Integrating dynamic features into machine learning models for predicting sewer network failures: a Random Forest approach

Oussama Aounali, Will Shepherd & Simon Tait

View supplementary material

Published online: 21 Nov 2025.

Submit your article to this journal

Article views: 71

View related articles

View Crossmark data

RESEARCH ARTICLE

Integrating dynamic features into machine learning models for predicting sewer network failures: a Random Forest approach

Oussama Aounali [ID], Will Shepherd [ID] and Simon Tait [ID]

School of Mechanical, Aerospace and Civil Engineering, University of Sheffield, Sheffield, UK

ABSTRACT

Sewer blockages and flooding remain persistent challenges. Water utilities often deploy labour-intensive, iterative approaches, such as repeated CCTV inspections and jetting, to detect and address these problems. Recently, machine learning (ML) based asset failure prediction has emerged as a cost-effective alternative, enabling proactive identification of vulnerable pipe sections that can then be the focus for inspection and intervention. Early ML-based predictive models primarily focused on non-dynamic factors, such as physical pipe attributes, while newer approaches have incorporated dynamic variables like rainfall and pipe flow, resulting in significant accuracy improvements. This study examines the integration of sediment transport mechanics, using network hydraulic model derived parameters, into a predictive Random Forest (RF) model. Results demonstrate that incorporating dynamic features representing sediment transport capacity and its spatial variation considerably enhances the RF model's predictive power, offering a more reliable tool for identifying and managing blockages, and flooding in combined sewer networks.

Introduction

Addressing sewer failures, such as localised flooding or structural failure has traditionally been a reactive process, focusing on interventions after failures occurs. This approach is not only costly but also ineffective in preventing environmental damage and service disruptions, which can result in significant costs for water utilities and adverse effects on public health and the environment (Draude et al. 2019; Rosin et al. 2022). As urbanisation and climate change exacerbate these challenges, there is an urgent need to develop more effective strategies for managing sewer failures.

Several factors commonly contribute to sewer failures, with pipe deterioration, environmental impacts, and system load variations among the most significant. Studies such as Jin and Mukherjee (2010) have highlighted issues like ageing infrastructure, root intrusions, and construction defects, often compounded by inadequate maintenance, as key causes of sewer system failures. Early work by Arthur, Crow, and Pedezert (2008) emphasized the vulnerability of combined sewer systems, where low-flow velocities, and small pipe diameters promote debris accumulation even during heavy rainfall. High population densities further exacerbate these issues, with rising wastewater volumes increasing the likelihood of sewer blockages causing localized flooding. Similarly, Sier and Lansey (2005) identified sanitary sewer overflows (SSOs) as major contributors to network failures, driven by infiltration, undersized systems, and structural defects. These studies underscore the importance of proactive maintenance and monitoring to detect and address vulnerabilities early.

While physical deterioration and environmental conditions are well-documented causes of sewer failures, operational factors like fat, oil, and grease (FOG) deposits require additional attention. Husain et al. (2014) explored how FOG from food service establishments and residences solidifies in pipes, creating dense obstructions. Although grease traps help mitigate this, their effectiveness is often reduced by high-temperature wastewater and poor maintenance. These findings emphasize the multifaceted nature of sewer failures, including physical, environmental, and operational factors.

Hydraulic parameters such as shear velocity play critical roles in sediment transport and sewer performance (Peng, Zhu, and Zhang 2024; Song et al. 2020). Shear velocity determines the capacity of sewer flows to

mobilize, transport or deposit sediments, which can influence blockages, and thus localized flooding. Despite its significance, shear velocity remains under-explored as a feature in Machine Learning (ML) predictive models. This study addresses this gap by examining the impact of incorporating shear velocity into Random Forest (RF) models for predicting sewer network failure at an individual asset level.

Traditionally, water utilities relied on public reports to manage sewer network problems, leading to delays and increased costs. Innovations like Closed-Circuit Television (CCTV) inspections have improved the detection of physical sewer defects and blockages by allowing detailed visual assessments, though they are labour-intensive and costly for extensive networks (Rosin et al. 2022). Acoustic sensors and IoT-based systems, such as ultrasonic water-level sensors with 4 G telemetry, offer non-intrusive, real-time monitoring, but they face, connectivity issues, and high false alarm rates (Bin Ali, Horoshenkov, and Tait 2011; Faris et al. 2024). While these technologies enhance blockage detection, they highlight the need for better predictive and cost-effective solutions.

Machine learning (ML) has emerged as a transformative approach, enabling proactive maintenance by predicting the likelihood of future sewer issues at individual assets. ML algorithms leverage historical and operational data to optimize resource allocation and reduce operational costs (Ribalta et al. 2023). For instance, Goodarzi and Vazirian (2024) applied a Support Vector Machine (SVM) model to predict sewer failures with 84% accuracy, while Bailey (2016) employed decision trees to improve predictions of recurrent failures. These models demonstrate the potential of ML to provide asset condition information to enable predictive maintenance strategies.

Recent advancements in predictive maintenance emphasize the integration of dynamic environmental variables with traditional static attributes (Nguyen and Seidu 2022). Factors like rainfall intensity, flow characteristics, and sediment transport significantly influence sewer performance, particularly in ageing or capacity-limited infrastructure (Draude et al. 2019; Muttil, Nasrin, and Sharma 2023). Among ML models, RF stands out for its accuracy in handling complex, nonlinear datasets typical of sewer network data (Vitorino et al. 2014).

RF is an ensemble learning method that combines multiple decision trees to improve prediction accuracy and robustness. Each tree is constructed using a random subset of the data, and at each node, a random selection of features is considered for splitting. This process reduces overfitting and ensures that the trees are diverse, which enhances the overall performance of the forest. RF models aggregate the predictions of individual trees by majority vote for classification or averaging for regression, ensuring stable predictions even in complex datasets. This combination of randomness and aggregation makes RF a versatile and powerful tool in predictive modelling (Breiman 2001).

A defining feature of RF is its ability to internally estimate model performance and feature importance using out-of-bag (OOB) data, eliminating the need for a separate validation set. Theoretical guarantees provided by the Strong Law of Large Numbers ensure that the generalisation error stabilises as the number of trees increases, preventing overfitting. Breiman (2001) highlights that the effectiveness of RF depends on the strength of the individual trees and the correlation between them where stronger trees and lower correlation lead to better predictions. This balance is achieved by the randomisation techniques employed in the tree-building process, making RF both efficient and reliable across various applications.

In the context of sewer systems, RF has been employed to predict sewer anomalies by analysing historical data and identifying patterns associated with failure events. For instance, Hassouna et al. (2019) utilised RF alongside decision trees and logistic regression to develop predictive models using real datasets from a major UK water and waste services provider. Their models achieved accuracies exceeding 89%.

Studies by Tavakoli, Sharifara, and Najafi (2020) and Okwori, Viklander, and Hedström (2021) highlight the effectiveness of RF in wastewater pipe and sewer blockage prediction. Tavakoli, Sharifara and Najafi demonstrated the RF's strong predictive power for wastewater pipe deterioration, successfully identifying priority pipes for inspection and maintenance. Similarly, Okwori, Viklander and Hedström employed RF ensembles to predict blockage recurrence in sewer networks, with accuracies between 60% and 80%, showcasing its capability in analysing diverse influencing factors such as pipe materials and environmental variables.

Earlier studies have widely applied RF models to predict sewer blockages by focusing on static features related to pipe characteristics. For example, Laakso et al. (2018) developed a model using attributes such as diameter, length, material, age, and gradient to evaluate sewer conditions, while Tavakoli, Sharifara, and Najafi (2020) employed similar physical pipe properties to predict deterioration.

However, these studies often fail to consider the considerable influence of environmental factors on sewer networks. For example, daily rainfall patterns significantly affect sewer performance. Heavy rainfall can

increase flow rates, leading to potential flooding and overloading of the sewer system. Although high flows might temporarily wash away debris, they can also carry high concentrations of sediments and organic matter into vulnerable pipe sections, creating new blockage risks. Conversely, during extended dry periods, sediment and debris are more likely to slowly accumulate, raising the risk of blockages, particularly in pipes with limited flow or smaller diameters (Marlow et al. 2011).

Although deep learning (Iqbal, Barthelemy, and Perez 2022; Jiang et al. 2021) and AI-driven (Patil et al., 2023) approaches have gained popularity in recent years, this study uses a RF model due to its strong performance on structured tabular data, its interpretability, and its proven effectiveness in previous sewer failure prediction research. RF models also require less training data and are more accessible for practical utility applications (Grinsztajn, Oyallon, and Varoquaux 2022).

This paper builds on prior research by incorporating shear velocity, an important yet underutilised hydraulic parameter associated with sediment entrainment and transport, into predictive models for sewer blockages, and local flooding. The study has three objectives: (1) to evaluate the predictive power of static attributes alone, (2) to assess improvements when dynamic hydraulic features are added, and (3) to analyse the specific contribution of derivatives of shear velocity that reflect local sediment transport capacity in identifying areas prone to sewer blockage and local flooding.

Methods

An overview of the methodological framework is shown in Figure 1. The workflow integrates multiple datasets describing physical pipe characteristics, operational interventions, and hydraulic model outputs, which are processed and combined prior to model development. The following sections describe the study area, data sources, and processing steps, and modelling approach in detail.

Catchment information

The study is conducted using data obtained from an urban catchment with an approximate size of $10 \, km^2$. The sewer network is predominantly a combined system flowing under gravity, with some outlying areas using pumped mains to transport flows to the treatment works. Due to the lack of specific sediment data for the catchment, it is assumed that blockages consist predominantly of fine sand, with a particle size ranging
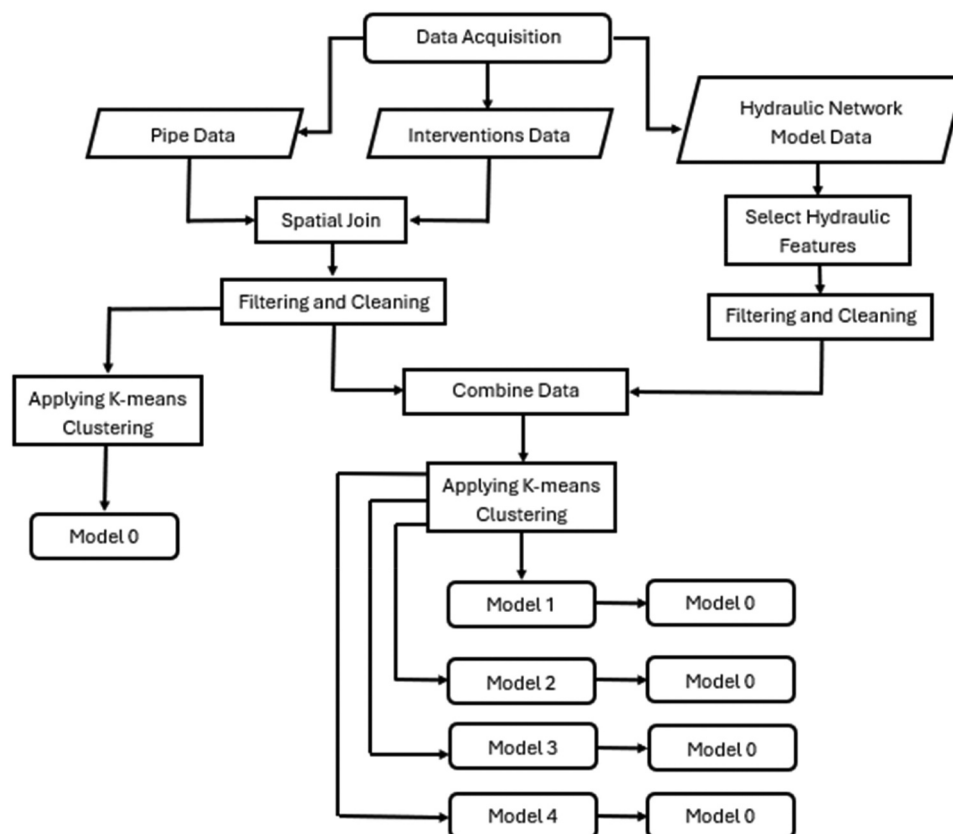

Figure 1. Methodological workflow for data processing, feature selection, and model development.

between 0.02 mm and 0.2 mm, this was defined as the most common type of sewer sediment in a wide-ranging review by Ashley et al. (2005). With this assumption in mind, it is also presumed that the catchment has not undergone significant changes or experienced any exceptional rainfall events that could substantially alter the expected sediment composition or distribution within the network during the data collection period.

Data processing

The data used in this study was sourced from the catchment area and comprised multiple datasets. The primary dataset included pipe attributes (diameter, length, and gradient), covering the entire network within the catchment. Diameters ranged between 100 mm and 1500 mm, the length ranged between 5.9 m and 105.6 m, and gradients ranged between 1/1000 and 1/10.

The second dataset consisted of 8 years of historical records from 2016 to 2023 of interventions related to observed system failure. This dataset is based on work orders raised for operational failures including blockages, flooding and a broad term of defects. The work orders have a geographic spot location, rather than being directly linked to a sewer asset.

A final dataset containing shear velocity (u∗) values extracted from a detailed Infoworks ICM 1D hydraulic model of the catchment, comprising 3270 nodes. This hydraulic model has been calibrated and validated by the water utility using water level and flow rate data collected during a 5-month flow survey during 2017–2018. The hydraulic model was run for the same 8-year period as the intervention data (2016 to 2023), during which historical work order records were available to identify and classify observed failures for model training, using rainfall radar data with a 5-minute time resolution. Results from the model were output at a 1-hour time step.

It is important to note that the dynamic hydraulic features used in this study, including shear velocity, were obtained from simulation outputs of a verified InfoWorks ICM model, and not from field-based sensors. This ensures the spatial data coverage is at a single asset level, but the approach remains practical and cost-effective by eliminating the need for widespread sensor deployment.

In sewer network design, Water UK recommends that the distance between adjacent manholes should not exceed 90 meters, ensuring no segment of a pipe is more than 45 meters from the nearest manhole (Water UK 2023). These principles inform the choice of spatial parameters for this analysis of pipe networks and their failures.

For this study, a spatial join was conducted in ArcGIS to link interventions (pipe failure) data with the nearest pipe segment. To determine the most suitable radius for analysis, three distances, 25, 50, and 75 meters, were evaluated. These radii (i.e. circular buffer distances used in the spatial join to locate and assign the nearest pipe failure to a certain pipe) were chosen based on standard design practices, where the typical spacing between manholes is generally within 25–50 meters, and distances up to 75 meters may be used in less common scenarios.

The evaluation revealed that a 25-meter radius captured a very limited number of connections, while a 75-meter radius resulted in challenges such as overlapping pipes and inconsistent linkages. A 50-meter radius was identified as the optimal distance, providing sufficient connections to address the analysis requirements without introducing overlaps or inconsistencies and aligned well with Water UK's design guidance. As a result, the 50-meter radius was used for the spatial join in this study.

Of the total pipe segments in this study, only about 1 in 20 experienced at least one recorded intervention during the 8-year period, highlighting a substantial imbalance between event and non-event classes. Before addressing this imbalance, the dataset was rigorously cleaned and pre-processed to ensure the RF classifier could function effectively. This process involved removing records with missing or incomplete values, ensuring all features were in the correct numerical format, and encoding categorical variables into numerical representations where applicable. The target variable, representing pipe issues, was encoded as 0 for pipes that had no events during the 8 years and 1 for pipes that experienced an event.

Previous studies on pipe condition assessment and failure prediction have predominantly utilized conventional sampling approaches such as Random Under-Sampling (RUS), Random Over-Sampling (ROS), Synthetic Minority Oversampling Technique (SMOTE), and class weighting to address data imbalance. Winkler et al. (2018) effectively combined RUS with boosted decision trees, demonstrating superior predictive performance in real-world scenarios. Similarly, Wijs et al. (2020) showed enhancements in pipe failure predictions by integrating SMOTE and class weighting with various ML algorithms. Robles-Velasco et al. (2021) highlighted the distinct advantages of under-sampling and over-sampling when combined with artificial neural networks (ANN), while Dimas, Nikolopoulos, and Makropoulos (2022) emphasized the necessity of over-sampling when dealing with imbalanced regression models for pipe failures. Fontecha et al. (2021) further illustrated the utility of over-sampling techniques in high-

resolution spatiotemporal analyses of urban sewer systems, reinforcing the benefits of balanced datasets.

Beig Zali et al. (2024) extended these conventional methods by incorporating domain-specific clustering of pipes based on attributes such as material, age, diameter, and length using K-Means, subsequently applying traditional sampling techniques to improve model performance. Inspired by this approach, this study adopts K-Means clustering as the primary undersampling method to address severe data imbalance. By grouping majority class samples and strategically selecting representative instances closest to cluster centroids, this approach ensures preservation of meaningful variability within non-event data.

Rainfall events were classified by identifying days with measurable rainfall and defining dry periods as those with no rainfall in the current and preceding 24 hours. For pipes with recorded events, the 90th percentile of hourly shear velocity values from the hydraulic model was calculated for the specific day on which the event occurred. For pipes without recorded events, daily 90th percentiles were calculated for the entire 8-year period, and the mean of these daily values was used. These summary values were used as model features, rather than inputting the raw time series directly.

Modelling methods

To evaluate predictive performance across multiple feature sets, the study developed five RF models focused exclusively on binary classification: identifying whether a pipe segment was likely to experience an issue (event) or not (no event). Each model incorporated a different combination of physical and hydraulic features.

Instead of a single train-test split, model performance was assessed using stratified 5-fold cross-validation, ensuring that class proportions were preserved in each fold. This approach provides a more robust estimate of generalisation performance and avoids bias associated with random splits (Salehinasab and Goshtasbi Rad 2025).

For each model, evaluation metrics included accuracy, precision, recall, F1-score, and ROC-AUC, calculated across folds. Confusion matrices were used to assess classification outcomes, while ROC curves provided insight into the model's ability to distinguish between classes. Additionally, feature importance scores derived from the trained RFs were used to interpret which inputs contributed most to the predictions.

Cleaned, filtered and balanced dataset sizes varied across models due to differences in feature availability, particularly for derived hydraulic features requiring upstream and downstream data. Final balanced datasets ranged between 242 and 988 records, depending on the feature combination used.

A central feature introduced in this study is using shear velocity and its derivatives within sewer pipes of the network, derived from shear stress values extracted from a calibrated sewer network model simulating both dry and wet weather performance over an 8-year period. This period aligns with the timeline of recorded interventions, ensuring that operational conditions at the time of an intervention could be estimated. Shear stress (τ) is a measure of the force per unit area exerted by the fluid flow against the pipe walls, which directly influences sediment movement and hence deposition patterns within the network. Assuming uniform, steady flow, it can be estimated using equation1:

$$\tau = \rho.g.R.S \tag{1}$$

where ρ represents fluid density, g is gravitational acceleration, R is the hydraulic radius, and S is the slope of the sewer line. High values of shear stress mean there is a greater potential to entrain and transport sediments at that location, which can help prevent the accumulation of debris (Peng, Zhu, and Zhang 2024).

Shear velocity (u_*) is derived from shear stress and given by:

$$u_* = \sqrt{\frac{\tau}{\rho}} \tag{2}$$

Thus, the shear velocity values could be indicative of the locations of the blockages. Peng, Zhu, and Zhang (2024) indicated that enhanced sedimentation was often observed in low gradient pipes.

Models

This study explored multiple iterations (five in total) of a predictive model, beginning with a base model that considers static features only. In each iteration, a different shear velocity metric was combined with the static features to assess its influence on pipe behaviour under varying conditions. Each iteration aimed to refine predictive capability by introducing elements that capture the interaction between hydraulic and structural factors.

Model 0: basic pipe attributes model:
Model 0 serves as the baseline model, focusing exclusively on the physical attributes of the pipes (diameter, length, and gradient) to establish how physical characteristics influence failure behaviour. These attributes were selected based on preliminary analyses, which indicated they provided good predictive performance

and were consistently available for most pipes in the dataset.

Model 1: difference of shear velocity at the ends of the pipe model

The first iteration (Model 1) focused on differences in shear velocity within the same pipe, comparing the upstream (u_* at usp) and downstream (u_* at dsp) values as demonstrated in Figure 2. The goal was to explore how localized hydraulic conditions within each pipe contribute to performance issues. While debris build-up was not explicitly modelled, significant differences in upstream and downstream shear velocity values could still highlight potential inefficiencies, such as areas prone to flow restriction or conditions favourable for sediment accumulation.

Model 2: difference of shear velocity in failure event flow and dry weather flow

The second iteration (Model 2) analysed the differences in shear velocity during wet events and dry weather flow.

- Failure event flow u_*: the u_* value chosen was the 90th percentile of u_* on the same day that any issue was reported, representing the impactful condition at the time of the observation without accounting for weather conditions.

- For pipes without issues, the median of the daily 90th percentile u_* values over an eight-year hydraulic simulation were used.

- Dry weather flow u_*: the values were based on days with no rainfall and no rainfall on the previous day, this u_* value was the 90th percentile during these days.

By comparing u_* values during these two scenarios (wet and dry weather), this iteration seeks to identify pipes that may be more susceptible to issues like erosion, sediment deposition, or blockages during high-stress events during dry weather or during rainfall events.

Model 3: difference of shear velocity at the upstream of the pipe and the downstream of the adjacent pipe – in-pipe streamwise variation:

This iteration (Model 3) examined the differences in shear velocity between the upstream end of one pipe (u_* at usp1) and the downstream end of the adjacent pipe that comes after it (u_* at dsp2) as demonstrated in Figure 3. The aim is to identify locations where the streamwise sediment transport capacity in the network changes rapidly. By focusing on these cross-pipe differences, the model highlights areas where significant shifts in transport dynamics might occur, which can influence sediment deposition and erosion within the network.
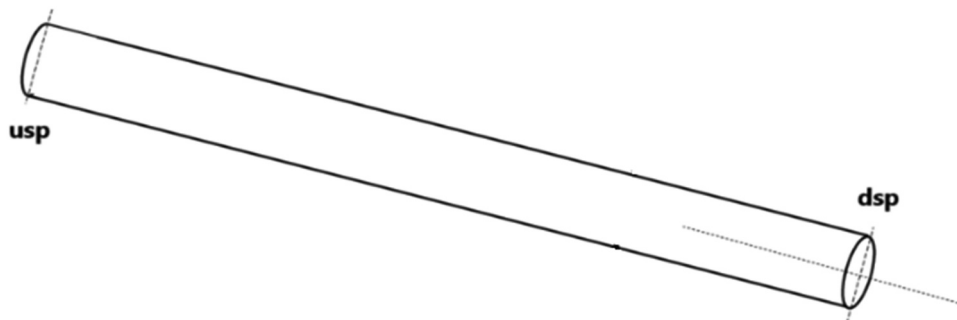

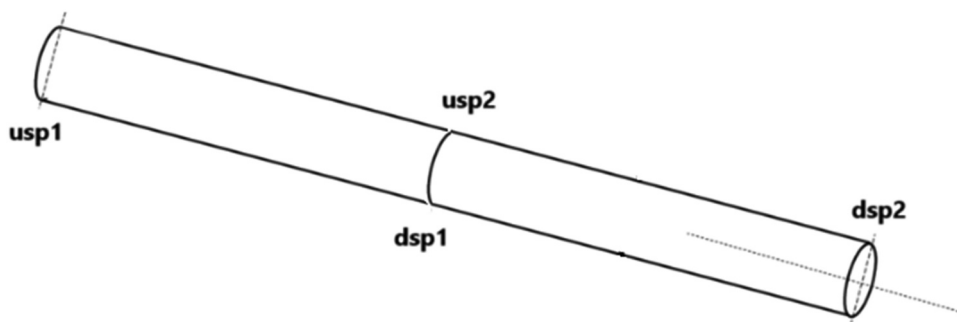**Figure 2.** Diagram for a sewer pipe illustrating both pipe ends.

Figure 3. Diagram for two consecutive sewer pipes illustrating all pipes ends.

Model 4: evaluation of sediment transport rates

This iteration (Model 4) focuses on evaluating sediment transport capacity between connected pipes in the hydraulic network. It examines how sediment mobilization and transport capacity are influenced by upstream and downstream flow conditions by introducing features derived from shear velocity values.

Due to the lack of specific sediment data for the catchment, it is assumed that the sediment consists predominantly of fine sand. This assumption simplifies the analysis, assuming the catchment is not undergoing any major hydrological changes.

$$\tau_{crit}^* = \frac{\tau_{crit}}{\rho.g.(s-1).d} \qquad (3)$$

Where (τ^*_{crit}) is the non-dimensional Shields parameter, (τ_{crit}) is boundary shear stress at the point of sediment entrainment, (ρ) is the fluid density, (g) is gravitational acceleration, (s) is the specific gravity of sediment taken as 2.65, and (d) is the sediment particle diameter assumed to be fine sand of 0.2 mm in diameter using Ashley et al. (2005) definition.

The Meyer-Peter Müller formula is widely used in granular sediment transport studies to estimate the sediment transport rate (q_s) under bedload conditions. It establishes that (q_s) is proportional to the excess shear stress raised to the power of 1.5:

$$q_s \propto (\tau - \tau_{crit})^{1.5} \qquad (4)$$

Substituting bed shear stress (τ) in terms of shear velocity (u_*), this proportionality becomes:

$$q_s \propto (u_*^2 - u_{*crit}^2)^{1.5} \qquad (5)$$

To understand the spatial significance of sediment transport estimates in terms of erosion or deposition potential, the Exner equation is used. It links the change in sediment transport rate along a spatial dimension to the rate of change in bed elevation:

$$\frac{\partial z_b}{\partial t} + \frac{1}{1-\lambda_p}.\frac{\partial q_s}{\partial x} = 0 \qquad (6)$$

Where (z_b) is the bed elevation, (t) is time, (λ_p) is the porosity of the sediment bed, and ($\partial q_s/\partial x$) is the spatial gradient of sediment transport rate.

The term ($\partial q_s/\partial x$) is particularly significant:

- When ($\partial q_s/\partial x$) > 0, there is a net increase in sediment transport downstream, indicating a higher risk of erosion as the material is removed from the pipe

- When ($\partial q_s/\partial x$) < 0, there is a net decrease in sediment transport downstream, suggesting a higher risk of deposition as material accumulates in the pipe.

This relationship allows the model to infer areas within the network that are prone to erosion or deposition based on the derived proportionality (5).

To evaluate sediment transport within and between connected pipes, three key features are introduced in this model:

The first feature is the average shear velocity within a pipe (u_{*avg}), calculated as the mean of the upstream and downstream values for the same pipe. For pipes with recorded issues, this value was taken from the specific day of the reported incident. For pipes without recorded issues, it was calculated as the mean of daily average u_* values over the eight-year simulation period, representing the baseline flow condition driving sediment mobilization where:

$$u_{*avg,p1} = \frac{u_{*usp1} + u_{*dsp1}}{2} \qquad (7)$$

The second feature is the sediment transport rate (q_s) within the pipe, derived from the proportionality (5). This feature represents the sediment mobilization potential across the pipe length.

The third feature (Δq_s) measures the change in sediment transport rate between the upstream pipe and the downstream pipe. It is calculated as:

$$\Delta q_s \propto \left[\left(u_{*avg,p2}^3 - u_{*crit}^3 \right) - \left(\left(u_{*avg,p1}^3 - u_{*crit}^3 \right) \right) \right] \qquad (8)$$

This feature provides insight into how sediment transport dynamics shift between connected pipes, highlighting regions with a higher likelihood of erosion ($\Delta q_s > 0$) or deposition ($\Delta q_s < 0$).

Summary of features across models

To highlight the progressive enhancement of the predictive models, Table 1 summarizes the features used in each iteration.

Key considerations

The initial dataset from the hydraulic network model provided shear velocity values at 5-minute intervals in each pipe. To effectively characterize the flow conditions, the 90th percentile of these u_* values were selected for two key reasons. First, higher shear velocities will be more effective for sediment mobilization. By focusing on the 90th percentile, the model emphasizes these critical periods, capturing the most influential

Table 1. Summary of input features used across model configurations.

Models	Features used
Model 0	Diameter, Length, Gradient
Model 1	Diameter, Length, Gradient, u_* usp, u_* dsp
Model 2	Diameter, Length, Gradient, u_* usp, u_* dsp, Δu_* (storm vs dry) upstream and downstream
Model 3	Diameter, Length, Gradient, u_* (usp1), u_* (dsp2), Δu_* (across pipe junction)
Model 4	Diameter, Length, Gradient, u_{*avg} (upstream & downstream pipes), $u_{*~avg}^3 - u_{*~crit}^3$ (upstream & downstream pipes), Δq_s (change in sediment transport capacity)

conditions during a day. Second, testing alternative percentiles, such as the median, showed minimal differences between wet and dry weather conditions, indicating an insignificant variation in flow regimes. In contrast, the 90th percentile effectively highlighted differences between wet and dry weather flows.

The code for the models that relied on adjacent pipes determines downstream pipes by matching the downstream node ID of a current pipe with the upstream node ID of other pipes in the dataset. This is done through a self-merge of the data, where each pipe's downstream node is compared with the upstream node of every other pipe. If a match is found, it indicates that the second pipe is directly connected downstream of the first. This method effectively maps the flow of water from one pipe to another in the network.

In most cases, if the network is linear, each pipe will have only one downstream connection. This means the downstream node of a pipe will match the upstream node of exactly one other pipe, resulting in a one-to-one relationship between pipes. The calculations, such as determining the difference in u_* values between the upstream and downstream pipes are straightforward in such cases because there is no ambiguity or need to choose between multiple downstream connections. This linearity ensures a clear and efficient analysis, reducing the complexity of the results.

However, in parts of the network where branching occurs and a pipe has multiple downstream connections, the code retains all possible connections. Each downstream pipe is treated as a separate case, resulting in multiple rows for the same upstream pipe in the dataset, which can lead to redundancy. To address this issue, the model compares the shear velocity values between the upstream and downstream pipes and selects the pipe with the highest difference to represent the connection allowing the model to focus on the connections with the greatest potential impact on sediment transport dynamics.

No feature scaling was applied, as RFs are unaffected by feature magnitude. Scaling offers no benefit for tree-based models (Ozsahin et al. 2022). Before training the predictive models, the dataset underwent a structured, model-specific preprocessing workflow to ensure consistency, reliability, and physical interpretability of

features. Records were excluded only when they lacked essential inputs required for constructing valid features, for instance, missing upstream or downstream shear velocity values in cases where streamwise gradients or adjacent-pipe comparisons were involved, or when the 90th percentile shear velocity value for the adjacent pipe on the event date was unavailable. Records with ambiguous or undefined static attributes (e.g. pipe diameter, or gradient) were also filtered to maintain categorical encoding consistency. Duplicate entries for the same pipe and issue type were removed to avoid overrepresentation. No imputation methods were applied, as dynamic hydraulic features are physically derived and context-sensitive; estimating them would risk distorting the underlying hydraulic behaviour and introduce unnecessary uncertainty. Although this filtration reduced the dataset size, it preserved the physical realism of the features and supported robust and interpretable model development. It is acknowledged that this trade-off may affect generalisability and limits predictions to pipes with known data, but it was essential to ensure that all model inputs were based on consistent, high-integrity information.

Performance evaluation

To evaluate the predictive performance of the models, we conducted stratified 5-fold cross-validation using accuracy, precision, recall, F1-score, and ROC-AUC as evaluation metrics. Given that each successive model (Models 1–4) incorporates new dynamic features beyond those used in Model 0, the inclusion of additional variables resulted in progressively smaller datasets due to the requirement for non-missing values across all features.

As a result, each model was trained and evaluated on a distinct subset of the data. For instance, Model 0 (which uses only static physical attributes: diameter, length, and gradient) included 988 assets, while Models 1 and 2 had 788 records each, and Models 3 and 4 had 242 and 540 records respectively. This variation makes direct comparisons across models based solely on accuracy or AUC potentially misleading due to differences in sample composition. To address this, a two-part evaluation strategy was adopted:

1- model evaluation on feature-filtered datasets

Each model was trained and validated on the subset of the data for which all required features were available. This ensured that every input to the model was based on complete, physically meaningful data without requiring imputation or approximation. By evaluating each model on its own valid dataset, we preserved the integrity of feature engineering choices and ensured that comparisons reflect the model's ability to learn from its specific input structure.

2- baseline evaluation using reduced datasets

To enable a consistent comparison across models with different feature sets, the baseline Model 0 was retrained on each of the filtered datasets used in Models 1 through 4, using only its original static features (diameter, length, and gradient). This allowed for a direct performance benchmark under identical data availability constraints. Any performance differences observed between a full model and its Model 0 counterpart can therefore be attributed to the impact of the additional features, rather than changes in dataset size or composition.

This dual approach allows for both practical evaluation of each model in its intended configuration and fair assessment of the added value brought by dynamic features. Evaluation outputs included standard classification metrics, confusion matrices, ROC curves, and feature importance plots derived from the RF algorithm.

Results

The results of the classification models are presented in Tables 2 and 3. Table 2 summarises the performance metrics for Models 0 through 4, each trained and validated on a distinct subset of the data based on the availability of required input features. Evaluation metrics include accuracy, precision, recall, F1-score, and ROC-AUC, computed using stratified 5-fold cross-validation.

Results are reported as the mean ± standard deviation across the five folds.

Table 3 presents Model 0 results using the reduced datasets for Models 1–4, enabling direct comparison under identical data availability conditions.

To illustrate this performance gap more intuitively, Figure 4 provides a direct visual comparison between Model 4 and its corresponding static baseline (Model 0 trained on the Model 4 dataset). The figure includes four panels for each model:

(a) cross-validation metrics,
(b) confusion matrices,
(c) ROC curves, and
(d) feature importance rankings.

This comparison highlights the gains in predictive performance and model confidence when dynamic variables are included in the input space.

Additional figures for the other models are provided in the Supplementary Materials:

- **Figure S1** presents cross-validation metrics, confusion matrices, ROC curves, and feature importance rankings for Models 0 through 4 using their full input feature sets.
- **Figure S2** contains the same visual outputs for Model 0 retrained on the reduced datasets of Models 1 through 4.

Discussion

This work aimed to quantify the relative improvement in predictive performance for sewer failures that occur when incorporating features derived from hydraulic and sediment transport theory into a machine learning framework. The results consistently support this objective.

Table 2. Performance metrics (mean ± standard deviation) for models 0–4.

Model	Accuracy	Precision	Recall	F1-scores	ROC AUV
Model 0	0.63 (±0.016)	0.63 (±0.016)	0.63 (±0.016)	0.63 (±0.016)	0.67 (±0.014)
Model 1	0.69 (±0.021)	0.69 (±0.021)	0.69 (±0.021)	0.69 (±0.021)	0.75 (±0.008)
Model 2	0.70 (±0.028)	0.70 (±0.027)	0.70 (±0.028)	0.70 (±0.029)	0.77 (±0.027)
Model 3	0.85 (±0.022)	0.85 (±0.019)	0.85 (±0.022)	0.85 (±0.023)	0.90 (±0.044)
Model 4	0.74 (±0.042)	0.74 (±0.043)	0.74 (±0.042)	0.74 (±0.042)	0.81 (±0.040)

Table 3. Performance of model 0 trained on reduced datasets from models 1–4.

Model	Using	Accuracy	Precision	Recall	F1-scores	ROC AUV
Model 0	Model 1 Dataset	0.66 (±0.017)	0.66 (±0.018)	0.66 (±0.017)	0.66 (±0.017)	0.70 (±0.027)
	Model 2 Dataset	0.63 (±0.021)	0.63 (±0.021)	0.63 (±0.021)	0.63 (±0.021)	0.67 (±0.022)
	Model 3 Dataset	0.68 (±0.046)	0.68 (±0.044)	0.68 (±0.046)	0.68 (±0.047)	0.70 (±0.022)
	Model 4 Dataset	0.60 (±0.027)	0.60 (±0.027)	0.60 (±0.027)	0.60 (±0.027)	0.62 (±0.045)

Model 0 Using Model 4 Dataset

```
Cross-Validation Metrics (mean ± std):
ACCURACY        : 0.598 ± 0.027
PRECISION_MACRO : 0.601 ± 0.027
RECALL_MACRO    : 0.598 ± 0.027
F1_MACRO        : 0.596 ± 0.027
ROC_AUC         : 0.615 ± 0.045
```

Model 4

```
Cross-Validation Metrics (mean ± std):
ACCURACY        : 0.735 ± 0.042
PRECISION_MACRO : 0.737 ± 0.043
RECALL_MACRO    : 0.735 ± 0.042
F1_MACRO        : 0.735 ± 0.042
ROC_AUC         : 0.811 ± 0.040
```
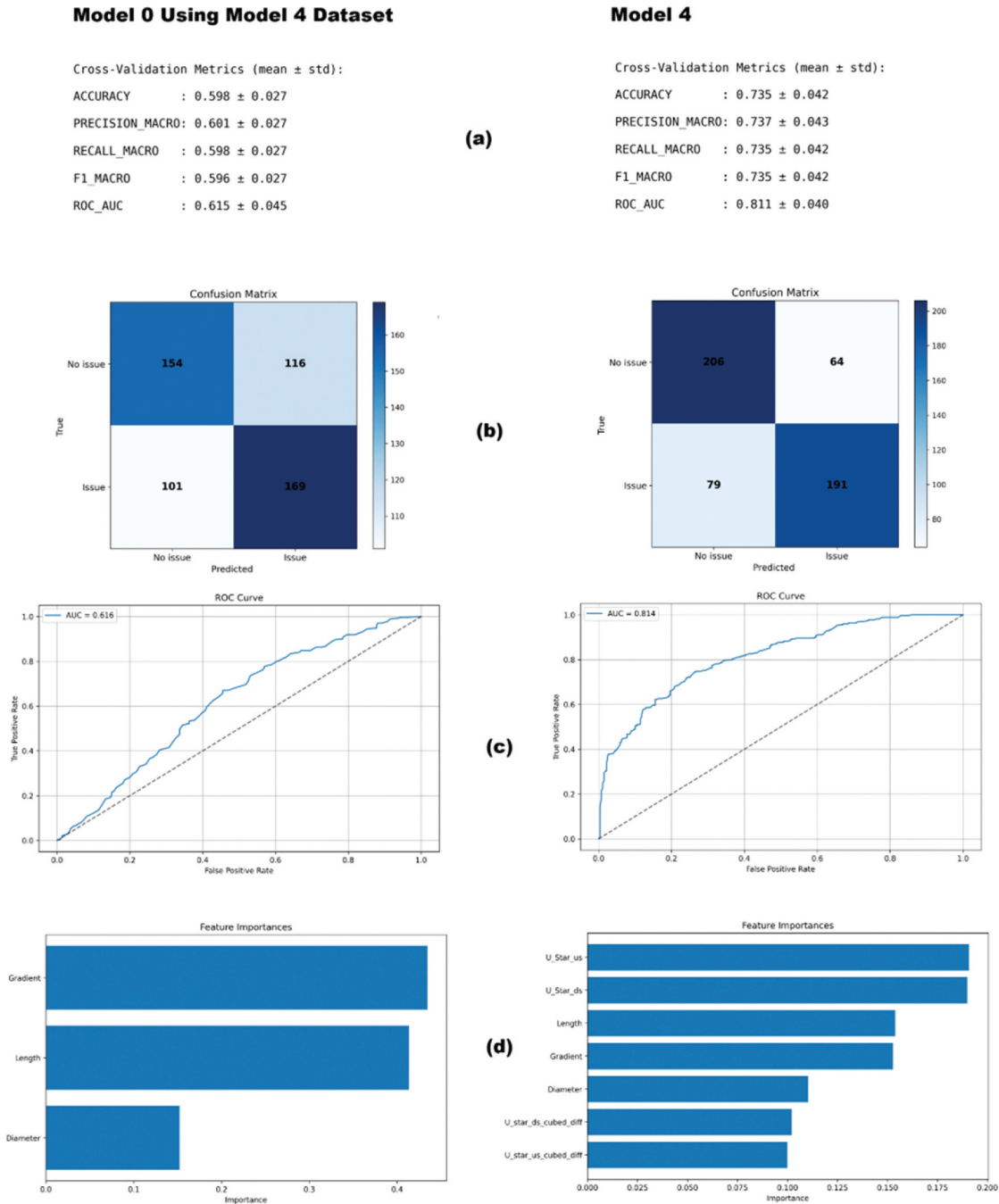
(a)

(b)

(c)

(d)

Figure 4. Model 4 vs. static baseline (model 0), metrics, confusion matrices, roc curves, and feature importance.

Model 0, using only static physical attributes such as pipe diameter, length, and gradient, provided a logical baseline. Its performance was relatively stable across datasets, but its limited feature scope meant it could not capture the operational and environmental variability that often precedes failure events. This aligns with findings from past studies where static-only models struggled to predict dynamic failure mechanisms such as sediment build-up or blockage recurrence during varying flow conditions.

Models 1 and 2 built upon this foundation by incorporating shear velocity (u_*) metrics. Model 1 introduced the difference in u_* between the upstream and downstream ends of the same pipe, providing a simple but informative representation of internal hydraulic conditions. Model 2 extended this by contrasting the shear velocity between wet and dry weather flow conditions. These additions allowed both models to reflect changes in sediment mobilization potential under different flow regimes. The

consistent performance improvements observed in both cases demonstrate the value of even basic hydrodynamic descriptors. When comparing Model 0 trained on the Model 2 dataset (63% accuracy) with the full-feature Model 2 (70% accuracy), the inclusion of rainfall-aware u_* variability improved the accuracy by 7%. A noticeable lift simply by including rainfall-aware u_* variability. These results underscore how relatively straightforward dynamic features can provide operational insight into blockage-prone behaviour.

Model 3 delivered the highest accuracy among all models (85%), using streamwise differences in shear velocity across consecutive pipes. This approach attempted to capture abrupt changes in sediment transport potential along the network, offering a proxy for identifying areas where erosion or deposition conditions might shift rapidly. However, these improvements came with limitations. For Model 3, calculating the streamwise gradients required valid 90th percentile u_* values for both the upstream end of the current pipe and the downstream end of the adjacent pipe on the specific incident date. If either value was missing or exceeded physical plausibility thresholds, the record was excluded. This strict, event-date-based matching substantially reduced the dataset size (from 988 records in Model 0 to only 242 in Model 3). This reduction likely excluded many of the more complex, variable sections of the network, meaning the resulting model may not generalize across all sewer conditions. Nevertheless, when we retrained Model 0 on the same filtered dataset, its accuracy reached only 68%, compared to 85% for the full-feature Model 3, an increase of nearly 17%. This delta confirms that the streamwise u_* features carried strong predictive value and were not simply exploiting quirks in a smaller dataset.

Model 4 approached the problem from a more physically grounded angle, introducing sediment transport rate estimations based on the Meyer-Peter Müller formula and the Exner equation. It considered both the average u_* within a pipe and the difference in sediment transport capacity between upstream and downstream pipes. These features directly relate to deposition and erosion tendencies, offering strong theoretical support for their relevance. Despite not achieving the same peak accuracy as Model 3, Model 4 maintained robust performance (increasing from 60% for model 0 using model 4 data to 74% for model 4) on a larger dataset and provided interpretable, physically meaningful predictors that align with observed failure mechanisms in real systems.

Feature importance rankings further reinforced the value of these dynamic additions. While traditional features like pipe length and gradient maintained relevance across most models, the new dynamic variables routinely ranked at or near the top in Models 1 through 4. For instance, features such as streamwise Δu_*, dry vs. wet u_* difference, and sediment transport rate change (Δq_s) consistently dominated the top positions.

Structural attributes such as pipe length and gradient consistently ranked among the most important features in all model iterations. This reflects their fundamental role in determining flow behaviour and sediment transport potential. Longer pipes, particularly those with shallow gradients, are more prone to debris accumulation and flow stagnation, increasing the likelihood of blockages or hydraulic inefficiencies. Moreover, these features are uniformly available across the dataset, enhancing their predictive strength relative to more sparsely distributed dynamic variables.

Model 3's drop in the importance of physical attributes like length likely resulted from the loss of data diversity in the reduced dataset, but even in that constrained setting, dynamic features remained dominant. This consistency across models suggests that hydrodynamic conditions are not just secondary effects but may be primary drivers in predicting failures related to sediment transport and blockages.

These findings validate the broader direction of incorporating hydraulic model outputs and sediment transport metrics into predictive sewer maintenance frameworks. While the models presented here can certainly be refined further with larger datasets, improved feature engineering, or alternate algorithms, the core message is 'dynamic system behaviour matters'. Its inclusion in modelling workflows offers measurable and interpretable improvements in identifying at-risk pipes. This understanding serves as a foundation for more operationally focused strategies, explored in the next section

Implications for practical application of applying dynamic RF models:

The integration of hydrodynamic features into machine learning models has direct operational implications for how utilities manage sewer networks. This study shows that using physically meaningful parameters (such as shear velocity and indicators of sediment transport potential) can enhance the ability to identify vulnerable assets under realistic flow conditions. These insights go beyond static risk assessments and support a more adaptive, data-informed approach to maintenance planning.

In practice, models that account for hydraulic variability can help shift decision-making from reactive to proactive. Instead of relying solely on CCTV schedules, historical complaints, or empirical heuristics, utilities could use these models to prioritise inspections based

on actual flow behaviours, highlighting pipes more likely to accumulate debris or experience sedimentation under current or seasonal conditions. This allows for more efficient allocation of field crews, potentially reducing emergency interventions and improving service reliability.

One key advantage of the approach used in this study is its compatibility with existing hydraulic simulation tools already in use across the water industry. Because the features are derived from model outputs rather than live sensors, utilities can extract risk-relevant data from routine simulations without requiring costly instrumentation upgrades. This makes it feasible to incorporate predictive modelling into routine asset planning workflows without introducing major overheads.

The structured nature of the feature design, particularly, features reflecting spatial gradients, flow condition contrasts, and estimated transport behaviour, also supports interpretability. Utilities and engineers can better understand not just where the risks are, but what mechanisms may be driving them. This opens the door for smarter root-cause analysis and targeted interventions, especially in complex, capacity-constrained networks.

While this study focused on classification, the structure of the models also lends itself to integration into broader decision-support tools. For instance, risk maps or ranking tables generated from the model outputs could inform funding decisions, pipe rehabilitation programmes, or capital planning prioritisation. Because the features are asset-level and physically grounded, they can easily be mapped and visualised in GIS tools or linked with existing maintenance records.

There is also the potential to apply this modelling framework as a foundation for early risk detection. Even without real-time sensing, simulations could be scheduled ahead of forecasted rainfall events, with the model used to highlight pipe sections at elevated risk based on projected hydraulic states. While this study did not focus on live forecasting, the structure of the features and model suggests strong potential for adaptation toward near real-time warnings and risk flags, particularly for high-priority areas such as properties with basements, schools, or critical infrastructure corridors.

Overall, this approach offers a scalable and physically meaningful method to inform sewer asset management. As utilities aim to modernise their networks under tighter budgets and climate uncertainty, tools that bridge hydraulic insight with machine learning will be increasingly important in delivering more resilient and responsive drainage systems.

Limitations of the approach

While this study demonstrates the potential of integrating dynamic hydraulic features into predictive models for sewer failure detection, several limitations should be acknowledged.

The analysis was based on historical data from a single urban catchment, which, although typical in its use of a combined and gravity-driven network layout common in the UK, may not reflect the full range of network configurations, or operational behaviours found elsewhere. As such, generalisability to other sewer systems, particularly those with distinct topographies, separate storm/sanitary systems, or different maintenance regimes, should be approached with caution.

Dataset preprocessing excluded records with missing or incomplete values, particularly where key hydraulic features (e.g. shear velocity) were unavailable. This led to smaller datasets for more complex models and may have introduced selection bias by removing pipes with less monitoring or incomplete metadata. While this step preserved physical interpretability, it also meant that some network sections (potentially those most difficult to assess in practice) were omitted from analysis; however, this also provides guidance on where asset databases could be enhanced to improve data quality for important parameters, thereby enabling better predictions.

To address class imbalance (where non-event cases outnumbered event cases by approximately 20:1), a cluster-based undersampling strategy was applied. While this improved class balance and model generalisation, the reduced number of samples, especially in Models 3 and 4 may limit the robustness of the findings. The smaller sample sizes in these models resulted from strict feature requirements (e.g. requiring upstream/downstream hydraulic matches), potentially excluding structurally important but less instrumented pipes.

Hydraulic parameters such as shear velocity were derived from simulation outputs of a calibrated hydraulic model, assuming uniform sediment properties (fine sand) and steady-state flow behaviour summarised via percentile-based metrics. While these simplifications provide consistency and computational efficiency, they may not fully capture transient hydraulic events, varying sediment loads, or localised debris formation, all of which could influence real-world failure patterns. Consequently, the results should be interpreted as

performance under expected rather than extreme conditions.

Although a single excess shear stress relationship has been used to indicate the transport capacity in individual pipes, we do appreciate that the sewer environment is complex, in terms of sediment diversity and environmental conditions. We have used sediment transport formulae that focussed on the transport capacity associated with bedload, as a rapid streamwise reduction in bedload transport capacity is often associated with severe in-pipe blockage. Suspended sediment, given the physical nature of its transport mode, is not often associated with in-pipe blockage. If the transport capacity of suspended sediment experiences a rapid reduction, it will just transition into bedload rather that transition into an in-pipe deposit. This is the reason we focussed on a bedload transport capacity relationship. Earlier studies that have studied the influence of biofilms (Schellart et al. 2005; Tait et al. 2003) and mixed grain size (Rushforth, Tait, and Saul 2003) have shown that these factors tend to influence the entrainment threshold of the sediment rather than the overall form of the transport. In the study here, a poor estimation of the critical shear stress may have a major impact on the difference in sediment transport capacity estimated between pipe sections if either pipe section was close to incipient motion. Examining the estimates, we made of the threshold of motion in the adjacent pipes, it was clear that in the locations that we predicted a rapid change streamwise transport capacity, neither pipe was close to threshold, so our estimate that in-pipe blockage development is likely to occur or not would be robust to uncertainties in estimating the sediment threshold of motion.

Alternative resampling methods were tested during preliminary experiments but did not yield stable results for this dataset, these approaches ultimately proved unsuitable due to a heightened risk of overfitting and the potential introduction of unrealistic patterns into the dataset. While these methods initially showed promising classification scores, closer inspection revealed inflated performance metrics that did not align with physically meaningful or interpretable outcomes. For this reason, the final approach relied on cluster-based undersampling of the majority class using K-means. This allowed the models to maintain generalisability by working with actual observed data, avoiding optimistic but misleading results driven by synthetic samples.

Additionally, the study attempted a novel approach using multiclass classification to distinguish between different failure types (such as blockages, defects, and localised flooding events) instead of treating all failures as a binary outcome. While conceptually valuable, the limited feature diversity and small sample size of labelled events made it difficult to produce stable or accurate predictions across all classes. This lack of discriminatory input features often led to overlapping model outputs and reduced precision, particularly in underrepresented failure categories. As such, the study retained a binary classification strategy for model simplicity and robustness.

Future directions

While the present study focused on same-day failure classification, preliminary testing also explored the models' ability to predict blockage-prone conditions based on data from preceding days. Models 3 and 4 maintained strong performance when supplied with shear velocity and sediment transport indicators from 24 to 72 hours prior to reported failures, suggesting that key hydraulic signals precede observed issues. This highlights the potential for applying similar models in a short-horizon risk forecasting context, enabling utilities to act before blockages fully manifest.

These early findings point to a promising direction: the development of predictive alarm systems built around critical hydraulic thresholds. Such systems could be integrated into operational workflows to flag high-risk pipes based on regularly updated hydraulic simulations. Although this study did not rely on real-time sensor data, future research could explore combining live flow or level measurements with model-based features to dynamically adjust risk profiles.

Additionally, expanding the dataset to include more diverse catchments, and incorporating data from additional weather events (such as prolonged dry spells, snowmelt periods, or extreme short-duration storms), may provide further improvements in accuracy and generalisability.

Multiclass classification holds significant potential for advancing targeted sewer maintenance strategies. In operational settings, distinguishing between failure types (such as blockages requiring jetting, structural defects needing excavation, or localised flooding that may require capacity intervention) can allow utilities to dispatch the appropriate crew and equipment from the outset. This precision-based approach could streamline on-site interventions, reduce unnecessary inspections, and lower overall maintenance costs. As future datasets grow in richness and granularity, incorporating additional operational, structural, and environmental predictors may support more reliable multiclass models. These could serve as intelligent

decision-support tools within digital twin frameworks or automated work-order systems, improving both the responsiveness and efficiency of urban drainage operations.

Conclusion

This study demonstrates the potential of integrating dynamic hydraulic features, specifically shear velocity metrics and sediment transport indicators, into RF models for predicting sewer network failures. By moving beyond traditional static attributes, the models developed here show that hydrodynamic conditions are critical predictors of operational performance failures.

Models incorporating upstream/downstream flow contrasts and streamwise gradients (particularly Models 3 and 4) showed significant improvements in predictive accuracy when applied to datasets with sufficient hydraulic detail. These improvements underscore the importance of capturing localised flow behaviour and sediment movement capacity when assessing failure risk.

Feature importance analysis revealed that while structural attributes like pipe gradient and length remain valuable, dynamic features consistently ranked among the most influential predictors. This shift toward physically meaningful, simulation-derived features reflects a broader opportunity for data-driven sewer asset management, particularly as more utilities gain access to hydraulic models.

Despite limitations related to dataset size, and modelling assumptions, the results provide strong evidence that incorporating hydrodynamic features leads to more accurate and interpretable predictions. Continued development in this space, supported by real-time data, broader catchment studies, and streamlined feature engineering, will support more targeted, proactive, and cost-effective maintenance strategies in modern wastewater networks.

Notations

τ	Shear stress (Pa or N/m^2).
ρ	Fluid density (kg/m^3).
g	Gravitational acceleration (9.81 m/s^2).
R	Hydraulic radius (m).
S	Slope of the pipe (unitless).
u_*	Shear velocity (m/s).
τ_{crit}	Critical shear stress for sediment mobilization (Pa).
τ_{*crit}	Critical Shields Parameter at sediment mobilisation (dimensionless)
s	Specific gravity of sediment (dimensionless).
d	Sediment particle diameter (m).
q_s	Sediment transport rate (kg/m/s).
u_{*crit}	Critical shear velocity - at sediment mobilisation (m/s)
z_b	Bed elevation (m).
t	time (s)
x	distance (m)
λ_p	Porosity of the sediment bed (dimensionless).
Δq_s	Change in sediment transport rate (kg/m/s).

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Oussama Aounali (iD) http://orcid.org/0009-0007-0648-0833
Will Shepherd (iD) http://orcid.org/0000-0003-4434-9442
Simon Tait (iD) http://orcid.org/0000-0002-0004-9555

Data availability statement

The data used in this study was supplied under licence from a water utility. Due to commercial restrictions supporting data is not available.

References

Arthur, S., H. Crow, and L. Pedezert. 2008 August. "Understanding Blockage Formation in Combined Sewer Networks." *Proceedings of the Institution of Civil Engineers-Water Management* 161 (4): 215–221. https://doi.org/10.1680/wama.2008.161.4.215.

Ashley, R. M., J. L. Bertrand-Krajewski, T. Hvitved-Jacobsen, and M. Verbanck, eds. 2005. *Solids in Sewers*. IWA Publishing. https://doi.org/10.2166/9781780402727.

Bailey, J. R. 2016. "Data Driven Models of Blockage Likelihood in the Wastewater Network." Masters thesis, *University of Exeter (United Kingdom)*. https://hdl.handle.net/10871/27354.

Beig Zali, R., M. Latifi, A. A. Javadi, and R. Farmani. 2024. "Semisupervised Clustering Approach for Pipe Failure Prediction with Imbalanced Data Set." *Journal of Water Resources Planning and Management* 150 (2): 04023078. https://doi.org/10.1061/jwrmd5.wreng-6263.

Bin Ali, M. T., K. V. Horoshenkov, and S. J. Tait. 2011. "Rapid Detection of Sewer Defects and Blockages Using Acoustic-Based Instrumentation." *Water Science and Technology* 64 (8): 1700–1707. https://doi.org/10.2166/wst.2011.183.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/a:1010933404324.

Dimas, P., D. Nikolopoulos, and C. Makropoulos. 2022. "Simulation Framework for Pipe Failure Detection and Replacement Scheduling Optimization." *Environmental Sciences Proceedings* 21 (1): 37. https://doi.org/10.3390/environsciproc2022021037.

Draude, S., E. Keedwell, R. Hiscock, and Z. Kapelan. 2019. "A Statistical Analysis on the Effect of Preceding Dry Weather on Sewer Blockages in South Wales." *Water Science and Technology* 80 (12): 2381–2391. https://doi.org/10.2166/wst.2020.063.

Faris, N., T. Zayed, E. Aghdam, A. Fares, and A. Alshami. 2024. "Real-Time Sanitary Sewer Blockage Detection System Using IoT." *Measurement* 226:114146. https://doi.org/10.1016/j.measurement.2024.114146.

Fontecha, J. E., P. Agarwal, M. N. Torres, S. Mukherjee, J. L. Walteros, and J. P. Rodríguez. 2021. "A Two-Stage Data-Driven Spatiotemporal Analysis to Predict Failure Risk of Urban Sewer Systems Leveraging Machine Learning Algorithms." *Risk Analysis* 41 (12): 2356–2391. https://doi.org/10.1111/risa.13742.

Goodarzi, M. R., and M. Vazirian. 2024. "A Machine Learning Approach for Predicting and Localizing the Failure and Damage Point in Sewer Networks Due to Pipe Properties." *Journal of Water and Health* 22 (3): 487–509. https://doi.org/10.2166/wh.2024.249.

Grinsztajn, L., E. Oyallon, and G. Varoquaux. 2022. "Why Do Tree-Based Models Still Outperform Deep Learning on Typical Tabular Data." *Advances in Neural Information Processing Systems* 35:507–520. https://doi.org/10.48550/arXiv.2207.08815.

Hassouna, M., M. Reis, M. Al Fairuz, and A. Tarhini. 2019. "Data-Driven Models for Sewer Blockage Prediction." 2019 International Conference on Computing, Electronics & Communications Engineering (iCCECE). London, UK, 68–72. IEEE. August. https://doi.org/10.1109/iCCECE46942.2019.8941848.

Husain, I. A., A. F. A. Ma, M. S. Jammi, M. E. Mirghani, Z. B. Zainudin, and A. Hoda. 2014. "Problems, Control, and Treatment of Fat, Oil, and Grease (FOG): A Review." *Journal of Oleo Science* 63 (8): 747–752. https://doi.org/10.5650/jos.ess13182.

Iqbal, U., J. Barthelemy, and P. Perez. 2022. "Prediction of Hydraulic Blockage at Culverts from a Single Image Using Deep Learning." *Neural Computing & Applications* 34 (23): 21101–21117. https://doi.org/10.1007/s00521-022-07593-8.

Jiang, Y., C. Li, L. Sun, D. Guo, Y. Zhang, and W. Wang. 2021. "A Deep Learning Algorithm for Multi-Source Data Fusion to Predict Water Quality of Urban Sewer Networks." *Journal of Cleaner Production* 318:128533. https://doi.org/10.1016/j.jclepro.2021.128533.

Jin, Y., and A. Mukherjee. 2010. "Modeling Blockage Failures in Sewer Systems to Support Maintenance Decision Making." *Journal of Performance of Constructed Facilities* 24 (6): 622–633. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000126.

Laakso, T., T. Kokkonen, I. Mellin, and R. Vahala. 2018. "Sewer Condition Prediction and Analysis of Explanatory Factors." *Water* 10 (9): 1239. https://doi.org/10.3390/W10091239.

Marlow, D. R., F. Boulaire, D. J. Beale, C. Grundy, and M. Moglia. 2011. "Sewer Performance Reporting: Factors That Influence Blockages." *Journal of Infrastructure Systems* 17 (1): 42–51. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000041.

Muttil, N., T. Nasrin, and A. K. Sharma. 2023. "Impacts of Extreme Rainfalls on Sewer Overflows and WSUD-Based Mitigation Strategies: A Review." *Water* 15 (3): 429. https://doi.org/10.3390/w15030429.

Nguyen, L. V., and R. Seidu. 2022. "Application of Regression-Based Machine Learning Algorithms in Sewer Condition Assessment for Ålesund City, Norway." *Water* 14 (24): 3993. https://doi.org/10.3390/w14243993.

Okwori, E., M. Viklander, and A. Hedström. 2021. "Spatial Heterogeneity Assessment of Factors Affecting Sewer Pipe Blockages and Predictions." *Water Research* 194:116934. https://doi.org/10.1016/j.watres.2021.116934.

Ozsahin, D. U., M. T. Mustapha, A. S. Mubarak, Z. S. Ameen, and B. Uzun. 2022. "Impact of Feature Scaling on Machine Learning Models for the Diagnosis of Diabetes." 2022 International Conference on Artificial Intelligence in Everything (AIE). Lefkosa, Cyprus, 87–94. IEEE. August. https://doi.org/10.1109/AIE57029.2022.00024.

Patil, R. R., R. K. Calay, M. Y. Mustafa, and S. M. Ansari. 2023. "ai-Driven High-Precision Model for Blockage Detection in Urban Wastewater Systems." *Electronics* 12 (17): 3606. https://doi.org/10.3390/electronics12173606.

Peng, X., D. Z. Zhu, and W. Zhang. 2024. "Transport of Non-Flushable Wipes in Sewers and Its Application in Sewer Management." *Journal of Cleaner Production* 434:139876. https://doi.org/10.1016/j.jclepro.2023.139876.

Ribalta, M., R. Bejar, C. Mateu, and E. Rubión. 2023. "Machine Learning Solutions in Sewer Systems: A Bibliometric Analysis." *Urban Water Journal* 20 (1): 1–14. https://doi.org/10.1080/1573062X.2022.2138460.

Robles-Velasco, A., C. Ramos-Salgado, J. Muñuzuri, and P. Cortés. 2021. "Artificial Neural Networks to Forecast Failures in Water Supply Pipes." *Sustainability* 13 (15): 8226 [online] https://doi.org/10.3390/su13158226.

Rosin, T. R., Z. Kapelan, E. Keedwell, and M. Romano. 2022. "Near Real-Time Detection of Blockages in the Proximity of Combined Sewer Overflows Using Evolutionary ANNs and Statistical Process Control." *Journal of Hydroinformatics* 24 (2): 259–273. https://doi.org/10.2166/hydro.2022.036.

Rushforth, P. J., S. J. Tait, and A. J. Saul. 2003. "Use of a Full-Scale Test Facility to Examine Sewer-Sediment Mobility." *Water and Environment Journal* 17 (1): 40–44. https://doi.org/10.1111/j.1747-6593.2003.tb00429.x.

Salehinasab, S., and E. Goshtasbi Rad. 2025. "Enhancing Pipe Failure Prediction Using Cluster-Specific Classification via K-Means Clustering." *Sustainable and Resilient Infrastructure*: 1–30. https://doi.org/10.1080/23789689.2025.2470504.

Schellart, A., R. Veldkamp, M. Klootwijk, F. H. L. R. Clemens, S. Tait, R. Ashley, and C. Howes. 2005. "Detailed Observation and Measurement of Sewer Sediment Erosion Under Aerobic and Anaerobic Conditions." *Water Science and Technology* 52 (3): 137–146. https://doi.org/10.2166/wst.2005.0070.

Sier, D. A., and K. Lansey. 2005. "Monitoring Sewage Networks for Sanitary Sewer Overflows." *Civil Engineering and Environmental Systems* 22 (2): 123–130. https://doi.org/10.1080/10286600500151460.

Song, Y. H., J. G. Joo, J. H. Lee, and D. G. Yoo. 2020. "Numerical Assessment of Shear Boundary Layer Formation in Sewer Systems with Fluid-Sediment Phases." *Water* 12 (5): 1332. https://doi.org/10.3390/w12051332.

Tait, S. J., R. M. Ashley, R. Verhoeven, F. H. L. R. Clemens, and L. Aanen. 2003. "Sewer Sediment Transport Studies Using an Environmentally Controlled Annular Flume." *Water Science and Technology* 47 (4): 51–60. https://doi.org/10.2166/wst.2003.0219.

Tavakoli, R., A. Sharifara, and M. Najafi. 2020 Pipelines 2020, edited by J. Felipe Pulido and Mark Poppe, pp. 90–102. Reston, VA: American Society of Civil Engineers. https://doi.org/10.1061/9780784483206.011.

Vitorino, D., S. T. Coelho, P. Santos, S. Sheets, B. Jurkovac, and C. Amado. 2014. "A Random Forest Algorithm Applied to Condition-Based Wastewater Deterioration Modeling and Forecasting." *Procedia Engineering* 89:401–410. https://doi.org/10.1016/j.proeng.2014.11.205.

Water UK. 2023. "Design and Construction Guidance for Foul and Surface Water Sewers Offered for Adoption Under the Code for Adoption Agreements for Water and Sewerage Companies Operating Wholly or Mainly in England ('The Code')." https://www.water.org.uk/sites/default/files/2023-11/SSG%20Appendix%20C%20-%20Design%20and%20Construction%20Guidance%20v2-3_0.pdf.

Wijs, R. J. A., G. F. Nane, G. Leontaris, T. R. W. Van Manen, and A. R. M. Wolfert. 2020. "Improving Subsurface Asset Failure Predictions for Utility Operators: A Unique Case Study on Cable and Pipe Failures Resulting from Excavation Work." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part A Civil Engineering* 6 (2): 05020002. https://doi.org/10.1061/AJRUA6.0001063.

Winkler, D., M. Haltmeier, M. Kleidorfer, W. Rauch, and F. Tscheikner-Gratl. 2018. "Pipe Failure Modelling for Water Distribution Networks Using Boosted Decision Trees." *Structure and Infrastructure Engineering* 14 (10): 1402–1411. https://doi.org/10.1080/15732479.2018.1443145.