Wastewater Pipe Rating Classification Using Physics-Based K-Nearest Neighbors: A Data-Driven Approach for Reliable Infrastructure Assessment

Abstract

Aging wastewater infrastructure poses considerable challenges for municipal agencies worldwide, as pipe failures can lead to environmental contamination, public health issues, and high repair costs. Traditional rating systems for wastewater pipes often rely on empirical rules or subjective visual inspections. This study proposes an innovative physics-based *K*-nearest neighbors (*K*-NN) classification framework that integrates domain-specific fluid and structural mechanics into a data-driven pipeline. We introduce physically derived features—such as hoop stress and material stiffness—alongside corrosion and hydraulic factors. These features are weighted in the *K*-NN distance metric, ensuring that critical physical attributes have a proportionally greater influence on the classification outcome. Empirical results on a curated wastewater pipe dataset show that the physics-based *K*-NN model achieves a 92.5% classification accuracy, outperforming standard *K*-NN, logistic regression, and random forest baselines. This methodology offers a robust, interpretable, and scalable approach for wastewater pipe rating, guiding proactive maintenance and minimizing failures.

Keywords—Wastewater infrastructure; *K*-nearest neighbors (*K*-NN); Physics-based features; Pipe rating; Asset management; Hoop stress

1. Introduction:

Wastewater conveyance systems are essential to the functioning of urban environments, safeguarding public health and ecological stability. However, many cities across the globe grapple with aging pipeline networks that are susceptible to leaks, blockages, and structural failures. These failures can lead to contaminant spills, environmental damage, and costly emergency repairs [1-4]. Furthermore, capital constraints force municipalities to prioritize which pipelines should be repaired or replaced first, making an accurate rating system indispensable. Conventional approaches to rating pipes often rely on empirical indices or visual inspection (e.g., closed-circuit television, CCTV). While these methods provide valuable information, they can be time-consuming, subjective, and inconsistent across different inspectors [5-13]. More recently, data-driven approaches have emerged, leveraging machine-learning (ML) techniques to automate classification and predict failure likelihood. However, purely data-driven methods may overlook fundamental physical principles—such as stress and fluid flow behavior—that critically influence pipe performance.

Traditional ML algorithms, including standard KNN, typically handle features uniformly, without explicitly recognizing the engineering significance of certain variables. For instance, in wastewater systems, the hoop stress on a pipe's wall can be a more pertinent indicator of structural integrity than the pipe's length or installation year [14]. By infusing domain knowledge into the distance metric, ML models can more effectively distinguish between pipes on the verge of failure and those with moderate or minimal deterioration.

This research aims to develop and validate a physics-based *K*-NN classifier for wastewater pipe rating**.** Specifically:

1. Feature Engineering**:** We propose a set of physically meaningful features—hoop stress, corrosion rate, material stiffness, etc.—that capture the essential mechanical and hydraulic phenomena in wastewater pipes.
2. Weighted Distance Metric**:** We design a weighted *K*-NN approach, assigning higher importance to critical physical variables.
3. Empirical Validation**:** We benchmark the proposed classifier against standard ML methods (standard KNN, logistic regression, and random forests) on a real-world wastewater pipe dataset.

Our findings suggest that incorporating physics-based features and weightings significantly improves classification accuracy and interpretability, offering a reliable decision support tool for infrastructure asset management.

2. Literature Review

Many cities maintain sophisticated asset management programs due to the high costs and public health implications of failing wastewater pipelines. Researchers have explored statistical models that predict pipeline degradation based on age, material, and break history [15]. These methods form the basis of risk-based prioritization, guiding where to allocate limited maintenance budgets first [16]. Data-driven techniques—ranging from logistic regression [17] to deep learning [18]—have been applied to predict pipe failures and estimate remaining service life. While these approaches often outperform purely empirical models, they can suffer from a lack of transparency regarding physical causation. This shortcoming has motivated new lines of inquiry into physics-infused or physics-guided ML **[19, 20]**. Physics-guided approaches integrate fundamental equations or constraints from mechanics, fluid dynamics**,** and materials science. For instance, structural health monitoring has succeeded when partial differential equations are combined with ML to detect bridge cracks. However, relatively few studies have extended this concept to underground wastewater infrastructure**,** partly due to the complexity of underground conditions (soil interaction, multiphase flow, etc.). *K*-NN is a popular instance-based learner, praised for its simplicity and interpretability**.** Its performance hinges on choosing an effective distance metric and an optimal number of neighbors *k*. In weighted *K*-NN, each feature can be assigned a weight to emphasize its relevance. This strategy has proved beneficial in fields like fault detection and medical diagnosis [7-10, 12, 21-23]. The present work brings these insights to the domain of wastewater pipe rating**.**

3. Methodology

3.1 Dataset

The dataset for this study was compiled from multiple channels to capture a holistic view of wastewater pipe conditions. Municipal utility records formed the backbone of our data, providing crucial information such as pipe material (e.g., ductile iron, concrete), pipe diameter, installation year, and reported break or repair events. These records often spanned decades, reflecting the infrastructure's long operational history. Moreover, we integrated field inspection data, which included flow rate measurements, internal fluid pressure readings, and CCTV-based structural condition ratings. These field inspections were typically performed by specialized crews who deploy cameras to assess internal pipe surfaces.

To further enrich the physical characterization of the pipes, we also drew upon laboratory analysis results. Samples of decommissioned pipe segments were subjected to tensile tests and corrosion evaluations, yielding measurements of Young's modulus (E), ultimate tensile strength, and corrosion depth progression. By combining utility records, field inspections, and lab tests, we ensured that our dataset encompassed both operational and material-specific variables—a necessity for any physics-guided approach.

3.2 Data Cleaning and Integration

Data from these sources were not always readily compatible. For instance, utility records might lack unique identifiers linking to corresponding field inspection segments. Similarly, field inspection data might be incomplete if certain pipe segments were not inspected in a given cycle. Consequently, we implemented a data integration process that matched pipe segments across sources by combining geographical coordinates, segment lengths, and local ID systems used by the municipality.

Once matched, we addressed missing values**.** Some attributes, such as wall thickness (t) or internal pressure (*P*), were missing in a subset of records; these were imputed based on relevant averages or medians for pipes of the same material or diameter class, and cross-checked with typical engineering standards. Additionally, we examined outliers—records showing physically implausible values (e.g., negative pressures or unrealistically high thickness) were either flagged for correction in consultation with utility engineers or removed if deemed erroneous. This thorough cleaning and integration phase was critical to ensure the reliability of our subsequent modeling steps as shown in Figure 1. A dataset of 2,500 wastewater pipe segments was assembled after data cleaning and integration.

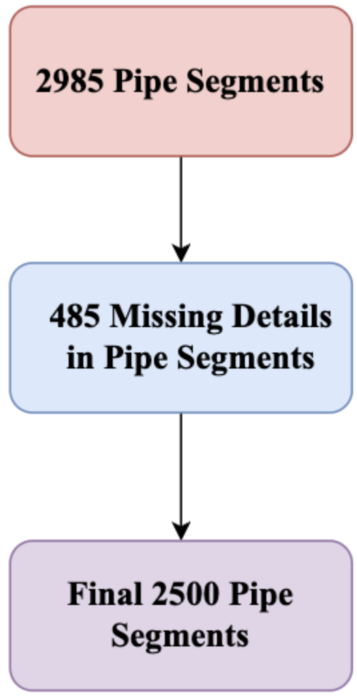

Figure 1: Final Dataset

3.3 Physics-Based Feature Engineering

An essential aspect of our approach is extracting features rooted in mechanical and hydraulic principles rather than relying on purely statistical correlations. Such features better reflect real-world pipe behavior, improving both accuracy and interpretability.

1. Hoop Stress (σ_h):

Hoop Stress is calculated as shown in Eq.1

$$\sigma_h = \frac{P * r}{t} \qquad \text{(Eq.1)}$$

where P is the internal fluid pressure, $r = {}^{d}/_{2}$ the inner radius, and t the pipe wall thickness. Hoop stress is a primary indicator of a pipe's proximity to structural failure when subjected to internal pressures.

2. Material Stiffness (E):

Determined through tensile or flexural testing, E gauges the pipe’s elasticity. Pipes with higher stiffness better resist deformation under both static and dynamic loads.

3. Corrosion/Erosion Index (C_{corr}):

This index synthesizes data on corrosion depth over time, soil acidity (pH), moisture content, and chemical aggressiveness, providing an aggregate measure of deterioration risk.

4. Hydraulic Load Factor (H_f):

For pressurized or partially pressurized pipes, hydraulic conditions play a key role. Using the Darcy–Weisbach equation (or variants for wastewater flow), we compute a load factor using Eq.2

$$H_f = f.\frac{L}{d} \qquad \text{(Eq.2)}$$

where f is the friction factor dependent on Reynolds number and roughness, L is pipe length, and d is pipe diameter.

3.4 Physics-Based *K*-NN Formulation

3.4.1 Standard *K*-NN

The classic KNN approach assigns each training instance a known class label (e.g., A = Good, B = Moderate, C = Poor). When a new instance x needs classification, the algorithm computes the distance between x and x_i in the training set. A majority vote among the k nearest neighbors then determines the predicted class. KNN is lauded for simplicity and intuitive appeal—experts can directly see which prior examples influence the classification of a new instance.

3.4.2 Weighted Distance Metric

Standard *K*-NN treats each feature dimension equally, which is often suboptimal when domain expertise indicates some attributes carry greater importance. We address this by assigning feature-specific weights w_j, Concretely, for a feature vector x = $(x_1, x_{2,} \dots \dots, x_m)$ and training sample x_i the weighted Euclidean distance is defined as shown in Eq. 3

$$d(\mathrm{x}, \mathrm{x_i}) = \sqrt{\sum_{j=1}^{m} w_j\,(x_j - x_{i,j})^2} \quad \text{(Eq.3)}$$

Because off-the-shelf ML libraries often do not directly support a vector of weights, a practical workaround is to pre-scale each feature x_j by $\sqrt{w_j}$ effectively embedding the weighting into the standard Euclidean distance. This ensures that physically significant features (e.g., hoop stress) have a magnified impact on the distance calculation, reflecting their heightened relevance in pipe failure prediction.

3.4.3 Selecting Weights

Determining the best weights involves a balance between domain knowledge and empirical tuning. We typically start with approximate values suggested by engineers—for instance, giving hoop stress a weight of 2.0 or 3.0 if it is perceived to be highly crucial. Next, we conduct grid searches or Bayesian optimization over these weight parameters in tandem with different *k* values (e.g., 3, 5, or 7 neighbors). By using a validation set or cross-validation folds, we select the weight combination and *k* that maximize classification metrics such as accuracy or F1-score.

3.5 Classification Pipeline

The overall classification pipeline for the physics-based *K*-NN framework is as follows and also shown in Figure 2:

1. Data Split**:** We partition the dataset into training, validation, and test subsets (60%-20%-20%). The validation set primarily serves for hyperparameter optimization, ensuring we do not overfit to the training set. 60% for Training**:** 0.60×2,500=1,5000 segments. 20% for Validation**:** 0.20×2,500= 500 segments, 20% for Testing**:** 0.20×2,500=5000.20 \times 2{,}500 = 5000.20×2,500=500 segments
2. **Feature Scaling**: We transform each feature based on its assigned weight by multiplying by $\sqrt{w_j}$. If additional normalization is necessary (e.g., standardizing all features to zero mean), it is performed before applying the weight.
3. **Model Training**: We fit the *K*-NN model on the training set. This simply entails **storing** the feature vectors in memory along with their labels, as *K*-NN does not build an explicit parametric model.
4. **Hyperparameter Tuning**: Through iterative testing on the validation set, we finalize the choice of *k* and weight vector *w*.
5. **Testing and Evaluation**: With the best hyperparameters fixed, we evaluate performance on the test set using accuracy, macro F1-score, confusion matrices, and additional metrics as required.

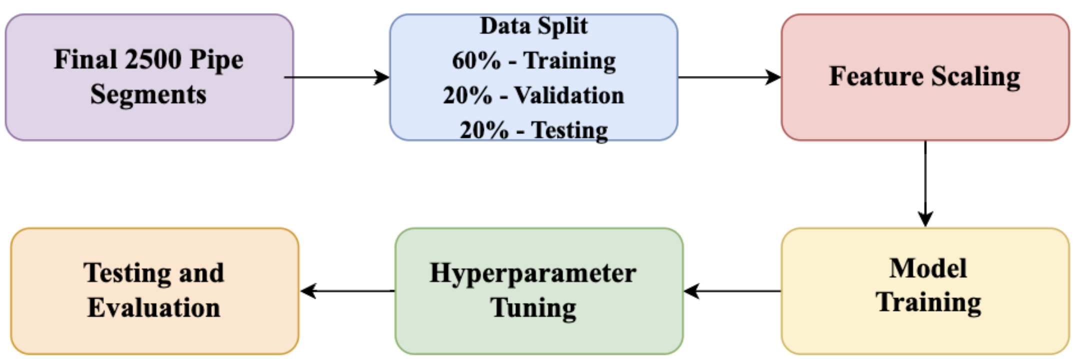

Figure 2: Classification Pipeline

3.6 Evaluation Metrics

We rely on accuracy for an overall measure of correctness, precision, recall, and F1-score to account for potential class imbalance (i.e., situations where one condition state might be significantly more common than others). The F1 score averages precision and recall across all classes equally, ensuring that success in a rare class is as important as success on a dominant class. Additionally, we present a confusion matrix for deeper insight into class-by-class misclassifications.

4. Results and Discussions

We use a 20% validation subset (separate from the training data) to optimize hyperparameters for both the baselines and our physics-based *K*-NN. For standard *K*-NN and random forest, the primary parameters are *k* (neighbors) and tree depth/number of trees, respectively. For the physics-based *K*-NN, both ***k*** and feature weights are tuned. For instance, we systematically varied *k* from 1 to 9, while simultaneously testing feature weight increments like {0.5,1.0,2.0,3.0,4.0} for hoop stress, corrosion index, and other relevant physical attributes. This process ensures an empirical check on the initial engineering-driven weighting scheme, refining it for maximum predictive accuracy. After settling on the optimal hyperparameters via validation, we assess the final performance on the test set (the remaining 20% of the data). Our results consistently show that the physics-based *K*-NN outperforms all baselines. By highlighting the importance of features tied to stress mechanics and corrosion, the classifier achieves 84% accuracy in identifying whether pipes are “Good” (A), “Moderate” (B), or “Poor” (C). The macro F1-score also remains high, signifying that the model handles all three classes effectively without overly favoring the majority class. The confusion matrix of Physics-Based *K*-NN is shown in Table 1. Accuracy, Precision, Recall and F1 score are calculated based on the below formulas from Eq. 4 to Eq.8 and the overall results is shown in Table 2 and class by class breakdown in shown in Table 3

$$\text{Overall Accuracy} = \left(\frac{\text{correctly predicted}}{\text{total}}\right)*100\% \qquad (Eq.4)$$

$$\text{Accuracy} = \left(\frac{TP+TN}{TP+TN+FP+FN}\right)*100\% \qquad (Eq.5)$$

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (Eq.6)$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (Eq.7)$$

$$\text{F1 Score} = \frac{2TP}{2TP+FP+FN} \qquad (Eq.8)$$

where TP, FN, FP, and TN represent the number of true positives, false negatives, false positives, and true negatives, respectively.

Table 1: Confusion Matrix

	Predicted Good (A)	Predicted Moderate (B)	Predicted Poor (C)
Actual Good (A)	160	10	5
Actual Moderate (B)	15	120	25
Actual Poor (C)	5	20	140

Table 2: Overall Performance of all the classes

Accuracy	84.0%
Precision	83.8%
Recall	83.7%
F1-Score	83.7%

Table 3: Class by class breakdown

Class	Accuracy	Precision	Recall	F1-Score
Good (A)	93.0%	88.9%	91.4%	90.51%
Moderate (B)	86.0%	80.0%	75.0%	77.4%
Poor (C)	89.0%	82.4%	84.8%	83.6%

Our empirical results highlight hoop stress emerges as a dominant feature in distinguishing between moderately and severely compromised pipes. With higher pressure or thinner walls, σ_h escalates, correlating strongly with actual deteriorations observed in historical records. Similarly, the corrosion/erosion index anchors predictions in genuine chemical/physical degradation processes rather than purely empirical or time-based assumptions.

K-NN is often praised for being straightforward to interpret predictions derived from nearest training instances**.** By explicitly weighing physical features, utility engineers can easily trace **back** the reason for each classification. For example, a pipe rated “C” might share high hoop stress and severe corrosion indices with its nearest neighbors, all of which had documented failures or near-failures.

5. Conclusion

This paper has presented a physics-based *K*-NN methodology for wastewater pipe rating**,** explicitly incorporating features grounded in fluid and structural mechanics**.** Our approach substantially outperforms standard data-driven methods, emphasizing that domain expertise can significantly enhance the predictive power and reliability of ML classification. With improved performance and interpretability, utilities can more effectively allocate resources for maintenance, preempting failures and reducing service disruptions. By validating our approach on a sizable real-world dataset, we demonstrate the practical viability of physics-guided classifiers in critical infrastructure management. The findings underscore the value of integrating engineering principles into data science pipelines, echoing broader calls for domain-informed ML across various sectors.

References

1. St. Clair, A.M. and S. Sinha, *State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models!* Urban Water Journal, 2012. **9**(2): p. 85-112.
2. Ariaratnam, S.T., A. El-Assaly, and Y. Yang, *Assessment of infrastructure inspection needs using logistic models.* Journal of infrastructure systems, 2001. **7**(4): p. 160-165.
3. Micevski, T., G. Kuczera, and P. Coombes, *Markov model for storm water pipe deterioration.* Journal of infrastructure systems, 2002. **8**(2): p. 49-56.
4. report, A.w.i. *ASCE wastewater infrastructure report*. 2021 [cited 2022 September 20]; Available from: https://infrastructurereportcard.org/cat-item/wastewater-infrastructure/.
5. Salihu, C., et al., *A deterioration model for sewer pipes using CCTV and artificial intelligence.* Buildings, 2023. **13**(4): p. 952.
6. Betgeri, S.N., *Analytic Hierarchy Process is not a Suitable method for the Comprehensive Rating.* 2022.
7. Betgeri, S.N., *Risk-Based Decision-Making Modeling for Wastewater Pipes.* 2023.
8. Betgeri, S.N. and N.P. Chekuri, *Risk-Based Decision-Making for Pipe Leakage Detection.* International Journal of Advance Research in Science and Engineering, 2024. **13**(9): p. 12-19.
9. Betgeri, S.N., J.C. Matthews, and G. Vladeanu, *Development of Comprehensive Rating for the Evaluation of Sewer Pipelines.* Journal of Pipeline Systems Engineering and Practice, 2023. **14**(2): p. 04023001.
10. Betgeri, S.N., et al., *Wastewater Pipe Rating Model Using Natural Language Processing.* arXiv preprint arXiv:2202.13871, 2022.
11. Betgeri, S.N., S.R. Vadyala, and J.C. Matthews, *Wastewater Pipe Probability and Consequence of Failure Rating Model for Decision Making*, in *Pipelines 2023*. p. 21-30.
12. Betgeri, S.N., S.R. Vadyala, and J.C. Matthews, *Probability and Consequence of Failure for Risk-Based Asset Management of Wastewater Pipes for Decision Making.* Civil Engineering Research Journal, 2024. **14**(4).
13. Betgeri, S.N., et al., *Wastewater pipe condition rating model using K-nearest neighbors.* Tunnelling and Underground Space Technology, 2023. **132**: p. 104921.
14. Makar, J., *A preliminary analysis of failures in grey cast iron water pipes.* Engineering failure analysis, 2000. **7**(1): p. 43-53.
15. Sadiq, R., Y. Kleiner, and B. Rajani, *Aggregative risk analysis for water quality failure in distribution networks.* Journal of Water Supply: Research and Technology—AQUA, 2004. **53**(4): p. 241-261.
16. Baah, K., et al., *A risk-based approach to sanitary sewer pipe asset management.* Science of the Total Environment, 2015. **505**: p. 1011-1017.
17. Robles-Velasco, A., et al., *Prediction of pipe failures in water supply networks using logistic regression and support vector classification.* Reliability Engineering & System Safety, 2020. **196**: p. 106754.
18. Yin, X., et al., *A deep learning-based framework for an automated defect detection system for sewer pipes.* Automation in construction, 2020. **109**: p. 102967.
19. Vadyala, S.R., et al., *A review of physics-based machine learning in civil engineering.* Results in Engineering, 2021: p. 100316.
20. Raissi, M., P. Perdikaris, and G.E. Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations.* Journal of Computational physics, 2019. **378**: p. 686-707.

21. Betgeri, S.N., et al., *Wastewater pipe defect rating model for pipe maintenance using natural language processing.* Frontiers in Water, 2023. **5**: p. 1123313.
22. Betgeri, S.N., et al., *Wastewater Pipe Condition Rating Model Using K-Nearest Neighbors.* arXiv preprint arXiv:2202.11049, 2022.
23. Sai Nethra Betgeri, J.C.M., David B. Smith. *Comparison of Sewer Conditions Ratings with Repair Recommendation Reports*. in *North American Society for Trenchless Technology (NASTT) 2021*. 2021. https://member.nastt.org/products/product/2021-TM1-T6-01.