Combining Predictions of Municipal Asset Conditions at the Segment Level to Determine Street Closures

Kelechukwu Tersoo Genger[a], Amin Hammad*[b]
Concordia Institute for Information Systems Engineering, Concordia University, Canada.
[a] t_genger@encs.concordia.ca, [b]amin.hammad@concordia.ca

Abstract. Intervention activities that involve cut and cover of the subsurface tend to have adverse effects on society and the economy. These negative impacts may include more traffic congestion, longer travel times, and increased noise and air pollution. To reduce the frequency of street closures for maintenance, repair, and rehabilitation work, a more comprehensive approach to intervention planning is necessary. This approach should be based on a unified classification model that assesses the conditions of various municipal assets (such as water and sewer pipes and pavements) at the segment level. By taking this holistic approach, we can minimize the negative impacts of excavation activities and reduce the need for future repairs and closures. This paper aims to develop three ensemble machine learning methods for classifying the conditions of underground municipal assets (i.e., pavement, water, and sewer pipes) within a segment and to determine street closures based on the interventions required. Based on these conditions and heuristics specific to each asset, the method determines segment-level interventions and the nature of the required street closures. The models have high accuracies ranging from 82.38% to 98.64%, and the segment-level intervention strategies have an accuracy of 79.92%. This research can help municipal decision-makers prioritize interventions, improve planning, and estimate the duration of street closures.

1 Introduction

A synchronized utility intervention (integrated asset management) is when the repair or renewal of spatially collocated assets is done simultaneously to reduce the socioeconomic impact as well as the cost of the intervention. This method of utility intervention introduces several challenges and therefore complicates the decision-making process. A major factor that determines the effectiveness of this mode of utility intervention is the level of coordination and data sharing among utility owners. In most cases, the level of coordination between municipal and private utility owners is inadequate (Abu-Samra *et al.*, 2018), even for spatially collocated utility assets. Although synchronized interventions are not practiced in several countries, they are encouraged (FCM and NRC, 2003). Some examples of municipalities in Canada currently practicing synchronized interventions include Sudbury and Hamilton in Ontario, Montreal in Quebec, Kelowna, Surrey in British Columbia, etc. Other examples are Bergen and Trondheim, in Norway and Trelleborg, Sweden (Braun, 2012; Chacon & Normand, 2016; FCM & NRC, 2003; Hafskjold, 2010; Hafskjold & Bertelsen, 2008).

In the short term, a certain level of sustainability can be achieved by conducting synchronized integrated interventions. However, in the long term, under certain conditions (e.g., high utility density, traffic density, etc.), a major shift can be made toward the sustainable placement of underground utilities using multi-purpose utility tunnels (MUTs). MUTs offer a long-term alternative by hosting utilities in an underground tunnel capable of housing several utilities in single or multiple compartments. By so doing, utilities are less vulnerable to damage, thus, increasing the lifespan of the utilities hosted in the tunnel. Also, expansion, inspection, and maintenance of underground utilities can be executed all year round with the possibility of eliminating social costs (Genger *et al.*, 2021).

Two common methods used to determine the need for intervention are failure prediction and condition classification. Weeraddana *et al.* (2019) used a supervised machine learning algorithm called random forest regression (RFR) to predict the likelihood of water mains failure, while Jafar et al. (2010) employed six artificial neural network (ANN) models to predict the failure of urban water mains. In terms of condition classification, many researchers have used binary classifications, such as good or bad, to classify asset conditions. However, using binary models may result in misclassification errors, which can increase economic losses due to prematurely replaced assets that were still in acceptable condition. Some researchers, such as Hernández *et al*. (2021), have classified sewer conditions into three classes (good, medium, and bad) using random forest and support vector machine (SVM)-based models at the asset and network levels for management and inspection purposes. Various combinations of features have been used to achieve pipe failure prediction or condition classification, and the accuracy of the machine learning models is dependent on the available features, as well as the ML algorithm selection, data preprocessing, hyperparameter tuning, and data quality.

Some researchers focused on predicting only one indicator and then used its value to determine pavement condition (Abdelaziz et al., 2020; Bashar & Torres-Machi, 2021; Kirbaş & Karaşahin, 2016; Zhou et al., 2021). Even though relationships exist between several performance indicators, using one indicator alone may not sufficiently capture the condition of a street segment, considering that different standards exist for acceptable indicator thresholds (Arhin et al., 2015).

Previous research on the prediction or classification of the conditions of municipal assets has primarily focused on evaluating each asset individually. However, it is crucial to view these assets as interconnected systems and coordinate their interventions to minimize user costs (Genger & Hammad, 2022). To achieve this, it is essential to evaluate the conditions of multiple assets within a segment instead of solely assessing each asset independently.

In previous studies, binary or multi-class classification models were used to assess individual infrastructure assets or networks in isolation. However, this research takes a different approach by conducting a multi-class classification of multiple assets, which are evaluated together at the segment level. Additionally, this research goes beyond determining segments for synchronized or unsynchronized interventions by proposing an alternative method, the MUT. Street segments requiring excavation-related interventions can be considered an opportunity to implement the MUT. Using this approach establishes a method of MUT location selection that is based on the condition of the asset conditions and their need for interventions.

The primary goal of this research is to forecast street closures by analyzing the combined conditions of municipal infrastructure assets located in the same segment. To achieve this objective, the research has two specific aims: (1) to create a machine learning (ML) technique for systematically classifying the condition of various underground municipal assets (including pavements, water and sewer pipes) that are spatially collocated within a segment; and (2) to employ a heuristic approach for determining street closures based on the synchronized or unsynchronized interventions at the segment level, which result from merging the interventions of individual assets within each segment.

2 Proposed Method

Figure 1 illustrates a methodology proposed to determine street closures at the segment level by integrating the conditions obtained from three separate ML models for each asset and then applying intervention strategies. The methodology begins with preparing the raw GIS data for each asset's failure, inspection, and network datasets, followed by executing various data preprocessing steps on the datasets. Next, the data is split into training/cross-validation and testing sets. The chosen features for each asset are then fed into their respective ML model, each of which comprises a voting-based ensemble ML algorithm. To obtain the highest accuracy, hyperparameter tuning is carried out on each ML model. Each model's output, which is the condition of the three assets, is assessed and combined at a segment level. The assets-to-segments mapping is established using each asset's unique identifier and the unique identifiers of the street segments where the assets are situated. Finally, the combined asset conditions are integrated with intervention strategies to determine the type of street closures required (partial or complete) and the necessity for synchronized or unsynchronized interventions.

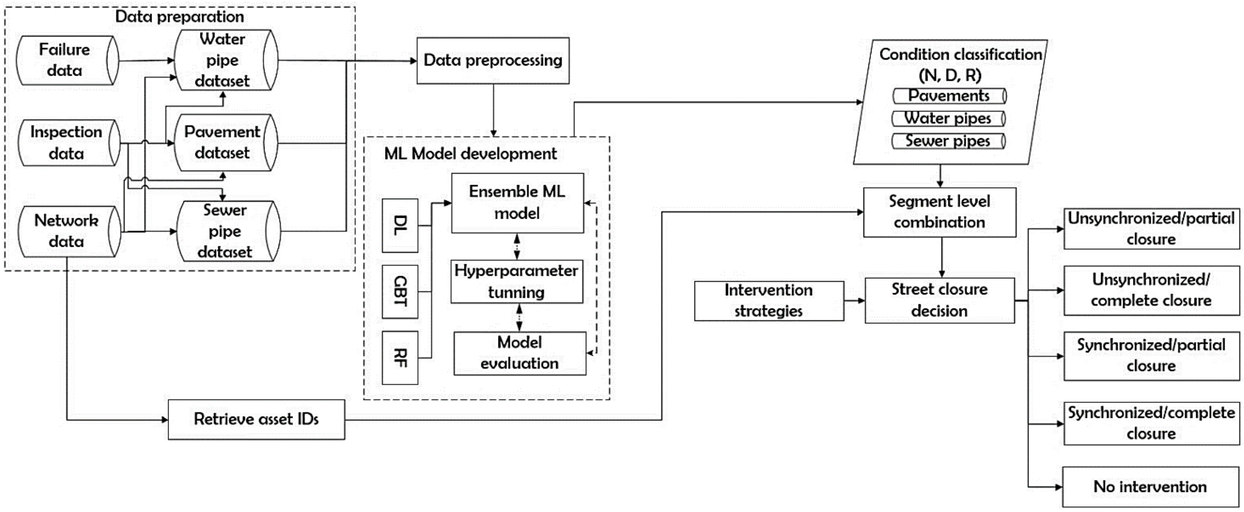

Figure 1: Proposed method.

2.1 Data Preprocessing

In the initial stage of the methodology, the datasets required for the machine learning process are created. The attributes to be included in the datasets are determined based on previous research, the quality of the data, and its availability. To this end, a significant amount of data related to the various characteristics of each asset is collected and utilized in the classification process. The features for each utility asset, and the data values obtained after the preparation phase, are presented in Table 1 to Table 3. The intersect tool of the ArcGIS geoprocessing toolbox was utilized to combine data from various sources using a unique identifier. The combined data underwent preprocessing to detect and eliminate duplicate, missing, erroneous, and outlier values from the training, validation, and testing datasets. To further improve the model's accuracy and reduce complexity, data preprocessing methods such as feature selection and dimensionality reduction were employed.

Feature selection eliminates redundant and noncrucial features while generating optimal features to achieve maximum accuracy. In addition, data normalization using the Z-transformation method was performed on numeric features. To avoid bias, the data was shuffled to ensure that each subset used in training/cross-validation and testing was representative of the overall data distribution. The Synthetic Minority Over-sampling TEchnique (SMOTE upsampling) was used to handle the class imbalance in the training dataset. This technique creates new instances of the minority class by finding a random neighbor for a subset of randomly selected instances using K nearest neighbors. Although the newly generated instances balance the missing instances, they do not provide any new information to the model.

2.2 ML Model Development

Following data preparation and preprocessing, the datasets were divided into two parts: training/cross-validation and testing datasets. The former constituted 70% of the data, while the latter comprised 30%. To enhance the learning model's performance, the cross-validation process was carried out on ten sets of mutually exclusive subsets with equal folds. Each fold was prepared using either shuffled or stratified sampling. To achieve the highest accuracy, a voting-based ensemble algorithm consisting of Random Forest (RF), Gradient Boosted Trees (GBT), and Deep Learning (DL) algorithms were used for each fold's training. This combination was necessary because the output of each algorithm varied significantly. Finally, hyperparameter optimization was conducted using the grid search technique for each base learner in the ensemble.

2.3 Condition Classification and Segment Level Classification

Multi-class classification is carried out to determine the conditions of each of the three assets from their independent models. The condition of each asset is labeled using a uniform scale of *N* (No intervention required, assets are unlikely to fail in the near future), *D* (intervention is desirable as assets have an estimated time of failure between 10 to 20 years), and *R* (immediate attention required i.e., assets have failed or assets likely to fail between 0 to 10 years). However, the focus lies on identifying street segments with conditions with a combination of classes *D* and *R* for all the assets. These conditions determine the required asset-level intervention strategies needed to restore the asset to an *N* condition. Each segment is assigned a unique identifier, and this ID is assigned to each asset located in the segment in addition to the unique identifier of the asset.

2.4 Intervention Strategies

Once the conditions of the assets have been classified, the next step is to determine the intervention strategies at both the asset and segment levels. Most street segments have multiple water and sewer pipes with different characteristics, such as diameter, length, and depth. In cases where the water and sewer pipes share the same right-of-way, the water pipes are usually buried above the sewer pipes to reduce contamination risks. Therefore, interventions on sewer pipes often require excavations below the water pipes, which creates an opportunity for synchronized interventions on water pipes. While this may increase intervention costs, it avoids future excavations for inevitable water pipe replacements.

The heuristics used to formulate the intervention strategies for each asset depend on various criteria or factors, such as the asset's current condition, performance indicators, and past intervention practices. Synchronized interventions also consider the combined strategies of the collocated assets at a segment level. Based on Montreal City's intervention guidelines, *R*-class pavements with a PCI ≤ 40 and IRI ≥ 6 will undergo reconstruction if the pavement is rigid or major rehabilitation if it is flexible. Major rehabilitation is required if the pavement condition is bad, and minor rehabilitation is necessary for *D*-class pavements. No intervention is needed for *N*-class pavements. However, these interventions depend on the intervention strategies of other collocated assets.

2.5 Street Closure Decisions

The output of the ML process (the asset condition classes) is combined with the criteria values extracted from the inspection data of each asset to determine the asset-level intervention strategies.

This study uses condition-based heuristics (Genger and Hammad, 2023) that are generally formulated by the asset owners to guide the decision-making process for the interventions of each asset. An example includes the use of a combination of the pavement condition index (PCI) and international roughness index (IRI) thresholds together with the pavement type as the criteria values used to construct the heuristics for determining either a *reconstruction*, *major* or *minor rehabilitation*, or *no intervention* on a street segment. Meanwhile, using the unique identifiers of the segments and the collocated assets, the asset-level interventions are then combined at the segment level to determine the nature of street closures.

3 Case Study

All ML processes were performed on a computer with the following specifications: Ubuntu 20.04.3LTS, AMD Ryzen Threadripper 3960x 24-core processor x 48 threads, and 251 GB memory. RapidMiner Studio Educational 9.10.001 is used to build all ML processes. The GBT and DL algorithms were executed using the H2O 3.30.0.1 (RapidMiner, 2021). The dataset covers all the boroughs in Montreal, but a portion of the data was held back from the Ville-Marie Borough to implement intervention strategies and street closures. To ensure that all three models could later be analyzed at a segment level, this reserved dataset was necessary, as many ML data processing and splitting techniques involve random sampling.

Data used in this research was obtained from three primary sources, including intervention and network data for all three assets and the failure data for only the water pipes. The summary of the datasets is presented in Table 1 to Table 3. Hyperparameter tuning was done on the parameters of all three models using a grid search. The models are trained using the combination of every single value (i.e., exhaustive search) in the search range, to find an optimal parameter set that is guided toward improving the accuracy of each model. All other parameter values for each algorithm retain their default values.

Table 1: Summary of pavement features

Feature names	**Values**
Number of road segments	15,603
Performance condition index (PCI) Average PCI	1-100 52
International roughness index (IRI) Average IRI	0 – 13.9 5
Rutting (m) Average rutting	0-50 4.79
Pavement coating	Asphalt, cobblestone, concrete, crushed stone
Average length (m)	120
Average surface area (m^2)	2,245.1
Category	Arterial, local
Pavement type	Rigid, flexible
Pavement condition	**N, D, R**

Table 2: Summary of water pipe features

Feature names	**Values**
Number of pipes	127,716

Mean age (yrs.), Std. Dev.	73.84 40.79
Major materials	Gray cast iron, ductile iron, copper, reinforced concrete, etc.
Jurisdiction	Local, metropolitan area, centre ville, etc.
Diameter range (mm)	15-3,900
Average pipe length (m)	215.4
Break rate (brk/km/yr) Average break rate	0-15 0.5
Break age	1–149
N_failures (Number of failures in the section) Average	0–15 0
R_life (age/estimated useful life) Average	0–1.68 0.55
T_length (Average total length of pipes in a section) (m)	286.3
N_p_bad (Number of pipes in a street section with a bad or very bad status) Average	0–8 0.21
N_p_segment (Number of pipes in a street section) Average	1-27 1.48
Pipe condition	**N, D, R**

Table 3: Summary of sewer pipe features

Feature names	**Values**
Number of pipes	119,857
Major materials	Reinforced concrete, grey font, brick, PVC, ductile iron, etc.
Pipe type	Combined, sanitary
Hierarchy	I, II, III
Installation year	1900-2015
Diameter range (mm)	75-5,325
Average length (m)	54.19
T_length (Average total length of pipes in a section) (m)	254.19
T_n_pipes (Average number of pipes in a section) (m)	5
R_life (Inspection age/estimated useful life) Average	0–3.48 0.45
Rem_life (Remaining life) Average	1-212 71
N_p_bad (Number of pipes in a street section with a bad or very bad status) Average	0-12 1.14
Inspection year	1993-2015
Jurisdiction	Local, arterial
Sewer condition	**N, D, R**

3.1 Pavement, Water and Sewer Pipe Classification Model Performances

Table 4 presents the results of the pavement, water and sewer pipes ensemble classification models. The table shows that using the ensemble algorithm on the pavement training and test datasets generated an accuracy of 98.66% and Kappa = 0.98. Applying the model to the test dataset generated an accuracy of 98.64%, Kappa = 0.98. The model's accuracy on the water pipe training and testing datasets is 97.27% and 96.37%, respectively. The test dataset's Kappa value (0.91) shows that only a small number of the expected classification is achieved by chance. The sewer model's training/cross-validation and test datasets accuracies are 86.22 and 82.38%, respectively. The corresponding Kappa values for both datasets are 0.79 and 0.67.

Table 4: Pavement, water and sewer pipes ensemble model performances

Asset type	**Data**	**Accuracy (%)**	**Kappa**
Pavement	Training dataset	98.66	0.98
	Test dataset	98.64	0.98
Water pipes	Training dataset	97.27	0.93
	Test dataset	96.37	0.91
Sewer pipes	Training dataset	86.22	0.79
	Test dataset	82.38	0.67

3.2 Intervention Strategies

Finally, the ML model was applied to the dataset used to test segment-level interventions and street closures. The accuracies of all ML models on this dataset are presented in Table 5. Based on the segment-level intervention strategies, Figure 2 is the predicted intervention map for street segments. The maps show street segments where no intervention, unsynchronized, or synchronized interventions are needed and the nature of the street closures. Street segments with no information on the collocated assets are labeled as no data. When comparing the actual and predicted sets of segment-level interventions, there was a 79.92% similarity.

Table 5: Accuracy of the models on the street closure dataset

Ville Marie subset data	**Accuracy (%)**
Pavement	96.79
Water pipe	94.70
Sewer pipes	70.17

4 Discussion

The results also show that some street segments undergo partial and complete street closures to accommodate the pipe rehabilitation phase and the subsequent pavement reconstruction or rehabilitation phase of the intervention. This type of closure reduces the accrued social cost on the street segment because road users can access the street during partial closures. The results also reveal street segments where the implementation of the MUT could serve as an alternative to the synchronized method. Although several criteria determine the placement of MUTs on a street segment (Genger *et al.*, 2021), this research can aid in identifying potential street segments which can be subsequently ranked using the criteria for determining the MUT placement.

Although synchronized interventions increase utility cost savings by reducing the number of repeated excavations, introducing the MUT as an alternative technique for street sections, where the combined condition of the assets is in a critical state, increases the lifespan and ease of maintenance of underground utilities. This method of utility placement increases sustainability while avoiding future excavations related to utility interventions in the implemented street segment.

GIS maps were used to display the street closures where interventions and subsequent street closures are imminent. These visualizations aid in traffic management (alternative route selection based on the impact on travel time) and intervention budget estimation (direct and social costs)

based on the conditions of the individual assets. By classifying the individual conditions of the pipes in a road segment, a more accurate intervention duration can be ascertained.
Table 6 shows the comparison of the results of each model to individual classification models in the literature review. The pavement and water models outperformed all the previous classification models, and the sewer condition model outperformed all the models except (Tavakoli *et al.*, 2020).

Table 6: Model accuracy comparison

Asset	Reference	Accuracy%	Accuracy % (This research)
Water pipes	Winkler *et al.* (2018)	96	96.37
	Robles-Velasco *et al.* (2020)	85	
	Kumar *et al.* (2018)	62	
Sewer pipes	Harvey and McBean (2014)	76	82.38
	Mohammadi *et al.* (2019)	81	
	Laakso *et al.* (2018)	62	
	Tavakoli *et al.* (2020)	93	
Pavement	Piryonesi & El-Diraby (2021)	88	98.64
	Hoang and Nguyen (2018)	87.5	

5 Conclusions and Future work

This research presents an approach for determining street closures based on the combined conditions of spatially collocated municipal infrastructure assets at the segment level. The use of ensemble ML methods for classifying multi-asset conditions while using a uniform scale for all assets made it applicable for enhancing synchronized interventions at the segment level.

Figure 2: Predicted segment intervention strategies and street closures

Notes: *N.I.: No intervention; S: Sewer pipe, P: Pavement; W: Water pipes; Sync: Synchronized; Unsync: Unsynchronized; RH: Rehabilitation; RC: Reconstruction, RP: Replacement; PC: Partial closure; CC: Complete closure; MUT: Multi-purpose utility tunnel*

The contributions of this research are as follows: (1) Developing an ML-based method for systematic condition classification of different spatially collocated underground municipal assets within a segment; and (2) Applying a heuristic approach for determining street closures based on the synchronized or unsynchronized interventions at the segment level induced by combining the interventions of individual assets within each segment.

This research was conducted on a dataset of 200 street segments in the City of Montreal and shows potential for scaling up to an entire city with similar asset features and intervention strategies. It could also include private infrastructure assets in addition to municipal assets. Limitations include the need for a balanced sewer pipe dataset with additional features such as slope and soil and the exclusion of other buried assets such as gas pipes and electrical cables. Coordination and data availability could be an issue when assets are not managed together. Tradeoffs such as budgetary constraints and organizational barriers should be addressed in synchronized interventions with multiple stakeholders. Future research should consider uncertainties and include risk assessment and probability of failure and ML methods.

References

Abdelaziz, N., Abd El-Hakim, R.T., El-Badawy, S.M. and Afify, H.A. (2020), "International Roughness Index prediction model for flexible pavements", *International Journal of Pavement Engineering*, Vol. 21 No. 1, pp. 88–99.

Abu-Samra, S., Ahmed, M., Hammad, A. and Zayed, T. (2018), "Multiobjective Framework for Managing Municipal Integrated Infrastructure", *Journal of Construction Engineering and Management*, Vol. 144 No. 1, available at:https://doi.org/10.1061/(ASCE)CO.1943-7862.0001402.

Arhin, S.A., Williams, L.N., Ribbiso, A. and Anderson, M.F. (2015), "Predicting Pavement Condition Index Using International Roughness Index in a Dense Urban Area", *Journal of Civil Engineering Research*, Vol. 10 No. 17, available at:https://doi.org/10.5923/j.jce.20150501.02.

Bashar, M.Z. and Torres-Machi, C. (2021), "Performance of Machine Learning Algorithms in Predicting the Pavement International Roughness Index", *Transportation Research Record: Journal of the Transportation Research Board*, SAGE PublicationsSage CA: Los Angeles, CA, Vol. 2675 No. 5, pp. 226–237.

Braun, A.J. (2012), "Planning an Integrated Utility /Roadway Project Brings Asset Management to Life", *Proceedings of the Water Environment Federation*, Water Environment Federation, Vol. 2003 No. 3, pp. 216–220.

Chacon, N. and Normand, H. (2016), *Plan D'Intervention Ville de Montreal*, Montreal.

FCM and NRC. (2003), "Coordinating infrastructure works: a best practice by the national guide to sustainable municipal infrastructure", *National Research Council of Canada. NRC Institute for Research in Construction*, No. 1, available at: https://fcm.ca/sites/default/files/documents/resources/guide/infraguide-coordinating-infrastructure-works-mamp.pdf (accessed 6 October 2020).

Genger, K.T. and Hammad, A. (2023), "Street closure prediction based on the combined conditions of spatially collocated municipal infrastructure assets at the segment level", *Expert Systems with Applications*, Vol. 219, p. 119671.

Genger, T.K. and Hammad, A. (2022), "Geospatial Visual Analytics for Supporting Decision Making for Underground Utility Integrated Interventions", *International Conference on Transportation and Development 2022: Application of Emerging Technologies - Selected Papers from the Proceedings of the International Conference on Transportation and Development 2022*, Vol. 6, American Society of Civil Engineers, Reston, VA, pp. 46–59.

Genger, T.K., Luo, Y. and Hammad, A. (2021), "Multi-criteria spatial analysis for location selection of multi-purpose utility tunnels", *Tunnelling and Underground Space Technology*, Pergamon, Vol. 115, p. 104073.

Hafskjold, L.S. (2010), *SIP-Future Rehabilitation Strategies for Physical Infrastructure: coordination of rehabilitation planning and measures-co-infrastructure interactions Water and Environment*, available at: https://www.sintef.no/globalassets/upload/byggforsk/infrastruktur/sip-rehab/3c027716-lsh-sip_coordination_rehab.pdf (accessed 30 July 2022).
Hafskjold, L.S. and Bertelsen, D. (2008), *SIP - Future Rehabilitation Strategies for Physical Infrastructure : Studieområde Nonnegata: Hva Var Planlagt Og Hva Ble Resultatet? Er Det Noe å Lære?*, SINTEF construction research, available at: https://www.nb.no/items/dde26cfbd022ebaaaf552c59df347a0c?page=0&searchText=oaiid:%22oai:nb.bibsys.no:990901344864702202%22 (accessed 30 July 2022).
Hernández, N., Caradot, N., Sonnenberg, H., Rouault, P. and Torres, A. (2021), "Optimizing SVM models as predicting tools for sewer pipes conditions in the two main cities in Colombia for different sewer asset management purposes", *Structure and Infrastructure Engineering*, Taylor and Francis Ltd., Vol. 17 No. 2, pp. 156–169.
Hoang, N.-D. and Nguyen, Q.-L. (2018), "A novel method for asphalt pavement crack classification based on image processing and machine learning", *Engineering with Computers 2018 35:2*, Springer, Vol. 35 No. 2, pp. 487–498.
Jafar, R., Shahrour, I. and Juran, I. (2010), "Application of Artificial Neural Networks (ANN) to model the failure of urban water mains", *Mathematical and Computer Modelling*, Vol. 51 No. 9–10, pp. 1170–1180.
Kumar, A., Ali Asad Rizvi, S., Brooks, B., Ali Vanderveld ShopRunner, R., Wilson, K.H., Kenney, C., Edelstein, S., *et al.* (2018), "Using Machine Learning to Assess the Risk of and Prevent Water Main Breaks", *ACM Reference Format: Breaks. In Proceedings of ACM SIGKDD*, Vol. 9, available at:https://doi.org/10.48550/arxiv.1805.03597.
Laakso, T., Kokkonen, T., Mellin, I. and Vahala, R. (2018), "Sewer Condition Prediction and Analysis of Explanatory Factors", *Water 2018, Vol. 10, Page 1239*, Multidisciplinary Digital Publishing Institute, Vol. 10 No. 9, p. 1239.
Mohammadi, M.M., Najafi, M., Tabesh, A., Riley, J. and Gruber, J. (2019), "Condition prediction of sanitary sewer pipes", *Pipelines 2019: Condition Assessment, Construction, and Rehabilitation - Proceedings of Sessions of the Pipelines 2019 Conference*, American Society of Civil Engineers, pp. 117–126.
Piryonesi, S.M. and El-Diraby, T.E. (2021), "Examining the relationship between two road performance indicators: Pavement condition index and international roughness index", *Transportation Geotechnics*, Elsevier, Vol. 26, p. 100441.
RapidMiner. (2021), "Gradient Boosted Trees", *RapidMiner*, available at: https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/trees/gradient_boosted_trees.html (accessed 17 February 2022).
Robles-Velasco, A., Cortés, P., Muñuzuri, J. and Onieva, L. (2020), "Prediction of pipe failures in water supply networks using logistic regression and support vector classification", *Reliability Engineering and System Safety*, Elsevier Ltd, Vol. 196, p. 106754.
Tavakoli, R., Sharifara, A. and Najafi, M. (2020), "Prediction of Pipe Failures in Wastewater Networks Using Random Forest Classification", *Pipelines 2020: Condition Assessment, Construction, Rehabilitation, and Trenchless Technologies - Proceedings of Sessions of the Pipelines 2020 Conference*, American Society of Civil Engineers (ASCE), pp. 90–102.
Weeraddana, D., Liang, B., Li, Z., Wang, Y., Chen, F., Bonazzi, L., Phillips, D., *et al.* (2019), "Machine learning for water mains maintenance", *Water E-Journal*, Vol. 4 No. 3, pp. 1–13.
Winkler, D., Haltmeier, M., Kleidorfer, M., Rauch, W. and Tscheikner-Gratl, F. (2018), "Pipe failure modelling for water distribution networks using boosted decision trees", *Structure and Infrastructure Engineering*, Vol. 14 No. 10, pp. 1402–1411.
Zhou, Q., Okte, E. and Al-Qadi, I.L. (2021), "Predicting Pavement Roughness Using Deep Learning Algorithms", *Transportation Research Record: Journal of the Transportation Research Board*, SAGE Publications, p. 036119812110237.