



# Support Vector Machines Used for the Prediction of the Structural Conditions of Pipes in Bogota's Sewer System\*

Aplicación de máquinas de soporte vectorial para predecir la condición estructural de las tuberías no inspeccionadas del sistema de alcantarillado de Bogotá

Submitted on: March 28, 2019 | Accepted on: February 14, 2020 | Published: June 17, 2021

**Nathalie Hernandez<sup>a</sup>**

Pontificia Universidad Javeriana, Colombia

ORCID: 0000-0001-5084-7937

**Nicolas Caradot**

Kompetenzzentrum Wasser Berlin, Germany

ORCID: 0000-0002-5252-4880

**Hauke Sonnenberg**

Kompetenzzentrum Wasser Berlin, Germany

ORCID: 0000-0001-9134-2871

**Pascale Rouault**

Kompetenzzentrum Wasser Berlin, Germany

ORCID: 0000-0003-3986-0123

**Andres Torres**

Pontificia Universidad Javeriana, Colombia

ORCID: 0000-0001-8693-8611

\* Research article

<sup>a</sup> Corresponding author. E-mail: [nathalie\\_hernandez@javeriana.edu.co](mailto:nathalie_hernandez@javeriana.edu.co)

DOI: <https://doi.org/10.11144/Javeriana.iued25.svmu>

## How to cite this article:

N. Hernandez, N. Caradot, H. Sonnenberg, P. Rouault, and A. Torres, "Support vector machines used for the prediction of the structural conditions of pipes in Bogota's sewer system," *Ing. Univ.*, vol. 25, 2021 [Online]. <https://doi.org/10.11144/Javeriana.iued25.svmu>

## Abstract

*Objective:* this paper focused on: (i) developing a deterioration model based on support vector machines (SVM) from its regression approach to separate the prediction of the structural condition of sewer pipes from a classification by grades and predict the scores obtained by failures found in CCTV inspections; and (ii) comparing the prediction results of the proposed model with the ones obtained by a deterioration model based on SVM classification tasks to explore the advantages and disadvantages of their predictions from different perspectives. *Materials and methods:* The sewer network of Bogota was the case study for this work in which a dataset consisting of the characteristics of 5031 pipes inspected by CCTV (obtained by GIS) was considered, as well as information on external variables (e.g., age, sewerage, and road type). Probability density functions (PDF) were used to convert the scores given by failures found in CCTV into structural grades. In addition, three techniques were used to evaluate the predictions from different perspectives: positive likelihood rate (PLR), performance curve and deviation analysis. *Results:* it was found that: (i) SVM-based deterioration model used from its regression approach is suitable to predict critical structural conditions of uninspected sewer pipes because this model showed a PLR value around 6.8 (the highest value among the predictions of all structural conditions for both models) and 74 % of successful predictions for the first 100 pipes with the highest probability of being in critical conditions; and (ii) SVM-based deterioration model used from its classification approach is suitable to predict other structural conditions because this model showed homogeneous PLR values for the prediction of all structural conditions (PLR values between 1.67 and 3.88) and deviation analysis results for all structural conditions are lower than the ones for the SVM-based model from its regression approach.

**Keywords:** Sewer asset management, structural condition, classification and regression models, support vector machines (SVM).

## Resumen

*Objetivo:* este artículo se centra en: (i) desarrollar un modelo de deterioro basado en máquinas de soporte vectorial (SVM) a partir de su enfoque regresivo para desligar la predicción de la condición estructural del alcantarillado de los grados de clasificación y predecir los puntajes dados por las fallas encontradas en las inspecciones CCTV; y (ii) comparar los resultados de predicción del modelo propuesto con aquellos resultados obtenidos de un modelo basado en SVM a partir de su enfoque de clasificación con el fin de explorar las ventajas y desventajas en sus predicciones bajo diferentes perspectivas. *Materiales y métodos:* el caso de estudio considerado fue la red de alcantarillado de Bogotá, el cual contaba con 5031 tuberías inspeccionadas, información de las características físicas de las tuberías e información de factores externos (p. e., edad, tipo de afluente y tipo de vía). Las funciones de densidad de probabilidad (FDP) se utilizaron para convertir los puntajes de las fallas encontradas en las inspecciones de CCTV en grados estructurales. Además, se utilizaron tres técnicas para evaluar las predicciones mediante diferentes perspectivas: tasas de verosimilitud positiva (TPR), curvas de rendimiento y análisis de desviación. *Resultados:* se encuentra que: (i) el modelo de deterioro basado en SVM a partir de su enfoque de regresión es apropiado para predecir condiciones estructurales críticas, ya que este modelo muestra valores de TPR alrededor de 6.8 (el valor más alto entre la predicción de todas las condiciones estructurales en ambos modelos) y 74 % de predicciones exitosas en las primeras 100 tuberías con más altas probabilidades de estar en condición crítica; y (ii) el modelo de deterioro basado en SVM a partir de su enfoque de clasificación es apropiado para predecir las otras condiciones estructurales, ya que este modelo muestra valores de PLR homogéneos para la predicción de todas las condiciones estructurales (entre 1.67 y 3.88) y las desviaciones entre lo observado y lo predicho son menores que aquellos resultados obtenidos del modelo SVM a partir de su enfoque de regresión.

**Palabras clave:** gestión de activos del alcantarillado, condición estructural, modelos de clasificación y regresión, máquinas de soporte vectorial (SVM).

## **Main Contributions**

- This work proposes a deterioration model based on SVM regression tasks to predict the structural condition of sewer assets. This model separates the structural prediction of sewer assets because it does not take into account the group of grades in the prediction. This method is developed on detail according to the scores of the sewer assets and the probability of belonging to a specific grade by using probability density functions.
- This paper compares two types of models based on SVM guided by two approaches (classification and regression), which made possible to explore the advantages and disadvantages of each model regarding to the management objectives of sewer asset management.
- The paper proposes an evaluation of model performance by means of two aims in order to fulfil two management objectives in sewer asset management
- Both deterioration models were applied to the case of Bogota with the purpose of supporting the decision-making process in sewer asset management

## **Introduction**

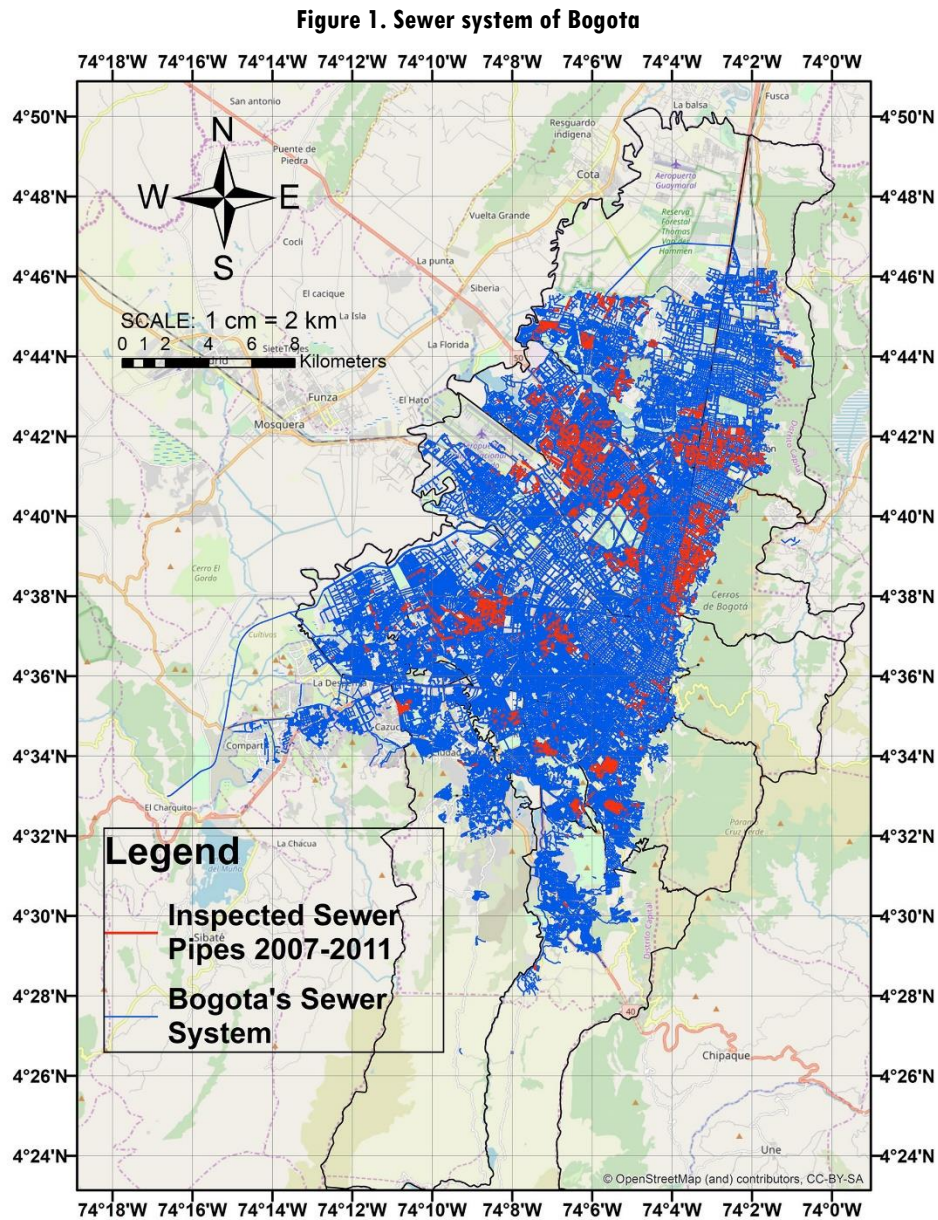
Urban drainage systems present alarming rates of aging and deterioration in both developed and developing countries [1], [2]. Currently, urban water system stakeholders are facing important challenges to achieve a rational, efficient, effective, and sustainable management and maintenance of this infrastructure while simultaneously considering the diversity of actors and constraints, such as budget limitations, environmental regulations, or urban water infrastructure benefits involved [3]–[6]. Deterioration models have been developed to support decision-making processes in sewer asset management. Statistical (i.e., logistic regression, survival models, Markov chain) [7], [4], [8] and machine learning (i.e., random forest, artificial neuronal networks, support vector machines) [9]–[12] methods have been the base for the development of these models. Particularly, some studies have shown that SVM are an effective method to carry out classification and regression tasks for prediction issues [13], [14]. In fact, deterioration models based on SVM have had successful predictions for case studies with low frequency of inspections [12], [15] using its classification task. Usually, SVM uses its classification approach in predicting the structural condition of sewer assets because most of the assessment standards classify the structural condition in a categorical variable (grade) [16]. For the case of Bogota, the mean annual inspection rate is less than 2 % [17] and the assessment standard used for qualifying the structural condition of the sewer system of Bogota rates the structural condition in two types of variables: numerical (scores) and categorical (grades) [18]. The deterioration models developed and applied in other studies usually are based on classification methods because of the classification in grades of structural condition [7], [4], [8]–[12], [15]. As the case of Bogota has a second rate,

as a numerical variable (scores), it sparks our interest to develop a deterioration model based on a regression approach using the scores, separating the aggrupation of structural condition from grades that could increase uncertainty in the prediction. Since deterioration models based on SVM have two prediction tasks and have had successful predictions with its classification task for case studies with low inspection rates, the authors propose a methodology based on the SVM's regression task to predict the structural condition of the sewer assets. The idea is exploring the advantages and disadvantages that an SVM model based on its regression (SVM\_R) approach could have compared to an SVM model based on its classification approach (SVM\_CL) to predict the structural condition of the sewer pipes. The evaluation of these predictions is based on two objectives: (i) Success in the prediction of all structural conditions for the management objectives at the network level (investment plans), and (ii) successful prediction in the critical conditions for the management objectives at the pipe level (rehabilitation plans). Therefore, this work suggests (i) a methodology, applying SVM from its regression approach (SVM\_R) to predict the structural condition of the uninspected pipes of the sewer system of Bogota and (ii) a comparison between the prediction results of the SVM\_R methodology and the prediction results using classification with SVM (SVM\_CL) for the same case study.

## **Materials and Methods**

### **Case Study: Sewer System of Bogota**

After data clean-up, the database of CCTV inspections performed in the sewer system of Bogota contains 5076 consistent inspections (representing 245 km) linked to 5031 sewer pipes during 2007 and 2011 (figure 1). The structural condition of the inspected pipes was evaluated using the local sewer assessment standard NS-058 [18]: To qualify the structural condition of each pipe scores are given to each structural failure found in a pipe and subsequently added up. These scores are ranked in 5 grades, from 1 to 5, 5 being the critical condition (total score  $\geq 165$  points). The original grades 3 and 4 were grouped in one because the number of pipes in these grades was 4 % and 13 %, respectively. Therefore, in this research, the structural condition was assumed to be one of four conditions (conditions 1 to 4).



Source: Own source.

### Support Vector Machines (SVM)

SVM is based on a supervised statistical learning method within the Kernel family of methods. This family consists of a class of algorithms for pattern analysis that finds and analyses general types of relations (e.g., clusters, ranges, principal component correlations, and classifications) in databases [19]. SVM is used to solve nonlinear classification problems using pattern recognition and function estimation. The principal problem addressed using SVM is the fit of a function describing a relation between an object X and a response Y. Initially, SVM was used for two-category classifications, where Y is the categorical vector;

however, there is an extension of the methodology that offers the possibility of regressions. SVM allows for classifications and regressions with parametric and nonparametric data [20]. The Kernlab library [21] was used with R software [22] to implement SVM in this research. Kernlab library proposes the kernel functions most typically used with SVM: Gaussian radial basis-RBF, Linear, Polynomial, Hyperbolic Tangent, Laplace, Bessel, and Anova [23]. After exploring these seven kernel functions (in both SVM approaches), Laplace Kernel function was chosen because of its better prediction of the structural conditions of the uninspected sewer pipes for both cases.

### **SVM Regression Approach Methodology**

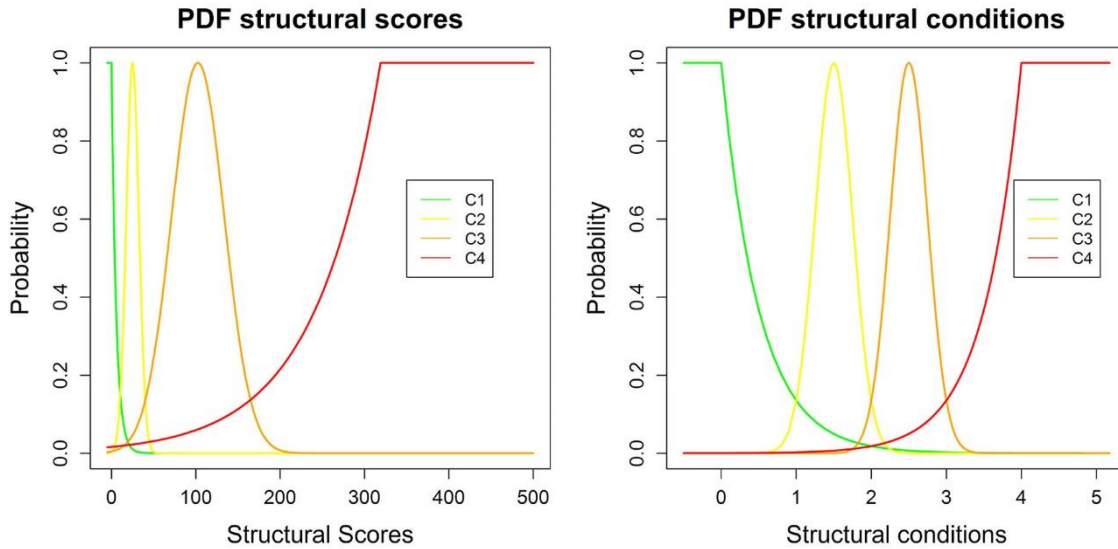
In order to apply the SVM regression approach, two thirds of the data were chosen randomly for calibration, and the rest was used for validation. The object X of the regression model contains the available variables that could influence the deterioration of sewer pipes [24]–[26], while the structural scores denoted its response Y. Therefore, the prediction results are represented by predicted structural scores.

For this case study, the variables included in the matrix of the object X in both SVM approaches were some sewer characteristics such as material (concrete, clay, and others), depth, length, slope, diameter, age, sewerage type (combined, wastewater, and stormwater), and road type (primary, intermediate, local, rural, and extension of a road). These variables were chosen according to previous studies where potential factors that could influence the structural condition of sewer pipes in the city of Bogota were identified [27]–[29].

Thanks to the fact that the structural condition of the sewer pipes in Bogota is assessed by both scores and grades [18], it is possible to observe two phenomena: Firstly, the mapping of score ranges to grades is not homogeneous (scores from 0 to 9 are mapped to grade 1, scores from 10 to 39 to grade 2, scores from 40 to 79 to grade 3, and so on); secondly, scores do not represent the same condition reliably despite being classified in the same particular condition. For example, for two pipes graded in excellent condition (grade 1) with a score of zero (0) for the first and ten (10) for the second one, it does not mean that both are at the same level of optimum condition, but it means that the first one does not have any structural failure and the second one has some superficial structural failures. For this reason, this methodology proposes to convert structural scores in probabilities using probability density functions (PDF). A PDF was assigned to each grade in order to describe the probability of being in this condition (figure 2, left): (i) for condition 1 ( $0 < \text{score} < 10$ ), a negative exponential distribution was assigned, because the lower the structural score, the higher the probability of being in excellent condition; (ii) for both conditions 2 ( $10 < \text{score} < 40$ ) and 3 ( $40 < \text{score} < 165$ ), symmetric normal or Gaussian distributions were assigned, because the probability of being in these conditions is higher at the midpoints (25 and 102.5 respectively) for these

ranks; (iii) for condition 4, a positive exponential distribution was assigned, because the higher the structural score is, the greater is the probability to be in critical condition. Then, the scores were normalized according to their grades or conditions (figure 2, right).

**Figure 2. PDFs proposed for different structural conditions (SVM\_R prediction)**



Source: Own source.

### Comparison between Regression and Classification SVM Prediction Results

The classification approach of SVM (SVM\_CL) was applied using the same data for calibration and validation as for SVM\_R. Also, SVM\_CL used the same available variables that were contained in object X for SVM\_R. However, the response Y was denoted by the structural conditions. In order to compare the prediction results of SVM\_CL with those of SVM\_R, the predicted grades or conditions (categorical variables) were converted to probabilities of being in a particular condition using dummy variables. If a pipe was predicted to be in critical condition, the probability of being in that condition would be 1, and the probabilities of being in other conditions would be 0.

Three techniques were used to analyze the prediction results from both SVM approaches: (i) ROC space [30]; (ii) Performance Curve [31]; (iii) deviation analysis [32]. The first one measures the True Positive Rate (TPR) versus False Positive Rate (FPR) of the predictions of the structural conditions. From these, the PLR (Positive Likelihood Ratio) is calculated, which quantifies how likely it is to have a positive prediction than a negative one (TPR/FPR). The second one identifies what percentage of pipes predicted in critical conditions was actually found to be in that condition (the percentage of the first pipes have the highest

probability of being in it). The last one analyses the deviation between the percentage of observed and predicted pipes of the four structural condition for each age period.

## Results

Tables 1 and 2 show the confusion matrices of the structural prediction in the validation data for SVM\_R and SVM\_CL approaches.

**Table 1. Confusion matrix, SVM\_R predictions**

<b>Observed Predicted</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>
C1	0	0	0	0
C2	494	141	78	42
C3	125	131	179	168
C4	14	19	46	148

Source: Own source.

**Table 2. Confusion matrix, SVM\_CL predictions**

<b>Observed Predicted</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>
C1	565	154	128	74
C2	20	30	12	7
C3	28	45	45	39
C4	44	51	113	230

Source: Own source.

As shown in the tables above, the SVM\_R approach (table 1) does not predict pipes in excellent condition (C1); it tends to predict pipes in a worse structural condition (C2 and C3). However, it does not predict many pipes in critical condition. The underestimation happens up to C2 and C3 conditions; SVM does not predict many pipes in critical condition (C4). On the other hand, SVM\_CL (table 2) tends to predict most pipes in excellent (C1) and critical conditions (C4).

According to the ROC space analysis (table 3), it was found in general terms that: (i) both SVM cases have PLR values (except C1 in SVM\_R) higher than one. This means that both methods have better predictions than a random one (True Positive Rate  $TPR \geq$  False positive Rate FPR), and (ii) the highest PLR value among the conditions is in the critical structural condition (C4) for both cases (being SVM\_R the highest). Nevertheless, SVM\_CL shows higher PLR values than SVM\_R for non-critical structural conditions (C1, C2, and C3). This



means that SVM\_CL gives more suitable predictions for the first three structural conditions whereas SVM\_R is suitable for the critical condition. When these results (table 3) are compared to the confusion matrices (tables 1 and 2), it is important to highlight that even SVM\_R predicted fewer pipes in critical condition than SVM\_CL; the SVM\_R predictions are more reliable than SVM\_CL for those pipes predicted in critical conditions.

**Table 3. ROC space results of SVM\_CL and SVM\_R approaches**

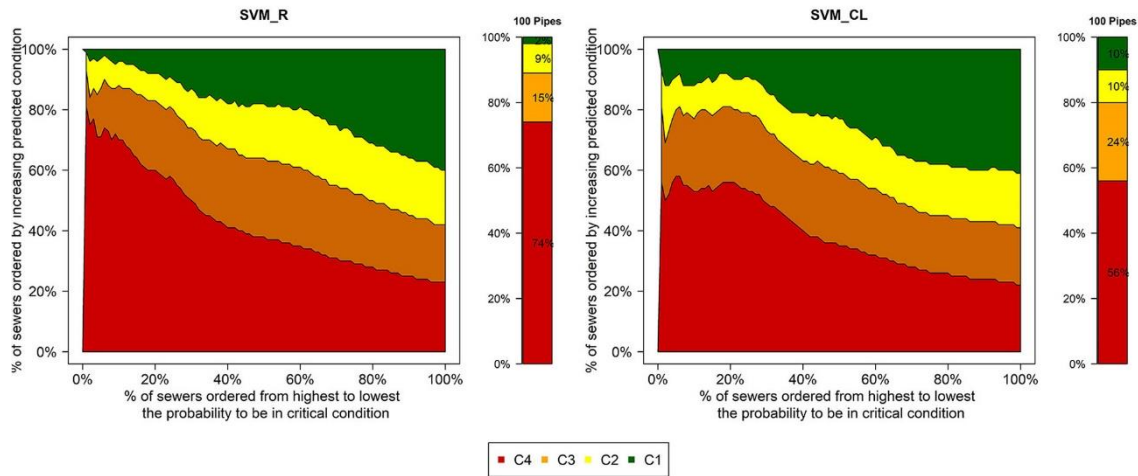
	SVM_R			SVM_CL		
	TPR	FPR	PLR	TPR	FPR	PLR
C1	0.00	0.00	NA	0.86	0.38	2.26
C2	0.48	0.47	1.02	0.11	0.03	3.67
C3	0.59	0.33	1.79	0.15	0.09	1.67
C4	0.41	0.06	6.83	0.66	0.17	3.88

Source: Own source.

The performance curves, presented in figure 3, show that (i) SVM\_R (figure 3, left) indicates a high percentage of success for pipes with high probability to be in critical condition (first percentages on the x-axis); likewise, it is possible to observe that the percentage of success decreases when the probability to be in critical condition also decreases; (ii) SVM\_CL (figure 3, right) shows two peaks of percentages of successful predictions located at 7 % and 22 % of the pipes with high probabilities to be in critical condition. This means that the pipes with a high probability to be in critical condition are not the ones with a high percent of success prediction.

The bar plots on the right side of each performance curve show a sample of the successful percentage for the first 100 pipes with high probabilities to be in critical condition (around 10 % of the validation data). The SVM\_R bar plot indicates that 74 % of pipes were correctly predicted to be in critical condition for the first 100 pipes with probability to be in that condition, while (ii) SVM\_CL correctly predicted 56 %. It means that, although both methods show better predictions than random, SVM\_R is more suitable to estimate and identify pipes in critical structural condition. Thanks to this analysis, the utility stakeholders of Bogota could use predictions of the SVM\_R model to design decision-making plans focusing on prioritizing the rehabilitation of the pipes in the most critical condition, taking into account the budget available for this activity. The model gives a success percentage of identifying pipes in critical condition according to the number (or percentage) of pipes that the utility stakeholders would be able to rehabilitate according to the assigned budget.

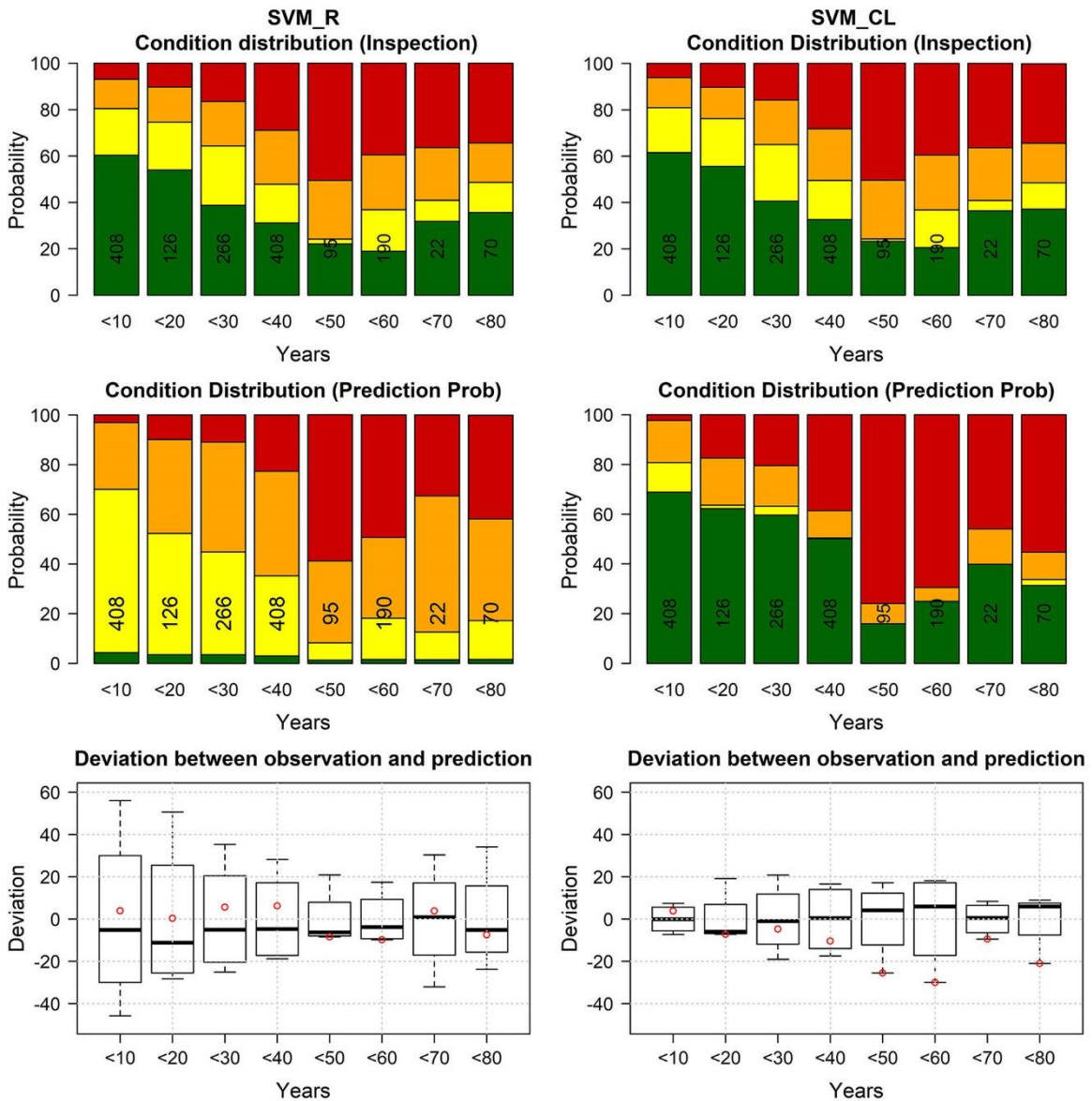
**Figure 3. Performance curves obtained for a sample of 100 pipes of SVM\_R (left) and SVM\_CL (right) prediction results**



Source: Own source.

Figure 4 shows the condition distributions with respect to the age of the pipes (periods of 10 years) comparing the observed and predicted databases (bar plots on the top and in the middle respectively, figure 3), as well as the deviation between these two databases (boxplots on the bottom) for both approaches (SVM\_R on the left, SVM\_CL on the right). Regarding the analysis of SVM\_R, it can be noticed that this approach tends to underestimate the structural conditions, in particular, pipes with excellent structural conditions (green strips, left) are predicted in acceptable and worse conditions (C2 and C3, yellow and orange strips, left), while SVM\_CL has an almost dichotomous prediction because it tends to predict the pipes either in excellent or critical conditions reducing the prediction of medium conditions (acceptable and worse conditions, yellow and orange strips respectively, right figures). These results confirm the analysis carried out for the results of the confusion matrices shown in tables 1 and 2.

Figure 4. Deviation analysis of SVM\_R (left) and SVM\_CL (right) prediction results vs inspection data results



Source: Own source.

According to the deviation figures (on the bottom of figure 4), it was found that (i) the boxplots for SVM\_R show a lower deviation ( $\pm 20\%$ ) for those pipes with ages between 50 and 60 years (on the bottom figure 3, left), while the boxplots for SVM\_CL show a lower deviation ( $\pm 20\%$ ) for newer pipes (< 40 years) and older pipes (70 and 80 years) (on the bottom figure 3, right); (ii) the deviation of critical structural condition (red circles in the deviation figures) is lower for SVM\_R ( $\pm 10\%$ ) in all pipes compared to SVM\_CL (more than 10% in pipes of 50, 60 years, and 80 years).

## **Conclusions**

This work aimed to propose a methodology using the regression approach of the Support Vector Machine (SVM) method in order to predict the structural condition of uninspected sewer pipes as a tool to support sewer asset management. Likewise, the prediction results of this methodology (SVM\_R) were compared to prediction results using the classification approach of SVM (SVM\_CL).

Furthermore, these approaches were accompanied by two different ways of converting the prediction results (scores or conditions) into probabilities of being in a particular condition. SVM\_R used probability distribution functions (PDF) and SVM\_CL used dummy variables. The integration between the SVM approaches and the converting ways indicate that (i) SVM\_R is a suitable method to predict the critical structural prediction since, according to the performance curve analysis, the higher the probability of being in critical condition, the higher the probability to predict this condition successfully, while this was not observed in SVM\_CL results; and (ii) SVM\_CL is a suitable method for the prediction of non-critical structural condition since it has a higher probability that these prediction results will be correct. The SVM classification approach has higher positive likelihood ratios (between 1.7 and 3.7) than the SVM regression approach (1.8 as maximum value).

The conclusion above was confirmed by the deviation analysis, which shows that SVM\_CL has lower deviations (+/- 20 %) when it analyses the four conditions in boxplots, while SVM\_R has lower deviations (+/-10 %) only for those pipes predicted in critical conditions. As a consequence of these results, the authors suggest that depending on the goals of the stakeholders both SVM approaches could be used with confident results. Applied to the case of Bogota, when the goal is to have a global overview of the structural state of the sewer system, the classification methodology of SVM could be a useful tool; conversely, when the goal is a high reliability in the estimation of the critical structural condition for purposes of prioritizing the management of pipes in critical conditions, the regression SVM approach methodology is suggested for this aim.

## **References**

- [1] B. C. Ferguson, R. R. Brown, and A. Deletic, "Diagnosing transformative change in urban water systems: Theories and frameworks," *Global Environ. Change*, vol. 23, no. 1, pp. 264–280, 2013. doi: 10.1016/j.gloenvcha.2012.07.008
- [2] H. Osman, "Agent-based simulation of urban infrastructure asset management activities," *Automat. Const.*, vol. 28, pp. 45–57, 2012. doi: 10.1016/j.autcon.2012.06.004

- [3] N. Caradot et al., “The benefits of deterioration modelling to support sewer asset management strategies,” presented at the 8th Int. Conf. *Sewer Process Networ.*, Rotterdam, The Netherlands, August 31–September 2, 2016.
- [4] H. S. Baik, H. S. Jeong, and D. M. Abraham, “Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems,” *J. Water Resour. Plan. Manag.*, vol. 132, no. 1, pp. 15–24, 2006. doi: 10.1061/(ASCE)0733-9496(2006)132:1(15)
- [5] M. A. Cardoso, M. S. Silva, S. T. Coelho, M. C. Almeida, and D. I. C. Covas, “Urban water infrastructure asset management: A structured approach in four water utilities,” *Water Sci. Technol.*, vol. 66, no. 12, pp. 2702–2711, 2012. doi: 10.2166/wst.2012.509
- [6] R. Younis and M. A. Knight, “Development and implementation of an asset management framework for wastewater collection networks,” *Tunnelling Underground Space Technol.*, vol. 39, pp. 130–143, 2014. doi: 10.1016/j.tust.2012.09.007
- [7] L. T. Wright, J. P. Heaney, and S. Dent, “Prioritizing sanitary sewers for rehabilitation using least-cost classifiers,” *J. Infrastructure Syst.*, vol. 12, no. 3, pp. 174–183, 2006. doi: 10.1061/(ASCE)1076-0342(2006)12:3(174)
- [8] Y. le Gat, “Modelling the deterioration process of drainage pipelines,” *Urban Water J.*, vol. 5, no. 2, pp. 97–106, 2008. doi: 10.1080/15730620801939398
- [9] M. M. Rokstad and R. M. Ugarelli, “Evaluating the role of deterioration models for condition assessment of sewers,” *J. Hydroinform.*, vol. 17, no. 5, pp. 789–804, 2015. doi: 10.2166/hydro.2015.122
- [10] D. H. Tran, A. W. M. Ng, and B. J. C. Perera, “Neural networks deterioration models for serviceability condition of buried stormwater pipes,” *Eng. Appl. Artificial Intelligence*, vol. 20, no. 8, pp. 1144–1151, 2007. doi: 10.1016/j.engappai.2007.02.005
- [11] D. H. Tran, A. W. M. Ng, B. J. C. Perera, S. Burn, and P. Davis, “Application of probabilistic neural networks in modelling structural deterioration of stormwater pipes,” *Urban Water J.*, vol. 3, no. 3, pp. 175–184, 2006. Available: <https://doi.org/10.1080/15730620600961684>
- [12] J. Mashford, D. Marlow, D. Tran, and R. May, “Prediction of sewer condition grade using support vector machines,” *J. Comput. Civil Eng.*, vol. 25, no. 4, pp. 283–290, 2010. doi: 10.1061/(ASCE)CP.1943-5487.0000089
- [13] K. Roushangar and R. Ghasempour, “Prediction of non-cohesive sediment transport in circular channels in deposition and limit of deposition states using SVM,” *Water Sci. Technol. Water Supply*, vol. 17, no. 2, pp. 537–551, 2017. doi: 10.2166/ws.2016.153
- [14] F. Granata, S. Papirio, G. Esposito, R. Gargano, and G. de Marinis, “Machine learning algorithms for the forecasting of wastewater quality indicators,” *Water*, vol. 9, no. 2, p. 105, 2017. doi: 10.3390/w9020105
- [15] N. Hernández, N. Caradot, H. Sonnerberg, P. Rouault, and A. Torres, “Support tools to predict the critical structural condition of uninspected sewer pipes in Bogota D. C.,” presented at *LESAM Water Wastewater Infrastructure Conf.*, Trondheim, Norway, 20–22 June, 2017.
- [16] R. Thornhill and P. Wildbore, “Sewer defect codes: Origin and destination,” *Underground Construction*, Apr. 2005. Available: <https://www.thefreelibrary.com/Sewer+defect+codes%3a+origin+and+destination.-a0132051614>
- [17] Gerencia Corporativa de Planeamiento y Control, “Informe del Balance General del PDD 2012-2016, ‘Bogotá Humana’”, Empresa de Acueducto y Alcantarillado de Bogotá (EAAB), Bogotá, Colombia, 2016. Available: [https://www.acueducto.com.co/wps/html/resources/empresa/2015/24\\_12\\_15\\_atenCiudadano/Informe\\_segplan\\_inv\\_y\\_gestion\\_corte\\_sept%2030\\_de\\_2015.pdf](https://www.acueducto.com.co/wps/html/resources/empresa/2015/24_12_15_atenCiudadano/Informe_segplan_inv_y_gestion_corte_sept%2030_de_2015.pdf)
- [18] Empresa de Acueducto y Alcantarillado de Bogotá (EAAB), “NS-058. Aspectos Técnicos para inspección y mantenimiento de redes y estructuras de alcantarillado”, Bogotá, Colombia, EAAB-E. S. P., 2001.

- [19] J. Shawe-Taylor and N. Cristianini, Kernel methods for pattern analysis. Cambridge: Cambridge University Press, 2004.
- [20] L. López-Kleine and A. Torres, “UV-vis in situ spectrometry data mining through linear and nonlinear analysis methods,” *Dyna*, vol. 81, no. 185, pp. 182–188, 2014. doi: 10.15446/dyna.v81n185.37718
- [21] A. Karatzoglou, A. Smola, K. Hornik, and A. Zeileis, “Kernlab-an S4 package for kernel methods in R,” *J. Stat. Softw.*, vol. 11, no. 9, pp.1–20, 2004. doi: 10.18637/jss.v011.i09
- [22] R. Core Team. (2019) [Online]. *R: A language and environment for statistical computing*,” *R Foundation for Statistical Computing*. Available: <https://www.gbif.org/es/tool/81287/r-a-language-and-environment-for-statistical-computing>
- [23] M. G. Genton, “Classes of kernels for machine learning: A statistics perspective,” *J. Machine Learning Res.*, vol. 2, pp. 299–312, Dec. 2001. Available: <http://www.jmlr.org/papers/volume2/genton01a/genton01a.pdf>
- [24] J. P. Davies, B. A. Clarke, J. T. Whiter, and R. J. Cunningham, “Factors influencing the structural deterioration and collapse of rigid sewer pipes,” *Urban Water*, vol. 3, nos. 1-2, pp.73–89, 2001. Available: [https://doi.org/10.1016/S1462-0758\(01\)00017-6](https://doi.org/10.1016/S1462-0758(01)00017-6)
- [25] R. Baur and R. Herz, “Selective inspection planning with ageing forecast for sewer types,” *Water Sci. Technol.*, vol. 46, nos. 6-7, pp. 389–396, 2002. doi: 10.2166/wst.2002.0704
- [26] F. Chughtai and T. Zayed, “Integrating WRc and CERIU condition assessment models and classification protocols for sewer pipelines,” *J. Infrastructure Syst.*, vol. 17, no. 3, pp. 129–136, 2011. DOI: 10.1061/(ASCE)IS.1943-555X.0000052
- [27] P. Niño, H. Angarita, D. Vargas, and A. Torres, “Identificación factores de riesgo para la gestión patrimonial óptima de sistemas de drenaje urbano: estudio piloto en la ciudad de Bogotá,” presented at *XXV Congreso Latinoamericano de Hidráulica*, San José, Costa Rica, september 9–12, 2012.
- [28] N. Hernandez, N. Obregón, and A. Torres, “Identificación de factores de tipo categórico relacionados con la condición estructural de tuberías de alcantarillado de Bogotá a partir de conceptos de entropía de la información,” *Ingenier. Solidaria*, vol. 12, no. 19, pp. 63–71, 2016. doi: 10.16925/in.v12i19.1194
- [29] L. López-Kleine, N. Hernández, and A. Torres, “Physical characteristics of pipes as indicators of structural state for decision-making considerations in sewer asset management,” *Ingenier. Invest.*, vol. 36, no. 3, pp. 15–21, 2016.
- [30] C. D. Brown and H. T. Davis, “Receiver operating characteristics curves and related decision measures: A tutorial,” *Chemometrics Intelligent Lab. Syst.*, vol. 80, no. 1, pp. 24–38, 2006. doi: 10.1016/j.chemolab.2005.05.004
- [31] A. Kästner, I. Kropp, M. Schöne, and N. Caradot, “D2.1 Sewer deterioration modelling, in Braunschweig Dissemination approved by technical committee members: I. Charron, J. M. Brossard, Veolia Environnement Recherche & Innovation (VERI) V. Perez, Veolia Water, Technical Direction C. Sardet, OEWA Wasser und Abwasser GmbH (OEWA) P. Chudoba, Veolia Water Europe E. Eckert, J. Wachnewski, Berliner Wasserbetriebe (BWB) C. Mesek, Stadtentwässerung Braunschweig (SE|BS) A. Hartmann, Kompetenzzentrum Wasser Berlin gGmbH (KWB)”, Kompetenzzentrum Wasser Berlin gGmbH, 2015.
- [32] N. Caradot et al., “The benefits of deterioration modelling to support sewer asset management strategies,” presented at *8th Int. Conf. Sewer Process Networ.*, August 31–September 2, 2016, Rotterdam, The Netherlands.