Sewer Deterioration Modeling

The Effect of Training a Random Forest Model on Logically Selected Data-groups

Hansen, Bolette Dybkjær; Rasmussen, Søren Højmark; Moeslund, Thomas B.; Uggerby, Mads; Jensen, David Getreuer

Published in:
Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020

DOI (link to publication from Publisher):
10.1016/j.procs.2020.08.031

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 176 (2020) 291–299

Procedia Computer Science

www.elsevier.com/locate/procedia

24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Sewer Deterioration Modeling: The Effect of Training a Random Forest Model on Logically Selected Data-groups

Bolette D. Hansen[a,b]*, Søren H. Rasmussen[b], Thomas B. Moeslund[a], Mads Uggerby[b] and David G. Jensen[b]

[a]*Aalborg University, Rendsburggade 14, Aalborg 9000, Denmark*
[b]*EnviDan A/S, Vejlsøvej 23, Silkeborg 8600, Denmark*

Abstract

Breakdown of sewers can induce significantly damage to roads and buildings placed upon it. For this reason, timely maintenance of the sewer system is essential. However, due to the under-ground position of the sewers they are very expensive to monitor, as this is done by CCTV inspection. Therefore, it is important to choose the right sewers for inspection and several decision-support tools have been developed to help the operators to select which sewers to inspect. These decision support tools all contain a model which predicts the condition of the sewers, and recently several models have been proposed in order to increase the performance. The scope of this paper is to investigate the effect of training a Random Forest model on logically selected groups of data, as opposed to training of a joined model on the full data set. The selected data groups were based on expert knowledge: The first data groups were based on the sewer material (concrete, plastic, clay, reinforced with lining and other material). The concrete data set was then further sub-divided into wastewater types (sewage, rain and combined) whereas the plastic data set was sub-divided into road classes. The results showed that the model trained on the full data set performed better than the models trained on logically selected data-groups as it encounters the heterogeneity of the data set. Furthermore, this answers an important question raised by end users of the deterioration models.

Keywords: Sewer deterioation; ageing; random forest

**E-mail address:* bdha@create.aau.dk

10.1016/j.procs.2020.08.031

1. Introduction

The sewer system is an essential infrastructure, but it is very expensive to build and maintain. However, maintenance of the system is important in order to account for the ageing sewer system and maintain functionality. If a sewer is not maintained in time it can break down and induce damage to roads and buildings placed upon it. Replacing the sewers after break-down have shown more expensive than to perform intime rehabilitation or replacement of the sewer [1]. However, due to the under-ground position of the sewers they are hard to monitor

Monitoring of the sewer system is often performed by Closed Circuit Television inspection (CCTV inspection). CCTV inspection is performed by sending a robot with camera through the sewer system and manually annotating all the events present in the sewers. This is very time consuming and thereby expensive for which reason effort is put into automate the inspection technics [2]. However, there is still a long way before these techniques can fully replace manual inspection. Therefore, the utilities need to prioritize, which sewers to inspect. In order to help the utilities identifying, which sewers to prioritize several decision-support tools have been developed [3]–[6]. These tools typically consist of a deterioration model, predicting the condition state of the sewer or the probability of the sewer to be in a certain condition and a consequence model for determining the criticality in case of sewer failure. By combining them the users have a risk based prioritizing tool. One decision support tools also includes an economic model for prioritizing the sewers [5].

Where the consequence models are based on easily accessible data such as distance to buildings, industry areas, hospitals, schools, beaches and roads, the deterioration of the sewers is harder to determine. It depends on several factors including both environmental, operational, hydraulic and physicality of the sewers. In order to optimize the prediction of the condition several deterioration models have been developed or evaluated since 2018 [3], [5]–[16]. The approaches for modeling the deterioration of the sewers have traditionally been split in to deterministic, statistically and artificial intelligence or machine learning approaches [17]. The models developed today are typically either based on statistical or machine learning methodologies. Since 2018 the following statistical approaches have been utilized: Bayesian Network [6], Gompertz models [8], [9], Logistic Regression [10], [14], and Monte Carlo Makov Chain [7], [12]. Likewise, since 2018, the following machine learning approaches have been utilized: Random Forest [8], [10], [14], [16], other forest based methods [13], Ant Colony Optimization [3], and Support Vector Machine [11].

Some of the methods have been tested in a way that makes them comparable with others by e.g. the sensitivity and specificity of the results [14]–[16], [18] or the number of true positive, true negative, false positives and false negatives [8], [10]. However, it is still hard to compare the different models as one might prioritize a high sensitivity over the specificity and opposite. Comparison of the models is further complicated by the privacy of the data sets. This means that the models are developed based on data sets with different ratio between good and bad pipes. An example of how the different data sets can give different results is presented by [10] who tested different models on both a German data set and at Columbian data set. Likewise, the deterioration of the pipes varies as for instance concrete pipes are very sensitive to corrosion due to sewerage [19]. In many cases one model is trained or calibrated to all the sewer pipes at the same time [6]–[8], [10], [11], [14], [16]. However [15] made Bayesian Logistic Regression models for concrete, clay, metallic and plastic pipes respectively, while others focused on specific material groups such as plastic [12] or the material groups with the most pipes represented [13]. This leaves the question of whether it is beneficial to develop a deterioration model based on specific data groups such as material, wastewater type and road type, or if it is sufficient with one model for all data groups. The answer to this question leads to another question, which is, whether or not a model should be evaluated according to the performance of separate data groups or the overall performance of the model. In the cross-field between data science and material science, these questions are sometimes raised by the end users of the deterioration models, such as decision makers at utilities. Therefore, investigation of this will help aligning the expectations between the end users of the deterioration models and the developers.

The scope of this paper is to investigate the effect of developing several Random Forest models based on logical grouped data sets compared to developing one general model for the whole data set. The use of machine learning methods in this field is still in its infancy and Random Forest is at present state of the art.

2. Method

2.1. Data and preprocessing

The data has been withdrawn from a database where 35 Danish utilities have entered the results of CCTV inspections performed on their sewer networks. Only pipes which allowed for extraction of all predicter and target variables were withdrawn from the database. Likewise, in order to avoid pipes with incorrect registrations only pipes with realistic registrations were withdrawn. Examples of suspicious data could be if a pipe did not fit into following spans: 0 years ≤ pipe age ≤ 169 years, 63 mm < pipe dimension < 3000 mm and 0.6 m < pipe depth < 10 m. furthermore, a datapoint would be considered suspicious if several pipes had the exact same damage percentage. The final data set consisted of 119.919 pipes.

The annotation of the occurrences in the CCTV inspections follow the Danish standard for CCTV inspection of sewers [20]. Hereafter, the damage percentage of the sewers was calculated. The damage percentage is a weighted evaluation of each occurrences' contribution to the total damage percentage [21]. It is worth noticing that, due to the way of calculating the damage percentage, it is possible to obtain a damage percentage above 100. Finally, the damage percentage is transformed to a continuous scale ranging from 0-10. The formula for this is inspired by the Danish standard for calculating the physical index [21] and can be seen in equation 1.

$$TS = 5 \cdot log(DP + 1) \tag{1}$$

In equation 1 TS refers to the transferred scale and DP refers to the damage percentage. The predictor variables available is the same as those presented by [16] which in general terms relate to: Dimension and length of the pipes, material type, wastewater type, if and how the sewer have been rehabilitated, surrounding areas (industry, city etc.), surrounding buildings and trees, soil types, road classes and position of the pipes. In addition to these parameters the slope and the ground water level have been found. As the slope is not accessible for all pipes, this data set contains a little less data than the data set used for [16]. The data set consists of both binary and continuous variables.

2.2. Logical data groups

The logical data groups in this study was found by consultation with construction engineers with several years of experience within sewer management. It was concluded that the main separation parameter should be the material. Due to a high number of concrete and plastic sewers, these data groups were further split into sub-groups. For concrete

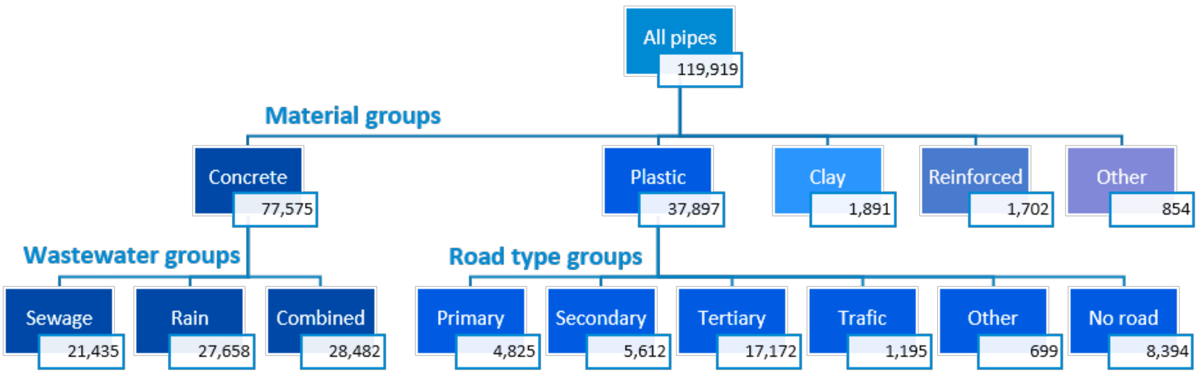

Fig. 1 Overview of the investigated data groups and the corresponding number of pipes.

pipes the sub-groups were based on the wastewater types running in the sewers (sewerage, rain and combined) and for the plastic pipe the sub-groups were based on the road type placed upon them (traffic road, primary road, secondary road, tertiary road and other road). An overview of all the logical data groups investigated in this study and corresponding number of pipes within the data group can be seen in Fig. 1.

2.3. Models

The model used in this study is a Random Forest model with the same settings as described in [16]. A benefit of using a Random Forest model is that it can handle data sets with both binary and continuous variables. Before training the model the data set for each data group is randomly split with 90 % for training and 10 % for test. The model is trained to predict the TS for each of the data groups. The model was set up using scikit-learn in Python [22].

2.4. Test

The models are trained to predict the TS, however, in order to compare the results with other methods the sensitivity and specificity of the models should be calculated. Furthermore, the precision of the models was calculated as this is important when the decision maker must prioritize which sewers to inspect. Another benefit of calculating the precision is that it can be used to account for the mixed distribution of good and bad pipes in different data groups when evaluating the models. In order to calculate the sensitivity, specificity and precision of the models all pipes with a TS ≥ 6 was considered to be in bad condition, and all pipes with a TS < 6 was considered to be in good condition. As earlier described the performance of different models can be very hard to compare due to the prioritization of sensitivity versus specificity. For this reason, the sensitivity of the models was fixed to 0.80 for testing purposes. 0.80 is based on a compromise presented by [14]. In order to adjust the sensitivity of a regression model the split between which pipes are predicted as being in a bad condition and which are predicted to be in a good condition can be changed till the sensitivity hits 0.80 as illustrated in Fig. 2. Finding the sensitivity of 0.80 was done using the Brents algorithm or Golden search for scalar optimization in the Python library SciPy [23]. If the two search methods came up with different sensitivity the one closest to 0.80 was chosen. If the sensitivity was too far away from 0.80, indicating that a local minimum was found for both methods, boundaries for the TS were set in order to find the correct minima. All the models were trained 10 times, and the mean and standard deviation for each data group was calculated.

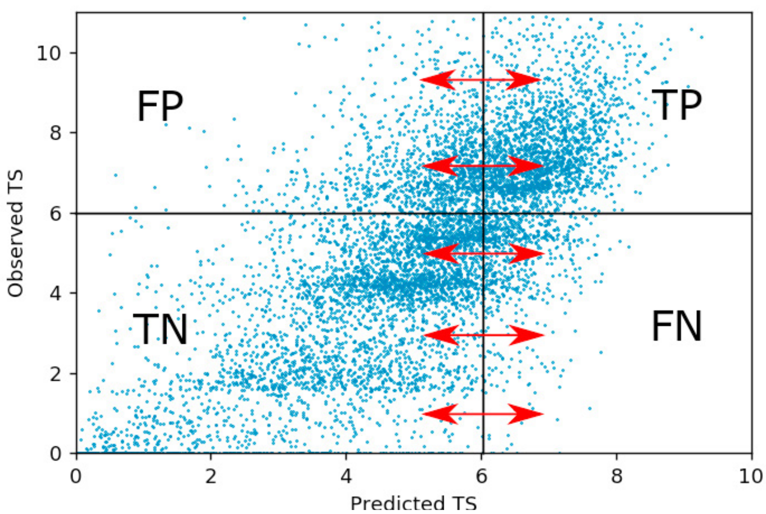

Fig. 2 Illustration of how the balance between sensitivity and specificity can be adjusted by changing the splitting point for the predictions. The figure is inspired from [16].

In order to compare the total performance of the models trained on the logically grouped data sets with the model trained on the full data set the weighted sensitivity, specificity and precision were found. This was done for each level of data groups presented in Fig. 1. The formulas for calculating the weighted sensitivity can be seen in equation 2.

$$WSens = \frac{TP_{DG1} + TP_{DG2} + \dots + TP_{DGn}}{TP_{Tot} + FN_{Tot}} \tag{2}$$

In equation 2 the WSens is the weighted sensitivity, TP_{DG} is the number of true positives in each data group, TP_{Tot} is the total number of true positive predictions and FN_{Total} is the total number of false negative. The formula for calculating the weighted specificity can be seen in equation 3.

$$WSpec = \frac{TN_{DG1} + TN_{DG2} + \dots + TN_{DGn}}{TN_{Tot} + FP_{Tot}} \tag{3}$$

In equation 3 WSpec is the weighted specificity, TN_{DG} is the number of true negatives, TN_{Tot} is the total number of true negative and FP_{Total} is the total number of false positive. The formula for calculating the weighted precision can be seen in equation 4.

$$WPrec = \frac{TP_{DG1} + TP_{DG2} + \dots + TP_{DGn}}{TP_{Tot} + FP_{Tot}} \tag{4}$$

In equation 4 WPrec is the weighted precision.

3. Results

The predictions for respectively the general model and the models trained on pipes within a specific material group can be seen in Fig. 3. From the figure it can be seen where the splitting point should be placed in order to obtain a sensitivity at 0.80 for each of the models.

An overview of the specificity and precision for each model can be seen in Table 1. This table also contains information on the performance of the general model for each data group when the splitting point is adjusted to the specific data groups. The model trained on the full data set obtained a sensitivity at 0.78 ±0.01 when the specificity is fixed at 0.80. The weighted specificity for the models trained on the logically grouped data set was calculated to be 0.73 ± 0.00. Furthermore, the general models obtain a precision of 0.62 ± 0.01 and thereby outperform the models trained on the material specific data sets as these obtain a precision of 0.57 ± 0.00.

3.1. Sub-division of the concrete group

An overview of the performance of the model trained on all the concrete pipes and the models trained on respectively the concrete pipes containing sewage, rain and combined water can be seen in Table 2. From the table it can be seen that the model trained on the concrete data set and models trained on subgroups of the concrete data set performs similar as they obtain a specificity of respectively 0.69 ± 0.00 and 0.69 ± 0.01 and a precision of respectively 0.65 ± 0.01 and 0.65 ± 0.00.

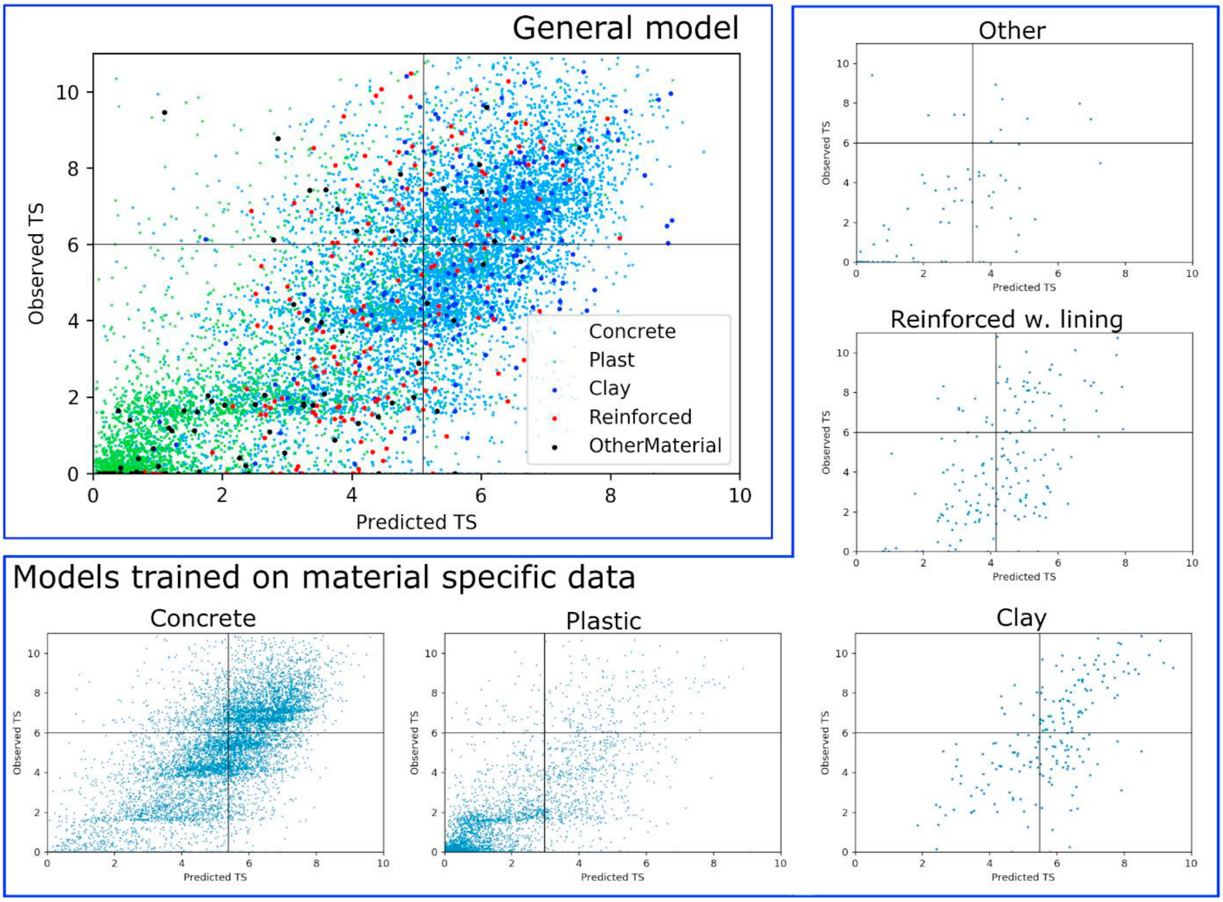

Fig. 3 Results for the general model trained on all the pipes and for the models trained on material specific data. The vertical lines in the graphs show the position of the splitting point in order to obtain a sensitivity of 0.80.

Table 1. Overview of the performance of the general model, the general model's performance on each material group and the performance of the models trained on material specific data when the sensitivity is fixed at 0.80. The mean and standard divination have been found by training each model 10 times.

	General model				Models trained on material specific data			
	Sensitivity	↑Specificity	↑Precision	Splitting point	Sensitivity	↑Specificity	↑Precision	Splitting point
All pipes	0.80 ± 0.00	0.78 ± 0.01	0.62 ± 0.01	5.18 ± 0.03	0.80 ± 0.00	0.73 ±0.00	0.57 ± 0.00	-
Concrete	0.80 ± 0.00	0.69 ± 0.02	0.65 ± 0.01	5.39 ± 0.02	0.80 ± 0.00	0.69 ± 0.01	0.65 ± 0.01	5.38 ± 0.02
Plastic	0.80 ± 0.00	0.82 ± 0.00	0.24 ± 0.02	2.86 ± 0.21	0.80 ± 0.01	0.83 ± 0.02	0.25 ± 0.02	2.93 ± 0.18
Clay	0.80 ± 0.01	0.60 ± 0.01	0.66 ± 0.07	5.44 ± 0.23	0.81 ± 0.02	0.59 ± 0.11	0.65 ± 0.07	5.63 ± 0.28
Reinforced w. lining	0.80 ± 0.01	0.52 ± 0.04	0.44 ± 0.05	4.10 ± 0.20	0.82 ± 0.03	0.55 ± 0.03	0.45 ± 0.04	4.08 ± 0.12
Other material	0.82 ± 0.05	0.76 ± 0.06	0.44 ± 0.08	3.68 ± 0.51	0.80 ± 0.05	0.80 ± 0.09	0.36 ± 0.10	3.33 ± 0.42

Table 2. Overview of the performance of the concrete model, the concrete model's performance on each wastewater group, and the performance of the models trained on concrete pipes containing sewage, rainwater and combined water respectively when the sensitivity is fixed at 0.80. The mean and standard divination have been found by training each model 10 times.

	Concrete model				Models trained on concrete and wastewater data			
	Sensitivity	↑Specificity	↑Precision	Splitting point	Sensitivity	↑Specificity	↑Precision	Splitting point
All concrete pipes	0.80 ± 0.00	0.69 ± 0.01	0.65 ± 0.01	5.38 ± 0.02	0.80 ± 0.00	0.69 ± 0.00	0.65 ± 0.00	-
Sewage	0.80 ± 0.00	0.69 ± 0.02	0.66 ± 0.02	5.49 ± 0.05	0.80 ± 0.00	0.69 ± 0.02	0.67 ± 0.01	5.50 ± 0.05
Rain	0.80 ± 0.00	0.72 ± 0.01	0.60 ± 0.01	5.12 ± 0.04	0.80 ± 0.00	0.71 ± 0.01	0.59 ± 0.01	5.06 ± 0.04
Combined	0.80 ± 0.00	0.65 ±0.02	0.67 ± 0.01	5.51 ± 0.04	0.80 ± 0.00	0.65 ± 0.01	0.67 ± 0.01	5.51 ± 0.03

3.2. Sub-division of the plastic group

It has turned out that it is not possible to make a fair evaluation of whether it is beneficial to split the plastic pipes into road groups with the amount of data available for this study. This is because that only very few bad pipes are available for some of the groups. Especially the category 'other road' is sensitive to this and in one case only one bad pipe was present in the test set. In 'other cases', it was discovered that when the sensitivity in one test was fixed at 0.80 the specificity could easily vary with 0.10 due to the small amount of bad pipes in the data group.

4. Discussion

Sewer management is a very hot topic with several new models being published since 2018. In this paper we investigated if it would be beneficial to develop one general model trained on a data set containing all the available types of pipes or if it would be more beneficial to split the data set into sub-data sets based on logically selected groups of data.

As seen in Table 1. there is no significant difference in the performances of the general model and the models trained on material specific data sets, when considering different material types and corresponding splitting points, as the variation in specificity and precision is within one standard deviation. The similar performance of the two approaches is underlined by the facts that 1) the splitting points for the two approaches are within one standard deviation when considering the different material groups and 2) the distribution of the predictions in Fig. 3 looks similar.

When considering the overall performance of the models, the general model, with a specificity at 0.78 ± 0.01 and a precision at 0.62 ± 0.01, significantly outperforms the models trained on the material specific data sets which obtains a specificity at 0.73 ± 0.00 and a precision at 0.57 ± 0.00. The reason for this is that the general model allows for a sensitivity > 0.80 on e.g. the concrete pipes and < 0.80 on e.g. the plastic pipes, and thereby encounters the fact that the plastic pipes generally are in a better condition than the concrete pipes. This is underlined by the fact that when using the general models' threshold, the sensitivity is 0.84 ± 0.00 for the concrete pipes and 0.30 ± 0.02 for the plastic pipes. Likewise, the specificity and precision for the concrete pipes, when using the general threshold, decreases to respectively 0.63 ± 0.01 and 0.62 ± 0.01 for the concrete pipes and increases to respectively 0.98 ± 0.00 and 0.54 ± 0.04 for the plastic pipes. Thereby the general model encounters the heterogeneity in the data, which is not accounted for when treating each data group individually.

The results for whether it makes sense to split the concrete data set into three sub-data sets for the wastewater types showed that it did not make significant difference. Likewise, it was investigated if it would be beneficial to split the plastic data set according to road type, however, the plastic pipes were in too good condition in order to make a fair comparison.

Table 3 Comparison of our results to the results obtained in the literature. *The numbers were found by calculating the weighted sensitivity and specificity using equation 2 and 3 respectively. ** The precisions for the different data groups were calculated based on confusion matrices presented by the authors

	Our models (best versions)			Best results from Kabir 2018 [15]		
	Sensitivity	↑Specificity	↑Precision	↑Sensitivity	↑Specificity	↑Precision**
All pipes	0.80 ± 0.00	0.78 ± 0.01	0.62 ± 0.01	0.69*	0.88*	0.28
Concrete	0.80 ± 0.00	0.69 ± 0.01	0.65 ± 0.01	0.64	0.89	0.18
Plastic	0.80 ± 0.00	0.83 ± 0.02	0.24 ± 0.02	0.50	0.98	0.40
Clay	0.80 ± 0.01	0.60 ± 0.01	0.66 ± 0.07	0.75	0.86	0.44
Reinforced w. lining	0.80 ± 0.01	0.55 ± 0.03	0.44 ± 0.05	-	-	-
Metallic	-	-	-	0.67	0.97	0.33
Other material	0.80 ± 0.05	0.80 ± 0.09	0.44 ± 0.08	-	-	-

The good performance of the general model according to the models trained on the logically grouped data sets is most likely due to the modelling method used in this paper: Random forest is an ensemble of different decision trees. The decision trees are randomly grown, but weight is put on the most informative splits, so that the most influential parameters are likely to contribute to the trees [24].

A comparison of our results to the results presented in the literature for different data groups can be seen in Table 3.

As seen in Table 3 it can be hard to compare our results to the results obtained by [15], as we have specified the sensitivity to be 0.80 whereas [15] has chosen the balance between sensitivity and specificity based on the form of the ROC curve. For clay pipes [15] performs better than our clay models do when considering specificity whereas our models have a better precision. Regarding the performance on the plastic pipes it is hard to compare our model to [15] as they only have four plastic pipes in bad condition. In all cases except for plastic, our models perform best regarding precision whereas Kabir has a higher sensitivity. The fact that we both obtain a higher sensitivity and precision than [15], while obtaining a lower specificity, indicates that our data set is less heterogeneous than the one used by [15]. The difference in heterogeneity in the data sets can be due to 1) the pipes in the data set used by [15] generally being in a better condition than the pipes in our data set or 2) that we use a broader definition of when a pipe is said to be in a bad condition than [15].

When comparing the results from the model trained on the full data set, without considering any sub-groups, our model performs slightly better than the best model presented by [16]. [16] used the same model setup as used in this article to obtain a specificity at 0.76 for their best model when the sensitivity was fixed at 0.80. This is most likely due to inclusion of the slope and groundwater level which were not included in [16]. Our model also outperforms [14] whom obtained a specificity of 0.47 when the sensitivity was set to 0.80.

5. Conclusion

This paper contributes to the general knowledge about development of deterioration models by investigating how the performance of the predictions is influenced by the training data set. It was investigated how the performance was affected when training on a full data set and when training on logically grouped sub-data sets. The results showed that there is no significant difference between the two approaches when considering their performance on specific data groups. Moreover, it was shown, that the general model performed better when considering the overall performance as it encounters the heterogeneity of the data.

References

[1] H. Korving, F. H. L. R. Clemens, og J. M. van Noortwijk, “Statistical Modeling of the Serviceability of Sewage Pumps”, *J. Hydraul. Eng.*, bd. 132, nr. 10, s. 1076–1085, okt. 2006.

[2] J. B. Haurum og T. B. Moeslund, “A Survey on Image-Based Automation of CCTV and SSET Sewer Inspections”, *Automation in*

Construction, bd. 111, s. 103061, mar. 2020.

[3] A. Altarabsheh, M. Ventresca, og A. Kandil, “New Approach for Critical Pipe Prioritization in Wastewater Asset Management Planning”, *J. Comput. Civ. Eng.*, bd. 32, nr. 5, s. 04018044, sep. 2018.

[4] M. Elmasry, A. Hawari, og T. Zayed, “An economic loss model for failure of sewer pipelines”, *Structure and Infrastructure Engineering*, bd. 14, nr. 10, s. 1312–1323, okt. 2018.

[5] M. Elmasry, T. Zayed, og A. Hawari, “Multi-Objective Optimization Model for Inspection Scheduling of Sewer Pipelines”, *J. Constr. Eng. Manage.*, bd. 145, nr. 2, s. 04018129, feb. 2019.

[6] S. M. Ghavami, Z. Borzooei, og J. Maleki, “An effective approach for assessing risk of failure in urban sewer pipelines using a combination of GIS and AHP-DEA”, *Process Safety and Environmental Protection*, bd. 133, s. 275–285, jan. 2020.

[7] N. Balekelayi og S. Tesfamariam, “Statistical Inference of Sewer Pipe Deterioration Using Bayesian Geoadditive Regression Model”, *J. Infrastruct. Syst.*, bd. 25, nr. 3, s. 04019021, sep. 2019.

[8] N. Caradot *m.fl.*, “Practical benchmarking of statistical and machine learning models for predicting the condition of sewer pipes in Berlin, Germany”, *Journal of Hydroinformatics*, bd. 20, nr. 5, s. 1131–1147, sep. 2018.

[9] N. Caradot *m.fl.*, “The influence of condition assessment uncertainties on sewer deterioration modelling”, *Structure and Infrastructure Engineering*, bd. 16, nr. 2, s. 287–296, feb. 2020.

[10] N. Hernández, N. Caradot, H. Sonnenberg, P. Rouault, og A. Torres, “Support tools to predict the critical structural condition of uninspected pipes for case studies of Germany and Colombia”, *Water Practice and Technology*, bd. 13, nr. 4, s. 794–802, dec. 2018.

[11] N. Hernández, N. Caradot, H. Sonnenberg, P. Rouault, og A. Torres, “Optimizing SVM Model as Predicting Model for Sewer Pipes in the Two Main Cities in Colombia”, i *New Trends in Urban Drainage Modelling*, G. Mannina, Red. Cham: Springer International Publishing, 2019, s. 926–931.

[12] P. Lin, X. Yuan, og E. Tovilla, “Integrative modeling of performance deterioration and maintenance effectiveness for infrastructure assets with missing condition data”, *Computer-Aided Civil and Infrastructure Engineering*, bd. 34, nr. 8, s. 677–695, aug. 2019.

[13] T. Laakso, T. Kokkonen, I. Mellin, og R. Vahala, “Sewer Life Span Prediction: Comparison of Methods and Assessment of the Sample Impact on the Results”, *Water*, bd. 11, nr. 12, s. 2657, dec. 2019.

[14] T. Laakso, T. Kokkonen, I. Mellin, og R. Vahala, “Sewer Condition Prediction and Analysis of Explanatory Factors”, *Water*, bd. 10, nr. 9, s. 1239, sep. 2018.

[15] G. Kabir, N. B. C. Balek, og S. Tesfamariam, “Sewer Structural Condition Prediction Integrating Bayesian Model Averaging with Logistic Regression”, *J. Perform. Constr. Facil.*, bd. 32, nr. 3, s. 04018019, jun. 2018.

[16] B. D. Hansen, D. Getreuer Jensen, S. H. Rasmussen, J. Tamouk, M. Uggerby, og T. B. Moeslund, “General Sewer Deterioration Model Using Random Forest”, i *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, Xiamen, China, dec. 2019, s. 834–841.

[17] E. V. Ana og W. Bauwens, “Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods”, *Urban Water Journal*, bd. 7, nr. 1, s. 47–59, feb. 2010.

[18] M. M. Rokstad og R. M. Ugarelli, “Evaluating the role of deterioration models for condition assessment of sewers”, *J Hydroinform*, bd. 17, nr. 5, s. 789–804, sep. 2015.

[19] X. Li *m.fl.*, “Evaluation of data-driven models for predicting the service life of concrete sewer pipes subjected to corrosion”, *Journal of Environmental Management*, bd. 234, s. 431–439, mar. 2019.

[20] B. Laden, DANVA, og Fotomanualgruppen, *Fotomanualen: TV-inspektion af afløbsledninger*. Skanderborg: DANVA, 2010.

[21] Dansk Vand- og Spildevandsforening og DANVA, *Fotomanualen: beregning af fysisk indeks ved TV-inspektion.* Skanderborg: Dansk Vand- og Spildevandsforening, 2005.

[22] F. Pedregosa *m.fl.*, “Scikit-learn: Machine Learning in Python”, *Journal of Machine Learning Research*, bd. 2011, nr. 12, s. 2825−2830.

[23] P. Virtanen *m.fl.*, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”, *Nature Methods*, bd. 17, s. 261–272, 2020.

[24] T. G. Dietterich, “An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization”, *Machine Learning*, bd. 40, nr. 2, s. 139–157, 2000.