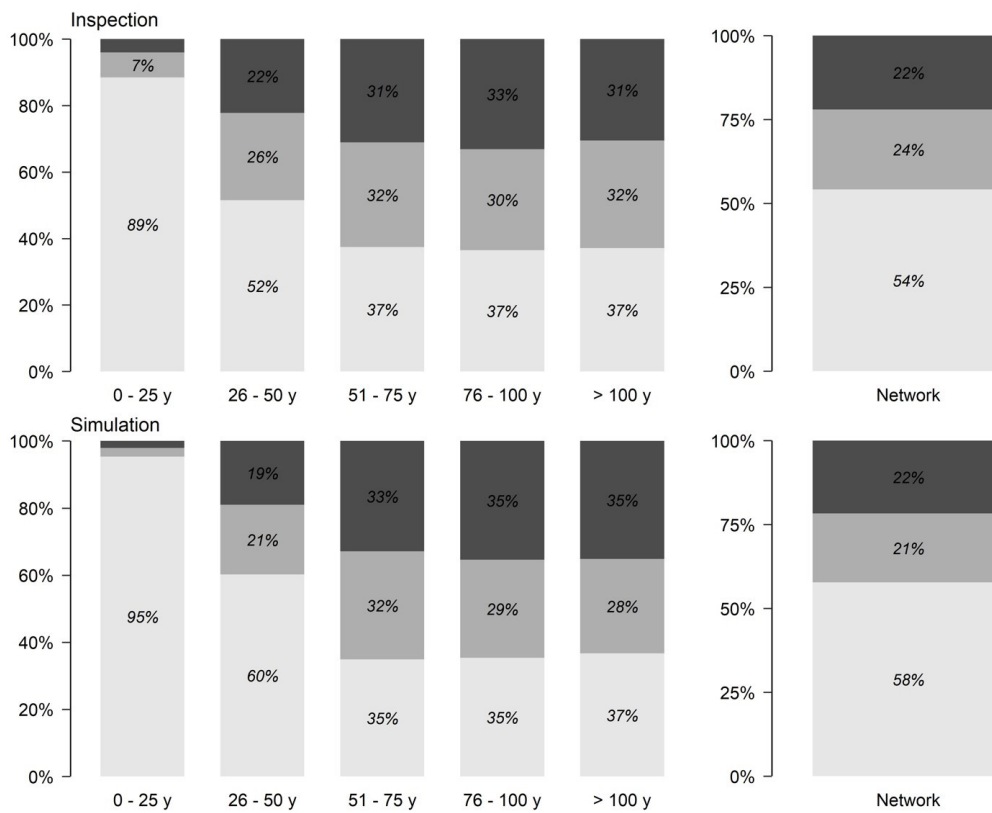


Figure 4.4. Inspected and predicted condition distributions with Random Forest for the entire network (right) and for each age group (left). The colors light grey, medium grey and dark grey represent good, medium and poor condition, respectively

4.3 Results and discussion

Table 4.3 summarizes the value of metrics obtained on the test dataset. The metrics have been calculated with the best models at both network and pipe levels. The deviations at network level (K1 to K6) are relatively low, below 5%. At pipe level, 64.0% of the pipes inspected in good condition have been predicted correctly (K7), 40.0% of the pipes inspected in medium condition have been predicted correctly (K8) and 66.7% of the pipes in poor condition have been predicted correctly (K9). 17.1% of the pipes inspected in medium condition and 9.5% of the pipes inspected in poor condition have been falsely predicted in good condition (K10 and K11). 28.3% of the pipes inspected in good condition have been wrongly predicted in poor condition (K12).

Table 4.3. Summary of performance metrics for the two models on the test dataset

				RF	GompitZ
Network Level	K1	Deviation of the condition distribution	$\in [0,100]$ Goal = minimize	3.5%	0.8%
	K2			3.4%	0.1%
	K3			0.1%	0.7%
	K4	Deviation of the condition distribution – 51-75 years only		2.0%	0.1%
	K5			0.1%	0%
	K6			1.9%	0.1%
Pipe Level	K7	True rate	$\in [0,100]$ Goal = maximize	64.0%	64.1%
	K8			40.0%	29.0%
	K9			66.7%	32.9%
	K10	False Negative rate	$\in [0,100]$ Goal = minimize	17.1%	42.8%
	K11			9.5%	38.0%
	K12	False Positive rate		28.3%	38.5%
Summary	K _{Network}	Summary metric for network level	$\in [0,100]$ Goal = minimize	2.3	0.5
	K _{Pipe}	Summary metric for pipe level		34.5	51.0

Figure 4.5 shows the distribution of the minimal depth of each variable among the 100 trees of the forest built for the pipe level. The most important variable for classification is the sewer age; in 95 of 100 trees, the age is the variable selected by the algorithm for the first split at the

root. The material, the shape and the type of effluent are the following most relevant variables; the material and the type of effluent are strongly correlated since most sanitary and storm water pipes are built of clay and concrete, respectively. The district shows also a strong influence in the model. The district is strongly correlated with pipes characteristics (for example pipes in the city centre are mainly combined sewers made of bricks and built in the 19th century) but also with other potential relevant deterioration factors such as the traffic load, the type of soil or the quality of the construction between the former east and west Berlin. The width, length and depth are also relevant for describing sewer deterioration but secondary compared to the material and district. Finally, the environmental features show little or no influence (type of soil, tree density, groundwater level).

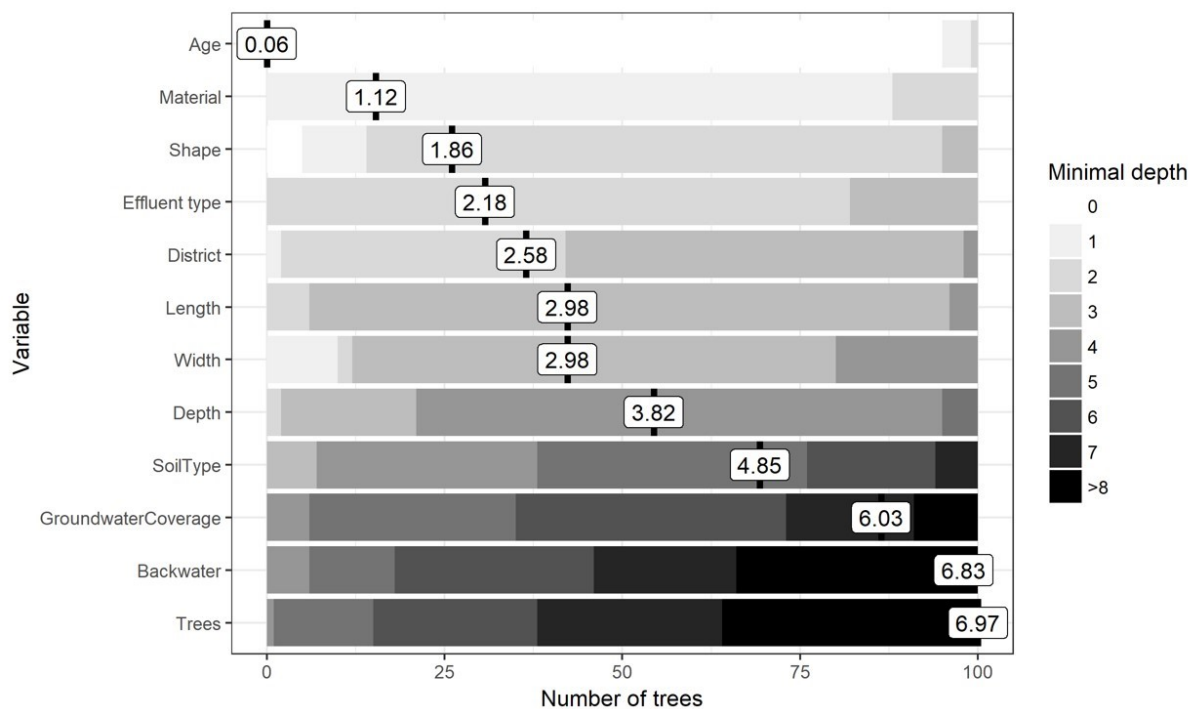


Figure 4.5. Distribution of the minimal depth of each variable among the 100 trees. The values in the bar plots indicate the mean minimal depth of each variable.

4.3.2 Markov chains and survival analysis

The cross-validation procedure has been performed on the training dataset for all combinations

of cohorts built with one to five categorical variables. The best cohort combination is composed of four categorical variables: the material, the district, the shape and the type of effluent. The consideration of additional numerical variables does not improve the model performance. Table 4.3 summarizes the value of metrics obtained on the test dataset. Figure 4.6 shows the inspected and predicted condition distributions of the network.

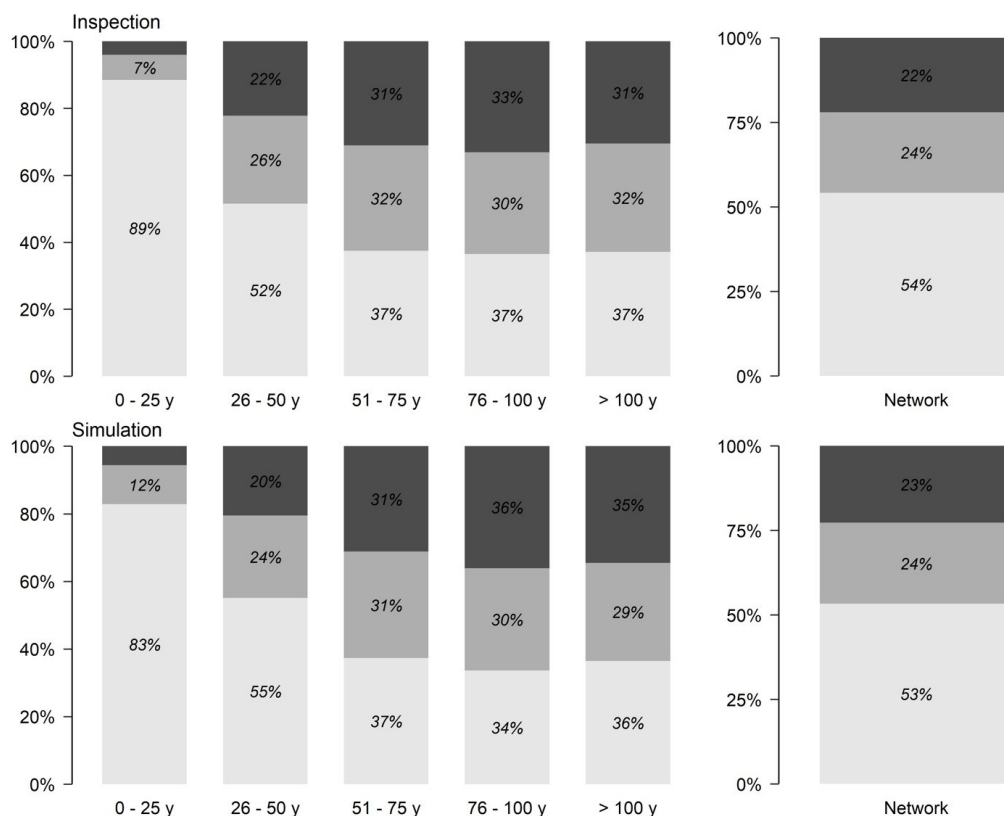


Figure 4.6. Inspected and predicted condition distributions with GompitZ for the entire network (right) and for each age group (left). The colors light grey, medium grey and dark grey represent good, medium and poor condition, respectively

4.3.3 Comparison of model outcomes

Network level

Random Forest and GompitZ give satisfactory outcomes at network level (Table 4.3). Deviations obtained with Random Forest are below 5%; deviations reached with GompitZ are even lower, below 1%. Similar results have been obtained using GompitZ on the Braunschweig case study (Chapter 3). Both models can reproduce accurately the condition distribution of the

entire network and for different age groups.

Pipe level

At pipe level, Random Forest performs better than GompitZ (Table 4.3). In particular, the True rates for pipes in medium and poor condition are 30% and 100% higher, respectively. The False Negative rates and False Positive rate are also minimized with Random Forest. It is interesting to note that the most relevant variables are the same for both models: the material, the district, the shape and the type of effluent.

Figure 4.7 plots the pipe level metrics of both models and compare model performance to the performances of a random and an ideal model.

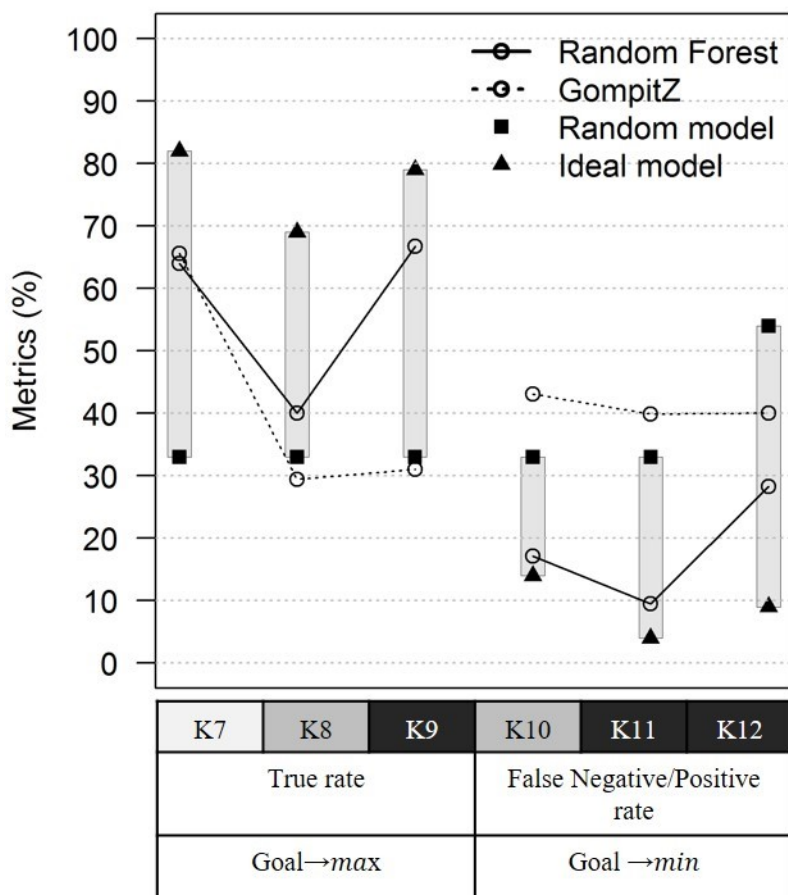


Figure 4.7. Comparison of model outcomes with a random and an ideal model

The following outcomes can be derived for the pipe level.

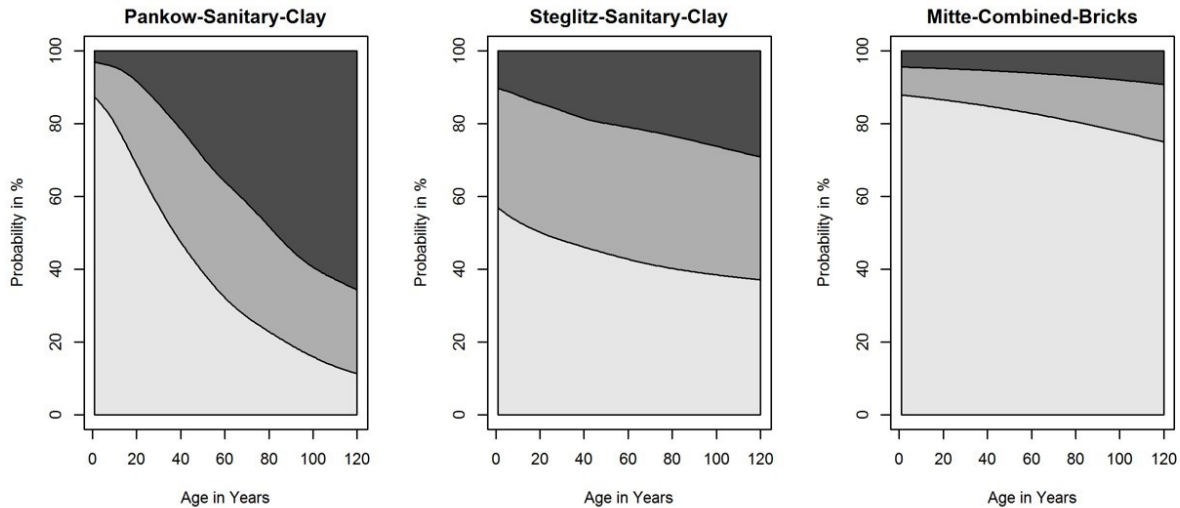
- GompitZ does not perform better than the random model, except for the simulation of pipes in good condition (K7). This result confirms the outcomes obtained in Braunschweig in chapter 3.
- The Random Forest performs much better than the random model. The performance is relatively good for the simulation of pipes in poor condition (K9) being close to the performance of the ideal model. This result is of interest since utilities might use modelling outcomes to identify pipes in poor condition and plan inspection programs. The False Negative rates (K10 and K11) are also very low, similar to the ideal model. On the other hand, the Random Forest model fails to identify pipes in medium condition (K8) and the False Negative rate is high compared to the ideal model (K12).

Simulation example

In order to visualize the outcomes of the Random Forest and GompitZ models, the condition of three single pipes with different characteristics have been simulated with both models (Figure 4.8). For example, the first column shows the deterioration process of a circular sanitary clay pipe situated in the Pankow district of Berlin (Figure 4.8 – left). The deterioration behaviour is similar with the two models: faster for clay pipes in the district “Pankow”, slower for clay pipes in the district “Steglitz” and much slower for brick pipes in the district “Mitte”. The deterioration is smoother with GompitZ since the survival function follows a Gompertz distribution (Le Gat, 2008). The deterioration with Random Forest is much sharper since the model learned from available data without the support of statistical regression. The sharpness of the Random Forest prediction is both the strength and the weakness of the model. It allows a more accurate prediction of the condition of specific pipes. However, it might also lead to doubtful predictions such as condition improvement along with pipe age. This leads to the

conclusion that the tested machine learning approach shall only be used for ad-hoc classification of the sewer pipes but not for long-term forecasts. This characteristic of machine learning models is still to be investigated in order to guarantee the plausibility of future predictions.

GompitZ - simulation of pipes condition



Random Forest - simulation of pipes condition

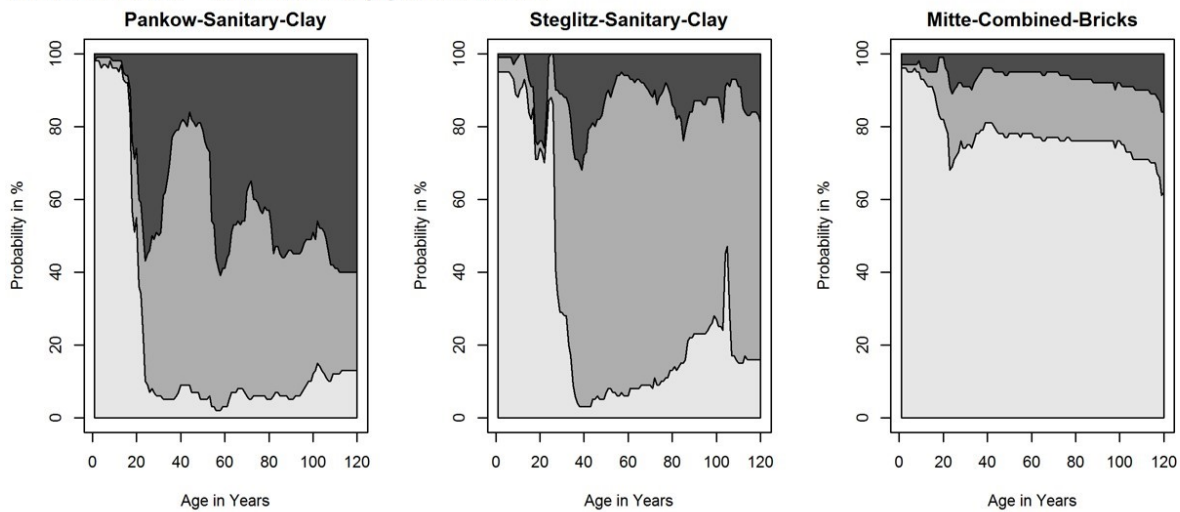


Figure 4.8. Example of simulation for three pipes with different characteristics (the title of each graph indicates pipe's characteristics in the following order: district – type of effluent – material). The colors light grey, medium grey and dark grey represent the probability for the pipe to be in good, medium and poor condition, respectively

4.4 Conclusion

This study aimed at assessing the performance of a statistical and a machine learning deterioration models using the extensive CCTV and GIS dataset of the city of Berlin, Germany. A set of straightforward and intuitive accessible metrics has been defined in dialogue with the local utility in order to evaluate sewer performance from an end user perspective. The selected metrics aim at convincing the municipality about the relevance or uselessness of using a given deterioration model to support asset management strategies. The main outcomes are summarized as follows.

- For both GompitZ and Random Forest models, the most important variables for calibration are the sewer age, the material, the shape and the type of effluent. The district is also a relevant variable being strongly correlated with pipes characteristics (for example pipes in the city centre are mainly combined sewers made of bricks and built in the 19th century) but also with other potential relevant deterioration factors such as the traffic load, the type of soil or the quality of the construction between the former east and west Berlin. The width, length and depth are also relevant for describing sewer deterioration but secondary compared to the material and district. Finally, the environmental features show little or no influence (type of soil, tree density, groundwater level).
- At network level, both Random Forest and GompitZ give satisfactory outcomes. Deviations between the predicted and inspected condition distributions, for the entire network and for different age groups, are <5% using Random Forest and even less (<1%) using GompitZ. Similar results have been obtained using GompitZ on the Braunschweig case study (Chapter 3). This result underlines the strong potential of both statistical and machine learning models to simulate the condition distribution of the network.

- At pipe level, GompitZ does not perform better than a simple random model, which attributes randomly a condition class to each inspected pipe. GompitZ is not able to simulate the condition of single pipes accurately.
- At pipe level, the Random Forest outperforms GompitZ. Random Forest performance is relatively good for the simulation of pipes in poor condition: 66.7% of the pipes inspected in poor condition have been predicted correctly and only 9.5% of the pipes inspected in poor condition have been falsely predicted in good condition. The True rate of Random Forest for pipes in poor condition (67%) is close to the True Positive rate of a CCTV inspection (79%). The Random Forest model shows a strong potential for supporting sewer operators in the identification of pipes in critical condition for inspection programs.
- The main weakness of the Random Forest model lies in its high False Positive rate: 28.3% of pipes predicted in poor condition are actually in good condition. This aspect of the performance might be improved in further studies by considering additional variables and testing other modelling approaches.
- Another weakness of the Random Forest model is the lack of physical information about pipe deterioration in the structure of the model. The model learns and reproduces the patterns observed in the inspection dataset. It can lead to doubtful prediction such as condition improvement along with pipe age. This leads to the conclusion that the tested machine learning approach shall only be used for ad-hoc classification of the sewer pipes but not for long-term forecasts into the future. This problem does not occur with GompitZ since the deterioration follows a GompertZ distribution that prevents any condition improvement. This aspect of machine learning should be carefully investigated before deploying such models in practice.

Chapter 5 Evaluation of uncertainties in sewer condition assessment

This chapter is an adapted version of the original paper:

Caradot N., Rouault P., Clemens F., Cherqui F. 2018. Evaluation of uncertainties in sewer condition assessment. *Structure and Infrastructure Engineering*, 14(2), 264-273.

<https://doi.org/10.1080/15732479.2017.1356858>

5.1 Overview and context of paper III

This paper has been published in the frame of the research project SEMA (2012-2016). Part of the analysis have been developed during a research exchange in Bogotá, Colombia, supported by DAAD with funds from the German Federal Ministry of Education and Research (BMBF).

This paper aims at

- Quantifying the uncertainty of the whole inspection process from CCTV inspection to condition assessment
- Determining the probability to underestimate, overestimate or accurately estimate the real condition of a pipe using CCTV inspection

This study introduces and demonstrates a methodology based on double inspections of the same pipes to determine the uncertainties of the structural condition assessment, i.e. the probability to estimate correctly the structural condition of a pipe from a CCTV inspection. The methodology is based on a non-linear optimization procedure coupled with a Monte-Carlo simulation and has been used to determine the probabilities of False Positive and False Negative by inspecting a pipe in a given condition.

Compared to the published version in the journal Structure and Infrastructure Engineering, a section has been added to evaluate the uncertainties for the inspection of the network of the city of Berlin.

5.2 Research method

5.2.1 List of notations

Name	Description	Status
N_{ij}	Number of pipes inspected twice, first in condition "i" and then in condition "j"	Known
N	$N = (N_{ij})$ 4x4 double inspection matrix	
$\sum N_{ij}$	Number of inspected pipes twice	
R_i	Number of pipes really in condition "i"	Unknown
R	$R = \{R_1, R_2, R_3, R_4\}$ vector	
$P(\beta = i)$	Probability for a pipe to be inspected in condition "i"	Unknown
$P(\alpha = i)$	Probability for a pipe to be really in condition "i"	Unknown
$P(\beta = i \alpha = j)$	Probability for a pipe to be inspected in condition "i" when really in condition "j"	Unknown
M	$M = (P(\beta = i \alpha = j))$ 4x4 uncertainty matrix	
$N_{estimated}$	$N_{estimated} = (N_{estimated_{i,j}})$ Estimated double inspection matrix	Unknown

5.2.2 General concept

Let's assume that each inspected pipe has a *real* structural condition that describes the rehabilitation needs. The *real* condition is defined as the sewer internal condition that would potentially lead to the best rehabilitation decision (from a structural perspective only, since other factors might be considered in the decision-making process). The *real* condition of the pipe is unfortunately unknown. The best estimation of the *real* condition of a pipe would be the average internal condition (or mode) obtained with a high (infinite) number of repeated CCTV inspections. In practice it is impossible to be sure to estimate correctly the real condition of a

pipe since there is no warranty that a pipe inspected twice (or even three or four times) in the same condition has been correctly inspected. Even if the inspected conditions are mutually consistent, they can be consistently wrong.

The *real* condition of the pipe can be estimated with an *inspected* condition, following the steps of CCTV visual inspection, sewer defect coding and sewer condition assessment (Figure 2.1). The *inspected* condition might estimate correctly the *real* condition but can also underestimate or overestimate the *real* condition since uncertainties affect each step of the condition assessment procedure.

Aim of the analysis is to determine the uncertainty matrix M .

$$M = (P(\beta = i|\alpha = j))_{i,j} \quad (5.1)$$

The probability $P(\beta = i|\alpha = j)$ expresses the conditional probability to be inspected in condition "i" when a pipe is really in condition "j". The term α indicates the real condition of a pipe (which is unknown). The probability for a pipe to be really in condition "j" is given by $P(\alpha = j)$. The term β indicates the inspected condition of a pipe (which is known). The probability for a pipe to be inspected in condition "i" is given by $P(\beta = i)$. To simplify the further development we consider four condition classes only, 4 being the worst condition indicating an urgent rehabilitation need.

$$i, j \in \{1,2,3,4\}$$

The probability $P(\beta = i|\alpha = j)$ is illustrated by the decision tree in Figure 5.1.

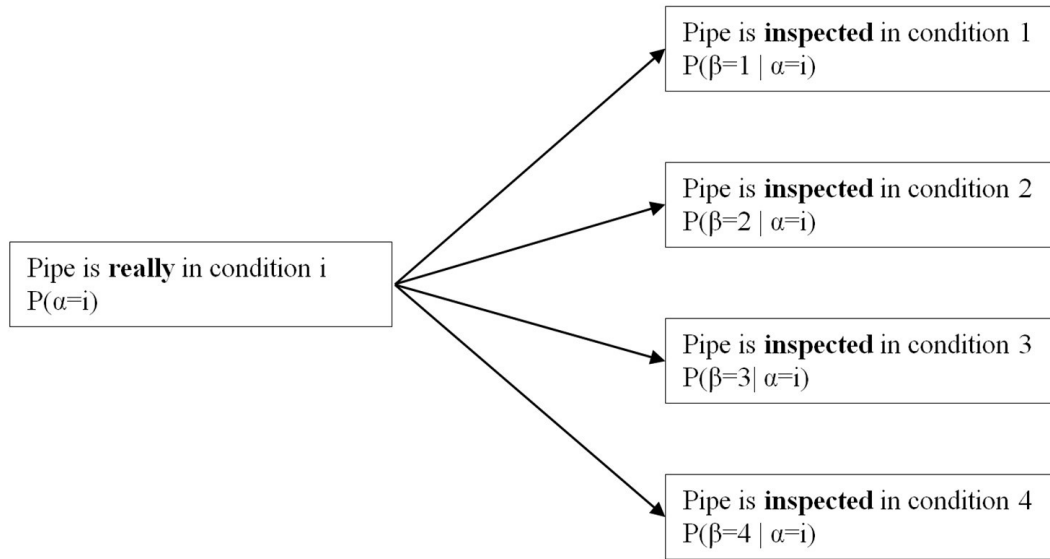


Figure 5.1. Decision tree for $P(\beta = i | \alpha = j)$.

M gives the probability to be inspected in condition "i" when a pipe is really in condition "j".

$$M = \begin{pmatrix} P(\beta = 1|\alpha = 1) & P(\beta = 1|\alpha = 2) & P(\beta = 1|\alpha = 3) & P(\beta = 1|\alpha = 4) \\ P(\beta = 2|\alpha = 1) & P(\beta = 2|\alpha = 2) & P(\beta = 2|\alpha = 3) & P(\beta = 2|\alpha = 4) \\ P(\beta = 3|\alpha = 1) & P(\beta = 3|\alpha = 2) & P(\beta = 3|\alpha = 3) & P(\beta = 3|\alpha = 4) \\ P(\beta = 4|\alpha = 1) & P(\beta = 4|\alpha = 2) & P(\beta = 4|\alpha = 3) & P(\beta = 4|\alpha = 4) \end{pmatrix} \quad (5.2)$$

We assume that M is an unknown constant and corresponds to the average uncertainty of the condition assessment procedure. This assumption is required because some factors that are known to influence the inspection results are not available. Visual inspection can be influenced by the experience of the operator and his age/sex/education (Hassan & Diab, 2010; Heidl *et al.*, 2013; Laofor & Peansupap, 2012); and by the condition of the investigation such as equipment used (Plihal *et al.*, 2014), cleaning condition, lightning, disturbances (Gramopadhye & Wilson, 1997).

The uncertainty matrix can be used to estimate the *inspected* condition distribution P from the *real* condition distribution R of the segments.

$$P = MR \quad (5.3)$$

Similarly, the *real* condition distribution can be estimated from an *inspected* condition distribution, providing that the matrix M is invertible, which would define R uniquely.

$$R = M^{-1}P \quad (5.4)$$

Therefore, if M is invertible (non-singular matrix with determinant not equal to 0), a given *inspected* condition distribution gives a unique *real* condition distribution.

R: number of pipes really in each condition

From a set of inspected pipes, the unknown number of pipes really in each condition is expressed as:

$$R = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix} \quad (5.5)$$

The total number of segments ($R_1 + R_2 + R_3 + R_4$) is equal to the number of segments inspected twice in the matrix N ($\sum_{i,j} N_{ij}$).

N: number of pipes inspected twice, first in condition "i" and then in condition "j"

From the inspection database, if several pipes have been inspected twice, we can build the double inspection matrix N such as $N = (N_{ij})$. Each cell of the matrix contains the number of pipes inspected twice, first in condition class "i" and then in condition class "j". We assume a short period between the repeated inspections so we can neglect sewer deterioration: in this case the matrix is (almost) symmetric and the repeated inspections are considered independent.

$$N_{ij} \approx N_{ji} \quad (5.6)$$

$$N = \begin{pmatrix} N_{11} & N_{21} & N_{31} & N_{41} \\ N_{12} & N_{22} & N_{32} & N_{42} \\ N_{13} & N_{23} & N_{33} & N_{43} \\ N_{14} & N_{24} & N_{34} & N_{44} \end{pmatrix} \approx \begin{pmatrix} N_{11} & - & - & - \\ N_{12} & N_{22} & - & - \\ N_{13} & N_{23} & N_{33} & - \\ N_{14} & N_{24} & N_{34} & N_{44} \end{pmatrix}$$

Relation between N, R and M

The values N_{ij} can be estimated using M and R with the following expression:

$$N_{ij} \cong \sum_{k=1}^4 R_k M_{ik} M_{jk}$$

$$N_{ij} \cong R_1 . P(\beta = i | \alpha = 1) . P(\beta = j | \alpha = 1) +$$

$$R_2 . P(\beta = i | \alpha = 2) . P(\beta = j | \alpha = 2) +$$

$$R_3 . P(\beta = i | \alpha = 3) . P(\beta = j | \alpha = 3) +$$

$$R_4 . P(\beta = i | \alpha = 4) . P(\beta = j | \alpha = 4) \quad (5.7)$$

The number of pipes inspected first in condition "i" and then in condition "j" is equal to the number of pipes in every condition inspected first in condition "i" and then in "j".

We are now able to create a system of 10 equations: 16 equations in total but since the matrix N is symmetric 6 equations are double. The system is composed of 20 variables with 15 degrees of freedom.

- R_4 can be estimated from R_1 , R_2 and R_3 as well as from the number of inspected pipes $\sum N_{ij}$. R has 4 variables with 3 degrees of freedom

$$R_4 = \sum N_{ij} - R_1 - R_2 - R_3 \quad (5.8)$$

- M can be estimated by knowing the 3x4 upper part of the matrix since the sum of columns is equal to 1. M has 16 variables with 12 degrees of freedom. If a pipe is really in condition "i", it will necessarily be inspected in condition 1, 2, 3 or 4:

$$P(\beta = 1 | \alpha = i) + P(\beta = 2 | \alpha = i) + P(\beta = 3 | \alpha = i) + P(\beta = 4 | \alpha = i) = 1 \quad (5.9)$$

Resolution of the system of non-linear equations

The system of 10 equations and 15 variables is resolved using the global optimization method

ISRES - Improved Stochastic Ranking Evolution Strategy (Runarsson & Yao, 2005). ISRES is available for different programming languages in the open-source library NLOpt for nonlinear optimization (Johnson, 2014). This method has been chosen for being a derivative-free optimization (no need to assess the gradient of the optimization function) and for supporting bound constraints as well as nonlinear inequality constraints. The nonlinear optimization method minimizes an objective function defined as the difference between the observed double inspections matrix and its estimation:

$$\min \sum (N_{ij} - N_{estimated_{i,j}})^2 \quad (5.10)$$

The method identifies the set of variables that minimize the objective function, i.e. that allow the most accurate estimation of the double inspection matrix (N_{ij}). Intuitively, for a given condition distribution (R) the optimisation method aims at finding the uncertainty values (M) that can explain the deviations of the double inspection matrix (N). Since R is unknown, the method attributes random values to R and then calculates the corresponding values of M to reproduce N .

The 15 variables are $P_{11}, P_{12}, P_{13}, P_{14}, P_{21}, P_{22}, P_{23}, P_{24}, P_{31}, P_{32}, P_{33}, P_{34}, R_1, R_2$ and R_3 . The variable constraints are defined as following:

- $P_{ij} \in [0,1]$ – Probabilities vary from 0 to 1
- $P_{ii} > P_{ji}$ – Probabilities to be inspected in the right condition is higher than the probability to be inspected in a wrong condition
- $R_i \in [0, \sum N_{ij}]$ – The number of pipes in each condition is smaller than the total number of pipes

The optimization is run 50 times using a Monte-Carlo simulation with a random selection of the starting variable values within the defined constraints field. The Monte-Carlo simulation aims at assessing the influence of the starting values on the stability of the results. Outcomes of the 50 Monte-Carlo runs are the mean and standard deviation of R and M .

Calculation of the probabilities of false positive and false negative

The matrix M is used to determine the probabilities of False Positive and False Negative by inspecting a pipe in a given condition. A False Negative (FN) occurs when the inspected condition overestimates the real condition: the inspected condition is better than the real condition (e.g. defects are actually present but have been missed by the operator). A False Positive occurs (FP) when the inspected condition underestimates the real condition: the inspected condition is worse than the real condition (e.g. the operator exaggerates the size of a crack).

$$P(FN|\alpha = i) = \sum_j P(\beta = j|\alpha = i) \text{ with } j < i \quad (5.11)$$

$$P(FP|\alpha = i) = \sum_j P(\beta = j|\alpha = i) \text{ with } j > i \quad (5.12)$$

5.3 Evaluation of uncertainties in Braunschweig

5.3.1 Data description

This study has been performed using the GIS and CCTV database of the city of Braunschweig in Germany (Refer to chapter 2.4 for a presentation of the case study and available data). After the data clean-up, the database contains 45,049 inspections with an inspected length of 1,784 km.

Figure 5.2 shows the number of pipes inspected every year and Figure 5.3 shows the number of inspections per pipe. Almost 50% of the pipes have been inspected at least twice.

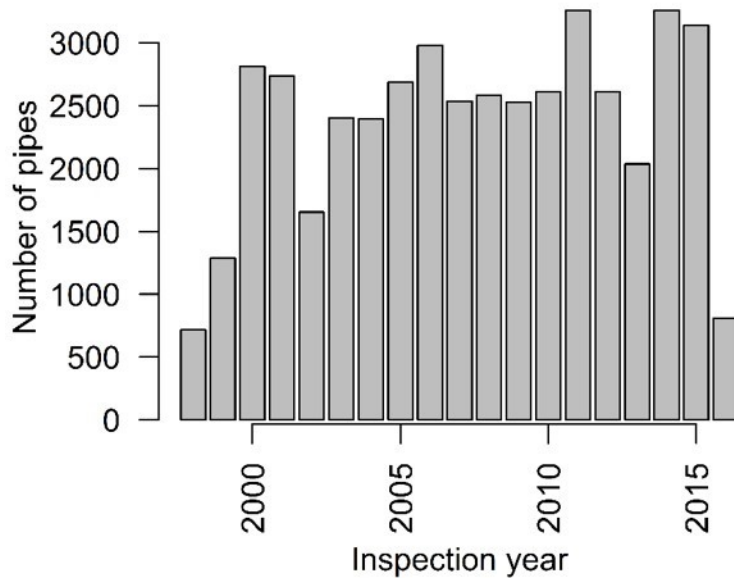


Figure 5.2. Number of pipes inspected every year.

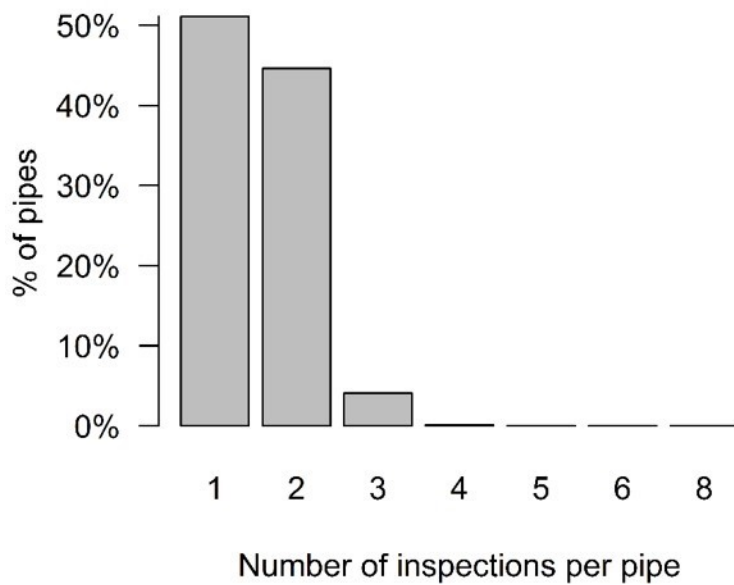


Figure 5.3. Proportion of pipe vs. number of inspections per pipe.

Figure 5.4 shows the distribution of durations between the repeated inspections. According to the local regulation, the entire network must be inspected every 10 years; indeed most second inspections have been performed between 8 and 10 years after the first inspection.

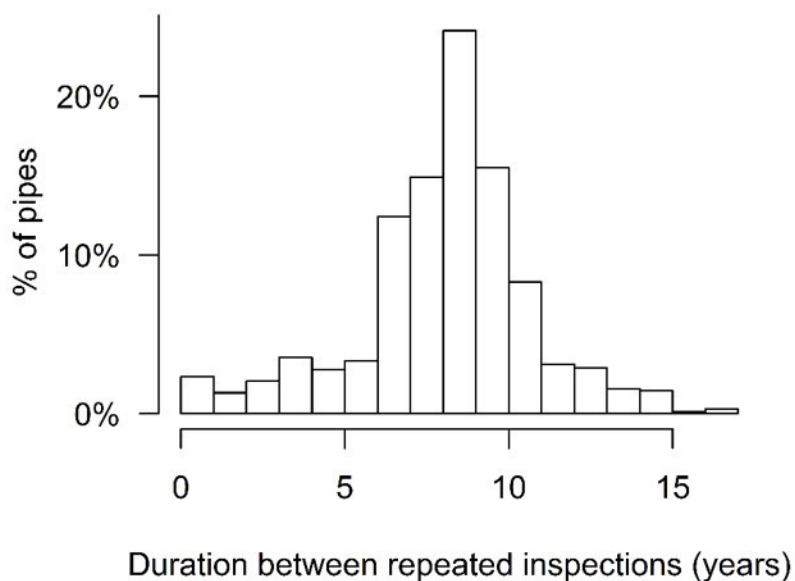


Figure 5.4. Percentage of pipes for different years between inspections.

The structural condition class of the inspected pipes is evaluated using an adaptation of the French classification methodology RERAU (based on dysfunction “COL” indicating the risk of collapse; see Ahmadi *et al.* 2014b; Le Gauffre *et al.* 2004). The aim of this methodology is to rank inspected sewer pipes based on the urgency of their rehabilitation needs. A structural condition class is assigned to each sewer segment (from manhole to manhole) on a four-grade scale (1 to 4, 4 being the worst condition and in need for immediate rehabilitation). The structural condition class is calculated using the characterization and quantification of sewer defects such as fissures, corrosion and surface damages that may lead to structural failure such as a pipe collapse.

The matrix of double inspection shows the number of pipes inspected first in condition "i" and then in condition "j" and can be expressed as:

$$N = \begin{pmatrix} 6125 & 484 & 644 & 106 \\ 1295 & 774 & 459 & 121 \\ 1190 & 502 & 1681 & 298 \\ 407 & 318 & 599 & 659 \end{pmatrix}$$

It is relevant to note that the off diagonal elements in the lower triangle of the matrix are larger than those in the upper triangle. For example, 1,295 pipes have been inspected first in condition

1 and then in condition 2 but only 484 pipes have been inspected first in 2 and then in 1. This behaviour is linked with the deterioration process between the repeated inspections. Most pipes have been inspected the second time more than 8 years after the first inspection and many of them switched in between to the next worst condition class.

Three reasons can explain a condition transition between first and second inspections: i) the condition of the second inspection is worst due to degradation; ii) the condition of the second inspection is better due to rehabilitation; iii) the condition has changed due to uncertainties in the condition assessment procedure. In order to observe the transitions due to uncertainties only, the pipes that undergo a condition transition due to reason i) or ii) must be removed from the database. Rehabilitated pipes have already been filtered out by preparing the dataset. In order to remove the influence of the deterioration process, the pipes with a period between repeated inspections higher than 3 years have been also filtered out. Considering the life duration of the pipes, we assume that the probability to observe a condition transition due to degradation within three years is not significant.

$$N = \begin{pmatrix} 340 & 39 & 48 & 18 \\ 38 & 37 & 24 & 9 \\ 45 & 33 & 111 & 27 \\ 13 & 10 & 28 & 71 \end{pmatrix}$$

The obtained matrix is almost symmetric: the average deviation between the down and upper part is small (11%) which indicates that the deterioration is not relevant anymore or at least less relevant than the uncertainty related to the inspection. Different periods have been tested (ranging from one to ten years). Within a period of 3 years we can assume that the deterioration is insignificant and that two inspections of a same pipe are statistically independent: the number of pipes inspected first in "i" and then in "j" is similar to the number of pipes inspected first in "j" and then in "i". Higher periods do not give such symmetrical matrix, and with a period smaller than 3 years, the matrix remains symmetric, however the number of pipes becomes too small to run the analysis. In order to apply the methodology the matrix is forced to be

symmetric. The mean of the lower and upper off-diagonal triangles of the original matrix is computed and used to replace both down and upper parts.

$$N = \begin{pmatrix} 340 & 38 & 46 & 16 \\ 38 & 37 & 28 & 10 \\ 46 & 28 & 111 & 28 \\ 16 & 10 & 28 & 71 \end{pmatrix}$$

The number of pipes with differences between structural condition classes is presented in Figure 5.5. About 65% of the pipes have been inspected twice in the same conditions whereas 35% of the pipes have a different condition between the repeated inspections.

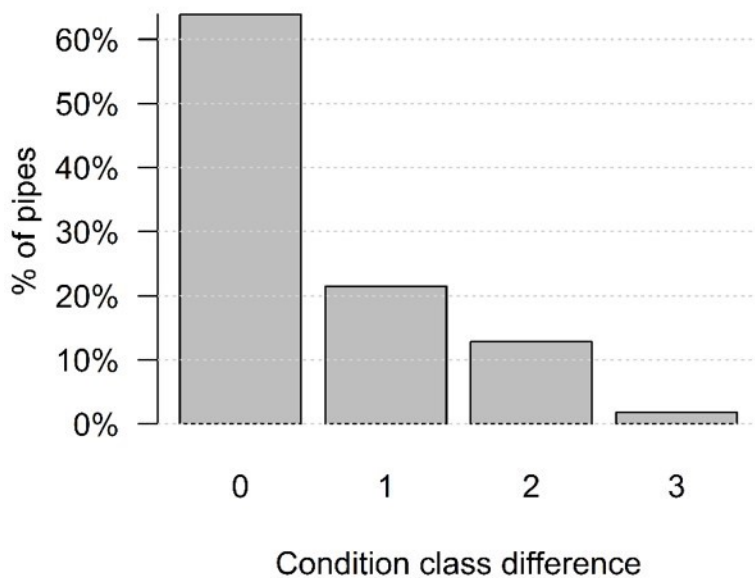


Figure 5.5. Percentage of pipes with differences between the structural condition classes.

5.3.2 Results and discussion

The optimisation procedure has been applied on the double inspection matrix N . Table 5.1 and Figure 5.6 summarizes the obtained mean and standard deviation for the variable R (real condition distribution of the segments) and M (uncertainty matrix).

Table 5.1. Outcomes (mean and standard deviation “sd”) from the optimisation procedure with 50 Monte-Carlo simulations

<p>M: uncertainty matrix</p> $\text{mean}(M) = \begin{pmatrix} 84.5 & 14.2 & 10.1 & 5.3 \\ 6.9 & 58.8 & 9.3 & 3.3 \\ 6.7 & 18.6 & 70.1 & 12.3 \\ 2 & 8.4 & 10.5 & 79.1 \end{pmatrix}$ $\text{sd}(M) = \begin{pmatrix} 3.3 & 7.7 & 6.9 & 2.9 \\ 1.2 & 8.7 & 4.6 & 2.3 \\ 2.7 & 10.4 & 6.4 & 3.8 \\ 0.7 & 3.6 & 2.8 & 1.8 \end{pmatrix}$	
<p>R: number of pipes really in each condition</p> $\text{mean}(R) = \begin{pmatrix} 469 \\ 101 \\ 213 \\ 108 \end{pmatrix}$ $\text{sd}(R) = \begin{pmatrix} 42 \\ 36 \\ 46 \\ 5 \end{pmatrix}$	
<p>N: number of pipes inspected twice, first in condition "i" and then in condition "j"</p> $\text{mean}(N_{\text{estimated}}) = \begin{pmatrix} 340 & 38 & 46 & 16 \\ 38 & 37 & 28 & 10 \\ 46 & 28 & 111 & 28 \\ 16 & 10 & 28 & 71 \end{pmatrix}$ $\text{sd}(N_{\text{estimated}}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ $N - N_{\text{estimated}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	

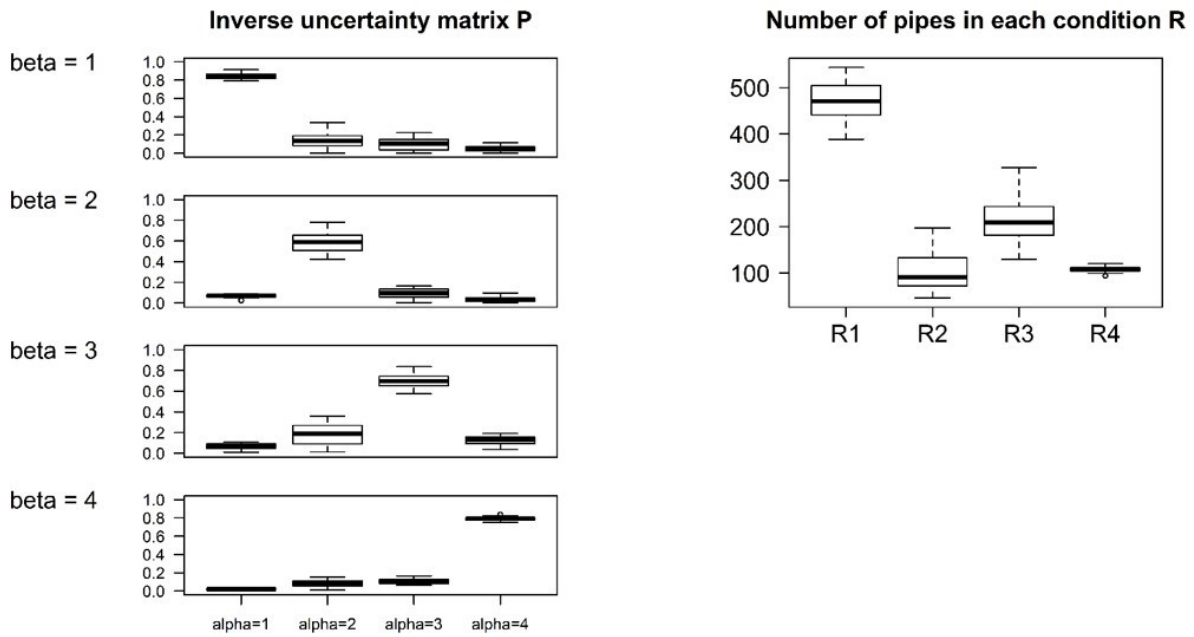


Figure 5.6. Deviations of M and R from 50 Monte-Carlo simulations; the left graphs plot the values of the four rows of the matrix P.

The optimization procedure was successfully applied. For each Monte-Carlo run, the objective function converged towards 0 ($<10^{-5}$) so $N = N_{estimated}$, independently from the variable starting values. The standard deviation of the obtained variables M and R is relatively low indicating that the optimization procedure finds a global optimum instead of multiple local optimums. M has been used to compute the probabilities of false positive and false negative with equations 5.11 and 5.12 (Table 5.2).

Table 5.2. Calculation of FP and FN for sewer condition assessment (in %); the condition ranges from 1 to 4, 4 being the worst condition indicating an urgent rehabilitation need

	Real condition i				
	1	2	3	4	
$P(FN \alpha = i)$	-	14.2	19.4	20.9	Inspected condition is better than the real condition
$P(\beta = i \alpha = i)$	84.5	58.8	70.1	79.1	Inspected condition is the real condition
$P(FP \alpha = i)$	15.5	27	10.5	-	Inspected condition is worse than the real condition

Several outcomes can be highlighted:

- The probability to be inspected in the right condition (P_{ii}) is higher for pipes in good condition (probability of 84.5% for pipes in condition 1) than for pipes in poor condition (probability of 79.1% for pipes in condition 4). It indicates that there is less uncertainty by assessing the condition of sewers with few defects only or without severe defects. This is an expected outcome: the inspectors that perform the defect coding are more prone to error when many defects are present as when sewers are in perfect condition.
- The probability to be inspected correctly in condition 2 is only 58.8% indicating high uncertainties especially for this condition class. This might be explained by the fact that only a few pipes are in condition 2 compared to pipes in condition 1 and 3. The condition classification method uses fine thresholds to separate conditions 1, 2 and 3. Given the high uncertainties in identifying pipes in condition 2, this condition might be merged with condition 1 without losing information.
- The probability to be inspected correctly in condition 4 is 79.1%. It means that there is a probability of 20.9% to be wrong by overestimating the real condition of the pipe (FN, i.e. inspected condition is better than the real condition). FN errors can have major consequences: major defects leading to failure or collapse might be missed (Ahmadi *et al.*, 2014b). Since rehabilitation programs are based on structural condition evaluation, the influence of this uncertainty on rehabilitation decisions remains to be evaluated.
- For pipes in poor condition (3 and 4), the probability of FN is significantly higher than the probability of FP. It means that there is a higher probability to be too optimistic (e.g. by missing defects). Dirksen *et al.* (2013) found out that the probability that the inspector fails to recognise the presence of a defect is significantly higher than the

probability that a defect is reported although it is not present: FN was in the order of 25% whereas FP in the order of few percent.

- For pipes in good condition (1 and 2), the probability of FP is significantly higher than the probability of FN. It means that there is a higher probability to be too pessimistic.

It is to note that these results describe the average uncertainty of sewer condition assessment. The condition assessment is far to be a homogeneous procedure with standard operating conditions. Several factors might influence the outcome: e.g. the operator skills, the inspection velocity, the light and cleaning quality, the resolution of the CCTV camera, etc. Further research might focus on the evaluation of uncertainty for specific operational conditions or highlight the factors that have most influence on the accuracy of sewer condition assessment.

5.4 Evaluation of uncertainties in Berlin

5.4.1 Data description

The study has been performed using the CCTV database of the city of Berlin in Germany (Refer to chapter 2.4 for a presentation of the case study and available data). After data clean-up, the database contains 124,450 inspections with a length of 5,222 km over 107,788 pipes. The number of inspections is higher than the number of pipes because several pipes have been inspected more than once.

Among the 124,450 available inspections, 13,753 segments have been inspected at least twice. Figure 5.7 shows the distribution of the inter-inspection time intervals. Most segments have been inspected twice with a time lag of 5-10 years between the two inspections (> 50%). It is also interesting to note that many pipes (13%) have been inspected twice in the same year. This is because the inspections are not recorded immediately in the database (due to delays in the data flow of the organisation and associated IT systems) so several inspection teams might inspect the same pipe, without knowing it has already been inspected.

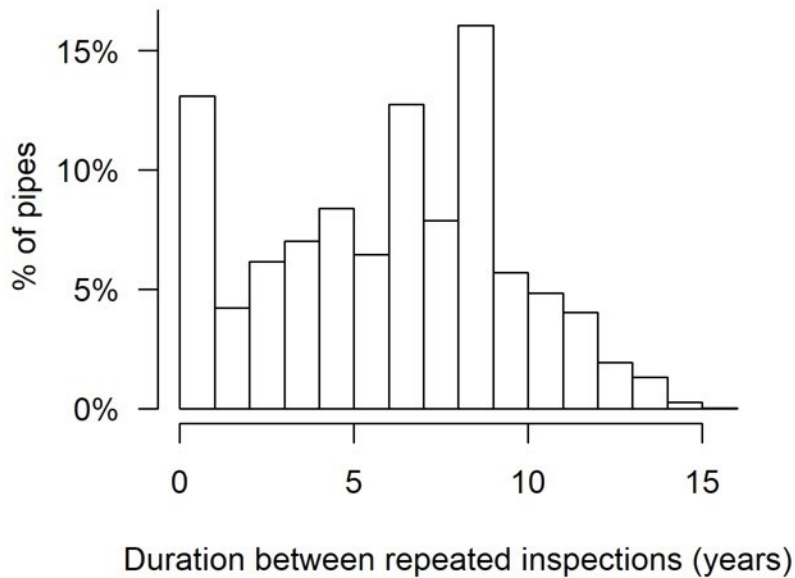


Figure 5.7. Histogram of the duration between repeated inspections.

In order to analyse only the uncertainties in condition assessment and reduce the influence of the deterioration process, segments with a period between the repeated inspections higher than 5 years have been filtered out. Considering the life duration of the segments, we assume that the probability to observe a condition transition due to degradation within 5 years is not significant. Figure 5.8 shows the deviations between first and second inspections: about 65% of the segments have been inspected twice in the same conditions whereas for 35% a change in condition class has been observed between the two inspections. The distribution of the deviations is symmetric which means that condition class deteriorations are observed as frequently as improvements. This observation supports our hypothesis that a timeframe of 5 years is suitable to observe the uncertainties and remove the influence of sewer deterioration.

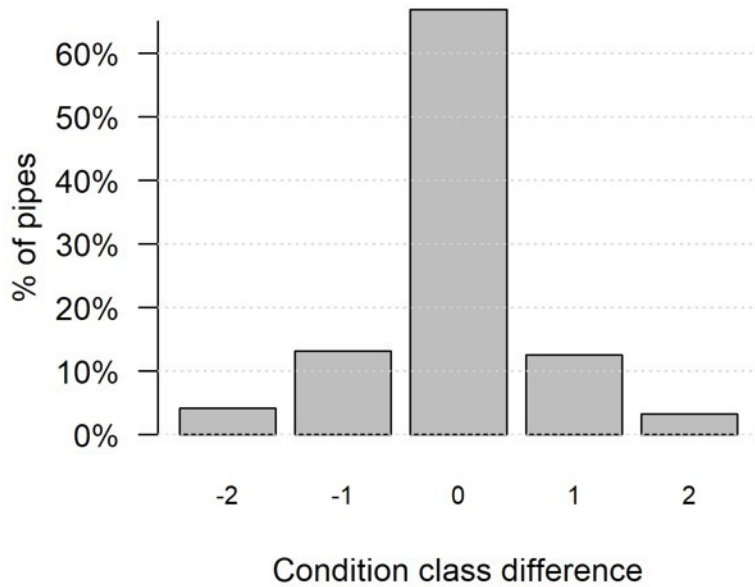


Figure 5.8. Deviations between first and second inspections.

The matrix of double inspection N shows the number of segments inspected first in condition "i" and then in condition "j":

$$N = \begin{pmatrix} 1862 & 338 & 193 \\ 373 & 506 & 275 \\ 152 & 211 & 745 \end{pmatrix}$$

For example, 338 segments have been inspected first in condition class 2 and then in condition class 1 (best condition). The matrix is almost symmetrical indicating that the repeated inspections are independent. A detailed analysis of the reasons for the observed deviations has been conducted by the water company. They are summarized as follows:

- Database errors: e.g. a rehabilitation (replacement, repair or liner) has been done between the two inspections but has not been properly documented.
- Environmental conditions: different environmental conditions between the two inspections are responsible for the different condition scores: e.g. variation of the water level, visibility of cracks due to higher groundwater level, etc.

- Subjectivity of the inspector: the different interpretation of the type, characterisation and quantification of defects by the operator leads to different condition scores.
- Goal of the inspection: the main purpose of sewer inspection is generally to assess the structural condition of the segment and record all defects. However, in some specific cases, inspections are performed for another specific purpose such as identifying house connections. These inspections are recorded in the database even if they should not be used to assess the structural condition.

Uncertainty sources can be responsible for both condition improvement and deterioration between the repeated inspections. In order to apply the methodology, the matrix is forced to be symmetric. The mean of the corresponding values above and below the diagonal of the matrix is computed and used to replace the original values.

$$N = \begin{pmatrix} 1862 & 356 & 172 \\ 356 & 506 & 243 \\ 172 & 243 & 745 \end{pmatrix}$$

It is interesting to note that that uncertainty in sewer condition assessment is not only due to the subjectivity of the inspectors but also depends on other reasons. In other words, the analysis of repeated inspections provides information on the uncertainty of the input data used for deterioration modelling rather than only on the uncertainty of the inspection procedure itself.

5.4.2 Results and discussion

The optimisation procedure has been applied on the double inspection matrix N . Table 5.3 summarized the mean and standard deviation obtained for the variable R (real condition distribution of the segments) and M (uncertainty matrix).

Table 5.3. Outcomes (mean and standard deviation “sd”) from the optimisation procedure with 1000 Monte-Carlo simulations

<i>M</i> : uncertainty matrix	
$mean(M) = \begin{pmatrix} 85.2 & 15.9 & 6.1 \\ 10.6 & 67.3 & 12.5 \\ 4.2 & 16.8 & 81.4 \end{pmatrix}$	$sd(M) = \begin{pmatrix} 4.0 & 9.5 & 4.1 \\ 3.4 & 8.4 & 5.8 \\ 1.8 & 7.7 & 6.0 \end{pmatrix}$
<i>R</i> : number of segments really in each condition	
$mean(R) = \begin{pmatrix} 2518 \\ 1055 \\ 1081 \end{pmatrix}$	$sd(R) = \begin{pmatrix} 282 \\ 321 \\ 190 \end{pmatrix}$
<i>N</i> : number of segments inspected twice, first in condition "i" and then in condition "j"	
$mean(N_{estimated}) = \begin{pmatrix} 1862 & 356 & 172 \\ 356 & 506 & 243 \\ 172 & 243 & 745 \end{pmatrix}$	$sd(N_{estimated}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
$N - N_{estimated} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	

For each Monte-Carlo run, the objective function converged towards 0 ($< 10^{-5}$), independently from the starting variable values. The mean value and standard deviation of the obtained variables *M* and *R* are relatively low, indicating that the optimisation procedure finds a global optimum instead of multiple local optimums. *M* has been used to compute the probabilities of false positive and false negative with equations 5.11 and 5.12 (Table 5.4).

Table 5.4. Calculation of FP and FN for sewer condition assessment (in %); the condition ranges from 1 to 3, 3 being the worst condition indicating an urgent rehabilitation need

	Real condition <i>i</i>			
	1	2	3	
$P(FN \alpha = i)$	-	15.9	18.6	Inspected condition is better than the real condition
$P(\beta = i \alpha = i)$	85.2	67.3	81.4	Inspected condition is the real condition
$P(FP \alpha = i)$	14.8	16.8	-	Inspected condition is worse than the real condition

Main outcomes can be highlighted as follows.

- For segments in good condition 1, the probability of assessing correctly the condition is 85.2%. On the other hand, the probability of to underestimate the condition is 14.8% (False Positive). In this case, the inspected condition is worse than the real condition, e.g. due to a pessimistic description of the observed defects.
- For segments in poor condition 3, the probability of assessing correctly the condition is 81.4%. On the other hand, the probability to overestimate the condition is 18.6% (False Negative). In this case, the inspected condition is better than the real condition, e.g. due to missing defects by the operator.
- The probability to inspect correctly a segment in intermediate condition 2 is lower because the assessment can lead to both False Positive and False Negative (overestimation or underestimation of the real condition). The intermediate condition is more ambiguous to assess as the segment neither is in perfect condition, nor already failed. Another reason can be the narrow range of defect characterisation and quantification that leads to intermediate condition 2.

5.5 Conclusion

This study introduced and demonstrated a methodology based on double inspections of the same pipes to determine the uncertainties of the structural condition assessment. The methodology is based on a non-linear optimization procedure coupled with a Monte-Carlo simulation and has been used to determine the probabilities of false positive and false negative by inspecting a pipe in a given condition. This study has been performed using the extensive CCTV databases of the cities of Braunschweig and Berlin in Germany. The main outcomes are summarized below:

- The probability to inspect correctly a pipe in poor condition is close to 80-85% and thus the probability to overestimate the condition of the pipe is close to 15-20% (False Negative, i.e. give a too optimistic image).
- The probability to inspect correctly a pipe in good condition is slightly higher than the probability to inspect correctly a pipe in poor condition. It indicates that there is less uncertainty by assessing the condition of sewers with few defects only or without severe defects. This is an expected outcome: the inspectors that perform the defect coding are more prone to error when many defects are present as when sewers are in perfect condition.
- The uncertainties in sewer condition assessment are not only due to the inspection procedure. The analysis of deviations in repeated inspections by the water utility of Berlin highlights further sources of uncertainties such as undocumented rehabilitation (i.e. the rehabilitation of segment has not been documented in the database) or inspections done on a specific purpose and leading to an incorrect inventory of defects, e.g. to identify house connections. The reduction of uncertainties could start by the improvement of data management procedures in order to be able to filter improper inspections before to calibrate sewer deterioration models.

The analysis of uncertainties on two German case studies deliver consistent outcomes. The methodology proposed could be used again to confirm these outcomes using other data and to assess uncertainties for condition assessment of other networks.

Chapter 6 The influence of condition assessment uncertainties in sewer deterioration modelling

This paper is an adapted version of the paper:

Caradot N., Riechel M., Rouault P., Caradot A., Lengemann N., Eckert E., Ringe A., Clemens F., Cherqui F. 2019. The influence of condition assessment uncertainties on sewer deterioration modelling. *Structure and Infrastructure Engineering*, accepted 05/2019.

6.1 Overview and context of paper IV

This paper has been prepared in the frame of the project SEMA-Berlin 2 (2018-2019) financed by the Berliner Wasserbetriebe. This project aims at developing practical deterioration models in Berlin in order to support the planning of sewer inspection strategies as well as long-term sewer rehabilitation and investment strategies.

This paper aims at assessing the influence of sewer condition uncertainty on the shape of the survival curves and the prediction outcomes of a deterioration model. First, the methodology proposed in chapter 5 has been applied to quantify uncertainties in sewer condition assessment from the analysis of a set of repeated inspections. Then, a method has been proposed to propagate uncertainties in the survival curves and predictions of a deterioration model. The deterioration model has been used to simulate simple long-term strategies and evaluate the impact of uncertainties over model prediction. The method has been demonstrated using the unique inspection dataset of the city of Berlin, Germany, where 13,753 segments have been inspected at least twice.

6.2 Material and method

6.2.1 Uncertainties in sewer condition assessment

Chapter 5 proposed a methodology to assess uncertainties of sewer condition assessment based on the analysis of repeated inspections of the same segments. The aim was to determine the probability to underestimate, overestimate or accurately estimate the real condition of a segment using CCTV inspection. The methodology assumes that each inspected segment has a *real* structural condition that describes the rehabilitation needs. The *real* condition is defined as the sewer internal condition that would lead to the best rehabilitation (or no-rehabilitation) decision. The *real* condition of the segment is unfortunately unknown but can be estimated with an *inspected* condition, following the steps of CCTV visual inspection, sewer defect coding and sewer condition assessment. The *inspected* condition might estimate correctly the *real* condition but can also underestimate or overestimate it since uncertainties affect each step of the condition assessment procedure.

Uncertainties can be expressed in a matrix $M = P(\beta = i|\alpha = j)_{i,j}$ which gives the conditional probability to be inspected in condition "i" when a segment is really in condition "j".

$$M = \begin{pmatrix} P(\beta = 1|\alpha = 1) & P(\beta = 1|\alpha = 2) & P(\beta = 1|\alpha = 3) \\ P(\beta = 2|\alpha = 1) & P(\beta = 2|\alpha = 2) & P(\beta = 2|\alpha = 3) \\ P(\beta = 3|\alpha = 1) & P(\beta = 3|\alpha = 2) & P(\beta = 3|\alpha = 3) \end{pmatrix} \quad (6.1)$$

The term α indicates the real condition of a segment (which is unknown). The term β indicates the inspected condition of a segment (which is known). The uncertainty matrix can be used to estimate the *inspected* condition distribution P from the *real* condition distribution R of the segments.

$$P = MR \quad (6.2)$$

Similarly, the *real* condition distribution can be estimated from an *inspected* condition

distribution, providing that the matrix M is invertible, which would define R uniquely.

$$R = M^{-1}P \quad (6.3)$$

Therefore, if M is invertible (non-singular matrix with determinant not equal to 0), a given *inspected* condition distribution gives a unique *real* condition distribution.

Chapter 5 describes the steps of the optimisation methodology, which leads to the determination of the uncertainty matrix M . The optimisation is run 1,000 times in order to deliver the mean and standard deviation of the uncertainty matrix M .

6.2.2 Propagation of the uncertainties in the calibration and prediction of a deterioration model

The obtained uncertainty matrix M can be used to propagate uncertainties in any survival model that predicts a probability class output (probability for a segment to be in a given condition class) from the segment's age and a set of numerical or categorical variables. The survival curves SC of a survival model give the probability to be in each condition at a given segment age T . For example, with three condition classes:

$$P(T) = (P_1(T), P_2(T), P_3(T)) = (SC_1(T), SC_2(T) - SC_1(T), 1 - SC_2(T)) \quad (6.4)$$

During the calibration procedure, survival curves are estimated from a set of inspected segments. They aim at reproducing the deterioration behaviour observed in the inspection dataset, i.e. the proportion of segments in each condition at a given age. After estimating the average uncertainty matrix, the survival curves can be corrected using equation 6.3 as following:

$$R(T) = M^{-1} P(T) \quad (6.5)$$

A confidence interval can be obtained from a Monte-Carlo simulation using the 1,000 uncertainty matrix M obtained in the optimisation procedure. The confidence interval can be

useful to propagate uncertainties in the prediction of the deterioration model. Monte-Carlo simulations allow simulating the impact of uncertainties on the strategic outcomes of the model.

First, the corrected survival curves are used to simulate the future evolution of the condition distribution of 1,000 random segments for a fictive do-nothing strategy without any rehabilitation actions. Over a given simulation period $[T_{start}; T_{end}]$, the model calculates, for each year T , the condition distribution, i.e. the proportion of pipes in each condition.

$$C(T) = (C_1(T), C_2(T), C_3(T)) \quad (6.6)$$

Then, the survival curves are used to simulate the necessary replacement rate to maintain the overall condition of the whole network (constant proportion of segments in poor condition) over time. The replacement rate is calculated as the percentage of pipes that shifted from the intermediate to the poor condition between the start year and a given year of the simulation, divided by the number of elapsed years. It provides the average yearly replacement rate to avoid the deterioration of the network.

$$replacement\ rate(T) = \frac{C_3(T) - C_3(T_{start})}{T - T_{start}} \quad (6.7)$$

6.2.3 Application on the sewer network of the city of Berlin

Description of the data

The study has been performed using the GIS and CCTV database of the city of Berlin in Germany (Refer to chapter 2.3.2 for a presentation of the case study and available data). After data clean-up, the database contains 124,450 inspections with a length of 5,222 km over 107,788 pipes.

Calibration of a deterioration model

Figure 6.1 shows the condition distribution of the dominating clay and concrete segments.

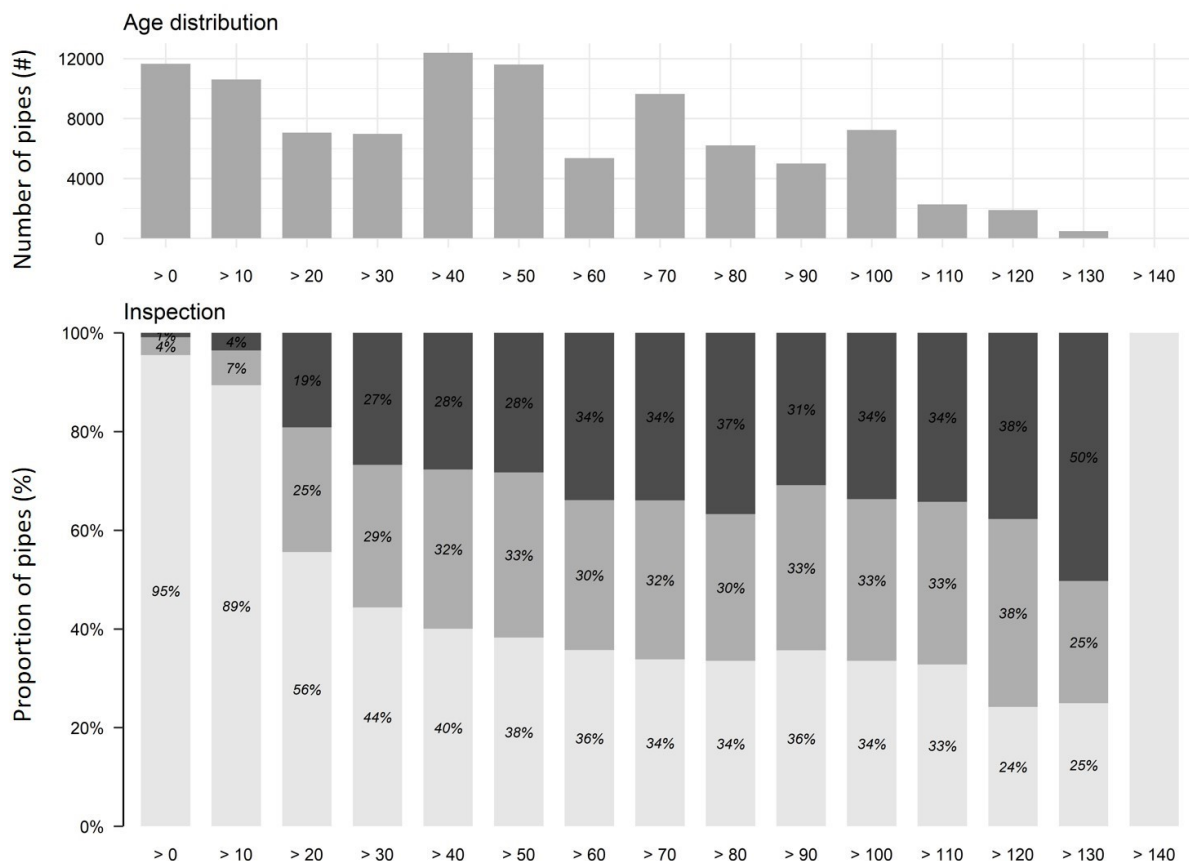


Figure 6.1. Condition distribution VS segment age for the group of concrete and clay segments in Berlin. The grey scale represents the three condition classes, dark grey being the proportion of segments in the worst condition (condition 3). For convenience, the labels of the age groups have been simplified: “>x” represents the interval [x ; x+10]

The condition is correlated with the segment’s age; old segments are generally in a worse condition than new segments. However, the condition of very old segments (> 90 years old) seems to improve slightly. This phenomenon could be related to the survival selection bias (Egger *et al.* 2013; Ouellet & Duchesne 2018). Inspection data tend to be biased as the observations are carried out in a restricted time window (for this study from 2001 to 2017). Most old or deteriorated segments have already been replaced, thus are not fully represented in the sample of inspection data. In order to remove (or at least reduce) the influence of the survival bias in model calibration, old segments (age > 70 years old for concrete pipes; age > 90 years old for clay pipes) have been removed from the dataset.

The model GompitZ (Le Gat, 2008) has been used to calibrate survival curves for one unique

cohort composed of concrete and clay segments. Survival curves have the mathematical form of a Gompertz distribution. They are calibrated with a regression procedure using the maximum likelihood estimation. Figure 6.2 shows the calibrated survival curves for the three condition classes. The curves reproduce the deterioration behaviour of concrete and clay segments observed in Figure 6.1. The influence of the survival bias has been estimated by calibrating the model with all data, i.e. without removing old pipes. As expected, the prediction is more pessimistic by removing old pipes from the database (e.g. +10% pipes in poor condition at 100 years; results not shown here).

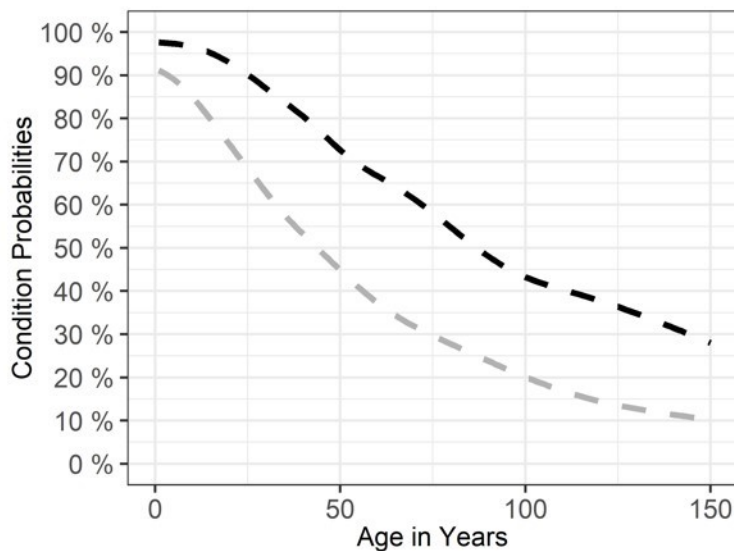


Figure 6.2. Calibrated survival curves for concrete and clay segments. The light grey curve shows transition from good condition (1) to intermediate (2) condition. The dark grey curve shows transition from intermediate condition (2) to poor (3) condition.

6.3 Results and discussion

6.3.1 Uncertainties in sewer condition assessment

The methodology described in chapter 5 (Caradot *et al.*, 2018a) has been applied using the inspection data of Berlin. The main uncertainty matrix M obtained from the optimisation procedure is as follows.

$$\text{mean}(M) = \begin{pmatrix} 85.2 & 15.9 & 6.1 \\ 10.6 & 67.3 & 12.5 \\ 4.2 & 16.8 & 81.4 \end{pmatrix} \quad (6.8)$$

6.3.2 Propagation of the uncertainties in a deterioration model

The uncertainty matrix is used to propagate uncertainties in the survival model and to assess their influence on the prediction of asset management strategies. Equation 6.5 is applied to correct the survival curves presented in Figure 6.2. For each age T , the corrected survival curves are obtained by multiplying the proportions of each condition indicated by the survival curves by the uncertainty matrix.

$$R(T) = M^{-1} P(T) \quad (6.9)$$

1,000 corrected survival curves are generated using the 1,000 uncertainty matrices obtained during the optimisation procedure. From the 1,000 corrected survival curves, the mean corrected survival curve and the boundaries of the 90% confidence interval are calculated.

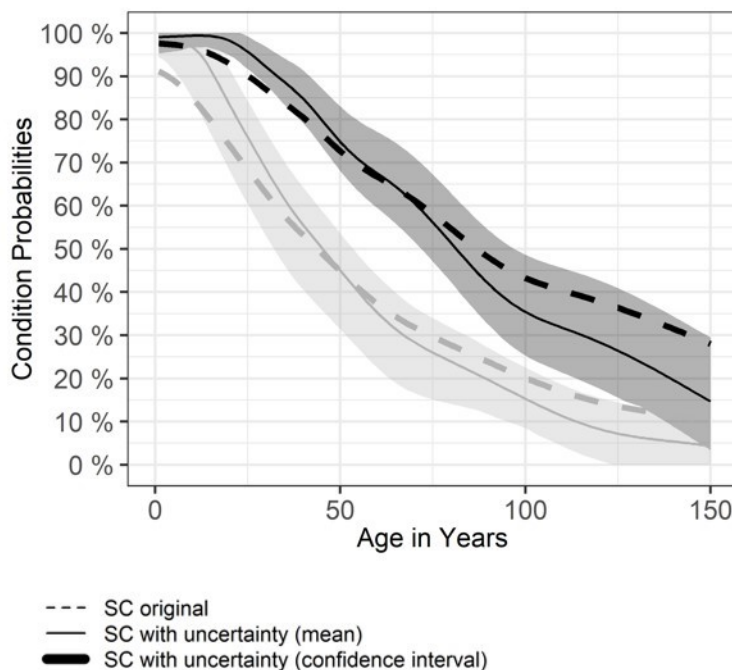


Figure 6.3. Original and corrected survival curves. Light grey curves show transition from good condition (1) to intermediate (2) condition. Dark grey curves show transition from intermediate condition (2) to poor (3) condition.

The confidence interval quantifies the uncertainty of the prediction, e.g. at 100 years, the proportions of segments in poor condition is $65\% \pm 12\%$. The mean corrected survival curves do not overlap with the original survival curves. It indicates that the propagation of uncertainties corrects a systematic error: the average prediction considering uncertainties (solid lines in Figure 6.3) is not equal to the prediction without considering uncertainties (dashed lines in Figure 6.3). This bias is related to the different probabilities of False Positive and False Negative. As discussed in the previous section, the most probable errors are False Positive for segments in good condition and False Negative for segments in poor condition. Most of the young segments (< 30 years) are in good condition and are thus more prone to False Positive than False Negative. There is a higher probability to be too pessimistic, so the corrected survival curve is more optimistic than the original survival curve. On the contrary, most of the old segments (> 75 years) are in poor condition and are thus more prone to False Negative than False Positive. There is a higher probability to be too optimistic, so the corrected survival curve is more pessimistic than the original survival curve.

It is also interesting to note that the slope of the original and corrected survival curve is not identical. The slope of the corrected survival curve is higher indicating that the systematic error increases with age. The correction leads to consider a faster network deterioration.

6.3.3 Propagation of uncertainties in asset management strategies

In the previous section, uncertainties from sewer condition assessment have been propagated in the calibrated survival curves of a deterioration model. Since models are used to simulate asset management strategies, it is of interest to assess the sensitivity of the predicted strategies to these uncertainties.

Impact of uncertainties on the predicted condition of the network

First, the survival curves have been used to simulate the future evolution of the condition

distribution of 1,000 random segments for a fictive do-nothing strategy, i.e. without any rehabilitation actions. The experiment has been run on 1,000 randomly selected segments to reduce the computation time of the Monte Carlo simulation. Figure 6.4a shows the evolution of condition distribution using the original survival curves and the corrected survival curves (with mean and 90% confidence interval). Figure 6.4b shows the deviation (in %) between the proportions of segments in poor condition (class 3) obtained with the original and corrected survival curves. The long-time horizon (> 100 years) is not realistic for planning strategies but interesting to understand the sensitivity of the models.

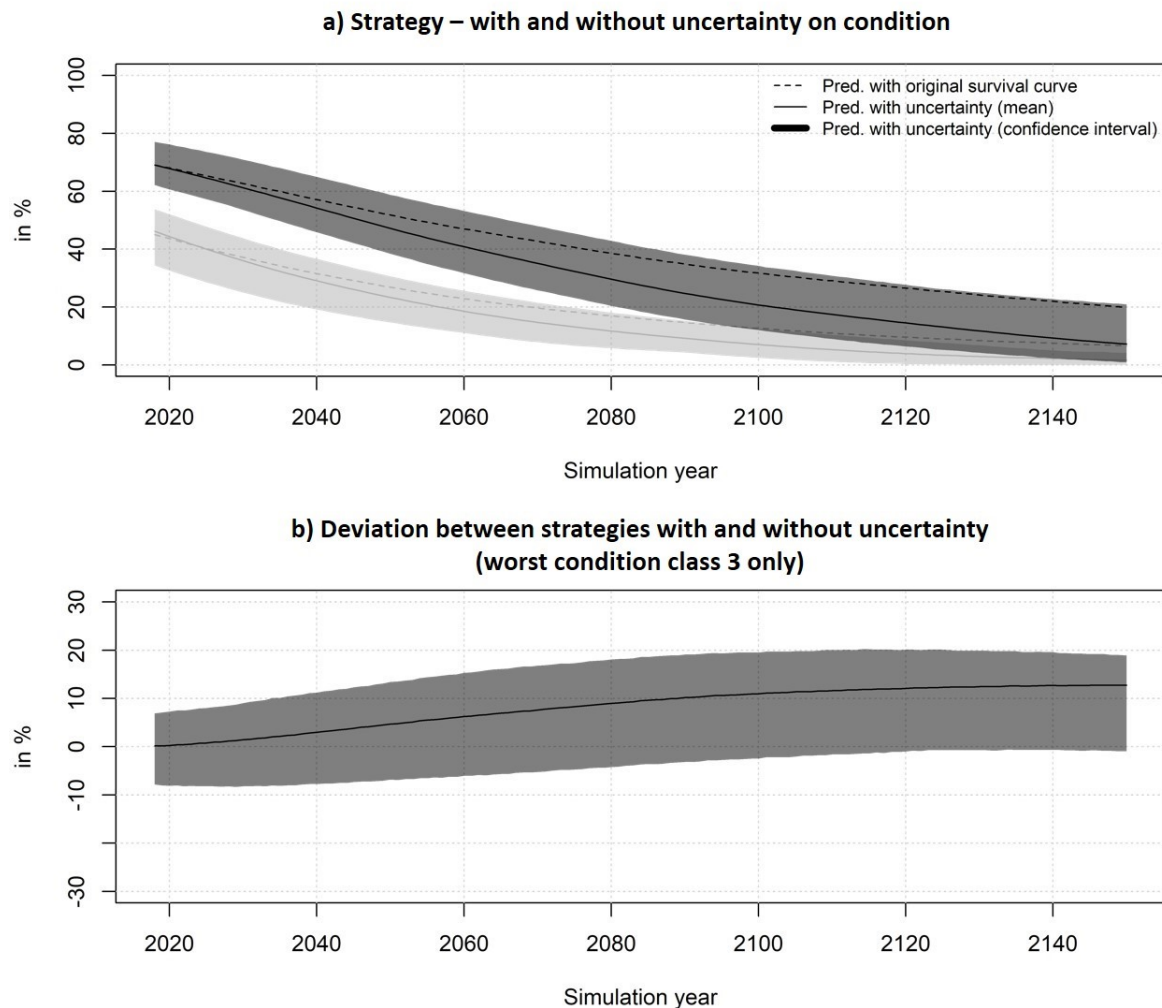


Figure 6.4. Impact of uncertainties on the prediction of a deterioration model. Simulation of a do-nothing strategy for 1000 random segments in Berlin. Light grey curves show the percentage of pipes in good condition (1). Dark grey curves show the percentage of pipes in good (1) and intermediate (2) conditions.

In 2060 (i.e. after 40 years of simulation), according to the original survival curves, the proportion of segments in poor condition would be 53% (Figure 6.4a). By considering uncertainties, the proportion rises to $59\% \pm 11\%$. (Figure 6.4a).

The following outcomes can be highlighted:

- The impact of the condition assessment uncertainties on a sewer asset management strategy is not negligible. Even at the start simulation year, the proportion of segments in poor condition can be estimated correctly within a range of $\pm 8\%$ (Figure 6.4b).
- The prediction uncertainty increases with the number of years simulated. In particular, the systematic error (bias between the average prediction considering uncertainties and the prediction without considering uncertainties) increases with the simulation year. In 2060 (i.e. after 40 years of simulation), the prediction uncertainty is in the range $[-12\%; +9\%]$ with a systematic uncertainty of $+6\%$. The systematic uncertainty would reach $+10\%$ only after 2090 (Figure 6.4b). When considering this time horizon, other uncertainties in operation conditions (e.g. climate change, demographic development) are likely to have much more influence on the outcomes than the condition assessment uncertainties, making predictions highly unrealistic.

Impact of uncertainties on the replacement rate

In a second experiment, the survival curves have been used to simulate the necessary replacement rate to maintain the network with a constant proportion of segments in poor condition over time, for the same subset of 1,000 random segments. Figure 6.5 shows the replacement rates obtained with the original and corrected survival curves.

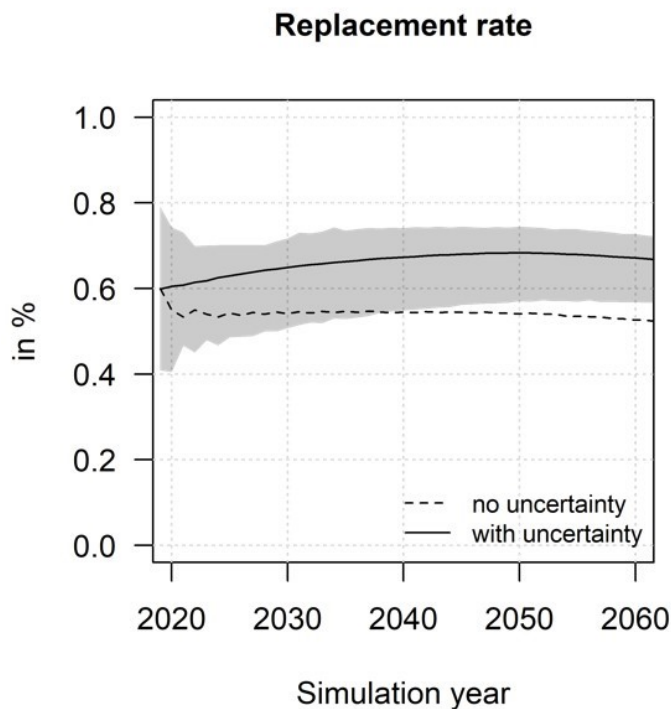


Figure 6.5. Impact of uncertainties on the required replacement rate to maintain the network with a constant proportion of segments in poor condition over time, for 1000 random segments.

Using the original survival curve, the necessary replacement rate ranges between 0.5% and 0.6%. Note that this is not the required replacement rate to avoid the global deterioration of the network (i.e. constant proportion of pipes in each condition) but only to maintain stable the proportion of segments in condition class 3 (poor condition). Considering uncertainties, the replacement rate is higher. Until 2050, the average replacement rate obtained by the original survival curve is 0.54%, and the corrected replacement rate ranges between [0.57%; 0.75%]. This result can be understood from the slopes of the survival curves. The slope of the corrected survival curves is higher than the slope of the original survival curves indicating a faster network deterioration. The consideration of uncertainties leads to a higher and more realistic replacement rate.

6.4 Conclusion

This paper analyses the influence of sewer condition uncertainties on the prediction of deterioration models. Uncertainties of condition assessment are a well-known issue for sewer operators who generally acknowledge the high subjectivity of the condition assessment procedure. First, the methodology proposed in chapter 5 has been applied to quantify uncertainties in sewer condition assessment from the analysis of a set of repeated inspections. The repeated inspections have been used to determine the uncertainty matrix, which quantifies the probability to inspect a segment correctly and to overestimate or underestimate its condition. Then, a method has been proposed to propagate uncertainties in the survival curves and predictions of a deterioration model. The deterioration model has been used to simulate simple long-term strategies and evaluate the impact of uncertainties over model prediction. The method has been demonstrated using the unique inspection dataset of the city of Berlin, Germany, where 13,753 segments have been inspected at least twice. The following outcomes can be highlighted.

- The probability to assess the correct condition of a segment in good condition is 85%; the probability to assess the correct condition of a segment in poor condition is 81%. The probability to assess the correct condition of a segment in the intermediate condition is lower and close to 67%.
- The propagation of uncertainties in the survival curves produces a confidence interval around the original survival curves. At 100 years, the uncertainty for the proportion of segments in poor condition is $\pm 12\%$.
- The analysis of this confidence interval highlights the presence of a systematic error: the mean corrected survival curves do not overlap with the original survival curves. This bias is related to the different probabilities of False Positive and False Negative.

- The impact of the uncertainties on the prediction of a deterioration model is not negligible. The systematic uncertainty increases with the simulation year. The analysis also shows that the required replacement rate to maintain a constant proportion of segments in poor condition is underestimated if the uncertainties are not included in the analysis.

Even influenced by uncertainties, deterioration models remain a powerful tool to assess the impact of future rehabilitation scenarios at network scale. However, the high uncertainties in deterioration modelling must be communicated to avoid the wrong interpretations of modelling outcomes and wrong management decisions. We recommend to focus on the mitigation of uncertainty sources and the visualisation of the remaining uncertainties in asset management tools to facilitate decision making in a highly uncertain context.

Chapter 7 Conclusion and perspectives

7.1 Conclusion

Insufficient investment represents a major challenge for the long-term management of urban drainage systems. In many cities worldwide, the underground infrastructure is nearing the end of its technical lifetime and will reach soon the age of massive renewal. Utilities are challenged to develop efficient rehabilitation strategies in order to keep the same level of service. Traditionally it has been economically feasible to apply reactive management strategies, repairing mainly when failures occur; however, this strategy will become less viable as the systems age and the funding gap increases (Rokstad & Ugarelli, 2015). In this context, a promising leverage for utilities is the improvement of technical asset management and, in particular, the use of predictive solutions to improve the efficiency of inspection and rehabilitation strategies.

CCTV inspection is used since the 1980's as industry standard for sewer system inspection and structural performance evaluation. Due to budget restrictions, CCTV inspection rates are generally low and municipalities tend to inspect only a small part of their network (Ahmadi *et al.*, 2014c; Harvey & Mc Bean, 2014; ONEMA, 2012). Furthermore, CCTV data as such (without further analysis) are insufficient to determine long term asset management strategies since they provide only a snapshot of the sewer condition at the date of inspection and no information regarding the remaining lifetime of the sewers. Since the definition of rehabilitation strategies is limited by the lack of information about sewer condition and remaining life, deterioration models have been developed to forecast the evolution of the system according to its current and past condition. Deterioration models can be used (i) to simulate the condition class of non-inspected pipes and (ii) to forecast the evolution of the network condition. If the possibilities of deterioration modelling to support inspection and rehabilitation programs are

now better understood, research efforts are still needed to investigate the performance of deterioration models and the benefits of modelling for planning mid to long-term rehabilitation budgets (Alegre & Matos, 2009; WERF, 2012). Deterioration models are still not commonly used by sewer operators and municipalities and the major gap between utilities' day-to-day management practices and the available modelling approaches remains significant.

One of the main factors hampering the uptake of deterioration modelling by utilities is the lack of real scale evidence of the tangible benefits provided (Scheidegger *et al.*, 2011; WERF, 2007). Most utilities are concerned by the minimum amount of CCTV data required and the relevance of using such models on their networks with limited data availability. The assessment of the influence of CCTV data availability on the reliability of deterioration modelling is still a key step to build the trust of utilities regarding modelling outcomes. Finally, most utilities acknowledge the uncertainties in the procedure of sewer condition assessment, mainly due to the subjectivity of the coding operator. There is still a need to quantify the uncertainty of the sewer condition assessment procedure and its influence on the outcomes of deterioration modelling.

This thesis has focused on the performance of sewer deterioration modelling in relation to variable CCTV data quality and quantity. Each research question is presented below, with the main conclusions.

What is the performance of sewer deterioration modelling in a case of high CCTV data availability?

A first methodological outcome of this thesis is the development of a set of straightforward and intuitive accessible metrics defined in dialogue with the Berlin local utility in order to evaluate sewer performance from an end-user perspective. The selected metrics aim at convincing the municipality about the relevance of using a given deterioration model to support asset

management strategies. The metrics assess model performance at two main levels: the network and the pipe levels. At the network level, the metrics indicate to which extent the model can predict the condition distribution of the entire network, i.e. the number of pipes in each condition. At pipe level, the metrics verify to which extent the model can predict correctly the inspected condition class of each single pipe. Both information are needed for different purposes: network level metrics show the model's relevance for supporting strategic rehabilitation planning; pipe level metrics illustrate the potential for supporting inspection strategies by identifying pipes in critical condition. These metrics could be used as benchmark in further research for evaluating and communicating the various facets of modelling performance.

Two deterioration models have been investigated using these new metrics. A statistical model (GompitZ) and a machine learning model (Random Forest) have been applied on the extensive CCTV and GIS dataset of the city of Berlin, Germany. The results show that **both machine learning and statistical models give satisfactory outcomes at network level**. Deviations between the predicted and inspected condition distributions, for the entire network and for different age groups, are below 5% using Random Forest and even lower (below 1%) using GompitZ. This result underlines the strong potential of both statistical and machine learning models to simulate the condition distribution of the network.

At pipe level, the machine learning model outperforms the statistical model. In the cases discussed in this thesis, Random Forest performance is relatively good for the simulation of pipes in poor condition: 66.7% of the pipes inspected in poor condition have been predicted correctly and only 9.5% of the pipes inspected in poor condition have been falsely predicted in good condition. The True rate of Random Forest for pipes in poor condition (67%) is close to the True Positive rate of a CCTV inspection (~80%, see chapter 5). The Random Forest model

shows a strong potential for supporting sewer operators in the identification of pipes in critical condition for inspection programs. One main weakness of the Random Forest model lies in its high False Positive rate: 28.3% of pipes predicted in poor condition are actually in good condition. This aspect of the performance might be improved in further studies by considering additional variables and testing other modelling approaches. Another major weakness of the Random Forest model is that it can lead to doubtful prediction such as condition improvement along with pipe age. The model learns and reproduces the patterns observed in the inspection dataset. This leads to the conclusion that the tested **machine learning approach shall only be used for ad-hoc classification of the sewer pipes but not for long-term prediction**. This problem does not occur with GompitZ since the deterioration follows a Gompertz distribution that prevents any condition improvement. This aspect of machine learning should be carefully investigated before deploying such models in practice.

What is the influence of CCTV data availability on the performance of sewer deterioration modelling?

The statistical model GompitZ has been applied using the extensive GIS and CCTV database of the city of Braunschweig in Germany. Its performance has been evaluated by calculating the deviation between the predicted and inspected condition distributions, with different sizes of random subsets. The assessment of model performance in Braunschweig confirms the outcomes obtained in Berlin. GompitZ can simulate the condition distribution of the network with low deviations <5% between inspected and predicted condition distributions. Even with 3% of data used for model calibration (about 1,000 pipes), GompitZ is able to simulate the number of pipes in poor condition with a deviation smaller than 10%. Based on this outcome, **a subset of at least 1,000 pipes (around 50 km) seems to be a mandatory minimum to simulate the network condition with relatively good accuracy**. These results obtained in Braunschweig confirm the results obtained by Tran (2016) in Australia and Ahmadi *et al.* (2016) who

recommend a minimum set of 700 and 1,000 pipes, respectively. However, this conclusion is only valid if the inspections used for model calibration were randomly selected over the entire network. The number of pipes required may depend on the sampling method used, particularly if pipes in poor condition are under-represented in the database (Ahmadi *et al.*, 2016).

Our results also indicate that **in the case of low data availability, GompitZ performs much better than a simple random model**. With 3% of data used for model calibration (about 1,000 pipes), the deviation obtained with GompitZ ranges between [-9.1%; 4.5%] whereas the deviation obtained with the random selection model ranges between [-29.7%; 30.3%].

What is the uncertainty of the sewer condition assessment procedure?

A methodology based on double inspections of the same pipes has been proposed to determine the uncertainties of the structural condition assessment. The methodology is based on a non-linear optimization procedure coupled with a Monte-Carlo simulation and has been used to determine the probabilities of false positive and false negative by inspecting a pipe in a given condition. This study has been performed using the CCTV database of the cities of Braunschweig and Berlin in Germany.

Results indicate that **the probability to inspect correctly a pipe in poor condition is close to 80-85%** and thus the probability to overestimate the condition of the pipe is close to 15-20% (False Negative, i.e. inspected condition is better than the real condition). The probability to inspect correctly a pipe in good condition is slightly higher than the probability to inspect correctly a pipe in poor condition. It indicates that there is less uncertainty by assessing the condition of sewers with few defects only or without severe defects. This is an expected outcome: the inspectors that perform the defect coding are more prone to error when many defects are present than when sewers are in perfect condition.

The analysis also showed that the **uncertainties related to sewer condition assessment are not only due to the inspection procedure**. The analysis of deviations in repeated inspections by the water utility of Berlin highlights further sources of uncertainties such as undocumented rehabilitation (i.e. the rehabilitation of segment has not been documented in the database) or inspections done on a specific purpose, e.g. to identify house connections. The reduction of uncertainties could start by the improvement of data management procedures in order to be able to filter improper inspections before calibrating sewer deterioration models.

Finally, the optimization approach developed is a strong methodological outcome of this thesis. The method could be further applied in other cities to assess sewer condition uncertainties locally, confirm the outcomes obtained in Braunschweig and Berlin and understand better the contribution of the various sources of uncertainties to the final uncertainty presented above.

What is the influence of CCTV data uncertainty on the performance of sewer deterioration modelling?

A method has been proposed to propagate uncertainties in the survival curves and predictions of a deterioration model. The deterioration model has been used to simulate simple long-term strategies and evaluate the impact of uncertainties in condition assessment. The method has been demonstrated using the inspection dataset of the city of Berlin, Germany. The propagation of uncertainties in the survival curves produces a confidence interval around the original survival curves. **At 100 years, the uncertainty for the proportion of segments in poor condition is $\pm 12\%$.**

The analysis of this confidence interval highlights **the presence of a systematic error**: the mean corrected survival curves do not overlap the original survival curves. **This bias is related to the different probabilities of False Positive and False Negative**. Most of the young segments (< 30 years) are in good condition and are thus more prone to False Positive than

False Negative. Most of the old segments (> 75 years) are in poor condition and are thus more prone to False Negative than False Positive. There is a higher probability to be too optimistic, so the corrected survival curve is more pessimistic than the original survival curve.

The impact of the uncertainties on the prediction of a deterioration model is not negligible.

The systematic uncertainty increases with the number of years simulated. The analysis also shows that **the required replacement rate to maintain a constant proportion of segments in poor condition is underestimated if the uncertainties are not included in the analysis.**

Even influenced by uncertainties, deterioration models remain a powerful tool to assess the impact of future rehabilitation scenarios at network scale. However, the high uncertainties in deterioration modelling must be communicated to avoid the wrong interpretations of modelling outcomes and wrong management decisions.

7.2 Perspectives

The analysis of the current state of the art and the outcomes of this thesis highlight the following research needs and perspectives:

- Further investigations are needed to quantify carefully each **uncertainty sources**, assess their cumulated propagation in deterioration models and find practical solutions to mitigate their impact on asset management decisions.
- In particular, studies should focus on the **communication of modelling outcomes and associated uncertainties** with the decision makers. Most modelling approaches developed in the research community are expert tools; the uptake of modelling approaches by utilities will require the development of user-friendly interfaces and communicative dashboard to facilitate discussions with the decision makers. Such

“simplification” must be done without removing information related to the quality of the prediction and the uncertainties associated.

- Regarding uncertainties, the **survival bias** seems to be a critical issue for the future development of deterioration models (Ouellet & Duchesne, 2018). Current models are expected to be optimistic because the observed pipes used for model calibration are only those that “survived” until the date of inspection, i.e. pipes that were not replaced before they reach their current degradation state (Le Gat, 2008). Since the models are calibrated using data concerning pipes that were in place at the date of inspection, they will inevitably underestimate the probability to be in a poor state, and consequently, overestimate the duration of useful life of pipes. Several studies already highlighted the existence of this bias using synthetic datasets (Ouellet & Duchesne, 2018) or by combining deterioration models with theoretical rehabilitation models (Egger *et al.*, 2013). Further work will be needed to quantify this bias using real datasets by considering the data from already replaced pipes in model calibration. Further investigations are also needed to propose practical methodologies to correct locally the survival bias to avoid the presence of systematic errors in long-term predictions. More generally, the communication of the survival bias issue might lead municipality in upgrading data management practices and improving their predictive abilities. A starting point would be the documentation of each rehabilitation action in the network and the systematic tracking of “old” inspections from rehabilitated pipes.
- Given the considerable annual investments for sewer rehabilitation, additional expenses on sewer inspection and data management for the reduction of model uncertainties might be beneficial to optimize the strategic planning of investments on the network. Further studies might investigate in detail the marginal benefits of reducing modelling

uncertainties in order to determine the appropriate level of expenses for sewer inspection and data management.

- Previous researches focused mainly on the development and test of deterioration modelling approaches. In the future, research will have to focus on the development of predictive decision-making tools to **support asset management strategies, i.e. based on risks or impacts prediction**. These methods use deterioration model to simulate sewer condition or probability of failure combined with a vulnerability or risk assessment module. For example, Baah *et al.* (2015) developed a method to assess the risk of pipe failure based on the combination of deterioration model prediction and consequence of failure using a weighted-sum scoring. Del Giudice *et al.* (2016) developed a statistical tool that may assist in recognizing critical sewers when limited information is available. Kessili & Bemamar (2016) proposed a methodology for the prioritization of sewer rehabilitation projects using multiple-criteria decision making. Further studies are needed to carefully assess and combine pipe condition prediction and vulnerability in a risk assessment framework. One major challenge in this area is to adapt the prediction to enable the risk approach (Le Gauffre *et al.*, 2007): this will require to predict distinctly several condition such as the physical condition (collapse risk), the operational condition (decrease of hydraulic capacity) or the environmental condition (infiltration / exfiltration risk).
- Researches should also aim at providing methods to **support the selection of cost-effective rehabilitation actions using Life Cycle Cost Analysis (LCC)**. These methods aim at defining cost-effective rehabilitation programs. They use discounted cost model to assess the cost of failure vs the costs of different rehabilitation alternatives. For example, Kleiner (2001b) developed a decision tool, which aims at finding the optimal rehabilitation schedule to minimize the total expected costs

associated with the asset throughout its life. Marzouk & Osama (2017) developed an optimization model using genetic algorithms to determine integrated rehabilitation plans. The optimization model considers rehabilitation alternatives, setting priorities for integrated rehabilitation, implementing optimization of renewal cost and defining the best replacement interval.

- Several tools have already been developed to **support the long-term investment planning of sewer networks**. They generally aim at assessing the influence of various rehabilitation scenarios over the condition and financial value of the network in a horizon of 10-40 years. Rehabilitation scenarios usually include the definition of annual investment and the allocation of different rehabilitation activities, such as repair, renovation and replacement. In Germany, several tools using a cohort survival or Markov model are proposed mainly by consulting offices (e.g. AQUA-WertMin, DynaStrat, KANEW-Z and STATUS). In general, these tools aim to predict rehabilitation needs and costs for different investment scenarios. They can be used to describe the relationship between budget allowance and the resulting development of the sewer network condition. In Portugal, the AWARE-P decision-support environment (Coelho *and al.*, 2013) proposes to simulate and prioritize competing rehabilitation scenarios over a set of performance indicators. In particular, these tools need to be further developed and enhanced in order to consider the deterioration of renovated (Liner) and repaired pipes. This issue needs to be addressed in order to build reliable asset management tools which consider the variety of potential rehabilitation actions. Next researches might focus on the development of such models and the description of the life duration of rehabilitation techniques.
- Finally, sewer networks are one of the many urban infrastructures. Drinking water networks, district heating, electricity in the underground or road and urban green

infrastructures on the surface might interact with the sewer network. Despite the fact that these infrastructures also require sound asset management strategies, their rehabilitation planning might benefit from synergies between the infrastructure in order mainly to mutualise rehabilitation costs and reduce nuisances to citizens. However the advantages and drawbacks of such synergies are not well known and utility practices vary strongly depending on the local context and political background. Several authors started to investigate this promising and complex topic of **integrated infrastructure planning** using deterioration models. Nafi & Kleiner (2010) proposed a methodology for the scheduling of water mains replacement taking into account economies of scale considerations as well as harmonization with other known infrastructure works. Carey & Lueke (2013) developed an optimized holistic approach for rehabilitation planning which allows the consideration of contiguity savings through the synchronization of rehabilitation projects with three infrastructure networks (roads, sewers and water distribution). Marzouk & Osama (2015) applied fuzzy logic to identify the optimum replacement time of different infrastructure networks (roads, sewers, water, gas and electric cables). Tscheikner-Gratl *et al.* (2015) and Tscheikner-Gratl *et al.* (2016) proposed a methodology for the integrated prioritisation of rehabilitation actions. Future researches might focus on quantifying the advantages and drawbacks of using deterioration models for integrated planning. In particular, research might investigate the influence of synergies on rehabilitation costs and service quality and identify the right level of cooperation between urban infrastructures in order to maximise the utility or city objectives.

Chapter 8 References

- Ahmadi, M., Cherqui F., De Massiac, J.-C. and Le Gauffre, P. 2014a. Influence of available data on sewer inspection program efficiency. *Urban Water Journal*, 11(8), 641-656
- Ahmadi, M., Cherqui, F., De Massiac, J.C, Werey, C., Lagoutte, S. & Le Gauffre, P. 2014b. Condition grading for dysfunction indicators in sewer asset management. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, 10(3), 346-358
- Ahmadi, M., Cherqui, F., De Massiac, J.C., Le Gauffre, P. 2014c. From sewer inspection programmes to rehabilitation needs: Research and results related to data quality and availability with the support of numerical experiment. *European Journal of Environmental and Civil Engineering*, 18(10), 1145-1156
- Ahmadi, M., Cherqui, F., De Massiac, J.C. & Le Gauffre, P. 2015. Benefits of using basic, imprecise or uncertain data for elaborating sewer inspection programmes. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, 11(3), 376-388
- Ahmadi, M., Cherqui, F., Aubin, J.-B. and Le Gauffre, P. 2016. Sewer asset management: impact of sample size and its characteristics on the calibration outcomes of a decision-making multivariate model. *Urban Water Journal*, 13(1), 41-56
- Alegre, H., Matos, R. 2009. Paving the way for a sustainable asset management of urban water infrastructures: outcomes of the 5th World Water Forum. *Water Asset Management International*, 5(2), 19-20
- American Water Works Association. 2012. *Buried No Longer: Confronting America's Water Infrastructure Challenge*. AWWA's infrastructure financing report, Denver, Colorado, United States
- Ana, E.V., Bauwens, W., Pessemier, M., Thoeye, C., Smolders, S., Boonen, I., and De Gueldre, G. 2009. An investigation of the factors influencing sewer structural deterioration. *Urban Water Journal*, 6(4), 303-312
- Ana, E. V., Bauwens, W. 2010. Modeling the structural deterioration of urban drainage pipes: The state-of-the-art in statistical methods. *Urban Water Journal*, 7(1), 47-59
- Ariaratnam, S., El-Assaly, A., and Yang, Y., 2001. Assessment of Infrastructure Inspection Needs Using Logistic Models. *Journal of Infrastructure Systems*, 7(4), 160-165

- ASCE. 2011. *Failure to act: the economic impact of current investment trends in water and wastewater treatment infrastructure*. American Society of Civil Engineers, Reston, Virginia, United States
- Ashley, R. M., Fraser, A., Burrows, R. and Blanksby, J. 2000. The management of sediment in combined sewers. *Urban Water*, 2(4), 263-275
- ATV M 143-2. 1999. *Inspektion, Instandsetzung, Sanierung und Erneuerung von Abwasserkanälen und -leitungen –Teil 2: Optische Inspektion*. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V., Hennef, Germany
- Baah, K., Dubey, B., Harvey, R., and McBean, E. 2015. A risk-based approach to sanitary sewer pipe asset management. *Science of the Total Environment*, 505, 1011-1017
- Baik, H-S., Jeong, HSD., Abraham, DM. 2006. Estimating Transition Probabilities in Markov Chain-Based Deterioration Models for Management of Wastewater Systems. *Journal of Water Resources Planning and Management*, 132(1), 15-24
- Baur, R., Herz, R. 2002. Selective inspection planning with ageing forecast for sewer types. *Water Science and Technology*, 46(6–7), 389-396
- Berger, C., Falk, C., Hetzel, F., Pinnekamp, J., Roder, S., Ruppelt, J. 2015. *Zustand der Kanalisation in Deutschland*. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V., Hennef, Germany
- Bergstra, J. and Bengio, Y. 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13, 281-305
- Bertrand-Krajewski, J.-L., Ellis, B. 2010. *Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems (APUSS)*. APUSS project report, IWA publishing
- Black and Veatch. 2013. *Strategic direction in the U.S. water industry*. Black and Veatch report, <https://www.bv.com/docs/reports-studies/sdr-water-industry.pdf>, retrieved on February 26 2019
- Breiman, L. 2001. Random forest. *Machine Learning*, 45, 5-32
- Breiman, L. 2002. *Manual on Setting Up, Using, And Understanding Random Forests V3.1*. https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf, retrieved on May 3, 2018
- Breiman, L., Friedman, J. H., Stone, C. J. and Ohlsen, R. A. 1984. *Classification and Regression Trees*. Taylor & Francis Ltd, United Kingdom.

- Brown, C. D., & Davis, H. T. 2006. Receiver operating characteristics curves and related decision measures: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, 80(1), 24-38
- Caradot, N., Sonnenberg, H., Kropp, I., Ringe, A., Denhez, S., Hartmann, A., Rouault, P. 2017. The relevance of sewer deterioration modelling to support asset management strategies. *Urban Water Journal*, 14(10), 1007-1015
- Caradot, N., Rouault, P., Clemens, F., Cherqui, F. 2018a. Evaluation of uncertainties in sewer condition assessment. *Structure and Infrastructure Engineering*, 14(2), 264-273
- Caradot, N., Riechel, M., Fesneau, M., Hernández, N., Torres, A., Sonnenberg, H., Eckert, E., Lengemann, N., Waschnewski, J., Rouault, P. 2018b. Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany. *Journal of Hydroinformatics*, 20(5), 1131-1147
- Caradot, N., Riechel, M., Rouault, P., Caradot, A., Lengemann, N., Eckert, E., Ringe, A., Clemens, F., Cherqui, F. 2019. The influence of condition assessment uncertainties on sewer deterioration modelling. *Structure and Infrastructure Engineering*, accepted 05/2019
- Carey, B. D. and Lueke, J. S. 2013. Optimized holistic municipal right-of-way capital improvement planning. *Canadian Journal of Civil Engineering*, 40(12), 1244-1251
- Carvalho, G., Amado, C., Brito, R.S., Coelho, S.T., Leitão, J.P. 2018. Analysing the importance of variables for sewer failure prediction. *Urban Water Journal*, 15(4), 338-345
- Chen, C. & Breiman, L. 2004 *Using Random Forest to Learn Imbalanced Data*. University of California, Berkeley, United States
- Cherqui, F., Gutierrez-Silva, C., Ahmadi, M., Aubin, J. B. and Le Gauffre, P. 2017. Trusting CCTV reports? An exploratory study of operators' ability to detect and recognize defects. In: *Proceedings of The Leading Edge Sustainable Asset Management of Water and Wastewater Infrastructure Conference – LESAM 2017*, 20-22 June, Trondheim, Norway
- Chughtai, F., Zayed, T. 2008. Infrastructure condition prediction models for sustainable sewer pipelines. *Journal of Performance of Constructed Facilities*, 22(5), 333-341
- Coelho, S.T., Vitorino, D., Alegre, H. 2013. AWARE-P: a system-based software for urban water IAM planning. In: *Proceedings of 2013 IWA LESAM*, Sydney, Australia

- Das, S., Bayat, A., Gay, L. F., Matthews, J. 2018. Productivity Analysis of Lateral CIPP Rehabilitation Process Using Symphony Simulation Modeling. *Journal of Pipeline Systems Engineering and Practice*, 9(1)
- Davies, J. P., Clarke, B. A., Whiter, J. T., Cunningham, R. J., 2001. Factors influencing the structural deterioration and collapse of rigid sewer pipes. *Urban Water*, 3(1-2), 73-89
- Del Giudice, G., Padulano, R., & Siciliano, D. 2016. Multivariate probability distribution for sewer system vulnerability assessment under data-limited conditions. *Water Science and Technology*, 73(4), 751-760
- Dey, A. K. and Kamioka, S. 2007. An integrated modeling approach to predict flooding on urban basin. *Water Science & Technology*, 55(4), 19-29
- Dirksen, J., Clemens, F.H.L.R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., Snaterse, C.T.M. 2013. The consistency of visual sewer inspection data. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, 9(3), 214-228
- Duchesne, S., Beardsell, G., Villeneuve, J. P., Toumbou, B., & Bouchard, K. 2013. A survival analysis model for sewer pipe structural deterioration. *Computer-Aided Civil and Infrastructure Engineering*, 28(2), 146-160
- Duchesne S., Bouchard K., Toumbou B. & Villeneuve J.-P. 2014. Assessing the impact of renewal scenarios on the global structural state of sewer pipe networks. *Canadian Journal of Civil Engineering*, 41(8), 761-768
- DWA-M 149-3. 2011. *Advisory Leaflet DWA-M 149-3 - Conditions and Assessment of Drain and Sewer Systems Outside Buildings – Part 3: Condition Classification and Assessment*. German original published 2007, English translation published 2011, DWA German Association for Water, Wastewater and Waste, Hennef
- Egger, C., Scheidegger, A., Reichert, P., & Maurer, M. 2013. Sewer deterioration modeling with condition data lacking historical records. *Water Research*, 47(17), 6762-6779
- EN 13508-2. 2011. *Condition of drain and sewer systems outside buildings – Part 2: Visual inspection coding system + appendix A1 (Consolidated version)*. European Committee for Standardization (CEN), Brussels, Belgium
- EPA. 2002. *Factsheet: Asset Management for Sewer Collection Systems*. United States Environmental Protection Agency (EPA), Washington, D.C., United States
- El-Housni, H., Ouellet, M. & Duchesne, S. 2018. Identification of most significant factors for modeling deterioration of sewer pipes. *Canadian Journal of Civil Engineering*, 45(3), 215-226

- Fuchs-Hanusch, D., Günther, M., Möderl, M., Muschalla, D. 2015. Cause and effect oriented sewer degradation evaluation to support scheduled inspection planning. *Water Science and Technology*, 72(7), 1176-83
- Gramopadhye, A. K. and Wilson, K. 1997. Noise, feedback training, and visual inspection performance. *International Journal of Industrial Ergonomics*, 20(3), 223-230
- Guo, W., Soibelman, L., Garrett Jr., J.H. 2009. Automated defect detection for sewer pipeline inspection and condition assessment. *Automation in Construction*, 18(5), 587-596
- Halfawy, M.R., Hengmeechai, J. 2014. Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine. *Automation in Construction*, 38, 1-13
- Harvey, R. R. and McBean, E. A. 2014. Predicting the structural condition of individual sanitary sewer pipes with random forests. *Canadian Journal of Civil Engineering*, 41(4), 294-303
- Hassan, M. H. and Diab, S. L. 2010. Visual inspection of products with geometrical quality characteristics of known tolerances. *Ain Shams Engineering Journal*, 1(1), 79-84
- Heidl, W., Thumfart, S., Lughofer, E., Eitzinger, C. and Klement, E. P. 2013. Machine learning based analysis of gender differences in visual inspection decision making. *Information Sciences*, 224, 62-76
- Hernández, N., Caradot, N., Sonnerberg, H., Rouault, P. and Torres, A. 2017. Support tools to predict the critical structural condition of uninspected sewer pipes in Bogota D.C. In: *Proceedings of the 7th IWA Leading Edge Strategic Asset Management Conference*, 20-22 June 2017, Trondheim, Norway
- Hernández, N., Caradot, N., Sonnenberg, H., Rouault, P. and Torres, A. 2018. Support tools to predict the critical structural condition of uninspected pipes for case studies of Germany and Colombia. *Water Practice & Technology*, 13(4), 794-802
- Hüben, S. 2002. *Einflüsse auf die Qualität der Zustandsklassifizierung von Kanalhaltungen*. Thesis (MSc), Rheinisch – Westfälisch Technische Hochschule Aachen, Germany
- IPK. 2014. *Forderungskatalog 2014: Forderungskatalog für funktionsfähige öffentliche und private Abwasseranlage*. Impulse Pro Kanalbau, Munich, Germany
- Ishwaran, H., Kogalur, B., Gorodeski, Z., Minn, J., Laue, S. 2010. High-Dimensional Variable Selection for Survival Data. *Journal of the American Statistical Association*, 105(489), 205-217
- Jiang, G., Keller, J., Bond, P.L., Yuan, Z. 2016. Predicting concrete corrosion of sewers using artificial neural network. *Water Research*, 92, 52-60

- Jin, Y. and Mukherjee, A. 2010. Modeling Blockage Failures in Sewer Systems to Support Maintenance Decision Making. *Journal of Performance of Constructed Facilities*, 24(6), 622-633
- Johnson, S.G. 2014. *The NLOpt nonlinear-optimization package*. Package documentation available under <http://ab-initio.mit.edu/nlopt>
- Kessili, A., & Benmamar, S. 2016. Prioritizing sewer rehabilitation projects using AHP-PROMETHEE II ranking method. *Water Science and Technology*, 73(2), 283-291
- KfW - Deutsches Institut für Urbanistik – Difu. 2016. *KfW-Kommunalpanel 2016*. Studie im Auftrag der KfW, Frankfurt am Main, Germany
- Kleiner, Y. & Rajani, B. 2001a. Comprehensive Review of Structural Deterioration of Water Mains: Statistical Models. *Urban Water Journal*, 3(3), 131-150
- Kleiner, Y. 2001b. Scheduling inspection and renewal of large infrastructure assets, *Journal of Infrastructure Systems*, 7(4) 136-143
- Kleiner, Y., Sadiq, R. and Rajani, B. 2006. Modelling the deterioration of buried infrastructure as a fuzzy Markov process. *Journal of Water Supply: Research and Technology - AQUA*, 55(2), 67-80
- Kley, G., Kropp, I., Schmidt, T., Caradot, N. 2013a. *Review of technologies and methodologies for sewer condition evaluation*. KWB project SEMA, Report 1.1, available on KWB website, retrieved on February 26 2019, 39p, Berlin, Germany
- Kley, G., Caradot, N. 2013b. *Review of sewer deterioration models*. KWB project SEMA, Report 1.2, available on KWB website, retrieved on February 26 2019, 43p, Berlin, Germany
- Korving, H. and Clemens, F. 2004. Reliability of coding of visual inspections of sewers. In: *Proceedings of 4th International Conference on Sewer Processes and Networks*, Funchal, Madeira, Portugal
- Kumar, S.S., Abraham, D.M., Jahanshahi, M.R., Iseley, T., Starr, J. 2018. Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks. *Automation in Construction*, 91, 273-283
- Laakso, T., Kokkonen, T., Mellin, I., Vahala, R. 2018. Sewer condition prediction and analysis of explanatory factors. *Water (Switzerland)*, 2018(10), 1-17
- Laofo, C. and Peansupap, V. 2012. Defect detection and quantification system to support subjective visual quality inspection via a digital image processing: A tiling work case study. *Automation in Construction*, 24, 160-174

- Large, A. 2016. *Une meilleure gestion patrimoniale des réseaux d'eau potable : le modèle de prévision du renouvellement à long terme OPTIMEAU*. PhD thesis, Université de Bordeaux, France
- Le Gat, Y. 2008. Modelling the deterioration process of drainage pipelines. *Urban Water Journal*, 5(2), 97-106
- Le Gat, Y., Kropp, I., Poulton, M. 2013. Is the service life of water distribution pipelines linked to their failure rate? *Water Science and Technology: Water Supply*, 13(2), 386-393
- Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C., Desmulliez, J. J. 2004. *Gestion patrimoniale des réseaux d'assainissement urbains, Guide méthodologique*. Paris: Lavoisier Tec&Doc, 416 p
- Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C., & Desmulliez, J.J. 2007. Performance indicators and multicriteria decision support for sewer asset management. *Journal of Infrastructure Systems*, 13(2), 105-114
- Li, D., Cong, A., Guo, S. 2019. Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification. *Automation in Construction*, 101, 199-208
- Liaw, A. and Wiener, M. 2002. Classification and regression by random forest. *R News*, 2/3, 18-22
- Marlow, D., Davis, P., Trans, D., Beale, D., Burn, S. and Urquhart, A. 2009. *Remaining Asset Life: A State of the Art Review - Strategic asset management and communication*. Report, Water Environment Research Foundation (WERF), Virginia, Unites States
- Marzouk, M. and Osama, A. 2015. Fuzzy approach for optimum replacement time of mixed infrastructures. *Civil Engineering and Environmental Systems*, 32(3), 269-280
- Marzouk, M. and Osama, A. 2017. Fuzzy-Based Methodology for Integrated Infrastructure Asset Management. *International Journal of Computational Intelligence Systems*, 10(1), 745-759.
- Mashford, J., Marlow, D., Tran, D., & May, R. 2010. Prediction of sewer condition grade using support vector machines. *Journal of Computing in Civil Engineering*, 25(4), 283-290
- Micevski, T., Kuczera, G. and Coombes, P. 2002. Markov Model for Storm Water Pipe Deterioration. *Journal of Infrastructure Systems*, 8(2), 49-56
- Morales, V. M., Mier, J. M. and Garcia, M. H. 2017. Innovative modeling framework for combined sewer overflows prediction. *Urban Water Journal*, 14(1), 97-111

- Müller, K. and Fischer, B. 2007. Objective condition assessment of sewer systems. In: *Proceedings of LESAM 2007 - 2nd Leading Edge Conference on Strategic Asset Management*, 17-19 October, Lisbon, Portugal,
- Nafi, A. and Kleiner, Y. 2010. Scheduling Renewal of Water Pipes While Considering Adjacency of Infrastructure Works and Economies of Scale. *Journal of Water Resources Planning and Management*, 136(5), 519-530
- Najafi, M. and Kulandaivel, G. 2005. Pipeline Condition Prediction Using Neural Network Models. In: *Pipeline Division Specialty Conference 2005*, 21-24 August, Houston, Texas, United States
- NASSCO. 2007. *Pipeline Assessment And Certificate Program (PACP)*. National Association of Sewer Service Companies, Marriottsville, United States
- Nariné Torres, M., Rodríguez, J.P., Leitão, J.P. 2017. Geostatistical analysis to identify characteristics involved in sewer pipes and urban tree interactions. *Urban Forestry & Urban Greening*, 25, 36-42
- Oelmann, M., Roters, B., Hoffjan, A., Hippe, M., Wedmann, T. 2017. Investitionsstau in der Abwasserentsorgung. *KA - Korrespondenz Abwasser, Abfall*, 2, 131-138
- ONEMA, 2012. *Observatory on public water and sanitation services, overview of services and of their performance*. French National Agency for Water and Aquatic Environments, Vincennes, France
- Ouellet, M. and Duchesne, S. 2018. A method to consider replaced pipes in modeling of sewer pipe deterioration. *Urban Water Journal*, 15(4), 392-397
- Plihal, H., Kuratko, A. and Ertl, T. 2014. Einsatz des elektronischen Spiegels für den betrieblichen Überblick im Kanalsystem. *Österreichische Wasser- und Abfallwirtschaft*, 66(3-4), 103-111
- Rodríguez, J.P., McIntyre, N., Díaz-Granados, M., Maksimović, Č. 2012. A database and model to support proactive management of sediment-related sewer blockages, *Water Research*, 46(15), 4571-4586
- Roghani, B., Cherqui, F., Ahmadi, M., Le Gauffre, P. & Tabesh, M. 2019. Dealing with uncertainty in sewer condition assessment: Impact on inspection programs, *Automation in Construction*, 103, 117-126.
- Rokstad, M. M., Le Gat, Y. and Ugarelli, R. 2014. On the uncertainty of probabilistic sewer deterioration forecast modelling. In: *Proceedings of 11th International Conference on Hydroinformatics*, 17-21 August, New York City, United States

- Rokstad, M. M., Ugarelli R. 2015 Evaluating the role of deterioration models for condition assessment of sewers. *Journal of Hydroinformatics*, 17(5), 789-804
- Runarsson, T.P. and Yao, X. 2005. Search biases in constrained evolutionary optimization. *IEEE Trans. on Systems, Man, and Cybernetics Part C: Applications and Reviews*, 35(2), 233-243
- Salman, B. 2010. *Infrastructure Management and Deterioration Risk Assessment of Wastewater Collection Systems*. PhD thesis, University of Cincinnati, United States
- Salman, B., Salem, O. 2012. Modeling Failure of Wastewater Collection Lines Using Various Section-Level Regression Models. *Journal of Infrastructure Systems*, 18(2), 146-154
- Santos, P., Amado, C., Coelho, S.T., Leitão, J.P. 2017. Stochastic data mining tools for pipe failure prediction. *Urban Water Journal*, 14(4), 343-353.
- Scheidegger, A., Hug, T., Rieckermann, J., and Maurer, M. 2011. Network condition simulator for benchmarking sewer deterioration models. *Water Research*, 45(16), 4983-4994
- Scheidegger, A., & Maurer, M. 2012. Identifying biases in deterioration models using synthetic sewer data. *Water Science and Technology*, 66(11), 2363-2369
- Scheidegger, A., Scholten, L., Maurer, M., Reichert, P. 2013. Extension of pipe failure models to consider the absence of data from replaced pipes. *Water Research*, 47(11), 3696-3705
- Schmidt, T., 2009. *Modellierung von Kanalalterungsprozessen auf der Basis von Zustandsdaten*. PhD thesis, Institut für Stadtbauwesen und Straßenbau, Technical University Dresden, Dresden, Germany
- SE|BS. 2019. Website: <http://www.stadtentwaesserung-braunschweig.de/rund-ums-abwasser/kanalnetz/>
- Sousa, V., Ferreira, F. M., Meireles, I., Almeida, N. and Saldanha Matos, J. 2014a. Comparison of the inspector and rating protocol uncertainty influence in the condition rating of sewers. *Water Science and Technology*, 69(4), 862-867
- Sousa, V., Matos, J.P., Matias, N. 2014b. Evaluation of artificial intelligence tool performance and uncertainty for predicting sewer structural condition. *Automation in Construction*, 44, 84-91
- SüwVO Abw. 2013. *Verordnung zur Selbstüberwachung von Abwasseranlagen - Selbstüberwachungsverordnung Abwasser – SüwVO Abw.* Gesetz- und Verordnungsblatt (GV. NRW.) Nr. 33, 601-612

- Syachrani, S., Jeong, H., and Chung, C., 2013. Decision tree based deterioration model for buried wastewater pipelines. *Journal of Performance of Constructed Facilities*, 27(5), 633-645
- Tran, D., Perera, A. W. M., Davis, P. 2006. Application of probabilistic neural networks in modeling structural deterioration of stormwater pipes. *Urban Water Journal*, 3(3), 175-184
- Tran, D., Ng, A. and Perera, B. 2007. Neural networks deterioration models for serviceability condition of buried stormwater pipes. *Engineering Applications of Artificial Intelligence*, 20(8), 1144-1151
- Tran, D. 2016. Sensitivity of Markov Model to Different Sampling Sizes of Condition Data. *Journal of Performance of Constructed Facilities*, 30(4)
- Tscheikner-Gratl, F., Sitzenfrei, R., Stibernitz, C., Rauch, W., and Kleidorfer, M. 2015. Integrated rehabilitation management by prioritization of rehabilitation areas for small and medium sized municipalities. In: *World Environmental and Water Resources Congress 2015: Floods, Droughts, and Ecosystems - Proceedings of the 2015 World Environmental and Water Resources Congress, 2045-2057*
- Tscheikner-Gratl, F., Sitzenfrei, R., Rauch, W. and Kleidorfer, M. 2016. Integrated rehabilitation planning of urban infrastructure systems using a street section priority model. *Urban Water Journal*, 13(1), 28-40
- Ugarelli, R.M., Selseth, I., Le Gat, Y., Rostum, J., Krogh, A.H. 2013. Wastewater pipes in Oslo: from condition monitoring to rehabilitation planning. *Water Practice and Technology*, 8(3-4), 487-494
- van der Steen, A. J., Dirksen, J. and Clemens, F. H. 2014. Visual sewer inspection: detail of coding system versus data quality? *Structure and Infrastructure Engineering*, 10(11), 1385-1393
- van Riel, W., Langeveld, J., Herder, P. and Clemens, F. 2014. Intuition and information in decision-making for sewer asset management. *Urban Water Journal*, 11(6), 506-518
- van Riel, W., van Bueren, E., Langeveld, J., Herder, P., Clemens, F. 2016. Decision-making for sewer asset management: Theory and practice. *Urban Water Journal*, 13(1), 57-68
- van Riel, W., Langeveld, J., Herder, P., Clemens, F. 2017. The influence of information quality on decision-making for networked infrastructure management. *Structure and Infrastructure Engineering*, 13(6), 696-708

- Vitorino, D., Coelho, S.T., Santos, P., Sheets, S., Jurkovic, B., Amado, C. 2014. A Random Forest Algorithm Applied to Condition-based Wastewater Deterioration Modeling and Forecasting. *Procedia Engineering*, 89, 401-410
- Werey, C., Rozan, A., Wittner, C., Le Gat, Y., Le Gauffre, P., Nirsimloo, K., Leclerc, C. 2012. Gestion patrimoniale des réseaux d'assainissement: de l'état des réseaux à la planification de leur réhabilitation – Outils, méthodes et perspectives. *Sciences Eaux & Territoires*, 9(4), 44-53
- WERF. 2007. *Infrastructure Management Final Report: Condition Assessment Strategies and Protocols for Water and Wastewater Utility Assets*. Water Environment Research Foundation, United States
- WERF. 2012. *Towards an economic decision methodology for remaining asset life*. Report SAM1R06g, Water Environment Research Foundation, United States
- WRc. 2013. *Manual of Sewer Condition Classification - 5th Edition*. Water Research Centre, UK
- Wirahadikusumah, R., Abraham, D., Iseley, T. 2001. Challenging issues in modeling deterioration of combined sewers. *Journal of Infrastructure Systems*, 7(2), 77-84
- Yazdi, J., Lee, E. and Kim, J. 2015. Stochastic Multiobjective Optimization Model for Urban Drainage Network Rehabilitation. *Journal of Water Resources Planning and Management*, 141(8)

FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : Caradot

DATE de SOUTENANCE : 27/06/2019

Prénoms : Nicolas Claude Jacques

TITRE : L'utilisation de modèles de détérioration pour l'élaboration de stratégies de gestion patrimoniale des réseaux d'assainissement

NATURE : Doctorat

Numéro d'ordre : 2019LYSEI034

Ecole doctorale : MEGA

Ecole Doctorale N° 162 (Mécanique, Energétique, Génie civil, Acoustique)

Spécialité : Génie Civil

RESUME :

Les infrastructures de collecte et de traitement des eaux usées représentent un investissement considérable pour les municipalités. La plupart des collectivités sont aujourd'hui confrontées au problème de la gestion durable de leur patrimoine vieillissant avec des besoins urgents de réhabilitation des infrastructures. Elles se doivent donc d'élaborer des stratégies de gestion patrimoniale efficace pour maintenir le niveau de service rendu et limiter l'augmentation des redevances.

L'inspection télévisée (ITV) est la méthode standard utilisée depuis des décennies pour l'évaluation de la condition structurelle des réseaux d'assainissement. Dû à l'absence de réglementation et aux coûts élevés d'inspection, la plupart des collectivités n'ont qu'une connaissance partielle de la condition de leur réseau et manquent d'information pour planifier efficacement les actions de maintenance et renouvellement. Pour pallier à cette situation, des modèles numériques de détérioration ont été développés pour simuler la condition des conduites non inspectées et prévoir l'évolution future de la condition des réseaux.

Un des obstacles majeurs à l'appropriation de ces outils par les collectivités est le manque de preuves à grande échelle de leur bonne performance. D'autre part, l'influence de la quantité de données disponibles pour étalonner les modèles sur la fiabilité des prédictions est mal connue. Les taux d'inspection étant relativement faibles, les collectivités sont rarement assurées d'avoir les données nécessaires pour développer des modèles fiables. Enfin, de nombreuses collectivités reconnaissent les incertitudes liées aux données d'inspection mais aucune étude n'a pu mettre en évidence l'influence de ces incertitudes sur les résultats de modélisation.

Ce travail de thèse vise à évaluer la performance des modèles de détérioration ainsi qu'à évaluer l'influence de la quantité de données disponibles sur la qualité de prédiction. Enfin, cette thèse a pour objectif de quantifier la fiabilité des données d'inspection et d'analyser l'influence de leur incertitude sur la performance des modèles de détérioration.

MOTS-CLÉS : réseau d'assainissement ; modèle ; inspection ; incertitudes ; gestion patrimoniale

Laboratoire (s) de recherche : DEEP

Directeurs de thèse: Frederic Cherqui, Francois Clemens (co-directeur)

Président de jury : Jean-Luc Bertrand-Krajewski

Composition du jury :

Jean-Luc Bertrand-Krajewski, Professeur (INSA Lyon, France)

Frédéric Cherqui, Maître de Conférences HDR (Université Lyon 1)

Francois Clemens, Professeur (TU Delft, Pays-Bas)

Sophie Duchesne, Professeur (INRS, Québec, Canada)

Andreas Hartmann, Directeur (SE|BS, Allemagne)

Yves Le Gat, Ingénieur en Chef des Ponts, des Eaux et des Forêts (IRSTEA, France)

João P. Leitão, Chargé de recherche (EAWAG, Suisse)

Pascale Rouault, Responsable de département (KWB, Allemagne)

Président de jury

Co-directeur de thèse

Co-directeur de thèse

Rapporteur

Invité

Examinateur

Rapporteur

Invitée