*Article*

# Sustainable and Explainable Machine Learning for Risk-Based Maintenance Prioritization of Airport Stormwater Pipe Networks

**Emanuel Muniez[1], Mateo Fernando García[2] and Faustus Ricardo Sulla[3],***

Independent Researchers

**Abstract**

Airport stormwater pipe networks operate under demanding conditions where infiltration, structural defects, debris accumulation, hydraulic inefficiencies, and water-quality concerns can gradually reduce system performance. Conventional inspection-based maintenance can identify these deficiencies, but reacting to large inspection datasets and determining which pipe segments require the earliest intervention remains challenging. This study proposes an explainable machine-learning framework for risk-based maintenance prioritization of airport stormwater infrastructure. An anonymized dataset containing pipe characteristics, inspection observations, hydraulic conditions, infiltration indicators, and water-quality measurements is used to develop and evaluate a proof-of-concept machine learning framework for identifying higher-priority maintenance locations. Random Forest and gradient-boosting algorithms are explored within the framework, while explainable-AI techniques are used to identify the factors influencing maintenance priority. The framework is intended to support faster engineering response, more targeted inspections and rehabilitation, reduced unnecessary intervention, and improved asset life-cycle management. The proposed approach provides a data-driven pathway toward more resilient and sustainable maintenance of underground drainage infrastructure in complex airport environments.

**Keywords:** Asset management ; Sustainable infrastructure ; Predictive maintenance ; Explainable AI ; Stormwater drainage ; Airport infrastructure ; Machine learning

---

## 1. Introduction

Underground stormwater pipe networks are among the most critical assets supporting airport infrastructure. These systems are responsible for safely conveying runoff, preventing localized flooding, protecting pavement structures, and maintaining uninterrupted airport operations during storm events. Because most components are buried, deterioration often progresses unnoticed until operational deficiencies or structural failures become evident. Delayed intervention can increase maintenance costs, reduce system reliability, shorten asset service life, and disrupt airport operations. Therefore, effective maintenance planning has become an essential component of sustainable airport infrastructure management **[1],[2]**.

Most airport drainage systems are maintained through periodic inspections and corrective actions after deficiencies have been identified. Although this approach has been widely adopted, it often results in delayed maintenance decisions because engineers must evaluate large volumes of inspection data before determining which assets require immediate attention. As airport infrastructure continues to age and maintenance resources remain limited, prioritizing pipe segments based solely on engineering judgment becomes increasingly difficult. A more systematic approach is needed to identify high-risk assets before their condition leads to operational or environmental consequences.

Recent advances in machine learning have created new opportunities for infrastructure asset management by enabling engineers to identify hidden patterns in large inspection datasets and estimate future asset conditions with greater accuracy. Data-driven models have been successfully applied to a wide range of civil engineering problems, including pipe damage classification, tunnel deformation prediction, and environmental system monitoring. These approaches can process complex relationships among multiple variables that are difficult to capture using conventional statistical or rule-based methods. As a result, machine learning has become an increasingly valuable tool for supporting engineering decisions while reducing uncertainty in maintenance planning **[3-5]**.

Despite these advances, relatively few studies have applied explainable machine learning to the prioritization of maintenance for underground pipe networks in airport environments. Existing research has largely focused on damage classification, condition assessment, or service quality evaluation, while less attention has been paid to transforming inspection data into practical maintenance priorities to support engineering decision-making. In addition, many machine learning models operate as black boxes, making it difficult for engineers to understand why a particular asset is identified as high risk. Improving model transparency is essential if artificial intelligence is to be adopted as a reliable decision-support tool for infrastructure management **[3-5]**.

This study presents an explainable machine learning framework for prioritizing maintenance activities within airport underground stormwater pipe networks. The proposed framework integrates infrastructure inspection data with supervised learning algorithms to predict maintenance priority and identify the factors that have the greatest influence on maintenance decisions. In addition to evaluating predictive performance, the study applies explainable artificial intelligence techniques to improve the transparency and interpretability of the developed models. The objective is not only to improve prediction accuracy but also to provide practical engineering insights that help infrastructure managers allocate maintenance resources more efficiently, reduce unnecessary interventions, and support sustainable asset management **[6]**.

The primary contribution of this study is the development of an explainable machine learning framework that integrates engineering inspection data with predictive analytics to support maintenance prioritization for underground pipe networks. Unlike conventional condition assessment methods, the proposed approach combines predictive capability with model interpretability, enabling maintenance decisions to be supported by transparent and data-driven evidence **[7]**.

## 2. Study Dataset

*2.1 Dataset Description*

The dataset used in this study consists of inspection and operational records collected from an airport underground stormwater pipe network. To protect the confidentiality of the infrastructure owner, all identifying information has been removed, and only engineering attributes relevant to the analysis are considered. The dataset contains information describing the physical condition, hydraulic performance, environmental observations, and maintenance-related characteristics of individual pipe segments. These variables provide a comprehensive representation of factors that influence maintenance requirements and asset condition. The available inspection records represent a case-study dataset intended to demonstrate the feasibility of the proposed explainable machine learning framework rather than to establish a universally applicable predictive model [4], [5].

*2.2 Data Preprocessing*

Data preprocessing plays an important role in improving the reliability and predictive capability of machine learning models. Raw inspection records often contain missing values, inconsistent formats, categorical variables, and attributes measured on different numerical scales. If these issues are not addressed before model development, they may introduce bias and reduce model performance. Therefore, the dataset was preprocessed through a sequence of data cleaning, transformation, and normalization procedures before model training.

Missing values were first examined to determine whether they resulted from incomplete inspections or unavailable measurements. Records with excessive missing information were excluded, while isolated missing values were treated using appropriate imputation techniques based on the characteristics of each variable. Duplicate observations and inconsistent records were also removed to improve data quality.

Categorical variables describing inspection observations and maintenance conditions were converted into numerical values using label encoding or one-hot encoding, depending on the number of categories and their characteristics. Numerical variables with different measurement ranges were normalized using Min-Max normalization to prevent variables with larger magnitudes from dominating the learning process. The normalized value of each feature was calculated as

$$x_i^{\{*\}} = frac\{x_i - x_{\{\backslash min\}}\}\{x_{\{\backslash max\}} - x_{\{\backslash min\}}\}$$

Where $x_i$ is the original value, $x_{\{\backslash min\}}$ are the minimum and maximum values of the corresponding feature, and $x_i^{\{*\}}$ is the normalized value.

Following preprocessing, the dataset was evaluated using cross-validation to reduce sensitivity to any single training-testing split. Given the limited number of pipe segments, this approach was used to provide a more reliable assessment of model stability and generalization [4], [5].

*2.3 Proposed Research Framework*

The overall methodology adopted in this study is illustrated in Figure 1. The framework begins with the collection and preprocessing of infrastructure inspection data, followed by feature engineering and machine learning model development. Multiple supervised learning algorithms are trained and evaluated using the same dataset to identify the model that provides the highest prediction accuracy. The selected model is then interpreted using explainable artificial intelligence techniques to identify the engineering variables that most strongly influence maintenance priority. Finally, the generated predictions are translated into a practical decision-support tool that assists engineers in prioritizing maintenance activities within airport underground stormwater pipe networks.

The proposed framework emphasizes both predictive accuracy and model interpretability. While achieving high prediction performance is important, maintenance decisions must also be understandable and technically defensible. Explainable artificial intelligence provides this capability by revealing how individual infrastructure characteristics contribute to the prediction process. Consequently, the framework supports more transparent engineering decisions while improving maintenance efficiency and promoting sustainable infrastructure management [3-5].

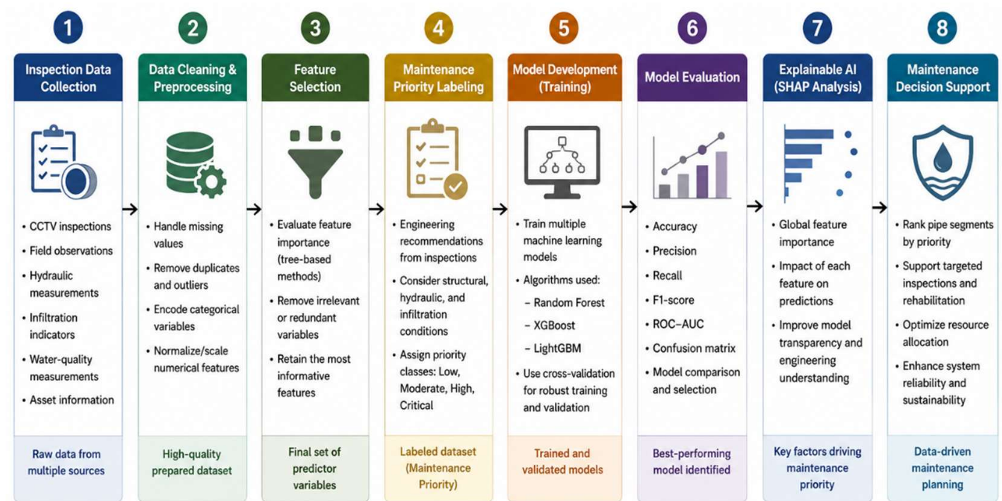


**Figure 1.** Proposed research framework for sustainable maintenance prioritization of stormwater pipe networks using explainable machine learning.

### *2.4 Feature Selection*

The predictive capability of a machine learning model largely depends on the quality and relevance of the input variables. Including irrelevant or redundant features may increase computational complexity, reduce model interpretability, and negatively affect prediction accuracy. Therefore, feature selection was performed to identify the variables that provide the greatest contribution to maintenance priority prediction while minimizing unnecessary information.

The candidate variables considered in this study represent multiple aspects of underground stormwater pipe performance, including structural condition, hydraulic characteristics, environmental observations, inspection findings, and maintenance history. These variables collectively describe the operational state of each pipe segment and provide sufficient information for the learning algorithms to recognize patterns associated with maintenance requirements.

The final predictor set included pipe characteristics, structural-condition indicators, hydraulic-condition indicators, and infiltration-related variables. Features with limited engineering relevance or redundant information were excluded before model training.

Rather than relying solely on statistical correlation, feature importance was evaluated during model development using tree-based learning algorithms. This approach allows the selected variables to reflect both linear and nonlinear relationships within the dataset, leading to a more representative description of the factors influencing maintenance priority. Similar feature selection strategies have been successfully applied in infrastructure management and explainable machine learning studies because they improve model robustness while maintaining engineering interpretability [3-5].

### *2.5 Machine Learning Model Development*

To establish an accurate and reliable maintenance prioritization framework, this study employs multiple supervised machine learning algorithms capable of identifying complex relationships between infrastructure characteristics and maintenance priority. Tree-based learning methods were selected because they effectively capture nonlinear interactions among variables, are robust to noisy engineering datasets, and provide reliable predictive performance for infrastructure management applications. In addition, these models naturally estimate feature importance, making them well suited for integration with explainable artificial intelligence techniques.

Three ensemble learning algorithms were considered in this study, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). These algorithms have demonstrated excellent predictive capability in engineering applications involving classification, condition assessment, and infrastructure risk analysis [5], [8], [9]. To provide a consistent evaluation, each model was trained using the same preprocessed dataset and evaluated using the same 5-fold cross-validation procedure. Their predictive performance was then compared using the classification metrics described in the following subsection.

Among the selected algorithms, Random Forest constructs an ensemble of decision trees and combines their predictions through majority voting. Each decision tree is developed using a random subset of observations and predictor variables, reducing overfitting while improving model stability. During tree construction, the quality of each split is commonly evaluated using the Gini impurity criterion, which is defined as

$$G = 1 - \sum_{i=1}^{C} p_i^2$$

where $p_i$ represents the probability of class $i$, and $C$ denotes the total number of output classes. Lower values of $G$ indicate purer node partitions.

XGBoost and LightGBM further improve predictive performance by sequentially constructing decision trees that minimize prediction errors generated by previous trees. Unlike conventional ensemble methods, gradient boosting optimizes an objective function that combines prediction loss with a regularization term to improve model generalization. The optimization objective can be expressed as

$$L = \sum_{i=1}^{n} l(y_i, \hat{}y_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where $L$ the overall objective function, $l$ represents the prediction loss, $f_k$ denotes the individual decision trees, and $\Omega(f_k)$ is the regularization function that controls model complexity. This optimization strategy enables gradient boosting algorithms to achieve high predictive accuracy while reducing the risk of overfitting, particularly when analyzing complex engineering datasets.

*2.6 Model Evaluation Metrics*

To provide a comprehensive assessment of predictive performance, the developed machine learning models were evaluated using several widely accepted classification metrics. Since maintenance prioritization requires not only accurate predictions but also reliable identification of high-risk pipe segments, relying on a single evaluation metric may lead to misleading conclusions. Therefore, this study employs five complementary performance measures, namely accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). Together, these metrics provide a balanced assessment of model performance and facilitate objective comparison among the developed algorithms.

Accuracy represents the proportion of correctly classified observations among all evaluated samples and is computed as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP denotes true positives, TN denotes true negatives, FP represents false positives, and FN represents false negatives.

Precision evaluates the reliability of positive predictions by measuring the proportion of correctly predicted positive observations. It is calculated as

$$Precision = \frac{TP}{TP + FP}$$

Recall, also referred to as sensitivity, measures the ability of the model to correctly identify observations belonging to the positive class. It is defined as

$$Recall = \frac{TP}{TP + FN}$$

The F1-score represents the harmonic mean of precision and recall and provides a balanced measure of classification performance when both false-positive and false-negative predictions are important.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Finally, the receiver operating characteristic (ROC) curve and its corresponding area under the curve (ROC-AUC) were employed to evaluate the discrimination capability of each model across different decision thresholds. Higher ROC-AUC values indicate stronger classification performance and greater capability to distinguish between different maintenance priority levels. These evaluation metrics have been widely adopted in infrastructure-related machine learning studies because they provide a reliable and comprehensive assessment of classification performance [4], [5].

### *2.7 Explainable Artificial Intelligence*

Although advanced machine learning algorithms can achieve high predictive accuracy, many of these models provide limited insight into the reasoning behind their predictions. This lack of transparency often reduces engineers' confidence in adopting artificial intelligence for infrastructure management, particularly when maintenance decisions involve critical assets and limited financial resources. Therefore, model interpretability has become an essential requirement for practical engineering applications.

To improve the transparency of the proposed maintenance prioritization framework, this study employs Shapley Additive explanations (SHAP), an explainable artificial intelligence technique that quantifies the contribution of each input feature to the prediction generated by the machine learning model. Unlike conventional feature importance methods, SHAP provides both global and local explanations, allowing engineers to understand the overall influence of each variable while also explaining the prediction associated with an individual pipe segment [5], [10], [11].

The SHAP value for an individual feature is calculated by estimating its average marginal contribution across all possible feature combinations. The mathematical formulation is expressed as

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

Where SHAP value represents the contribution of feature $i$ , $F$ denotes the complete set of input features, $S$ is a subset of features excluding feature $i$ , and $f$ represents the prediction generated by the trained machine learning model.

The resulting SHAP values provide a quantitative explanation of the model predictions by identifying the variables that increase or decrease the predicted maintenance priority. Consequently, infrastructure managers can not only identify pipe segments requiring immediate intervention but also understand the engineering factors responsible for those recommendations. This level of interpretability improves confidence in the proposed framework and supports more transparent and defensible maintenance decision-making.

### *2.8 Maintenance Priority Classification*

To translate model predictions into practical maintenance actions, the predicted outputs were grouped into maintenance priority levels. This step converts numerical model results into categories that are easier for engineers and asset managers to interpret during planning. The proposed framework considers four levels: low, moderate, high, and critical priority. These levels reflect the urgency of intervention based on the predicted condition and associated risk indicators.

Maintenance priority classes were assigned using the engineering recommendations contained in the inspection records. Pipe segments with no significant deficiencies or only minor observations were categorized as low or moderate priority, while segments requiring cleaning, sealing, repair, lining, or rehabilitation were categorized as high or critical priority. The classification considered the combined structural, hydraulic, and infiltration conditions of each pipe segment rather than a single defect indicator.

## 3. Results and Discussion

### *3.1 Source Characteristics and Observed Maintenance Conditions*

The dataset analyzed represented 21 underground stormwater pipe segments with varying dimensions, material conditions, hydraulic characteristics, and observed deficiencies. Inspection records showed considerable variation across the network, ranging from segments with relatively minor infiltration staining to locations exhibiting multiple interacting deficiencies, including groundwater intrusion, ponding, debris accumulation, structural deformation, deteriorated lining, and unlined lateral connections. Although the dataset is relatively small, it represents a realistic engineering case study and provides sufficient information to evaluate the feasibility of the proposed maintenance prioritization framework.

Infiltration-related indicators were among the most frequently observed conditions. Moisture staining, localized wet areas, joint leakage, and groundwater entry through lateral or underdrain connections were identified across several portions of the network. Hydraulic deficiencies were also evident, particularly where debris, uneven pipe inverts, or localized deformation restricted flow and created conditions favorable for standing water. These observations are important from a maintenance perspective because individual defects may have limited impact when considered independently, while combinations of infiltration, restricted flow, and structural deterioration can substantially increase the urgency of intervention.

The resulting feature set was therefore organized into four principal groups: pipe characteristics, structural condition, hydraulic condition, and infiltration-related indicators. Maintenance recommendations derived from the inspection records were used as the engineering basis for defining the target maintenance-priority categories. This structure allowed the predictive framework to evaluate maintenance needs using combinations of observed conditions rather than relying on a single defect indicator.

Table 1. Summary of input variables considered for maintenance-priority assessment.

| Feature Group | Representative Variables |
|---|---|
| Pipe characteristics | Diameter, length, material, lining condition |
| Structural condition | Cracking, deformation, liner deterioration |
| Hydraulic condition | Ponding, debris, obstruction, uneven invert |
| Infiltration condition | Staining, wet spots, leakage, lateral and underdrain intrusion |

### *3.2 Maintenance Priority Distribution and Model Development*

The inspection records showed that maintenance needs were not evenly distributed across the evaluated pipe segments. Some segments exhibited only minor staining or isolated hydraulic deficiencies, while others contained combinations of active infiltration, structural damage, ponding, debris accumulation, or defective connections. In several locations, the inspection recommendations specifically called for sealing, heavy cleaning, lateral lining, or additional rehabilitation, indicating a higher level of maintenance urgency.

To convert these engineering observations into a machine learning target, the maintenance recommendations were grouped into four priority classes: low, moderate, high, and critical. Segments with no significant defects or only minor observations were assigned to the lower categories, while assets requiring active repair, sealing, cleaning, or rehabilitation were assigned to higher priority levels. The classification was based on the combined condition of each pipe segment rather than on a single defect indicator.

The preprocessed dataset was then used to train Random Forest, XGBoost, and Light GBM models. Each algorithm received the same input features and target classes to ensure a consistent comparison. Because the available dataset represents a relatively small engineering case study, model development was treated as a proof-of-concept application. Cross-validation was therefore emphasized to reduce sensitivity to a single training-testing split and to provide a more reliable indication of model stability.

To quantitatively evaluate the predictive performance of the developed models, Random Forest, XGBoost, and LightGBM were compared using accuracy, precision, recall, F1-score, and ROC-AUC. Model performance was evaluated using 5-fold cross-validation to account for the limited size of the case-study dataset and to provide a consistent comparison across the three algorithms. The resulting performance metrics are summarized in Table 2.

Table 2. Performance comparison of the machine learning models using 5-fold cross-validation.

| Model | Accuracy | Precision | Recall | F1-score | ROC-AUC |
|---|---|---|---|---|---|
| **Random Forest** | 0.524 | 0.516 | 0.458 | 0.462 | 0.747 |
| **XGBoost** | 0.524 | 0.413 | 0.479 | 0.439 | 0.695 |
| **LightGBM** | 0.238 | 0.113 | 0.185 | 0.132 | 0.245 |

The results indicate that Random Forest provided the strongest overall performance among the evaluated models, achieving an accuracy of 0.524, an F1-score of 0.462, and a ROC-AUC of 0.747. XGBoost produced a similar overall accuracy but slightly lower F1-score and ROC-AUC values, while LightGBM showed comparatively weaker performance for the limited case-study dataset. Based on these results, Random Forest was selected for the subsequent SHAP analysis.

### *3.3 Explainable AI Analysis*

To improve the interpretability of the maintenance-priority predictions, SHAP analysis was applied to the best-performing machine learning model. The analysis was used to quantify the contribution of individual input features to the predicted maintenance priority and to identify the conditions that most strongly influenced the model predictions. Figure 2 presents the SHAP summary plot, where features are ranked according to their overall influence on model output. This analysis provides a transparent connection between the predictive model and the engineering conditions observed during pipe inspections.

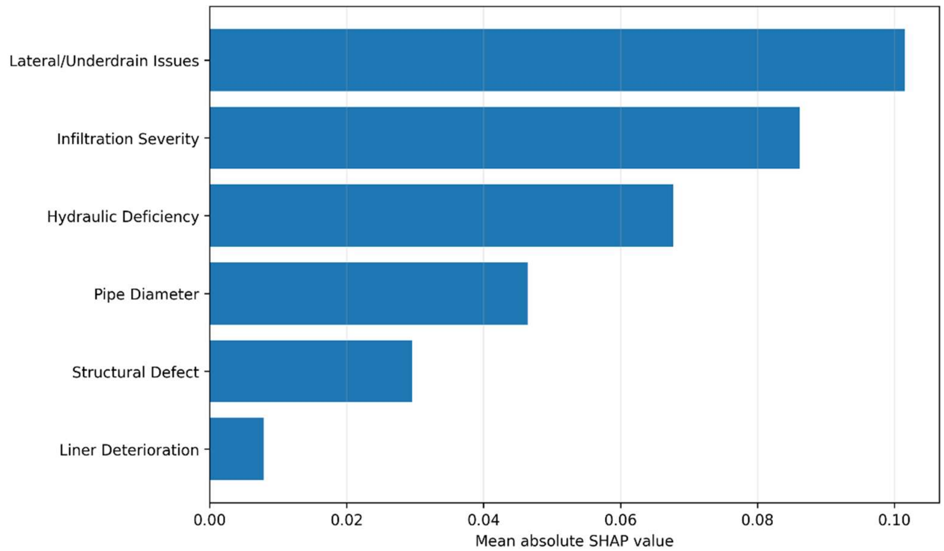


**Figure 2.** SHAP feature-importance analysis showing the relative influence of input variables on maintenance-priority predictions.

SHAP analysis was applied to the best-performing model to evaluate the relative influence of the input variables on maintenance-priority predictions. The results indicate that lateral and underdrain connection issues, infiltration severity, and hydraulic deficiencies were the most influential factors in the model output. Pipe diameter and structural defects showed a moderate influence, while liner deterioration had a comparatively lower contribution. These results suggest that maintenance priority is primarily driven by conditions that create pathways for groundwater intrusion and reduce hydraulic performance, rather than by a single structural deficiency alone.

### *3.4 Engineering Interpretation of Maintenance Priority Factors*

The inspection records indicate that maintenance priority is influenced by the interaction of multiple infrastructure characteristics rather than by any single observed defect. Pipe segments exhibiting isolated deficiencies, such as minor staining or localized debris, generally require routine maintenance. In contrast, segments containing simultaneous structural, hydraulic, and infiltration-related deficiencies represent a substantially higher maintenance risk because multiple deterioration mechanisms may accelerate future degradation.

Among the observed conditions, lateral and underdrain connection issues and infiltration-related defects appear to have the greatest influence on maintenance priority. Groundwater intrusion through joints, defective lateral connections, and underdrain interfaces introduces continuous moisture into the pipe system, increasing the likelihood of material deterioration, sediment transport, and long-term structural degradation. These observations are consistent with previous studies identifying infiltration as one of the primary factors affecting the long-term sustainability of underground drainage infrastructure [1].

Hydraulic deficiencies also play an important role in maintenance prioritization. Ponding, debris accumulation, localized sediment deposits, and restricted flow capacity reduce the hydraulic efficiency of the drainage network and increase the probability of localized flooding during major storm events. When these deficiencies occur together with structural deterioration, maintenance requirements become considerably more urgent because the resulting conditions may accelerate additional damage.

The explainable machine learning framework proposed in this study provides an advantage over conventional inspection review by evaluating the combined influence of these interacting variables rather than assessing each observation independently. This capability enables infrastructure managers to identify pipe segments that exhibit the highest overall maintenance risk and to prioritize interventions before deterioration progresses to more severe conditions. Consequently, maintenance planning becomes more proactive, resource allocation becomes more efficient, and rehabilitation activities can be scheduled based on measurable engineering evidence rather than subjective judgment alone.

### *3.5 Implications for Sustainable Airport Infrastructure Management*

The proposed framework demonstrates how explainable machine learning can support a transition from reactive maintenance toward a predictive and risk-based maintenance strategy for airport stormwater infrastructure. Rather than responding to failures after operational impacts have occurred, infrastructure managers can identify pipe segments with elevated maintenance priority using inspection data and predictive analytics. This proactive approach enables maintenance activities to be scheduled before deterioration reaches a critical stage, reducing both operational disruptions and long-term rehabilitation costs.

Compared with conventional inspection-based maintenance planning, the proposed framework provides a more systematic approach for prioritizing maintenance activities by simultaneously considering multiple structural and hydraulic indicators. Unlike traditional engineering assessments that typically evaluate individual defects independently, the proposed methodology integrates multiple condition variables within a single predictive framework while maintaining model transparency through explainable artificial intelligence. This capability supports more informed maintenance planning and demonstrates the potential of combining engineering knowledge with data-driven decision support for airport infrastructure management [12], [13].

From a sustainability perspective, prioritizing maintenance based on predicted risk provides several practical benefits. First, timely intervention can extend the service life of underground assets by addressing localized defects before they develop into extensive structural failures. Second, optimized maintenance scheduling reduces unnecessary inspections and rehabilitation activities, allowing agencies to allocate financial and human resources more efficiently. Finally, maintaining the hydraulic performance of stormwater systems contributes to improved environmental protection by reducing the potential for localized flooding, erosion, and uncontrolled pollutant transport within airport facilities [10].

An additional advantage of the proposed framework is its transparency. Conventional machine learning models often produce accurate predictions without providing sufficient information regarding the factors influencing those predictions. By incorporating explainable artificial intelligence, the proposed framework enables engineers to identify the variables contributing most strongly to maintenance priority. This capability increases confidence in the prediction process and allows maintenance recommendations to be supported by measurable engineering evidence rather than solely by algorithmic output.

Although the present study focuses on an airport stormwater drainage network, the proposed methodology is sufficiently flexible to be adapted to other buried infrastructure systems, including sanitary sewer networks, water distribution systems, and transportation drainage facilities. As inspection databases continue to expand through routine asset management programs, the predictive capability of the framework is expected to improve further. Consequently, the integration of explainable machine learning with engineering inspection data represents a promising direction for achieving more resilient, sustainable, and data-driven infrastructure management.

## 4. Conclusions

This study presented an explainable machine learning framework for prioritizing maintenance activities within airport underground stormwater pipe networks. By integrating infrastructure inspection records with supervised machine learning algorithms, the proposed framework provides a systematic approach for identifying pipe segments requiring earlier maintenance intervention. Unlike conventional inspection-based practices that rely primarily on engineering judgment, the proposed methodology integrates multiple infrastructure characteristics to support objective, data-driven maintenance decisions.

The incorporation of explainable artificial intelligence represents a key contribution of this study. In addition to predicting maintenance priority, the framework provides transparent explanations of the engineering factors influencing each prediction. This capability improves confidence in machine learning recommendations and allows infrastructure managers to better understand the relationship between observed defects and maintenance requirements.

The findings also demonstrate the potential of predictive analytics to support sustainable infrastructure management. By identifying higher-risk assets before significant deterioration occurs, airport agencies can optimize inspection schedules, allocate maintenance resources more efficiently, extend infrastructure service life, and reduce the likelihood of costly emergency repairs. These improvements contribute to more resilient stormwater systems while supporting long-term operational and environmental sustainability.

This study has several limitations. First, the proposed framework was evaluated using a case-study dataset from a single airport stormwater drainage network, which may limit the generalizability of the findings to other infrastructure systems. Second, the available inspection records represent a relatively limited number of assets, and larger datasets would improve model training and validation. Finally, the proposed framework should be validated using additional inspection datasets from different airports and environmental conditions to further evaluate its robustness and transferability.

The proposed framework demonstrates that integrating explainable machine learning with engineering inspection data can support more transparent and data-driven maintenance prioritization for airport stormwater pipe systems. By combining predictive modeling with model interpretability, the methodology enables infrastructure managers to better understand the factors influencing maintenance decisions while supporting more efficient allocation of inspection and rehabilitation resources. Although additional validation using larger and more diverse datasets is needed, the proposed approach provides a practical foundation for future intelligent asset management systems and contributes to the development of more sustainable airport infrastructure.

## Author Contributions

Conceptualization, Emanuel Muniez; methodology, Emanuel Muniez; investigation, Mateo Fernando García; writing, review and editing, Emanuel Muniez; Faustus Ricardo Sulla; supervision, Mateo Fernando García. All authors have read and agreed to the published version of the manuscript.

## Funding

This research received no external funding.

## Data Availability Statement

The original infrastructure inspection records used in this study are not publicly available due to confidentiality and ownership restrictions. A supplementary anonymized proof-of-concept dataset containing the variables and assumptions used to demonstrate the proposed machine learning framework is provided with this study.

## Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Iwanek, M. Hidden Risks to Sustainable Operation of Water Systems: Suffosion Process Triggered by Pipe Leakage. *Sustainability* 2026, *18*, 7593. https://doi.org/10.3390/su18157593

[2] Saliba, S.M.; Bowes, B.D.; Adams, S.; Beling, P.A.; Goodall, J.L. Deep Reinforcement Learning with Uncertain Data for Real-Time Stormwater System Control and Flood Mitigation. *Water* 2020, *12*, 3222. https://doi.org/10.3390/w12113222

[3] Kutyłowska, M.; Cieżak, W. Classification of Water Pipe Damage Types Using Random Forest. *Sustainability* 2026, *18*, 5101. https://doi.org/10.3390/su18105101

[4] Darabi, H., Mohamadi, S., Karimidastenaei, Z. *et al.* Prediction of daily suspended sediment load (SSL) using new optimization algorithms and soft computing models. *Soft Comput* 25, 7609–7626 (2021). https://doi.org/10.1007/s00500-021-05721-5

[5] Mirzamohammadi, M., Haghighatjoo, M., Taheri, H.J. *et al.* Physics-informed and transferable explainable AI framework for reliable prediction of tunnel-induced settlement in coastal environments. *Discov Artif Intell* 6, 558 (2026). https://doi.org/10.1007/s44163-026-01108-3

[6] Zhao, Y.; Zhou, S.; Zhao, C.; Valeo, C. The Influence of Geotextile Type and Position in a Porous Asphalt Pavement System on Pb (II) Removal from Stormwater. *Water* 2018, *10*, 1205. https://doi.org/10.3390/w10091205

[7] Wang, X.; Ma, X.; Li, Z. Research on SLAM and Path Planning Method of Inspection Robot in Complex Scenarios. *Electronics* 2023, *12*, 2178. https://doi.org/10.3390/electronics12102178

[8] Liu,J.;Xiao,Y.;Tan,L. (2025). Application of Machine Learning Model in Fraud Identification: A Comparative Study of CatBoost, XGBoost and LightGBM. Applied and Computational Engineering,119,135-141. https://doi.org/10.54254/2755-2721/2025.21637

[9] I. Ali, G. Kabir, U. M. Modibbo, F. Goodarzian, and E. D. R. S. Gonzalez, "Advances in big data optimization models, algorithms, and applications," *Decision Analytics Journal*, vol. 12, p. 100491, Sep. 2024, https://doi.org/10.1016/j.dajour.2024.100491

[10] Towsif Khan, S.; Chapa, F.; Hack, J. Highly Resolved Rainfall-Runoff Simulation of Retrofitted Green Stormwater Infrastructure at the Micro-Watershed Scale. *Land* 2020, *9*, 339. https://doi.org/10.3390/land9090339

[11] Guha, S.; Kodipalli, A.; Fernandes, S.L.; Dasar, S. Explainable AI for Interpretation of Ovarian Tumor Classification Using Enhanced ResNet50. *Diagnostics* 2024, *14*, 1567. https://doi.org/10.3390/diagnostics14141567

[12] S. Lin, "A multiple criteria decision-making model for enhancing informative service quality at airports," *Decision Analytics Journal*, vol. 12, p. 100487, Sep. 2024, https://doi.org/10.1016/j.dajour.2024.100487

[13] H. J. Taheri, "A KPI-Driven Dashboard Framework for Enhanced Safety Management in Airfield Construction," *EngrXiv*, Sep. 2025. https://doi.org/10.31224/5432