# Comparing the utility of decision trees and support vector machines when planning inspections of linear sewer infrastructure

Robert Richard Harvey and Edward Arthur McBean

## ABSTRACT

Closed-circuit television inspection technology is traditionally used to identify aging sewer pipes requiring rehabilitation. While these inspections provide essential information on the condition of pipes hidden from day-to-day view, they are expensive and often limited to small portions of an entire sewer system. Municipalities may benefit from utilizing predictive analytics to leverage existing inspection datasets so that reliable predictions of condition are available for pipes that have not yet been inspected. The predictive capabilities of data mining systems, namely support vector machines (SVMs) and decision tree classifiers, are demonstrated using a case study of sanitary sewer pipe inspection data collected by the municipality of Guelph, Ontario, Canada. The modeling algorithms are implemented using open-source software and are tuned to counteract the negative impact on predictive performance resulting from class imbalance common within pipe inspection datasets. The decision tree classifier outperforms SVM for this classification task – achieving an acceptable area under the receiver operating characteristic curve of 0.77 and an overall accuracy of 76% on a stratified test set. Although predicting individual pipe condition is a notoriously difficult task, decision trees are found to be a useful screening tool for planning future inspection-related activities.

**Key words** | decision tree, infrastructure, inspection, predictive analytics, sewer

**Robert Richard Harvey** (corresponding author)
**Edward Arthur McBean**
School of Engineering,
University of Guelph,
Guelph,
ON,
Canada N1G 2W1
E-mail: *rharvey@uoguelph.ca*

## INTRODUCTION

Sewers play a crucial role in the day-to-day operations of modern cities as they convey enormous volumes of wastewater to treatment centers for safe processing and disposal. Individual pipes within a sewer system often remain in operation well beyond their typical design life of 50–75 years and a growing list of evidence suggests unchecked deterioration of this infrastructure is causing significant harm to the natural environment.

The cracks, holes and fractures that form in aging sewer pipes allow raw, untreated wastewater to leak/exfiltrate into the surrounding environment. As an example, the exfiltration rate for the approximately 850,000 km of sewer pipe currently in operation in Germany is reportedly 6% of the average daily flow, equating to 300 million $m^3$ of leakage per year (Scheyer *et al.* 1991). Exfiltration rates in the UK have been reported to be in the range of 1–5% of average daily flow (Anderson *et al.* 1995; Bishop *et al.* 1998; Yang *et al.* 1999; Ellis *et al.* 2004). Exfiltration rates in North America are also reportedly high as a study carried out by Amick & Burgess (2000) indicates 56% of the average flow was leaking out of sewer pipes studied in California, 49% in Baltimore, 35% in Kentucky and 16% in Washington.

Leaking sewers pose a significant environmental threat as a range of contaminants can be found in typical sewer flows (e.g. bacteria and pathogenic microorganisms) and it has generally been concluded that the poor quality of urban groundwater aquifers in the UK is due to elevated rates of sewer exfiltration (Price & Reed 1989; Reynolds & Barrett 2003). In a study of sandstone aquifers underlying

doi: 10.2166/hydro.2014.007

the cities of Birmingham and Nottingham in the UK, sewage-derived contaminants were found capable of reaching depths of 60–91 m (Powell *et al.* 2003). These findings were supported by the work of Morris *et al.* (2005), where fecal indicators were detected in an aquifer 60 m below ground level. Leaking sewers were the source of X-ray contrast media and boron in groundwater studied in Wolf *et al.* (2004) and the presence of viruses in 33 municipal supply wells studied in Hunt *et al.* (2010) was attributed to the close proximity of each well to aging sanitary sewer pipes. Approximately 90% of 22 drinking water wells recently studied in Guelph, Ontario, Canada contained at least one sewage-derived contaminant and 45% of the wells exhibited human enteric viruses derived from the exfiltration of domestic sewage flows (Allen 2013).

Infiltration of groundwater into sanitary sewer pipes through defective joints, cracks and fractures is also cause for concern as these pipes are only designed to convey wastewater. Excess amounts of infiltration can overload the sanitary sewer system during periods of wet weather, thereby forcing a system overflow. Sanitary sewer overflows are a widespread problem in North America, with an estimated 75,000 occurring every year in the USA, resulting in an annual discharge of several billion gallons of untreated wastewater into the environment (EPA 2004). An increase in wastewater treatment costs will also be directly related to the amount of water that enters a sanitary sewer system as a result of infiltration. The pumps involved with wastewater treatment will need to work harder to handle the increased load which puts unneeded strain on these expensive pumps and shortens their life expectancy. Infiltration is also known to increase the failure probability of adjacent infrastructure, such as paved roads (Kuo *et al.* 2005; Karpf & Krebs 2011). The potential damage from the collapse of an aging sewer pipe can therefore be significant, as in addition to environmental pollution, the collapse may also cause severe interruptions to service and traffic (Zhao 1998). Furthermore, an unexpected sewer pipe collapse can result in expensive emergency repairs having an average unit cost 3.6 times higher than the average unit cost of non-emergency rehabilitation (Zhao & Rajani 2002).

The general consensus among the wastewater industry is that an 'out-of-sight out-of-mind' approach to sewer management only serves to accelerate the inevitable deterioration of sewer pipes and exacerbates exfiltration and infiltration-related issues. For example, the portion of American sewer pipes in 'poor,' 'very poor,' and 'life-elapsed' condition has been projected to increase from 10% of total system size in 2000 to 44% by 2020 if existing sewers are extended to meet increased population growth, but there is no renewal or replacement of the existing pipes (EPA 2002). Alternatively, a more proactive approach to management based on assembling a dataset of pipe condition using visual inspection techniques can help maintain the effectiveness of existing sanitary sewer systems.

Closed-circuit television (CCTV) inspection can be used to visually investigate the condition of individual pipes in a sanitary sewer system. Although CCTV inspections provide valuable information required for rehabilitation planning purposes, they are time-consuming and expensive. As a result, most municipalities are forced to limit inspection-related work to portions of their entire system. Proactive asset management can be supported by modeling techniques that extract information from existing inspection datasets so that predictions of condition can be made for pipes that have not yet been inspected. Efficient approaches need to be made available to municipalities seeking to learn from existing inspection datasets as many existing modeling techniques (e.g. multiple linear regression, logistic regression and Markov chains) are often incapable of reliably predicting individual pipe condition.

The capabilities of support vector machines (SVMs) and decision tree classifiers for predicting individual sewer pipe condition are evaluated in this paper. A sanitary sewer pipe condition dataset collected by the City of Guelph, Ontario, Canada from 2008 to 2011 is used to demonstrate the process of implementing tools for predictive analytics. The described modeling framework represents a simple, yet powerful approach for gaining knowledge of the sewer pipe deterioration process that is novel as it provides a solution to class imbalance problems common in sewer inspection datasets. The concept of a decision tree classifier is general and the resulting model is interpretable by those unfamiliar with the data mining process, a characteristic lacking when other complex statistical or machine learning approaches are applied.

## BACKGROUND INFORMATION ON SEWER PIPE CONDITION MODELING

The majority of sewer deterioration models have been foundationally based on statistical theory. Binary logistic regression models were developed to predict deficiency probability for sewers in Edmonton, Alberta but no goodness-of-fit tests were carried out (Ariaratnam *et al.* 2001). Binary logistic regression categorized the condition of sewer pipe segments in Phoenix, Arizona, but no indication of predictive capability for individual pipes was provided (Koo & Ariaratnam 2006). Bayesian logic and expert opinion were used to develop a sewer cataloguing retrieval and prioritization system presented in Merrill *et al.* (2003). Wright *et al.* (2006) indicate a linear regression model greatly under-estimated the length of deficient pipe and a logistic discriminant model was questionable on a pipe-by-pipe basis in a California sewer system (87% of acceptable pipes misclassified as being deficient). Multiple regression models were trained to predict the condition of concrete, asbestos cement and PVC sewers in Pierrefonds, Quebec and Niagara Falls, Ontario using a small dataset (Chugthai 2007; Chugthai & Zayed 2008). Ordinal regression models developed in Younis & Knight (2010) predicted network-level sewer condition in Niagara Falls, Ontario. Statistical deterioration models developed in Opila (2011) had low $R^2$ values in the range of 0.12 to 0.17. Binary logistic regression models developed in Ens (2012) were incapable of reliably predicting sewer condition.

Deterioration curves for cohorts of sewer pipes in Germany were developed using survival models (Baur & Herz 2002). Survival models tend to underestimate the number of pipes in the poorest condition states (Ana & Bauwens 2010). Markov methods described in Micevski *et al.* (2002) were used to model condition for groups of stormwater pipes in Australia. Similar models developed by Baik *et al.* (2006) using the results of a condition survey of 90 km of sewer pipe in San Diego, California were found to be unsatisfactory as goodness-of-fit scores were low. Markov models and ordinal regression models developed in Tran *et al.* (2008) were deemed unsuitable for pipe-level predictions.

A case-based reasoning approach is described in Fenner *et al.* (2007) where information on condition, performance (e.g. number of previous complaints) and management outcomes (e.g. intervene or non-intervene) for a small number of pipes can be used to proactively manage other pipes in a network. Deterioration models have also been developed using data mining techniques derived from the fields of artificial intelligence and machine learning. Neural networks have been used to predict the frequency and timing of failure in water distribution systems (Tabesh *et al.* 2009; Harvey *et al.* 2014) and the condition of pipes in sewer and stormwater systems (Najafi & Kulandaivel 2005; Tran 2007; Khan *et al.* 2010). A historical database of customer complaints related to pipe blockages was used in Arthur *et al.* (2009) to develop a methodology that incorporated consequence and likelihood of pipe failure to prioritise sewerage maintenance for systems in the UK. A novel data mining technique, Evolutionary Polynomial Regression (EPR), was developed in Savic *et al.* (2006) to identify sewer pipes most likely to fail for a confidential location. EPR was also used in Berardi *et al.* (2008) to predict pipe bursts in a water distribution system and in Berardi *et al.* (2009) to plan sewer pipe inspections based on pipe-specific attributes, condition predicted by a Markov model and an expected cost of failure. Generalized pipe failure prediction models were developed in Savic *et al.* (2009) using an EPR approach that considered data from a number of individual water and sewer systems. SVMs were presented as an alternative to neural networks for predicting sewer condition in South Australia in the work of Mashford *et al.* (2011). SVMs, first developed in the 1970s (Vapnik 1979, 1999), are based on an algorithm that constructs a linear model called the 'maximum margin hyperplane', which is a line (in two dimensions) or a flat plane (in multiple dimensions) that provides the greatest separation between instances with different values of the target variable. Datasets containing instances that cannot be easily separated with a straight line are projected into a higher-dimensional space using a kernel function.

Models developed using neural networks or support vector machines are inherently black boxes, where relationships between inputs and outputs are deeply embedded within the model. A more transparent approach to predicting pipe condition would be decision tree classifiers, such as those developed using the classification and regression tree (CART) algorithm developed by Breiman *et al.* (1984). The CART algorithm is capable of extracting information from mixed datasets (i.e. those datasets containing numerical, categorical and missing data) and requires very little in the way of data pre-processing (Han *et al.* 2006) – characteristics that make the algorithm of potential utility when extracting information hidden within existing pipe inspection datasets. In general terms, the CART algorithm extracts information embedded within an existing dataset for knowledge discovery purposes. The resulting model is presented in a tree-like structure consisting of a root node (containing all the instances in the dataset used to train the model) and branches (illustrating the influence of various input predictors on the target class). The CART algorithm constructs a decision tree classifier using a divide-and-conquer approach, where the branches of the tree are optimally selected so that they lead to homogeneous subsets of data that have a larger proportion of one class than another. Optimal splits for the decision tree are determined by first sorting all the available instances based on their predictor value. A contingency table is generated for each split point (Table 1), and a Gini purity criterion is determined for each potential split for a two-class problem

$$\text{Gini(prior to split)} = 2\left(\frac{n_{1+}}{n}\right)\left(\frac{n_{2+}}{n}\right) \tag{1}$$

$$\text{Gini(after the split)} = 2\left[\left(\frac{n_{11}}{n}\right)\left(\frac{n_{12}}{n_{+1}}\right) + \left(\frac{n_{21}}{n}\right)\left(\frac{n_{22}}{n_{+2}}\right)\right] \tag{2}$$

where $n_{11}$ = the count of class 1 instances greater than the split point, $n_{21}$ = the count of class 1 instances less than or equal to the split point, $n_{1+}$ = the total number of instances in class 1, $n_{12}$ = the count of class 2 instances greater than the split point, $n_{22}$ = the count of class 2 instances less than or equal to the split point, $n_{2+}$ = the total number of instances in class 2, $n_{+1}$ = the total number of instances greater than the split point, $n_{+2}$ = the total number of instances less than or equal to the split point, and $n$ = the total number of instances. The CART algorithm evaluates all possible split points and partitions the dataset using the split point resulting in the minimum Gini purity criterion (Kuhn & Johnson 2013). The tree grows larger as this process continues for each newly created partition until the data cannot be split any further. Eventually, branches of the tree are pruned away using a complexity parameter that is a function of the number of leaves in the tree and the resulting accuracy. Decision tree algorithms are popular within the fields of marketing, finance and medicine but have not yet been used to predict sewer pipe condition using attributes only available prior to condition inspection. Sewer collapse rate and blockage rate deterioration models were developed in Heywood *et al.* (2007) using decision trees. Decision trees were used in Oliveira *et al.* (2007) as part of an exploratory analysis of the relationship between defect severity and pipe attributes. Decision trees in Syachrani *et al.* (2012) were used to visualize the relationship between operational condition and pipe material on an estimated 'real age', reflecting an adjusted pipe age based on location and operational conditions. Jung *et al.* (2012) used decision trees to identify important attributes associated with high-density regions of defective pipes in a mid-sized city.

**Table 1** | Contingency table for split points in the decision tree

| | Class 1 | Class 2 | |
|---|---|---|---|
| >split | $n_{11}$ | $n_{12}$ | $n_{+1}$ |
| ≤split | $n_{21}$ | $n_{22}$ | $n_{+2}$ |
| | $n_{1+}$ | $n_{2+}$ | $n$ |

## CASE STUDY

The municipality of Guelph, Ontario, Canada relies on a 515 km long sanitary sewer system consisting of 7,446 gravity pipes, 43 siphons and 33 pressurized forcemains. The average sewer pipe age in Guelph is 38 years old, including the approximately 1,900 pipes in operation for more than 50 years and the approximately 300 pipes transporting sewage for more than 100 years. Pipe diameters in the sanitary sewer system range from 100 to 1,650 mm and approximately 82% of all pipes have a diameter between 200 and 300 mm. Burial depths range from 0.4 to 10 m (average burial depth is 3.2 m). The proportion of total system length grouped according to pipe material is as follows: asbestos cement (11%), concrete (16.1%), polyvinyl chloride (PVC) (44.6%), reinforced concrete (3.1%), and vitrified clay (25.2%). In general, the oldest pipes in the system are vitrified clay and the majority of pipes constructed within the past 30 years are made of PVC.

The City of Guelph retained an engineering consultancy from 2008 to 2011 to CCTV inspect a portion of their sanitary sewer system and assist in the development of a capital rehabilitation/replacement program for their linear wastewater infrastructure. Pipes were selected for inspection using expert opinion, where pipes older than 50 years received the majority of inspection effort as it was expected these would be in the poorest structural condition. Structural defects were identified using the third edition of the Water Research Center Manual of Sewer Condition Classification (WRc MSCC) (WRc 1996) and severity scores were assigned to the defects using the fourth edition of the Water Research Center Sewerage Rehabilitation Manual (WRc SRM) (WRc 2001). The engineering consultancy assigned each inspected pipe an internal condition grade (ICG) of 1, 2, 3, 4 or 5 based on thresholds established in the WRc SRM for the highest severity scores accumulated in any one meter length of the pipe. Comprehensive quality assurance/quality control (QA/QC) was carried out to ensure accuracy of inspection data.

An analysis of detailed CCTV inspection records indicates the average inspected pipe length was 68 m and 33, 59 and 80% of all structural defects occur within 10, 20, and 30 m of the nearest manhole, respectively. These findings have implications for future inspection-related activity as there is potential to reduce inspection costs by implementing zoom-camera technology. Zoom cameras capture video of the pipe interior using a camera mounted on a pole that is lowered into a manhole but have an effective useable distance of approximately 30 m. Zoom cameras are capable of inspecting up to 1,800 m of pipe per day at a cost of approximately $0.90/m (EPA 2010). CCTV on the other hand can only cover 500 m of pipe per day with an average cost for utilities that outsource inspection of $3.96/m (EPA 2010). While the technology is incapable of providing the same detailed visual evaluation as CCTV, zoom cameras may potentially expedite investigations of uninspected pipe given most structural defects are within the sight distance limit of the technology.

## DATA MINING METHODOLOGY

### Selection of a classification target

Data mining analysis presented herein deals exclusively with 123 km of gravity sanitary sewer inspection data collected from 2008 to 2011 (1,825 individual pipes) performed solely by the primary inspection sub-contractor. A complete list of modeling inputs/attributes contained within the inspection dataset used for model development is presented in Table 2. Stratified random sampling was

**Table 2** | Attributes available for data mining

| Attribute | Type | Description |
|---|---|---|
| Material | Nominal | Pipe material being either asbestos cement, concrete, PVC, reinforced concrete or vitrified clay |
| Age | Numeric | Age at the time of inspection (range: 1–108 years, mean = 50.1 years) |
| Type | Nominal | Type of sewer (trunk or branch) |
| Diameter | Numeric | Pipe diameter (150–900 mm, mean = 270 mm) |
| Length | Numeric | Pipe length (1.70–198.57 m, mean = 68 m) |
| Slope | Numeric | Pipe slope (0–9.97 m per 100 m, mean = 1.47 m per 100 m) |
| Down elevation | Numeric | Downstream invert elevation (304.40–355.95 m, mean = 329 m) |
| Depth | Numeric | Pipe burial depth (0.65–8.88 m, mean = 3.2 m) |
| Road coverage | Numeric | Portion of the pipe covered by a roadway (0–100%, mean = 52%) |
| Watermain breaks (3 m) | Numeric | Number of historical water main breaks within 3 m of the pipe (0–7 breaks, mean = 0.2) |
| Structural condition | Nominal | The target class of whether a pipe is in good (ICG of 1–2–3) or bad (ICG 4–5) structural condition |

Attributes 1–10 represent an input predictor dataset that has irrelevant predictors already removed (i.e. none of the input predictors have near-zero variance, nor are there any between predictor correlations).

used to partition the inspection dataset into separate training, evaluation and test sets using a 70-10-20 split ratio. The inspection dataset is 'class imbalanced', with the majority of inspected pipes assigned an ICG of 1 (acceptable condition), 2 (minimal collapse risk but potential for further degradation) or 3 (collapse unlikely but further deterioration likely). Very few inspected pipes were assigned an ICG of 4 (collapse likely in the near future) or 5 (collapsed or collapse imminent) (Table 3). Class imbalance is common within inspection datasets, as pipes prone to failure tend to have already failed and been replaced prior to CCTV inspection, causing the number of observations for poor condition pipes to be underestimated (Ana & Bauwens 2010). Class imbalance significantly compromises the ability of most algorithms to construct useful models for all condition classes, as model-building efforts focus on correctly classifying pipes in the majority class. As an example, an initial SVM model trained to assign pipes to one of the five ICGs assigned every pipe with an ICG 4 or 5 to the majority classes ICG 1, 2 or 3 and was therefore unsuitable for planning future CCTV work. Similar complications posed by class imbalance were reported in the work of Salman (2010), when a variety of statistical models were investigated for sewers in Cincinnati, Ohio using pipe-specific attributes (e.g. size, length, slope, age, burial depth, and material). The available data violated the proportional odds assumption necessary for ordinal regression model development. The validation set overall accuracy of a multinomial logistic regression model was 53 and 66% for binary logistic regression. A study carried out using class imbalanced datasets from two Belgian municipalities indicated Markov models were inaccurate at the individual pipe level, binary logistic models were 20% accurate for pipes in a failed condition state, and neural network models were unsuitable for forecasting pipe deterioration (Ana 2009).

Techniques currently available for accommodating class-imbalance in data mining applications have primarily been developed for two-class problems. Dealing with class imbalance when more than two classes are involved is considerably more difficult and remains an active area of research within the data mining community (Han *et al.* 2006). This necessitates transforming pipe condition into a two-class format. Guelph considers the relatively rare categories of ICG 4 and 5 to be of great interest as their defects pose an immediate threat to the environment and surrounding infrastructure (although ICG 4 pipes may have some remaining life, it is unwise from both an environmental and economic perspective to delay their rehabilitation). An alternative classification task can therefore be established whereby pipes are classified as being in either a good (ICG 1–3) or bad (ICG 4–5) structural condition state. The rate of bad pipes in the dataset after this transformation is approximately 11% (training set: 141 bad and 1,137 good pipes, evaluation set: 20 bad and 163 good pipes and the test set: 40 bad and 324 good pipes).

## Implementing the algorithms

The SVM model was developed using the 'kernlab' package (Karatzoglou *et al.* 2004) and tuned using the 'caret' package (Kuhn 2013) within the open source software environment *R*. The general purpose radial basis function (RBF) kernel was

**Table 3** | A summary of ICG by pipe material

| | | ICG | | | | | Structural condition | |
|---|---|---|---|---|---|---|---|---|
| **Material** | **Diameter (mm)** | **1** | **2** | **3** | **4** | **5** | **Good** | **Bad** |
| Asbestos cement | 200–400 | 373 | 14 | 24 | 5 | 0 | 411 | 5 |
| Concrete | 200–900 | 364 | 79 | 73 | 36 | 7 | 516 | 43 |
| PVC | 200–450 | 110 | 9 | 1 | 2 | 1 | 120 | 3 |
| Reinforced concrete | 300–900 | 56 | 8 | 0 | 0 | 0 | 64 | 0 |
| Vitrified clay | 200–450 | 213 | 135 | 165 | 121 | 29 | 513 | 150 |
| Total | 150–900 | 1116 | 245 | 263 | 164 | 37 | 1624 | 201 |

used to project the data

$$k(a, b) = \exp\left(-\sigma\|a - b\|^2\right) \tag{3}$$

where $k$ is the kernel function, $a$ and $b$ represent two instances of pipe condition and $\sigma$ is a hyper parameter that is automatically determined using an algorithm within the 'kernlab' package. An additional cost parameter ($C$) included as part of the SVM optimization objective function for the RBF kernel was optimally determined using three repeats of 10-fold cross-validation of the training dataset. Predictor inputs presented to the SVM were centered and scaled to have a mean of zero and standard deviation of one. Recursive feature elimination was implemented to identify the predictive benefit of removing any non-informative predictors.

The CART algorithm was implemented using the 'rpart' (Therneau *et al.* 2006) and 'caret' (Kuhn 2013) packages for *R*. Optimal tree size was determined using three repeats of 10-fold cross-validation of the training dataset. Decision tree algorithms are largely insensitive to the characteristics of the predictor data, and no data pre-processing was required to improve predictive performance for the given set of input predictors. Whereas the predictive success of SVMs may hinge on presenting an appropriate subset of features to the algorithm, decision trees implicitly perform feature selection (ensuring predictors that do not contribute to the predictive power of the model are ignored).

## Evaluating predictive performance

Continuous valued predictions in the form of a class membership probability between 0 and 1 generated by the models are used to establish predictions of pipe condition (good vs. bad) using a default classification threshold/cutoff of 0.50 and predictive performance on a dataset with known class labels can be evaluated using the confusion matrix shown in Table 4. When pipes in bad condition are considered the positive class of interest, correctly classified pipes are represented by TP = true positive (pipe actually in bad condition correctly predicted to be in bad condition) and TN = true negative (pipe actually in good condition correctly predicted to be good). Incorrect classifications are represented by FP = false positive (pipe predicted to be bad, when in fact, it is not), and FN = false negative (pipe predicted to be good, when it is actually bad – essentially, saying there is nothing to worry about, when there actually is). Using these definitions, a model's predictive accuracy is defined by

**Table 4** | Confusion matrix for a binary classification task

| | | Predicted condition | |
|---|---|---|---|
| | | **Bad (ICG 4–5)** | **Good (ICG 1–3)** |
| Actual condition | Bad (ICG 4-5) | True positive | False negative |
| | Good (ICG 1-3) | False positive | True negative |

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{4}$$

When dealing with class imbalance, accuracy on its own can be misleading as a trivial classifier that predicts every pipe as belonging to the majority good class can achieve high accuracy. The accuracy metric assumes false positives and false negative errors have the same costs, when in reality false negatives will have more severe consequences. As a result, a series of alternative metrics should be used when evaluating predictive capability

$$\text{True positive rate} = \text{TPR} = \text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{True negative rate} = \text{TNR} = \text{specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \tag{6}$$

$$\text{False positive rate} = \text{FPR} = 1 - \text{specificity} \tag{7}$$

$$\text{False negative rate} = \text{FNR} = 1 - \text{sensitivity} \tag{8}$$

A useful visual tool for contrasting the predictive performance of two models is the receiver operating characteristic (ROC) curve, which is a plot of the TPR and FPR achieved across the full continuum of probability thresholds that could be used to make discrete class predictions. The area under the ROC curve can be used to gauge model performance, where perfect models have an area under the ROC curve of 1 and random models have an area under the ROC curve close to 0.5 (Fawcett 2006). An area under the ROC curve greater than 0.7 on a stratified test set would be considered acceptable (Hosmer & Lemeshow 2000). The model with the largest area under the ROC curve can be considered to be most effective for the classification task at hand.

The default threshold of 0.50 used to determine discrete class predictions tends to be unsuitable when working with class imbalance. The models were tuned to enhance predictive accuracy on the positive minority class using an alternative classification threshold that effectively changed the definition of a predicted event. The evaluation set ROC curve can be used to derive a new cut-off, where the threshold closest to the upper left corner of the ROC curve is optimal (as it is closest to a perfect model). This threshold moving approach has been shown to outperform some other popular class-imbalance learning techniques (e.g. down-sampling and up-sampling) (Han *et al.* 2006).

## RESULTS

### Support vector machine model

The SVM model achieving the highest area under the ROC curve during three repeats of 10-fold cross-validation of the training dataset had the following parameters: full-set of input predictors, $\sigma = 0.09$, $C = 256$, support vectors = 381, and area under ROC curve = 0.69. The test set confusion matrix using this optimal SVM design and the default classification threshold of 0.50 is shown in Table 5(a). Although overall accuracy was 89%, the confusion matrix indicates the algorithm focused entirely on correctly classifying the majority class, resulting in a true positive rate of 0%.

An optimal classification threshold derived from the point closest to the top left of the evaluation set ROC curve was 0.11 (Figure 1). The test set confusion matrix obtained when using this optimal threshold to reclassify predictions of condition in the test set is shown in Table 5(b). This optimal threshold results in the following performance metrics on the test set: TPR = 83%, TNR = 54%, accuracy = 58% and an area under the ROC curve = 0.72.

### Decision tree classifier

The decision tree classifier achieving the highest area under the ROC curve during three repeats of 10-fold cross-validation of the training dataset had a complexity parameter for pruning of 0.002 and achieved a cross-validated area under the ROC curve of 0.71. A visual depiction of the decision tree is provided in Figure 2. The test set confusion matrix using this optimal SVM design and the default classification threshold of 0.50 is shown in Table 6 (a). Although overall accuracy was 89%, the confusion matrix indicates the algorithm focused on correctly classifying the majority class, resulting in a true positive rate of only 5%.

**Table 5** | SVM model confusion matrix for test set

**(a) Test set (Cutoff = 0.50)**

| | | Predicted condition | |
|---|---|---|---|
| | | **Bad (+)** | **Good (−)** |
| | **Bad (+)** | 0 | 40 |
| **Actual condition** | **Good (−)** | 1 | 323 |

TPR = 0/(0 + 40) = 0%

TNR = 323/(1 + 323) = 99.9%

FPR = 1/(1 + 323) = 0.01%

FNR = 40/(0 + 40) = 100%

Accuracy = (0 + 323)/(363) = 89%

Area under ROC curve = 0.73

**(b) Test set (Cutoff = 0.11)**

| | | Predicted condition | |
|---|---|---|---|
| | | **Bad (+)** | **Good (−)** |
| | **Bad (+)** | 33 | 7 |
| **Actual condition** | **Good (−)** | 148 | 176 |

TPR = 33/(33 + 7) = 83%

TNR = 176/(148 + 176) = 54%

FPR = 148/(148 + 176) = 46%

FNR = 7/(33 + 7) = 17%

Accuracy = (33 + 176)/(363) = 58%

Area under ROC curve = 0.73

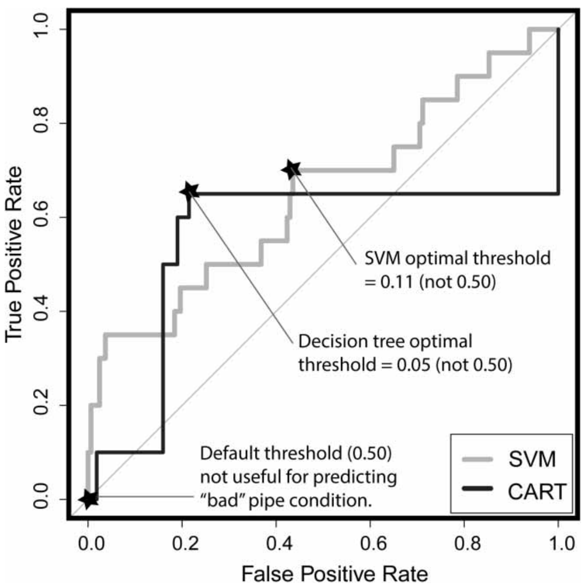


**Figure 1** | The evaluation set ROC curves developed for the SVM and decision tree models.

The evaluation set area under the ROC curve for the decision tree model was 0.75 and an optimal classification threshold derived from the point closest to the top left of the evaluation set ROC curve was 0.05 (Figure 1). The test set confusion matrix obtained when using this optimal threshold (Table 6(b)) indicates the following performance metrics: TPR 78%, TNR 76%, accuracy = 76% and area under the ROC curve = 0.77.

## DISCUSSION

### Comparing the SVM and decision tree predictive models

In terms of the area under the ROC curve metric, the decision tree model performed better than the SVM model during cross-validation of the training set (0.71 vs. 0.69) and during model testing (0.77 vs. 0.72). The decision tree model also achieved a higher level of accuracy on the test dataset than the SVM model (76% vs. 58%).

In terms of knowledge discovery, the branches and leaves of the decision tree provide insight into the role of various pipe-specific attributes on determining pipe condition. The root of the decision tree shown in Figure 2 indicates 11% of the pipes in the training set were in bad condition and 89% were in good condition. The first split of the tree off of this root node illustrates the influence of time on pipe condition, with 5% of inspected pipes less than 50 years old being in bad condition (node 2) compared to 24% of those more than 50 years old found to be in bad condition (leaf 3). Nodes 4 and 5 of the decision tree indicate pipe burial depth influences the likelihood of older pipes being in a structurally deteriorated condition state. Pipes more than 50 years old with burial depths less than 1.9 m have a 42% chance of being in bad condition (node 4) compared to a 21% chance for pipes buried at greater depths (node 5). The literature indicates pipe defect rates and the likelihood of collapse tend to increase as burial depths decrease (Cullen 1982; Jones 1984; O'Reilly *et al.* 1989; Fenner & Sweeting 2000; Savic *et al.* 2006; Berardi *et al.* 2009). Water and wastewater infrastructure deterioration in the municipality are related, as sewer pipes more than 50 years old, with a burial depth less than 1.9 m that have experienced at least one nearby watermain break have an 80% likelihood of being in bad condition (leaf 6), compared to a 33% chance for similar sewer pipes with no nearby watermain breaks (leaf 7). Structural condition is influenced by the length of an individual pipe, where 24% of older pipes with a burial depth greater than 1.9 m and a length greater than 33 m are in bad condition (node 8). This is significantly higher than the 5% of similar pipes with a length less than 33 m that are in bad condition (leaf 9). Longer pipe are known to be more vulnerable to differential settlement and have more possible locations of pipe failure (e.g. joints) (Ana *et al.* 2008). Node 10 and leaf indicate the influence of pipe diameter on pipe condition, where smaller diameter pipes have a 26% chance of being in bad condition, compared to a 12% chance for larger diameter pipes. It is possible that larger diameter pipes have a reduced likelihood of being in bad condition due to the tendency for larger diameter pipes to be installed by experienced personnel, reducing the likelihood of defects related to installation error (Davies *et al.* 2001; Savic *et al.* 2006). An additional split of smaller diameter pipes (leaves 12 and 13) again indicates that a history of nearby watermain failures can increase the likelihood of an older sewer pipe being in bad condition.

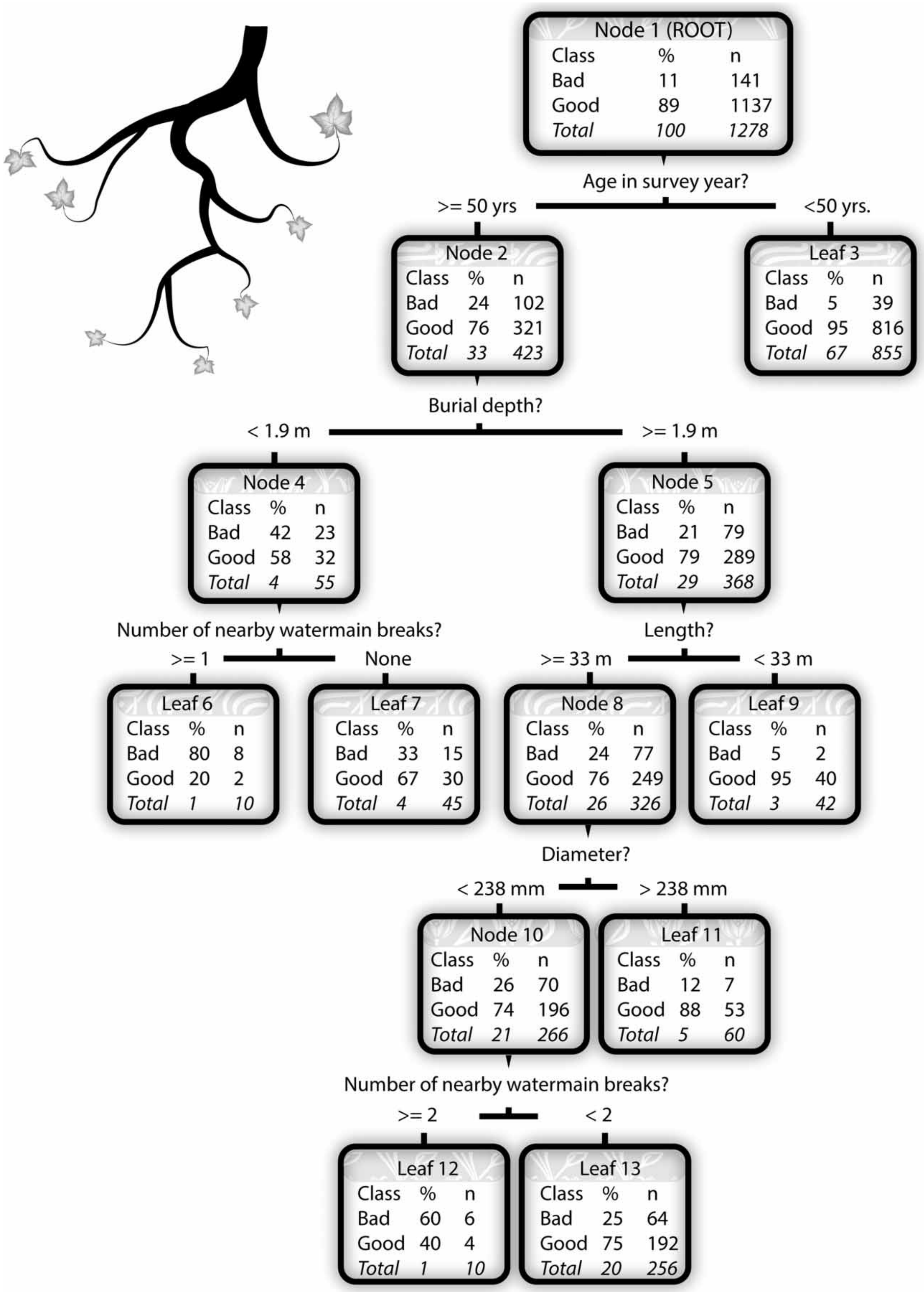


**Figure 2** | A visual depiction of the decision tree predictive model for sewer pipe condition.

**Table 6** | Decision tree model confusion matrix for test set

**(a) Test set (Cutoff = 0.50)**

| | | Predicted condition | |
|---|---|---|---|
| | | **Bad (+)** | **Good (–)** |
| | **Bad (+)** | **1** | **39** |
| **Actual condition** | **Good (–)** | **0** | **324** |

TPR = 1/(1 + 39) = 3%

TNR = 324/(0 + 324) = 100%

FPR = 0/(0 + 324) = 0%

FNR = 39/(1 + 39) = 97%

Accuracy = (1 + 324)/(363) = 89%

Area under ROC curve = 0.78

**(b) Test set (Cutoff = 0.05)**

| | | Predicted condition | |
|---|---|---|---|
| | | **Bad (+)** | **Good (–)** |
| | **Bad (+)** | **31** | **9** |
| **Actual condition** | **Good (–)** | **79** | **245** |

TPR = 31/(31 + 9) = 78%

TNR = 245/(79 + 245) = 76%

FPR = 79/(79 + 245) = 24%

FNR = 9/(31 + 9) = 22%

Accuracy = (31 + 245)/(363) = 76%

Area under ROC curve = 0.78

## Screening pipes for future inspection

The decision tree classifier can be used to predict the condition of uninspected pipes in the Guelph system. Of the more than 4,000 uninspected pipes (the large majority of which are less than 50 years old), a total of 876 pipes are classified as being in bad condition. For the next round of inspection, the municipality can choose to inspect all 876 predicted bad condition pipes, knowing that some of them will actually be in good condition. Alternatively, the municipality can use the class probability scores output by the decision tree to select a subgroup of predictions that have a higher proportion of bad pipes than in the original dataset. This approach is similar to one taken when advertising firms, looking to decrease marketing costs, use predictive analytics to identify customers that are most likely to respond to mail-out campaigns. Considering the 365 instances in the test set, if all predictions of pipe condition are sorted in order of decreasing probability of being bad, the gain for a given sample size (percentile) is

$$\text{Gain} = \left(\frac{\#\,\text{of pipes in the sample that are actually bad}}{\text{Total}\,\#\,\text{of bad pipes}}\right) \times 100\%$$

Figure 3 presents the gain associated with various sampling scenarios, indicating the top 20% of pipes sorted by decreasing propensity contain approximately 68% of all bad pipes in the dataset. The top 73 predictions (20% × 365) in the test set contain 27 pipes that were actually observed to be bad. The entire test set only contained 40 bad pipes, so 20% of the decision tree's top predictions capture 27/40 = 68% of all the bad pipes. The diagonal line indicates the expected response when inspections are carried out without the predictive model (i.e. using the same expert opinion used when planning the first round of inspections), indicating 248 or 68% of all 365 pipes in the test set would need to be inspected to capture 68% of all the bad pipes. The model provides an opportunity to carry out significantly fewer inspections to identify the same number of bad pipes, resulting in considerable time and cost savings.

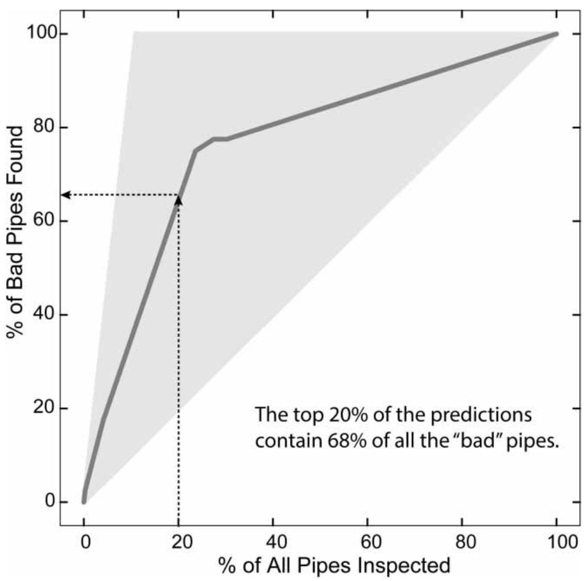


**Figure 3** | Gain chart for the decision tree model.

## The impact of dataset size

A sensitivity analysis can be carried out to examine the potential to develop decision tree models using less extensive datasets as such models would be useful for those municipalities in the early stages of a planned inspection programe. Guelph's original inspection dataset can be split into two – the first containing 30% of the original dataset (representing a smaller, hypothetical inspection programe of 547 pipes) and the second containing the other 70% of the original dataset (representing a hypothetical set of 1,278 uninspected pipes). Given that the condition of the pipes in the latter dataset is available for analysis (141 bad and 1,137 good pipes), it can be used to gauge the predictive utility for smaller datasets

The hypothetical inspection of programe of 547 pipes would have identified 347 good pipes and only 35 bad pipes, and the CART algorithm can use these instances of pipe condition to construct a decision tree classifier. Although there were fewer instances of bad pipes available for learning purposes, which inevitably lowered the predictive capabilities of the model when compared to a model developed with a larger dataset (three repeats of 10-fold cross-validation area under $\mathrm{ROC} = 0.65$), the trained model was found to be an effective tool for screening pipes for condition inspection. The decision tree model achieved a true positive rate of 70% and a true negative rate of 70% for the 1,278 pipes in the hypothetical set of uninspected pipes (for which, in this sensitivity analysis, the condition state is known already). Rather than inspecting every pipe predicted to be bad, the top $20\% \times 1{,}278$ pipes $= 256$ inspections would have identified 73 pipes that were actually in bad condition. There were only 141 bad pipes in the dataset, therefore inspecting only 20% of all the pipes captured $73/141 = 52\%$ of all the remaining bad pipes in the network that have not yet been inspected. The decision tree provides an opportunity to perform 256 inspections instead of $52\% \times 1{,}278 = 665$ inspections (the number required to achieve the same result if data mining was not carried out), indicating the algorithm continues to be valuable even for smaller datasets. Given that the average CCTV inspection cost in North America is approximately \$2/m and the average inspected pipe length in Guelph was approximately 70 m, the opportunity to capture more information on bad condition in fewer inspections can result in considerable cost savings to the municipality. For this hypothetical example, implementing predictive analytics after the first round of inspections are performed so that 73 bad condition pipes can be identified in 256 inspections instead of 665 inspections, representing an inspection cost difference of approximately \$35,800 vs. \$93,100.

## Improving overall accuracy and future directions of research

The decision tree classifier was developed without information related to soil condition in Guelph and this likely had an impact on predictive accuracy. Soil corrosion potential (a product of particle size, uniformity, organic content, etc.) has been known to affect the rate of exterior pipe corrosion. Municipalities looking to implement decision-tree based predictive analytics should include detailed soil records if they are available as this information may increase the predictive capability of a developed model. Information on maintenance activities for individual assets would also potentially improve predictive performance for any developed model. As class-imbalance within inspection datasets provides an obstacle for model development, municipalities should direct inspection efforts towards identifying pipes that are most likely in bad condition. Beyond providing information required for planning necessary rehabilitative action, the new bad pipe condition data would help to alleviate class-imbalance and increase the likelihood the data mining algorithm can learn to accurately predict pipe condition.

Although the developed decision tree model is easier to interpret than many existing deterioration modeling techniques, the use of a single decision tree results in a general trade-off between predictive performance and the simplicity of the overall model structure. As such, future modeling efforts can be directed towards other data mining systems (e.g. random forests that combine the predictive power of many individual decision trees) that may potentially improve predictive accuracy as they are known to be effective when dealing with imbalanced datasets. Investigations into the utility of decision-tree based predictive models developed within a risk assessment framework are recommended as a future direction for research – where predicted condition paired to risk-based concepts (e.g. WRc critical sewers (WRc 2001) or other utility-specific asset importance criteria) might enhance the process of screening pipes for inspection based on both the likelihood and consequence of failure.

## CONCLUSIONS

An examination of CCTV inspections carried out in Guelph, Ontario indicates the most common structural defects within Guelph's sanitary sewer pipes are cracks, fractures and defective joints, representing 42, 38 and 11% of all recorded structural defects, respectively. Inspection records indicate 33.2, 59.3 and 80.3% of all recorded structural defects occur within 10, 20, and 30 m of the nearest manhole, respectively. These findings suggest future inspection efforts in Guelph may benefit from zoom-camera inspection technologies that are faster than CCTV but often have their utility limited unless sight distance is less than 30 m.

Many proposed approaches to modeling sewer pipe deterioration are unsuited to predicting individual pipe condition and may provide municipalities with a spurious impression of their true predictive power. The predictive capabilities of SVMs and decision tree classifiers are evaluated. Decision trees are found to be a simple and effective method of gaining deeper insight into the influence of pipe-specific parameters on the structural condition of individual pipes. Transforming the classification task into a binary format of good vs. bad pipe condition and then adjusting thresholds used to make classification decisions accommodates class imbalance common within pipe inspection datasets.

The predictive capabilities of the models were evaluated using a variety of metrics suitable for assessing imbalanced datasets. The developed decision tree had a higher area under the ROC curve on a stratified test set than the SVM model. The decision tree achieved an acceptable cross-validated area under the ROC curve of 0.71 during training and a test set accuracy of 76%, true positive rate = 78% and a true negative rate = 76%. Overall the model obtained a high prediction rate for bad pipes without sacrificing a reasonable level of accuracy for the good pipes, suggesting the model can guide inspection future work towards uninspected pipes that have a high likelihood of being in bad structural condition.

The trained decision tree classifier indicates pipe age is important for determining the structural condition of sanitary sewers in Guelph, with pipes more than 50 years old having a significantly greater chance of being in bad structural condition than newer pipes in the same sanitary system. The decision tree also illustrates the increased likelihood of pipe failure associated with shallow burial depths, longer pipe lengths, smaller diameters, and the presence of nearby watermain failures for select subsets of sanitary pipes. The developed model provides the municipality with an opportunity to learn from an existing inspection dataset so that bad pipe yield can be significantly higher during future inspection programs. Overall, the combination of inspection and efficiently implemented, open-source data mining techniques presented in this paper may result in significant cost savings for municipalities looking to ensure publically owned sewer systems are being effectively managed.

## ACKNOWLEDGEMENTS

The authors would like to thank the City of Guelph for the provision of the sanitary sewer inspection records. Thank you to the Department of Geography and the Data Resource Center for their assistance with geospatial analysis. The University of Guelph, the Natural Sciences and Engineering Council of Canada and the Canada Research Chairs program funded this research.

## REFERENCES

Allen, A. 2013 Vulnerability of a Fractured Bedrock Aquifer to Emerging Sewage-Derived Contaminants and their use as Indicators of Virus Contamination. Master of Science, Environmental Science, University of Guelph, Guelph, Ontario, Canada.

Amick, R. & Burgess, E. 2000 Exfiltration in Sewer Systems. Report EPA/600/R-01/034. National Risk Management Research Laboratory, United States Environmental Protection Agency, Cincinnati, Ohio.

Ana, E. 2009 Sewer Asset Management – Sewer Structural Deterioration Modeling and Multi-criteria Decision Making in Sewer Rehabilitation Projects Prioritization. Doctor in Engineering, Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels.

Ana, E. & Bauwens, W. 2010 Modeling the structural deterioration of urban drainage pipes. *Urban Water J.* **79** (1), 1069–1079.

Ana, E., Bauwens, W., Pessemier, M., Thoeye, C., Smolders, S., Boonen, I. & Gueldre, G. D. 2008 Investigating the effects of specific attributes on sewer aging – a Belgian case study. In: *11th International Conference on Urban Drainage*, Edinburgh, Scotland, pp. 1–10.

Anderson, G., Misstear, B. & White, M. 1995 *Reliability of Sewers in Environmentally Vulnerable Areas – Technical Report 44*. Construction Industry Research and Information Association, London, UK.

Ariaratnam, S., ElAssaly, A. & Tang, Y. 2001 Assessment of infrastructure inspection needs using logistic models. *J. Infrastruct. Syst.* **7** (4), 160–165.

Arthur, S., Crow, H., Pedezert, L. & Karikas, N. 2009 The holistic prioritisation of proactive sewer maintenance. *Water Sci. Technol.* **59** (7), 1385–1396.

Baik, H., Jeong, H. & Abraham, D. 2006 Estimating transition probabilities in Markov-chain based deterioration models for management of wastewater systems. *J. Water Resour. Plan. Manage.* **132** (1), 15–24.

Baur, R. & Herz, R. 2002 Selective inspection planning with ageing forecast for sewer types. *Water Sci. Technol.* **46** (6), 389–396.

Berardi, L., Kapelan, Z., Giustolisi, O. & Savic, D. 2008 Development of pipe deterioration models for water distribution systems using EPR. *J. Hydroinform.* **10** (2), 113–126.

Berardi, L., Giustolisi, O., Savic, D. A. & Kapelan, Z. 2009 An effective multi-objective approach to prioritisation of sewer pipe inspection. *Water Sci. Technol.* **60**, 841–850.

Bishop, P., Misstear, B., White, M. & Harding, N. 1998 Impacts of sewers on groundwater quality. *J. Inst. Water Environ. Manage.* **12** (3), 216–223.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. 1984 *Classification and Regression Trees*. CRC Press, Boca Raton, FL.

Chugthai, F. 2007 Integrated Condition Assessment Models for Sustainable Sewer Pipelines. M.A. Sc., Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada.

Chugthai, F. & Zayed, T. 2008 Infrastructure condition prediction models for sustainable sewer pipelines. *J. Perform. Construct. Facil.* **22** (5), 333–341.

Cullen, N. 1982 The sewer deterioration problem – evidence from collapse studies. In: *Restoration of Sewerage Systems*. Thomas Telford, London, UK.

Davies, J., Clarke, B., Whiter, J. & Cunningham, R. 2001 Factors influencing the structural deterioration and collapse of rigid sewer pipes. *Urban Water J.* **3** (1), 73–89.

Ellis, J., Revitt, D., Blackwood, D. & Gilmour, D. 2004 Leaky sewers: assessing the hydrology and impact of exfiltration in urban sewers. In: *Hydrology: Science and Practice for the 21st Century* (B. Webb, ed.). British Hydrological Society, London, pp. 266–271.

Ens, A. 2012 Development of a Flexible Framework for Deterioration Modeling in Infrastructure Asset Management. Master of Applied Science, Department of Civil Engineering, University of Toronto, Toronto, Ontario, Canada.

EPA 2002 *The Clean Water and Drinking Water Infrastructure Gap Analysis*. United States Environmental Protection Agency – Office of Water, Washington, DC.

EPA 2004 Report to Congress – impacts and control of CSOs and SSOs. Report EPA 833-R-04-001. United States Environmental Protection Agency, Office of Water, Washington, DC.

EPA 2010 Condition assessment of wastewater collection systems. Report EPA-600-R-10-0. United States Environmental Protection Agency.

Fawcett, T. 2006 An introduction to ROC analysis. *Pattern Recognit. Lett.* **27** (8), 861–874.

Fenner, R. & Sweeting, L. 2000 A new approach for directing proactive sewer maintenance. *Water Maritime Eng. J.* **124** (2), 67–78.

Fenner, R., McFarland, G. & Thorne, O. 2007 Case-based reasoning approach for managing sewerage assets. *Proc. ICE-Water Manage.* **160** (1), 15–24.

Han, J., Kamber, M. & Pei, J. 2006 *Data Mining – Concepts and Techniques*. Morgan Kaufmann Publishers, Waltham, MA, USA.

Harvey, R., McBean, E. & Gharabaghi, B. 2014 Predicting the timing of watermain failure using artificial neural networks. *ASCE J. Water Resour. Plann. Manage.* **140** (4), 425–434.

Heywood, G., Pearman, C. & Lumbers, J. 2007 Advances in the modelling and optimization of sewerage infrastructure investment planning. In: *International Water Association Leading Conference on Strategic Asset Management*, Lisbon, Portugal, pp. 1–14.

Hosmer, D. & Lemeshow, S. 2000 *Applied Logistic Regression*. Wiley Publishing, New York, USA.

Hunt, R., Bouchardt, M., Richards, K. & Spencer, S. 2010 Assessment of sewer source contamination of drinking water wells using tracers and human enteric viruses. *Environ. Sci. Technol.* **44** (20), 7956–7963.

Jones, G. 1984 The structural deterioration of sewers. In: *International Conference on the Planning, Construction, Maintenance and Operation of Sewerage Systems*, Reading, UK, p. 499.

Jung, J., Garrett, J., Soibelman, L. & Lipkin, K. 2012 Application of classification models and spatial clustering analysis to a sewage collection system of a mid-sized city. In: *International Conference on Computing in Civil Engineering* (R. Issa, ed.). ASCE, Clearwater Beach, Florida, pp. 537–544.

Karatzoglou, A., Smola, A., Hornik, K. & Zeileis, A. 2004 kernlab – An s4 package for kernel methods in R. *J. Statist. Softw.* **11** (9), 1–20.

Karpf, C. & Krebs, P. 2011 Quantification of groundwater infiltration and surface water inflows in urban sewer networks based on a multiple model approach. *Water Res.* **45** (10), 3129–3136.

Khan, Z., Zayed, T. & Moselhi, O. 2010 Structural condition assessment of sewer pipelines. *J. Perform. Construct. Facil.* **24** (2), 170–179.

Koo, D. & Ariaratnam, S. 2006 Innovative method for assessment of underground sewer pipe condition. *Autom. Construct.* **15** (4), 479–488.

Kuhn, M. 2013 The caret package. http://caret.r-forge.r-project.org/ (accessed October 1, 2013).

Kuhn, M. & Johnson, K. 2013 *Applied Predictive Modeling*. Springer Science and Business Media, New York.

Kuo, S., Zhao, L., Mahgoub, H. & Suarez, P. 2005 *Investigation of Ground Penetrating Radar for Detection of Leaking Pipelines under Roadway Pavements and Development of Fiber-wrapping Repair Technique – Report Number 48*. University of Central Florida – The Center for Advanced Transportation Systems, Florida, USA.

Mashford, J., Marlow, D., Tran, D. & May, R. 2011 Prediction of sewer condition grade using support vector machines. *J. Comput. Civil Eng.* **25** (4), 283–290.

Merrill, S., Lukas, A., Palmer, R. & Hahn, M. 2003 Development of a tool to prioritize sewer inspections. Report 97-CTS-7. Water Environment Research Federation.

Micevski, T., Kuczera, G. & Coombes, P. 2002 Markov model for stormwater pipe deterioration. *J. Infrastruct. Syst.* **8** (2), 49–56.

Morris, B., Rueedi, J., Cronin, A. & Whitehead, E. 2005 Case study city Doncaster: field investigations – final report. In: EU Project for Assessing and Improving Sustainability of Urban Water Resources and Systems.

Najafi, M. & Kulandaivel, G. 2005 Pipeline condition prediction using neural network models. In: *Pipelines 2005 – Optimizing Pipeline Design, Operations and Maintenance*. American Society of Civil Engineers, Houston, Texas, USA, pp. 767–775.

O'Reilly, M., Rosbrook, R., Cox, G. & McCloskey, A. 1989 *Analysis of Defects in 180 km of Pipe Sewers in the Southern Water Authority – Technical Report 172*. Transport and Road Research Laboratory, UK.

Oliveira, D., Guo, W., Soibelman, L. & Garrett Jr, J. G. 2007 Spatial data management and analysis in sewer system condition assessment: an overview (L. Soibelman & B. Akinci, eds). *International Workshop on Computing in Civil Engineering*. American Society of Civil Engineers, Pittsburgh, PA, pp. 391–398.

Opila, M. 2011 Structural Condition Scoring of Buried Sewer Pipes for Risk-based Decision Making. PhD, Department of Civil and Environmental Engineering, University of Delaware, Newark, Delaware.

Powell, K., Taylor, R., Cronin, A., Barrett, M., Pedley, S., Sellwood, J., Trowsdale, S. & Lerner, D. 2003 Microbial contamination of two urban sandstone aquifers in the United Kingdom. *Water Resour.* **37** (2), 339–352.

Price, M. & Reed, D. 1989 The influence of mains leakage and urban drainage on groundwater levels beneath conurbations in the UK. *Proc. Inst. Civil Eng. Part 1* **86**, 31–39.

Reynolds, J. & Barrett, M. 2003 A review of the effects of sewer leakage on groundwater quality. *Water Environ. J.* **17** (1), 34–39.

Salman, B. 2010 Infrastructure Management and Deterioration Risk Assessment of Wastewater Collection Systems. PhD, Department of Civil and Environmental Engineering, University of Cincinnati, Ohio.

Savic, D., Giustolisi, O., Berardi, L., Shepherd, W., Djordjevic, S. & Saul, A. 2006 Modelling sewer failure by evolutionary computing. *Water Manage. J.* **159** (2), 111–118.

Savic, D., Giustolisi, O. & Laucelli, D. 2009 Asset deterioration analysis using multi-utility data and multi-objective data mining. *J. Hydroinform.* **11** (3–4), 211–224.

Scheyer, R., Milde, G. & Milde, K. 1991 Development of aquifer protection policy in Germany. In: Proceedings IWEM Annual Symposium on Groundwater Pollution and Aquifer Protection in Europe. Palais des Congrès, Paris, France.

Syachrani, S., Jeong, H. & Chung, C. 2012 Decision tree based deterioration model for buried wastewater pipelines. *J. Perform. Construct. Facil.* **27** (5), 633–645.

Tabesh, M., Soltani, J., Farmani, R. & Savic, D. A. 2009 Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. *J. Hydroinform.* **11** (1), 1–17.

Therneau, T. M., Atkinson, B. & Ripley, M. B. 2006 The rpart package. http://cran.r-project.org/web/packages/rpart/rpart.pdf. Retrieved October 6, 2006.

Tran, D. 2007 Investigation of Deterioration Models for Stormwater Pipe Systems. PhD Thesis, Victoria University, Melbourne.

Tran, D., Ng, A., McManus, K. & Burns, S. 2008 Prediction models for serviceability deterioration of stormwater pipes. *Struct. Infrastruct. Eng.* **4**, 287–295.

Vapnik, V. 1979 *Estimation of Dependencies Based on Empirical Data* [English translation]. Springer-Verlag, New York.

Vapnik, V. 1999 *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.

Wolf, L., Held, I., Eiswirth, M. & Hotzel, H. 2004 Impact of leaky sewers on groundwater quality. *Acta Hydrochim. Hydrobiol.* **32** (4–5), 361–373.

WRc 1996 *Manual of Sewer Condition Classification*, 3rd edn. Water Research Centre, Swindon, UK.

WRc 2001 *Sewerage Rehabilitation Manual*, 4th edn. Water Research Center, Marlow, UK.

Wright, L., Heaney, J. & Dent, S. 2006 Prioritizing sanitary sewers for rehabilitation using least-cost classifiers. *ASCE J. Infrastruct. Syst.* **12** (3), 174–183.

Yang, Y., Lerner, D., Barrett, M. & Tellam, J. 1999 Quantification of groundwater recharge in the city of Nottingham. *Environ. Geol.* **38** (3), 183–198.

Younis, R. & Knight, M. 2010 A probability model for investigating the trend of structural deterioration in wastewater pipelines. *Tunnel. Underground Space Technol.* **25** (6), 670–680.

Zhao, J. 1998 Trunk sewers in Canada. In: APWA International Public Works Congress NRCC/CPWA Seminar Series – Innovations in Urban Infrastructure, Las Vegas, pp. 75–89.

Zhao, J. & Rajani, B. 2002 *Construction and Rehabilitation Costs for Buried Pipe with a Focus on Trenchless Technologies*. NRC Institute for Research in Construction, Canada.

First received 8 January 2014; accepted in revised form 31 March 2014. Available online 12 May 2014