https://doi.org/10.71573/z36dgw29

Application of a predictive machine-learning model to forecast sewer's pipes condition. A case study in Lausanne, Switzerland

Francesco Del Punta [1,*] https://orcid.org/0009-0008-1303-6869,
Hauke Sonnenberg [1] https://orcid.org/0000-0001-9134-2871, *Antoine Daurat* [1], *Yoann Sadowski* [2],
Frederic Cherqui [3] https://orcid.org/0000-0001-8836-3970 & *Nicolas Caradot* [1] https://orcid.org/0000-0002-5252-4880

[1]*Kompetenzzentrum Wasser Berlin, Grunewaldstraße 61-62, 10825 Berlin, Germany*
[2]*Ville de Lausanne - Service de l 'eau, Rue de Genève 36, 7416-1001 Lausanne, Switzerland*
[3]*Institut National des Sciences Appliquées - Lyon, 20 avenue Albert Einstein, 69621 Villeurbanne*

**Corresponding author email: francesco.del-punta@kompetenz-wasser.de*

Abstract

This study explores the application of a machine learning model, specifically a Random Forest classifier, to predict the condition of uninspected pipes using available structural, operational, and environmental data. Originally developed for Berlin, Germany, the model has been adapted and applied to the sewer network of Lausanne, Switzerland. Model performance was evaluated using custom metrics, with results compared to previous applications in Berlin. Despite challenges related to class imbalance, the model demonstrated promising accuracy, supporting its potential as a decision-making tool for inspection prioritization.

Highlights

- Predictive machine-learning models help utilities prioritize sewer's inspection strategy.
- A model originally developed for the city of Berlin was adapted to Lausanne's sewer.
- Class-imbalance in the training phase can limit the performance of the model.

Introduction

Wastewater utilities often face challenges in assessing the condition of their sewer networks, with CCTV inspections being the most common approach (Roghani et al., 2019). While effective, this method requires careful planning to avoid inspecting pipes not in need of short-term intervention (Tscheikner-Gratl et al., 2020). To address this issue, predictive machine-learning models offer a promising solution by helping utilities prioritize inspections and focus on pipes with the highest likelihood of critical damage. This study focuses on the application of a machine learning model to the sewer network data of the city of Lausanne, Switzerland.

Originally developed and applied in Berlin in collaboration with the Berlin Water Utility (BWB), the model predicts the sewer condition of uninspected pipes, learning from inspection data provided by the utility. In this paper, we outline the model's methodology, focusing on training and performance evaluation and compare the obtained results in Lausanne with the model's performance obtained in Berlin.

The challenges encountered during its adaptation to Lausanne are highlighted.

Methodology

The developed application provides a framework to evaluate the condition of sewer pipes and identify those requiring urgent rehabilitation. It leverages a machine learning (ML) Random Forest model, to predict the condition of uninspected pipes based on observed aging patterns.

Random Forest is a robust classification algorithm composed of multiple decision trees. Each tree iteratively splits the data into smaller subsets using decision rules that maximize classification. By aggregating predictions through majority voting, the ensemble model reduces the instability of individual trees.

During training, the model learns patterns linking input features to pipe conditions, with class weighting helping to address class imbalance. Model performance is evaluated comparing predictions to observed conditions through metrics derived from the confusion matrix, such as recall value per class, false negative rates (ratio of bad conditions pipes misclassified as good), and false positive rates (ratio of pipes incorrectly flagged as at risk).

A custom global error metric $\boldsymbol{k}$ has been used as a first screening tool in the calibration step to find the best performing set of hyperparameters. $\boldsymbol{k}$ is computed from the three recall values of each class and considers the errors (false negatives and false positives) for the “bad” condition class (Caradot et al., 2018). These metrics reflect key practical objectives: minimizing the underestimation of poor conditions and avoiding unnecessary inspections.

Case Study

The city of Lausanne, through the public company *Service de l'Eau*, manages the city’s water networks serving a population of approximately 145,000 residents. The structural condition of the pipelines is categorized into five levels numbered from 0 (very poor condition) to 4 (very good condition). For the purposes of this study, the condition scored have been aggregated into three classes: *good*, *medium* and *bad*. Table 1 provides a list of the main variables used to train the algorithm and forecast pipes’ condition.

Table 1. Explicatory variables used in Lausanne for the condition simulation**.**

Name	Definition	Type	Unit
Length	Length of the pipe from one manhole to the next	Numerical	m
Width	Width of the pipe’s cross-section	Numerical	cm
Height	Height of the pipe’s cross-section	Numerical	cm
Lining	Information about whether the pipe has undergone lining	Categorical	yes/no
Construction year	Installation date of the pipe from which the age can be derived	Numerical	dd-mm-yyyy
Shape	Shape of the pipe’s cross-section (e.g. circle, egg-shaped…)	Categorical	-
Material	Pipe’s material	Categorical	-
Sewerage Type	Type of effluent flowing in the pipe (storm, sanitary, combined)	Categorical	-

The portion of the network for which a valid inspection record is available is slightly less than 60%. The corresponding CCTV material have been evaluated by three different agents: human operators right after the inspection, an image recognition software capable of detecting and classify damages from the videos and experts who have derived a condition score from previously recorded material. Overall, the observed conditions are highly unevenly distributed (see Fig. 1) with a minimal representation of the “bad” class.

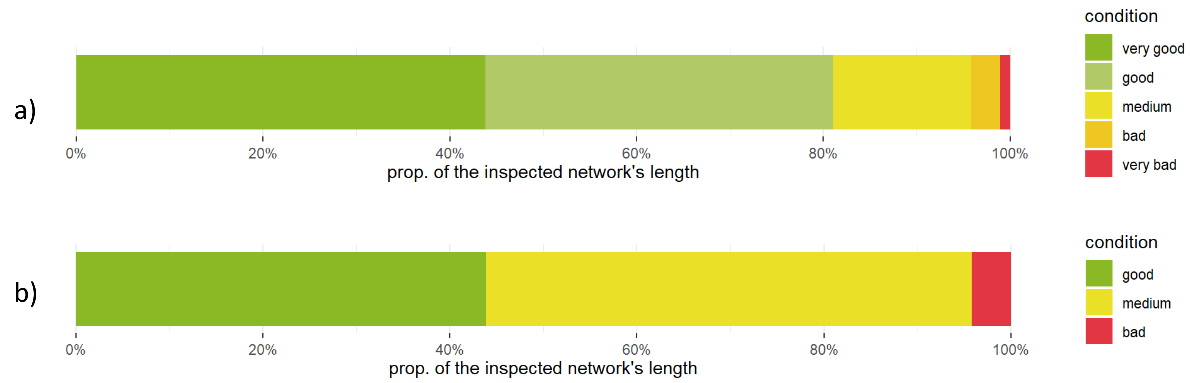

Figure 1. Distribution of the observed condition classes among the inspection data provided by the city of Lausanne (a) and class aggregation strategy adopted for the condition simulation (b).

The model autonomously tests available variables, selecting those that improve classification. Key factors such as pipe age, material, and cross-section dimensions are expected to be particularly relevant. Pipes are treated as segments between two access manholes, and only inspections covering the entire pipe length are included in training.

A dedicated module was developed to filter non-compliant data, correct formats, and standardize datasets. For example, incomplete inspection records or records missing inspection date have been removed, as well as pipes with non-plausible dimensioning. Missing construction age is one of the most common reasons for rejecting data points.

This process is highly customized and iterative. While ensuring data consistency is crucial, overly strict filtering can excessively reduce the training set, hindering the model's ability to generalize.

Results and discussion

The model's performance was evaluated using a cross-validation with 3 folds, with results summarized in a confusion matrix (Table 2). Recall and precision have been computed both globally and separately for the three condition classes and used for assessing model performance.

Additionally, a cumulative distribution plot was generated, showing the observed condition classes sorted by the model's predicted ranking. This visualization underscores the benefits of using predictive models, as pipes predicted to be in poor condition show a higher proportion of observed bad condition compared to a random selection approach (Fig. 2).

One of the key challenges encountered is the strong class imbalance within the dataset. The underrepresentation of the "bad" condition class strongly limits the model's predictive capacity for this category, despite the application of class weighting during training. In addition to this imbalance, the very low absolute number of pipes in bad condition may further impair the model's ability to learn reliable patterns for this class. This results in a model that is accurate in finding the pipes in bad condition (high recall value) but at the same time tends to have a high false negative rate (low precision).

It also emerged that including inspections assessed by different sources (e.g., human operators and artificial intelligence algorithms) in the model training does not improve predictive performance. On the contrary, it introduces a potential source of uncertainty that may affect the model's reliability.

To further assess the model's reliability, its performance in Lausanne was compared to results obtained in Berlin using the same evaluation metrics (Table 3). The comparison indicates promising results, with accuracy levels in Lausanne aligning with those observed in previous applications, reinforcing the model's adaptability across different urban sewer networks.

		predictions			
		good	medium	bad	sum
observations	g	564	284	123	971
	m	58	860	199	1117
	b	3	4	85	92
sum		625	1148	407	

Table 2. Confusion matrix of model's validation

metric	value Lausanne		value Berlin
Recall *good*	564/971	0.58	0.64
Precision *good*	564/625	0.90	0.85
Recall *medium*	860/1117	0.77	0.40
Precision *medium*	860/1148	0.75	0.39
Recall *bad*	85/92	0.92	0.67
Precision *bad*	85/407	0.21	0.42
False negative rate (*bad*)	7/92	0.08	0.33
False positive rate (*bad*)	322/407	0.79	0.58

Table 3. Evaluation metrics for Lausanne case study and earlier application in Berlin

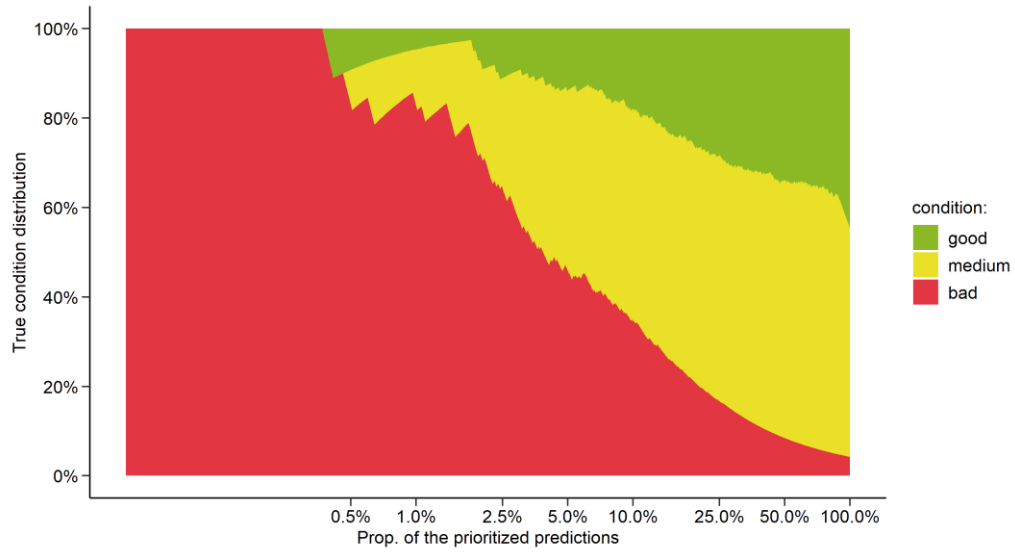

Figure 2. Cumulative distribution of observed conditions for inspected pipes, ranked by predicted condition (red: bad, yellow: medium, green: good). The graph illustrates how prioritizing inspections based on the model's ranking increases the detection of pipes in poor condition. In contrast, a random inspection strategy would yield a detection rate proportional to the overall condition distribution across all inspections (compare with Fig. 1).

To enhance model interpretability, feature importance scores from the Random Forest model were analysed. Variables such as pipe age at inspection, length and material emerged as the most influential predictors (Figure 3).

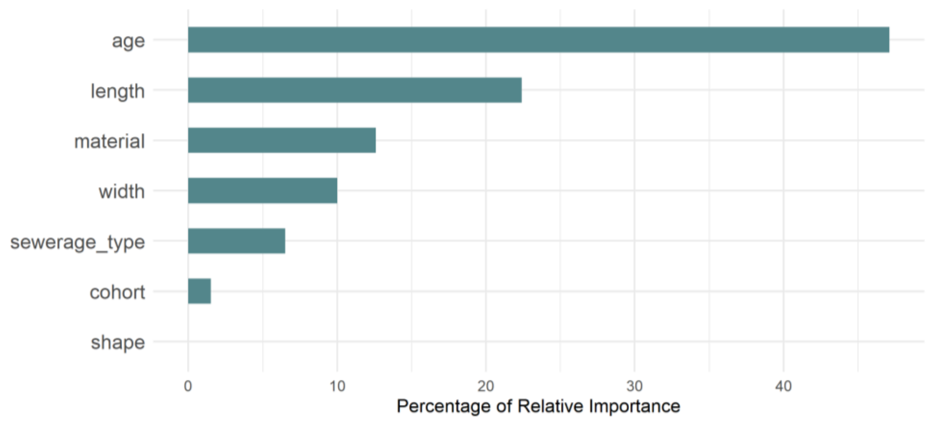

Figure 3. Importance of variables by mean decrease in Gini Index

Conclusions and future work

Overall, the study demonstrates the feasibility of applying a predictive ML-Model to forecast sewer condition in Lausanne. The findings highlight the benefits of using such models for inspection prioritization effectively reducing inspections of pipes that do not need urgent intervention. The model shows a good performance in finding pipes in bad condition but the false-negative rate remains high. We anticipate that increasing the number of inspections available for training would improve model performance. In cases of data scarcity, alternative strategies for enhanced class balancing should be explored.

Future research could consider binary modelling explicitly trained to detect a particular condition class or focus on the opportunity of training the models with additional foreign inspection data (i.e. from other cities than the one for which the simulation is carried on). Investigating the feasibility of a generalized model applicable across multiple cities could provide valuable insights into the transferability of predictive approaches in sewer condition assessment (Skjelde et al., 2024).

In addition, exploring alternative predictive models, including both traditional machine learning algorithms and hybrid approaches could offer important insights into the robustness and adaptability of such tools (Nguyen et al., 2023). Comparing their performance with that of random forests may shed light on model-specific strengths and limitations, particularly in terms of transferability across different urban contexts and data availability scenarios.

Acknowledgement

The present study is part of the SEMAplus project. We thank Service de l'Eau for their valuable support, for providing the data, and for collaborating with us to ensure their usability.

References

Caradot et al. (2018) Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany. *Journal of Hydroinformatics*; 20 (5): 1131–1147. doi: https://doi.org/10.2166/hydro.2018.217

Roghani et al. (2019) Dealing with uncertainty in sewer condition assessment: Impact on inspection programs, Automation in Construction, Pages 117-126, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2019.03.012.

Tscheikner-Gratl et al. (2019) Sewer asset management – state of the art and research needs. Urban Water Journal, VOL. 16, NO. 9, 662–675, doi: https://doi.org/10.1080/1573062X.2020.1713382

Skjelde et al. (2024) Transferability and comparability of sewer deterioration models – a case study on Norwegian sewer data. Urban Water Journal, 21:10, 1242-1253, doi: https://doi.org/10.1080/1573062X.2024.2423391

Nguyen, L. V, Razak, S., 2023. Predicting sewer structural condition using hybrid machine learning algorithms. Urban Water J. 20, 882–896. https://doi.org/10.1080/1573062X.2023.2217430