



Application of regression methods for classification of sewers' damages

Małgorzata Kutylowska¹ · Dariusz Kowalski²

Received: 17 December 2020 / Accepted: 30 August 2021 / Published online: 3 September 2021
© The Author(s) 2021

Abstract

The paper presents possibilities of application of selected regression methods (classification trees, support vector machines, *K*-nearest neighbours, artificial networks) for classification of sewers' damages. Operational data from the time span 2006–2011 obtained from water utility were used for deterioration analysis. On the basis of the following independent variables, the modelling was carried out: diameter, depth, year of construction, material and season of damage's occurring. The following kinds of damages were classified: corrosion, crack, longitudinal crack, displacement, unsealing, failure, collapse. The main aim of the paper was to check if prediction methodology could be useful for classification of different kinds of sewers' damages. The obtained results pointed out that proposed classification methods are not appropriable in quality analysis of registered damages of sewers. Moreover, it is recommended for water and sewerage companies to register types of failures using unified notation which make easier preliminary classification before applying modelling approach. The calculations were performed in Statistica 13.1 software.

Keywords Failures · Sewers · Modelling · Prediction methods

Introduction

Sewage removal systems belong to the critical infrastructure, and their proper operation and maintenance are very important tasks from the point of view of both the operators and the users. For many years in Poland and in the world, quantitative research has been conducted to analyse in depth the causes and effects of failures and the number of failures and abnormalities occurring in sewers (Kuliczowska 2017, 2016; Caradot et al. 2018; Miszta-Kruk 2016; Ana et al. 2009). The quality and moisture of soil surrounding the buried infrastructure have a great influence on the technical condition of pipes (Iwanek 2018). Much attention is devoted to problems relating to the renovation or upgrading of sewers (Ariaratnam and MacLeod 2002; Madryas and Wysocki

2008; Duchesne et al. 2014) and to planning and carrying out video camera pipe inspections (Plihal et al. 2016; Harwey and McBean 2014).

The main aim of this paper is to show how selected regression methods could be used to classify and predict types of sewer damage. Recently, an attempt was made to use modelling based on the multicriteria decision process to select sewer sections for replacement and renovation (Kessili and Benmamar 2016). Moreover, artificial intelligence assumptions (Tran et al. 2006) were used to analyse the failure rate of a storm sewer system. Such a system (drainage structure at urban areas) should be designed on the basis of the latest data concerning the amount of precipitation in the analysed area (Wartalska et al. 2020). Prediction methods, used to model phenomena which are difficult to describe because of the dynamically changing interdependences between the variables, are currently increasingly often used in the broadly understood environmental engineering (Czapczuk et al. 2015), e.g. to analyse the failure frequency of pumping units (Bevilacqua et al. 2003), to locate water-pipe network leakages (Candelieri et al. 2014), in hydraulic calculations of water distribution networks (Czapczuk et al. 2017; Piasecki et al. 2018) and to predict the hydraulic conductivity of the ground (Elbisy 2015). The prediction methods

✉ Małgorzata Kutylowska
malgorzata.kutylowska@pwr.edu.pl

¹ Faculty of Environmental Engineering, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

² Faculty of Environmental Engineering, Lublin University of Technology, ul. Nadbystrzycka 40B, 20-618 Lublin, Poland

include: classification trees (CT), support vector machines (SVM), K -nearest neighbours (KNN) and artificial neural networks (ANN). The particular methods are based on different assumptions and entail the proper selection of model parameters. In the case of artificial neural networks, it is necessary to determine the number of neurons in the hidden layer and select an appropriate training method. Classification trees are based on the proper division of variables at the particular tree levels. The KNN method uses the notion of the distance of the dependent variable from the independent variable. The SVM method is based on the notion of decision spaces. The prediction of variables (by means of regression and classification methods) is a two-stage process consisting in creating a model describing the investigated phenomenon and then in applying the model to predict the selected dependent variable on the basis of the independent variables (predictors). More information about the above methods can be found in (Hastie et al. 2008; Breiman et al. 1984; Cristianini and Shawe-Taylor 2000). Further, in this paper the terms “regression method” and “modelling method” are used interchangeably. The sewerage systems are relatively old and exploiters all over the world as well as scientist should pay more attention to proper maintenance, rehabilitation and active damage localization. Investigations using statistical models in the assessment of damage to sewers have been done in the world, e.g. sewer pipe condition was established by means of neural networks (Hassan et al. 2019), diagnosis of sewer defects was based on support vector machines (Yang and Su 2008), random forest was used to model sewer deterioration (Hansen et al. 2020). Nevertheless, the literature on the subject indicates that so far broadly understood regression and classification methods (prediction methods) have not been used to evaluate the condition and failure frequency of sewerage networks in Poland. Considering this fact, the authors decided to undertake research in this area. Proper appointment of the most important parameters of models and selection of number and kinds of independent variables influencing predicted values of dependent variable seem to be the crucial problems during mathematical modelling using machine learning approach. Moreover, it is necessary to check if application of selected prediction methods could be useful for rational management of buried infrastructure. Carried studies pointed out that appropriate arrangement of operational data registered in water and sewerage companies is necessary and should result in greater possibilities of using such information for construction of reliability model in the future.

Material and methods

Operational data for the years 2006–2011, obtained from one of the water and sewerage companies in a selected Polish town (number of citizens exceeds 300,000), concerning the recorded sewerage network failures were used to classify and predict types of sewerage network damage by means of the four regression methods mentioned above. Sewerage network is built in gravitational separate system. There are several pumping stations due to differences of altitude. Sanitary sewers (the total length was equal to almost 568 km at the end of 2011) collect and transport sewage to wastewater treatment plant. The length of sewer laterals amounted nearly to 198 km at the end of 2011. Some parts of the network were constructed in the beginning of XX century. Sewerage network is still expanding and renovated, and is made mainly from vitrified clay, concrete, reinforced concrete and PVC. Almost 81% of the total number of damages occurred on sewers made from vitrified clay. The rest failures were registered on concrete sewers (ca. 16%) and on PVC conduits (ca. 3%). The most deteriorated (circa 60% of all failures) were sewers built in the time span 1970–1990. The average value of sewers' failure rate equalled to 0,0444 fail./(km·a). Considering their different roles in the system and different damage occurrence determinants, separate classification models were built for sewers and sewer laterals. Such abnormalities in sewerage pipe functioning as clogups, chokings or blockages were not taken into account. The focus was on typical damage, such as pipe breakage, corrosion and loss of integrity (leakage) and sewer cave-in. In some cases, the type of damage was not precisely defined. It was merely noted in the documentation that a failure occurred. In such cases, the general term “damage” was used without going into details as to the type and quality of the damage. That assumptions mean that not all damages were taken into account during statistical modelling. For further investigations, the total number of sewers' failures varied between 8 in 2011 and 29 in 2006. The type of damage (TD) was the dependent variable while the independent variables (predictors), on the basis of which the classification was made, were the following parameters describing sewerage conduits: (1) quantitative parameters—conduit diameter (D), conduit laying depth (LD), year of conduit construction (YC) and (2) qualitative parameters—season in which failure occurred (SE: winter (W), summer (SM)) and conduit material (M : vitrified clay (VC), cast iron (CI), concrete (C), PVC). The summer (non-heating) season comprised the months from March to October inclusive. Because of the relatively small sample size, V -fold cross-validation was used. In the case of the CT method, the following model parameters were adopted: equal wrong

classification cost, Gini coefficient (a measure of random variable distribution concentration) as goodness of fit, a priori estimated probability, tree pruning at wrong classification, node count-5, maximum 10 tree levels and maximum 1000 nodes. Four types of basis functions: linear (*L*), sigmoidal (*S*), polynomial (*P*) and radial basis functions (RBF) were used in SVM modelling. The maximum number of iterations-1000, the capacity-10, training ended at an error of 0.001. Rescaling of predictors and data narrowing were used. The number of support vectors and localized vectors is given when analysing the results for the selected SVM model. Four distance measures: Euclidean distance (KNN-E), squared Euclidean distance (KNN-E2), Manhattan distance (KNN-M) and Chebyshev distance (KNN-C) were used in the KNN method. The number of nearest neighbours was determined by the trial-and-error method to be 10. The other nearest neighbour values would result in worse classification results. ANN models based on radial basis functions, in which the minimum and maximum number of hidden neurons amounted to, respectively, 10 and 14 (sewers) and 6 and 8 (sewer laterals), and multilayer perceptrons for which the minimum and maximum number of hidden neurons amounted to 1 and 20, were built. Sum square and mutual entropy as the error function and the following activation functions: linear, logistic, hyperbolic tangent, exponential and sine functions were used. Altogether 20 ANN models were built. 95 failures of sewers and 39 failures of sewer laterals occurred in the analysed period. The whole data set was randomly divided into a training set (70% of the

cases) and testing set (30% of the cases). The training set data were used to build models while the testing set data were used to verify the modelling and its quality. Tables 1 and 2 show dependent variable and predictor values for, respectively, training and testing. The calculations were performed using Statistica 13.1.

Results and discussion

Figures 1 and 2 present the classification accuracy results for all the types of damage to the sewers and to the sewer laterals for the modelling methods: CT; SVM-L, SVM-S, SVM-P, SVM-RBF (the support vector machine method based on linear, sigmoidal, polynomial and radial basis functions); ANN-4 and ANN-5 (the selected optimal ANN models). The results of classification by means of the *K*-nearest neighbours method are for only the testing data. Therefore, they are presented in a tabular form (Table 3). The results for the training data for the KNN model are not available in Statistica.

The model and method most suitable for the classification of the types of damage to the considered sewerage conduits were selected on the basis of an analysis of classification accuracy. The latter is a ratio (expressed in per cent) of the accurately classified failures (with no division into the particular types) to all the failures, separately for the training data and the testing data. Another parameter which should be taken into account when selecting a prediction method is the architecture of the given model.

Table 1 Ranges of predictors and dependent variable—sewers

SE	<i>M</i>	LD, m	<i>D</i> , mm	YC	TD
<i>Training set (67 cases)</i>					
W, SM	C, VC, PVC	1.0–6.5	110–600	1926–2001	Corrosion, breakage, longitudinal breakage, shift, leakage, damage, cave-in
<i>Testing set (28 cases)</i>					
W, SM	C, VC, PVC	1.2–4.2	150–450	1926–1995	Corrosion, breakage, longitudinal breakage, shift, leakage, damage, cave-in

Table 2 Ranges of predictors and dependent variable—sewer laterals

SE	<i>M</i>	LD, m	<i>D</i> , mm	YC	TD
<i>Training set (29 cases)</i>					
W, SM	C, VC, PVC, CI	1.2–4.3	110–200	1926–1995	Breakage, longitudinal breakage, shift, leakage, damage, cave-in
<i>Testing set (10 cases)</i>					
W, SM	VC, PVC, CI	1.2–3.5	150–250	1938–2008	Longitudinal breakage, leakage, damage, cave-in

Fig. 1 Classification accuracy of sewers' damages

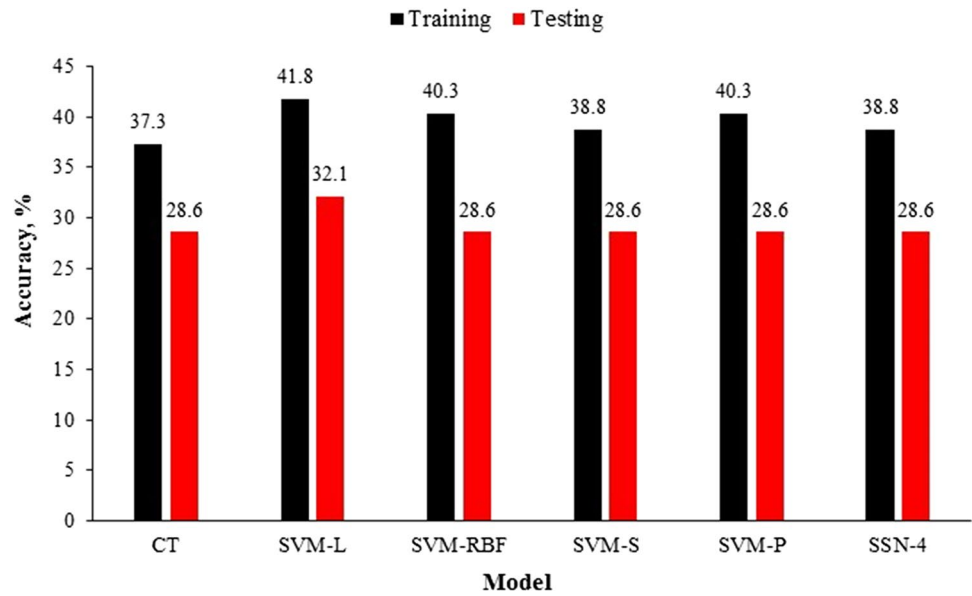


Fig. 2 Classification accuracy of sewer laterals' damages

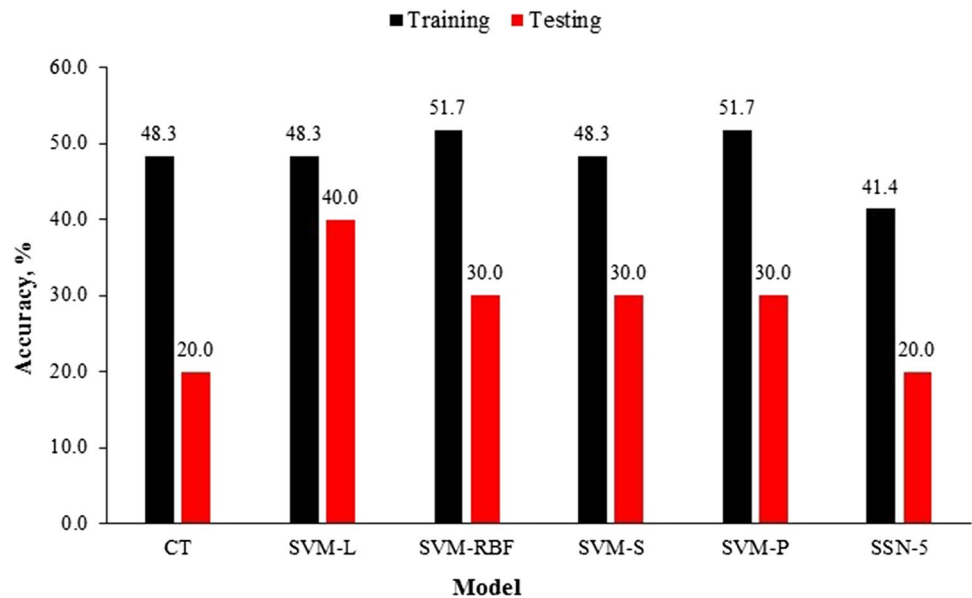


Table 3 Classification accuracy (%) of damages—KNN method

Model	Sewers	Sewer laterals
KNN-E	24.1	25.0
KNN-E2	24.1	25.0
KNN-M	24.1	16.7
KNN-C	24.1	16.7

However, since four different methods were analysed as part of this study, a comparison of their structures is not a clear-cut task. Each of the regression methods is based on different assumptions and comprises different training

parameters, which makes a comparative analysis quite difficult. Therefore, only a preliminary comparison based on the classification accuracy mentioned above was made.

An analysis of Fig. 1 and Table 3 shows that none of the proposed methods is a tool accurately classifying types of sewer damage. In the KNN models, the classification accuracy results are the same for different measures of distance, which is surprising since each of the distance measures uses a different relation between the dependent variable and the predictors. This problem looked differently when predicting water conduit damage intensity (Kutyłowska 2018). Thus, the above results put a question mark over the use of the KNN method for both regression

and classification tasks. The best classification results were obtained using the support vector machine method based on the linear basis function (SVM-L) (Fig. 1). The damage type classification accuracy amounted to just under 42% and 32% for, respectively, the training data and the testing data. Detailed results of the classification of the particular types of damage by the SVM-L method, for which the number of support vectors and localized vectors amounted to, respectively, 67 and 0, are presented in Table 4. Even though the general classification accuracy result is at the level of 42% (training), the detailed results are not so promising. Only corrosion and cave-in were correctly classified by the model, whereas such abnormalities as shift, breakage and general damage were incorrectly classified. Unfortunately, the classification results for the training data (Table 5) are even worse. The few types of damage which were correctly classified are in bold. The

testing data set contained 28 cases of which only 9 were unequivocally classified by the SVM-L model.

From the two longitudinal breakages, one was classified correctly while corrosion and cave-in were assigned correctly in 100%. Table 5 shows results for only the selected model. The results of classification by means of other methods and models were even worse.

The failures of the sewer laterals were somewhat better classified (Fig. 2). Also SVM-L was chosen as the optimal model since the classification results for both the training and testing data were acceptable for all the types of damage. Even if accuracy for the training data was higher in the case of SVM-RBF and SVM-P, the results for the testing data were slightly worse than for the model based on the linear basis function. Similarly as in the case of sewers, the SVM-L model had no localized vectors while the number of support vectors was 29, i.e. it was equal to the sample size. Tables 6

Table 4 Classification of sewers’ damages—model SVM-L, training set

Class name	Totally	Correct	Incorrect	Correct, %	Incorrect, %
Corrosion	4	4	0	100.0	0.0
Shift	11	0	11	0.0	100.0
Breakage	4	0	4	0.0	100.0
Longitudinal breakage	6	2	4	33.3	66.7
Leakage	11	1	10	9.1	90.9
Damage	8	0	8	0.0	100.0
Cave-in	23	21	2	91.3	8.7

Table 5 Classification of sewers’ damages—model SVM-L, testing set

Observed	Model SVM-L						
	Damage	Longitudinal breakage	Leakage	Cave-in	Shift	Corrosion	Breakage
Damage (6)	0	0	0	6	0	0	0
Longitudinal breakage (2)	0	1	0	1	0	0	0
Leakage (6)	0	1	0	3	0	2	0
Cave-in (6)	0	0	0	5	0	1	0
Shift (2)	0	0	0	2	0	0	0
Corrosion (3)	0	0	0	0	0	3	0
Breakage (3)	0	0	0	3	0	0	0

Table 6 Classification of sewer laterals’ damages—model SVM-L, training set

Class name	Totally	Correct	Incorrect	Correct, %	Incorrect, %
Shift	3	0	3	0.0	100.0
Breakage	1	0	1	0.0	100.0
Longitudinal breakage	2	1	1	50.0	50.0
Leakage	7	1	6	14.3	85.7
Damage	7	7	0	100.0	0.0
Cave-in	9	5	4	55.6	44.4
Shift	3	0	3	0.0	100.0

Table 7 Classification of sewer laterals' damages—model SVM-L, testing set

Observed	Model SVM-L			
	Leakage	Cave-in	Damage	Longitudinal breakage
Leakage (5)	2	0	2	1
Cave-in (2)	0	0	1	1
Damage (2)	0	0	1	1
Longitudinal breakage (1)	0	0	0	1

and 7 show classification accuracy for the particular types of damage for the training data and the testing data.

Only the abnormality generally referred to as “damage” was correctly classified in all the considered cases. Cave-in and longitudinal breakage were classified correctly in 50%, whereas most of the other types of damage were not predicted (classified) correctly by the chosen model. In the case the results for testing, three of the four types of abnormalities (leakage, cave-in, damage) were classified correctly in about 50%, which is not a wholly satisfying result. One should look for the causes of so poor results of classification by means of the selected regression methods in the peculiarities of the problem, as discussed below. In order to thoroughly investigate the problem of the unsatisfactory results of the classification of types of damage, one can use a table containing operational data. Table 8 shows a few illustrative cases (for training and testing) for sewer laterals. This quite representative sample is enough to demonstrate certain interdependences (or their absence) which had an unquestionable influence on the quality of the modelling of the classification problems by means of the regression methods.

An analysis of the exemplary operational data (Table 8) indicates their high variability. For example, a leakage

occurred in a vitrified clay sewer lateral 150 mm in diameter laid 2.5 m under the ground surface in the non-heating season in 1980, whereas in the case of another sewer lateral made of the same material and having the same diameter, but laid 1.6 m below the ground in 1997, a shift of the conduit occurred. The classification results indicate that the selected regression methods are not flexible enough to correctly classify damage when there is a considerable discrepancy between the operational data used to build the models. Similar problems are encountered in the case of operational data and the modelling of types of sewer damage. The variability of the dependence between, e.g., the conduit material, its diameter and the abnormality to which it was subject in the analysed period, or its absence is quite normal from the engineering and operational point of view. Nevertheless, such problems as indicated above should have no influence on the quality of modelling and the accuracy of classification, especially that as many as four different regression methods were selected for the modelling. The results presented in this paper are one of the first attempts at classifying damage types by means of regression methods, but the obtained results are not promising. Perhaps in the future this problem should be approached differently, i.e. a preliminary classification of damage should be made depending on what class of abnormalities it belong to (e.g. a class of structural damage, a class of damage affecting conduit operation hydraulics, etc.), which could improve the quality of the modelling. Moreover, a separate problem is the damage classification system used in water and sewerage companies. Its standardization and the use of precisely defined abnormalities and damage types would make it possible to obtain a more representative sample, which could significantly improve the quality of the results of analyses carried out using the prediction methods.

Table 8 Exemplary operational data—sewer laterals

Sample	Material	Diameter, mm	Depth, m	Season	Year of construction	Type of damage
Training	VC	150	2.5	SM	1980	Leakage
Training	VC	150	1.6	SM	1977	Shift
Training	VC	150	1.8	SM	1971	Cave-in
Training	VC	150	1.2	SM	1960	Leakage
Training	VC	150	1.4	SM	1995	Damage
Training	VC	150	1.8	SM	1982	Damage
Testing	VC	150	3.0	SM	1938	Leakage
Training	CI	150	2.5	SM	1978	Shift
Training	VC	150	1.3	SM	1965	Cave-in
Training	VC	150	1.5	SM	1981	Breakage
Testing	VC	150	2.2	W	1994	Cave-in

Conclusions

The obtained results show that the proposed regression methods (classification trees, the support vector machine method, the K -nearest neighbours method, artificial neural networks) do not classify types of damage to sewers in a satisfactory way since the accuracy of the classification for all the types of abnormalities did not exceed 50% for the adopted SVM-L model. At the current stage of this research, one cannot unreservedly recommend any of the prediction methods for use in the qualitative evaluation of the condition and damageability of sewerage networks. Undoubtedly further, more advanced, research on the classification and prediction of qualitative variables is needed, also based on other operational data. Moreover, it seems necessary to standardize the system of classification and notation of damage types not only within a particular water and sewerage company, but also for the whole country. Of course, there are world guidelines and Polish recommendations proposed by Kuliczkowski and Kuliczowska (2007) for classifying abnormalities occurring in sewerage networks, but their implementation in many water and sewerage companies still poses a difficulty.

Funding The author(s) received no specific funding for this work.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ana E, Bauwens W, Pessemier M, Thoeye C, Smolders S, Boonen I, De Gueldre G (2009) An investigation of the factors influencing sewer structural deterioration. *Urban Water J* 6(4):303–312. <https://doi.org/10.1080/15730620902810902>
- Ariaratnam ST, MacLeod CW (2002) Financial outlay modeling for a local sewer rehabilitation strategy. *J Constr Eng Manag* 128(6):486–495. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2002\)128:6\(486\)](https://doi.org/10.1061/(ASCE)0733-9364(2002)128:6(486))
- Bevilacqua M, Braglia M, Montanari M (2003) The classification and regression tree approach to pump failure rate analysis. *Reliab Eng Syst Saf* 79(1):59–67. [https://doi.org/10.1016/S0951-8320\(02\)00180-1](https://doi.org/10.1016/S0951-8320(02)00180-1)
- Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Chapman and Hall/CRC, Boca Raton
- Candelieri A, Soldi D, Conti D, Archetii F (2014) Analytical leakages localization in water distribution networks through spectral clustering and support vector machines. The Icewater Approach. *Procedia Eng* 89:1080–1088. <https://doi.org/10.1016/j.proeng.2014.11.228>
- Caradot N, Rouault P, Clemens F, Cherqui F (2018) Evaluation of uncertainties in sewer condition assessment. *Struct Infrastruct Eng* 14(2):264–273. <https://doi.org/10.1080/15732479.2017.1356858>
- Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
- Czapczuk A, Dawidowicz J, Piekarski J (2015) Artificial intelligence methods in the design and operation of water supply systems. *Annu Set Environ Prot* 17(2):1527–1544
- Czapczuk A, Dawidowicz J, Piekarski J (2017) Application of multilayer perceptron for the calculation of pressure losses in water supply lines. *Annu Set Environ Prot* 19:200–210
- Duchesne S, Bouchard K, Toumbou B, Villeneuve JP (2014) Assessing the impact of renewal scenarios on the global structural state of sewer pipe networks. *Can J Civ Eng* 41:761–768. <https://doi.org/10.1139/cjce-2014-0037>
- Elbisy MS (2015) Support vector machine and regression analysis to predict the field hydraulic conductivity of sandy soil. *KSCE J Civ Eng* 19(7):2307–2316. <https://doi.org/10.1007/s12205-015-0210-x>
- Hansen BD, Rasmussen SH, Moeslund TB, Uggerby M, Jensen DG (2020) Sewer deterioration modeling: the effect of training a random forest model on logically selected data-groups. *Procedia Comput Sci* 176:291–299. <https://doi.org/10.1016/j.procs.2020.08.031>
- Harvey R, McBean E (2014) Comparing the utility of decision trees and support vector machines when planning inspections of linear sewer infrastructure. *J Hydroinf* 16(6):1265–1279. <https://doi.org/10.2166/hydro.2014.007>
- Hassan SI, Dang MN, Mehmood I, Im S, Choi C, Kang J, Park YS, Moon H (2019) Underground sewer pipe condition assessment based on convolutional neural networks. *Autom Constr* 106:1028–1049. <https://doi.org/10.1016/j.autcon.2019.102849>
- Hastie T, Tibshirani R, Friedman J (2008) The elements of statistical learning. Data mining, inference and prediction. Springer series in statistics. Springer, New York, NY. <https://doi.org/10.1007/978-0-387-84858-7>
- Iwanek M (2018) Numerical investigations of moisture distribution in a selected anisotropic soil medium. *Eurasian Soil Sci* 51(1):73–80. <https://doi.org/10.1134/S1064229318010088>
- Kessili A, Benmamar S (2016) Prioritizing sewer rehabilitation projects using AHP-PROMETHEE II ranking method. *Water Sci Technol* 73(2):283–291. <https://doi.org/10.2166/wst.2015.490>
- Kuliczowska E (2016) Risk of structural failure in concrete sewers due to internal corrosion. *Eng Fail Anal* 66:110–119. <https://doi.org/10.1016/j.engfailanal.2016.04.026>
- Kuliczowska E (2017) Environmental and structural risk assessment of long operated vitrified clay sewers. *Environ Prot Eng* 44(2):53–67. <https://doi.org/10.5277/epe170205>
- Kuliczkowski A, Kuliczowska E (2007) Critical comments regarding applied classifications of sewers' damages. *Instal* 4:42–47
- Kutyłowska M (2018) K-Nearest Neighbours method as a tool for failure rate prediction. *Periodica Polytechnica Civ Eng* 62(2):318–322. <https://doi.org/10.3311/PPci.10045>
- Madryas C, Wysocki L (2008) Renovation of brick interceptor sewers. *Tunn Undergr Space Technol* 23:718–726. <https://doi.org/10.1016/j.tust.2007.12.007>
- Miszta-Kruk K (2016) Reliability and failure rate analysis of pressure, vacuum and gravity sewer systems based on operating data. *Eng Fail Anal* 61:37–45. <https://doi.org/10.1016/j.engfailanal.2015.07.034>

- Piasecki A, Jurasz J, Kaźmierczak B (2018) Forecasting daily water consumption: a case study in town, Poland. *Periodica Polytechnica-Civ Eng* 62(3):818–824. <https://doi.org/10.3311/PPci.11930>
- Plihal H, Kretschmer F, Bin Ali MT, See CH, Romanova A, Horoshenkov KV, Ertl Th (2016) A novel method for rapid inspection of sewer networks: combining acoustic and optical means. *Urban Water J* 13(1):3–14. <https://doi.org/10.1080/1573062X.2015.1076857>
- Tran DH, Ng AWM, Perera BJC, Burn S, Davis P (2006) Application of probabilistic neural networks in modelling structural deterioration of stormwater pipes. *Urban Water J* 3(3):175–184. <https://doi.org/10.1080/15730620600961684>
- Wartalska K, Kaźmierczak B, Nowakowska M, Kotowski A (2020) Analysis of hyetographs for drainage system modelling. *Water* 12(1):1–21. <https://doi.org/10.3390/w12010149>
- Yang MD, Su TC (2008) Automated diagnosis of sewer pipe defects based on machine learning approaches. *Expert Syst Appl* 35:1327–1337. <https://doi.org/10.1016/j.eswa.2007.08.013>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.