ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ & ΠΕΡΙΒΑΛΛΟΝΤΟΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Διερεύνηση της αξιοπιστίας μοντέλων υπολογισμού τρωτότητας σε αγωγούς ακαθάρτων. Εφαρμογή σε πόλη της Βόρειας Γερμανίας.

Μαρία Ξενοχρήστου

Επιβλέπων: Χρήστος Μακρόπουλος,

Επίκουρος καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2015

Φωτογραφία εξωφύλλου: ‘Light at the end of the tunnel’,

The chimney on Troopers Hill, Bristol

Πηγή: http://panthera-lee.deviantart.com/art/Light-at-the-End-of-the-Tunnel-205911061

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ & ΠΕΡΙΒΑΛΛΟΝΤΟΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Διερεύνηση της αξιοπιστίας μοντέλων υπολογισμού τρωτότητας σε αγωγούς ακαθάρτων. Εφαρμογή σε πόλη της Βόρειας Γερμανίας.

Μαρία Ξενοχρήστου

Επιβλέπων: Χρήστος Μακρόπουλος,

Επίκουρος καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2015

Acknowledgment

The current thesis was written in collaboration with the German research institute KompetenzZentrum Wasser Berlin as part of project SEMA (Asset Management Strategies for Sewer Systems). Therefore I feel the need to write this part in English, in order to express my gratitude to all the wonderful people and excellent scientists I had the chance to work with during my stay in Berlin.

First of all I would like to thank my supervisor in Germany, Nicolas Caradot, for his guidance and support, as well as his patience and help in every aspect of this project. I feel grateful not only for the opportunity he gave me to work on such an interesting topic, but also for our excellent collaboration in Berlin, as well as in Athens, that still inspires and motivates me for my next steps.

Next, I would like to thank my supervisor in Greece, assistant professor Christos Makropoulos, for his inspiring lectures, which triggered my interest in the topic, as well as his assistance on this thesis.

I would also like to thank Hauke Sonnenberg, for introducing me to R programming language, and for his constant support on any technical issues that arose, as well as assistant professor Konstantinos Mamasis for his spontaneous and at the same time vital help when it was mostly needed.

Lastly, I would like to thank my friend Victor for all the interesting ideas and discussions, and for speaking English to me at lunch in order not to improve my German, as well as all the people from KWB for their kindness and understanding. I would also like to thank Liam, for proofreading this document and offering me technical support, as well as his own laptop, in order to complete this thesis.

And of course I would like to thank my family for providing me with financial support for a bit too long.

This document was originally written in English and later translated to Greek as a shorter version of the original text. It is my sincere hope, leaving National Technical University of Athens, and feeling extremely grateful for the knowledge it provided me with, that future students will be allowed to hand their original work in the universal language of science.

Περιεχόμενα

Περίληψη

Καθώς η ανησυχία των αρμόδιων φορέων για την κατάσταση των υποδομών αποχέτευσης ομβρίων και ακαθάρτων αυξάνει, η ανάπτυξη αξιόπιστων και αποδοτικών στρατηγικών διαχείρισης γίνεται ολοένα και πιο απαραίτητη. Σκοπός αυτής της εργασίας είναι να ερευνήσει σε βάθος και να αξιολογήσει τις προοπτικές των υπαρχόντων μοντέλων υπολογισμού τρωτότητας να εκτιμήσουν το επίπεδο φθοράς των υπόγειων αγωγών. Για το σκοπό αυτό, η επίδοση του GompitZ, ενός μαθηματικού μοντέλου Markov, ερευνάται λεπτομερώς, χρησιμοποιώντας μια εκτενή βάση δεδομένων μιας Γερμανικής πόλης (31.394 οπτικές επιθεωρήσεις). Αρχικά, οι παράγοντες που συνεισφέρουν στην αλλοίωση της ποιότητας των αγωγών του δικτύου εντοπίζονται, και η επιρροή τους στη βαθμονόμηση και προγνωστική ακρίβεια του μοντέλου αξιολογείται. Το 70% των διαθέσιμων δεδομένων χρησιμοποιείται για τη βαθμονόμηση του μοντέλου, ενώ το υπόλοιπο 30% για τον έλεγχο των αποτελεσμάτων. Στο πρώτο στάδιο, η επίδοση του μοντέλου αξιολογήθηκε χωρίς τη χρήση μεταβλητών (age model). Στη συνέχεια, οι παράγοντες που διαπιστώθηκε ότι σχετίζονται με τη διάρκεια ζωής του δικτύου χρησιμοποιήθηκαν ως επεξηγηματικές μεταβλητές. Καθώς τα βραχυπρόθεσμα προγράμματα αποκατάστασης εστιάζουν στην επιδιόρθωση και αντικατάσταση των αγωγών που βρίσκονται σε κρίσιμη κατάσταση, μια καινούρια μέθοδος αξιολόγησης εισήχθη με σκοπό την εκτίμηση της ικανότητας του μοντέλου να εντοπίσει τους αγωγούς που βρίσκονται στην δυσμενέστερη κατάσταση. Σύμφωνα με την παραπάνω μέθοδο, η ευαισθησία του μοντέλου στο μέγεθος του συνόλου δεδομένων εισόδου (χρησιμοποιώντας 100%, 20%, και 10% του διαθέσιμου δείγματος) διερευνήθηκε χρησιμοποιώντας προσομοίωση Monte Carlo.

Abstract

As the concern of municipalities over the condition of drainage infrastructures increases, the development of reliable and cost-effective asset management strategies becomes of utmost importance. The aim of this study is to investigate and evaluate the potentials of sewer deterioration models to predict the future condition of urban drainage pipelines. In order to achieve this, the performance of GompitZ, a Markov-based sewer deterioration model is investigated, using an extensive sewer and CCTV dataset from a German case study (31,394 classified inspections). After the configuration and processing of the available data, a preliminary analysis was performed in order to identify the most influential deterioration factors and carefully assess their influence on the model's calibration and accuracy of prediction. The model was calibrated using 70% of the available data, whilst the remaining 30% was used to validate it. Its performance was assessed using the most influential deterioration factors identified above as explanatory covariates, as well as no covariates. Since short-term rehabilitation programs focus on repair and replacement of pipes in poor condition, a new assessment method was introduced to evaluate the ability of the model to identify the most deteriorated sewers. In accordance to this method, the model's sensitivity to the size of the available dataset (using 100%, 20%, and 10% of the sample) was investigated using Monte Carlo simulation.

1 Εισαγωγή

Το δίκτυο αποχέτευσης ομβρίων και ακαθάρτων αποτελεί ένα αναπόσπαστο κομμάτι των αστικών υποδομών και η δυσλειτουργία ή και αστοχία των μελών του μπορεί να θέσει σε κίνδυνο την περιβαλλοντική ισορροπία, καθώς και τη δημόσια υγεία και ασφάλεια (Becker *et al.*, 2009). Οι συνέπειες μιας αστοχίας σε ένα μεγάλο αγωγό μπορεί να περιλαμβάνουν τη διακοπή της κυκλοφορίας, λόγω καταστροφών σε κεντρικές οδικές αρτηρίες, κινδύνους για την υγεία, καθώς και ένα τεράστιο κόστος επιδιόρθωσης των βλαβών που προκλήθηκαν, επιπρόσθετο του κόστους αποκατάστασης του αγωγού (WERF, 2013).

Την ίδια στιγμή, διάφορες έρευνες που διεξήχθησαν στο Ηνωμένο Βασίλειο, την Αυστραλία και τις Ηνωμένες Πολιτείες, επιστούν την προσοχή στην επιβαρυμένη κατάσταση των υπόγειων αγωγών, που ολοένα και χειροτερεύει. Στη Γερμανία, μια έρευνα υπολόγισε ότι το 17% περίπου των αγωγών έχουν σοβαρές βλάβες και απαιτούν άμεση ή βραχυπρόθεσμη αποκατάσταση (DWA, 2009).

Ωστόσο, η επένδυση για τις ανάγκες διατήρησης και αποκατάστασης του δικτύου αδυνατεί να παρέχει στους χρήστες το απαιτούμενο επίπεδο εξυπηρέτησης (WERF, 2007). Σύμφωνα με έρευνα η οποία διεξήχθη από την αμερικανική υπηρεσία περιβαλλοντικής προστασίας (EPA, 2008), οι συνολικές ανάγκες χρηματοδότησης για αντικατάσταση, αποκατάσταση και επέκταση του υπάρχοντος δικτύου συλλογής, καθώς και για την κατασκευή καινούριων δικτύων, για μια περίοδο 20 ετών στις Ηνωμένες Πολιτείες, ανέρχονται σε 65,3 δισεκατομμύρια δολάρια. Το ποσό αυτό αντιστοιχεί στο 32,2% των συνολικών αναγκών των δημόσιων υπηρεσιών για την επεξεργασία και συλλογή των λυμάτων.

Για να αποφευχθούν οι συνέπειες μιας δυνητικής αστοχίας και να επιτραπεί ο εκ των προτέρων σχεδιασμός μελλοντικών επενδύσεων, είναι απαραίτητη η ανάπτυξη αποτελεσματικών στρατηγικών διαχείρισης, που θα επιτρέψουν την προνοητική εφαρμογή σχεδίων αποκατάστασης εκεί που απαιτείται, πριν υπάρξει κάποιο ορατό πρόβλημα.

1.1 Θεωρητικό υπόβαθρο

Η ανάπτυξη και ο σχεδιασμός αποτελεσματικών στρατηγικών διαχείρισης προϋποθέτουν καλή γνώση της κατάστασης του συστήματος, καθώς και των αδύναμων σημείων του. Στα πλαίσια αυτής της προσπάθειας, συχνές οπτικές επιθεωρήσεις μέσω κλειστού κυκλώματος τηλεόρασης (CCTV) είναι απαραίτητες, ώστε να εντοπιστούν εσωτερικές αστοχίες, και να εκτιμηθεί η διάρκεια ζωής των αγωγών.

Στις περισσότερες χώρες, οι φθορές στους αγωγούς οι οποίες καταγράφονται κατά τη διάρκεια επιθεωρήσεων με οπτικά μέσα, κωδικοποιούνται σύμφωνα με τυποποιημένα συστήματα και η συνολική κατάσταση του αγωγού αξιολογείται χρησιμοποιώντας μια μεθοδολογία κατάταξης (για παράδειγμα στη Γαλλία, Le Gauffre *et al.,* 2004· στο Ηνωμένο Βασίλειο, WRc, 2004· στις Ηνωμένες Πολιτείες, NASSCO, 2007· στη Γερμανία, DWA M149-3, 2011).

Παρότι οι οπτικές επιθεωρήσεις είναι ο πιο διαδεδομένος τρόπος παρατήρησης δυσλειτουργιών στο εσωτερικό των αγωγών, εξακολουθούν να αποτελούν μια αρκετά χρονοβόρα και δαπανηρή διαδικασία, η οποία είναι επιρρεπής σε λάθη, λόγω της υποκειμενικότητας της ανθρώπινης κρίσης. Σαν αποτέλεσμα, τα ποσοστά τους είναι σε γενικές γραμμές πολύ χαμηλά και η κατάσταση ενός μικρού μόνο μέρους του δικτύου είναι γνωστή (ONEMA, 2012).

Καθώς η ανάπτυξη στρατηγικών αποκατάστασης περιορίζεται από την έλλειψη πληροφοριών σχετικά με την κατάσταση των αγωγών, ένα πλήθος μαθηματικών μοντέλων αναπτύχθηκε με στόχο την προσομοίωση της εξέλιξης των δικτύων αποστράγγισης, σύμφωνα με την παρούσα και περασμένη κατάστασή του (Tran, 2007· Chughtai και Zayed, 2008· Le Gat, 2008· Ana *et al.,* 2009· Salman, 2010· Khan *et al*., 2010· Ens, 2012· Ahmadi *et al.,* 2013). Τα παραπάνω μοντέλα μπορούν να χρησιμοποιηθούν είτε (α) για την προσομοίωση της κατάστασης των αγωγών που δεν έχουν επιθεωρηθεί, είτε (β) για την πρόγνωση της εξέλιξης της κατάστασης ολόκληρου του συστήματος.

Ωστόσο, τα περισσότερα από αυτά αδυνατούν να αποδείξουν ότι μπορούν να προβλέψουν επαρκώς μελλοντικές συνθήκες (Ana και Bauwens, 2010· Scheidegger *et al.,* 2011). Η ακρίβεια, καθώς και η ευαισθησία των μοντέλων στα διαθέσιμα δεδομένα εισόδου, αλλά και σε παραμέτρους που καθορίζονται από το χρήστη, δεν έχει ακόμα αποσαφηνιστεί. Διάφορες μέθοδοι μοντελοποίησης είναι σήμερα διαθέσιμες, αλλά όχι ευρέως χρησιμοποιούμενες από δήμους και αρμόδιους φορείς για το σχεδιασμό κατάλληλων στρατηγικών (Kley και Caradot, 2013). Η ενίσχυση της εμπιστοσύνης του τελικού χρήστη στα μοντέλα προσομοίωσης τρωτότητας θα οδηγούσε με βεβαιότητα σε διεύρυνση της χρήσης τους.

1.2 Επισκόπηση της εργασίας

Η παρούσα εργασία στοχεύει στην επίλυση του παραπάνω προβλήματος απαντώντας σε τρεις πρακτικές ερωτήσεις:

- Πώς αξιολογείται η ακρίβεια πρόγνωσης ενός μαθηματικού μοντέλου; Εφόσον οι στρατηγικές αποκατάστασης εστιάζουν στον εντοπισμό των αγωγών που βρίσκονται στη δυσμενέστερη κατάσταση, μια ουσιαστική αξιολόγηση θα εστίαζε στους κρίσιμους αγωγούς που το μοντέλο είναι ικανό να εντοπίσει, αντί για το σύνολο των ακριβών προγνώσεων.
- Πώς μπορούν να εντοπιστούν οι βέλτιστες παράμετροι, και ποιά η επιρροή τους στην προγνωστική ακρίβεια του μοντέλου; Αν και η βέλτιστη θεώρηση μεταβλητών μπορεί να ποικίλει σε κάθε περίπτωση, η ενσωμάτωση στο μοντέλο παραγόντων που επηρεάζουν το ρυθμό φθοράς του αγωγού μπορεί να έχει σημαντική επίδραση στην ποιότητα της πρόγνωσης, η οποία μένει να εξακριβωθεί.
- Τι ακρίβεια πρόγνωσης μπορεί να επιτευχθεί χρησιμοποιώντας ένα εκτενές σύνολο οπτικών παρατηρήσεων και ποία είναι η ευαισθησία του μοντέλου στο μέγεθος του εν λόγω συνόλου; Οι περισσότερες πόλεις διαθέτουν περιορισμένη γνώση της κατάστασης του υπόγειου δικτύου τους, γεγονός που επιβάλλει τη διερεύνηση των περιορισμών που αυτό μπορεί να θέτει στην ανάπτυξη αξιόπιστων μοντέλων.

Για να απαντηθούν οι παραπάνω ερωτήσεις, ένα καινούριο εργαλείο αναπτύχθηκε, βασισμένο στο μοντέλο Gompitz, και τη γλώσσα προγραμματισμού R, το οποίο περιλαμβάνει ενότητες για προετοιμασία και επεξεργασία των δεδομένων, καθώς και αξιολόγηση των αποτελεσμάτων.

Η απόδοση του μοντέλου εξετάζεται αρχικά χρησιμοποιώντας ένα εκτενές σύνολο δεδομένων μίας γερμανικής πόλης (31.394 οπτικές καταγραφές), χωρίς την εισαγωγή μεταβλητών (age model). Εν συνεχεία, αναπτύσσεται και εφαρμόζεται μια μεθοδολογία με σκοπό τον εντοπισμό των σημαντικότερων μεταβλητών και την αξιολόγηση της επιρροής τους στην απόδοση του μοντέλου. Τέλος, η ποιότητα πρόβλεψης αξιολογείται χρησιμοποιώντας προσομοίωση Monte Carlo (1000 δοκιμές) για διαφορετικά μεγέθη δεδομένων, χρησιμοποιώντας το 100%, 20%, και 10% του συνόλου.

Τα τελικά συμπεράσματα που προκύπτουν από την παραπάνω ανάλυση αφορούν την απόδοση του μοντέλου, σε συνδυασμό με την εξάρτησή της από τις μεταβλητές εισόδου, καθώς και την ευαισθησία του μοντέλου στο μέγεθος του διαθέσιμου συνόλου δεδομένων.

2 Μέθοδοι μαθηματικής προσομοίωσης τρωτότητας αγωγών

Μια μεγάλη ποικιλία μοντέλων έχει αναπτυχθεί, με σκοπό την προσομοίωση και πρόγνωση της τρωτότητας των αστικών δικτύων αποστράγγισης. Σύμφωνα με τη βασική αρχή λειτουργίας τους, μπορούν να ταξινομηθούν σε τρεις κατηγορίες: ντετερμινιστικά, στατιστικά και μοντέλα τεχνητής νοημοσύνης (Kley και Caradot, 2013). Ενώ τα ντετερμινιστικά και τα στατιστικά μοντέλα συσχετίζουν τους παράγοντες φθοράς με την κατάσταση του αγωγού χρησιμοποιώντας την αντίστοιχη μαθηματική εξίσωση, τα μοντέλα τεχνητής νοημοσύνης επιχειρούν να προσομοιώσουν τη λειτουργία του ανθρώπινου νευρικού συστήματος, διερευνώντας τρόπους και δημιουργώντας μαθηματικές σχέσεις μεταξύ των εξαρτημένων (κατάσταση αγωγού) και των ανεξάρτητων (παράγοντες φθοράς) μεταβλητών (WERF, 2009).

Στα στατιστικά μοντέλα, η θεωρία των αλυσίδων Markov είναι μια μεθοδολογία που χρησιμοποιείται ευρέως για την πρόγνωση της μελλοντικής κατάστασης διαφόρων τύπων υποδομών (Le Gat, 2008). Η εν λόγω θεωρία περιγράφει τη συμπεριφορά ενός συστήματος που περνάει από ένα πεπερασμένο πλήθος καταστάσεων. Σε κάθε χρονικό βήμα, το σύστημα μπορεί να αλλάξει την κατάστασή του από την τρέχουσα σε μια χειρότερη, ή να παραμείνει στην ίδια κατάσταση, σύμφωνα με μια δεδομένη πιθανότητα. Διάφοροι συγγραφείς έχουν ήδη επιδείξει εφαρμογές του μοντέλου Markov στην πρόγνωση της δομικής αλλοίωσης των υπόγειων αγωγών (Mehle *et al.*, 2001· Wirahadikusumah *et al.*, 2001· Micevski *et al.*, 2002· Baik *et al.*, 2006· Tran, 2007· Ana, 2009· Ugarelli *et al.*, 2013). Διάφορα εργαλεία που βασίζονται στη θεωρία Markov και απευθύνονται σε αγωγούς έχουν αναπτυχθεί τα τελευταία 15 χρόνια (KANEW, Kropp και Baur, 2005· GompitZ, Le Gat, 2008· STATUS, Stein και Gedheri, 2009).

2.1 Το μοντέλο GompitZ

Στην παρούσα εργασία, το GompitZ, ένα στατιστικό μοντέλο Markov, χρησιμοποιήθηκε για τη μοντελοποίηση της δομικής αλλοίωσης των αγωγών αποχέτευσης. Το GompitZ αναπτύχθηκε από τον Le Gat (2008) στο πλαίσιο του ευρωπαϊκού ερευνητικού προγράμματος CareS (Saegrov, 2006). Μία εισαγωγή πάνω στο θεωρητικό υπόβαθρο του μοντέλου βρίσκεται στο άρθρο του Le Gat (2008), "Modelling the deterioration process of drainage pipelines".

Το GompitZ χωρίζεται σε δύο ενότητες, σχεδιασμένες για βαθμονόμηση του μοντέλου και μελλοντική πρόγνωση. Για τη βαθμονόμηση απαιτούνται οι παρακάτω πληροφορίες:

- Μία κλάση για κάθε αγωγό: τα οπτικά δεδομένα από τις επιθεωρήσεις μεταφράζονται αρχικά σε κώδικες ελαττωμάτων και στη συνέχεια μέσω ενός αλγόριθμου ταξινόμησης σε μια κλάση για κάθε αγωγό. Αυτό μπορεί να γίνει είτε σε σχέση με την επιτακτικότητα για μέτρα αποκατάστασης (priority-based methodologies), είτε σε σχέση με το μήκος του αγωγού που χρειάζεται αποκατάσταση (substance-based methodologies) (Kley *et al.,* 2013). Η κλάση αυτή είναι χαρακτηριστική της υφιστάμενης κατάστασης του αγωγού, όσον αφορά τη δομή του, τη λειτουργία του, ή και τα δύο.
- Μεταβλητές του μοντέλου: μία λίστα παραγόντων στους οποίους περιλαμβάνονται κατασκευαστικά χαρακτηριστικά (υλικό, διάμετρος, είδος αποβλήτου), καθώς και περιβαλλοντικοί παράγοντες (τοποθεσία, ύψος υδροφόρου ορίζοντα) χρησιμοποιούνται ως επεξηγηματικές μεταβλητές, καθώς θεωρείται πιθανό να επηρεάζουν καθοριστικά το ρυθμό επιδείνωσης της κατάστασης των αγωγών (Davies *et al.,* 2001· Ana *et al.,* 2009).
- Συμπεριφορά των μεταβλητών: μια χαρακτηριστική τιμή (status) καθορίζεται για κάθε μεταβλητή, η οποία αντιπροσωπεύει τον τρόπο με τον οποίο η συγκεκριμένη μεταβλητή επιδρά στη φθορά των αγωγών: καμία επίδραση (τιμή 0), επίδραση στην αρχική κατάσταση του αγωγού (τιμή 1), επίδραση στη διεργασία επιδείνωσης (τιμή 2), ή επίδραση και στα δύο παραπάνω (τιμή 3).

Το μοντέλο υπολογίζει χρονικά εξαρτημένες πιθανότητες μετάβασης, δηλαδή πιθανότητες για κάθε αγωγό να παραμείνει στην ίδια κατάσταση ή να μεταβεί στην αμέσως επόμενη, δυσμενέστερη φάση. Οι πιθανότητες μετάβασης εξαρτώνται από το σύνολο των μεταβλητών που χρησιμοποιούνται για τη βαθμονόμηση του μοντέλου (Le Gat, 2008). Το αποτέλεσμα είναι για κάθε αγωγό ένα διάνυσμα πιθανοτήτων $\boldsymbol{p} = (p_1(t), p_2(t),\ldots, p_n(t))$ για κάθε χρονική στιγμή *t*, όπου *n* είναι ο αριθμός των κλάσεων που προσδιορίστηκε από τον αλγόριθμο ταξινόμησης. Το παραπάνω διάνυσμα αντιπροσωπεύει την πιθανότητα $p_n(t)$ ο αγωγός να βρίσκεται στην κατάσταση *n* τη χρονική στιγμή *t*. O Le Gat (2008) προτείνει τη σύνθεση των πιθανοτήτων σε ένα δείκτη πρόγνωσης, που μπορεί να χρησιμοποιηθεί για την κατάταξη των αγωγών σύμφωνα με την αναμενόμενη κατάστασή τους. Ο δείκτης πρόγνωσης (prediction index-*PI*) υπολογίζεται σαν το βαθμωτό γινόμενο του διανύσματος πιθανοτήτων και ενός διανύσματος βαρύτητας που επιλέγεται από το χρήστη:

$$PI = \boldsymbol{p}^{\boldsymbol{T}}\boldsymbol{\kappa}, \tag{1}$$

Όπου $\boldsymbol{p}^{\boldsymbol{T}}$ είναι το ανεστραμμένο διάνυσμα πιθανοτήτων (π.χ. ($p_1(t)$, $p_2(t)$, $p_3(t)$, $p_4(t)$)) και $\boldsymbol{\kappa}$ ένα διάνυσμα βαρύτητας, το οποίο περιέχει ένα βάρος για κάθε πιθανότητα (π.χ. (0, 1, 2, 3)). Καθώς μεγαλύτερα βάρη (2, 3) απονέμονται στις πιθανότητες να είναι στις δυσμενέστερες καταστάσεις ($p_3(t)$, $p_4(t)$), ο δείκτης πρόγνωσης αυξάνει, καθώς η κατάσταση του αγωγού επιδεινώνεται. Ο δείκτης πρόγνωσης είναι μια πολύ χρήσιμη ένδειξη της κατάστασης των αγωγών, καθώς επιτρέπει την κατάταξή τους και άρα την ιεράρχηση των μέτρων αποκατάστασης.

3 Μεθοδολογία αξιολόγησης

Λίγες μελέτες επιδίωξαν να αξιολογήσουν την αξιοπιστία της πρόγνωσης των μοντέλων τρωτότητας (Tran, 2007· Chughtai και Zayed, 2008· Le Gat, 2008· Ana, 2009· Salman, 2010· Khan *et al.,* 2010· Ens, 2012· Werey *et al.,* 2012). Τα αποτελέσματά τους δεν είναι άμεσα συγκρίσιμα καθώς (α) τα δεδομένα που χρησιμοποιήθηκαν για τη βαθμονόμηση του μοντέλου παρουσιάζουν σημαντικές διαφορές σε κάθε περίπτωση (ποσοστό του δικτύου που επιθεωρήθηκε, είδος των μεταβλητών που χρησιμοποιήθηκαν) και (β) οι μέθοδοι οι οποίες χρησιμοποιήθηκαν για την αξιολόγηση των αποτελεσμάτων διαφέρουν (Kley και Caradot, 2013). Ενώ κάποιες μελέτες (Ana, 2009) βρήκαν τα μοντέλα Markov ανεπαρκή, άλλες (Tran, 2007· Le Gat, 2008) απέδειξαν την αποτελεσματικότητά τους στην εξεύρεση των αγωγών που βρίσκονται σε κρίσιμη κατάσταση.

Η διαδικασία επικύρωσης των αποτελεσμάτων έχει ως στόχο να διερευνήσει αν το μοντέλο έχει ένα ικανοποιητικό εύρος αξιοπιστίας, σύμφωνο με το πεδίο εφαρμογής του (Schlesinger *et al.*, 1979· Sargent, 1999). Η αξιοπιστία ενός μαθηματικού μοντέλου αξιολογείται συγκρίνοντας τα αποτελέσματα της πρόγνωσης με πραγματικά δεδομένα, τα οποία έχουν προκύψει από παρατηρήσεις. Τα δεδομένα πάνω στα οποία θα γίνει η πρόγνωση θα πρέπει να είναι τελείως ανεξάρτητα και επομένως δε μπορούν να χρησιμοποιηθούν για την παραμετροποίησή του. Αυτό επιτυγχάνεται χωρίζοντας τα διαθέσιμα δεδομένα σε δύο σύνολα, τα οποία χρησιμοποιούνται για τη βαθμονόμηση και την επικύρωση του μοντέλου (περίπου 70-80% και 30-20% αντίστοιχα). Για τις ανάγκες της παρούσας εργασίας, τα υποσύνολα θα καλούνται υποσύνολο βαθμονόμησης και υποσύνολο ελέγχου αντίστοιχα.

Τα βραχυπρόθεσμα προγράμματα αποκατάστασης στοχεύουν στην επιδιόρθωση και αντικατάσταση των αγωγών που βρίσκονται στη δυσμενέστερη κατάσταση. Επομένως το μοντέλο αξιολογήθηκε σύμφωνα με την ικανότητά του να εντοπίζει τους αγωγούς με τις μεγαλύτερες φθορές. Η ποιότητα της πρόγνωσης, σε αυτή την περίπτωση, δεν εξαρτάται από το σύνολο των αγωγών για τους οποίους αποδείχθηκε ακριβής, αλλά από το ποσοστό των αγωγών σε κρίσιμη κατάσταση που το μοντέλο κατάφερε να εντοπίσει.

Μία χαρακτηριστική εφαρμογή του προβλήματος αποτελεί η προσομοίωση μέρους ενός δικτύου για το οποίο δεν έχουν ληφθεί δεδομένα. Σε αυτή την περίπτωση, οι αρμόδιοι φορείς θα μπορούσαν να χρησιμοποιήσουν τα αποτελέσματα ώστε να εστιάσουν ενδεχόμενες μελλοντικές επιθεωρήσεις στους αγωγούς οι οποίοι προγνώστηκαν σε πιο επιδεινωμένη κατάσταση. Δύο μεθοδολογίες (καμπύλες απόδοσης 1 και 2) χρησιμοποιήθηκαν για τη διερεύνηση της αποδοτικότητας του μοντέλου με βάση το σκοπό αυτό.

3.1 Καμπύλη απόδοσης 1

Ο Le Gat (2008) κατασκεύασε μια καμπύλη ώστε να προσδιορίσει τί ποσοστό από τους αγωγούς οι οποίοι βρέθηκαν σε δυσμενή κατάσταση εντοπίστηκε από το μοντέλο GompitZ, σε κάθε υποσύνολο αγωγών με τους χειρότερους δείκτες πρόγνωσης.

Η καμπύλη αυτή κατασκευάζεται όπως παρακάτω:

- Βήμα 1: το μοντέλο βαθμονομείται χρησιμοποιώντας το αντίστοιχο υποσύνολο δεδομένων και εν συνεχεία χρησιμοποιείται για την πρόγνωση της κατάστασης των αγωγών του υποσυνόλου ελέγχου (οι οποίοι δε χρησιμοποιήθηκαν για τη βαθμονόμησή του), τη χρονιά που επιθεωρήθηκαν.
- Βήμα 2: οι αγωγοί κατατάσσονται με βάση το δείκτη πρόγνωσής τους από αυτόν με τις περισσότερες σε αυτόν με τις λιγότερες φθορές. Ο αριθμός κατάταξης κάθε αγωγού διαιρείται στη συνέχεια με το συνολικό αριθμό των αγωγών που βρίσκονται στο υποσύνολο ελέγχου, με αποτέλεσμα μια χαρακτηριστική τιμή σχετικής κατάταξης για κάθε αγωγό.
- Βήμα 3: για κάθε κατάταξη, οι πραγματικές κλάσεις των αγωγών, όπως αυτές καταγράφηκαν κατά τη διάρκεια επιθεωρήσεων με οπτικά μέσα, αναλύονται: το σύνολο των αγωγών που βρίσκονται στη δυσμενέστερη κλάση (σε αυτή την περίπτωση κλάση 4) και εντοπίστηκαν από το μοντέλο, διαιρείται με το σύνολο των αγωγών του υποσυνόλου ελέγχου, οι οποίοι κατατάχθηκαν στην κλάση 4 κατά την επιθεώρησή τους. Το αποτέλεσμα αντιπροσωπεύει το ποσοστό των αγωγών κλάσης 4 (μεταξύ όλων των αγωγών που ανήκουν στη συγκεκριμένη κλάση), οι οποίοι εντοπίστηκαν από το μοντέλο στο αντίστοιχο ποσοστό που εξετάζεται.
- Βήμα 4: το σχετικό ποσοστό των αγωγών που ανήκουν στην κλάση 4 (Βήμα 3) χαράσσεται ως προς τη σχετική κατάταξη του αγωγού για την προγνωθείσα κατάστασή του (Βήμα 2).

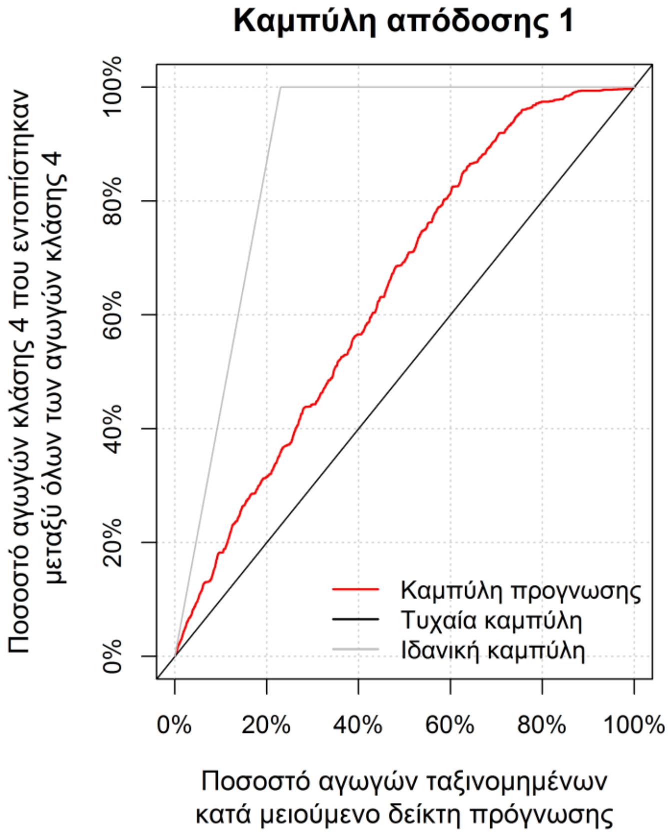
Τελικά, η ποιότητα πρόγνωσης του μοντέλου μπορεί να αξιολογηθεί από την καμπύλη απόδοσης 1, σε σχέση με την αντίστοιχη ιδανική (βέλτιστο σενάριο), και τυχαία (χείριστο σενάριο) καμπύλη (Σχήμα 1). Η ιδανική καμπύλη αναπαριστά τη μορφή της καμπύλης απόδοσης 1, για την περίπτωση μιας απόλυτα ακριβούς πρόβλεψης, ενώ η τυχαία καμπύλη αντιπροσωπεύει το αποτέλεσμα για μια τυχαία κατάταξη των αγωγών, δηλαδή στην περίπτωση ενός μοντέλου χωρίς καμία προγνωστική αξία. Όσο η καμπύλη απομακρύνεται από την τυχαία της μορφή (χ=ψ), τόσο πιο αποτελεσματικό είναι το μοντέλο. Αντίστοιχα, όσο πιο κοντά βρίσκεται στην ιδανική καμπύλη, τόσο καλύτερη η πρόγνωση.

Στο παράδειγμα που παρουσιάζεται στο Σχήμα 1, το 23% των αγωγών του υποσυνόλου ελέγχου ανήκουν στην κλάση 4. Άρα ένα απόλυτα ακριβές μοντέλο θα εντόπιζε όλους τους αγωγούς που βρίσκονται στην κλάση 4, μεταξύ του 23% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση (Σχήμα 1, ιδανική καμπύλη). Από την άλλη, ένα μοντέλο χωρίς καμία προγνωστική αξία θα εντόπιζε (λόγω τυχαιότητας) περίπου το 23% όλων των αγωγών κλάσης 4, μεταξύ του 23% των αγωγών οι οποίοι προγνώστηκαν στη δυσμενέστερη κατάσταση (Σχήμα 1, τυχαία καμπύλη).

Εφόσον το ποσοστό των αγωγών που βρίσκονται στη δυσμενέστερη κλάση καθορίζει τη μορφή της καμπύλης, συμπεράσματα μπορούν να εξαχθούν μόνο συγκρίνοντας την παραχθείσα καμπύλη με την ιδανική για κάθε περίπτωση.

Αν και η καμπύλη απόδοσης 1 αποτελεί μια ισχυρή ένδειξη της απόδοσης του μοντέλου, στερείται πρακτικής αξίας. Στις περισσότερες περιπτώσεις, οι αρμόδιοι φορείς έχουν περιορισμένους πόρους στη διάθεσή τους, για τη συντήρηση του αστικού δικτύου αποχέτευσης και άρα μπορούν να αποκαταστήσουν ένα συγκεκριμένο μόνο αριθμό ελαττωματικών αγωγών. Κατά συνέπεια, δεν ενδιαφέρονται απαραίτητα για τον εντοπισμό και την επισκευή κάθε φθαρμένου αγωγού, αλλά μόνο εκείνων που βρίσκονται σε κρισιμότερη κατάσταση.

Για την υποστήριξη του σχεδιασμού και της πραγματοποίησης επιθεωρήσεων στους αγωγούς με τις μεγαλύτερες φθορές, το μοντέλο θα πρέπει να μπορεί να εντοπίσει όσο το δυνατόν περισσότερους εξ'αυτών, μέσα σε ένα πολύ μικρό ποσοστό των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, δηλαδή εκείνων που έλαβαν τους χειρότερους δείκτες πρόγνωσης. Μια δεύτερη μέθοδος αξιολόγησης αναπτύχθηκε με σκοπό την αξιοποίηση αυτού του κριτηρίου.

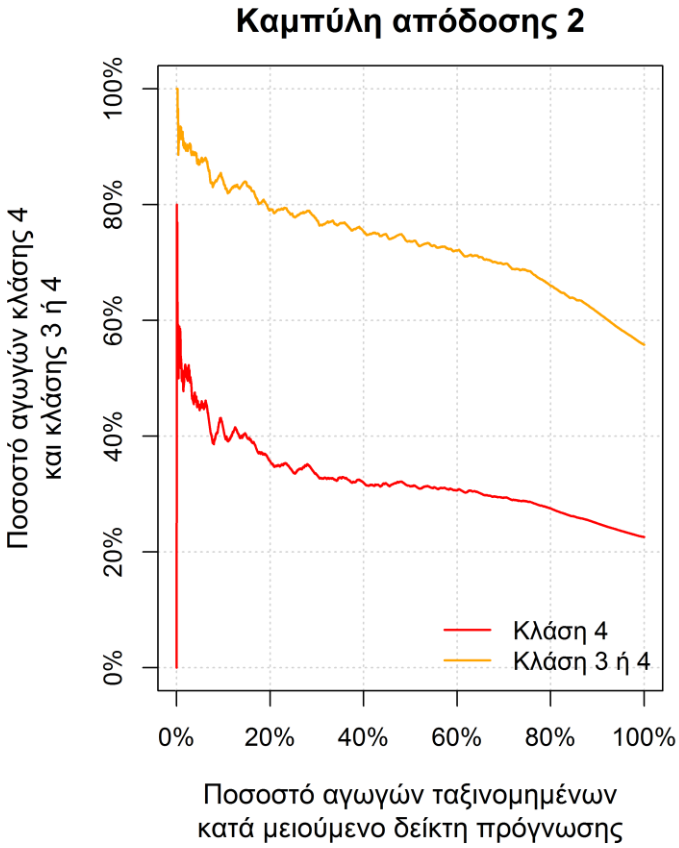


Σχήμα 1: Καμπύλη απόδοσης 1 για ένα υποσύνολο ελέγχου όπου το 23% των αγωγών βρίσκονται στη δυσμενέστερη κλάση (κλάση 4). Η καμπύλη πρόγνωσης ερμηνεύεται ως ακολούθως: μεταξύ του χ% των αγωγών οι οποίοι προγνώστηκαν στη δυσμενέστερη κατάσταση, εντοπίστηκε ψ% του συνόλου των αγωγών κλάσης 4. Η καμπύλη πρόγνωσης βρίσκεται μεταξύ της ιδανικής (βέλτιστο σενάριο) και της τυχαίας (χείριστο σενάριο) καμπύλης.

3.2 Καμπύλη απόδοσης 2

Η δεύτερη μέθοδος επιχειρεί να εντοπίσει τί ποσοστό από τους αγωγούς που προγνώστηκαν στη δυσμενέστερη κατάσταση, ήταν όντως σε κρίσιμη κατάσταση. Η κατασκευή της γίνεται με παρόμοιο τρόπο με την καμπύλη πρόγνωσης 1. Ωστόσο, το βήμα 3 διαφοροποιείται: για κάθε τιμή κατάταξης, ο αριθμός των αγωγών που βρέθηκαν στην κλάση 4, διαιρείται με το σύνολο των αγωγών που αντιστοιχούν στη συγκεκριμένη κατάταξη (αντί για το σύνολο των αγωγών κλάσης 4) (Σχήμα 2).

Η ίδια διαδικασία εφαρμόζεται για την κατασκευή της δεύτερης καμπύλης του γραφήματος, αλλά σε αυτή την περίπτωση η κλάση 3 λαμβάνεται υπόψη μαζί με την 4. Το αποτέλεσμα είναι μια καμπύλη, η οποία αναπαριστά την πραγματική κατάσταση των αγωγών οι οποίοι προγνώστηκαν από το μοντέλο στη δυσμενέστερη κατάσταση.



Σχήμα 2: Καμπύλη απόδοσης 2 για ένα υποσύνολο ελέγχου όπου το 23% των αγωγών βρίσκονται στη δυσμενέστερη κλάση (κλάση 4). Η καμπύλη πρόγνωσης ερμηνεύεται ως ακολούθως: το χ% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται από ψ_1% αγωγών κλάσης 4 (κόκκινη καμπύλη) και ψ_2% αγωγών κλάσης 3 ή 4 (πορτοκαλί καμπύλη).

Η απόδοση του μοντέλου μπορεί να εκτιμηθεί γραφικά, αναλύοντας την κλίση της καμπύλης. Στην περίπτωση ενός μοντέλου χωρίς καμία προγνωστική αξία (τυχαία επιλογή), το υποσύνολο που περιέχει χ% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, θα είχε περίπου την ίδια κατανομή κλάσεων με το υποσύνολο ελέγχου. Η καμπύλη που αναπαριστά μια τυχαία επιλογή αγωγών, θα ήταν οριζόντια. Από την άλλη, ένα τέλειο μοντέλο θα έβρισκε μόνο αγωγούς κλάσης 4, μεταξύ των αγωγών με τις δυσμενέστερες προβλέψεις (ψ=100%).

Θεωρώντας το 100% των αγωγών με τις δυσμενέστερες προβλέψεις (άρα όλους τους αγωγούς), η κατανομή κλάσεων που φαίνεται στο γράφημα είναι η κατανομή του υποσυνόλου ελέγχου, με 23% των αγωγών να βρίσκονται στην κλάση 4.

4 Επισκόπηση και πρωτογενής επεξεργασία της βάσης δεδομένων

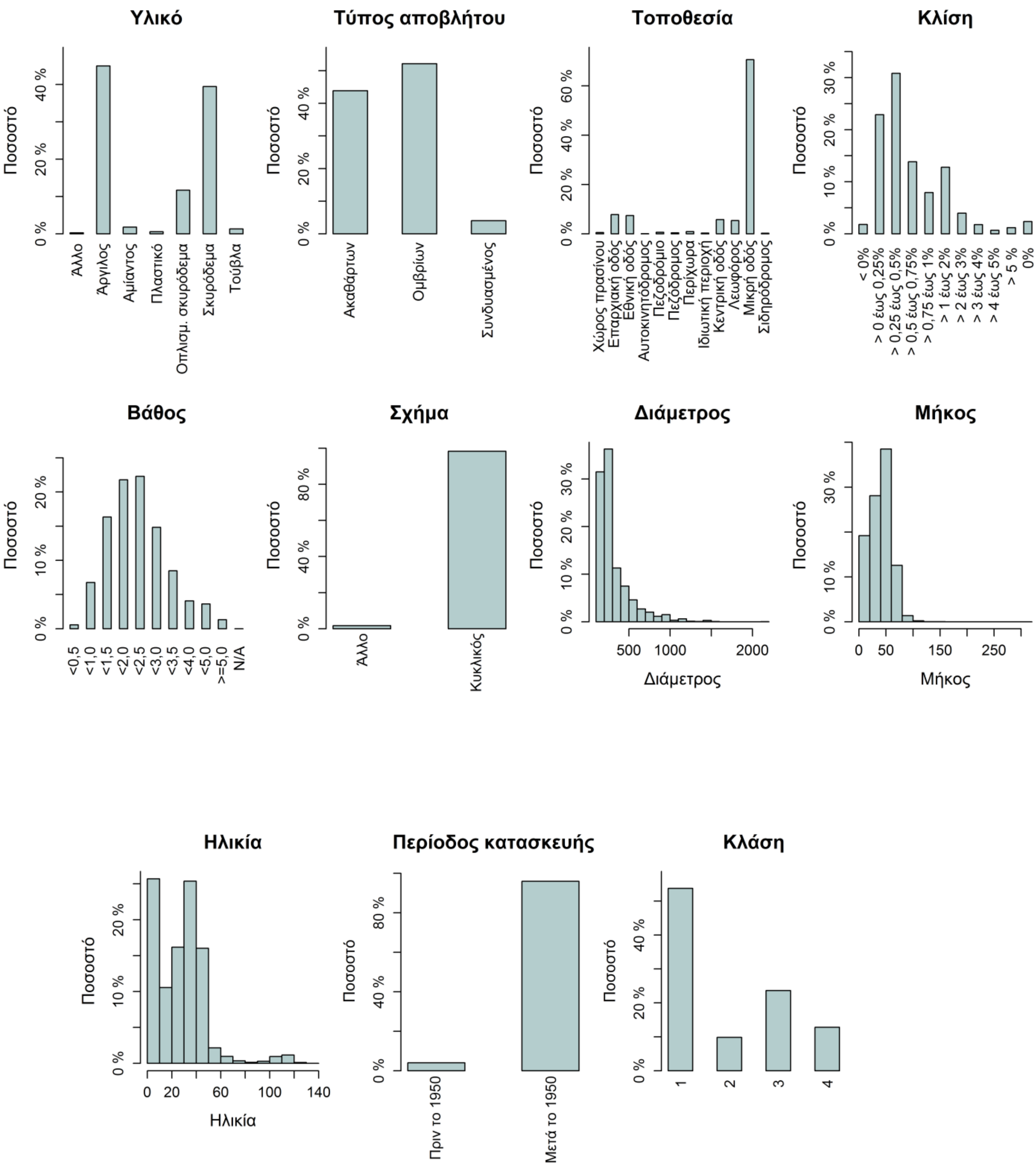
Το μοντέλο δοκιμάστηκε και αξιολογήθηκε χρησιμοποιώντας μία βάση δεδομένων η οποία περιέχει στοιχεία οπτικών επιθεωρήσεων αλλά και κατασκευαστικών χαρακτηριστικών του δικτύου αποστράγγισης μιας Γερμανικής πόλης. Αυτή η πόλη αποτέλεσε ιδανικό δείγμα μελέτης, καθώς ολόκληρο το αποχετευτικό σύστημα (43.000 αγωγοί/ 1.700 χιλιόμετρα) έχει επιθεωρηθεί τουλάχιστον μία φορά και πάνω από το 50% έχει επιθεωρηθεί δύο ή περισσότερες.

4.1 Επισκόπηση των δεδομένων

Μετά την επεξεργασία των διαθέσιμων δεδομένων, οι δοκιμές του μοντέλου εκτελέστηκαν σε ένα σύνολο 31.394 καταγραφών. Τα αποτελέσματα των επιθεωρήσεων αξιολογήθηκαν σύμφωνα με ένα σύστημα κατάταξης με βάση την προτεραιότητά τους, ή με άλλα λόγια με βάση το κατεπείγον της αναγκαιότητάς τους για μέτρα αποκατάστασης (priority-based).

Η μεθοδολογία που χρησιμοποιήθηκε στην παρούσα εργασία αναπτύχθηκε στα πλαίσια του γαλλικού προγράμματος RERAU (Le Gauffre *et al.,* 2004) και αποδίδει ένα σκορ από 1 έως 4, με το 4 να αντιπροσωπεύει τη δυσμενέστερη κατάσταση (άμεση αποκατάσταση είναι απαραίτητη). Σύμφωνα με τον οδηγό του RERAU (Le Gauffre *et al.,* 2004, 2007), δέκα τύποι δυσλειτουργιών (π.χ. διείσδυση υδάτων, έμφραξη), οι οποίοι προέρχονται από βλάβες, καταγράφονται, καταλήγοντας σε οχτώ επιπτώσεις (π.χ. μόλυνση των επιφανειακών υδάτων, διακοπή των επιφανειακών δραστηριοτήτων). Η ανάγκη αποκατάστασης αξιολογείται με βάση κριτήρια, τα οποία καθορίζουν τη σχέση μεταξύ των δυσλειτουργιών οι οποίες καταγράφονται σε συγκεκριμένα τμήματα του αγωγού, και των επιπτώσεων στις οποίες μπορεί να καταλήξουν, αν δεν αντιμετωπιστούν. Η παραπάνω μεθοδολογία κατάταξης λαμβάνει υπόψη μόνο τη δομική κατάσταση του αγωγού.

Τα βασικά χαρακτηριστικά των αγωγών που περιέχονται στη βάση δεδομένων φαίνονται στο Σχήμα 3.

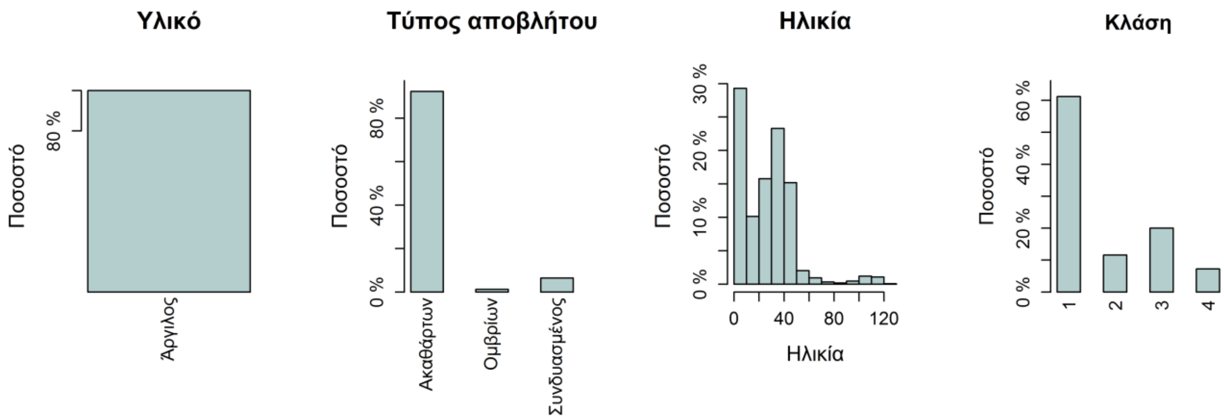


Σχήμα 3: Κατανομή των χαρακτηριστικών για το σύνολο των αγωγών.

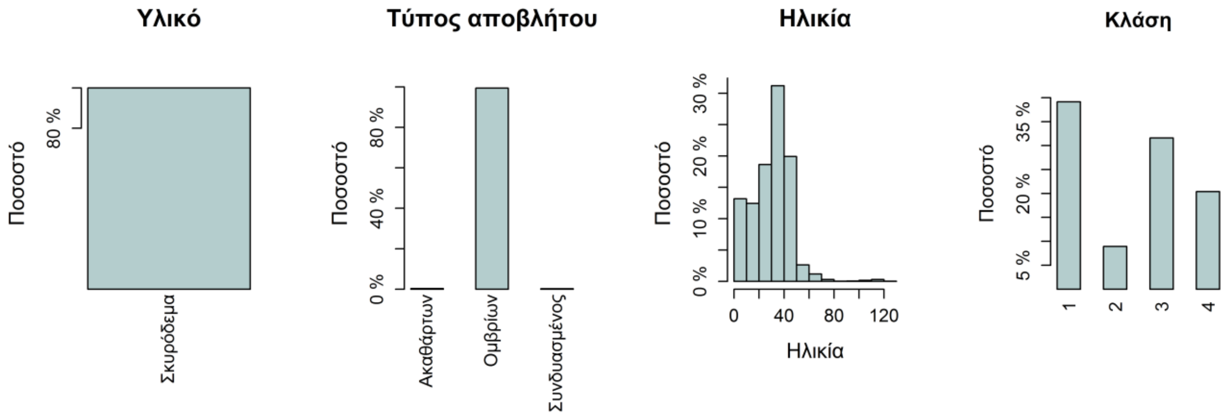
Η πλειονότητά τους έχει κατασκευαστεί από άργιλο ή σκυρόδεμα και εξυπηρετεί είτε ανάγκες αποχέτευσης ακαθάρτων είτε στράγγισης ομβρίων. Καθώς οι περισσότεροι αγωγοί (71%) είναι κατασκευασμένοι κάτω από μικρές οδούς, δεν υποβάλλονται σε μεγάλες πιέσεις από την υπερκείμενη κίνηση. Οι κλίσεις τους φαίνονται να είναι ως επί το πλείστον ήπιες, με το 76% να έχουν κλίσεις μικρότερες του 1% και το 20% περίπου να έχουν κλίσεις μεταξύ 1% και 5%. Από τους αγωγούς, περίπου οι μισοί είναι εγκατεστημένοι σε μεσαίο βάθος (2,5-5m), ενώ το 46% βρίσκονται θαμμένοι σε βάθος μικρότερο των δύο μέτρων. Η βάση δεδομένων αποτελείται κυρίως από νέους, κυκλικούς αγωγούς, οι περισσότεροι από τους οποίους (96%) έχουν κατασκευασθεί μετά το 1950. Η πλειοψηφία τους έχει διάμετρο μέχρι 500mm και μήκος μέχρι 80m, με το 40% περίπου να έχουν μήκος μεταξύ 40 και 60 μέτρων. Οι περισσότεροι αγωγοί είχαν ηλικία μέχρι 50 ετών, κατά την επιθεώρησή τους. Σχεδόν το 60% αυτών βρέθηκαν σε ικανοποιητική κατάσταση (κλάσεις 1 και 2), ενώ περισσότερο από το 10% είχαν άμεση ανάγκη μέτρων αποκατάστασης (κλάση 4).

Όπως είναι αναμενόμενο, το υλικό και το είδος του αποβλήτου παρουσιάζουν ισχυρή αλληλεξάρτηση. Σχεδόν το 90% των αργιλικών αγωγών είναι αγωγοί αποχέτευσης ακαθάρτων, με ένα μικρότερο ποσοστό να χρησιμοποιούνται παράλληλα και ως αγωγοί ομβρίων. Το σκυρόδεμα από την άλλη μεριά, χρησιμοποιείται σχεδόν αποκλειστικά σε αγωγούς αποχέτευσης ομβρίων. Η ισχυρή συσχέτιση του υλικού του αγωγού και του τύπου αποβλήτου που εξυπηρετεί, καθώς και οι αντίστοιχες κατανομές της ηλικίας και των κλάσεων για κάθε περίπτωση, φαίνονται στα σχήματα 4 και 5.

Ένα μεγάλο ποσοστό αργιλικών αγωγών (σχεδόν το 30%) έχουν κατασκευαστεί την τελευταία δεκαετία, ενώ υπάρχει και ένα πολύ μικρό ποσοστό αγωγών πολύ μεγάλης ηλικίας (100-120 ετών). Η πλειοψηφία των αγωγών σκυροδέματος από την άλλη, φαίνεται να είναι μεγαλύτερης ηλικίας, με το 70% να είναι μεταξύ 30 και 50 ετών, ηλικίες που εκπροσωπούν το 54% αντίστοιχα των αργιλικών αγωγών. Το μεγάλο ποσοστό πολύ νέων αργιλικών αγωγών (<10 ετών) ίσως ευθύνεται και για την επιβαρυμένη κατάσταση των αγωγών σκυροδέματος σε σχέση με τους αργιλικούς· περισσότερο από το 50% αυτών βρέθηκαν σε κλάση 3 ή 4 και άρα απαιτούν βραχυπρόθεσμη αποκατάσταση. Η πλειοψηφία των αργιλικών αγωγών από την άλλη (70%) βρέθηκαν σε ικανοποιητική κατάσταση (κλάση 1 ή 2).



Σχήμα 4: Κατανομή χαρακτηριστικών στους αργιλικούς αγωγούς.



Σχήμα 5: Κατανομή χαρακτηριστικών στους αγωγούς σκυροδέματος.

4.2 Επεξεργασία των δεδομένων

Για τη διευκόλυνση του χειρισμού των δεδομένων και την πραγματοποίηση της αξιολόγησης του μοντέλου, τα δεδομένα καταγράφηκαν σε ένα υπολογιστικό φύλλο και έγιναν διαθέσιμα στη γλώσσα προγραμματισμού R.

Όπως αναφέρθηκε προηγουμένως, κάποια χαρακτηριστικά αντιστοιχούν σε πολύ μικρούς πληθυσμούς αγωγών (π.χ. σχήμα, περίοδος κατασκευής). Τα δεδομένα αυτά δε συγκροτούν αντιπροσωπευτικές ομάδες του πληθυσμού, γεγονός που καθιστά προβληματικό το χειρισμό τους, την αξιοπιστία τους σε στατιστικές δοκιμές, και την ερμηνεία των αποτελεσμάτων. Για αυτούς τους λόγους, οι παρακάτω αγωγοί εξαιρέθηκαν από οποιαδήποτε περαιτέρω ανάλυση:

- Αγωγοί από αμίαντο, τούβλα, οπλισμένο σκυρόδεμα, πλαστικό και άλλα. Μόνο αγωγοί από άργιλο και σκυρόδεμα συμπεριλήφθηκαν στα δεδομένα.
- Μη κυκλικοί αγωγοί

Τελικά, η ποιότητα πρόγνωσης αξιολογήθηκε με βάση ένα δείγμα 31.394 καταγραφών.

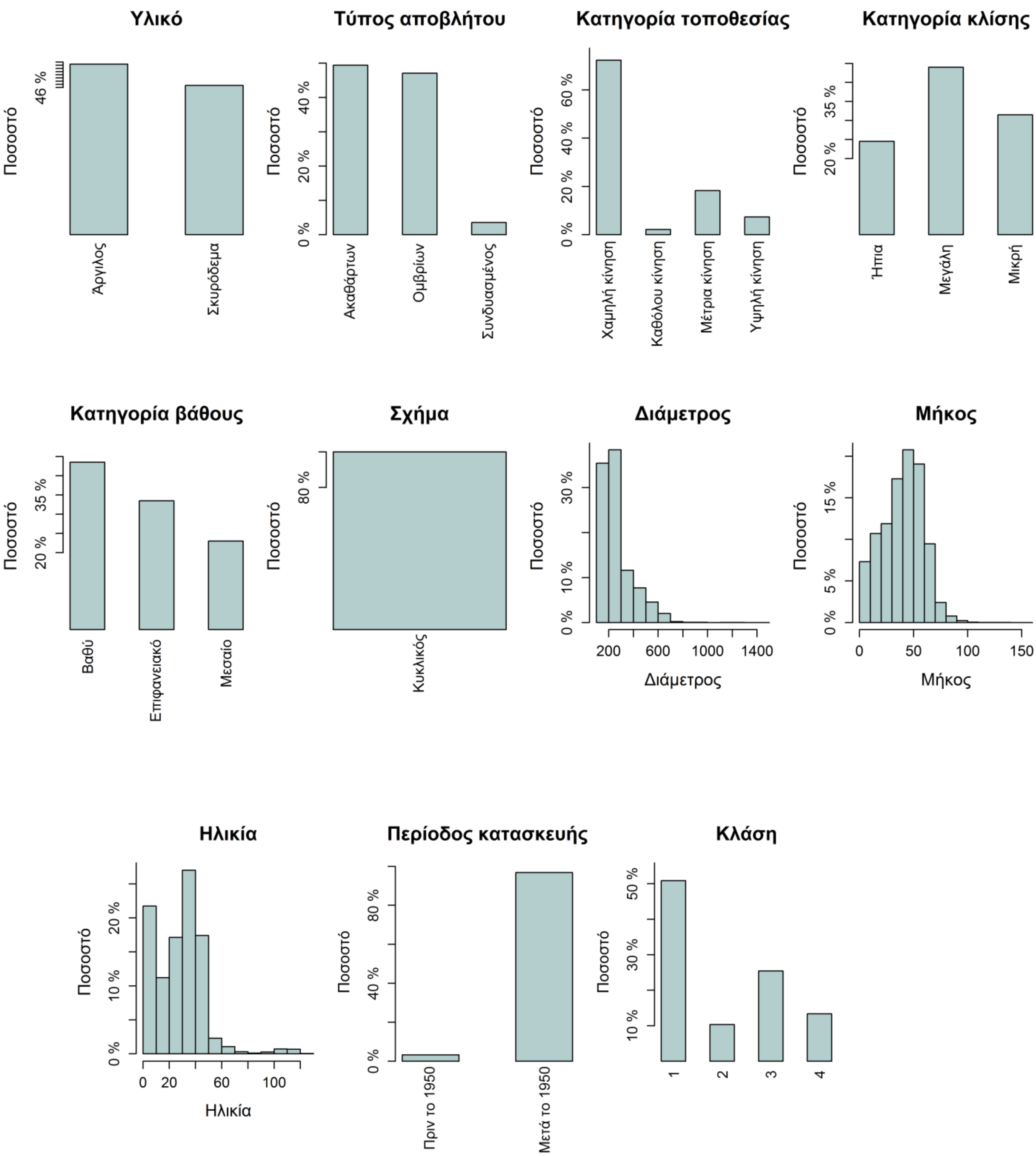
Οι αγωγοί οι οποίοι κατασκευάστηκαν πριν το 1950 περιλήφθηκαν στα δεδομένα αλλά εξαιρέθηκαν από όλες τις στατιστικές δοκιμές, και τις δοκιμές του μοντέλου, καθώς ο πληθυσμός τους (<5%) δε θεωρείται στατιστικά σημαντικός.

Ομαδοποίηση των δεδομένων

Για τη διευκόλυνση του χειρισμού των δεδομένων και την ελαχιστοποίηση του χρόνου επεξεργασίας από την R, εφαρμόστηκε μια περαιτέρω ομαδοποίηση ορισμένων χαρακτηριστικών (Σχήμα 6):

- Κλίση: τρείς νέες κατηγορίες δημιουργήθηκαν, ήπιες κλίσεις (<0%, 0% έως 0,25%), μεσαίες κλίσεις (>0,25% έως 0,5%) και απότομες κλίσεις (>0,5%).
- Βάθος εγκατάστασης: παρομοίως, τρεις νέες κατηγορίες δημιουργήθηκαν για το βάθος εγκατάστασης, επιφανειακό (<2,0 μ), μεσαίο (2 έως 2,5 μ) και βαθύ (>2,5 μ).
- Τοποθεσία: λόγω του μεγάλου εύρους τοποθεσιών που περιλαμβάνουν ένα αρκετά μικρό δείγμα τιμών, έγινε μια περαιτέρω κατηγοριοποίηση σε τέσσερις νέες ομάδες, σύμφωνα με το επίπεδο της κυκλοφοριακής έντασης της υπερκείμενης οδού: υψηλή κίνηση (αυτοκινητόδρομοι, λεωφόροι, σιδηρόδρομοι, εθνικοί οδοί), μέτρια κίνηση (κεντρικοί οδοί), χαμηλή κίνηση (μικροί οδοί, επαρχιακές οδοί, περίχωρα) και καθόλου κίνηση (πεζόδρομοι, χώροι πρασίνου, ιδιωτικοί χώροι, πεζοδρόμια).

Η τελική κατανομή όλων των χαρακτηριστικών φαίνεται στο Σχήμα 6.



Σχήμα 6: Κατανομή των χαρακτηριστικών των αγωγών μετά την ομαδοποίηση των δεδομένων.

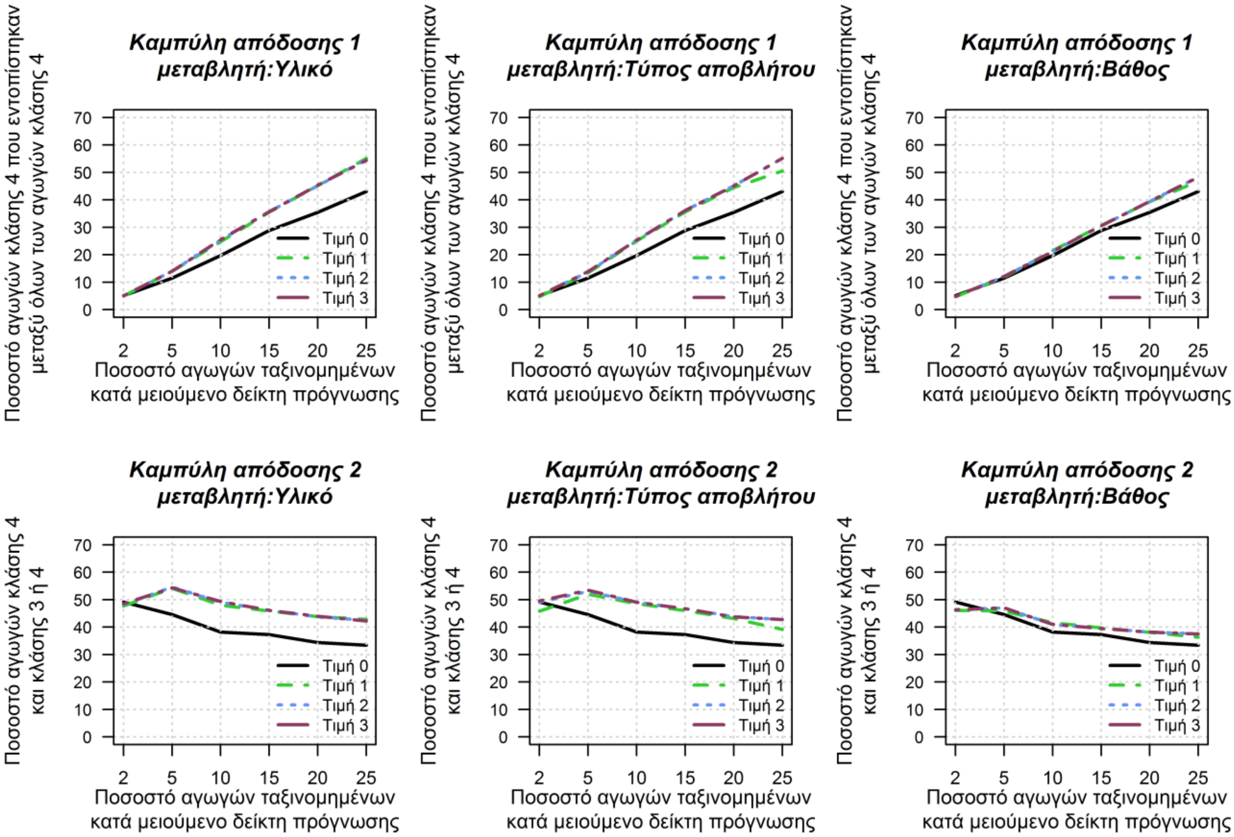
5 Παραμετροποίηση και εφαρμογή του μοντέλου

5.1 Εντοπισμός των κατάλληλων μεταβλητών και των αντίστοιχων τιμών τους

Για την αξιολόγηση της επιρροής κάθε μεταβλητής στη διεργασία φθοράς, αναπτύχθηκε μια μεθοδολογία βασισμένη στην τελική ποιότητα πρόγνωσης. Τα διαθέσιμα δεδομένα χωρίστηκαν ακόμη μια φορά χρονολογικά σε δυο υποσύνολα, αφιερωμένα στη βαθμονόμηση και την επικύρωση του μοντέλου (70% και 30% αντίστοιχα).

Τα κατασκευαστικά χαρακτηριστικά των αγωγών, όπως αυτά περιγράφηκαν παραπάνω (το υλικό, το είδος του αποβλήτου, η τοποθεσία, η κλίση, το βάθος, η διάμετρος, και το μήκος) ενσωματώθηκαν ως μεταβλητές στο μοντέλο, καθώς υποτέθηκε ότι μπορούν να επηρεάσουν το χρόνο ζωής τους. Σε κάθε μία από τις επτά μεταβλητές ανατέθηκε μία τιμή από 0 έως 3, υποδεικνύοντας τη μορφή συνεισφοράς της μεταβλητής στη φθορά του αγωγού (καμία επιρροή, επιρροή μόνο στη φάση κατασκευής, επιρροή μόνο στη φάση λειτουργίας, ή επιρροή και στις δύο φάσεις, αντίστοιχα). Για κάθε μία από τις τέσσερις θεωρήσεις των μεταβλητών στο μοντέλο, οι δύο καμπύλες απόδοσης σχεδιάστηκαν. Στη συνέχεια, οι τιμές οι οποίες αντιστοιχούν στο 2%, 5%, 10%, 15%, 20%, και 25% των αγωγών οι οποίοι προγνώστηκαν στην κρισιμότερη κατάσταση, εξήχθησαν από τις δύο καμπύλες απόδοσης και σχεδιάσθηκαν σε ένα κοινό γράφημα για κάθε μία από τις τέσσερις θεωρήσεις των μεταβλητών. Οι τέσσερις καμπύλες, μία για κάθε τιμή της μεταβλητής (0 έως 3), και οι αντίστοιχες τιμές τους φαίνονται στα Σχήματα 7 και 8.

Για κάθε μεταβλητή, η καμπύλη για την οποία μεγιστοποιείται η απόδοση του μοντέλου, υποδεικνύει τη σωστή θεώρηση της μεταβλητής. Αν όλες οι καμπύλες βρίσκονται κοντά η μία στην άλλη, η μεταβλητή έχει πολύ μικρή ή καθόλου επιρροή. Από την άλλη, αν τουλάχιστον μία από τις καμπύλες δείχνει μία ξεκάθαρη αύξηση της προγνωστικής ποιότητας του μοντέλου (βρίσκεται ψηλότερα από τις υπόλοιπες), η μεταβλητή θα εφαρμοσθεί στο μοντέλο με την αντίστοιχη τιμή. Σε περίπτωση που αυτή η καμπύλη αντιστοιχεί στην τιμή 0, η μεταβλητή δεν περιλαμβάνεται στο μοντέλο, καθώς η ενσωμάτωσή της δε βελτιώνει την ποιότητα πρόγνωσης.

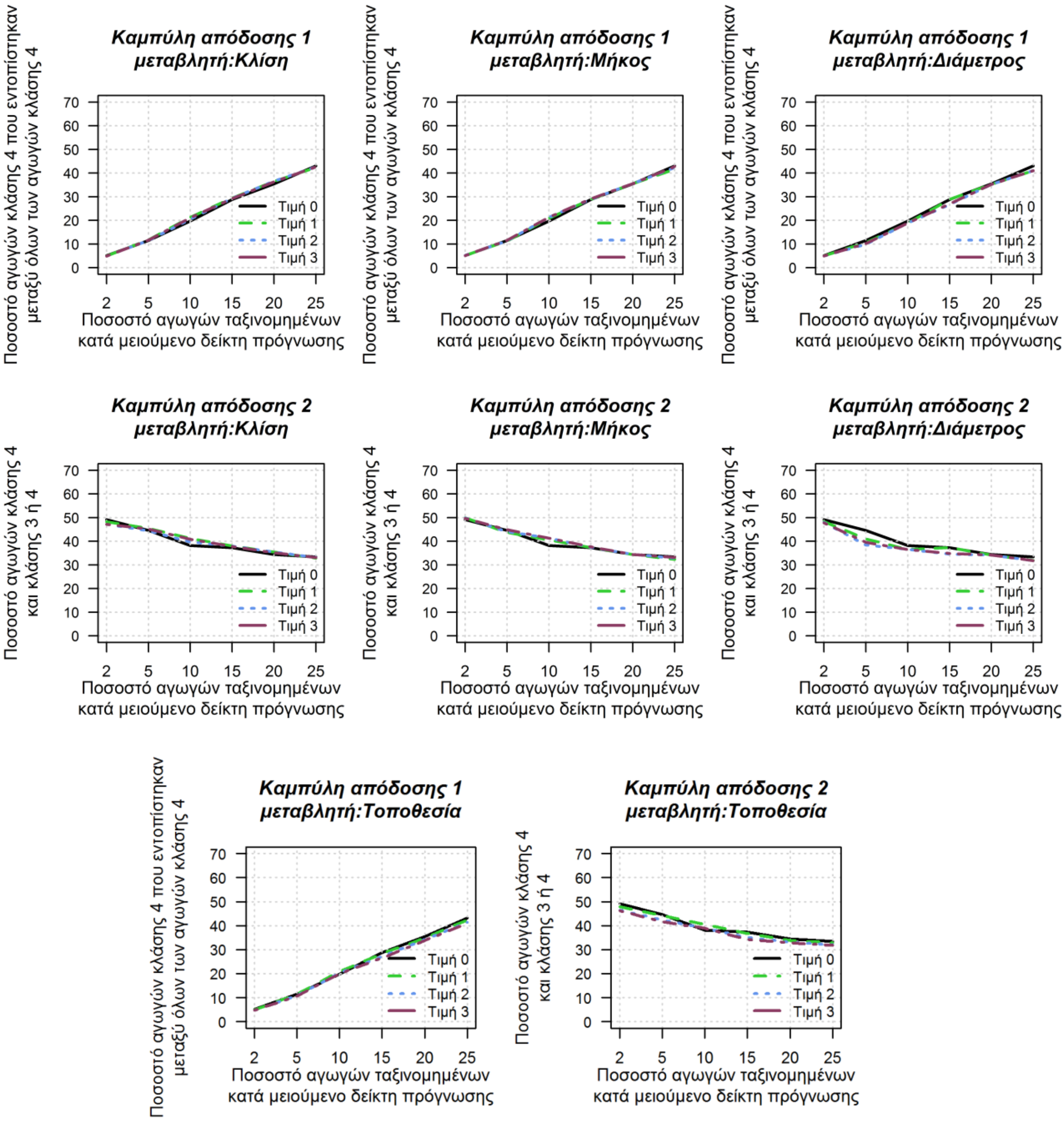


Σχήμα 7: Μεταβλητές που επηρεάζουν την ποιότητα πρόγνωσης.

Τα αποτελέσματα υποδηλώνουν ότι το υλικό του αγωγού και το είδος του αποβλήτου είναι οι μεταβλητές με τη μεγαλύτερη επιρροή (Σχήμα 7), ενώ κανένας άλλος παράγοντας δε φαίνεται να έχει σημαντική επίδραση (Σχήμα 8). Για αυτές τις δύο μεταβλητές, η ποιότητα πρόγνωσης βελτιώνεται σημαντικά με την ενσωμάτωσή τους στο μοντέλο, ενώ είναι σχεδόν ασήμαντο αν θα εφαρμοστούν με την τιμή 1, 2 ή 3. Χωρίς τη θεώρηση του υλικού σαν μεταβλητή (τιμή 0), το 42% όλων των αγωγών κλάσης 4 εντοπίζεται μεταξύ του 25% των αγωγών οι οποίοι προγνώστηκαν στη δυσμενέστερη κατάσταση (Σχήμα 7, καμπύλη απόδοσης 1, μεταβλητή: υλικό). Όταν το υλικό ενσωματώνεται στο μοντέλο (με τιμή 1, 2 ή 3), ο αριθμός αυτός αυξάνει στο 54%.

Ενώ το βάθος φαίνεται να έχει ένα μικρό θετικό αντίκτυπο για τα μεγαλύτερα ποσοστά (25%), δεν είναι ακόμα σαφές αν μπορεί όντως να προσφέρει στο τελικό αποτέλεσμα, σε συνδυασμό με άλλες μεταβλητές (Σχήμα 7).

Η θεώρηση των υπολοίπων παραγόντων (κλίση, μήκος, διάμετρος, τοποθεσία), επίσης δε φαίνεται να βελτιώνει την απόδοση του μοντέλου· στις περισσότερες περιπτώσεις, οι καμπύλες πέφτουν ακριβώς η μία πάνω στην άλλη, ενώ στην περίπτωση της διαμέτρου, η ενσωμάτωσή της προκαλεί ακόμη και μείωση της απόδοσης του μοντέλου, αφού η καμπύλη που αντιστοιχεί στην τιμή μηδέν βρίσκεται ψηλότερα (Σχήμα 8).



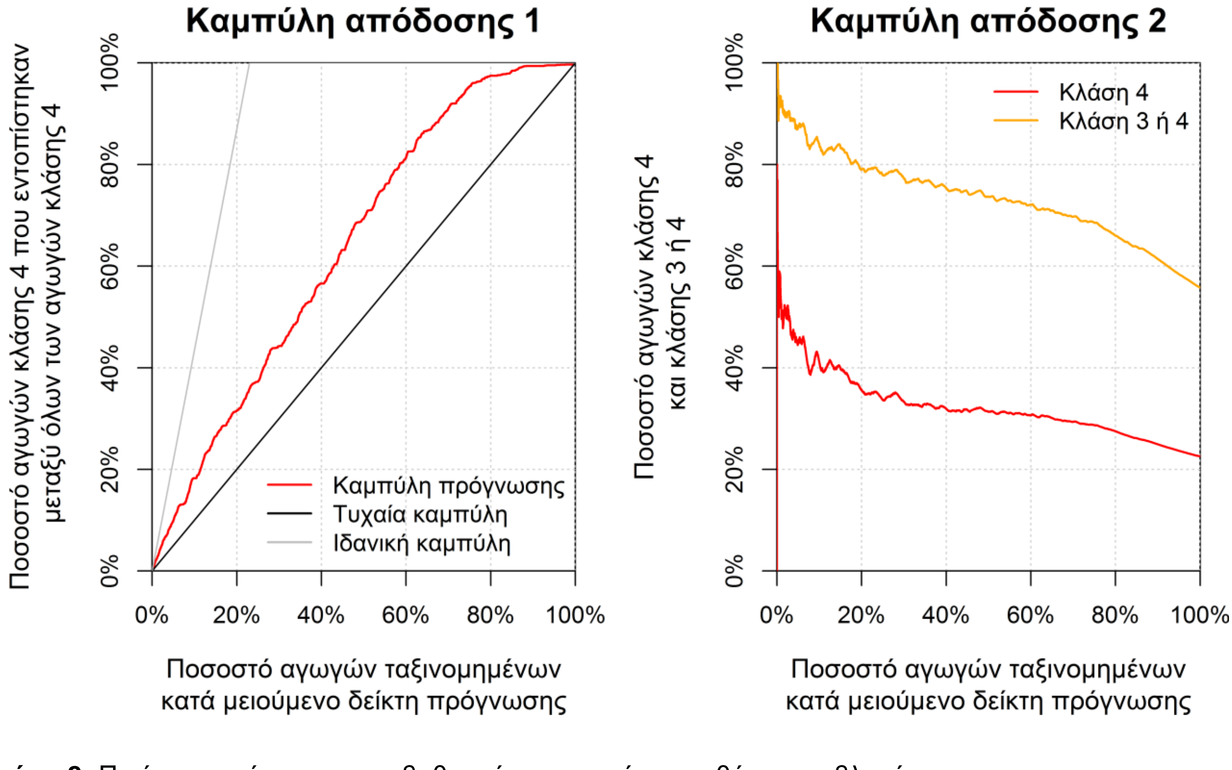
Σχήμα 8: Μεταβλητές που δεν επηρεάζουν την ποιότητα πρόγνωσης.

5.2 Προσομοίωση με χρονολογική επιλογή υποσυνόλων βαθμονόμησης και ελέγχου

5.2.1 Χωρίς μεταβλητές

Η ποιότητα της πρόβλεψης αξιολογήθηκε αρχικά χωρίς μεταβλητές (age model). Σε αυτή την περίπτωση ο μόνος παράγοντας που λαμβάνεται υπόψη κατά την προσομοίωση της κατάστασης των αγωγών είναι η ηλικία τους στο χρόνο για τον οποίο γίνεται η πρόγνωση.

Η βάση δεδομένων χωρίστηκε χρονολογικά σε ένα υποσύνολο βαθμονόμησης και ένα υποσύνολο ελέγχου: το 70% των παλαιότερων καταγραφών χρησιμοποιήθηκε για τη βαθμονόμηση του μοντέλου (21.976 επιθεωρήσεις), ενώ το νεότερο 30% για την αξιολόγησή του (9.418 καταγραφές). Το Σχήμα 9 αναπαριστά την τελική ποιότητα πρόγνωσης, όπως αυτή προκύπτει από τις καμπύλες απόδοσης.



Σχήμα 9: Ποιότητα πρόγνωσης για βαθμονόμηση χωρίς προσθήκη μεταβλητών.

Μεταξύ του 20% των αγωγών οι οποίοι προγνώστηκαν στην κρισιμότερη κατάσταση, εντοπίστηκε το 31% του συνόλου των αγωγών κλάσης 4 (Σχήμα 9 – Καμπύλη απόδοσης 1). Επιπλέον, το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται κατά 36% από αγωγούς κλάσης 4 και κατά 79% από αγωγούς κλάσης 3 ή 4 (Σχήμα 9 – Καμπύλη απόδοσης 2).

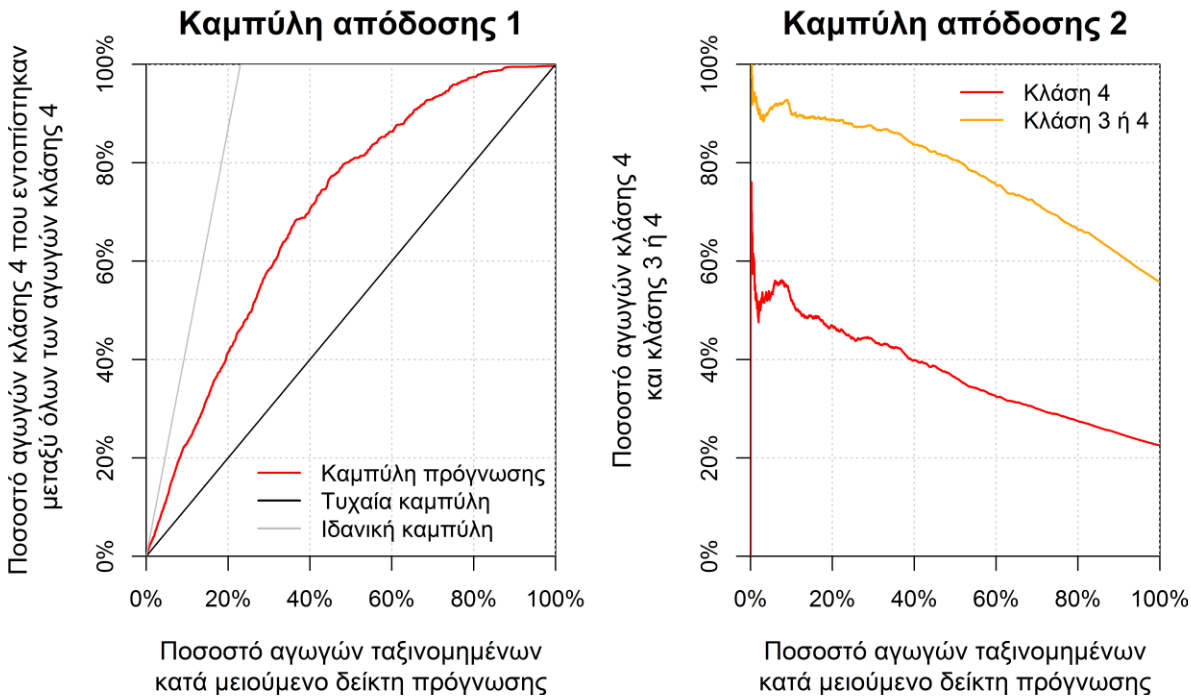
Τα αποτελέσματα δείχνουν ότι το μοντέλο μπορεί όντως να υποστηρίξει τους αρμόδιους φορείς στον εντοπισμό των αγωγών με τις μεγαλύτερες φθορές. Ωστόσο, η ηλικία των αγωγών δεν είναι η μόνη μεταβλητή που επηρεάζει την ταχύτητα φθοράς τους. Τα ποσοστά επιδείνωσης μπορεί να ποικίλλουν σημαντικά μεταξύ αγωγών, με βάση διάφορους κατασκευαστικούς, λειτουργικούς, και περιβαλλοντικούς παράγοντες (υλικό, βάθος εγκατάστασης, διάμετρος, κ.α.) που τους επηρεάζουν (Davies *et al.*, 2001). Άρα, η θεώρηση περαιτέρω μεταβλητών στο μοντέλο, που μπορούν να εξηγήσουν τη διαφορετική συμπεριφορά αγωγών με διαφορετικά χαρακτηριστικά, ίσως βελτιώσει σημαντικά την ποιότητα πρόγνωσης.

5.2.2 Με μεταβλητή το υλικό κατασκευής

Μετά την αναγνώριση του υλικού κατασκευής και του είδους του αποβλήτου ως τους παράγοντες που μπορούν να συνεισφέρουν στη βελτίωση της προγνωστικής ακρίβειας, το μοντέλο βαθμονομήθηκε ακόμη μια φορά, θεωρώντας το υλικό σα μεταβλητή με τιμή 2. Εφόσον το υλικό και το είδος των λυμάτων είναι ισχυρά συσχετισμένα, η ενσωμάτωση και των δύο μεταβλητών στο ίδιο μοντέλο δε βελτιώνει περαιτέρω την ποιότητα πρόγνωσης. Τα αποτελέσματα του νέου μοντέλου φαίνονται στο Σχήμα 10.

Σύμφωνα με τις καμπύλες απόδοσης, η ποιότητα πρόγνωσης βελτιώνεται σημαντικά μετά την ενσωμάτωση του υλικού σα μεταβλητή στο μοντέλο. Καθώς οι καμπύλες των σχημάτων 7 και 10 έχουν προκύψει από το ίδιο σύνολο δεδομένων, τα αποτελέσματά τους είναι άμεσα συγκρίσιμα. Στην περίπτωση του σχήματος 10, και οι δύο καμπύλες φτάνουν σε υψηλότερα ποσοστά, και άρα καλύτερα αποτελέσματα.

Μεταξύ του 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, εντοπίστηκε το 40% των αγωγών κλάσης 4 (Σχήμα 10 – Καμπύλη απόδοσης 1). Επιπρόσθετα, το υποσύνολο που περιέχει το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται κατά 47% από αγωγούς κλάσης 4 και κατά 90% από αγωγούς κλάσης 3 ή 4 (Σχήμα 10 – Καμπύλη απόδοσης 2). Με άλλα λόγια, αν επιθεωρούνταν το 20% των αγωγών οι οποίοι προγνώστηκαν στην κρισιμότερη κατάσταση, το 47% των αγωγών θα βρίσκονταν στην κλάση 4, το 43% στην κλάση 3 και μόνο το 10% θα βρίσκονταν σε κλάση 1 ή 2.



Σχήμα 10: Ποιότητα πρόγνωσης για βαθμονόμηση του μοντέλου με ενσωμάτωση του υλικού σα μεταβλητή με τιμή 2.

5.3 Προσομοίωση με τη μέθοδο Monte Carlo

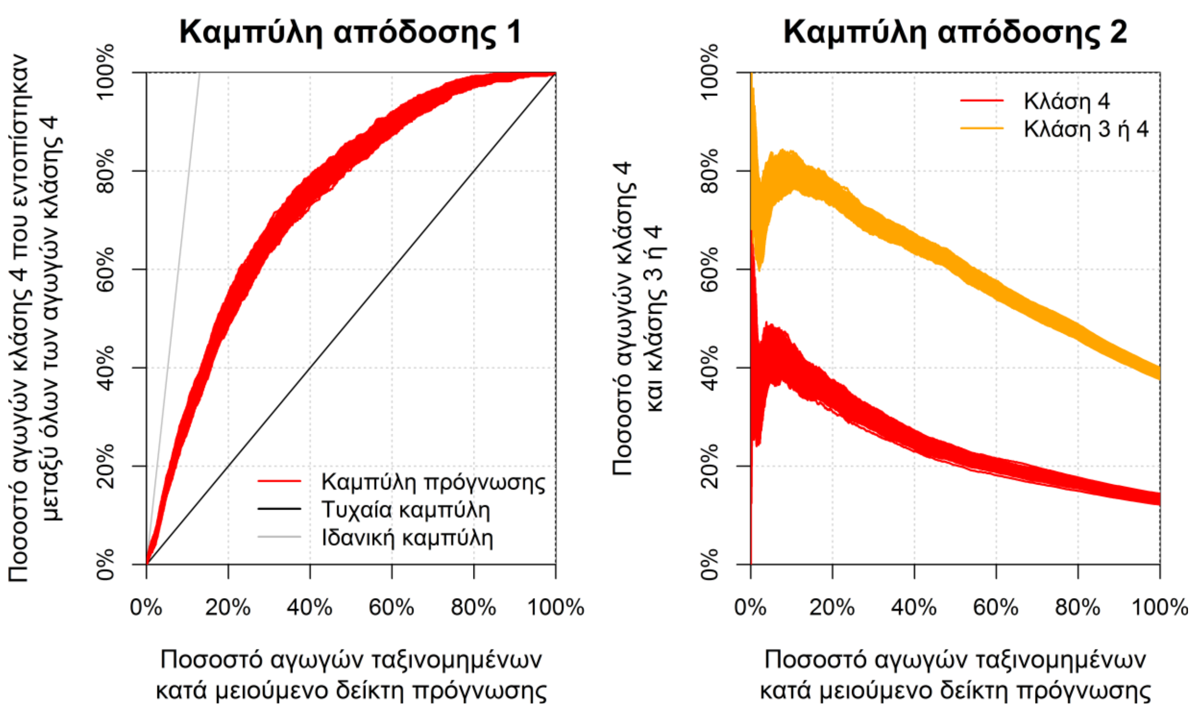
Στη συνέχεια, έγινε έλεγχος της επιρροής του συνόλου των διαθέσιμων δεδομένων στην αξιοπιστία του μοντέλου, με τη μέθοδο Monte Carlo. Η επιλογή των συνόλων βαθμονόμησης και ελέγχου έγινε με τυχαίο τρόπο από το σύνολο των διαθέσιμων δεδομένων και με αναλογία 70%-30% αντίστοιχα.

Το μοντέλο, το οποίο εξακολουθεί να περιλαμβάνει το υλικό σα μεταβλητή με τιμή 2, δοκιμάστηκε για 1000 τυχαίες επιλογές και διαφορετικά μεγέθη συνόλων, για 100%, 20% και 10% της αρχικής βάσης δεδομένων.

5.3.1 Προσομοίωση με το 100% των διαθέσιμων δεδομένων

Οι μορφές και των δύο καμπυλών απόδοσης είναι παρόμοιες με εκείνες οι οποίες προέκυψαν με χρονολογική επιλογή των υποσυνόλων βαθμονόμησης και ελέγχου (Σχήμα 10), μαρτυρώντας παρόμοια ποιότητα πρόγνωσης. Ωστόσο, η μέση κατανομή κλάσεων του υποσυνόλου ελέγχου, χρησιμοποιώντας προσομοίωση Monte Carlo, διαφέρει από την κατανομή κλάσεων του υποσυνόλου ελέγχου, χρησιμοποιώντας χρονολογική κατάταξη (Σχήματα 10 και 11 – Καμπύλη απόδοσης 2, τιμή ψ για χ = 100%).

Αυτό εξηγεί τη διαφορά στα αποτελέσματα, η οποία προκύπτει αναλύοντας συγκεκριμένες τιμές των καμπυλών· μεταξύ του 20% των αγωγών οι οποίοι προγνώστηκαν στη δυσμενέστερη κατάσταση, εντοπίστηκε περισσότερο από το 50% όλων των αγωγών κλάσης 4 (Σχήμα 11 – Καμπύλη απόδοσης 1). Επιπρόσθετα, το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται από περίπου 35% αγωγών κλάσης 4 και 75% αγωγών κλάσης 3 ή 4 (Σχήμα 11 – Καμπύλη απόδοσης 2).

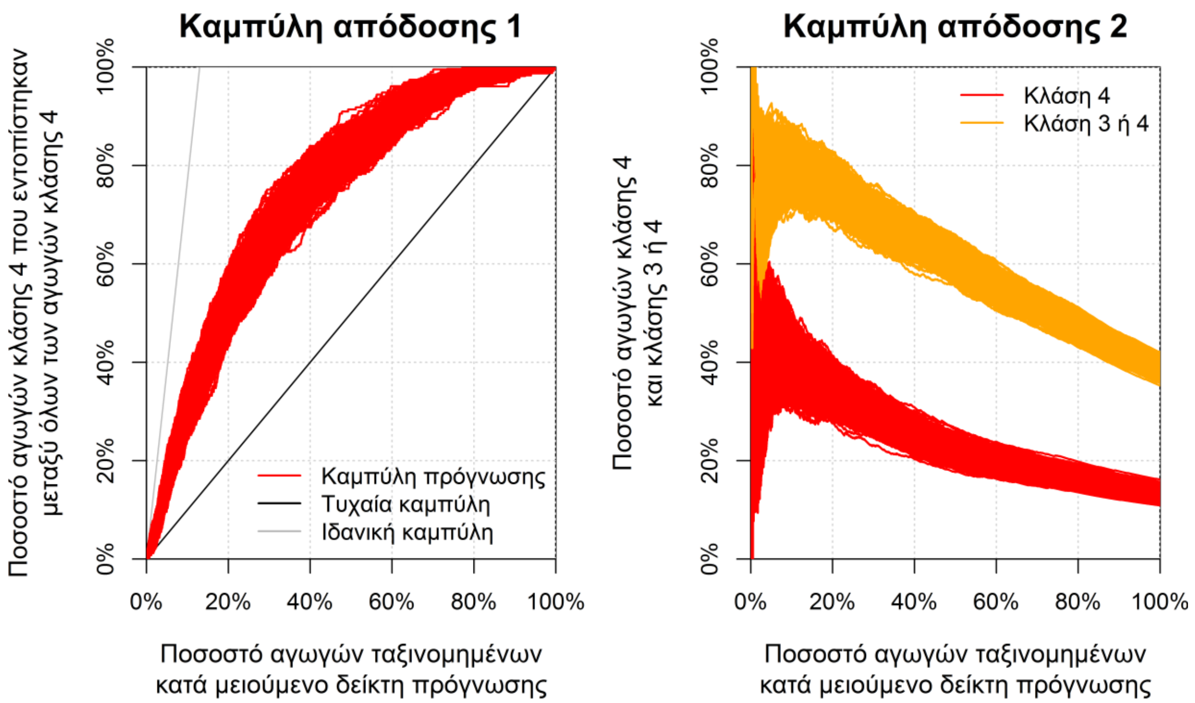


Σχήμα 11: Ποιότητα πρόγνωσης για βαθμονόμηση του μοντέλου με προσθήκη του υλικού σα μεταβλητή με τιμή 2. Το μοντέλο δοκιμάστηκε χρησιμοποιώντας το 100% των διαθέσιμων δεδομένων για 1000 προσομοιώσεις Monte Carlo.

5.3.2 Προσομοίωση με το 20% των διαθέσιμων δεδομένων

Το μοντέλο δοκιμάστηκε ακόμα μια φορά, χρησιμοποιώντας μόνο το 20% των διαθέσιμων δεδομένων (από το οποίο το 70% χρησιμοποιήθηκε για τη βαθμονόμηση και το 30% για τον έλεγχο του μοντέλου).

Γενικά, η μορφή και των δύο καμπυλών απόδοσης 1 και 2 είναι παρόμοια με αυτή που προέκυψε δοκιμάζοντας το μοντέλο με το 100% της βάσης δεδομένων. Σχετικά με την καμπύλη απόδοσης 1, το διάστημα εμπιστοσύνης γύρω από το μέσο ποσοστό αγωγών κλάσης 4, φαίνεται να διπλασιάζεται για όλο το μήκος της. Στην περίπτωση της καμπύλης απόδοσης 2, το μέσο ποσοστό των αγωγών κλάσης 4, και 3 ή 4, παραμένει το ίδιο, ενώ το διάστημα εμπιστοσύνης γύρω από τη μέση τιμή αυξάνει σημαντικά για τα μικρότερα ποσοστά. Καθώς η καμπύλη πλησιάζει το 100% η διαφορά μικραίνει.



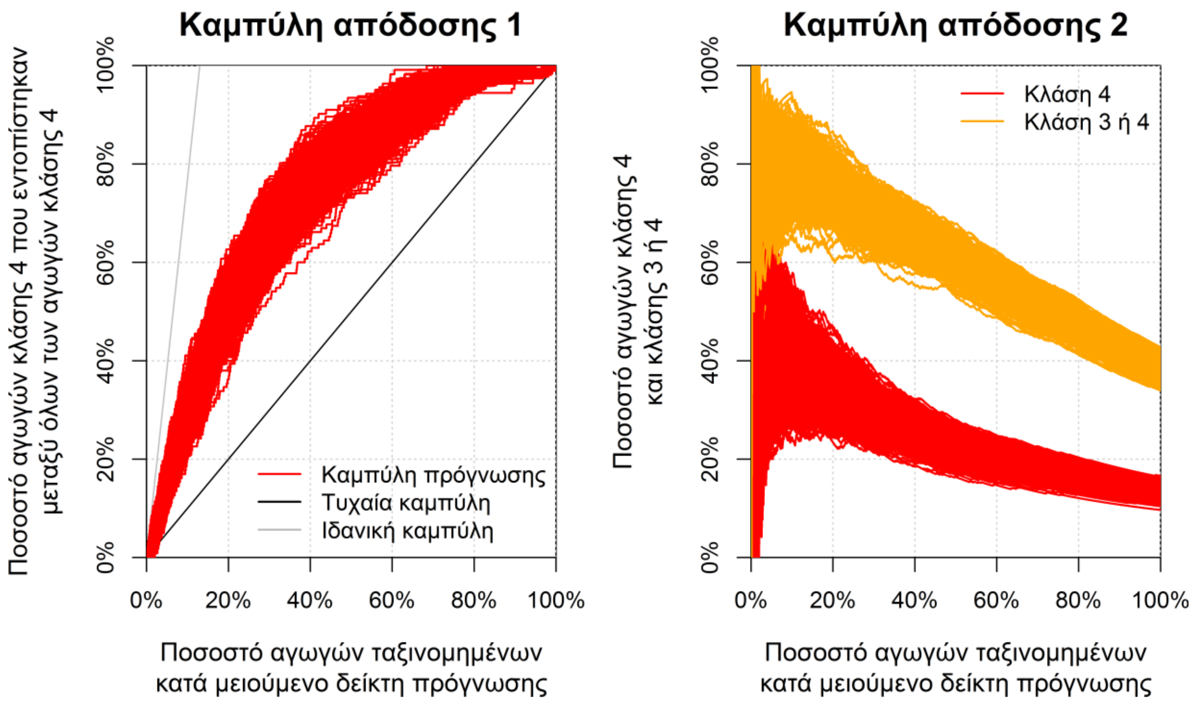
Σχήμα 12: Ποιότητα πρόγνωσης για βαθμονόμηση του μοντέλου με προσθήκη του υλικού σα μεταβλητή με τιμή 2. Το μοντέλο δοκιμάστηκε χρησιμοποιώντας το 20% των διαθέσιμων δεδομένων για 1000 προσομοιώσεις Monte Carlo.

Τα αποτελέσματα δείχνουν ότι το μέγεθος του υποσυνόλου βαθμονόμησης επηρεάζει την ποιότητα πρόγνωσης του μοντέλου· μεταξύ του 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, το ποσοστό των αγωγών κλάσης 4 που εντοπίστηκε, ποικίλει μεταξύ 40% και 60% (Σχήμα 12 – Καμπύλη απόδοσης 1). Επιπρόσθετα, το υποσύνολο που περιέχει το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται από 28% έως 41% αγωγών κλάσης 4, και 68% με 82% αγωγών κλάσης 3 ή 4 (Σχήμα 12 – Καμπύλη απόδοσης 2). Όταν εξεταστεί μόνο το 5% των δυσμενέστερων καταστάσεων (το οποίο στις περισσότερες περιπτώσεις υπερβαίνει τον ετήσιο προϋπολογισμό ενός δήμου για αποκατάσταση του δικτύου αποστράγγισης), ο αριθμός των αγωγών κλάσης 4 κυμαίνεται μεταξύ 30% και 60%, ενώ ο αριθμός των αγωγών κλάσης 3 ή 4 μεταξύ 70% και 90%.

5.3.3 Προσομοίωση με 10% των διαθέσιμων δεδομένων

Για μια ακόμη φορά, η μορφή των καμπυλών απόδοσης 1 και 2 είναι παρόμοια με αυτή που προέκυψε δοκιμάζοντας το μοντέλο με το 100% και 20% των δεδομένων. Ωστόσο, το διάστημα εμπιστοσύνης γύρω από τις καμπύλες αυξάνεται περαιτέρω (Σχήμα 13).

Μεταξύ του 20% των αγωγών με τις δυσμενέστερες προβλέψεις, το ποσοστό των αγωγών κλάσης 4 που εντοπίστηκαν κυμαίνεται από 40% έως 65% (Σχήμα 13 – Καμπύλη απόδοσης 1). Ακόμα, το 20% των δυσμενέστερων προγνώσεων αποτελείται από 22% έως 48% αγωγών κλάσης 4, και 62% έως 88% αγωγών κλάσης 3 ή 4 (Σχήμα 13 – Καμπύλη απόδοσης 2). Αν εξεταστεί το 5% των δυσμενέστερων προγνώσεων, ο αριθμός των αγωγών κλάσης 4 που εντοπίστηκαν κυμαίνεται μεταξύ 20% και 65%, ενώ ο αριθμός των αγωγών κλάσης 3 ή 4 κυμαίνεται μεταξύ 62% και 97%.



Σχήμα 13: Ποιότητα πρόγνωσης για βαθμονόμηση του μοντέλου με προσθήκη του υλικού σα μεταβλητή με τιμή 2. Το μοντέλο δοκιμάστηκε χρησιμοποιώντας το 10% των διαθέσιμων δεδομένων για 1000 προσομοιώσεις Monte Carlo.

Εφόσον η βάση δεδομένων είναι ιδιαίτερα εκτεταμένη (31.394 καταγραφές), κατά την επιλογή τυχαίων υποσυνόλων βαθμονόμησης μεταξύ του 100% των δεδομένων, τα υποσύνολα παρουσιάζουν μικρές διαφορές, όσον αφορά τους τύπους αγωγών που περιλαμβάνουν και επομένως η ποιότητα πρόβλεψης έχει μικρές μεταβολές. Όταν το σύνολο των δεδομένων το οποίο έχει αφιερωθεί για τις δοκιμές του μοντέλου μειώνεται στο 20% και εν συνεχεία στο 10%, οι διαφορές μεταξύ των τυχαίων υποσυνόλων γίνονται σημαντικές. Αυτό οδηγεί σε διαφορές στην απόδοση του μοντέλου.

5.4 Προσομοίωση αργιλικών και σκυροδετημένων αγωγών

Εφόσον το υλικό κατασκευής είναι, όπως αποδείχθηκε παραπάνω, ο σημαντικότερος παράγοντας που επηρεάζει την ταχύτητα επιδείνωσης της κατάστασης των υπόγειων αγωγών, η επίδραση του κάθε υλικού ερευνήθηκε περαιτέρω. Το σύνολο των αγωγών που περιλαμβάνονται στη βάση δεδομένων διαιρέθηκαν σε δύο ομάδες, με βάση το υλικό κατασκευής τους. Τα υποσύνολα βαθμονόμησης και ελέγχου επελέχθησαν χρονολογικά από τα δεδομένα που περιλαμβάνονται σε κάθε ομάδα.

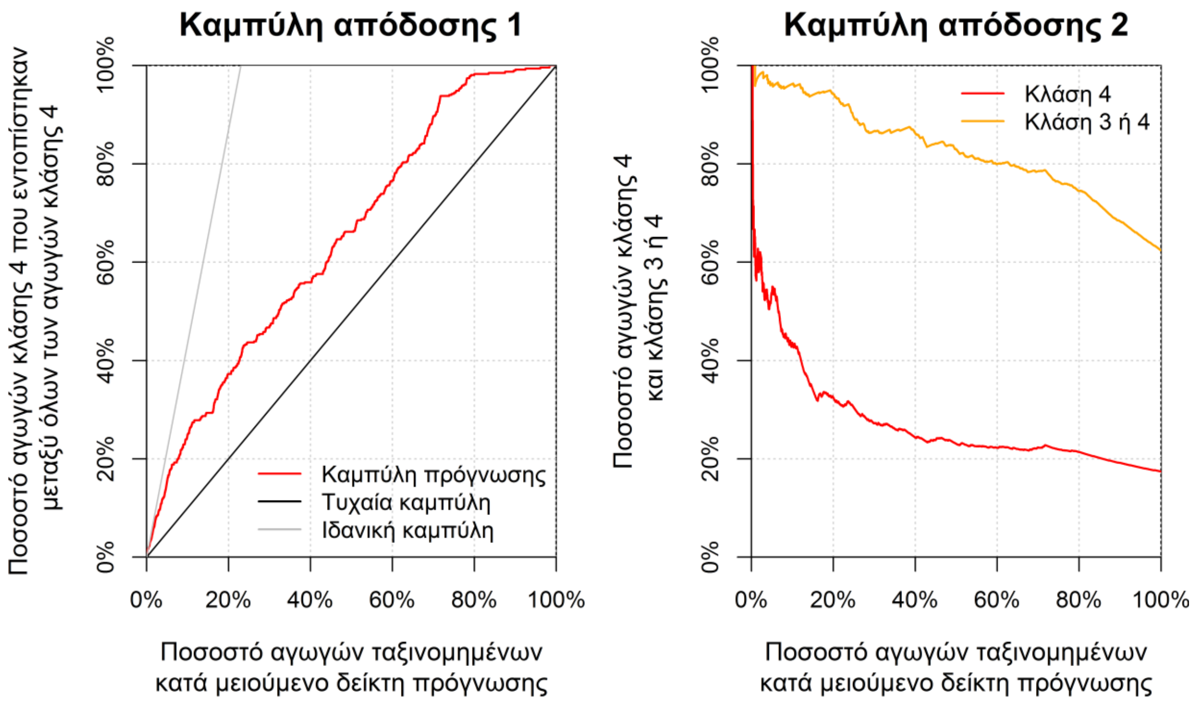
Για να καταστεί δυνατή η απευθείας σύγκριση των αποτελεσμάτων μεταξύ των δύο υλικών, τα υποσύνολα ελέγχου για κάθε περίπτωση επελέχθησαν με σκοπό να περιέχουν ακριβώς την ίδια κατανομή κλάσεων. Αυτό επετεύχθη αφαιρώντας τους αγωγούς μίας κλάσης της μίας εκ των δύο ομάδων, οι οποίοι υπερέβαιναν τον αντίστοιχο αριθμό της ίδιας κλάσης που περιελάμβανε η άλλη ομάδα. Τελικά, το υποσύνολο ελέγχου και για τα δύο υλικά περιέχει 764 αγωγούς κλάσης 1, 244 κλάσης 2, 1205 κλάσης 3, και 467 αγωγούς κλάσης 4. Οι παραπάνω πληθυσμοί αντιστοιχούν σε μια κατανομή με 28,5%, 9,1%, 45%, και 17,4% των αγωγών να ανήκουν στις κλάσεις 1, 2, 3, και 4, αντίστοιχα.

5.4.1 Προσομοίωση αργιλικών αγωγών

Τα αποτελέσματα της προσομοίωσης της ομάδας των αργιλικών αγωγών φαίνονται στο Σχήμα 14.

Μεταξύ του 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, εντοπίστηκε το 38% των αγωγών κλάσης 4 (Σχήμα 14 – Καμπύλη απόδοσης 1). Επιπρόσθετα, το υποσύνολο που περιέχει το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται κατά 32% από αγωγούς κλάσης 4 και κατά 94% από αγωγούς κλάσης 3 ή 4 (Σχήμα 14 – Καμπύλη απόδοσης 2).

Παρότι τα αποτελέσματα δεν είναι ευθέως συγκρίσιμα με όσα παρουσιάστηκαν παραπάνω, καθώς η κατανομή των κλάσεων μεταξύ των υποσυνόλων διαφέρει, η ποιότητα πρόγνωσης για τους αργιλικούς αγωγούς μοιάζει να είναι πολύ κοντά σε αυτή του συνόλου των αγωγών, όταν το υλικό κατασκευής ενσωματώθηκε σα μεταβλητή με τιμή 2.



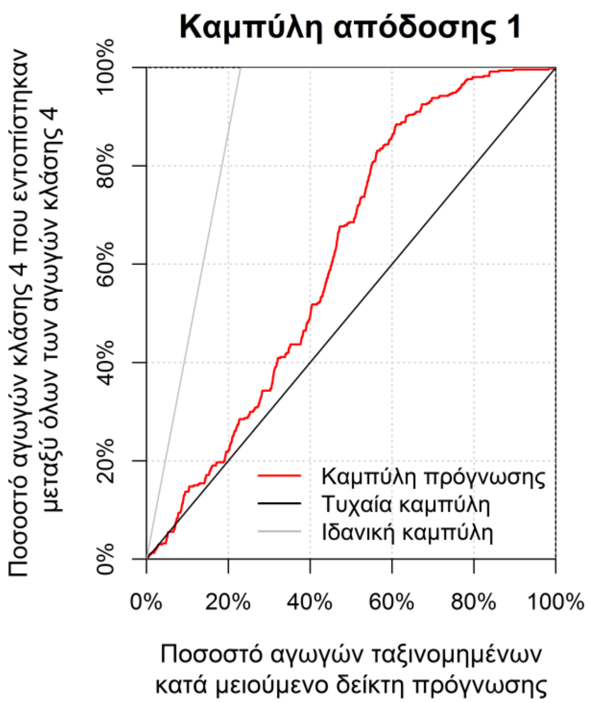
Σχήμα 14: Ποιότητα πρόγνωσης για αργιλικούς αγωγούς και βαθμονόμηση χωρίς προσθήκη μεταβλητών.

Τα παραπάνω αποτελέσματα φαίνονται μάλλον λογικά, καθώς η ενσωμάτωση στο μοντέλο του υλικού σα μεταβλητή, οδηγεί ουσιαστικά στη διαίρεση των δεδομένων σε δύο ομάδες οι οποίες μελετώνται και παραμετροποιούνται ξεχωριστά από το μοντέλο ως προς τη συμπεριφορά τους.

5.4.2 Προσομοίωση αγωγών από σκυρόδεμα

Οι καμπύλες απόδοσης 1 και 2 για τους αγωγούς από σκυρόδεμα, χωρίς καμία προσθήκη μεταβλητών στο μοντέλο, είναι διαθέσιμες στο Σχήμα 15.

Μεταξύ του 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, εντοπίστηκε μόλις το 20% των αγωγών κλάσης 4 (Σχήμα 15 – Καμπύλη απόδοσης 1), ποσοστό που αντιστοιχεί σε μία τυχαία επιλογή αγωγών, χωρίς χρήση του μοντέλου. Επιπρόσθετα, το υποσύνολο που περιέχει το 20% των αγωγών που προγνώστηκαν στη δυσμενέστερη κατάσταση, αποτελείται κατά 20% από αγωγούς κλάσης 4 και κατά 79% από αγωγούς κλάσης 3 ή 4 (Σχήμα 14 – Καμπύλη απόδοσης 2).



Σχήμα 15: Ποιότητα πρόγνωσης για σκυροδετημένους αγωγούς και βαθμονόμηση χωρίς προσθήκη μεταβλητών.

Η καμπύλη απόδοσης 2 επιβεβαιώνει τα αποτελέσματα της πρώτης καμπύλης. Το 20% των αγωγών με τις χειρότερες προβλέψεις αποτελείται κατά 20% από αγωγούς κλάσης 4, το οποίο είναι οριακά μεγαλύτερο από αυτό που αντιστοιχεί στο σύνολο των αγωγών (18%). Τα αποτελέσματα δεν παρουσιάζουν σημαντική βελτίωση όταν οι κλάσεις 3 και 4 λαμβάνονται υπόψη μαζί.

5.4.3 Συμπεράσματα

Για την κατανόηση αυτού του γεγονότος, είναι εξίσου σημαντικό να γίνει κατανοητή η διαφορετική φύση των δύο ομάδων αγωγών. Όπως έχει ήδη αποδειχθεί, οι αγωγοί από σκυρόδεμα που βρίσκονται στη βάση δεδομένων, χρησιμοποιούνται για στράγγιση ομβρίων, και άρα επηρεάζονται από την αβεβαιότητα που περικλείεται σε αυτή τη διαδικασία.

Ενώ στους αγωγούς ακαθάρτων, τα λύματα μεταφέρονται μέσω ενός δικτύου αγωγών που συνδέουν τον κεντρικό αγωγό με μία (τις περισσότερες φορές) οικιακή πηγή, οι αγωγοί ομβρίων παραλαμβάνουν το επιπλέον νερό που καταλήγει κυρίως λόγω βροχοπτώσεων στο αστικό περιβάλλον, και άρα είναι εύκολα προσβάσιμοι σε μία ποικιλία στερεών, όπως ιζήματα, απόβλητα, άμμο, καθώς και κλαδιά δέντρων και φύλλα, τα οποία μπορεί να προκαλέσουν προβλήματα στο εσωτερικό του αγωγου.

Το μοντέλο μπορεί να προβλέψει τη συμπεριφορά ενός αγωγού ακαθάρτων, καθώς αναμένεται να αναλάβει ένα σταθερό όγκο λυμάτων στη διάρκεια ζωής του, και άρα η κατάστασή του εξαρτάται κυρίως από την ηλικία του. Ωστόσο, τυχαία γεγονότα όπως έντονες αλλαγές στη ροή, καθώς και μια ποικιλία στερεών, τα οποία μπορεί να εισέλθουν σε έναν αγωγό ομβρίων, μπορούν να επιβαρύνουν την κατάστασή του με απρόβλεπτο τρόπο.

6 Συζήτηση

6.1 Σύνοψη και συμπεράσματα

Η ποιότητα πρόβλεψης του μοντέλου GompitZ αξιολογήθηκε χρησιμοποιώντας μια εκτενή βάση δεδομένων μιας γερμανικής πόλης, η οποία περιλαμβάνει στοιχεία για τα κατασκευαστικά χαρακτηριστικά και την κατάσταση των αγωγών του δικτύου αποχέτευσης. Το παραπάνω σύνολο δεδομένων είναι αξιοσημείωτο τόσο για το μέγεθός του, όσο και για την αντιστοιχία του στο 100% των αγωγών του υπόγειου δικτύου.

Εφόσον τα βραχυπρόθεσμα προγράμματα αποκατάστασης εστιάζουν στην επισκευή και αντικατάσταση των αγωγών που βρίσκονται στη δυσμενέστερη κατάσταση, το μοντέλο αξιολογήθηκε σύμφωνα με την ικανότητά του να εντοπίζει τους κρίσιμους αγωγούς. Δύο μεθοδολογίες χρησιμοποιήθηκαν για την αξιολόγηση της απόδοσης του μοντέλου σύμφωνα με το παραπάνω κριτήριο.

Τα αποτελέσματα αναδεικνύουν τις δυνατότητες των μοντέλων προσομοίωσης φθοράς να υποστηρίξουν προγράμματα επιθεωρήσεων και αποκατάστασης σε αγωγούς αποχέτευσης. Ακόμα και λαμβάνοντας υποψη τις αβεβαιότητες που εισέρχονται στο πρόβλημα, τα αποτελέσματα του μοντέλου μπορούν να υποστηρίξουν τους αρμόδιους φορείς στην ανάπτυξη αποτελεσματικότερων στρατηγικών.

Ωστόσο, η επιρροή όλων των μεταβλητών που εισέρχονται στο πρόβλημα πρέπει να ληφθεί σοβαρά υπόψη. Το υλικό κατασκευής για παράδειγμα, φαίνεται να έχει σημαντική επίδραση όχι μόνο στην ταχύτητα φθοράς των υπόγειων αγωγών, αλλά και στην ικανότητα του μοντέλου να προγνώσει τη συμπεριφορά τους. Ενώ το μοντέλο φαίνεται να λειτουργεί για τους αργιλικούς αγωγούς ακαθάρτων, είναι μάλλον ανίκανο να προγνώσει τη φθορά των σκυροδετημένων αγωγών ομβρίων. Η συμπεριφορά των αγωγών από σκυρόδεμα μοιάζει να είναι μάλλον απρόβλεπτη, καθώς ο παράγοντας της ηλικίας απεδείχθη ανεπαρκής για μία ακριβή πρόβλεψη.

Από την άλλη, τα συμπεράσματα της προσομοίωσης Monte Carlo υπογραμμίζουν την ευαισθησία της ποιότητας πρόγνωσης στο μέγεθος και την ποιότητα του υποσυνόλου βαθμονόμησης. Και για τις δύο καμπύλες, το διάστημα εμπιστοσύνης γύρω από τη μέση ποιότητα πρόγνωσης αυξάνεται σημαντικά όταν το μέγεθος του συνόλου δεδομένων μειώνεται.

Τέλος, τα αποτελέσματα υπογραμμίζουν επίσης την ανάγκη ενσωμάτωσης μεταβλητών, με σκοπό τη βελτίωση της απόδοσης του μοντέλου. Ωστόσο, ενώ η σωστή επιλογή μεταβλητών και των αντίστοιχων τιμών τους μπορεί να αυξήσει σημαντικά την ποιότητα πρόγνωσης, η εφαρμογή ενός μεγάλου αριθμού μεταβλητών, χωρίς πρακτική αξία, καθώς επίσης και μια εσφαλμένη ερμηνεία των τιμών τους, μπορεί να οδηγήσει στη δημιουργία ενός υπερβολικά πολύπλοκου μοντέλου που είναι χρονοβόρο και παράγει αμφισβητήσιμα αποτελέσματα. Η μεθοδολογία που προτάθηκε παραπάνω μπορεί να εφαρμοστεί και σε άλλες μελέτες για τον εντοπισμό των σχετικότερων μεταβλητών, σε αντιστοιχία με τα διαθέσιμα δεδομένα. Στην παρούσα μελέτη, το υλικό του αγωγού και το είδος του αποβλήτου βρέθηκαν να είναι οι μόνες σημαντικές μεταβλητές.

Όσον αφορά τις καμπύλες απόδοσης, ένα ζήτημα που προέκυψε ήταν η εξάρτησή τους από τη βάση δεδομένων. Φαίνεται ότι ενώ το υποσύνολο βαθμονόμησης επηρεάζει την ποιότητα πρόγνωσης, το υποσύνολο ελέγχου, και πιο συγκεκριμένα η κατανομή κλάσεων σε αυτό, επηρεάζει τις τιμές της καμπύλης. Όπως είναι αναμενόμενο, όσο μεγαλύτερο είναι το ποσοστό των αγωγών σε δυσμενή κατάσταση στο υποσύνολο ελέγχου, τόσο μεγαλύτερο είναι το ποσοστό που το μοντέλο θα μπορέσει να εντοπίσει. Άρα, οι τιμές της καμπύλης αλλάζουν ανάλογα με τα δεδομένα, και ανεξάρτητα της ποιότητας πρόγνωσης. Δυστυχώς αυτό σημαίνει ότι η σύγκριση μεταξύ καμπυλών που έχουν προκύψει από διαφορετικά σύνολα δεδομένων είναι προβληματική. Η απόδοση του μοντέλου θα πρέπει να αξιολογηθεί σύμφωνα με το σχετικό ποσοστό αγωγών που κατάφερε να εντοπίσει, συγκριτικά με το σύνολο των αγωγών σε δυσμενή κατάσταση, αντί για την απόλυτη τιμή.

6.2 Προτάσεις για περαιτέρω έρευνα

Πολλές ερωτήσεις μένουν να απαντηθούν. Η παρούσα μελέτη έγινε χρησιμοποιώντας ένα εκτενές σύνολο παρατηρήσεων. Ωστόσο, στις περισσότερες πόλεις, μόνο μικρά ποσοστά του δικτύου αποχέτευσης έχουν επιθεωρηθεί και δεδομένα που αφορούν κατασκευαστικά χαρακτηριστικά του δικτύου λείπουν (π.χ. άγνωστο υλικό κατασκευής).

Η ευαισθησία των μοντέλων φθοράς στην ποιότητα των δεδομένων εισόδου (αβεβαιότητα οπτικών παρατηρήσεων όπως υπογραμμίσθηκε πρόσφατα από τους Dirksen et al., 2013 και van der Steen et al., 2013), καθώς επίσης και η επιρροή της κατανομής κλάσεων της διαθέσιμης βάσης δεδομένων θα πρέπει να αξιολογηθούν προσεκτικά. Ένας δείκτης θα πρέπει να εξαχθεί και από τις δύο καμπύλες, ο οποίος θα επιτρέψει τη σύγκριση των αποτελεσμάτων μεταξύ καμπυλών που προήλθαν από διαφορετικές βάσεις δεδομένων.

Τέλος, όπως φάνηκε παραπάνω, για τους αγωγούς των οποίων η συμπεριφορά φαίνεται να μην ακολουθεί συγκεκριμένα πρότυπα, κανένα μοντέλο δε μπορεί να προβλέψει την εξέλιξη της κατάστασής τους. Αυτό θα μπορούσε να βελτιωθεί, αν περισσότερες μεταβλητές οι οποίες σχετίζονται με τη φθορά των αγωγών εισέρχονταν στο πρόβλημα. Για παράδειγμα, η τοποθεσία ενός σκυροδετημένου αγωγού ομβρίων θα μπορούσε να εξηγήσει την ποσότητα στερεών και νερού που ο αγωγός αναμένεται να αναλάβει.

Αναφορές

Ahmadi, M., Cherqui, F., De Massiac, J.C., και Le Gauffre P., 2013. Influence of available data on sewer inspection program efficiency, *Urban Water* [online]. Διαθέσιμο από: http://dx.Doi.Org/10.1080/1573062x.2013.831910 [Τελευταία πρόσβαση 30 Απριλίου 2014].

Ana, E.V., 2009. *Sewer asset management - sewer structural deterioration modelling and multi-criteria decision making in sewer rehabilitation projects prioritization*. Διδακτορική διατριβή. Vrije Universiteit Brussel.

Ana, E.V., Bauwens, W., Pessemier, M., Thoeye, C., Smolders, S., Boonen, I., και De Gueldre, G., 2009. An investigation of the factors influencing sewer structural deterioration. *Urban Water Journal*, 6 (4), 303-312.

Ana, E.V. και 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.

Baik, H.S., Jeong, H.S., και Abraham, D.M., 2006. Estimating Transition Probabilities in Markov Chain-Based Deterioration Models for Management of Wastewater Systems. *Water Resources Planning and Management*, 132 (1), 15-24.

Becker, G., Boduroglu, H., Camarinopoulos, S., Frondistou-Yannas, S., Gedikli, A., Kallidromitis, V.G., Kampranis, D., και Sanna, C., 2009. Structural assessment and upgrading of sewers based on inspection results. *Infrastructure Systems*, 15 (4), 321-329.

Chugtai, F. και Zayed, T., 2008. Infrastructure condition prediction models for sustainable sewer pipelines. *Performance of constructed facilities*, 22 (5), 333-341.

Dirksen, J., Clemens, F.H.L.R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., και Snaterse, C.T.M., 2011. The consistency of visual sewer inspection data. *Structure and Infrastructure Engineering*, 9(3), 214-228.

Davies, J.P., Clarke, B.A., Whiter, J.T., και Cunningham, R.J., 2001. Factors influencing the structural deterioration and collapse of rigid sewer pipes. *Urban Water*, 3 (1-2), 73-89.

DWA, 2009. *Zustand der Kanalisation in Deutschland, Ergebnisse der DWA-Umfrage 2009*, Germany: Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.

DWA, 2011. *Advisory Leaflet DWA-M 149-3 - Conditions and Assessment of Drain and Sewer Systems Outside Buildings – Part 3: Condition Classification and Assessment*, 2007. Hennef: DWA Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.

Ens, A., 2012. *Development of a Flexible Framework for Deterioration Modelling in Infrastructure Asset Management*. Thesis (MSc). University of Toronto.

EPA, 2008. *Clean Watersheds Needs Survey 2004 Report to Congress.* US: Environmental Protection Agency.

Khan, Z., Zayed, T., και Moselhi, O., 2010. Structural Condition Assessment of Sewer Pipelines. *Performance of constructed facilities,* 24 (2), 170-179.

Kley, G., Kropp, I., Schmidt, T., και Caradot, N., 2013. *SEMA Report: D 1.1 Review of Available Technologies and Methodologies for Sewer Condition Evaluation* [online]. Βερολίνο, KompetenzZentrum Wasser Berlin. Διαθέσιμο από: http://www.kompetenz-wasser.de/fileadmin/user_upload/pdf/forschung/SEMA/D11_SEMA_Review_of_technologies_and_methodologies_for sewer_condition_evaluation.pdf [Τελευταία πρόσβαση 28 Απριλίου 2014].

Kley G. και Caradot N., 2013. *SEMA Report: D 1.2 Review of Sewer Deterioration Models* [online]. Βερολίνο, KompetenzZentrum Wasser Berlin. Διαθέσιμο από: http://www.kompetenz-wasser.de/fileadmin/user_upload/pdf/forschung/SEMA/D12_SEMA_Review_of_sewer_deterioration_models.pdf [Τελευταία πρόσβαση 28 Απριλίου 2014].

Kropp, I. και Baur, R., 2005. Integrated failure forecasting model for the strategic rehabilitation planning process. *Water Supply,* 5 (2), 1-8.

Le Gat, Y., 2008. Modelling the deterioration process of drainage pipelines. *Urban Water Journal*, 5 (2), 97-106.

Le Gauffre, P., *et al.*, 2004. *Gestion patrimoniale des réseaux d'assainissement urbains, Guide méthodologique.* Παρίσι: Lavoisier Tec&Doc.

Mehle, J., O´Keefe, P., και Wrase, P., 2001. *An Examination of Methods for Condition Rating of Sewer Pipelines*. Thesis (MSc). University of Minnesota.

Micevski, T., Kuszera, G., και Coombes, P., 2002. Markov model for storm water pipe deterioration. *Infrastructure Systems*, 8(2), 49-56.

NASSCO (National Association of Sewer Service Companies), 2007. Pipeline Assessment And Certificate Program (PACP).

ONEMA, 2012. *Observatory on public water and sanitation services, overview of services and of their performance*. Γαλλία: French National Agency for Water and Aquatic Environments.

Saegrov, S., 2006. *CARE – S: Computer Aided Rehabilitation of Sewer and Storm Water Networks*. Εκδόσεις IWA.

Salman, B., 2010. *Infrastructure Management and Deterioration Risk Assessment of Wastewater Collection Systems.* Διδακτορική διατριβή. University of Cincinnati.

Sargent, R.G., 1998. *Verification and Validation of Simulation Models* [online]. New York, Syracuse University. Διαθέσιμο από: http://surface.syr.edu/eecs/7 [Τελευταία πρόσβαση 6 Μαΐου 2014].

Scheidegger, A., Hug, T., Rieckermann, J., και Maurer, M., 2011. Network condition simulator for benchmarking sewer deterioration models. *Water Research*, 45 (16), 4983-4994.

Schlesinger, S., 1979. Terminology for model credibility. *Simulation*, 32 (3), 103-104.

Stein, R. και Gedheri, S., 2009. *Wertermittlung von Abwassernetzen,* Germany: Stein & Partner GmbH.

Tran, H.D., 2007. *Investigation of Deterioraton Models for Stormwater Pipe Systems.* Διδακτορική διατριβή. Victoria University.

Ugarelli, R.M., Selseth, I., Le Gat, Y., Rostum, J., και Krogh, A.H., 2013. Wastewater pipes in Oslo: from condition monitoring to rehabilitation planning, *Water Practice and Techonology*, 8 (3-4), 487-494.

Van der Steen, A.J., Dirksen, J., και Clemens, F.H.L.R., 2013. Visual sewer inspection: detail of coding system versus data quality?. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance* [online]. Διαθέσιμο από:

http://www.tandfonline.com/doi/abs/10.1080/15732479.2013.816974#.U2kJw1dV9K0 [Τελευταία πρόσβαση 6 Μαΐου 2014]

Werey, C., Rozan, A., Wittner, C., Le Gat, Y., Le Gauffre, P., Nirsimloo, K., και 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 et Territoires* [online], 9, 44-53. Διαθέσιμο από: http://www.set-revue.fr/gestion-patrimoniale-des-reseaux-d-assainissement-de-l-etat-des-reseaux-la-planification-de-leur-reh [Τελευταία πρόσβαση 6 Μαΐου 2014]

WERF, 2007. *Infrastructure Management Final Report: Condition Assessment Strategies and Protocols for Water and Wastewater Utility Assets.* US: Water Environment Research Foundation.

WERF, 2009. *Strategic Asset Management Final Report: Remaining Asset Life: A State of the Art Review.* US: Water Environment Research Foundation.

WERF, 2013. *Infrastructure Final Report : Cost Information for Wastewater Pipelines*. US: Water Environment Research Foundation.

Wirahadikusumah, R., Abraham, D., και Iseley, T., 2001. Challenging issues in modeling deterioration of combined sewers. *Infrastructure Systems*, 7, 77-84.

WRc, 2004. Manual of Sewer Condition Classification – 4η Έκδοση. UK: Water Research Centre.

1 Introduction

1.1 Background

Although a functioning sewer network is essential to the preservation of the environmental welfare and public health, traditional strategies, as well as the funds dedicated to the maintenance of drainage systems, contradict that fact (Salman, 2010). Due to the ‘invisibility’ of these structures, their maintenance is often neglected, resulting in a costly rehabilitation, as well as various side effects to the surrounding population and the environment (WEF-ASCE, 1994). According to Ariaratnam et al. (2001) those costs increase further, since the rehabilitation strategies focus mostly on short-term solutions rather than long-term needs.

Several infrastructure studies conducted in the United Kingdom, Australia, and the United States, highlight the on-going deterioration of critical assets in the water and wastewater system. In Germany, a national study estimates that about 17% of the sewers have severe defects and should be immediately or in the short term rehabilitated (DWA, 2009).

At the same time, the investment for maintenance and replacement needs cannot provide users with the required level of service (WERF, 2007). According to a needs survey conducted by EPA (2008), total funding requirements for replacement, rehabilitation, and expansion of existing collection systems, along with constructions of new collection systems, for a 20 year period in the United States, rise up to $65.3 billion, i.e. 32.2% of the total needs of public agencies for wastewater treatment and collection.

In order to prevent failures in the sewerage system that would result in severe problems and significant costs for a municipality, as well as consequences for the population and the environment, the condition of the sewers should be inspected before a catastrophic event occurs (Baur and Herz, 2002). In several German states, local regulations oblige sewer operators to inspect the entire network once every ten years (e.g. North Rhine-Westphalia). In other states, sewer operators have to evaluate and report on the sewers’ condition based on the recorded defects. The most common visual inspection technique uses CCTV cameras, which are driven inside the pipeline. The obtained images are then evaluated and the camera operator assigns a code to each observed defect, representing its severity. All the codes that are assigned to a sewer pipeline are then combined and evaluated according to a classification algorithm, in order to assess the pipeline's condition.

Although this procedure can give a fairly accurate overview of the network's condition, it requires sensitive and expensive equipment, while it is subjective to the experience of the operator and human judgement. The new technology and tools have improved significantly the accuracy as well as the time efficiency of these methods, however, visual inspections on hundreds of kilometres of sewers is still a questionable, fund- and time- consuming procedure.

Mainly due to budget restrictions, inspection rates are generally very low and the condition of only a small part of the system can be known (ONEMA, 2012). A target-driven inspection technique that focuses on more regular inspections of the critical assets that are more prone to deterioration would increase significantly the efficiency of sewer inspections (Baur and Herz, 2002). There is therefore the need to develop a decision-support tool that will enable an inspection planning based on the sewers that are in priority of intervention measures (Chughtai and Zayed, 2008).

Within that scope, more and more municipalities turn to pro-active rehabilitation strategies in order to maintain their buried assets (McDonald and Zhao, 2001· Sousa, 2014). These strategies will enable the development of infrastructure asset management programs that can prioritise rehabilitation needs and define the best distribution of funding available for this purpose (Salman and Salem, 2012).

1.2 Asset management

According to the US Environmental Protection Agency (EPA), asset management is a combination of practices applied to maintain the desired level of service in the most cost-effective manner. This includes rehabilitation, repair, and replacement of defective components, as well as owning and operational costs. It covers the whole life cycle of the assets in a way that pursues sustainability.

Asset management strategies can benefit municipalities in different ways (New Mexico Environmental Finance Center, 2006· EPA):

- Increasing the knowledge of the overall system's condition and its critical parts
- Providing greater ability to plan future investments

- Providing better operational efficiency
- Ensuring the system meets the required and regulated level of service
- Reducing overall costs
- Increasing the safety of the assets
- Prolonging the life of the assets through efficient rehabilitation and maintenance

One of the first asset management programs was suggested by the Water Research Centre and was based on inspections made by Closed-Circuit Television cameras (CCTV). While this strategy was originally implemented only on the critical assets, whose failure would result in great economic damage, it was later expanded to the whole network. Later, more complex modelling techniques were developed that implemented a variety of factors, in order to support decision making and help prioritising rehabilitation (Sousa, 2014).

A considerable amount of research has been dedicated to support this purpose. It mainly focuses on two aspects: (a) simulating the deterioration of pipelines using modelling techniques, and (b) investigating the influence of a variety of factors (mainly sewer features and environmental factors) on the deterioration of drainage pipelines (Khan, 2010).

Models are typically used to simulate the condition class of non-inspected pipes and forecast the evolution of the system's condition. However, most of them fail to show that they can adequately forecast future conditions (Ana and Bauwens, 2010· Scheidegger et al., 2011). The accuracy as well as the sensitivity of the models to both the available input data and the user-defined parameters are yet to be demonstrated. Several modelling approaches are now available but not commonly used by sewer operators and municipalities to support strategies (Kley and Caradot, 2013). Building the end users' confidence in deterioration models would certainly lead to an expanded use of them.

1.3 Research objectives

This thesis aims to test and evaluate the suitability of sewer deterioration models to support inspection and rehabilitation strategies, as well as carefully assess the influence of a variety of factors on the model's quality of prediction. Sewer operators and municipalities need to be fully aware of the model's capabilities and weaknesses, as well as its dependence on the input data, in accordance to a successful operation and interpretation.

In order to achieve that, this study focuses on addressing three practical questions:

- How does one assess the prediction accuracy of deterioration models? The answer to this question depends on the model's aim. Since pro-active rehabilitation strategies focus on identifying the most deteriorated sewers, a meaningful model assessment would focus on the critical pipes that the model is able to identify, rather than the total amount of accurate predictions.
- How does one identify the most relevant covariates and what is their influence on the model predictions? While the right set of covariates might vary for each individual case study, implementing explanatory factors inside the model can have a significant effect on the quality of prediction that remains to be assessed. Several studies have attempted to associate a variety of factors with the deterioration of drainage pipelines. The knowledge of these factors would lead to a much more efficient and target-driven construction of sewer pipelines, as well as an improved accuracy of prediction, as these factors can explain deterioration patterns.
- What prediction accuracy can be reached using an extensive dataset of CCTV records, and what is the model's sensitivity to the size of the available dataset? Most cities have a limited amount of inspection data. It is therefore important to know what limitations that may put on the development of trustworthy models. If the model's performance is only assessed for an exceptional case, where 100% of the system has been inspected, results can be misleading regarding the applicability of the model in other cases.

In order to answer these questions, several practical and technical difficulties had to be overcome. Since the main aim of this study is to evaluate the performance of deterioration models, several methods were evaluated in order to assess their suitability in answering the question. The best assessment method is in accordance with the model's objective, while minimising the amount of uncertainties that arise from its use.

Furthermore, the maximum performance that the model can reach should not be limited by an incorrect interpretation of covariates, nor a limited amount of input data. According to Le Gat (2008), the selected methodology (GompitZ) is rather data 'hungry', as it requires a great amount of data for calibration. In order to overcome this problem and test the model for different amounts of calibration data, there was the need for an extensive dataset of sewer inspections.

In order to ensure the maximum performance is reached, the model is calibrated for different sets of covariates and amount of input data and the optimum set of parameters is determined. Therefore a safe conclusion can be made regarding the question: What are the potentials of sewer deterioration models?

1.4 Methodology

A deterioration tool was developed using the R programming language, based on the GompitZ model (Le Gat 2008), accompanied with modules for data preparation and model performance assessment.

The model's prediction quality was assessed according to two performance curves. The first one was developed by Le Gat to assess the performance of GompitZ model. The second one was developed within the frame of this study, with respect to the model's intended use, as it was claimed by municipalities and in literature. These two curves were used to assess the model's performance for different sets of parameters and input data. First, the model's performance was tested considering the age to be the only relevant covariate.

Next, three methods were developed for assessing and comparing the influence of each factor on the deterioration:

- First, a correlation analysis was performed in order to investigate patterns occurring between the deterioration factors and the sewer's condition. The higher the correlation between a factor and the condition class, the stronger the influence of the covariate on the deterioration.
- Next, the data was separated into two sets, dedicated to model calibration and validation (70% and 30% respectively). For each run, a different factor was used as the sole covariate in the model, in order to determine its influence on the calibration. If sewers with different characteristics (e.g. different materials) are predicted to have different deterioration speeds, it means that the corresponding factor (the material) has an influence on the deterioration.
- Lastly, the prediction accuracy of the model was tested for every consideration of each covariate.

The results of the three above methods were compared in order to arrive to conclusions.

Finally, the prediction accuracy that can be reached using an extensive dataset of CCTV records was identified. The quality of prediction was assessed using Monte Carlo simulation (1000 runs) for different sizes of datasets. Conclusions were made regarding the model's performance and its sensitivity to the covariates as well as the size of the available dataset.

1.5 Thesis overview

The thesis is organised into seven chapters.

Chapter 1 offers an introduction to the topic. It develops the problematic of the study and the current difficulties it will try to address. A short description of the methodology as well as the structure of the thesis is also provided.

Chapter 2 focuses on the theoretical background of sewer deterioration. It offers a description of the most influential deterioration factors according to the literature, as well as a detailed guide through the condition assessment of sewer pipelines, from the sewer inspection methods to the defect coding and condition classification.

Chapter 3 describes the state of the art of sewer deterioration modelling. It offers an overview and a comparison between the different model types, i.e. deterministic, statistical, and artificial intelligence models, focusing on Markov chain theory and the GompitZ model. Finally, it summarises the results of previous studies that were made on the topic.

Chapter 4 describes the process of collecting and processing the data that was used in the case study. It includes descriptive statistics and analysis of the available data.

Chapter 5 develops the methodology that was followed to produce the results. It summarises the model development within R and analyses the methods that were used in order to identify the most influential deterioration factors and assess the model's performance.

Chapter 6 lists the results of the study. The outcomes of the different methodologies are compared with each other in order to arrive to conclusions.

Chapter 7 summarises the results and conclusions of the study and makes suggestions for further research.

2 Introduction to Sewer Networks: Analysis of Deterioration Factors, Monitoring and Condition Assessment

This chapter focuses on the theoretical background of sewer deterioration, from the contributing factors to sewer inspections and condition assessment. First, a literature review of the pipe features and environmental characteristics that have an effect on the deterioration of sewer pipelines is provided. Next, the different steps of the condition assessment procedure are analysed, including a variety of inspection techniques, defect coding systems and condition classification algorithms.

2.1 Overview

2.1.1 Terminology

Several definitions exist to describe the assets of a sewer network. According to the UK Water Research Centre glossary, a sewer is 'a pipe conveying wastewater from buildings comprising more than one curtilage'. In the following, a sewer pipeline is defined as the sewer section from manhole to manhole. The terms 'pipeline' and 'sewer pipeline' are used interchangeably.

Unlike the word sewer, 'pipe' is a fairly general term that can receive a variety of definitions. In this thesis, a 'pipe' refers to a hollow cylinder used to construct the sewer. A sewer pipeline consists of several pipes linked together.

Lastly, sewerage refers to the system of sewers belonging to a particular locality, whereas sewage is the refuse matter conveyed in sewers (Oxford dictionary, sewerage, sewage).

2.1.2 Structure of sewer networks

A sewer network is composed of pipes with the aim of efficiently conveying a type of sewage from the source to the desired destination, i.e. a treatment plant or a disposal point. A sewer pipe can carry two types of sewage: (a) wastewater, containing mainly used water from domestic sources and partly treated water from businesses and industries, as well as (b) storm water, i.e. the excessive amount of water resulting from rainfalls or the melting of snow (Ana, 2009). In some cases the same sewer can carry both types of sewage (combined sewer) (Figure 2.1).

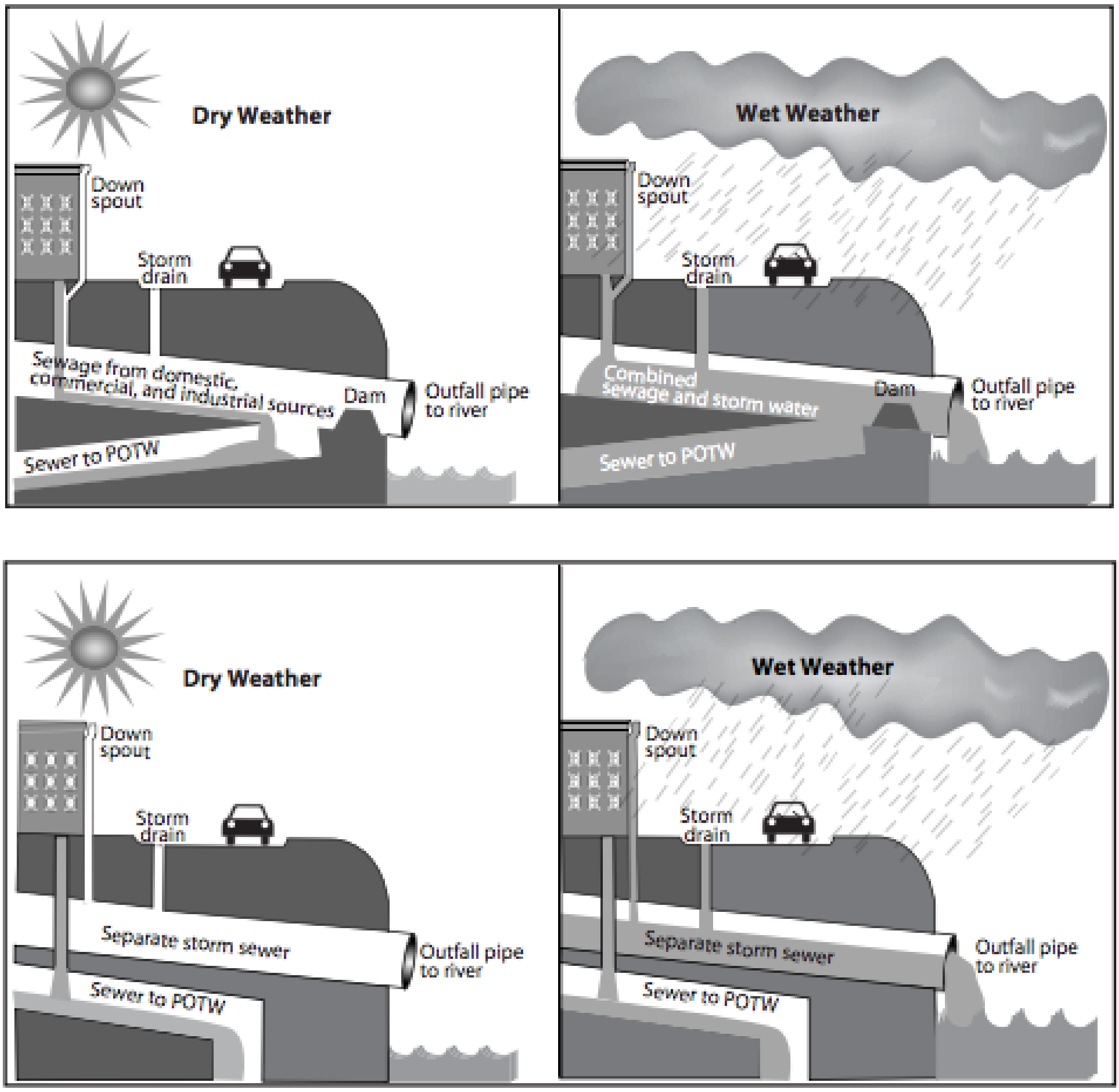


Figure 2.1: Typical combined and separate (sanitary and storm water) sewer systems in dry and wet weather (Source: EPA, 2004)

As water distribution systems continue to age, their structural condition, as well as hydraulic capacity, becomes increasingly worse (Al-Barqawi and Zayed, 2006):

- The structural condition represents the load bearing capacity of a sewer pipeline. Structural defects can appear as observed cracks (along with their size and density), breaks (displaced or missing pieces), and changes in the geometry of the sewer, faulty joints, or even partial collapse of the structure.

- The hydraulic capacity deals with the operational efficiency of the pipeline. The amount of sewage carried in the sewer can be significantly altered by deformations of the original structure (structural deterioration), blockage of the inner capacity (e.g. by tree roots or sediments), or water infiltration and exfiltration.

The maintenance of both the structural and the hydraulic condition of the pipes to their initial standards is essential in order to preserve the required level of service. For the purposes of maintenance, inspection and cleaning, sewer pipes can be accessed through manholes. These are chambers or shafts expanding from the ground surface to the sewer (Schladweiler, 2008).

2.2 Sewer deterioration factors

As it was already mentioned, sewer pipes degrade with time, but not always at the same speed. The life of sewers may vary greatly, and is based on the special circumstances that occur in each case. That means an older pipe is not necessarily in worse condition than a younger one.

There are various factors that have been associated with the structural and hydraulic deterioration of sewer systems. According to their nature, they can be divided into two categories: (a) pipe features, and (b) environmental factors. An overview of the most important deterioration factors is provided in the following.

2.2.1 Pipe Features

Pipe features refer to the construction characteristics of the pipe, defined by the producer. They are the easiest to access and investigate, since they are often made available by the sewer operators. The deterioration factors included in this category are the construction period, material, sewer function, pipe size and shape, installation depth, length, and slope:

Construction period

The installation year of a pipe indicates not only the age of the sewer but also the level of craftsmanship of that specific time. The construction period has been examined thoroughly as a deterioration factor and it’s proven to be highly correlated to the sewer’s condition state (O’Reilly et al., 1989· Ariaratnam et al., 2001· Baur and Herz, 2002· Al-Barqawi and Zayed, 2006). Baur and Herz (2002) and Al-Barqawi and Zayed (2006) found that the failure rate for pipes constructed after WWII was much higher than for the ones constructed before. On the other hand, O’Reilly et al. (1989) suggested that sewers built after WWII, deteriorated slower than the ones constructed before.

Material

Sewers are constructed by a variety of materials. Depending on the location, the availability of resources, the tradition, the installation year and the other sewer features (e.g. the size and the type of sewage), the appropriate material is selected. Most commonly used materials are vitrified clay and concrete, followed by brick, cast iron, asbestos cement, and plastic (PVC). Depending on its strength and durability, as well as its resistance to chemical erosion, the material can offer a different life expectancy to the sewer. Several studies have demonstrated the high relevance of the material to sewer ageing (O’Reilly et al., 1989· Fenner et al., 2000· Ana, 2009· Salman, 2010· Khan, 2010). O’Reilly et al. (1989) found out that cast iron and asbestos cement appear to have the least amount of defects. Fenner et al. (2000) demonstrated that brick pipes are associated with higher failure rates, which was confirmed by Ana (2009), by proving that brick pipes deteriorate significantly faster than concrete pipes. Baur and Herz (2002) also demonstrated the superiority of concrete and stoneware pipes. They found that these pipes require a significantly greater amount of time (approximately 150 years) to transit to a worse condition state than pipes made of PVC (36 years).

Type of sewage

The sewage type was found to have a significant effect on the deterioration, although different studies produced very contradictory results (O’Reilly et al., 1989· Fenner et al., 2000· Ariaratnam et al., 2001· Baur and Herz, 2002· Al-Barqawi and Zayed, 2006· Ana, 2009). The type of sewage that is conveyed in the sewer (waste water, storm water, or both) can degrade its material (Hahn et al., 2002), and accelerate its structural deterioration. O’Reilly et al. (1989) identified higher defect rates for storm water than sanitary sewers with combined sewers lying in the middle. That was explained by the installation depth of storm water pipes, that are typically installed much shallower than sanitary pipes, as well as the large flow fluctuations. Fenner et al. (2000) contradicted that fact by proving that sanitary sewers appear to be the most deteriorated. Baur and Herz (2002) concluded that combined sewers have the least amount of defects. Ana (2009) came to different conclusions for different case studies.

Pipe size

In the case of circular pipes, the size is represented by their diameter. Several studies investigated the influence of pipe size to the sewers' degradation, although results are very contradictory. O'Reilly et al. (1989) proved that medium sized pipes (300-700 mm) have the most defects compared to smaller or larger pipes. On the other hand, Fenner et al. (2000), Ariaratnam et al. (2001), Baur and Herz (2002), and Ana (2009) suggested that smaller diameters have a higher probability of failure. Khan et al. (2010) found a subtle negative impact on the structural condition for pipes larger than 600 mm.

Installation Depth

Sewer pipes are traditionally buried in the ground, so their installation depth can influence their degradation rate. Pipes that are installed very close to the surface (less than 2 meters depth) are proven to deteriorate faster (Fenner et al., 2000· Davies et al., 2001· Kawabata et al., 2003· Ana, 2009). This negative influence is mainly caused by the increased stress of surface factors such as road traffic (O'Reilly et al., 1989· Davies et al., 2001· Ana et al., 2009) as well as variations in the soil's moisture (Jones, 1984). O'Reilly et al. (1989) suggested that the defect rates begin increasing again for pipes that are buried very deep in the ground (under 5.5 meters). That was explained by the increase in the soil pressure or problems arising from the construction of sewers in headings (Davies et al., 2001).

Length

The length of a sewer pipeline is defined as the sewer section from manhole to manhole. The impact of length on degradation was rarely investigated. Some studies indicate that longer sewers deteriorate faster (Fenner et al., 2000· Ana et al., 2009), whereas others (Ana et al., 2009) found the impact of length on the deterioration insignificant.

Pipe shape

The vast majority of sewer pipes are circular. However, there is also a minority of egg-shaped, arch-shaped, and rectangular pipes or even pipes with an irregular shape. Very few studies investigated the influence of the pipe shape on the deterioration process. Baur and Herz (2002) found out that egg-shaped sewers are less prone to structural deterioration. Ana et al. (2009) confirmed this result by identifying decreased deterioration rates for egg-shaped pipes in comparison to circular and arch-shaped.

Slope

Common pipelines have a slope of up to 10%, with the great majority being between 0% and 5%. Flat slopes have been associated with increased rates of corrosion (Ablin and Kinshella, 2004), as well as clogging, due to lower velocity (Ana et al., 2009). On the other hand, very steep slopes increase the danger of abrasion and erosion, due to the high velocity rates (Chugtai and Zayed, 2008). Baur and Herz (2002) found out that sewers with average slopes (1%-5%) are proved to have the lowest deterioration rates. Fenner et al. (2000) and Ana (2009) found out that steep gradients have a positive effect on the pipe's life expectancy.

2.2.2 Environmental factors

The environmental factors influencing pipe degradation are characteristics of the sewer's environment. Unlike the pipe features, information about these factors is not always easy to access. Often, the sewer operators don't collect information about them since they are not directly related to the sewer's operational activity. However, they might be available by other municipal services (Kley and Caradot, 2013).

Traffic load

The type of land use and traffic intensity above the sewer can aggravate its condition by subjecting it to greater amount of pressure that can deform some construction materials. However, it is not clear whether that will ultimately have an effect on the sewer, since there is more care and supervision for sewers that are expected to receive greater pressure. O'Reilly et al. (1989) found an approximately equal deterioration rate for different kinds of overlying roads and a surprisingly increased defect rate for sewers operating under gardens, which could be due to construction works related to houses. Baur and Herz (2002) and Ana (2009) found the least amount of defects on sewers that were laid under main streets undergoing high traffic intensity.

2.2.3 Overview and conclusions

A summary of results from five case studies that facilitates the comparison between them is provided in Table 2.1. While the sample sizes as well as the sewer characteristics, locations, and methods vary greatly, in some cases the studies resulted in the same conclusions while in others, results are very controversial.

Out of the five case studies that are under comparison, two (O'Reilly et al., 1989· Fenner et al., 2000) were conducted in the United Kingdom, one (Baur and Herz, 2002) in Germany and two (Ana, 2009) in the Netherlands. The inspected lengths, sewer characteristics and assessment methodologies for each case study are available in Table 2.1.

It is remarkable, that even using the same method (logistic regression model), Ana (2009), came to different conclusions regarding the influence of most deterioration factors, when using a different dataset. Only the results about the pipe material and size seem to be in agreement for the two Belgian case studies, in Antwerp and Leuven.

Taking this into consideration, it does not come as a surprise that many of the case studies often contradict one another. While some assessment methods might be more suitable than others and some datasets might be more statistically significant, the location and special features of the sewers in each case study, as well as the method that was used, have an influence that cannot be overlooked.

However, there are some patterns that occur and are difficult to ignore even in such varying datasets. Concrete pipes that are laid under busy streets have the least amount of defects, as a result of most case studies. On the other hand, the pipes that were consistently found to be the most deteriorated, were shallow buried, brick pipes with small diameters that are buried under streets undergoing none to medium traffic.

Table 2.1: Influence of sewer deterioration factors on sewer degradation, as they appear in literature

Factors	Reference	Structural Deterioration				
		O'Reilly et al. 1989	Fenner et al. 2000	Baur and Herz 2002	Ana 2009	
Case Study	Location	Southern Water Authority, UK	Central England, UK	Dresden, Germany	Antwerp, Belgium	Leuven, Belgium
	Inspected Sewer Length	180 km	>2000 km	37.8 km 2.7%	63 km	50 km
	Sewer Characteristics	77% Vitrified Clay, 13% Concrete, 80% Sanitary, 9.5% Combined	-	61% Concrete, 36% Stoneware, 3% PVC	Brick, Combined, Young Sewers, Flat Slopes	Concrete, Combined Flat Slopes
	Method	Statistics	Bayesian Model	Cohort Survival Model	Logistic Regression Model	Logistic Regression Model
Construction Period	Positive	After 1944	-	Before 1900	After 1960	Before 1940
	Negative	Before 1944	-	1940-2002	Before 1960	After 1940
Pipe Material	Positive	Cast Iron, Asbestos Cement, Concrete	-	Concrete Stoneware	Concrete	Concrete
	Negative	Pitch Fibre, Vitrified Clay	Bricks	PVC	Bricks	Bricks
Type of Sewage	Positive	Sanitary	-	Combined	Storm Water	Sanitary, Combined
	Negative	Storm Water	Sanitary	Sanitary, Storm Water	Combined	Storm Water
Pipe Size (Diameter mm)	Positive	100-300, 501-550 750+	>225	>1000	Bigger	Bigger
	Negative	300-700	<150	<300	Smaller	Smaller
Installation Depth (m)	Positive	5.5	2-4	-	Deep	Not Significant
	Negative	>6	0-1	-	Shallow	Not Significant
Pipe Length	Positive	-	Shorter	-	Not Significant	Shorter
	Negative	-	Longer	-	Not Significant	Longer
Location	Positive	Railway, Footpath, Trunk Road	-	Main Streets	High Traffic	-
	Negative	Garden, Premises, Principal Road	-	Side Streets	Low-Medium Traffic	-
Pipe Shape	Positive	-	Undecided	Egg-Shaped	Not Significant	Egg-Shaped
	Negative	-	Undecided	Circular	Not Significant	Arch-Shaped Circular
Pipe Slope	Positive	-	Steep Gradients	1-5%	Undecided	Steep Gradients
	Negative	-	Slack Gradients	>5%, <1%	Undecided	Slack Gradients

2.3 Monitoring and Assessing the Sewer Pipeline's Condition

In order to locate the sewers that require rehabilitation, regular inspections of the system are necessary in order to identify sewer failures and assess the remaining life of the assets. In most countries, pipe defects recorded during the inspections are coded according to standard coding systems and the overall condition of the sewer pipeline is evaluated using a classification methodology (e.g. in France, Le Gauffre et al., 2004· in the UK, WRc, 2004· in the US, NASSCO, 2007· in Germany, DWA M149-3, 2011). The result is a score assigned to each sewer pipeline that represents its condition and can be further used for development of new tools and decision-making.

2.3.1 Sewer Inspection Methods

There is a wide range of inspection techniques that can be used alone or combined to identify different kinds of defects (Figure 2.2). The most common are the visual inspections, mainly performed by CCTV cameras, driven from one manhole to the next.

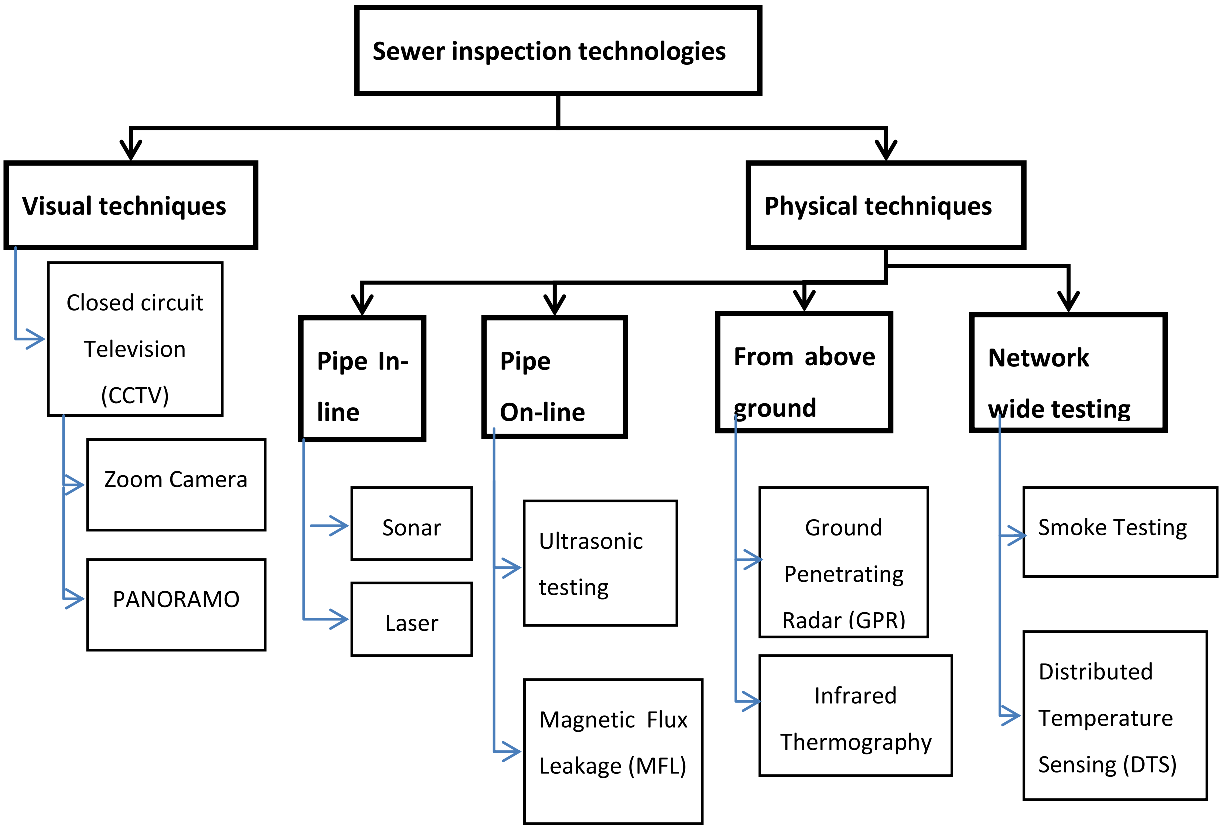


Figure 2.2: Overview of main sewer inspection technologies (Source: Kley et al., 2013)

Complementary to the visual methods, a variety of physical techniques that target a specific kind of defect can be applied to obtain more precise information about the severity of the problem (Kley et al., 2013):

- Sonar and Lasers detect changes in the pipe geometry
- Ultrasonic and Magnetic Flux Leakage (MFL) are on-line inspection techniques, which means they require contact with the pipe wall. They identify problems as corrosion and cracks and measure the wall thickness
- Ground Penetrating Radar (GPR) and Infrared Thermography are used to locate sewer leaks and bedding conditions
- Smoke Testing and Distributed Temperature Sensing (DTS) identify cross-connections and sewer infiltration

2.3.1.1 CCTV Cameras

As previously mentioned, visual inspections represent the majority of sewer inspections, as they enable the detailed observation and storage of pipe defects (Salman, 2010). They are mainly performed by CCTV robots, a construction consisting of a CCTV (Closed Circuit Television) camera and a tractor (Figure 2.3). The camera is accompanied by lighting and attached on the tractor, which makes it convenient for the operator to move and control along the pipe.

Figure 2.3: CCTV robot, “Digi Sewer” by IPEC (Source: www.ipec.at)

The inspection equipment also consists of a service truck containing the necessary tools (power supply, video monitoring and recording equipment) that allow the operator to control the procedure and evaluate the camera output.

How does it work?

The CCTV robot is driven into the sewer with the aid of a cable and a winch (Figure 2.4). While it is moving inside the sewer, the camera records photos and videos of the inside lining of the pipe. The operator can stop the procedure at any time to examine and record defects. The analysis of the optical material by the inspection staff leads to the defect codification. The whole procedure of evaluation and codification of the visual material is performed manually, mostly on site, and is therefore prone to subjectivity and errors (Kley et al., 2013).

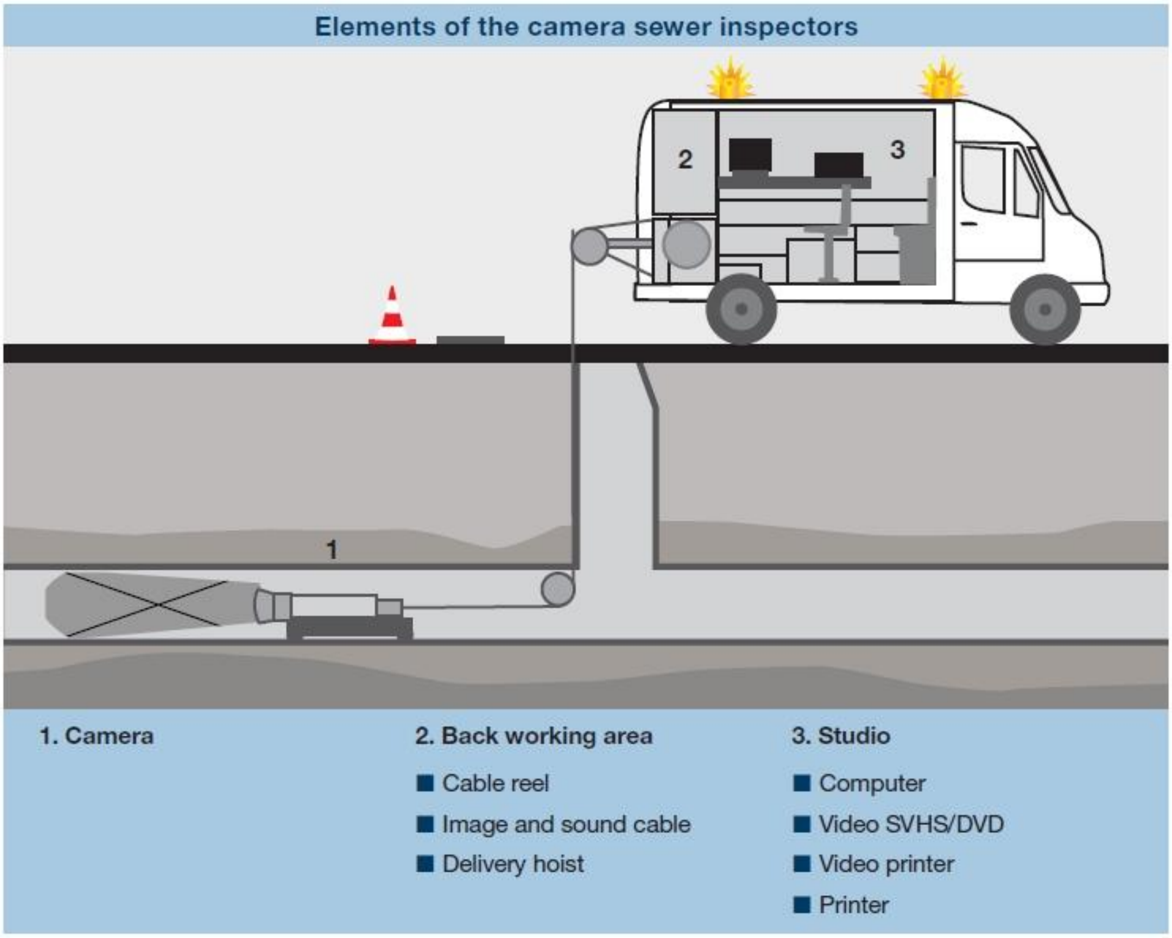


Figure 2.4: CCTV inspection procedure and components (Source: FSCM, 2009)

While CCTV cameras allow the operator to move freely inside the sewer, zoom and look backwards, they have some limitations that question their ease of use. CCTV doesn't allow the user to inspect what is not directly visible through a camera lens. For example they cannot identify defects that are hidden behind obstacles, or under the sewage flow (in the case of non-dewatered sewers) (Hao et al., 2011). They also cannot provide any information regarding external conditions, such as the wall's thickness or the bedding condition (Kley et al., 2013).

In order to improve the performance of visual inspections, sewer inspection companies introduced two new camera types, the zoom camera and the PANORAMO:

- The zoom camera provides a quick overview of the inner condition by recording photos and videos without moving through the pipe. It is lowered through a manhole and remains stationary during the whole procedure.
- PANORAMO was introduced by IBAK (Helmut Hunger GmbH, Germany). It uses two high resolution, digital cameras, one in the front and one in the back that provide an all-round scan of the entire interior in one vertical run (IBAK, 2013).

2.3.2 Defect coding

Defect coding is a standardised procedure defined in Europe by CEN (French: Comité Européen de Normalisation, European Committee for Standardisation). The current directive that applies to all EU-members is the 'EN 13508-2:2003+A1:2011, Investigation and assessment of drain and sewer systems outside buildings - Part 2: Visual inspection coding system'. The suggestions of the directive were (or have to be) implemented and adjusted by the local committee of each member-state. An example of the codification system as it was introduced in the EN 13508-2, 2011 is provided in Table 2.2.

Defect coding is a vital part of the condition classification procedure, as it translates the visual data from the CCTV inspections to a coding system in order to be handled, evaluated and saved for further use. However, it's also subjective as it depends on the personal view, experience and qualifications of the operating staff. According to Dirksen et al. (2013) the possibility of an incorrect observation (recognition and/or description of the defect) is over 50%. Ertl et al. (2007) supported the development of an automatic procedure that will eliminate the subjectivity of the human factor (as described by Fischer et al., 2006) but the effectiveness of it is yet to be demonstrated. An alternative way is to ensure the quality of results produced by the human personnel, by providing education, and control of the results.

Table 2.2: Defect coding examples as introduced in EN 13508-2, 2011

Main code	Characterisation	Quantification	Position	Defect on connection	Longitudinal position	Photo reference	Video reference	Remarks
BAB	C, A	06	3	-	5,50	F16	12:23	-
BBF	B	-	3	-	5,50	F16	12:23	-
Crack(BAB); visible open (C); longitudinal direction (A); 6 mm wide; 3 o'clock, 5,50m after the starting position; on photo 16; in the video from 12:23 Infiltration (BBF); drippy (B); 3 o'clock, 5,50m after the starting position; on photo 16; in the video from 12:23								

2.3.3 Condition Classification Methods

A condition classification method evaluates the defect codes that were assigned to a pipeline during the inspection and translates them to a condition grade that is representative of the pipe's condition. It consists of an algorithm that combines the defect codes with different weights and results to one unique value indicating the overall structural condition, hydraulic condition or both.

Various classification methodologies have been developed, with different interpretations of defect coding. The highest significance can be given to the severity of the defects, the number, the density or a combination of the above mentioned. Alternatively, more recent approaches to sewer classification evaluate the repair length and rehabilitation method as the most important indicators (Kley et al., 2013).

Different classification methods have been developed by companies and national research centers to evaluate these defects in the most effective way. The methods presented below are the French methodology RERAU, the German guideline DWA M 149-3 (2011) and an adapted asset-value based methodology (DWA T4, 2012). According to the nature of their assigned scores, they can be divided into priority-based and substance-based methodologies.

2.3.3.1 Priority-based

As the name suggests, priority-based methodologies classify the pipes according to their urgency (priority) for rehabilitation. They calculate and assign a score to the sewer based on the most severe damage recorded inside the pipeline and the density of the observed defects. Sometimes a risk assessment is also considered in the calculation of the final score (Le Gauffre et al., 2004). Pipes thats failure would result in greater consequences for the environment and the society (underwater pollution, flooding or collapse of main roads) receive a higher priority score. Among the several priority-based methodologies that have been developed, two of them will be further analysed:

RERAU, France

RERAU (Rehabilitation of Urban Sewer Networks) is a national project, created and supported by the French state in order to collect and expand knowledge on sewer rehabilitation. It brings together twenty-two partners working on several research topics. Within the frame of the project, a methodology was developed that translates the defect codes that were assigned to each sewer pipeline to condition classes from 1 to 4, with 4 representing the worst condition. According to the RERAU guide (Le Gauffre et al., 2004, 2007), ten types of dysfunctions (resulting from the defects) are recorded, contributing to eight impacts (Table 2.3):

- Dysfunctions: infiltration (INF), exfiltration (EXF), decrease of hydraulic capacity (HYD), sand silting (SAN), blockage (BLO), destabilisation of ground-pipe system (SPD), ongoing corrosion (COR), ongoing degradation from roots intrusion (ROO), ongoing degradation from abrasion (ABR) and risk of collapse (COL).
- Impacts: pollution of surface waters (POL), pollution of ground and groundwater (POG), nuisances of a hydraulic nature (NUH), disruption of surface activities (TRA), damages to the built environment (DAB), network operation surplus costs (OCS), treatment plant operating surplus costs (OCP), cost associated to shortened lifetime (SLC).

Rehabilitation needs are assessed by decision criteria that define the relationship between dysfunctions recorded on specific sewer segments and the impact they may result in if not treated. For example R/POL/SPI/INF is a criterion that represents the risk of pollution (POL) caused by infiltration (INF) due to dry weather spillages (SPI) (Le Gauffre et al., 2009). Finally, a condition grade is assigned to each criterion, representing the urgency for rehabilitation.

Table 2.3: Impacts and dysfunctions links (Source: Le Gauffre et al., 2009)

DYSFUNCTIONS

IMPACTS

	INF	EXF	HYD	SAN	BLO	SPD	COR	ROO	ABR	COL
POL	**x**		**x**	**x**	**x**					
POG		**x**	**x**	**x**	**x**					
NUH			**x**	**x**	**x**					
TRA				**x**	**x**					**x**
DAB		**x**								**x**
OCS	**x**			**x**	**x**					
OCP	**x**									
SLC						**x**	**x**	**x**	**x**	

An example:

INF2 is the infiltration that would be expected as a result of the observed contributing conditions around the sewer (water level, type of backfill). INF3 is the infiltration as observed during CCTV inspections of the hydraulic condition in the sewer segment. Each of these dysfunctions is given a performance indicator. The combination of the scores for INF2 and INF3 results to the score of INF4, which represents the overall condition of the dysfunction infiltration. Similarly, the amount of abnormal spillages (SPI) combined with the sensitivity of the receiving waters (V-SPI) summarises the risk of pollution (POL). Finally, the intensity of the dysfunction (infiltration) is considered together with the impact of the result (pollution), which in turn provides the overall performance indicator of the specific criteria (Figure 2.5) (Le Gauffre et al., 2009).

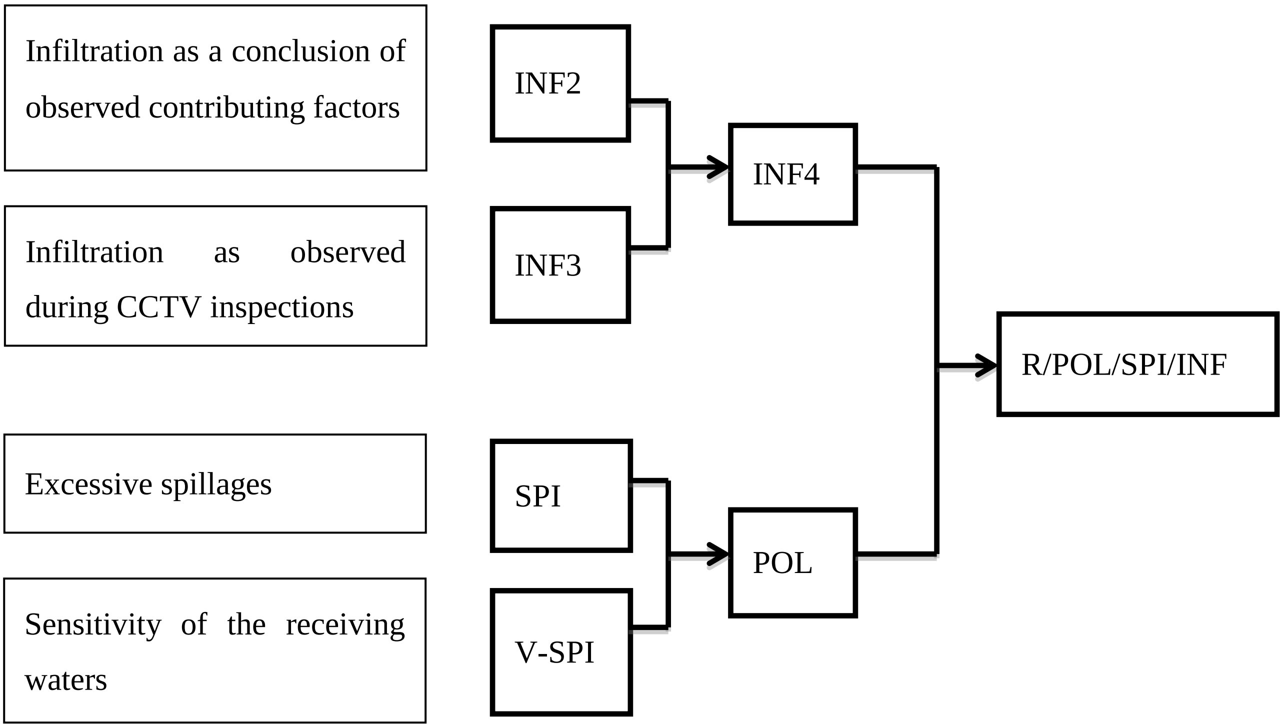


Figure 2.5: Indicators contributing to criterion R/POL/SPI/INF (adapted by Le Gauffre et al., 2009)

An example on how to combine the different indicators is provided in Table 2.4.

Table 2.4: Example of aggregation table (adapted by Le Gauffre et al., 2009)

		PI2			
		G1	**G2**	**G3**	**G4**
	G1	1	1	2	2
PI1	**G2**	1	2	2	3
	G3	2	2	3	4
	G4	2	3	4	4

The numbers indicate the priority of rehabilitation whereas PI (Performance Indicator) represents the different dysfunctions and impacts mentioned above (INF2, SPI, POL). The full guide is available in the RERAU guidebook by Le Gauffre et al. (2004).

Based on the methodology described above, several companies developed software for condition assessment, such as 'G2C environment' within the framework of INDIGAU research program, as well as 'Veolia Eau' with "OctaVE".

DWA-M 149-3, Germany

DWA-M 149-3 was developed by the German organisation DWA, which promotes research and international cooperation. Similarly to RERAU, the objective of this methodology is to rank the sewers according to their urgency for intervention measures. Factors that influence the prioritisation scheme are the most severe defect found in the sewer segment, along with the density of the observed defects and the corresponding impact to the environmental and social wellbeing.

DWA-M 149-3 evaluates both the operational and structural condition of a sewer. It assigns a grade for each defect (from 0 to 4 with 0 representing the worst condition) indicating its influence on three fundamental requirements: Leaktightness (L), Stability (S) and Operational Safety (O). A classification example, as introduced in the DWA-M 149-3 guidebook 'Conditions and Assessment of Drain and Sewer Systems Outside Buildings – Part 3: Condition Classification and Assessment, November 2007' is provided in Table 2.5. The absence of mortar (BAE) affects the Leaktightness and Stability of the sewer but not its operational safety. According to the millimetres of the sewer that are affected, a condition class is assigned for each requirement. If the absence of mortar expands for example within 25 mm of the sewer, a condition class of 4 is assigned for Leaktightness and 3 for Stability (Table 2.5).

The worst condition class that was assigned to each requirement of each sewer is combined with the density of the defects to give the final condition class. In case of a boundary condition, additional criteria are considered. Lastly, the final condition grade that was assigned to each requirement is merged into a rehabilitation coefficient based on the requirement that was challenged the most. The result of that rehabilitation coefficient is an overall condition rating for the sewer, indicating the time when rehabilitation actions will be required (Kley et al., 2013).

Table 2.5: Example of classification table for missing mortar (Source: DWA-M 149-3, page 29, Table A.6)

Main Code	Characterisation		Requirements			Unit	Condition Class				
	Ch1	Ch2	L	S	O		0	1	2	3	4
BAE	-	-	x			mm			≥100		<100
				x		mm			≥100	≥10 <100	<10

DWA-M 149-3 has been adjusted to the modern codification suggested by EN 13508-2 (2011). Therefore the sewers that were inspected with previous defect coding systems have to be translated in order to be evaluated and classified with this methodology.

2.3.3.2 Substance-based

Unlike priority classification, substance-based methodologies consider the length of pipe that requires rehabilitation, rather than the urgency. They use as criterion the type of intervention needed: replacement, renovation or repair. The condition grade that is assigned to the sewer indicates the length that should be rehabilitated. Defects that are close to each other and apply to the same repair length are considered together, reducing the intervention costs. Therefore a sewer with one major defect or several defects within a small distance requires a repair within a limited length and will result in a better average score. On the other hand, a sewer with a lot of minor defects spread over its length will probably require an expensive replacement as the most cost-effective solution and thus result in a worse substance score.

DWA T4, Germany

DWA T4 is a substance-based classification methodology introduced by DWA. Similarly to DWA-M 149-3, it assigns a condition grade to every defect according to its impact on each of the three sewer requirements: Leaktightness (L), Stability (S) and Operational Safety (O). A specific repair length is then defined for each defect and represented on a profile view of the sewer (Figure 2.6).

Typically, lengths of 1m and 4m are assigned to defects that can be repaired with no-dig or open-trench technologies, respectively. In case of defects that are close to each other and result in overlapping repair lengths, the worst condition defines the repair length and defects are rehabilitated together (Kley et al., 2013).

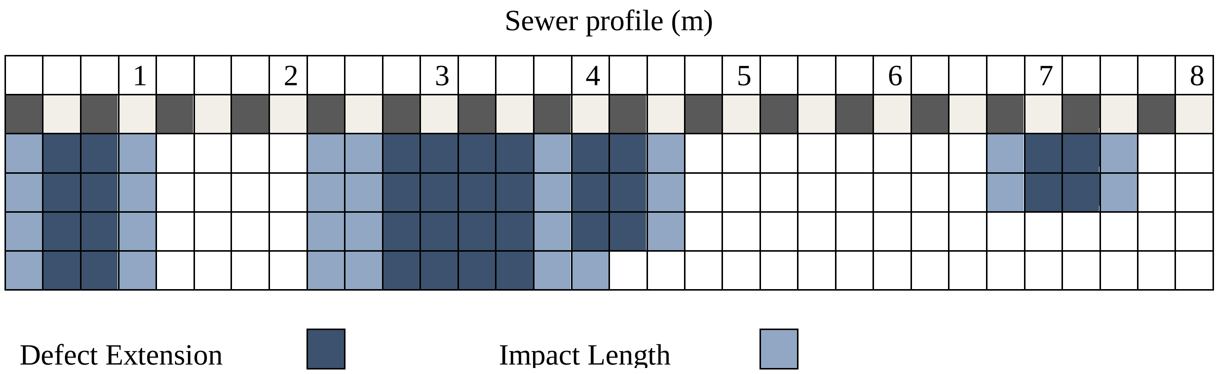


Figure 2.6: Attribution of repair lengths along the sewer (Source: project SEMA, Kley et al., 2013)

Finally, a total length is defined for each condition grade and a weight indicating the severity of the condition (with 1 representing the worst interim condition). The repair density is calculated as the ratio of the repair lengths to the total sewer length. For densities greater than 30%, a replacement of the sewer is suggested as the most cost-effective solution. For lower values, repair and restoration actions can be applied (Kley et al., 2013). The full guideline is published and provided by DWA as 'Leitfaden zur strategischen Sanierungsplanung von Entwässerungssystemen außerhalb von Gebäuden - T4/2012'.

3 Sewer deterioration modelling

Since rehabilitation strategies are limited by the lack of information on sewer condition, deterioration models have been developed to forecast the evolution of the system according to its current and past condition (Tran, 2007· Chughtai and Zayed, 2008· Le Gat, 2008· Ana et al., 2009· Salman, 2010· Khan et al., 2010· Ens, 2012· Ahmadi et al., 2013). This chapter offers an overview of the most common modelling techniques, focusing on Markov chain theory and the GompitZ model. In addition, it provides an overview of their validation results, as they were concluded from different case studies.

3.1 Overview of modelling techniques

A wide variety of models have been developed in order to simulate and forecast the ageing of urban drainage systems. According to their main operating principle, they can be classified in three categories: deterministic, statistical and artificial intelligence models (Tran, 2007). While deterministic and statistical models relate the deterioration factors to the pipe's condition using the corresponding mathematical functions, artificial intelligence models attempt to simulate the function of the human nervous system, by investigating manners and creating mathematical relations between the dependent (condition classes) and the independent (deterioration factors) variable (WERF, 2009).

3.1.1 Deterministic models

As it was already mentioned, deterministic models are used when the relationship between the components is or is assumed to be known. They are based on an already existing understanding of the deterioration patterns and assume a mathematical relation described by an equation. The parameters of the equation are calculated to provide the best fit between predicted and observed values (WERF, 2009).

According to a review conducted by WERF (2009), deterministic models can be applied with an empirical or physical approach. Empirical deterministic models can only be applied to cohorts of assets (groups that are suspected to have a homogenous behaviour) while physical can only be applied to individual assets. Both require a dataset of historical as well as asset related data. Examples of deterministic models are linear and exponential (Tran, 2007).

Deterministic models describe the first attempts to simulate the deterioration of infrastructure facilities, due to their simplicity and capability to form a direct relation between cause and effect (Tran, 2007). However, they are not appropriate to model the deterioration of assets due to the following:

- While they can successfully describe simpler procedures (e.g. corrosion), the deterioration of the drainage system follows complicated patterns that depend on a large amount of factors and are difficult to understand and describe (Schmidt, 2009),
- Linear and exponential models are not suitable to model the discrete condition states describing the deterioration of sewer pipelines (Tran, 2007),
- Deterministic models do not integrate the inherent randomness of failure events (Morcous et al., 2002).

Although in literature there are several applications of empirical deterministic models to water pipelines, there are no commercial tools known (WERF, 2009). An application of deterministic models to detect failure rates in water pipes can be found in Kleiner and Rajani (2001).

3.1.2 Statistical models

Similar to deterministic, statistical models are considered model-driven, since they are based on an already known relation that is assumed to occur between input and output factors. They provide a more realistic modelling approach of sewer deterioration than deterministic models, since the deterioration of assets is treated as a stochastic process and the outcome is not a quantitative but a probability value; this might include a binary outcome, multiple category responses or a transition probabilities matrix (Tran, 2007).

The parameters of the model represent the impact of the corresponding covariate and are calculated by calibrating the model using maximum-likelihood estimation methods in order to minimise the error between predicted and observed values (WERF, 2009). A large dataset of historical data as well as asset related data is necessary in order to calibrate the model. The asset related data is used as explanatory covariates in the model, as it is suspected to drive the deterioration process.

Statistical models have been extensively used in literature to model the deterioration of sewer pipelines (Tran, 2007· Chughtai and Zayed, 2008· Le Gat, 2008· Ana et al., 2009· Salman, 2010· Ens, 2012· Salman and Salem, 2012), due to their relative simplicity and effectiveness answering the problem. They take into account the probabilistic nature of the deterioration process, as well as the random effects. Examples of statistical models are the cohort survival model, the Markov and semi-Markov model (using Markov chain theory), the ordinal and logistic regression analysis, as well as the linear and multiple discriminant analysis.

Statistical models have been primarily used as cohort-type models. In accordance with a successful calibration, the cohorts must be small enough to have homogeneous deterioration behaviour, but at the same time large enough to provide sufficient amount of data for the calibration (WERF, 2009). However, municipalities do not always have this amount of data available and therefore their use becomes limited by the amount of knowledge on sewer condition.

Commercial tools based on statistical models are the KANEW (2003), AwwaRF (2005), GompitZ (2008), and STATUS (2009).

3.1.3 Artificial intelligence models

Unlike deterministic and statistical models, artificial intelligence models are data-driven, which means that no assumptions are made regarding the model's structure. They aim to mimic the function of the human nervous system, creating relations by learning and generalising. Artificial intelligence models are classified as 'black box' models, since they produce an output from the provided input without any knowledge of the model's structure, i.e. the procedure according to which the model produces the corresponding result (Tran, 2007).

While these models do not depend on any mathematical relation and can automatically detect non-linear relations, the lack of an underlying model has the danger of over fitting (the model is too dependent on the provided input data). In addition, artificial intelligence models are time-consuming and require an extensive dataset.

According to Tran (2007), three artificial intelligence techniques have been used for modelling the deterioration of infrastructure facilities: case-based reasoning (CBR), fuzzy set theory, and neural networks (NNs). Case-based reasoning is a problem-solving regime based on the experience of previous cases that simulates human judgement. It requires an extensive and diverse database in order to solve the case by adjusting and retrieving information from the case library. Fuzzy set theory is mainly used under data scarcity or when the available information is expressed in qualitative terms (e.g. "poor", "good"). Lastly, neural network models imitate the function of the human brain by creating a network of artificial neurons, each of which receives and outputs a signal. They learn the behaviour patterns using the input data and produce a relation between cause and effect (Kley and Caradot, 2013).

Commercial tools based on artificial intelligence models include SPSS, NeuralWare, and Neuroshell2, as well as some MATLAB applications.

3.2 GompitZ: A Markov chain model for sewer deterioration

In statistical modelling, Markov chain theory is a widely used methodology for the prediction of the future condition of many types of infrastructure (Le Gat, 2008). It describes the behaviour of a system that passes through a finite amount of condition states. At each time step, the system may change its condition from the current to a worse one or remain in the same condition, according to a given probability. Several authors have already demonstrated applications of the Markov model to predict sewer pipes' structural deterioration (Mehle et al., 2001· Wirahadikusumah et al., 2001· Micevski et al., 2002· Baik et al., 2006· Tran, 2007· Ana, 2009· Ugarelli et al., 2013). Several tools using Markov theory have been developed for sewer systems during the last 15 years, including KANEW (Kropp and Baur, 2005), GompitZ (Le Gat, 2008), and STATUS (Stein and Gedheri, 2009).

GompitZ is a Markov-based statistical tool introduced by Yves Le Gat (2008) with the aim of modelling the deterioration of urban drainage infrastructures. It was developed in France within the European research project CareS (Saegrov, 2006), as a computer program written in the C programming language. GompitZ is based on a mixed multi-state deterioration process that uses non-homogeneous Markov chains (NHMC) (Le Gat, 2008).

An introduction to the theoretical background of the model was published by Le Gat as "Modelling the deterioration process of urban drainage pipelines" in 2008 as well as a detailed user manual "Modelling the Degradation of Sewer Pipelines in CareS: The GompitZ Tool User's Guide – version 2.08" in 2011. In this chapter the basic features of the model are presented.

3.2.1 Model calibration

The model consists of a set of equations that can be used to describe the condition of the pipelines at a finite point in the future. During the calibration step, the model uses survival functions to calculate time-dependent transition probabilities, i.e. probabilities for each pipe to remain in the same condition or move on to the next worse condition state. The model's transition probabilities are calculated using Gompit analysis (an extension of probit analysis that uses the Gompertz distribution). The model was named after Gompit analysis, also accounting for the fact it can handle covariates, denoted in literature by the vector Z, resulting in the name GompitZ.

In order to perform the calibration, GompitZ requires the following information:

- A condition class for each pipe, that represents its condition at the time of the inspection.
- Covariates of the model, i.e. explanatory variables that are suspected to partially explain the deterioration process (Davies et al., 2001· Ana et al., 2009).
- A status for each covariate that defines its influence on the sewer's degradation: no influence (status 0), influence on the initial state only (status 1), influence on the deterioration process only (status 2), or both (status 3).

The transition probabilities are conditional on the values of the set of covariates, as well as the pipe's initial state (inspected condition class) and the random effects' distribution (IFF-Individual Frailty Factor). The IFF integrates the randomness of the failure events. During the calibration of the model, the parameters are calculated by maximising the logarithm of the marginal likelihood of the observations (Le Gat, 2011). They modify the survival functions (Figure 3.1) in order to give the best fit among the input data provided to calibrate the model. In case the consideration of a covariate does not have any explanatory value for the deterioration, the parameter is set to 0 and the covariate is not considered in the calibration process.

Figure 3.1: Example of survival curves, as they were calculated by GompitZ model, based on the Gompertz survival functions. Each colour represents one of the four condition classes, with red representing the worst condition (4) and green representing the best (1). Yellow and orange represent intermediary condition classes 2 and 3. The x axis demonstrates the amount of years it takes to move on to the next condition state, while the y axis demonstrates the percentage of the network that will have passed to the next state at the corresponding amount of time (y axis).

3.2.2 Model prediction

The prediction module is executed after the calibration and uses all the information that was provided or produced during the calibration. A prediction can be made for pipes that were inspected before but also for pipes with a missing inspection.

The model outputs for each pipe a vector of condition probabilities $\boldsymbol{p}$ = (p1(t), p2(t),…, pn(t)) at each year *t*, where n is the number of condition classes that are assigned by the condition classification algorithm. Le Gat (2008) suggests synthesising the condition probabilities in a prediction index, useful to rank the pipes according to their predicted condition. The prediction index (PI) is calculated as the scalar product of the condition probability vector and a condition score vector:

$$PI = \boldsymbol{p}^T \boldsymbol{\kappa},$$

With $\boldsymbol{p^T}$ being the transposed condition probability vector (e.g. (p1(t), p2(t), p3(t), p4(t))) and $\boldsymbol{\kappa}$ a condition score vector containing user-defined weights for each probability (e.g. (0, 1, 2, 3)). Since a higher weight (2, 3) is assigned to the probabilities to be in the most deteriorated conditions (p3(t), p4(t)), the prediction index increases as the condition of the pipe becomes worse.

An example:

If the transposed condition probability vector equalled to p^T = (0, 0.17, 0.30, 0.43), and the condition score vector κ = (0, 1, 2, 3), the prediction index would result in

$$PI = 0*0 + 0.17*1 + 0.30*2 + 0.43*3 = 2.06$$

This value is characteristic of the sewer's condition and has a much greater applicability compared to the vector of condition probabilities.

The prediction module of GompitZ, also allows testing different rehabilitation strategies. These strategies are defined by a number n, which has to be provided as additional information to the model. The number n can indicate one of the following: "do-nothing" strategy in case no rehabilitation measures are taken (n=0), length-driven strategy, which focuses on rehabilitating each year the most deteriorated pipes, up to a maximum total specified length (n=1), budget-driven strategy, which focuses on rehabilitating each year the most deteriorated pipes, up to a maximum total cost (n=2), and optimisation strategy, which targets on bringing the system in a specified condition probability vector and maintaining it there (n=3).

3.3 Model evaluation techniques and application in three case studies

The validation of the model is the final and most crucial step, as it determines the applicability and usefulness of the tool to predict the future condition of the sewer pipelines. The results of the validation can be used to demonstrate the model's value as well as its uncertainties to the end users. Only when the operators are fully aware of the model's strengths and limitations, they are able to use the available tools in the most effective way.

3.3.1 Evaluation Techniques

The evaluation of the model requires a comparison between the prediction results, which were produced by the model, and the observed values, which resulted from actual inspections performed inside the sewers. Typically, the available data is divided into two sets, dedicated to model calibration and validation. After the model is calibrated, a prediction is made on the pipelines of the validation dataset, in order to predict their condition in the year of their inspection. The prediction results are then cross-validated with the inspected ones (for the years when there is a recorded inspection) and conclusions are made regarding the model's predictive value.

Chi-square goodness-of-fit test

The goodness-of-fit (based on the Pearson's chi-square), tests how accurately the model's predictions fit the observations and is probably the most commonly used statistical test to determine the prediction quality of deterioration models (Tran, 2007· Ana, 2009· Ens, 2012). It tests the null hypothesis that the observed frequency matches the expected (predicted) frequency (Micevski et al., 2002). The chi-square is calculated as

$$\chi^2 = \Sigma (O_i - E_i)^2 / E_i ,$$

Where Oi is the observed frequency and Ei the expected frequency of pipes identified in each condition i (Tran, 2007). The chi-square is compared to a critical value for 95% confidence level and two degrees of freedom. If the chi-square is higher than the critical value, the null hypothesis is rejected. That means that the observed and predicted values differ.

Confusion Matrix

The confusion matrix is a simple method, especially useful when the model outputs discrete values, e.g. sewer pipelines' condition classes (Tran, 2007· Ana, 2009). It demonstrates directly the number of accurate predictions for each category (Table 3.1). The cells marked with green colour show the accurate predictions (pipes predicted and inspected in the same condition). The red cells represent the number of pipes that were predicted in a better condition than they actually were (underestimated). Lastly, the yellow cells represent the number of pipes that were predicted in a worse condition than they actually were (overestimated).

OSR and FNR values

The OSR (Overall Success Rate) and FNR (False Negative Rate) can be calculated from the confusion matrix (Table 3.1) using the following equations:

$$OSR = (V_{11} + V_{22} + V_{33} + V_{44}) / (Sum1 + Sum2 + Sum3 + Sum4)$$

$$FNR = (V_{31} + V_{32}) / (V_{31} + V_{32} + V_{33})$$

The OSR indicates the model's accuracy to predict the conditions of individual pipes while the FNR represents the risk of using the model (Tran, 2007).

Table 3.1: Confusion matrix

Predicted / Observed	**1**	**2**	**3**	**4**	Sum
1	V11	V12	V13	V14	Sum1
2	V21	V22	V23	V24	Sum2
3	V31	V32	V33	V34	Sum3
4	V41	V42	V43	V44	Sum4

Performance Curve

Le Gat (2008) developed the sewer deterioration model GompitZ, as well as a method to assess its performance. This method consists of a curve indicating the amount of sewer pipelines in the worst condition that the model was able to identify within each percentage of the worst predicted cases (Figure 3.4).

The x-axis represents the total percentage of pipelines, ordered from the worst to the best-predicted condition, while the y-axis shows the cumulated percentage of the sewer pipelines inspected in the worst condition. That means, for every correct prediction, the curve moves up vertically until all the sewers inspected in the worst condition have been identified (the curve reaches 1.0 or 100%).

The identity line represents a model with no predictive value. The curve shows graphically what is the advantage in identifying the most deteriorated sewer pipelines when using a deterioration model, compared to a random selection of pipes (no-model strategy) (Figure 3.4).

Sewer ageing profile

The sewer ageing profile is a graphical assessment method (Ana, 2009· Ugarelli et al., 2013). It is constructed by plotting the age of the sewer (x-axis) against the predicted and inspected conditions for each time point in the future (y-axis). The sewer ageing profile is a generalisation of the confusion matrix. It shows an overview of the accurately predicted condition states. Though, instead of one specific moment in the future, it shows the evolution of the system's condition in the future (Ana, 2009).

The sewer ageing profile is also a graphical indicator of the model's goodness-of-fit. The closer the predicted conditions are to the inspected ones for each period of time, the better the performance of the model.

3.3.2 Application and results

A lot of studies have attempted to determine the accuracy and applicability of sewer deterioration models. Some of them focused on the evaluation of only one model, while others attempted a comparison in order to determine the most suitable one, for different kinds of approaches. The assessment methods that were used as well as the results and conclusions of three case studies are presented in the following.

Case study in Dandenong, Australia (2007)

Tran (2007) tested the capability of five deterioration models to predict both the structural and the hydraulic condition of the system, for a case study in the City of Greater Dandenong, Australia. The sample consisted of 417 pipelines, including data for their structural and hydraulic condition as well as a variety of deterioration factors (pipe size, age, location). Three statistical models (Markov, multiple discriminant and ordered probit analysis), and two artificial intelligence models (neural network and probabilistic neural network) were under examination.

A goodness-of-fit test was implemented to evaluate the capability of all five models to predict condition changes in pipe cohorts (Table 3.2). Next, Tran (2007) assessed the model performance for individual pipes (the Markov model was excluded from that comparison), based on the overall success (OSR) and false negative rate (FNR). The models were tested in order to determine their accuracy predicting the structural and hydraulic condition of the pipelines.

Results lead to the same conclusions; the Markov model had the lowest chi-square value and was therefore the best model predicting condition changes of pipe populations. Artificial intelligence models were also proven suitable, as they passed the goodness-of-fit test (Table 3.2).

Table 3.2: Chi-square values calculated while testing different models (adapted by Tran, 2007)

Deterioration models	Chi-square values χ^2 ($\leq \chi^2_{0.05}$=5.99)			
	Structural models		Hydraulic models	
	Calibration dataset	Validation dataset	Calibration dataset	Validation dataset
Markov Model	0.22-1.38	0.34-1.53	0.06-0.16	0.09-0.21
MDDM	12.9	14.5	6.01	6.92
OPDM	7.21	7.35	5.89	6.36
NNDM	2.13-2.95	2.57-4.16	2.24-5.59	2.55-5.89
PNNDM	1.97	4.21	2.17	3.34

Concerning their accuracy predicting the deterioration of individual pipes, the two artificial intelligence models were found to be the most accurate, as they had the highest OSR and lowest FNR values (Figures 3.2 and 3.3).

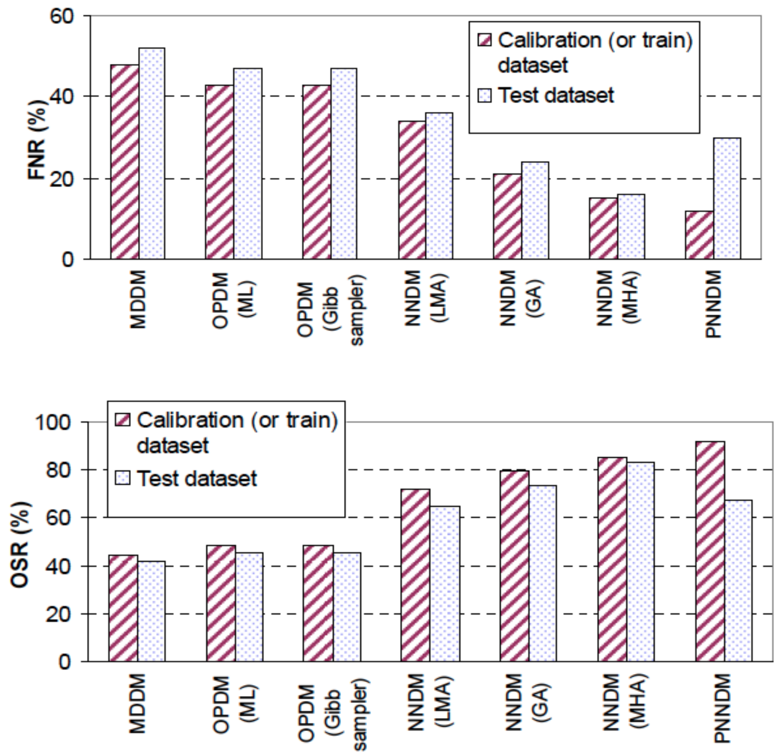


Figure 3.2: FNR and OSR rates for structural models (Source: Tran, 2007)

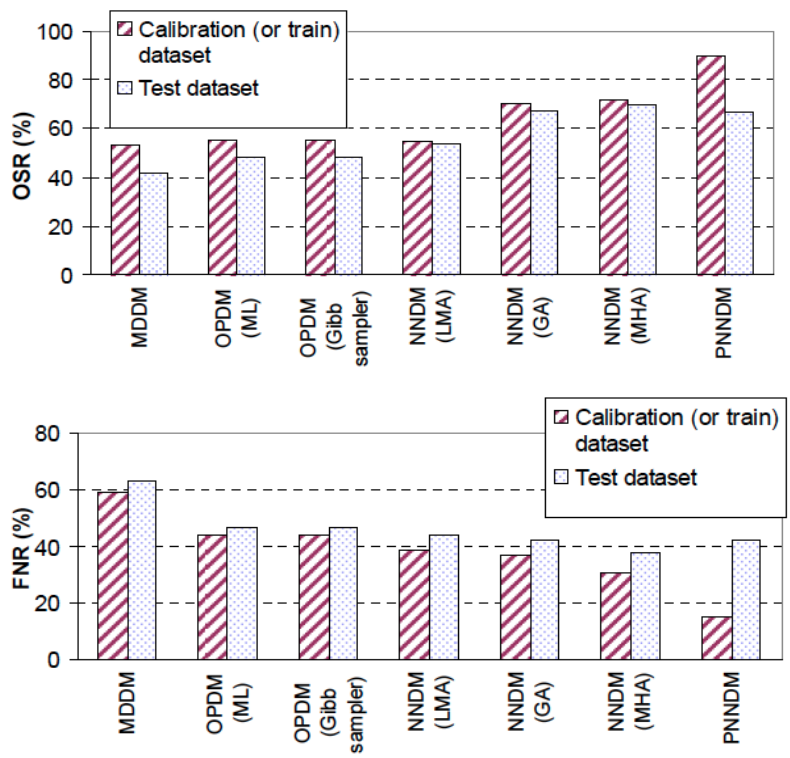


Figure 3.3: FNR and OSR rates for hydraulic models (Source: Tran, 2007)

Case study in Dresden, Germany (2008)

Le Gat (2008) used an extensive dataset acquired from the city of Dresden, Germany, in order to test the prediction accuracy of GompitZ model. Results are presented only for concrete sewers, corresponding to 21,966 pipes, inspected between 1997 and 2003. The diameter, the type of effluent, and the installation period were considered as risk factors associated with the deterioration.

Results highlight the potential of the model identifying the most deteriorated sewers. The performance curve, when compared to the identity line (representing the shape of the curve in case no model is used) indicates a significant increase in the percentage of the worst inspected cases that are identified within the worst predicted cases.

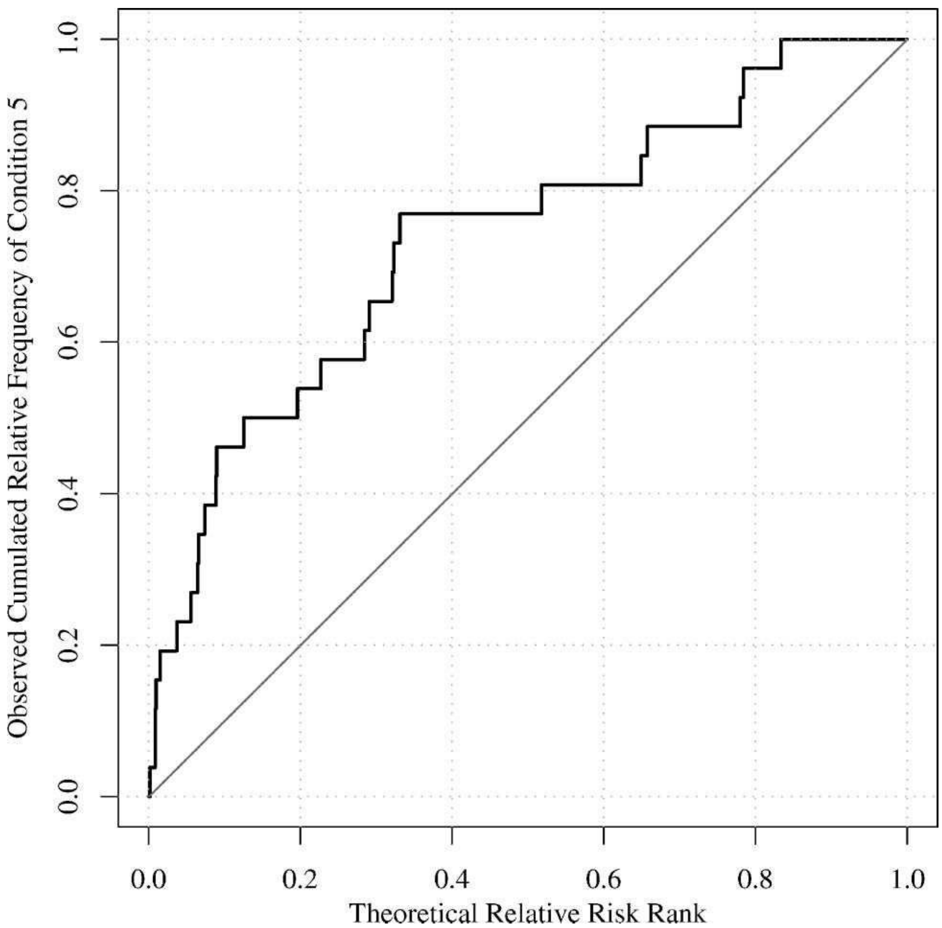


Figure 3.4: Performance curve for a case study in Dresden (Source: Le Gat, 2008)

Case study in Oslo, Norway (2013)

Ugarelli et al. (2013) also tested the accuracy of the deterioration model GompitZ (Le Gat, 2008), for a case study in Norway, Oslo. The dataset he used for the test consisted solely of concrete pipes with a diameter of less than 600mm. In order to compare the model's predictions with the inspected values obtained by CCTV cameras, they used three sewer-ageing profiles (Figure 3.5).

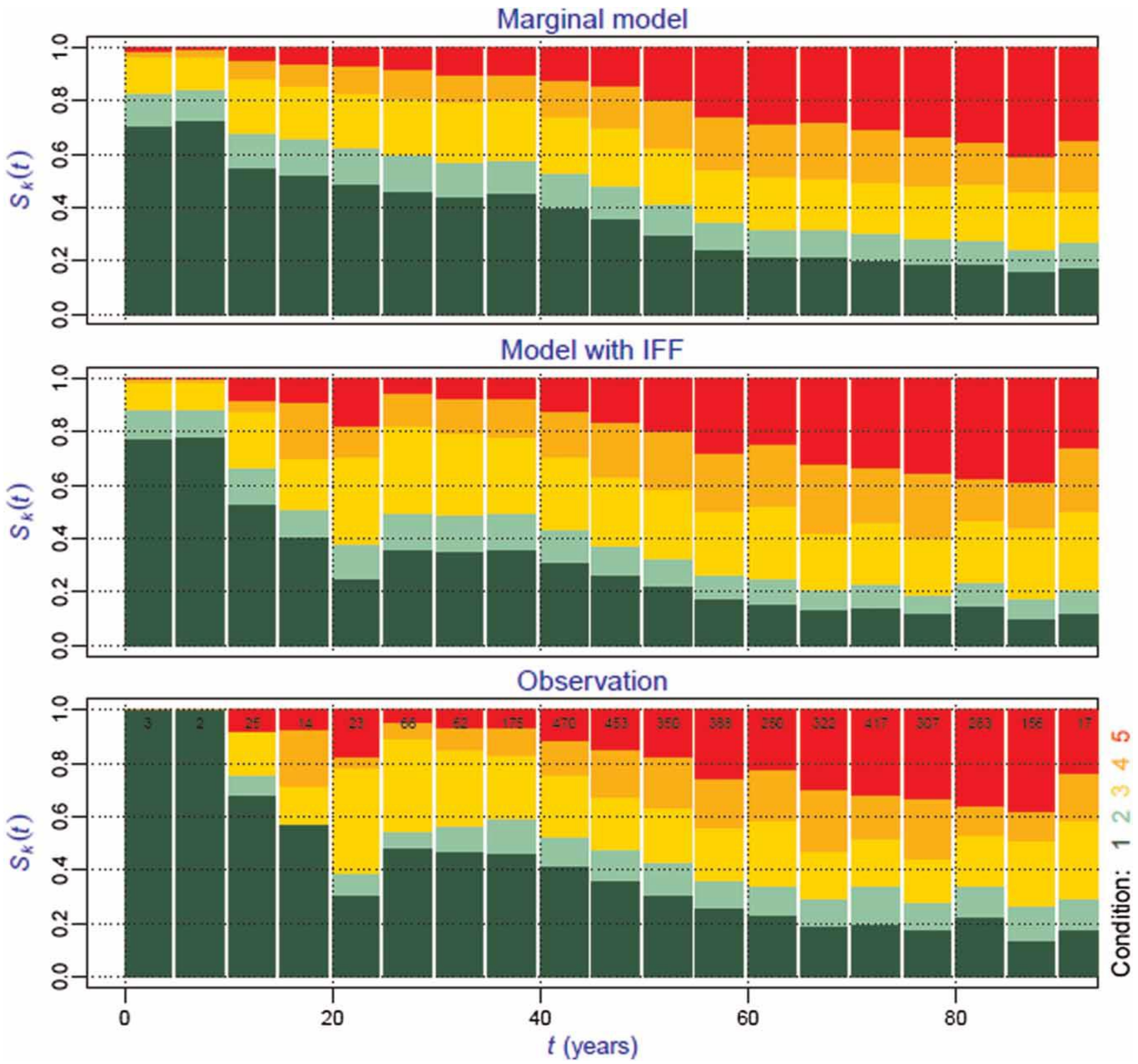


Figure 3.5: Three sewer ageing profiles for a case study in Oslo (Source: Ugarelli, 2013)

The first one represents the predictions of the marginal model, i.e. a model that calculates the probabilities of a sewer pipeline to be in each condition based on the average transition time from one condition state to the next worse one (calculated during the calibration of the model). The second model (model with IFF) takes also into account the random events, represented by an individual frailty factor (IFF) (Le Gat, 2008). The third model represents the sewer ageing profile as it was recorded during the inspections.

Results are available in Figure 3.5. The model considering the IFF (Model with IFF) produces results very similar to the inspected ones (Observation), demonstrating the model's power to accurately predict the future condition of drainage pipelines.

4 A German Case Study: Collection and Process of Data

Within the frame of this study, the performance of GompitZ model was tested and evaluated using an extensive sewer and CCTV dataset from a German city. This city is an ideal case study, since the entire sewer system (43,000 pipes/ 1,700 km) has already been inspected once and over 50% has been inspected at least twice since 1998. This chapter presents an overview of the collection as well as the handling of the available data, along with analysis and descriptive statistics.

4.1 Collection of Data

4.1.1 Overview

The raw data consists of 43,687 pipes with a total length of 1,766 km. After the evaluation of the available dataset, the model testing was finally executed on 31,394 classified inspections. The inspections were performed in accordance to local regulations, which require the entire system to be inspected every ten years. No specific rule or criteria has been applied to define the inspection programs, therefore the dataset consists of random inspections.

The sewers were inspected using either classic, or Panoramo CCTV cameras and defects were identified and coded by an operator according to a German standard (ATV, 1999). The German codes were translated to the European standard 13508-2 and classified by KWB/OEWA and Veolia Eau using the French methodology RERAU (Le Gauffre et al., 2004). The classification focuses only on the structural condition and attributes a score from 1 to 4 with 4 being the worst condition (immediate rehabilitation is needed).

4.1.2 Descriptive statistics

The main characteristics of the pipes in this case study can be seen in Figure 4.1. The sewers are mainly constructed of two materials, clay and concrete. The vast majority are either sanitary or storm water pipes. As most of them (71%) are constructed under a minor road, they are not under a great amount of pressure from the overlying traffic. Their slopes appear to be mostly slack, with 76% having a slope of less than 1% and 19% having a slope between 1 and 5%. Only half of the pipes are installed at an average depth (2.5-5m), while 46% lie at a depth of less than 2m. Almost the whole dataset consists of circular pipes (98%), while only 2% are other shapes (ovoid, arched, rectangular, or with no specific shape). The majority of them have a diameter of up to 500mm and a length of up to 80m with around 40% being between 40 and 60m long.

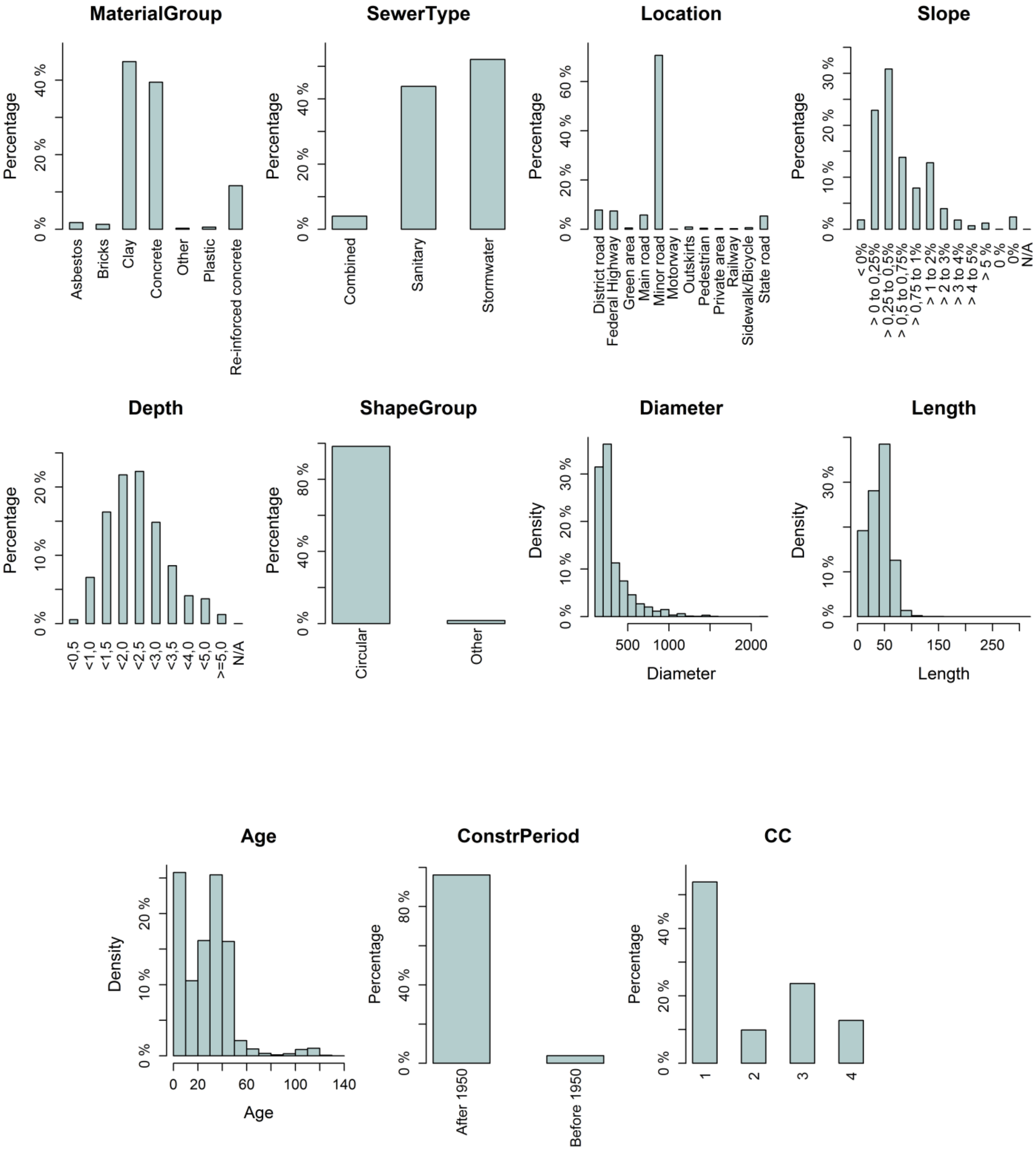


Figure 4.1: Distribution of pipe characteristics within the sewer system.

Regarding their age, the dataset consists mostly of young pipes, 96% of which were constructed after 1950. Most of them were not more than 50 years old during their inspection, while almost 25% were less than 10 years old. As a result, more than half of the sewer pipelines (55%) were found to be in a very good condition (condition class 1), while 10% were found in a good working condition (condition class 2). Only 35% of the sewer pipelines required short-term rehabilitation measures (condition classes 3 and 4).

In order to identify patterns associated with the most deteriorated pipes, a distribution of sewer characteristics is also provided for the pipelines that were identified in the worst condition (condition 4) (Figure 4.2). Observing a higher density of specific characteristics among the most deteriorated pipes would imply a relation between these characteristics and the pipe degradation.

As expected, a relation between the age of the sewer and the severity of its condition is immediately visible. In addition, concrete storm water pipes seem to be more deteriorated and thus have a much greater share of the most deteriorated cases (Figure 4.2). The percentage of clay sanitary pipes on the other hand, drop significantly.

Other than that, no strong patterns are observed and therefore no other relations can be assumed between pipe characteristics and condition class. The first indication that material and sewer type have an effect on the degradation process is yet to be demonstrated by further results.

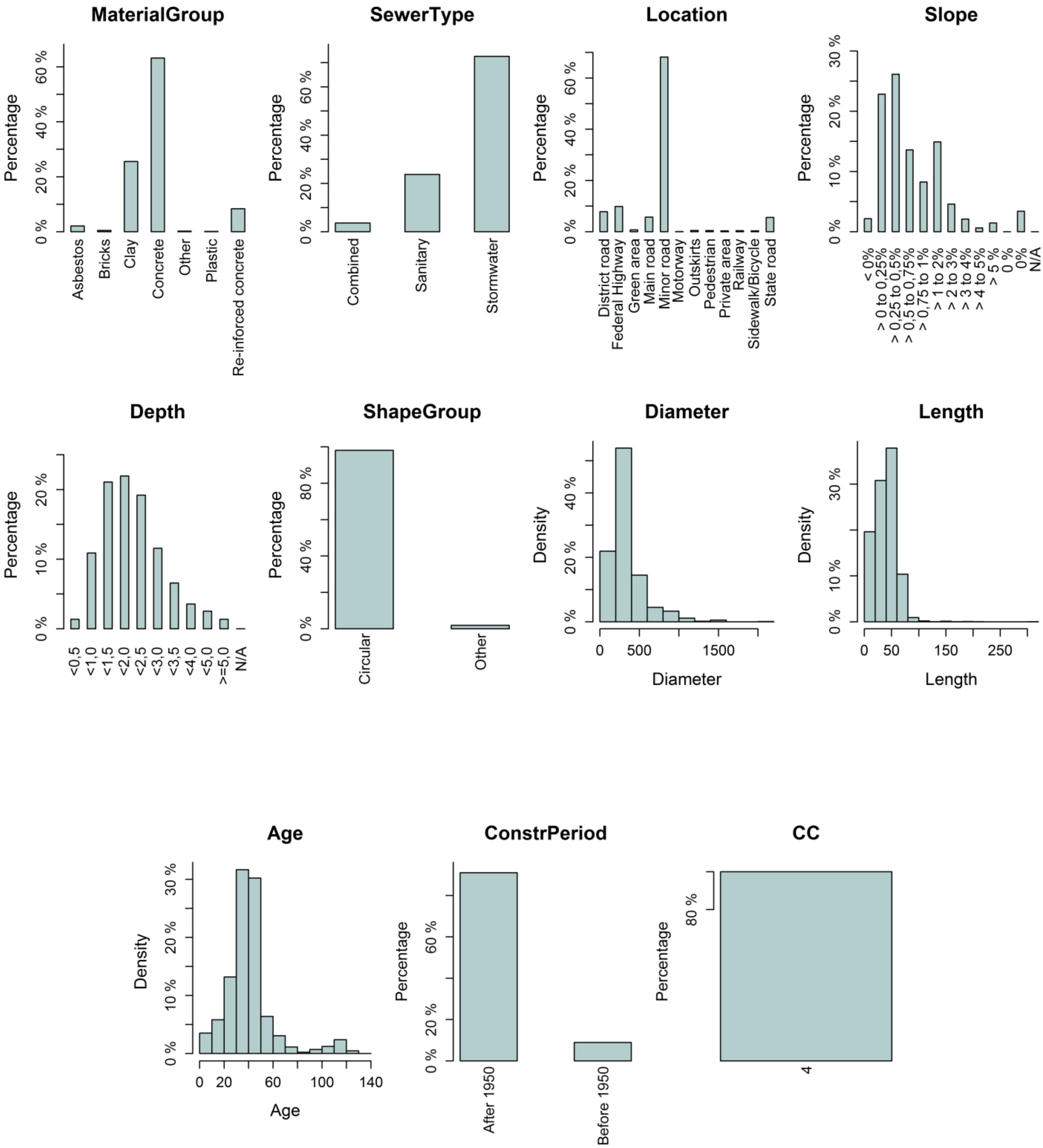


Figure 4.2: Distribution of pipe characteristics within all the sewers in condition 4.

In order to better understand the patterns that occur for different materials, the sewer type, age, and condition class distribution is plotted for clay (Figure 4.3) and concrete (Figure 4.4) pipes separately.

As usual, material and sewer type are strongly correlated to each other. Almost 90% of the clay pipes are sanitary with a smaller percentage being combined. Concrete on the other hand, is the dominating material used in storm water sewers (Figures 4.3 and 4.4).

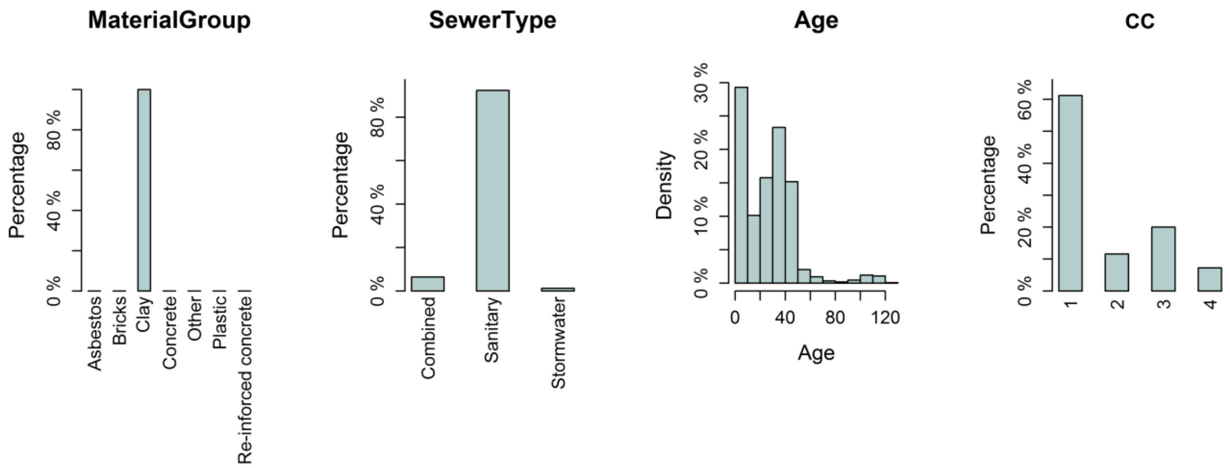


Figure 4.3: Distribution of sewer characteristics within clay pipes

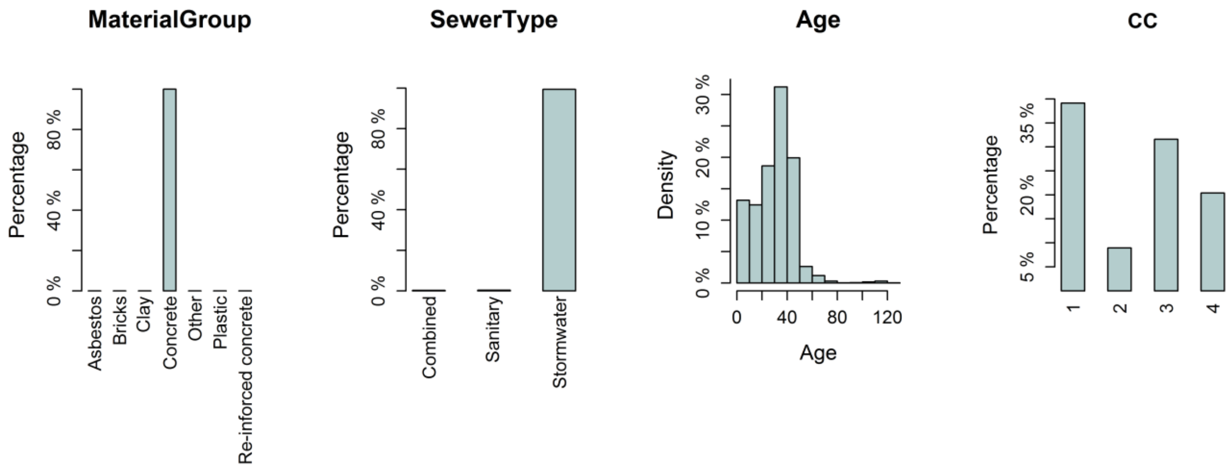


Figure 4.4: Distribution of sewer characteristics within concrete pipes

Clay sanitary pipes have a higher density of assets up to 10 years old (almost 30%), as well as a small percentage of very old sewers (100-120 years old). Concrete storm water pipes are generally older, with 70% being between 30 and 50 years old, whereas these ages cover only 54% of clay pipes. This high percentage of young clay pipes (<10 years old) might be responsible for the fact that concrete pipes are in a much more deteriorated condition; more than 50% of them were inspected in condition 3 or 4 and therefore require short-term rehabilitation. The majority of clay pipes on the other hand (70%) were inspected in a satisfactory condition (condition class 1 or 2).

4.2 Data Processing

In order to facilitate the handling of data and enable model testing, the dataset was recorded in a spreadsheet and imported to R software.

4.2.1 Overview

During the first run of data processing the following pipes were excluded:

- Pipes no longer in use
- Pressurised pipes with a missing inspection
- Pipes with a diameter smaller than 150mm
- Pipes with a missing installation year

After that initial evaluation, 33,368 pipes remained in the dataset, with a length of 1,302.9 km. The data was further processed by KWB/OEWA and Veolia Eau. During the classification of the sewers:

- 12,155 inspections were found to have inconsistent inspection codes
- 3,109 inspections could not be linked to an active sewer
- 1,708 inspections had a negative inspection date
- 689 inspections did not correspond to a real inspection

All the previous mentioned inspections were also excluded from the analysis dataset. After the final processing, the dataset consisted of 37,658 classified inspections, corresponding to 27,228 active pipelines.

As it was mentioned above, some sewer characteristics corresponded to very small populations (e.g. pipe shape, construction period). This data was difficult to handle and utilise, since it does not form a representative part of the population. Therefore their use in statistical tests was questionable and their interpretation among the results could be problematic. For these reasons, the following data was excluded from further testing:

- Sewers made of asbestos, bricks, reinforced concrete, plastic and others. Only clay and concrete sewers are included in the dataset
- Overflow sewers
- Non-circular sewers

Finally, the quality of prediction was assessed using a sample of 31,394 inspections.

Pipes that were built before 1950 were included in the dataset but excluded from any statistical and modelling tests since their population (<5%) was not considered statistically significant (96% of the sewer pipelines were built after 1950).

4.2.2 Grouping of data

In order to facilitate the handling of data and minimise the processing time within R software, a grouping of data characteristics was also applied (Figure 4.5):

- Slope: three new categories were created, Slack slopes (<0%, 0% to 0,25%), Medium slopes (>0,25 to 0,5%) and Steep slopes (>0,5%).
- Depth: similarly, three new categories were created for depth, Shallow (<2,0 m), Medium (2,0 to 2,5 m) and Deep (>2,5 m).
- Location: due to the high number of location categories that contain a rather small sample of values, the location sample was further divided into four new groups, according to the level of the traffic intensity of the overlying road: High Traffic (motorway, federal highway, railway), Medium Traffic (main road, state road), Low Traffic (minor road, district road) and No Traffic (outskirts, pedestrian, green area, private area, sidewalk/bicycle).

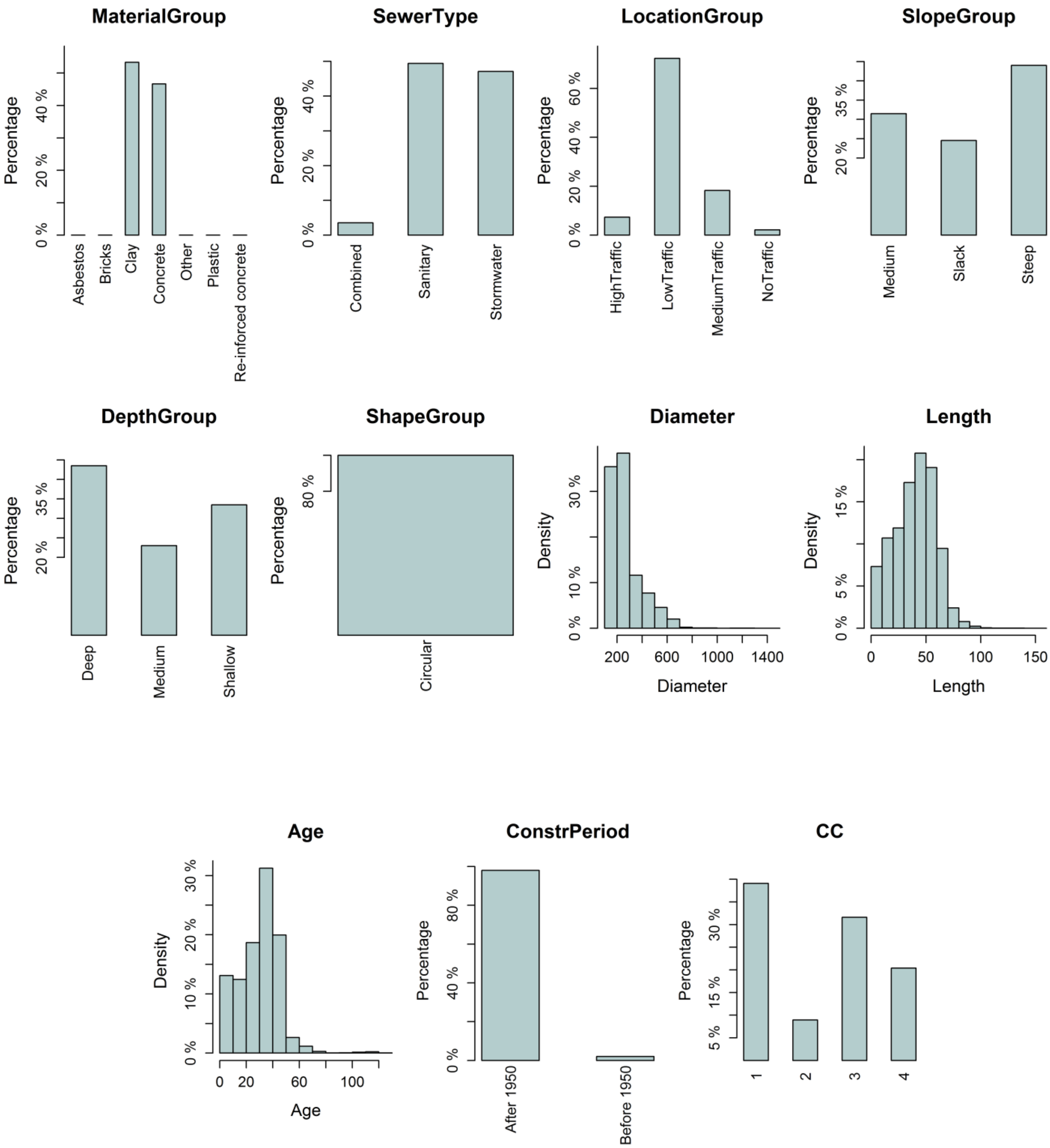


Figure 4.5: Distribution of the pipe characteristics after the grouping of data.

4.2.3 Descriptive statistics

Since the age is the most decisive factor regarding the sewers' condition, the relation between age and condition class (CC) was further investigated. Figures 4.6 and 4.7 show the distribution of the system's condition classes by age of the corresponding components.

Figure 4.6 demonstrates the distribution of conditions within the network that correspond to every age and pipe length. There is a significant amount of newly constructed sewers in the dataset (>5km) that are, as expected in an excellent condition. As the sewers get older, their condition becomes increasingly worse.

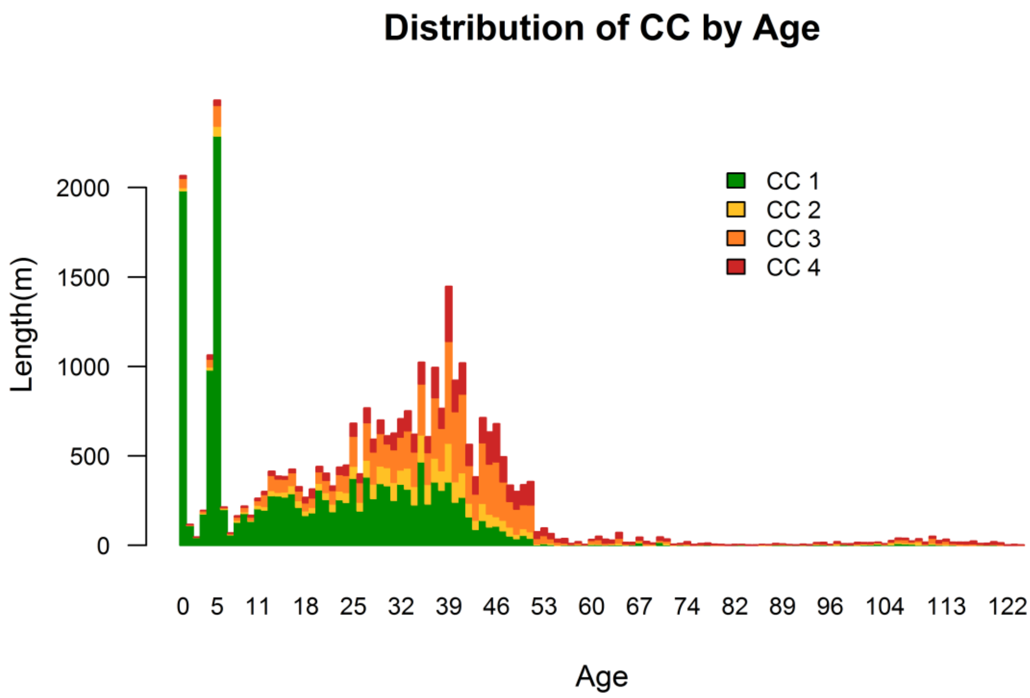


Figure 4.6: Distribution of the sewer system's condition classes by age and length.

Figure 4.7 allows a closer look into the previous figure. Instead of the pipe length, the percentage of sewers in each condition is now demonstrated by their age. While results seem reasonable for sewers up to 55 years old, there is a surprising increase of sewers in condition 1 among the oldest sewers, especially the ones between 80 and 110 years old. That can only be explained by the assumption that these sewers were replaced during their lifetime but their replacement was not registered. According to Figure 4.6, there are very few sewers aged more than 51 years old, and therefore Figure 4.7 alone, could be slightly misleading, since where it appears to represent the majority of sewers (e.g. all of the sewers aged 90 years old are in condition 1), that might in fact correspond to a single pipeline.

As it was already pointed out, there is a great amount of uncertainties inserted into the problem. Even if the condition class is accurately recorded, and the age of the sewer is stated correctly, if the sewer has been replaced and that fact was not registered in the dataset, the corresponding entry is false and has a negative effect on the final result. Since errors and uncertainties are an undeniable fact when the human factor interferes with the data, these mistakes will be accepted in this thesis as a normal occurrence that is present in most datasets that the model will be applied on.

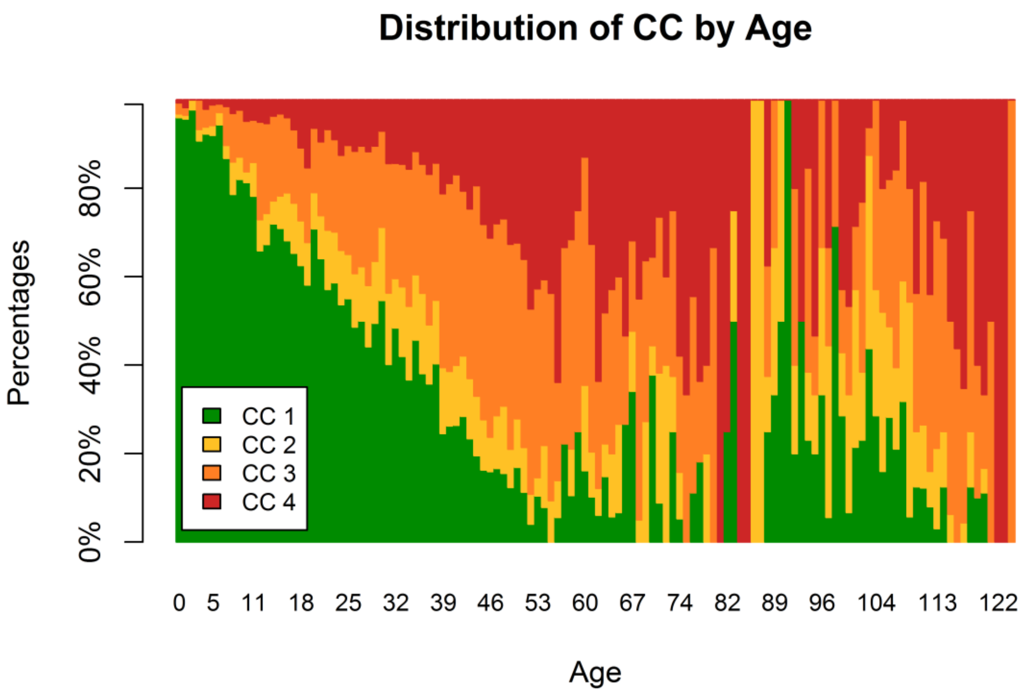


Figure 4.7: Distribution of the sewers' condition classes by their age and percentages.

Finally, the distribution of materials was plotted against their installation year (Figure 4.8). The construction of the sewerage system, as it appears in Figure 4.8, started slowly after the end of World War II, in 1945, followed by a boom in construction ten years later, in 1955, which lasted till 1982. Construction slowed down after this point, only to start growing again after 1990 and until 2000.

Traditionally, the material that was used for the few sewers that were built before 1930 was clay. Although concrete pipes never seem to have dominated clay, the construction of pipes for the next 70 years, till 2000, is almost equally shared by clay and concrete. Clay is used again almost exclusively for pipes that were built after 2000 (Figure 4.8). These results explain what was already noted above; although the two materials share a quite similar age distribution, clay pipes have a small percentage of very old components, as well as a large amount of newly built ones.

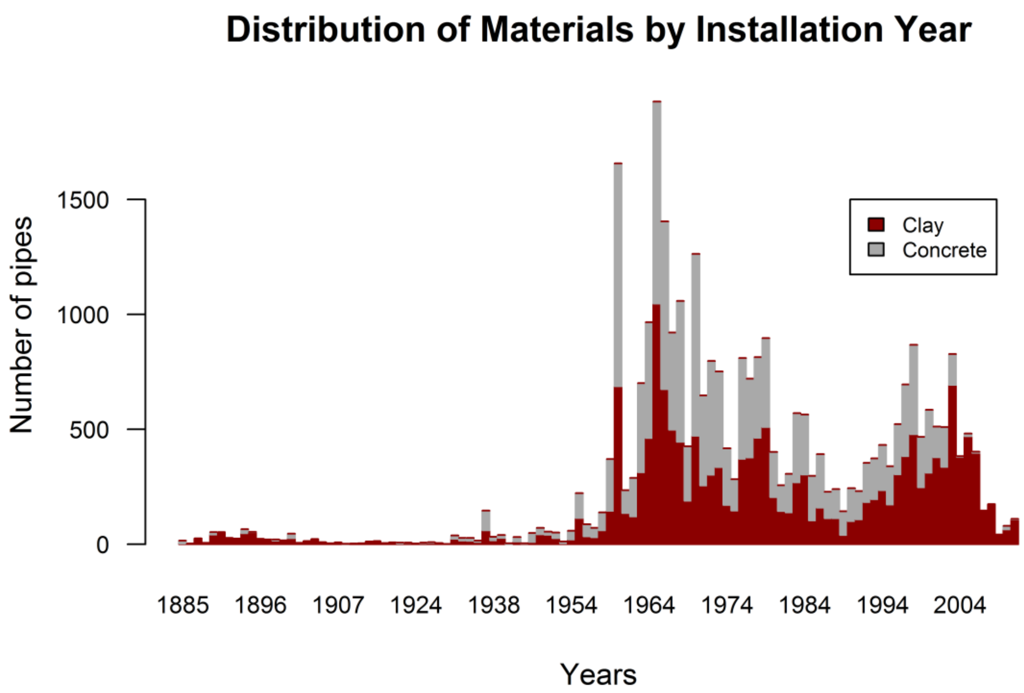


Figure 4.8: Number of clay and concrete pipes installed from 1885 to 2011.

5 Methodology: Modelling and Performance Assessment

In order to answer the questions described in chapter 1, a deterioration tool was implemented in R based on the GompitZ model (Le Gat, 2008). This chapter presents the different steps of the methodology that was developed in order to run the model and assess its performance, for different configurations of input data.

5.1 Overview

R is an open-source programming language that is widely used within the field of statistical computing. GompitZ was implemented in R, along with other steps for data preparation and model performance assessment (Figure 5.1).

The methodology is based on an exchange of information between the modules developed in R and GompitZ. R was used in order to simplify and automate the procedure of handling the input and output of the model. The automation of this procedure as well as the efficient cooperation between a deterioration model and a programming language can extend the potential of deterioration models, as well as minimise processing time.

The main steps of the model calibration and prediction are performed by GompitZ, whereas the rest of the steps, concerning data preparation and handling, as well as model performance assessment are executed within R. The execution is based on two main procedures: (i) the model calibration, which calculates the parameters of the model with respect to the provided input data and (ii) the validation of the results, which determines the model's efficiency by comparing the predicted and inspected condition of the sewers.

As a first step, all the input data is imported and manipulated in R software. The basic information required by the model is the age and the condition of the sewer. However, a set of deterioration factors can also be implemented as explanatory covariates in the model, in order to improve the quality of prediction (Figure 5.1).

Next, R calls GompitZ to run the calibration. The parameters of the model are calculated and returned to R software. In this step, the calibration results can be further handled in order to perform modelling tests or become directly available to the prediction module of GompitZ.

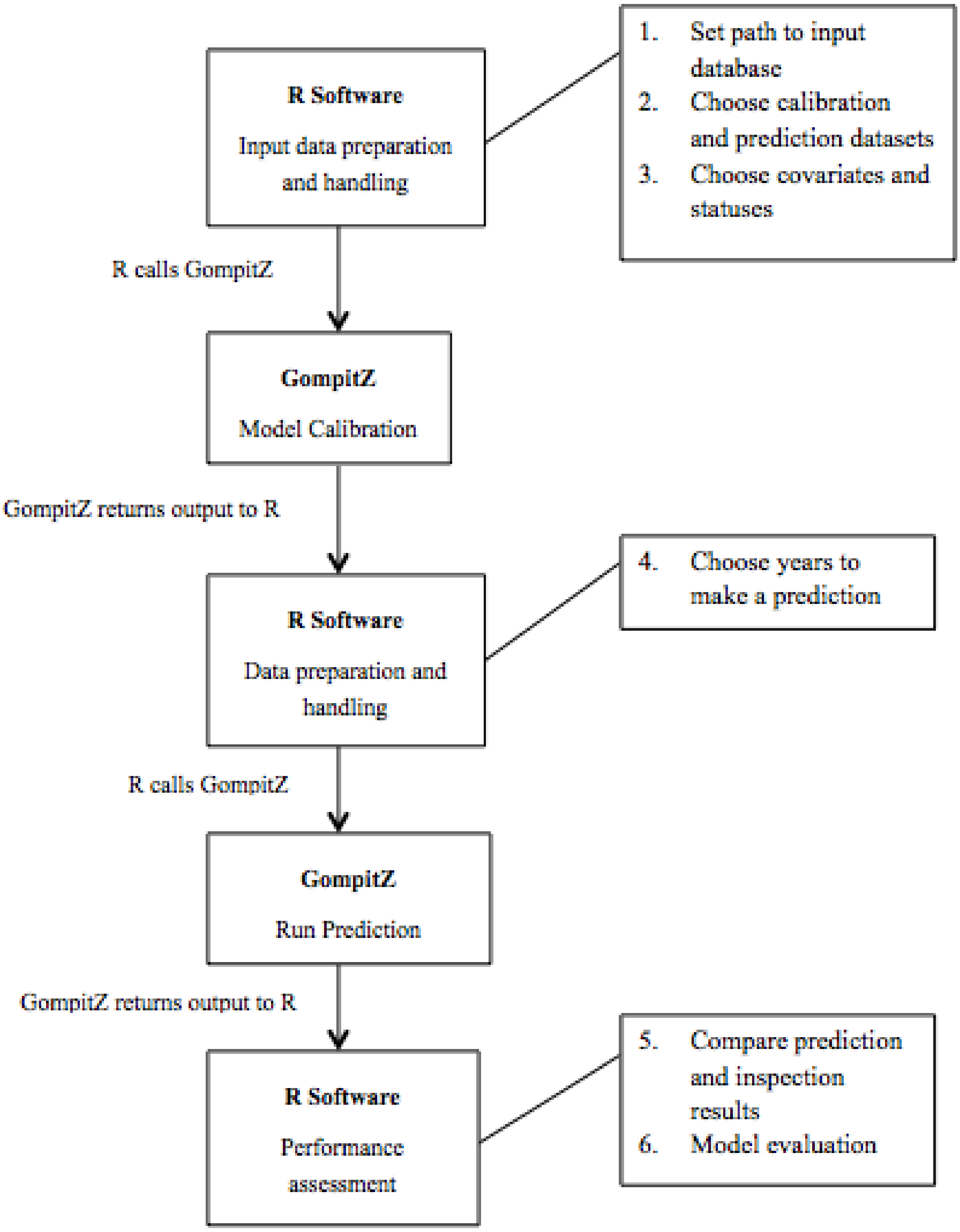


Figure 5.1: Development of the deterioration model in R software.

The prediction accuracy of the model is evaluated by comparing the results with real data, i.e. observed values. Two different selections of calibration and validation datasets were tested; a chronological selection (assigns the oldest inspections to the calibration dataset, while using the new ones to validate the model), and a Monte Carlo simulation, drawing random selections of calibration and validation datasets.

The range of years, as well as the sewers, for which a prediction is required are defined in this step. Since the range of inspections occured between 1998 and 2013, a prediction is made for these years, in order to enable the comparison between the predicted and inspected condition. Next, R calls GompitZ to run the prediction module. Finally, the results are once again returned to R in order to perform the model evaluation, based on the selected method.

5.2 Model validation

The validation process aims to demonstrate if the model has a satisfactory range of accuracy consistent with its intended application (Schlesinger et al., 1979· Sargent, 1998).

Since short-term rehabilitation programs focus on repair and replacement of pipes in poor conditions, the model has been assessed according to its ability to identify the most deteriorated pipes. A typical model application would be the simulation of non-inspected pipes for a city in which only part of the system has been inspected. Municipalities can use the results to focus future inspections on the pipes predicted in poor condition and support the planning of rehabilitation programs. Two methodologies were used to demonstrate the model's performance according to this objective.

5.2.1 Performance curve 1

Le Gat (2008) introduced a performance curve to determine the percentage of pipes in the worst inspected condition that the model manages to predict within the worst predicted cases. In the example demonstrated in Figure 5.2, within the x% of the worst predicted cases (worst prediction index) the model manages to identify three pipes inspected in condition 4 out of a total of four pipes in condition 4, existing within the exemplary dataset (75%).

	Sewer ID	PredIndex	CC
	21	3.00	4
x%	63	2.77	3
	34	2.26	4
	94	2.20	4
	101	1.95	2
	52	1.80	4
	13	1.58	2
	5	1.46	3
	42	1.14	3
	78	1.00	3

Figure 5.2: Sewer pipes ranked by improving predicted condition (reducing PredIndex), along with their corresponding inspected condition class (CC). In the example demonstrated in Figure 5.2, within the x% of the worst predicted cases, 75% of all the pipes in condition 4 are identified.

Performance curve 1 is produced as follows:

- Step 1: the model is calibrated using the corresponding dataset. The model is then applied against the pipes of the validation dataset in order to make a prediction for the year of their inspection.
- Step 2: the pipes are ordered from the most severely to the least deteriorated in respect to their prediction index. The rank of each pipe is then divided by the total number of pipes in the validation dataset, resulting in a relative rank number.
- Step 3: for each rank, the inspected condition classes are analysed: the cumulative number of pipes inspected in the worst condition (in this case condition 4), identified by the model, is divided by the total number of pipes in the validation dataset that were inspected in condition 4. Results indicate the percentage of pipes in condition 4 that were identified within the corresponding percentage.
- Step 4: the relative percentage of pipes inspected in condition 4 (Step 3) is plotted against the relative rank of the pipe's predicted condition (Step 2).

The quality of the model's predictions can be demonstrated by performance curve 1, in close examination to the corresponding ideal (best case), and random (worst case) curve (Figure 5.3). The ideal curve represents the shape of performance curve 1 for the case of a perfect prediction, whereas the random curve demonstrates the obtained results in case of a random pipe order, i.e. a model with no predictive value. The further the curve moves from the random curve (identity line), the more efficient the model. On the other hand, the closer it is to the ideal curve, the better.

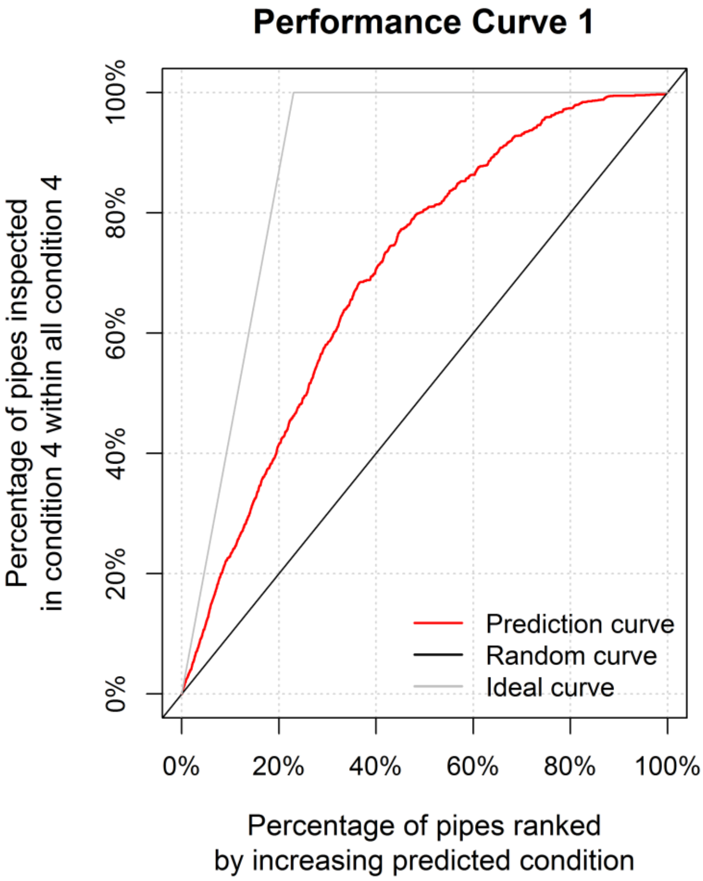


Figure 5.3: Example of performance curve 1 for an exemplary dataset with 23% of pipes in condition class 4.

In the example presented in Figure 5.3, 23% of the pipes in the validation dataset were inspected in condition 4. Therefore, a perfect model would identify all pipes in condition 4 within 23% of the worst predicted cases (Figure 5.3, Ideal curve). On the other hand, a model with no predictive value would identify 23% of all pipes in condition 4 within 23% of the worst predicted cases (Figure 5.2, Random curve). Since the amount of pipes in the worst condition defines the curve's shape, conclusions can only be made by comparing the plotted curve to the ideal curve for each individual case.

Although performance curve 1 is a strong indicator of the model's performance, it lacks practical value. In most cases, municipalities have a restrained budget at their disposal when maintaining the city's drainage infrastructure and can rehabilitate only a limited amount of defective sewers. Therefore they are not necessarily interested in identifying and repairing every pipe in poor condition but only the most deteriorated ones.

In order to plan and perform inspections in the most deteriorated sewers, the model should be able to detect as many pipes in poor condition as possible within a small percentage of the pipes predicted in the worst conditions, i.e. the pipes predicted with the worst prediction indexes. In order to fill this gap, a second assessment methodology was developed.

5.2.2 Performance curve 2

The second method aims to identify what percentage of the pipes predicted in the worst conditions were inspected in poor condition. In the example demonstrated in Figure 5.4, within the x% of the worst predicted cases (worst prediction index), the model manages to identify three (out of four corresponding to x%) pipes inspected in condition 4.

	Sewer ID	PredIndex	CC
x%	21	3.00	4
	63	2.77	3
	34	2.26	4
	94	2.20	4
	101	1.95	2
	52	1.80	4
	13	1.58	2
	5	1.46	3
	42	1.14	3
	78	1.00	3

Figure 5.4: Sewer pipes ranked by improving predicted condition (reducing PredIndex), along with their corresponding inspected condition class (CC). In the example demonstrated in Figure 5.4, within the x% of the worst predicted cases, 75% of the pipes the model identifies are in condition 4 and 25% are condition 3.

Performance curve 2 is constructed similarly to performance curve 1. However, step 3 is different: for each rank, the number of pipes inspected in condition 4 is divided by the total number of pipes considered within the rank (instead of the total number of pipes inspected in condition 4) (Figure 5.5). The same procedure is applied to construct the second (orange) curve of the graph, although in this case conditions 3 and 4 are considered together. The result is a curve demonstrating the real condition distribution of the pipes predicted in the worst conditions.

In the case of a model with no predictive value (random selection), a subset containing x% of the pipes predicted in the worst conditions would have approximately the same condition distribution as the validation dataset. The curve representing a random selection of pipes is represented by a horizontal line, indicating the total percentage of sewers in condition 4 within the validation dataset. On the other hand, a perfect model would find only pipes in condition 4 within the worst predicted cases.

Considering 100% of the pipes predicted in the worst conditions, i.e. all pipes, the condition distribution indicated in the graph is the condition distribution of the validation dataset, with 23% of pipes in condition 4.

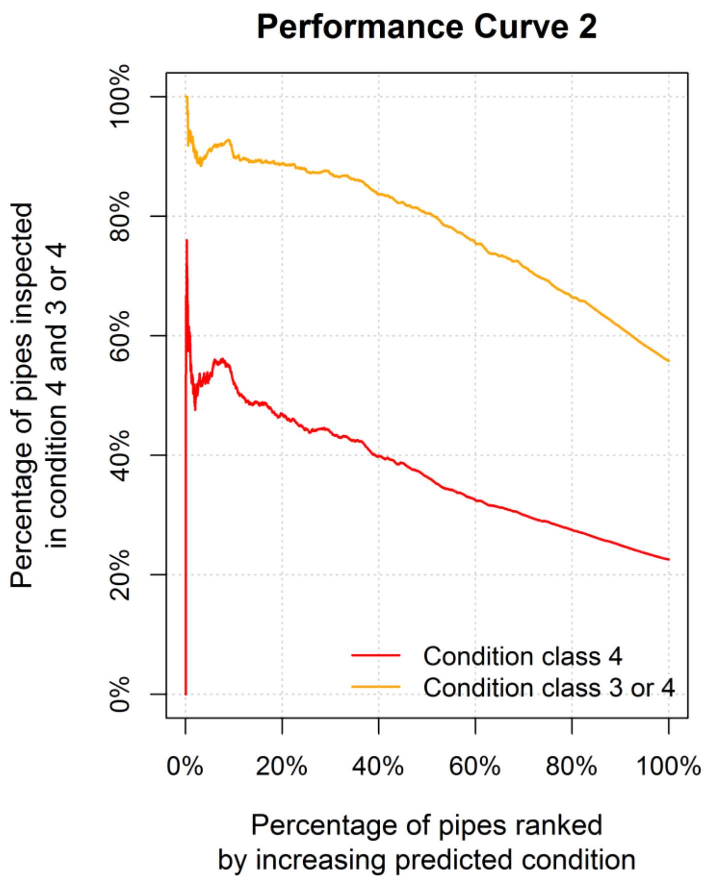


Figure 5.5: Example of performance curve 2 for an exemplary dataset with 23% of pipes in condition class 4.

5.3 Assessment of the most important covariates

In order to assess the quality of the model's predictions, it is necessary to identify the covariates that maximise that performance. Three tests were performed in order to determine the factors that have the greatest influence on the deterioration of the sewerage system, and thus have an explanatory value that can improve the accuracy of prediction:

- A correlation analysis
- A test to assess the influence of the factors on the calibration
- A test to assess the influence of the factors on the model's predictions

5.3.1 Correlation analysis

As preliminary analysis of the input data, the deterioration factors were tested for statistically significant relations (a) with the independent variable (in this case the condition class) and (b) with each other.

The problem that arises here derives from the diversity of information that is provided by the different deterioration factors, and complicates the comparison between them. Depending on the nature of their assigned values, each factor can be classified as interval or categorical:

- Factors described by interval values give a scalar indication of a characteristic. The age, the length, and the diameter of a sewer pipe are interval variables. That means that they represent measurements where the difference between two values is meaningful (e.g. the difference between 20 and 40 meters is the same as the difference between 40 and 60). The Pearson's (or Spearman's, depending on the assumptions as explained later on) correlation can be used to extract information about the relation between interval variables.
- Factors described by categorical values define grouping criteria according to the presence or lack of a specific characteristic. In this thesis, most of the deterioration factors are treated as categorical, since some interval variables were divided into groups (see Chapter 4). The material, the type of sewage, the installation depth, the slope, and the location are treated as categorical variables. The cross-table analysis is used to identify relations between the above deterioration factors.

Pearson's and Spearman's rank

The Pearson's correlation is the most common and strongest tool to identify relationships between interval variables. However, it requires the following assumptions: (a) a linear relationship between the variables, (b) lack of outliers, (c) an approximately normal distribution of variables, and (d) homoscedasticity of the data.

On the other hand, the Spearman's rank correlation coefficient is not affected by non-linearity or outliers, it does require though a monotonic relation between the variables and a general lack of duplicate values.

Spearman's test consists of a correlation matrix, indicating the strength of relationship between each pair of variables, as well as a test for significance. The correlation matrix shows if there is some statistical correlation, whereas the p-value of the significance test shows if the correlation is statistically significant at 95% confidence level.

Cross-Table analysis

Cross tabulation is a statistical process that compares categorical data by creating a comparison table. Each cell of the table includes the number of pipes that are characterised by both factors (demonstrated in the row and line that cross the cell), as well as a chi-square statistic that indicates the strength of this relationship. A p-value is produced as a result of the chi-square value combined with the degrees of freedom. The smaller the p-value the greater the correlation between the variables.

The p-value represents the probability that the observed relation between the variables is false. In most cases, a p-value of 0.05 is the acceptable limit in order to consider the assumption true. This value corresponds to a 5% probability of error.

Kruskal-Wallis one-way analysis of variance

The Kruskal–Wallis one-way analysis of variance is a non-parametric test that compares the medians of several groups and concludes if the samples originate from the same distribution. It shows if the differences between the groups are so large that they are unlikely to have occurred by chance. The dependent variable has to be at least ordinal and the independent variable has to be categorical. The parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA).

Without assuming the data to be normally distributed, it tests at 5% significance level if, for example, the age of the sewer has identical distributions for every material. The null hypothesis is that it has identical populations. To test this hypothesis, the independent data for the different materials are compared. If the p-value turns out to be zero, the null hypothesis is rejected. At 5% significance level, the age of the sewer for different materials does not form identical populations.

This test is appropriate for use under the following circumstances: (a) three or more conditions are compared, (b) each condition is performed by a different group of participants, and (c) the data does not meet the requirements for a parametric test. However, the test does assume an identically shaped and scaled distribution for each group, except for any difference in medians. If the dataset meets the requirements for a parametric test, it is better to use one-way Analysis Of Variance (ANOVA), since it is more powerful than the Kruskal-Wallis test.

5.3.2 Influence of the factors on model calibration

As previously mentioned, during the calibration of the model, GompitZ calculates the optimum coefficients for each deterioration factor in order to achieve the best model fit. In case the consideration of the covariate does not improve the fit of the survival curves describing the deterioration of the sewer pipelines, the coefficients are set to zero and the covariate is not considered in the calibration process.

In order to test if a covariate improves the model's goodness of fit, one deterioration factor was implemented at a time as a covariate in the model. Next, a prediction was made for the sewer pipelines of the validation dataset. The prediction index calculated for each sewer pipeline was finally plotted against the age (Figure 5.4).

If the covariate has an explanatory value, the curves for each characteristic (stormwater, sanitary, combined) of the same factor (sewer type) demonstrate the different deterioration behaviours defined by the different pipe characteristics. The further the curves move apart from each other, the stronger the influence of the corresponding covariate. If all the curves overly on each other, every category of this specific factor has the same deterioration speed and therefore the covariate can be neglected.

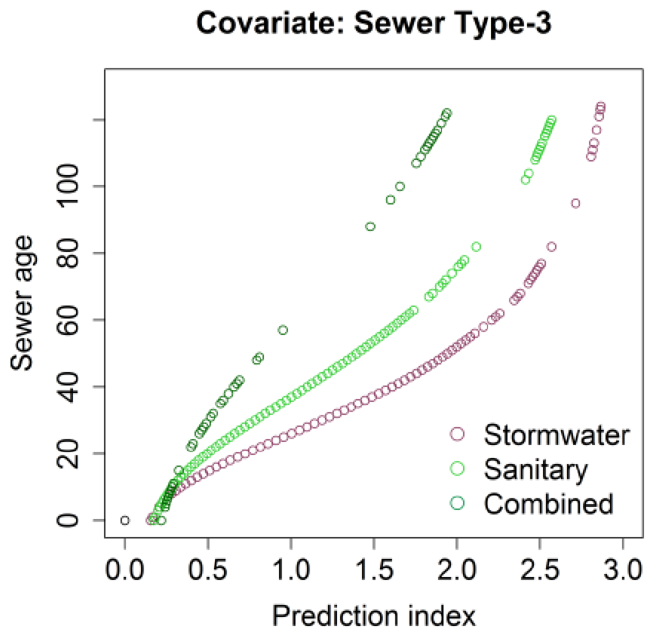


Figure 5.4: Curves indicating different deterioration patterns for different sewer types

For example, if stormwater and sanitary pipes of the same age are predicted with a different prediction index, the conclusion is that the material has an influence on the deterioration. When a curve has a steeper gradient (Figure 5.4, combined pipes), it means that the sewer age increases quickly compared to the prediction index. Therefore, a small change in the pipe's condition corresponds to a rather big change of age. On the other hand, when the curve has a very slack gradient (Figure 5.4, stromwater pipes), it shows that there are significant changes in the prediction index, for minor changes in the pipe's age.

5.3.3 Influence of the factors on the accuracy of prediction

In order to assess the influence of each covariate, a methodology was also developed based on the resulting quality of prediction. One deterioration factor was considered at a time as a covariate in the model, with a status from 0 to 3. For each of these four considerations, the two performance curves were plotted. Next, the values corresponding to 2%, 5%, 10%, 15%, 20% and 25% of the pipes predicted in the worst conditions were extracted from both performance curves and plotted in a new graph for each of the four considerations of the covariate.

For each factor, the curve that maximises the model's performance indicates the appropriate consideration of the covariate inside the model. If all the curves are close to each other, the covariate has little to no influence. On the other hand, if at least one curve shows a clear increase on the quality of prediction (lies higher than the others) the covariate will be implemented in the model with its corresponding status. In case that curve corresponds to status 0, the covariate is not implemented in the model, since its consideration does not improve the model's efficiency.

In figure 5.5, the performance curves 1 and 2 were plotted for four considerations of the sewer type as a covariate with status 0, 1, 2 and 3. The curve representing status 0 lies lower than the rest, indicating a reduced quality of prediction when the covariate is not considered in the model. On the other hand, the model's predictive value hardly changes when the covariate is considered with a status from 1 to 3 (the curves almost overly on each other).

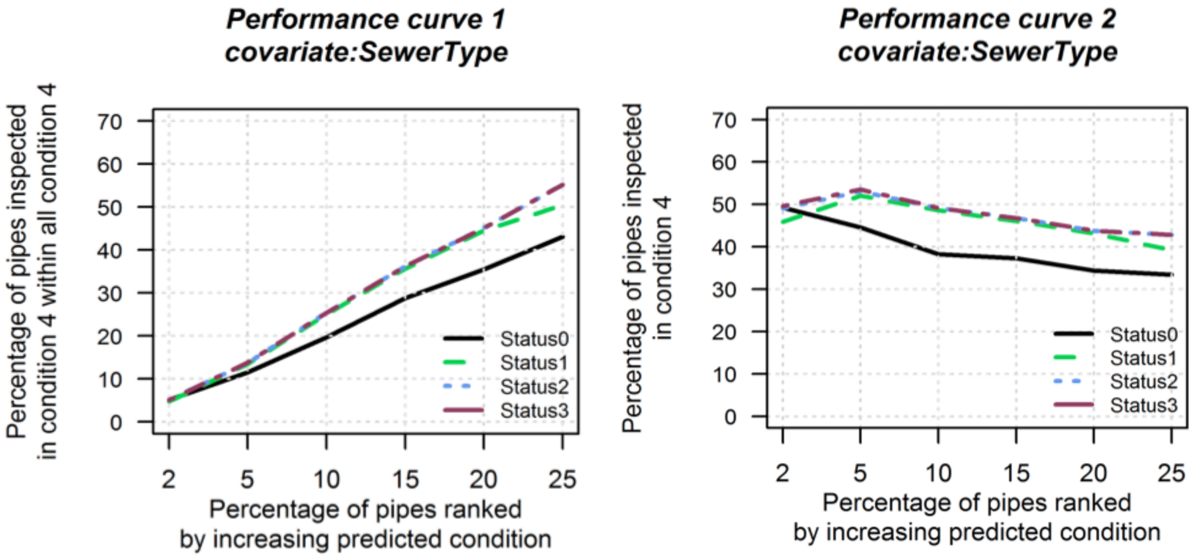


Figure 5.5: Curves indicating different prediction quality for different considerations of sewer type.

5.4 Monte Carlo sensitivity analysis

Monte Carlo is a widely used methodology in the fields of science, engineering, finance, and statistics that aims to solve a numerical problem through a large number of random experiments (Kroese D.P. et al., 2011). In the development of computational models, Monte Carlo simulations find an excellent application determining the model output uncertainty originating from input variable and parameter uncertainties (Verbeeck H. et al., 2006).

Unfortunately, this analysis is often neglected when evaluating complex computational models (Benke K.K. et al., 2007). In order to conclude on the model's general value, its dependencies on parameters that are conditional on each case study had to be carefully assessed.

When it comes to sewer deterioration modelling, the amount of data available to calibrate the model is often scarce, or bias. Most cities cannot support the inspection of the whole system and therefore focus the inspections only on the critical sewers. In the present study, the influence of the calibration dataset was investigated for different sizes and selections of input datasets, running 1000 Monte Carlo simulations.

As it was already mentioned, the data was divided into two datasets, dedicated to model calibration and prediction (70% and 30% of the data, respectively). The Monte Carlo simulations were performed in order to determine if the selection of datasets among all data influences the results. For each selected set of data, 1000 random selections of calibration and validation datasets were performed, and the corresponding prediction quality for each case was plotted in the same graph. The variation of the curves is in accordance with the level of influence that the dataset selection has on the model. A great variance among the curves indicates a very different prediction quality resulting from different selections of datasets among the same set of data. On the other hand, if the curves do not fall far from each other, it means that the model is rather robust against random data selection.

The same test was applied for 20% and 10% of all available data, in order to check the effect of data scarcity on the development of trustworthy models.

6 Results: Analysis of deterioration factors and identification of the most relevant covariates

This chapter presents the results of the study. The best combination of covariates is determined and the accuracy of the model's predictions is assessed for different parameters and sizes of input data.

6.1 Correlation analysis

Eight deterioration factors, i.e. the construction material, sewer type, diameter, length, location, depth, slope, and pipe age were examined in order to determine their correlation to each other, as well as their influence on the sewer's condition.

6.1.1 Spearman's rank

The assumptions required to perform a Pearson's correlation were tested and not fulfilled. Therefore a Spearman's rank test was used to identify correlations among the interval variables, i.e. the pipe age, length, diameter and condition class (CC). In the following test, the condition class is treated as an interval variable, although the definition of condition classes is arbitrary and does not imply that there is a solid stable difference between each condition rank. The results of the Spearman's correlation are available in Table 6.1.

The following figure indicates a strong relationship between age and condition class (Table 6.1 – Correlation Matrix) that is present (rather than due to chance) in the population (Table 6.1 – Test for Significance) and a rather non-existent correlation between the rest of the variables.

However, Spearman's test is a measure of the monotonic relationship between paired data and thus a value close to 0 does not imply a complete lack of relationship but rather a lack of monotonic relationship. When tested, the assumption of monotonicity between the variables failed to be fulfilled and thus the above results should be handled with care.

Table 6.1: Spearman's test in R. Correlation matrix and test for significance.

Spearman's Test

Correlation Matrix

	Age	CC	Diameter	Length
Age	1	0.51	-0.06	-0.02
CC	0.51	1	-0.01	0.03
Diameter	-0.06	-0.01	1	0.04
Length	-0.02	0.03	0.04	1

Test for Significance

	Age	CC	Diameter	Length
Age	0	0	0	0
CC	0	0	0.03	0
Diameter	0	0.03	0	0
Length	0	0	0	0

6.1.2 Cross-Table analysis

The cross-table analysis was used to identify relations between the sewer characteristics described by categorical variables, i.e. the construction material, sewer type, slope, depth, location and condition class (CC). The condition class is now treated as a categorical variable. The results of the cross-table analysis are available in Table 6.2. The p-value resulted from the combination of the chi-square value and the degrees of freedom (df). The smaller it is, the higher the correlation between the variables.

According to the above table, the factors that appear to have the strongest correlation with the condition class, and therefore the strongest influence on the deterioration, are by order of influence: the material, the sewer type and the depth. The strongest correlation between other factors is identified between the material and the sewer type, followed by a significantly strong relation between material and location, and material and depth (Table 6.2).

Table 6.2: Cross-Table analysis results.

Cross Table Analysis Results

Covariate1	Covariate2	chisquare	df	pvalue
MaterialGroup	MaterialGroup	161682.00	36	0.000000e+00
MaterialGroup	SewerType	28644.84	18	0.000000e+00
MaterialGroup	Location	1933.67	66	0.000000e+00
MaterialGroup	Depth	5520.95	54	0.000000e+00
MaterialGroup	CC	2032.08	18	0.000000e+00
SewerType	SewerType	80841.00	9	0.000000e+00
SewerType	Depth	5895.88	27	0.000000e+00
Location	Location	296417.00	121	0.000000e+00
Slope	Slope	269470.00	100	0.000000e+00
Depth	Depth	242523.00	81	0.000000e+00
MaterialGroup	Slope	1580.20	60	9.266639e-291
Location	Depth	1701.20	99	5.874687e-290
SewerType	Location	1200.92	33	3.864584e-231
SewerType	CC	808.12	9	3.802594e-168
Slope	Depth	1060.42	90	1.662545e-165
Location	Slope	874.75	110	2.244882e-119
Depth	CC	638.88	27	2.276080e-117
SewerType	Slope	531.97	30	3.275636e-93
Slope	CC	200.52	30	3.959878e-27
Location	CC	189.61	33	6.732112e-24

6.1.3 Kruskal-Wallis and one-way analysis of variance

Once again, the data did not meet the requirements for a parametric test (ANOVA). A Kruskal-Wallis test was used instead, in order to enable the comparison between sewer pipe characteristics described by interval and categorical variables.

A summary of the Kruskal-Wallis test is demonstrated in Table 6.3. Similarly to the cross table analysis, the result is a chi-square value that combined to the degrees of freedom, results to a p-value. According to Kruskal-Wallis, the factors with the strongest relation to the condition class are the age, the material, the sewer type and the depth. The strongest correlation between the rest of the factors is identified between material and age and material and diameter.

Table 6.3: Kruskal-Wallis analysis results.

Kruskal-Wallis Test

Covariate1	Covariate2	chisquare	df	pvalue
Age	MaterialGroup	2593.75	6	0.000000e+00
Age	CC	6627.77	3	0.000000e+00
CC	MaterialGroup	1765.31	6	0.000000e+00
CC	CC	26946.00	3	0.000000e+00
Diameter	MaterialGroup	12728.18	6	0.000000e+00
Diameter	SewerType	11927.47	3	0.000000e+00
Diameter	Location	1886.34	11	0.000000e+00
Diameter	Slope	2993.32	10	0.000000e+00
Length	Slope	2265.48	10	0.000000e+00
Diameter	Depth	1164.08	9	6.879353e-245
CC	SewerType	601.61	3	4.516780e-130
CC	Depth	528.28	9	5.037709e-108
Length	Depth	411.59	9	4.612846e-83
Age	Location	356.97	11	8.118753e-70
Length	MaterialGroup	211.64	6	6.314357e-43
Age	SewerType	168.30	3	2.965395e-36
Diameter	CC	145.10	3	3.009154e-31
Length	CC	144.69	3	3.689401e-31
Length	SewerType	84.04	3	4.163208e-18
CC	Slope	105.33	10	4.649338e-18
Length	Location	103.73	11	3.254463e-17
CC	Location	91.03	11	1.045275e-14
Age	Slope	63.66	10	7.297218e-10
Age	Depth	55.83	9	8.443945e-09

6.1.4 Conclusions

The results of the three tests, although they are not clear (in the sense that Cross-Table analysis and Kruskal-Wallis indicate a significant p-value for every combination of covariates), they do come in total agreement, pointing at age, material, and sewer type as the most influential deterioration factors, followed by the installation depth. Since these factors appear to have the strongest influence on the sewer's condition, they should also have an explanatory value for the deterioration and could thus improve the model's predictions. In the following, the same factors are tested for their influence on the model's calibration module and their contribution to the accuracy of prediction.

6.2 Influence of the factors on model calibration

In the following, the prediction index, calculated by the prediction module of GompitZ, was plotted against the pipe's age in order to visualise the evolution of the pipe's condition in conjunction with its age. The title of each graph demonstrates the covariate and status that was implemented for each model simulation.

While some pipes seem to have different deterioration speeds depending on their special features (Figure 6.4), others have no effect on the deterioration speed (Figure 6.6).

The material seems to have a significant influence on the deterioration, as clay pipes deteriorate slower than concrete pipes (Figure 6.4 – Material) for all considerations of the variable. When the prediction index calculated by the model is plotted against the age, concrete pipes always have a worst (greater) prediction index than clay pipes of the same age. Similarly, clay pipes are much older than concrete pipes predicted to be in the same condition (same prediction index).

That means that during the calibration of the model, the sewer pipelines of different materials were found to follow different deterioration patterns, which make the material a significant deterioration factor.

Material and sewer type are strongly correlated as it was pointed out before. As a result, the sewer type also appears to drive the deterioration process. Stormwater sewers have the fastest deterioration rates, while combined have the lowest, with sanitary lying in the middle.

For both material and type of sewage, the distinction between the different sewer characteristics becomes significant when the deterioration factor is considered in the model with status 2 or 3, while it is milder when a status of 1 is assigned. These two factors are most efficiently considered when they are assumed to have an effect on the degradation process (status 2), rather than the initial state (status 1). The consideration of both effects (status 3) does not seem to further differentiate the deterioration speed of the different types of pipelines (the plotted curves for status 2 and 3 are almost identical) (Figure 6.4 – Sewer Type).

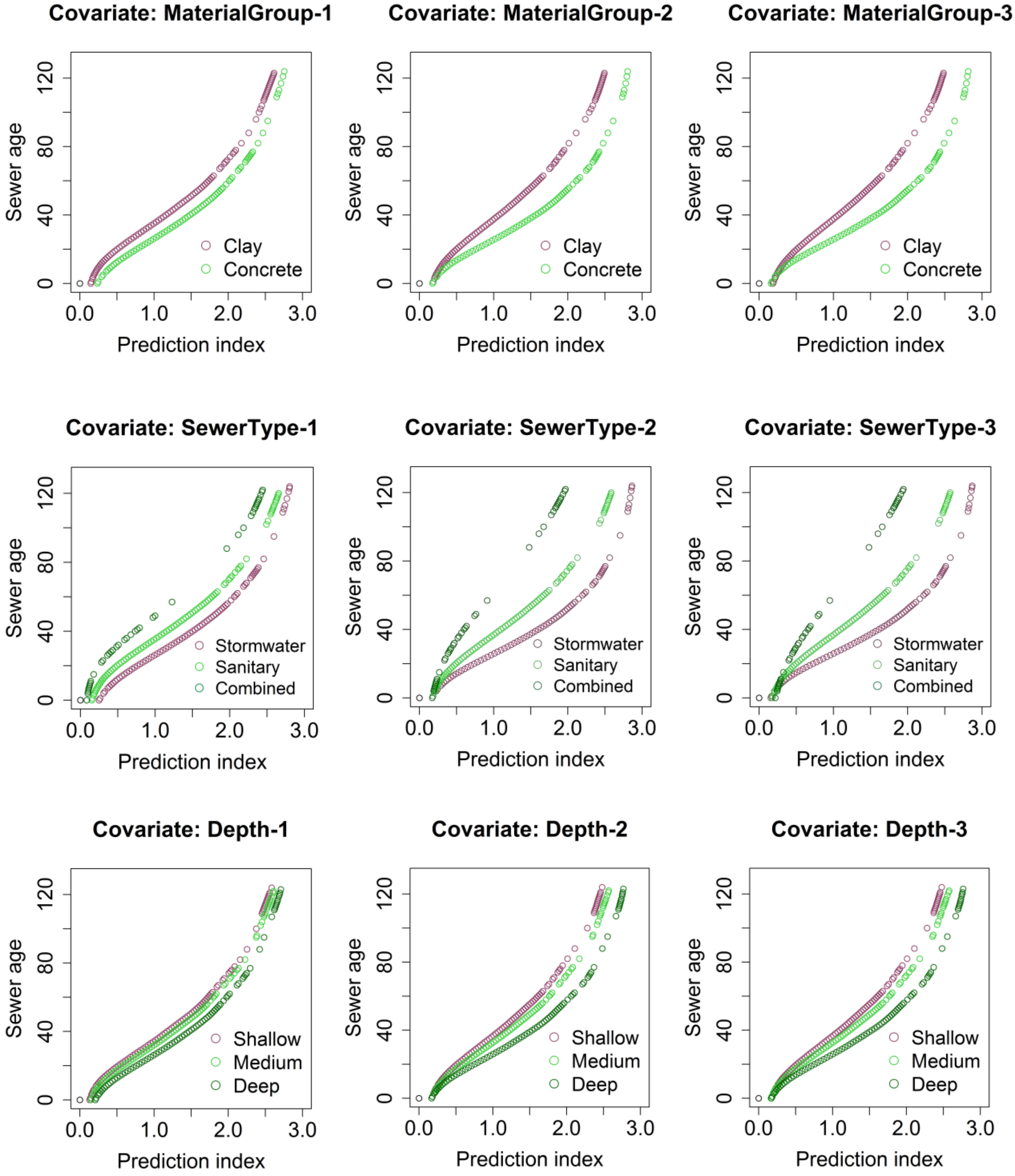


Figure 6.4: Covariates with an influence on the calibration.

In the case of the installation depth, the curves are not strongly distinguished, indicating a much more subtle effect on the deterioration. Sewers that are laid at a deep or medium depth deteriorate slower than shallow sewers (less than 2 m depth). The best consideration of the covariate is again a status 2 (influence on the degradation process), although no consideration of the covariate might also be an acceptable solution.

Lastly, the coefficients calculated for the location, when it is implemented with a status of 1 or 2, seem to be set very close to 0, since the different subcategories of the factor (overlying road with high traffic, medium traffic, etc.) follow very similar deterioration patterns (the curves almost overly on each other) (Figure 6.5).

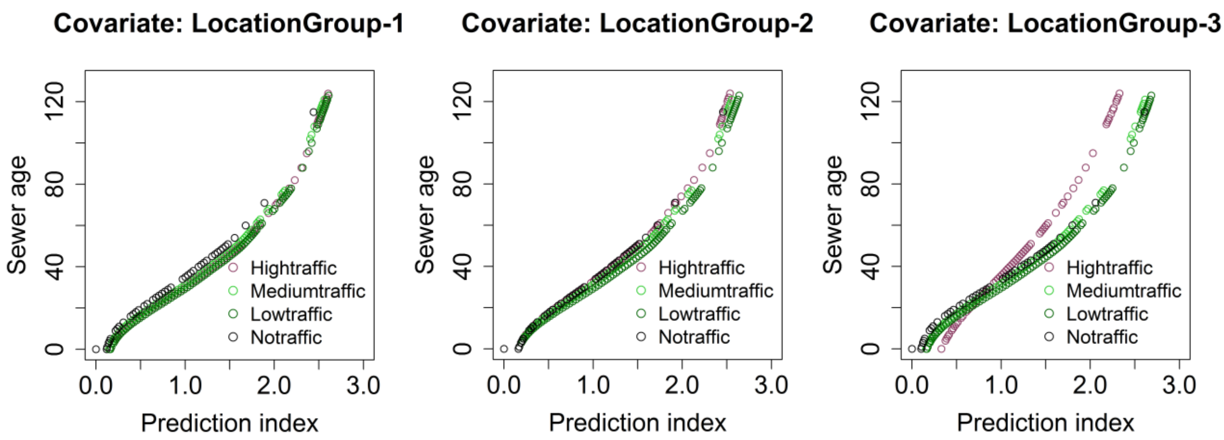


Figure 6.5: Influence of the location on the calibration.

However, when both effects of the factor are considered together (status 3), the sewer pipelines buried under roads that undertake high traffic dissociate their deterioration behaviour from the rest of the sewers, since they seem to have the lowest deterioration speed. However, this behaviour does not appear for the whole range of ages, but only for sewers older than 25 years old. For younger sewers, the effect of the location seems to be rather negative, since the sewers that are laid under high traffic have a worse prediction index than the rest. Additionally, the effect of the pipe location appears only for one type of overlying traffic and should therefore be neglected.

Unlike the previously mentioned, the pipe slope, length and diameter were found to have absolutely no influence, as their parameters seem to have been set very close to 0. In these cases, the sewers corresponding to different characteristics (e.g. different slopes, different diameters) overly on each other, indicating no special patterns in their deterioration behaviour (Figure 6.6).

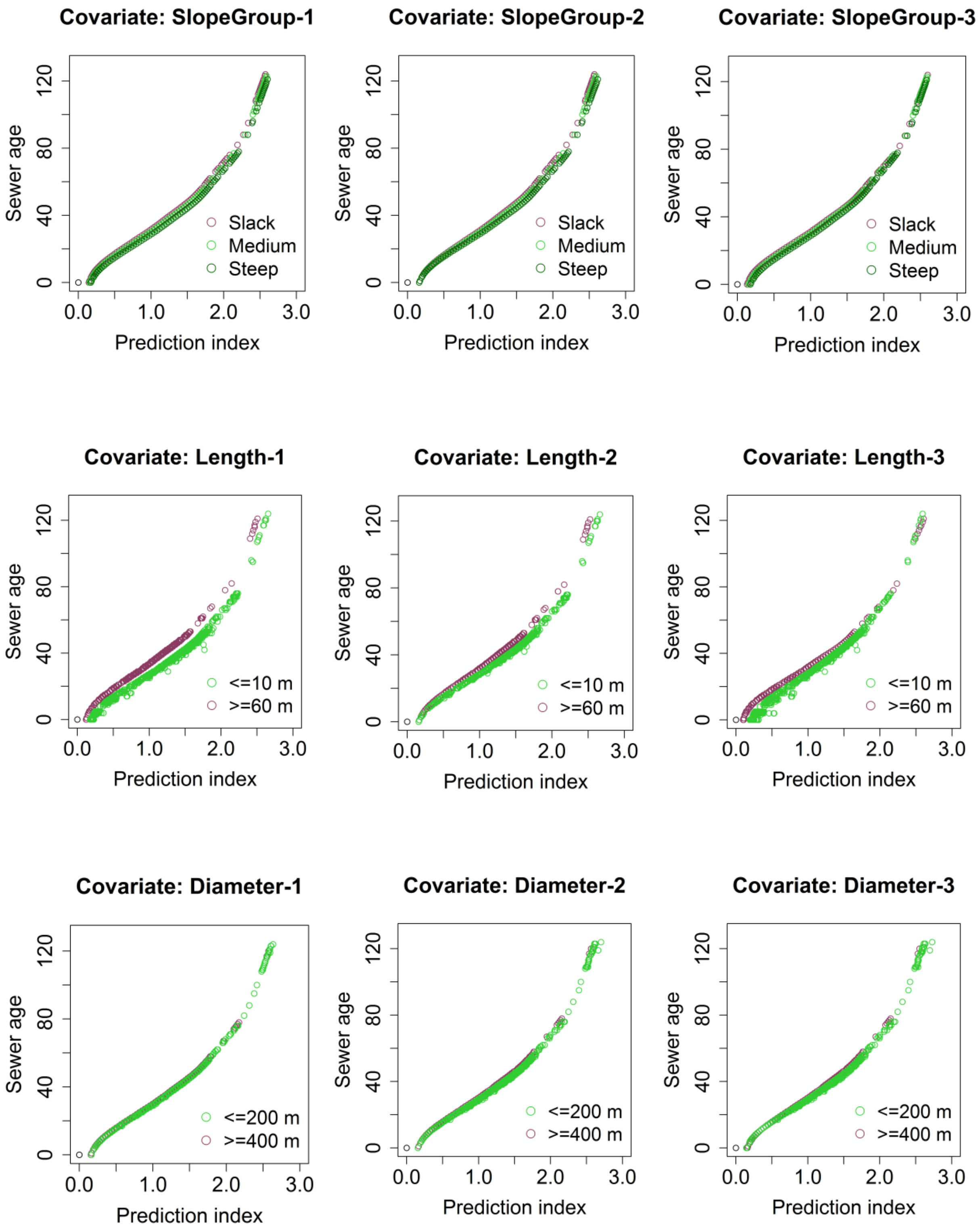


Figure 6.6: Covariates with no influence on the calibration process.

The above method verified the results of the correlation analysis. The factors that were found to have an influence on the calibration and thus on the deterioration are once again the material, sewer type, and installation depth, while no other covariate was considered significant.

6.3 Influence of the factors on the accuracy of prediction

Finally, the two performance curves were plotted for all considerations of each covariate.

Results indicate that the pipe material and type of sewage are the most influential covariates (Figure 6.7) while no other factor seems to have a significant effect (Figure 6.8). For these two covariates, the quality of prediction increases significantly when they are considered in the model, while it's almost irrelevant if they are considered with status 1, 2 or 3.

Without consideration of the material as a covariate (status 0), 42% of all the pipes in condition 4 are identified within 25% of the pipes predicted in the worst condition (Figure 6.7, performance curve 1, covariate: material). Considering material in the model (with a status of 1, 2 or 3) this number rises to 54%.

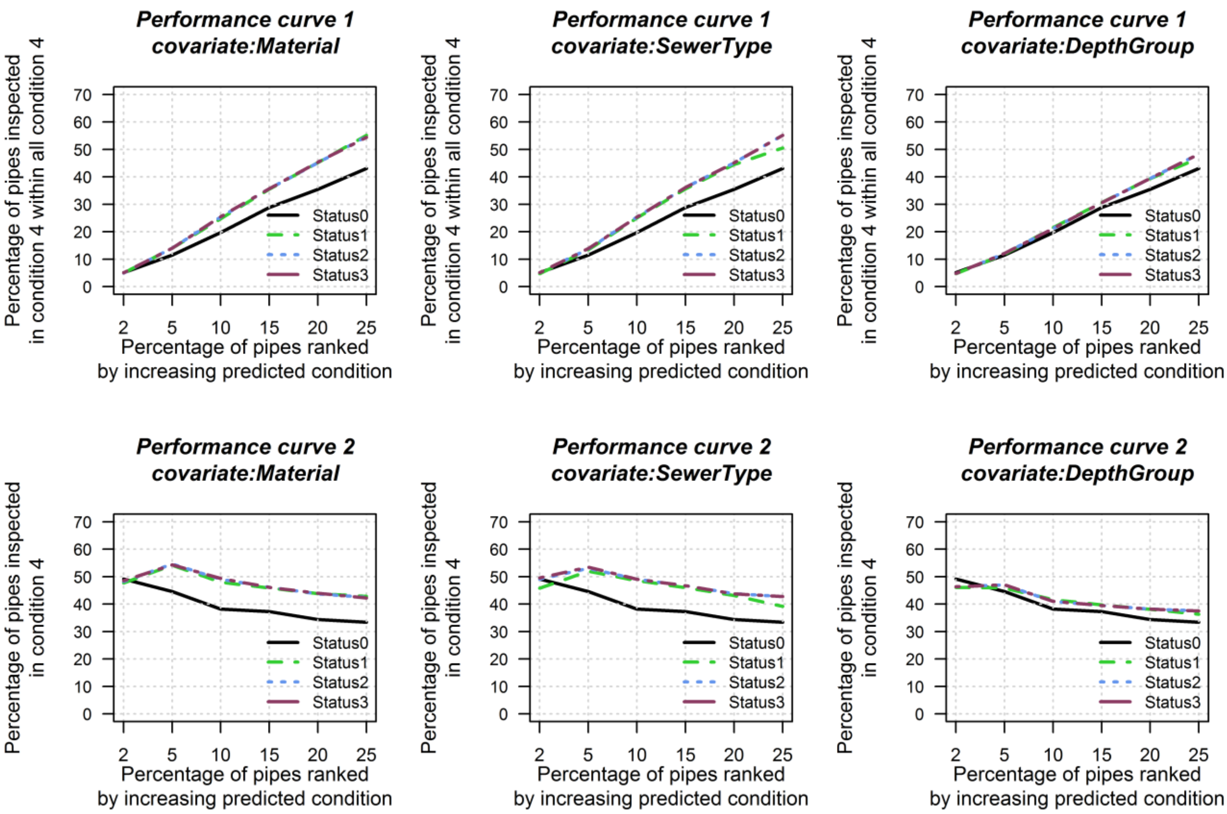


Figure 6.7: Covariates with an influence on the prediction.

The depth has also a subtle positive impact on the quality of prediction. However, it is not considered significant enough to be implemented in the model.

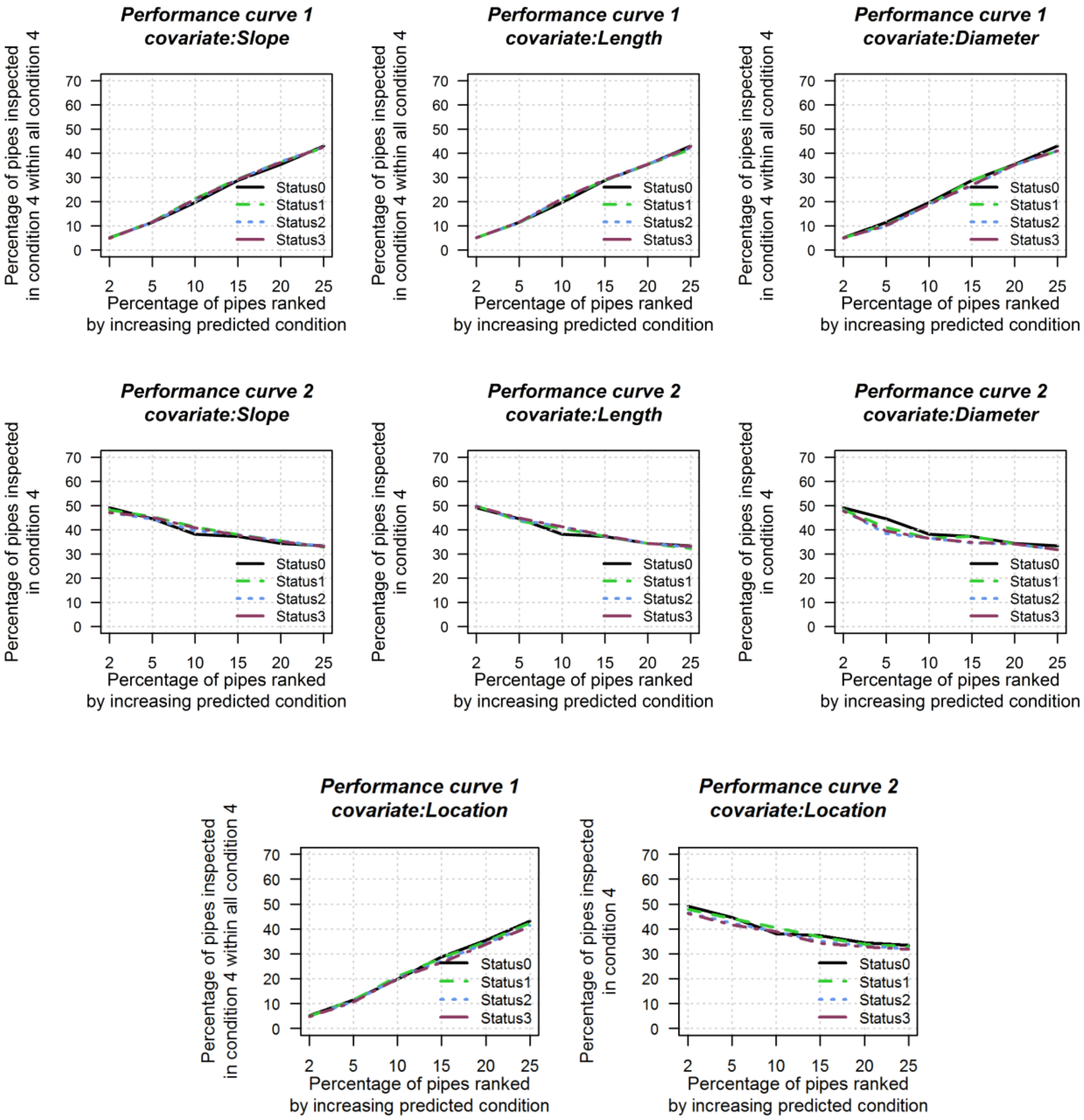


Figure 6.8: Covariates with no influence on the prediction.

The consideration of the remaining factors (diameter, slope, length, location) does not seem to improve the performance of the model. In most cases, the curves fall almost exactly on each other. In the case of the diameter and the construction period, their implementation even reduces the model's efficiency, since the curve corresponding to status 0 lies higher than the rest (Figure 6.8).

7 Modelling results

Finally, the two assessment methods described in Chapter 5, i.e. performance curves 1 and 2, were plotted for different considerations of input data in order to determine the model's prediction quality in each case. The available data was separated again into calibration and validation datasets, first chronologically (from the oldest to the most recent inspection) and then using a Monte Carlo simulation, drawing random selections of calibration and validation datasets (1000 runs). The first 70% was used to calibrate the model (21,976 inspections), while the rest 30% to validate it (9,418 inspections). The prediction quality for two cohorts, clay and concrete pipes was compared.

7.1 Modelling results for chronological selection of datasets

7.1.1 Modelling results without consideration of covariates

The model was calibrated without any consideration of covariates (age model). The survival functions following the Gompertz distribution that were calculated for no consideration of covariates were plotted in Figure 7.1.

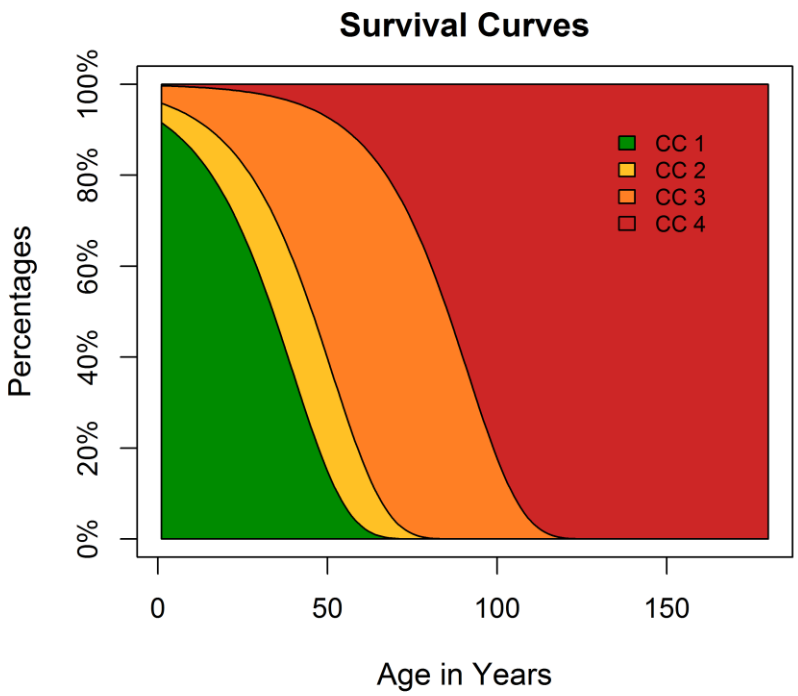


Figure 7.1: Survival curves plotted for all pipes and no consideration of covariates.

According to the survival curves plotted in Figure 7.1, it takes more than 50 years for the majority of the network to move to condition state 2 and 3, while it takes about 120 years until all pipes are in condition class 4.

Next, a prediction of the pipes' condition at the year of their inspection was made for the pipes of the validation dataset. Figure 7.2 shows the quality of the obtained results; within 10% of the pipes predicted in the worst conditions, 18% of all pipes inspected in condition 4 were identified (Figure 7.2 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of 42% of pipes inspected in condition 4 and 85% of pipes in condition 3 or 4 (Figure 7.2 - Performance curve 2).

Results indicate that the model can indeed support municipalities in the identification of the most deteriorated sewers. However, the pipe age is not the only relevant variable to describe sewer degradation. Deterioration rates can vary significantly between pipes depending on construction, operational and environmental factors (Davies et al., 2001). The consideration of further variables in the model may increase the quality of prediction.

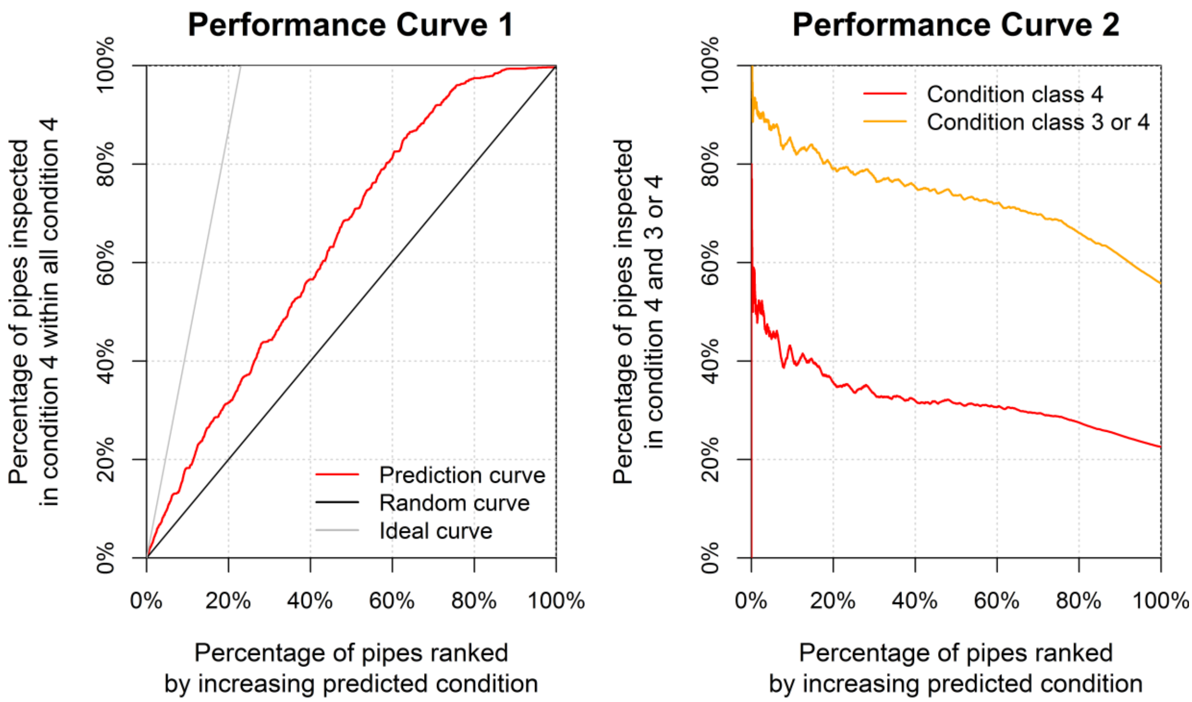


Figure 7.2: Quality of prediction obtained on the validation dataset for calibration without consideration of covariates.

7.1.2 Modelling results with consideration of the material as a covariate with status 2

The model was calibrated again with consideration of the material as the sole covariate with status 2. The new survival curves that were plotted are available in Figure 7.3.

The first thing that becomes apparent from the plotted survival functions is that with the current covariate implementation, the pipes' life expectancy seems to have reduced. The system needs less amount of time to move on to the next, worse condition state. According to Figure 7.3, all sewers older than 100 years old have moved to condition class 4, whereas when the model is calibrated with no covariates, the border age expands to 120 years old.

According to the performance curves, the prediction quality increases significantly after the material is considered as a covariate in the model. Since the performance curves in Figure 7.2 and 7.4 have originated from the same dataset, the results are directly comparable. The curves that were produced after material was included as a covariate in the calibration reach higher percentages and therefore indicate better results.

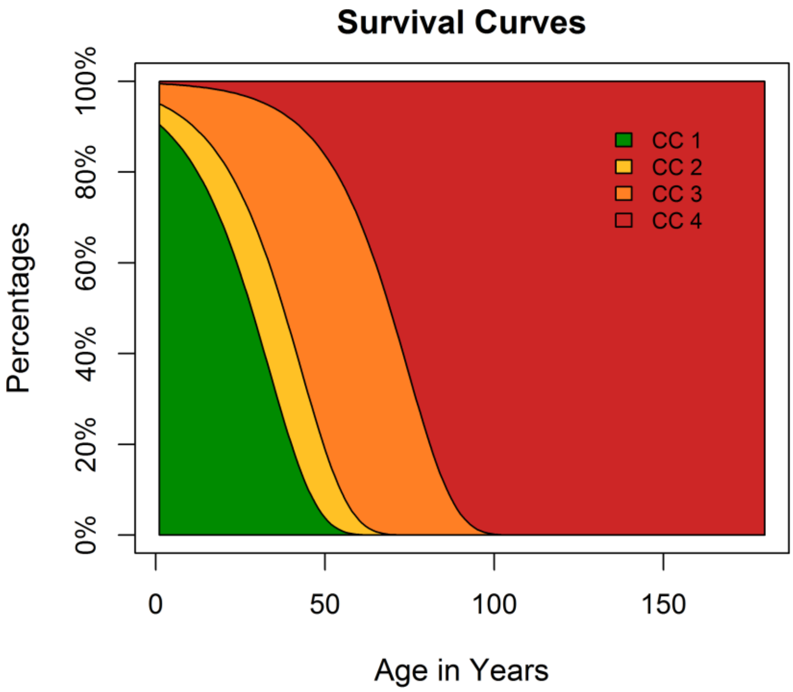


Figure 7.3: Survival curves plotted for all pipes and consideration of the material as a covariate with status 2.

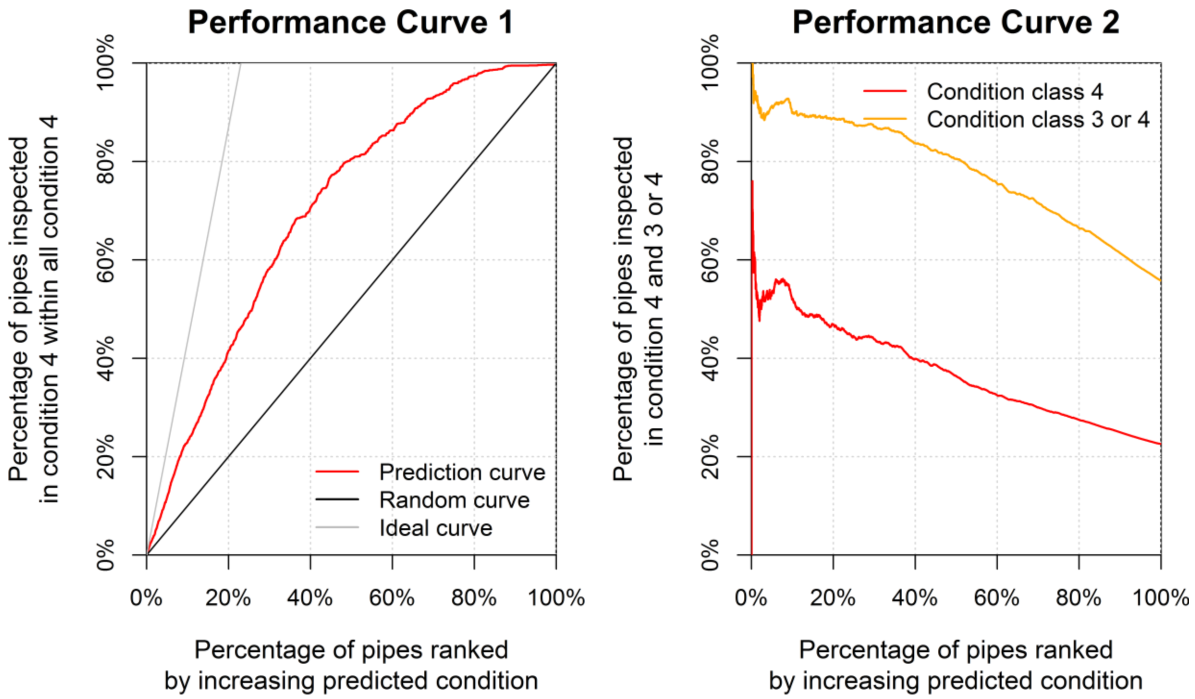


Figure 7.4: Quality of prediction obtained on the validation dataset for calibration with consideration of the material as a covariate with status 2.

Specifically, within 10% of the pipes predicted in the worst conditions, 24% of all pipes inspected in condition 4 were identified (Figure 7.4 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is now composed of 53% of pipes inspected in condition 4 and 92% of pipes inspected in condition 3 or 4 (Figure 7.4 - Performance curve 2).

This means that if inspections were performed on 10% of the pipes predicted in the worst conditions, 53% of pipes would be found in condition 4, 39% in condition 3 and only 8% would be found in condition 1 or 2.

7.2 Monte Carlo Simulation

7.2.1 Modelling results with consideration of the material as a covariate with status 2 (using 100% of the available data)

In order to assess the influence of the data selection on the quality of prediction, the model, including material as a covariate, was run 1000 times using Monte Carlo simulation (Figure 7.5). For each run, the calibration and validation datasets were selected randomly from all available data.

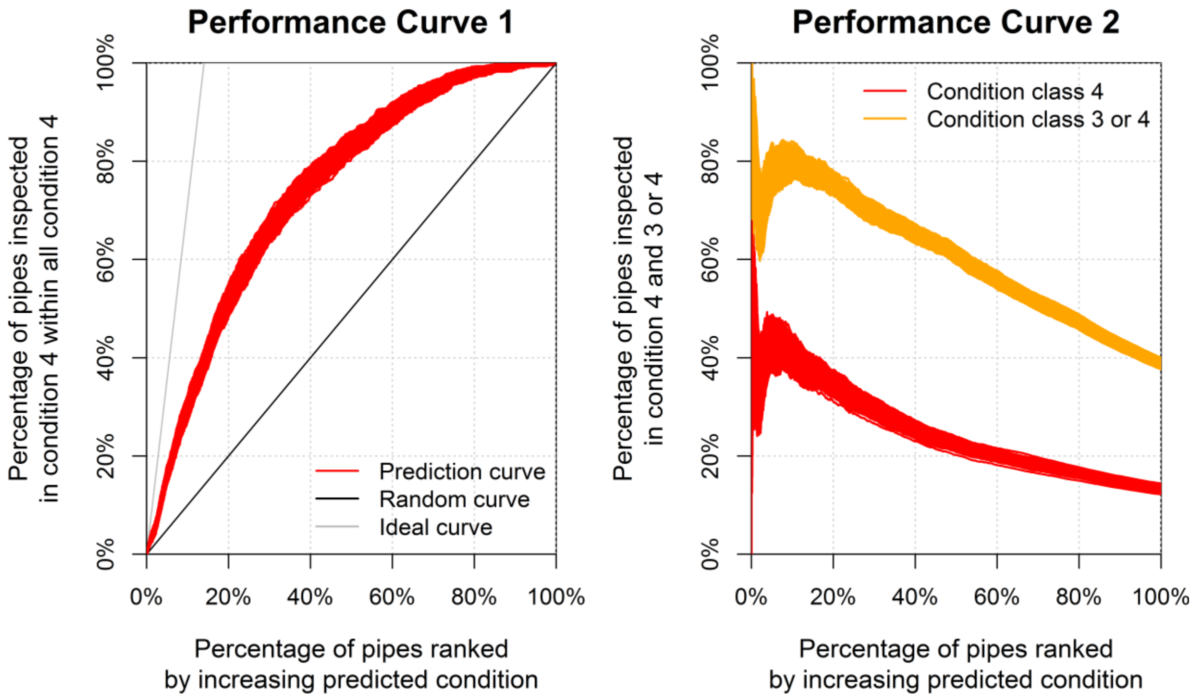


Figure 7.5: Quality of prediction obtained on the validation dataset considering material as a covariate with status 2. The model was tested using 100% of the data for 1000 Monte Carlo simulations.

The shape of both performance curves is similar to the ones obtained with a chronological selection of calibration and validation datasets, indicating a similar quality of prediction. However, the average condition distribution of the validation dataset, using Monte Carlo simulation, differs from the condition distribution of the validation dataset, using chronological selection (Figures 7.4 and 7.5 – Performance curve 2, read at x = 100%). The validation dataset that was selected chronologically, is composed of 23% of pipes in condition 4, whereas when it was selected randomly, that number turns to an average of 14%. Since the targeted dataset (pipes in condition 4) is reduced, it becomes more difficult to identify its components and therefore the percentage identified by the model within the same subset is expected to be reduced, without implying that the model's performance is reduced as well.

That explains the difference of results obtained by analysing specific values of the curves; within 10% of the pipes predicted in the worst conditions, more than 30% of all pipes inspected in condition 4 were identified (Figure 7.5 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of an average of 40% of pipes inspected in condition 4 and 80% of pipes inspected in condition 3 or 4 (Figure 7.5 - Performance curve 2).

7.2.2 *Modelling results with consideration of the material as a covariate with status 2 (using only 20% of the available data)*

In order to assess the influence of the size of the calibration dataset on the quality of prediction, the model was once again tested, but this time using only 20% of the available data. The model, including material as a covariate, was run 1000 times using Monte Carlo simulation (Figure 7.6).

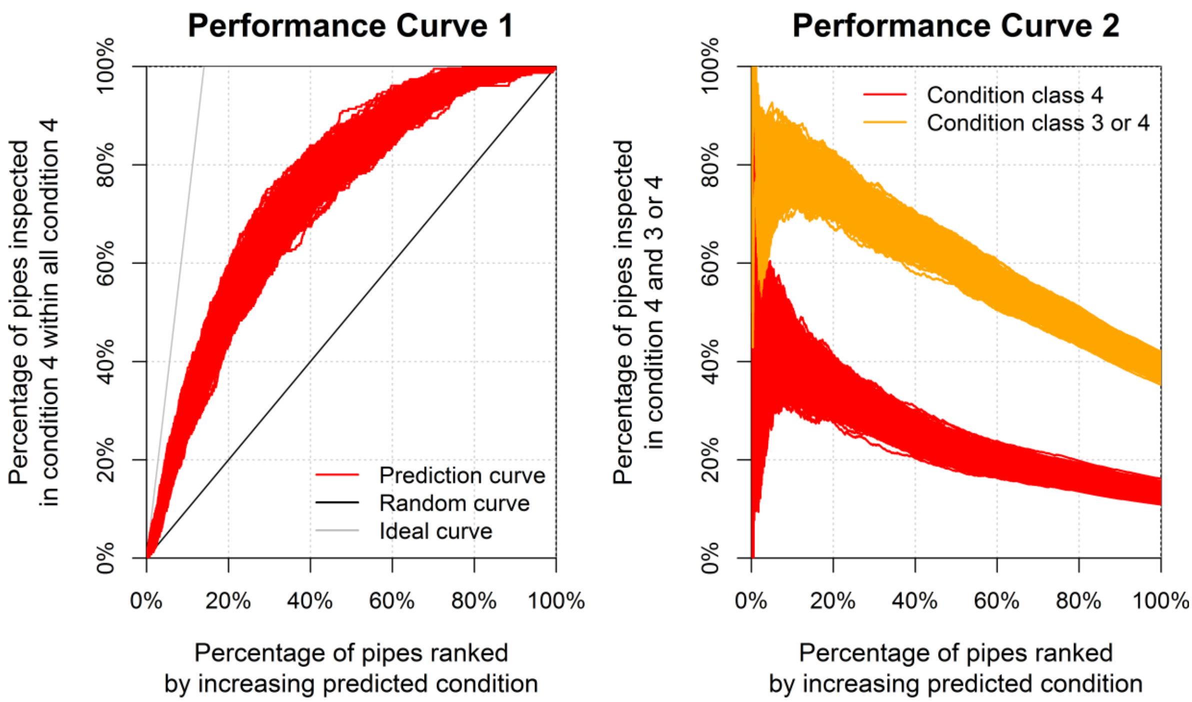


Figure 7.6: Quality of prediction obtained on the validation dataset, considering material as a covariate with status 2. The model was tested using 20% of the data for 1000 Monte Carlo simulations.

In general, the shape of both performance curves 1 and 2 is similar to the ones obtained when testing the model with 100% of the data. Regarding performance curve 1, the confidence interval around the average percentage of sewers identified in condition 4 seems to double throughout the curve (Figure 7.6 - Performance curve 1). In the case of performance curve 2, the average percentage of sewers inspected in condition 4, and 3 or 4, remains the same, while the confidence interval around the average value increases significantly for the smaller percentages (0% to 20%). As the curve approaches to 100%, the difference becomes much more subtle (Figure 7.6 - Performance curve 2).

Results indicate that the calibration dataset has an influence on the model's predictions; within 10% of the pipes predicted in the worst conditions, the percentage of pipes inspected in condition 4 that were identified varies between 30% and 40% (Figure 7.6 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of 30% to 50% of pipes inspected in condition 4, and 70% to 90% of pipes inspected in condition 3 or 4 (Figure 7.6 - Performance curve 2).

When examining the 5% of the worst predicted cases (which in most cases even exceeds the annual budget of a municipality for rehabilitation of the drainage system), the number of sewers identified in condition 4 varies between 30% and 60%, whilst the number of sewers in condition 3 or 4 vary between 70% and 90%.

7.2.3 Modelling results with consideration of the material as a covariate with status 2 (using only 10% of the available data)

While the general shape of both performance curves 1 and 2 is once again very similar to the ones obtained by testing the model with 20% of the data, the confidence interval around the average value increases further (Figure 7.7).

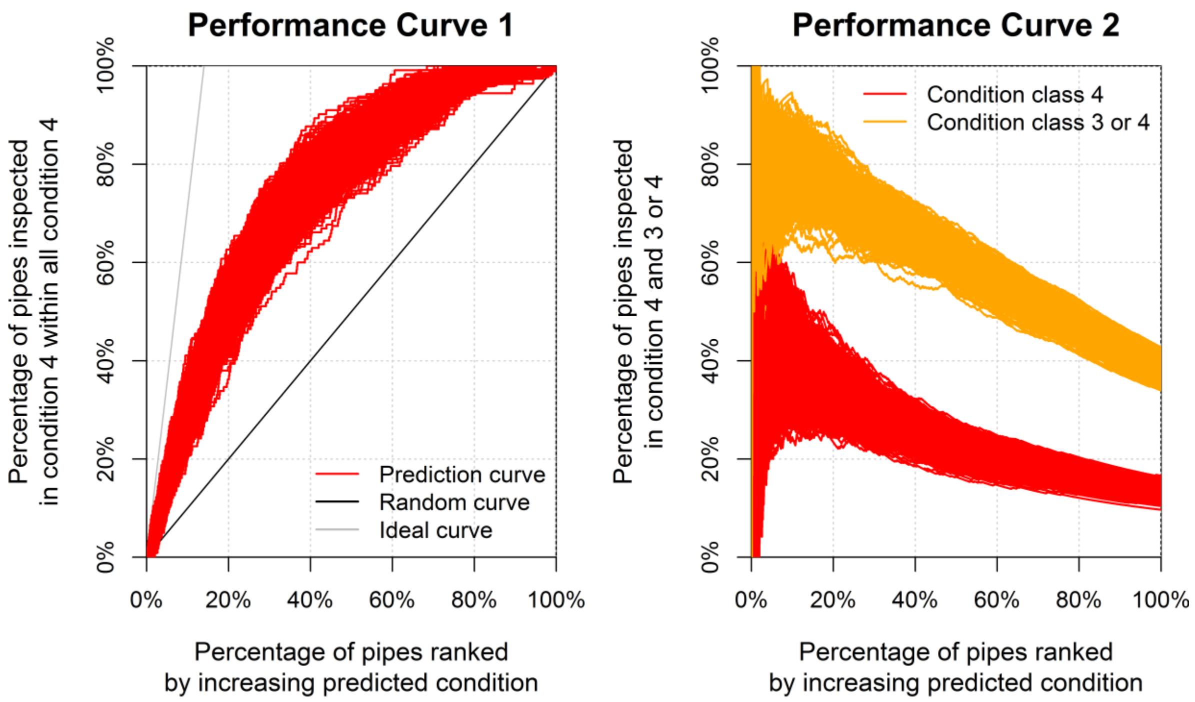


Figure 7.7: Quality of prediction obtained on the validation dataset, considering material as a covariate with status 2. The model was tested using 10% of the data for 1000 Monte Carlo simulations.

Within 10% of the pipes predicted in the worst conditions, the percentage of pipes inspected in condition 4 that were identified varies between 20% and 40% (Figure 7.7 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of 22% to

55% of pipes inspected in condition 4, and 65% to 95% of pipes inspected in condition 3 or 4 (Figure 7.7 - Performance curve 2). When examining the 5% of the worst predicted cases, the number of sewers identified in condition 4 varies between 20% and 62%, whilst the number of sewers in condition 3 or 4 varies between 62% and 97%.

Since the calibration dataset is particularly extensive (31,394 inspections), when selecting random data for calibration, all datasets are very similar in terms of the types of pipelines they contain and therefore the quality of prediction shows small variations. When the amount of data dedicated to model testing is reduced to 20% and then to 10%, the differences among the random datasets become significant, since there is a much smaller data pool to choose from. This leads to differences in the model's performance. The random datasets chosen at each time determine the model's accuracy.

Since the value of the curve regards mainly the smaller percentages, which contain the pipes predicted in the worst conditions, the impact of the input dataset is maximised in the beginning of the curve.

7.3 Modelling results for clay and concrete pipes

Since the material is the most important factor that affects sewer deterioration, the effect of each material was further investigated. The data was divided with respect to the corresponding material into two datasets, containing only clay or only concrete pipes. The calibration and validation datasets were chosen chronologically among all data contained in each group. In order to enable the direct comparison of prediction results between clay and concrete pipes, the validation datasets for each case were manipulated in order to contain the exact same distribution of condition classes. This was done by excluding the pipes of a condition class from either dataset when they exceeded the corresponding amount in the other dataset. Finally, the validation dataset for both materials contained 764 pipes in condition class 1, 244 in condition 2, 1205 in condition 3 and 467 in condition 4. These populations correspond to a distribution of 28.5%, 9.1%, 45% and 17.4% of pipes in condition 1, 2, 3 and 4, respectively.

7.3.1 Modelling results for clay pipes

The survival functions calculated for the subset of clay pipes are demonstrated in Figure 7.8. They are almost identical to the survival curves acquired for no consideration of covariates, since they need a relatively longer amount of time to move on to the next, worse, state.

Next, a prediction of the pipes' condition at the year of their inspection was made for the pipes of the validation dataset. Within 10% of the pipes predicted in the worst conditions, 25% of all pipes inspected in condition 4 were identified (Figure 7.9 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of 45% of pipes inspected in condition 4 and 97% of pipes in condition 3 or 4 (Figure 7.9 - Performance curve 2).

While results are not directly comparable to the ones produced above, since the distribution of condition classes among the datasets differs, the prediction quality for clay pipes seems to be rather similar to the one identified for all pipes, when the material was implemented as a covariate with status 2.

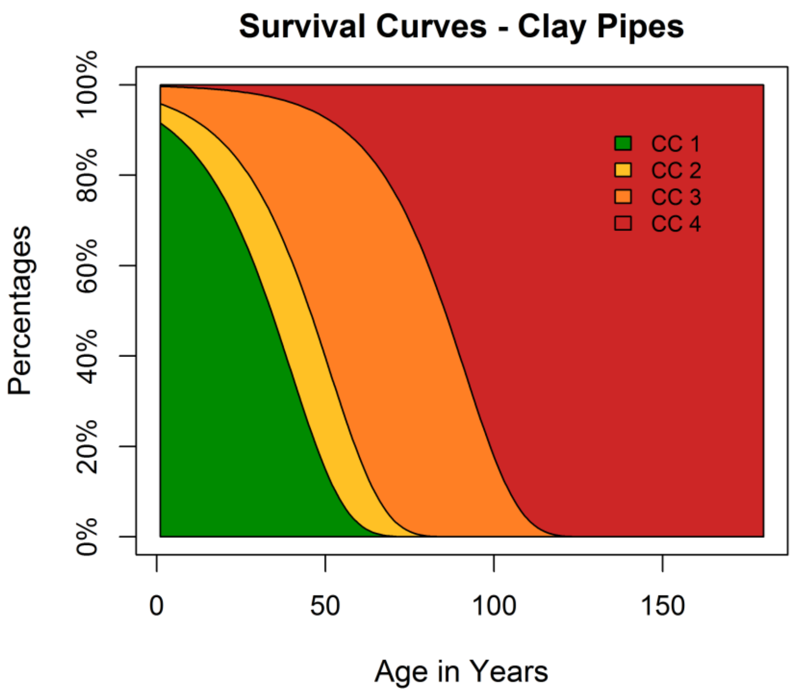


Figure 7.8: Survival curves plotted for clay pipes and no consideration of covariates.

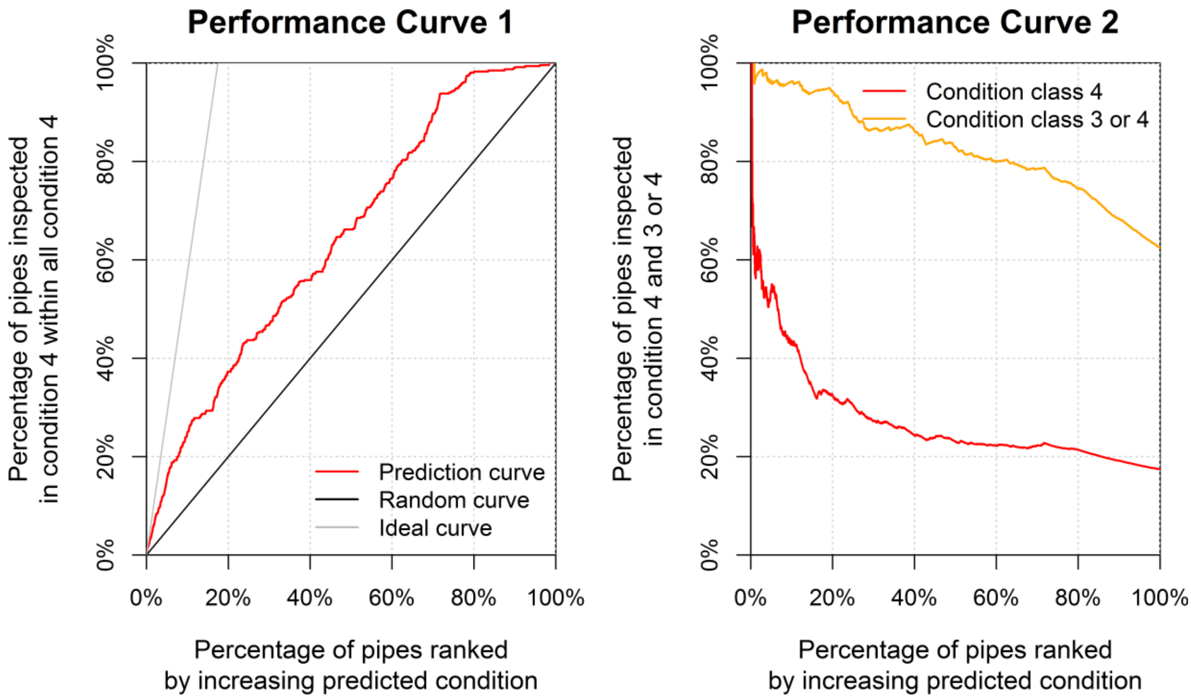


Figure 7.9: Quality of prediction obtained on the validation dataset only for clay pipes for calibration without consideration of covariates.

7.3.2 Modelling results for concrete pipes

As it was already noted, clay pipes are in a better general condition than concrete pipes. This fact also reflects to the survival curves plotted for the two materials (Figures 7.8 and 7.10).

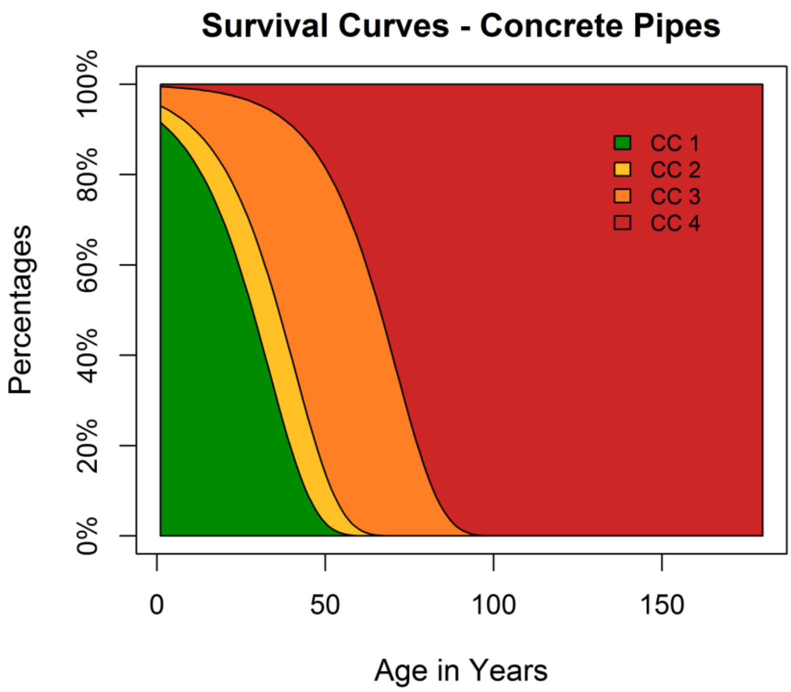


Figure 7.10: Survival curves plotted for concrete pipes and no consideration of covariates.

According to Figure 7.10, 100% of the concrete pipes have moved to condition class 4 after approximately 90 years after their construction, whereas clay pipes (Figure 7.8) need 35 years more, i.e. 39% more time until all of them have moved to the worst condition.

Finally, the performance curves 1 and 2, only for concrete pipes and with no implementation of covariates are available in Figure 7.11.

Within 10% of the pipes predicted in the worst conditions, 18% of all pipes inspected in condition 4 were identified (Figure 7.11 - Performance curve 1). In addition, the subset of 10% of the worst predicted cases is composed of 25% of pipes inspected in condition 4 and 84% of pipes in condition 3 or 4 (Figure 7.11 - Performance curve 2).

Results indicate that the model is not efficient for predicting the future condition of concrete pipes. The first, and also critical, 20% of the worst predicted cases demonstrated in performance curve 1 (Figure 7.11) contains only 20% of all pipes in condition 4, the same amount that we would find if we were performing random inspections.

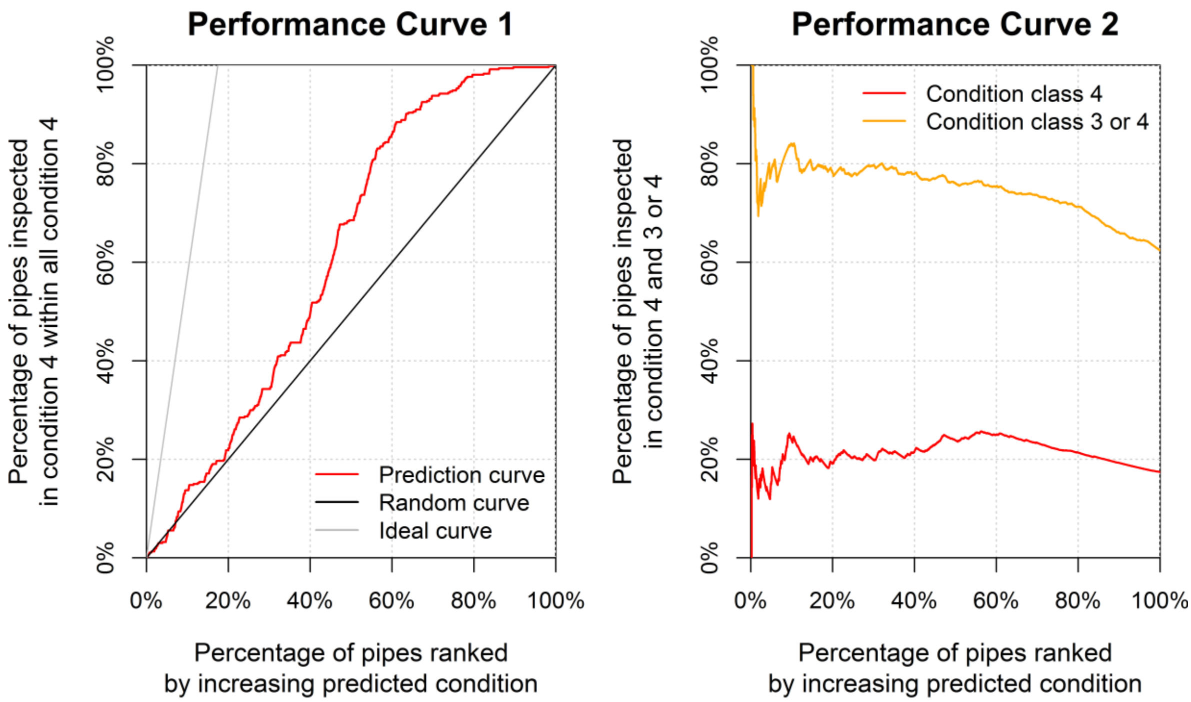


Figure 7.11: Quality of prediction obtained on the validation dataset only for concrete pipes for calibration without consideration of covariates.

At the same time, performance curve 2 verifies the results of the first curve. The 20% of the worst predicted cases is composed of 20% of pipes in condition 4, which is hardly higher than the percentage of pipes in condition 4 corresponding to all the pipes. Results are not improved considerably, even when condition classes 3 and 4 are considered together.

7.3.3 Conclusions

Surprisingly enough, the material of the sewer seems to have a great impact not just on the deterioration speed of the pipes, but also on the model's capability to predict deterioration patterns. While the model seems to be working for clay sanitary pipes, it is rather useless to predict the deterioration of concrete stormwater pipes. The deterioration behaviour of concrete pipes seems rather unpredictable, as the age factor is proven insufficient for an accurate prediction.

In order to understand that, it is essential to also understand the different nature of these two subsets of pipes. As it was already demonstrated, the concrete pipes of this dataset facilitate needs of rainwater drainage, and therefore this fact might be due to the unexpectable nature of storm water pipes themselves. The model can predict the degradation pattern of a sanitary pipe, that is expected to receive a stable amount of sewage during its lifetime, and therefore its degradation depends mostly on its age. However, random events, including large flow fluctuations, as well as a variety of solids that can enter a stormwater pipe, can accelerate its deterioration, and thus make it difficult to predict. While sanitary pipes receive wastewater, through a network of pipes that connects the main sewer to a (most of the times) domestic source, storm water pipes receive the excessive amount of water from the streets, and are therefore easily accessible to a variety of solids, including sediments, waste, sand, as well as tree branches and leaves that can challenge the inner capacity of the sewer.

8 Conclusions

The quality of prediction of the sewer deterioration model GompitZ was evaluated using an extensive inspection dataset from a German case study. Results demonstrate the potentials of deterioration models, as well as the need for further research.

8.1 Summary and results

In order to highlight their importance, sewer networks are often called the 'lifeline' of a municipality (Ariaratnam et al., 2001). Since they carry the domestic waste and rainwater, their malfunction can cause severe problems to the population and the environment. Therefore pro-active strategies are the favoured solution in order to minimise the consequences of a failed component.

The application of asset management strategies for sewer maintenance and rehabilitation requires knowledge over the condition of sewer pipelines, typically acquired by visual inspections of the network. However, this procedure is expensive, time-consuming, and the results' accuracy can be questioned due to uncertainties involving technical difficulties and human judgement.

Sewer deterioration models aim to enable an effective, pro-active maintenance of the network, while minimising the inspection costs. However, their use is rather limited. Few studies tried to assess the prediction quality of deterioration models and determine their value for municipalities. Demonstrating the use of these models to sewer operators would increase the trust in them and therefore their range of use.

This thesis evaluates the performance of GompitZ (Le Gat, 2008), a Markov-based sewer deterioration model, using an extensive dataset from a German city (31,394 classified inspections). Since short-term rehabilitation programs focus on repair and replacement of pipes in poor condition, two methodologies were used, in order to assess the model's performance according to its capacity to identify the most deteriorated pipes.

Modelling results are encouraging, even if somewhat controversial. Most tests verified the potentials of sewer deterioration models to support inspection and rehabilitation programs. Even under consideration of model uncertainties, the results can still support sewer operators to define more efficient strategies. However, other factors might as well interfere with the model’s ability to identify the most deteriorated sewers. The inherit randomness of failure events that might occur for some types of pipes makes the prediction of their behaviour a challenging procedure. In this study, the model’s capacity to identify the most deteriorated sewers differs significantly for different pipe materials. While the model proved very useful in identifying the deterioration patterns of clay pipes it was rather incapable of predicting failures in concrete pipes.

On the other hand, the results of the Monte Carlo simulation underline the sensitivity of the model’s quality of prediction to the size and quality of the available data. When random calibration and validation datasets are selected among the data, the quality of prediction varies significantly for different datasets, when these do not correspond to a representative population of pipes, i.e. when the size of the original dataset is decreased considerably. For both performance curves, the confidence interval around the average quality of prediction increases significantly when the size of the dataset decreases. That dependency of the model should be taken into account when making future predictions. When the size of the dataset is not significant, the model results can be questioned.

Results also underline the model’s dependency on a set of user-defined parameters, including a set of covariates suspecting to influence the deterioration, as well as a status defining the type of influence. However, while the correct selection of covariates and statuses can significantly increase the quality of prediction, the implementation of an unreasonable amount of covariates that have no explanatory value, as well as an incorrect interpretation of statuses, can result in an overly complicated model that is time consuming and produces questionable results. The methodology proposed in chapter 5.2 can be applied to other case studies to identify the most relevant user-defined parameters with respect to the available dataset. In this case study, the pipe material and type of sewage were found to be the only relevant covariates.

8.2 Discussion

Testing the prediction quality of sewer deterioration models is necessary but it can also be a very challenging and problematic procedure. GompitZ model requires a large database to perform the model calibration, as well as a well-defined set of covariates. Most municipalities don’t have access to such an extensive dataset and cannot perform model predictions including only the pipelines they operate.

In addition, GompitZ model does not output a predicted condition class for each sewer in a defined time point in the future. It outputs a condition probability vector, containing a probability for each sewer to be in each condition class. This does not allow a direct comparison of the predicted and inspected values. Results have to be manipulated in a way that enables the model validation. In this thesis, both assessment methodologies are based on the extraction of a prediction index from the probability vector that allows the ranking of the pipelines according to their predicted future condition. This method has some obvious disadvantages, since some information is lost during that translation and its accuracy can be questioned. However, it imports less uncertainty than the translation of the condition probabilities to one single condition class, since this method eliminates the influence of the rest of the probabilities (it equals to a probability vector that assigns a probability of 100% to one sole class).

The performance curve 2, which was developed in the frame of this thesis, seems to add significant value for municipalities and sewer operators compared to performance curve 1. It identifies the real condition distribution among the sewers that were predicted to be in the worst condition states. In other words, if a municipality inspects the x% of the sewers predicted with the worst prediction indexes, they will know what percentage of pipes they are going to find in each condition, according to the model's prediction. That is a very useful indicator, since it directly addresses a practical problem. While performance curve 1 has a scientific value in determining the model's performance, sewer operators rarely have the funds to rehabilitate every pipeline in poor condition. Therefore the percentage of sewers inspected in the worst condition that the model is able to identify among all sewers in critical condition has more of a scientific and less of a practical value.

Regarding the performance curves, another issue that came up was their dependence on the calibration dataset, and specifically its condition distribution. As expected, the higher the percentage of sewers in poor conditions within the calibration dataset, the higher the percentage the model is able to identify. Unfortunately, that means that a comparison between curves that have originated from different datasets can be very problematic. The model's performance should be assessed based on the relative amount of sewers that it was able to identify compared to the total amount of sewers in poor condition.

8.3 Recommendations for further research

Many questions remain to be answered. This study was performed using an extensive inspection dataset. However, in most cities, only small percentages of the sewer system have already been inspected and pipe characteristics data may be missing (e.g. unknown pipe material).

The sensitivity of deterioration models to input data quality (uncertainty of CCTV data as underlined recently by Dirksen et al., 2013 and van der Steen et al., 2013), as well as the influence of the condition distribution of the available dataset needs to be carefully assessed. An indicator should be extracted from both performance curves that will normalise the results and enable the comparison between curves originating from different datasets.

Lastly, as it was concluded above, for pipes that their deterioration does not follow any specific patterns, there is no model that can successfully predict their condition. This could be potentially improved if more variables that are related to the pipe's degradation are inserted in the problem. For example the location of a concrete, storm water pipe could explain the amount of solids and rainwater the pipe is expected to undertake.

References

Ablin, R.L. and Kinshella, P., 2004. Dude, where's my pipe – Accelerated corrosion rate threatens Phoenix sewers, Pipeline engineering and construction [online], 1-8. Available from: http://dx.doi.org/10.1061/40745(146)81 [Accessed 23 August 2014].

Ahmadi, M., Cherqui, F., De Massiac, J.C., and Le Gauffre P., 2013. Influence of available data on sewer inspection program efficiency, Urban Water [online]. Available from: http://dx.Doi.Org/10.1080/1573062x.2013.831910 [Accessed 30 April 2014].

Al-Barqawi, H. and Zayed, T., 2006. Condition rating model for underground infrastructure sustainable water mains. Journal of Performance of Constructed Facilities, 20 (2), 126-135.

Ana, E.V., 2009. Sewer asset management - sewer structural deterioration modelling and multi-criteria decision making in sewer rehabilitation projects prioritization. Thesis (PhD). Vrije Universiteit Brussel.

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. and 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.M., El-Assaly, A., Associate Members, ASCE, and Yang Y., 2001. Assessment of infrastructure inspection needs using logistic models. Journal of Infrastructure Systems, 7(4), 160-165.

AwwaRF, 2005. Distribution system security primer for water utilities. USA: AwwaRF.

Baik, H.S., Jeong, H.S., and Abraham, D.M., 2006. Estimating Transition Probabilities in Markov Chain-Based Deterioration Models for Management of Wastewater Systems. Water Resources Planning and Management, 132 (1), 15-24.

Baur, R. and Herz, R., 2002. Selective inspection planning with aging forecast for sewer types.

Becker, G., Boduroglu, H., Camarinopoulos, S., Frondistou-Yannas, S., Gedikli, A., Kallidromitis, V.G., Kampranis, D., and Sanna, C., 2009. Structural assessment and upgrading of sewers based on inspection results. Infrastructure Systems, 15 (4), 321-329.

Chugtai, F. and Zayed, T., 2008. Infrastructure condition prediction models for sustainable sewer pipelines. Performance of constructed facilities, 22 (5), 333-341.

Davies, J.P., Clarke, B.A., Whiter, J.T., and Cunningham, R.J., 2001. Factors influencing the structural deterioration and collapse of rigid sewer pipes. Urban Water, 3 (1-2), 73-89.

Dirksen, J., Clemens, F.H.L.R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., and Snaterse, C.T.M., 2013. The consistency of visual sewer inspection data. Structure and Infrastructure Engineering, 9(3), 214-228.

DWA, 2009. Zustand der Kanalisation in Deutschland, Ergebnisse der DWA-Umfrage 2009, Germany: Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.

DWA, 2011. Advisory Leaflet DWA-M 149-3 - Conditions and Assessment of Drain and Sewer Systems Outside Buildings – Part 3: Condition Classification and Assessment, 2007. Hennef: DWA Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.

DWA, 2012. Leitfaden zur strategischen Sanierungsplanung von Entwässerungssystemen außerhalb von Gebäuden - T4/2012. Hennef: DWA Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V., 978-3-942964-58-6.

Ens, A., 2012. Development of a Flexible Framework for Deterioration Modelling in Infrastructure Asset Management. Thesis (MSc). University of Toronto.

EPA, 2008. Clean Watersheds Needs Survey 2004 Report to Congress. US: Environmental Protection Agency.

EPA. Asset Management. Available from: http://water.epa.gov/infrastructure/sustain/asset_management.cfm [Accessed 23 August 2014].

Ertl, Th., Gangl, G., Bölke, K.P., Kretschmer, F., 2007. Implementing quality management and EN 13508-2 for CCTV sewer inspection in Austria. Novatech.

Fenner, R.A., Sweeting, L., and Marriott, M.J., 2000. A new approach for directing proactive sewer maintenance. Water and Maritime Engineering, 142(2), 67-77.

Fischer, B., Hunger, W., Lehmann, T. M., Müller, K. and Schäfer, T., 2006. Objective Condition

Establishment of Sewer Systems. Proceedings of the 2nd IWA Conference on Sewer Operation and Maintenance – SOM06, Vienna.

Hahn, M.A., Palmer, R.N., Merrill, M.S., and Lukas, A.B., 2002. Expert system for prioritizing the inspection of sewers: knowledge base formulation and evaluation. Journal of Water Resources Planning and Management, 128 (2),121–129.

Hao, T., Rogers, C.D.F., Metje, N., Chapman, D.N., Muggleton, J.M., Foo, K.Y., Wang, P., Pennock, S.R., Atkins, P.R., Swingler, S.G., Parker, J., Costello, S.B., Burrow, M.P.N., Anspach, J.H., Armitage, R.J., Cohn, A.G., Goddard, K., Lewin, P.L., Orlando, G., Redfern, M.A., Royal, A.C.D., and Saul, A.J., 2011. Condition assessment of the buried utility service infrastructure, Tunnelling and Underground Space Technology, 28, 331-344.

IBAK. IBAK PANORAMO [online]. Available from: http://www.ibak.de/en/produkte/ibak_show/frontenddetail/product/panoramo/ [Accessed 24 August 2014].

IBAK, 2010. PANORAMO SI 3D Optical Manhole Scanner [online]. Available from: http://www.rapidview.com/panoramo_si.html [Accessed 24 August 2014].

Jones, G.M.A., 1984. The structural deterioration of sewers. International conference on the planning, construction, maintenance, and operation of sewerage systems, Reading, UK.

Kawabata, T., Uchida, K., Ariyoshi, M., Nakase, H., Mohri, Y., and Ling, H., 2003. D.E.M. Analysis on Behavior of Shallowly Buried Pipe Subject to Traffic Loads, Pipeline Analysis, 1218-1227.

Khan, Z., Zayed, T., and Moselhi, O., 2010. Structural Condition Assessment of Sewer Pipelines. Performance of constructed facilities, 24 (2), 170-179.

Kleiner, Y. and Rajani, B., 2001. Comprehensive review of structural deterioration of water mains: statistical models. Urban Water, 3(3), 131-150.

Kley, G., Kropp, I., Schmidt, T., and Caradot, N., 2013. SEMA Report: D 1.1 Review of Available Technologies and Methodologies for Sewer Condition Evaluation [online]. Berlin, KompetenzZentrum Wasser Berlin. Available from: http://www.kompetenz-wasser.de/fileadmin/user_upload/pdf/forschung/SEMA/D11_SEMA_Review_of_technologies_and_methodologies_for sewer_condition_evaluation.pdf [Accessed 28 April 2014].

Kley G. and Caradot N., 2013. SEMA Report: D 1.2 Review of Sewer Deterioration Models [online]. Berlin, KompetenzZentrum Wasser Berlin. Available from: http://www.kompetenz-wasser.de/fileadmin/user_upload/pdf/forschung/SEMA/D12_SEMA_Review_of_sewer_deterioration_models.pdf [Accessed 28 April 2014].

Kropp, I. and Baur, R., 2005. Integrated failure forecasting model for the strategic rehabilitation planning process. Water Supply, 5 (2), 1-8.

Le Gat, Y., 2008. Modelling the deterioration process of drainage pipelines. Urban Water Journal, 5 (2), 97-106.

Le Gat, Y., 2011. Modelling the degradation of sewer pipelines in CareS : The GompitZ tool user’s guide - version 2.08. Cemagref/REBX.

Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C., and Desmulliez, J. J., 2004. Gestion patrimoniale des réseaux d'assainissement urbains, Guide méthodologique. Paris: Lavoisier Tec&Doc, 416.

Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C., and Desmulliez, J. J., 2007. Performance Indicators and Multicriteria Decision Support for Sewer Asset Management. Infrastructure Systems, 13(2), 105-114.

Le Gauffre, P. and Cherqui, F., 2009. Sewer rehabilitation criteria evaluated by fusion of fuzzy indicators. LESAM 2009, Miami, United States.

McDonald, S.E., and Zhao, J.Q., 2001. Condition assessment and rehabilitation of large sewers. International Conference on Underground Infrastructure Research, University of Waterloo, 361–369.

Mehle, J., O´Keefe, P., and Wrase, P., 2001. An Examination of Methods for Condition Rating of Sewer Pipelines. Thesis (MSc). University of Minnesota.

Micevski, T., Kuszera, G., and Coombes, P., 2002. Markov model for storm water pipe deterioration. Infrastructure Systems, 8(2), 49-56.

Morcous, G., Rivard, H., and Hanna, A.M., 2002. Modeling Bridge Deterioration Using Case-Based Reasoning. Journal of Infrastructure Systems, ASCE, 8(3), 86-95.

NASSCO (National Association of Sewer Service Companies), 2007. Pipeline Assessment And Certificate Program (PACP).

New Mexico Environmental Finance Center, 2006. Asset management: a guide for water and wastewater systems. Available from: http://www.nmenv.state.nm.us/dwb/assistance/documents/AssetManagementGuide.pdf [Accessed 24 August 2014].

ONEMA, 2012. Observatory on public water and sanitation services, overview of services and of their performance. France: French National Agency for Water and Aquatic Environments.

O'Reilly, M.P., Rosbrook, R.B., Cox, G.C., and McCloskey, A., 1989. Analysis of defects in 180km of pipe sewers in southern water authority. Transport and road research laboratory.

Oxford dictionary, 2014. Oxford: Europe.

Saegrov, S., 2006. CARE – S: Computer Aided Rehabilitation of Sewer and Storm Water Networks. IWA Publishing.

Salman, B., 2010. Infrastructure Management and Deterioration Risk Assessment of Wastewater Collection Systems. Thesis (PhD). University of Cincinnati.

Salman, B. and Salem O., 2012. Modeling failure of wastewater collection lines using various section-level regression models. Journal of Infrastructure Systems. 18(2), 146-154.

Sargent, R.G., 1998. Verification and Validation of Simulation Models [online]. New York, Syracuse University. Available from: http://surface.syr.edu/eecs/7 [Accessed 6 May 2014].

Scheidegger, A., Hug, T., Rieckermann, J., and Maurer, M., 2011. Network condition simulator for benchmarking sewer deterioration models. Water Research, 45 (16), 4983-4994.

Schladweiler, J.C., 2002. Tracking Down the Roots of Our Sanitary Sewers. Available from: http://www.sewerhistory.org/chronos/roots.htm [accessed 23 August 2014].

Schlesinger, S., 1979. Terminology for model credibility. Simulation, 32 (3), 103-104.

Schmidt, T., 2009. Modellierung von Kanalalterungsprozessen auf der Basis von Zustandsdaten. Thesis (PhD), Institut für Stadtbauwesen und Straßenbau, Technical University Dresden.

Sousa, V., Matos, J.P., Matias, N., 2014. Evaluation of artificial intelligence tool performance and uncertainty predicting sewer structural condition. Automation in Construction, 44, 84-91.

Stein, R. and Gedheri, S., 2009. Wertermittlung von Abwassernetzen, Germany: Stein & Partner GmbH.

Tran, H.D., 2007. Investigation of Deterioraton Models for Stormwater Pipe Systems. Thesis (PhD). Victoria University.

Ugarelli, R.M., Selseth, I., Le Gat, Y., Rostum, J., and Krogh, A.H., 2013. Wastewater pipes in Oslo: from condition monitoring to rehabilitation planning, Water Practice and Techonology, 8 (3-4), 487-494.

Van der Steen, A.J., Dirksen, J., and Clemens, F.H.L.R., 2013. Visual sewer inspection: detail of coding system versus data quality?. Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance [online]. Available from: http://www.tandfonline.com/doi/abs/10.1080/15732479.2013.816974#.U2kJw1dV9K0 [Accessed 6 May 2014].

WEF–ASCE, 1994. Existing sewer evaluation and rehabilitation. Alexandria: ASCE, ISBN-13: 978-0784400494.

Werey, C., Rozan, A., Wittner, C., Le Gat, Y., Le Gauffre, P., Nirsimloo, K., and 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 et Territoires [online], 9, 44-53. Available from : http://www.set-revue.fr/gestion-patrimoniale-des-reseaux-d-assainissement-de-l-etat-des-reseaux-la-planification-de-leur-reh [Accessed 6 May 2014]

WERF, 2007. Infrastructure Management Final Report: Condition Assessment Strategies and Protocols for Water and Wastewater Utility Assets. US: Water Environment Research Foundation.

WERF, 2009. Strategic Asset Management Final Report: Remaining Asset Life: A State of the Art Review. US: Water Environment Research Foundation.

WERF, 2013. Infrastructure Final Report : Cost Information for Wastewater Pipelines. US: Water Environment Research Foundation.

Wirahadikusumah, R., Abraham, D., and Iseley, T., 2001. Challenging issues in modeling deterioration of combined sewers. Infrastructure Systems, 7, 77-84.

WRc, 2004. Manual of Sewer Condition Classification - 4th Edition. UK: Water Research Centre.

WRc, 2013. Sewerage Risk Management, Glossary and Acronyms. Available from: http://srm.wrcplc.co.uk/glossary.aspx [Accessed 24 August 2014].