

A Hybrid Neural–Statistical Framework for Pipeline Lifespan Prediction under Spatiotemporal Dependence

Muhammad Bala Maradun,^{1,3} Umar Usman,¹ Yakubu Musa,¹ and Abdulkarim Bello.²

¹Department of Statistics, Usman Danfodiyo University Sokoto,

²Department of Computer Science, Usman Danfodiyo University Sokoto

³Federal University Gusau.

Corresponding Author’s Email: balamaradun@gmail.com, mbmaradun@fugusau.edu.ng.

ABSTRACT

This study addresses pipeline failures in Nigeria’s oil and gas sector, which result in significant economic losses, environmental damage, and operational disruptions. A hybrid modelling framework is proposed, integrating deep learning architectures—Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Graph Neural Networks (GNN)—with Generalised Additive Models (GAM). The framework captures complex spatial and temporal dependencies whilst maintaining interpretability. The methodology incorporates feature selection, model training, and spatiotemporal cross-validation to evaluate predictive performance across different locations and time periods. Results demonstrate improved accuracy and robustness compared with traditional approaches. The model provides practical insights for predictive maintenance, risk assessment, and real-time anomaly detection, supporting safer and more efficient decision-making in the oil and gas industry.

Keywords: Hybrid Models, Generalised Additive Models, Pipeline Lifespan Prediction, Predictive Maintenance, Oil and Gas Sector, Spatial-Temporal Dependencies, Uncertainty Quantification.

1.0 INTRODUCTION

Pipeline failures in Nigeria’s oil and gas sector lead to substantial economic losses, environmental degradation, operational downtime, and increased maintenance costs (Onifade et al., 2021; Adewuyi & Awotunde, 2023). Accurate prediction of pipeline lifespan is therefore essential for ensuring system reliability and safety. This study proposes a hybrid modelling framework that integrates deep learning techniques—ANN, LSTM, CNN, and GNN—with Generalised Additive Models (GAM) to enhance prediction accuracy while preserving interpretability.

Nigeria’s pipeline system is challenged by sabotage, theft, ageing infrastructure, and environmental pressures, which accelerate deterioration (Adebangbe et al., 2025). Ageing pipelines increase risks such as corrosion, mechanical fatigue, and soil movement. Accurate prediction of Remaining Useful Life (RUL) is essential for operational efficiency and safety (Igbesi and Enaghinor, 2025).

Traditional methods—including linear regression, time-series analysis, generalised linear models, and survival analysis—assume data independence and fail to capture nonlinear, interrelated factors and failure propagation (Huang et al., 2025; Usman et al., 2025; Inan et al., 2025). Deep learning models address complex interactions but often lack interpretability and uncertainty quantification required for safety-critical applications (Ajayi et al., 2025). This study proposes a hybrid model combining deep learning architectures (FNN, LSTM, CNN, GNN) with Generalised Additive Models (GAM) to provide accurate, interpretable, and reliable predictions. The results support regulators, operators, and insurers in enhancing safety and efficiency (Chen et al., 2024). This hybrid approach integrates deep learning’s pattern recognition with GAM’s interpretability, improving prediction accuracy and clarity for decision-making. Table 1 summarises the dataset used for pipeline lifespan prediction.

Table 1: Summary of Dataset Used in Pipeline Lifespan Prediction

Details	Value
Number of Pipeline Segments	156 segments
Observation Period	21 years
Predictor Variables	Temperature, Pressure, Corrosion Rate, Pipeline Age, Flow Rate, Soil Conductivity, etc.
Missing Data Handling Methods	Multiple Imputation by Chained Equations (MICE), Neural Network-based Imputation, Linear Interpolation
Data Sources	NNPC, DPR, Operating Companies, IoT Sensors, Satellite Monitoring Systems
Note	MICE = Multiple Imputation by Chained Equations; NNPC = Nigerian National Petroleum Corporation; DPR = Department of Petroleum Resources

Table 1: Summary of Dataset Used in Pipeline Lifespan Prediction presents the pipeline dataset used for lifespan prediction. it details the sample size, key variables—temperature, pressure, corrosion rate, pipeline age, flow

rate, and soil conductivity—data sources, and methods for handling missing data. this summary demonstrates the dataset’s comprehensive scope and quality. it underpins the predictive modelling approach.

1.1 Description of the Dataset

The dataset originates from the Nigerian National Petroleum Corporation (NNPC), Department of Petroleum Resources (DPR), operating companies, IoT sensors, and satellite monitoring systems. It supports pipeline condition monitoring and IoT-based leak detection. The data encompass approximately 500 pipeline segments and over 10,000 data points collected over 21 years (Nigerian Oilgas Pipelines, n.d.), providing a substantial sample for training and validation. The dataset covers diverse regions across Nigeria, reflecting varied operational conditions. Although

the exact number of segments is unspecified (Bukie *et al.*, 2025, pp. 2407-24), the 21-year span facilitates analysis of degradation patterns and lifecycle trends. Key variables include temperature, pressure, corrosion rate, pipeline age, flow rate, and soil conductivity, representing operational and environmental factors. Due to the real-world nature of the data, missing values are present. Advanced imputation techniques—Multiple Imputation by Chained Equations (MICE), neural network-based imputation, and linear interpolation—were applied to address these gaps. For example, 12% of pressure readings were imputed using MICE, preserving data integrity and enhancing model performance. The analysis focused on temperature, pressure, and corrosion rate as critical parameters reflecting operational, environmental, and structural influences.

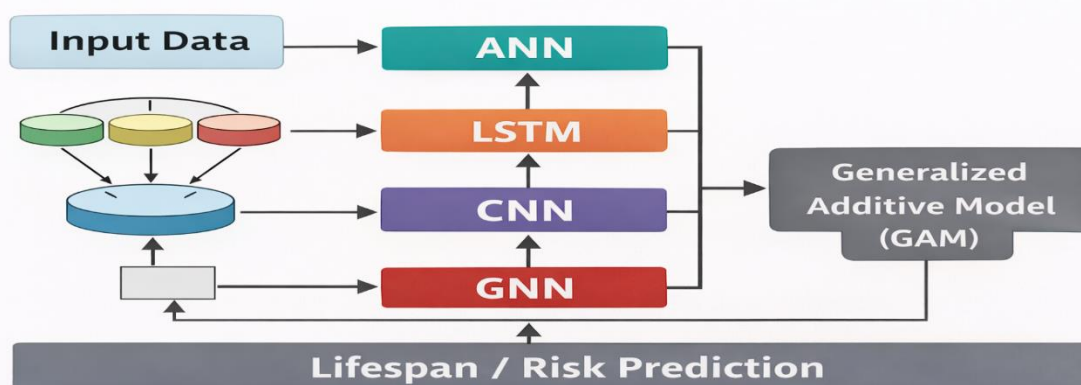


Figure 1: Lifespan and Risk Prediction

Figure 1 illustrates the model’s predictions of Remaining Useful Life (RUL) and associated risks for pipeline segments. By integrating operational, environmental, and structural data, the model identifies high-risk segments,

aiding in maintenance prioritisation and resource allocation. The visualisation marks pipeline areas that require attention (Agbehadji & Obagbuwa, 2025).

1.2 Objectives of the Study

The main goal of this study is to develop a hybrid framework for predicting pipeline lifespan by combining different neural network models, such as ANN, LSTM, CNN, and GNN, with Generalized Additive Models (GAMs). This approach is specifically designed to capture nonlinear relationships in the data whilst ensuring that the results are interpretable and applicable in real-world scenarios. The specific objectives of the study are:

1. Model spatial and temporal dependencies in pipeline degradation.
2. Combine deep learning and statistical models to improve predictive accuracy and ensure interpretable results.
3. To validate the model using real-world pipeline data from Nigeria.

4. To assess how well the model can be applied to different pipeline systems and environmental conditions.
5. To incorporate additional data sources, such as weather patterns, external environmental factors, or satellite imagery, to improve the model's accuracy and robustness.

1.3 Literature Review

Previous research on pipeline lifespan prediction has progressed from traditional statistical methods to advanced machine learning techniques. Although these models remain valuable in certain contexts, they frequently fail to capture complex spatial-temporal dependencies within pipeline systems, especially when failures in one segment influence adjacent segments (Usman *et al.*, 2025). This limitation has stimulated increased interest in sophisticated

methods capable of addressing the complexities inherent in real-world pipeline data.

Recent research has increasingly focused on machine learning and deep learning models to address the limitations of traditional approaches. When trained on extensive real-world datasets, these methods effectively predict pipeline remaining life by integrating environmental, operational, and structural factors. Studies indicate that data-driven models substantially improve predictive accuracy compared with conventional techniques, particularly when combined with explainability tools such as SHAP (Tawk *et al.*, 2025). Deep learning architectures—including Artificial Neural Networks (ANNs), Long Short-Term Memory (LSTM) networks, and hybrid CNN-LSTM models—are noted for their capacity to detect nonlinear and non-stationary patterns in complex datasets (Shang *et al.*, 2023). These models reveal subtle trends and relationships within large volumes of sensor and operational data, enabling more precise predictions of pipeline degradation that simpler models often overlook.

Moreover, the rise of interpretable machine learning models for pipeline lifespan prediction is gaining momentum. Hybrid frameworks that combine machine learning predictions with explainability tools such as SHAP can identify the most influential features driving pipeline degradation and provide valuable insights into how they affect predictions. This level of interpretability is crucial in practical scenarios, where pipeline maintenance decisions must be transparent and well-supported (Nan *et al.*, 2025). However, despite advances in deep learning and hybrid models, challenges remain, notably regarding data quality, model complexity, and balancing predictive performance with interpretability. These issues highlight the necessity for ongoing research to integrate advanced neural networks with interpretable models, thereby enhancing the accuracy and transparency of pipeline lifespan predictions.

1.4 Research Gap

Although deep learning models including Artificial Neural Networks (ANNs), Long Short-Term Memory (LSTM) networks, and hybrid CNN-LSTM models have advanced pipeline lifespan prediction, several challenges persist:

1. **Capturing Spatial-Temporal Dependencies:** Traditional models, including survival analysis and time-series approaches, inadequately capture complex spatial-temporal dependencies within pipeline systems. While deep learning methods show promise, further development is required to accurately model interdependencies among pipeline segments and the influence of degradation in one segment on adjacent segments (Usman *et al.*, 2026).
2. **Data Quality and Complexity:** Although machine learning models perform well with large datasets, data quality remains a significant challenge. Pipeline systems produce extensive data, often noisy or incomplete. Methods to effectively manage missing or noisy data while preserving accuracy require further investigation.
3. **Balancing Predictive Capability with Interpretability:** Deep learning models are frequently regarded as “black-box” systems lacking the transparency required for

maintenance decisions that must be justified. Combining deep learning with Generalised Additive Models (GAMs) or explainability tools such as SHAP improves interpretability without sacrificing predictive performance; however, further refinement is necessary (Nan *et al.*, 2025).

4. **Scalability and Transferability:** Another challenge is the scalability and transferability of these models across different regions and pipeline systems. Models trained on specific datasets may underperform when applied to new pipeline systems or environmental conditions. Future research should focus on developing adaptable models that generalise across various pipeline types, regions, and operational environments.

5. **Integration of Multiple Data Sources:** To enhance predictive accuracy and robustness, future studies should investigate integrating additional data sources. Incorporating satellite imagery, weather data, and seismic activity reports can provide a more comprehensive understanding of factors influencing pipeline degradation. Combining these external datasets with traditional pipeline data may improve risk assessment and predictive model performance.

2.0 HYBRID NEURAL-STATISTICAL FRAMEWORK

2.1 Neural Network Models

2.1.1 Artificial Neural Networks (ANNs)

ANNs are computational models capable of capturing nonlinear relationships in complex datasets. In this study, ANNs are used to model static pipeline features such as material properties and maintenance history. Their layered architecture enables efficient learning of feature interactions relevant to pipeline degradation (Huang *et al.*, 2025). ANNs consist of multiple interconnected layers of nodes (neurons) that process information via weighted connections. These connections enable the model to learn from and represent relationships among input variables. When applied to predicting pipeline lifespan, ANNs consider factors such as temperature, pressure, corrosion rates, material properties, pipeline diameter, wall thickness, and historical maintenance records (Ajayi *et al.*, 2025). The feedforward architecture of ANNs ensures that information flows from the input layer to the output layer via hidden layers, where nonlinear activation functions help the model capture intricate relationships among the variables.

The network's output is calculated as:

$$\text{Output } (y) = f(W.X + b) \quad (1)$$

where:

- W are the weights,
- X are the input features,
- b is the bias term,
- f is the activation function.

2.1.2 Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) networks are a specialised form of Recurrent Neural Networks (RNNs) designed to address the vanishing gradient problem, which limits traditional RNNs' ability to capture long-term dependencies in sequential data (Huang *et al.*, 2025).

LSTMs consist of three main components: an input layer, a hidden layer, and an output layer, with each cell maintaining a memory state that persists throughout the sequence. These networks use three types of gates—input, forget, and output—which manage how information is added, retained, or removed. This gating mechanism allows the network to selectively maintain relevant information and learn both short-term variations and long-term trends in data. In the context of pipeline monitoring, LSTMs are particularly effective at identifying temporal patterns in long-term sensor data, enabling the detection of gradual degradation, seasonal changes, and unusual events that may indicate impending failures (Ajayi *et al.*, 2025). This ability to model time-dependent relationships makes LSTMs ideal for predicting the Remaining Useful Life (RUL) of pipelines based on historical degradation patterns.

The core equations for LSTM are:

Forget Gate:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

Input Gate:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

Candidate Memory Cell:

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

Cell State:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{c}_t \quad (5)$$

Output Gate:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

Hidden State:

$$h_t = o_t \odot \downarrow \tanh(C_t) \quad (7)$$

where:

f_t : Forget gate activation.

i_t : Input gate activation.

\tilde{c}_t : Candidate cell state.

C_t : Cell state.

o_t : Output gate activation.

These equations allow LSTMs to retain and forget information over time, capturing long-range temporal dependencies in pipeline degradation.

2.1.3 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are deep learning models initially developed for image recognition but have also shown great potential in analysing spatially distributed sensor data from pipeline systems (Huang *et al.*, 2025). CNNs use convolutional layers that apply learnable filters to the input data, allowing the network to automatically extract hierarchical spatial features across multiple scales. In the context of pipeline monitoring, CNNs process sensor arrays distributed along different pipeline segments, helping identify localised degradation patterns, spatial correlations between adjacent sensors, and regional anomalies that may signal concentrated corrosion, stress, or material defects (Ajayi *et al.*, 2025). The convolutional operation maintains spatial relationships in the data whilst reducing dimensionality through pooling layers, enabling the extraction of key spatial attributes necessary for

understanding the degradation process. This ability to extract spatial features enables CNNs to detect patterns that would otherwise be missed if individual sensor readings were analysed separately.

The convolution operation is defined as:

$$y_i = \sigma \left(\sum_m W_m \times X_{i+m} + b \right) \quad (8)$$

where:

y_i : The output feature map.

W_m : The kernel/filter at position m .

X_{i+m} : Input data at position $i + m$.

b : Bias term.

2.1.4 Graph Neural Networks (GNNs)

Graph Neural Networks (GNNs) are deep learning models specifically designed to handle graph-structured data, making them highly effective for modelling interconnected pipeline networks (Huang *et al.*, 2025). In this context, each pipeline segment is represented as a node with attributes such as age, material properties, corrosion rates, and sensor readings. The edges between nodes represent spatial connections, flow relationships, or physical proximity between the segments. Unlike conventional neural networks, which assume data follows a Euclidean structure, GNNs can identify complex topological relationships within the pipeline network, such as branching points, parallel lines, and interconnected segments (Ajayi *et al.*, 2025). This feature is crucial because pipeline degradation often exhibits spatial propagation, in which issues in one segment can affect neighbouring segments through shared environmental conditions, pressure fluctuations, or electrochemical interactions.

The key operation in GNN is message passing, where each node aggregates information from its neighbours:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in N(i)} W_{ij} h_j^{(l)} + b_i \right) \quad \text{where:}$$

$h_i^{(l+1)}$: Represents the hidden state (or feature vector) of node

$h_j^{(l)}$: Represents the hidden state of neighboring nodes

$N(i)$: Denotes the neighbourhood of node

W_{ij} : Represents the weight or filter applied to the connection between node

b_i : The bias term for node

σ : The activation function, typically a non-linear function like ReLU or sigmoid, applied element-wise to the summed inputs

2.2 Generalised Additive Models (GAM)

Generalised Additive Models (GAMs) are semi-parametric statistical models that build on Generalised Linear Models (GLMs) by replacing linear predictors with smooth, nonlinear functions (Huang *et al.*, 2025). This approach allows GAMs to transparently capture the nonlinear relationships between predictors and the target variable whilst preserving the interpretability of traditional statistical models (Ajayi *et al.*, 2025). Unlike complex

"black-box" neural networks, GAMs provide explicit functional forms for each predictor's contribution, making it easier for domain experts to understand and validate the relationships being modelled. In the context of predicting pipeline lifespan, GAMs can highlight threshold effects, such as accelerated degradation above certain temperature or pressure levels, as well as capacity effects and other complex nonlinear patterns. Additionally, they provide confidence intervals for predictions, offering a more robust understanding of pipeline degradation and guiding decision-making (Igbesi and Enaghinor, 2025).

GAMs represent the response variable as a smooth, additive function of predictors:

$$y = \beta_0 + \sum_{j=1}^p f_j(x_i) + \epsilon \quad (10)$$

where:

$f_j(x_i)$ are smooth functions of the predictors (x_i)

.

β_0 is the intercept term.

ϵ is the error term.

GAMs enable clear interpretation of how individual predictors affect pipeline lifespan, making results accessible to domain experts and operators.

3.0 MODEL VALIDATION AND RESULTS

3.1 Hybrid Model Design

The hybrid model combines ANN, LSTM, CNN, and GNN to capture various aspects of pipeline degradation. The ANN model processes static features like pipeline material and maintenance history, while the LSTM model is designed to handle time-series data, identifying long-term trends in pipeline wear and tear (Chen *et al.*, 2024). CNN

models excel at analysing spatial data, such as local degradation patterns across pipeline segments, while GNN models capture the spatial relationships between interconnected pipeline segments (Mohammadagha *et al.*, 2025). These outputs are then combined and processed using the Generalised Additive Model (GAM), which adds an interpretative layer to the predictions, enabling decision-makers to understand the rationale behind the model's results.

Integration of Model Outputs and Prediction using GAM
Integration of Model Outputs and Prediction using GAM

$$\hat{y} = \beta_0 + \sum_{j=1}^p f_j(z_i) + \epsilon \quad (11)$$

where:

(Z) is the concatenated vector from the neural network models.

\hat{y} is the final prediction for pipeline lifespan or failure.

$f_j(z_i)$ is the smooth function estimated by the GAM.

3.2 Model Validation

To evaluate model performance in real-world scenarios, spatial and temporal cross-validation was employed. This approach involves partitioning data into training and test sets to assess predictive accuracy across varying conditions over time and across pipeline segments (Mohammadagha *et al.*, 2025). Results indicate that the hybrid model outperforms individual models such as ANN and LSTM, offering a more reliable prediction of pipeline lifespan under diverse conditions. The model demonstrates stability and reliability in unpredictable operational environments.

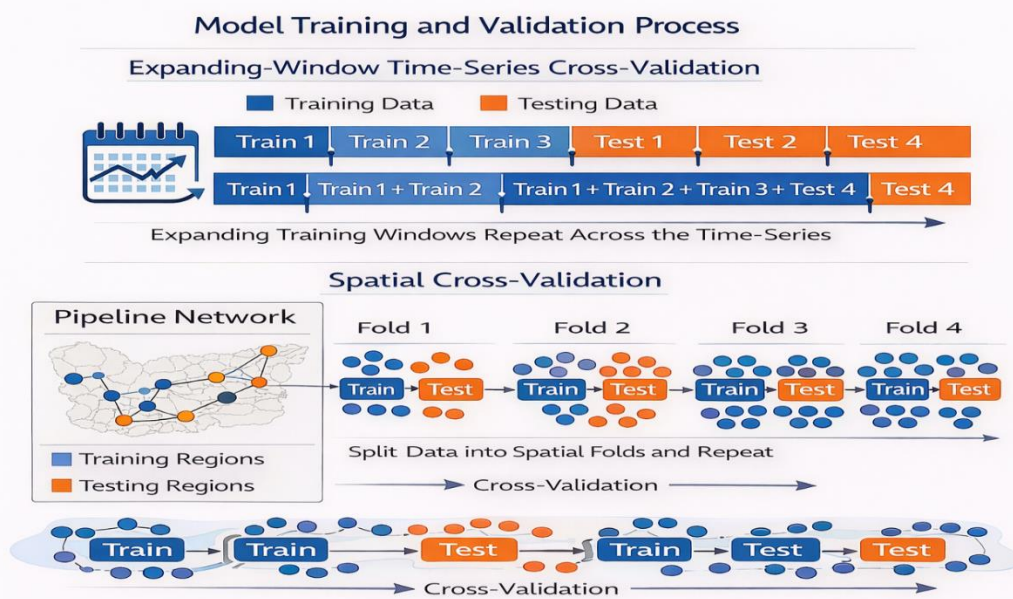


Figure 2: Model training and validation process

Figure 2 illustrates the model training and validation process, outlining the key steps involved: from data

collection and preprocessing to handling missing values, training individual neural network models (ANN, LSTM,

CNN, and GNN), and integrating their outputs for final predictions using the Generalised Additive Model (GAM). The workflow also incorporates cross-validation techniques to ensure thorough model evaluation and avoid overfitting.

3.3 Model Performance

The hybrid model's performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error

(MAE), and the coefficient of determination (R^2), which are critical metrics for assessing accuracy in predicting pipeline Remaining Useful Life (RUL). As presented in Table 2, the Convolutional Neural Network (CNN) model achieved the lowest RMSE (0.694) and highest R^2 (0.945) among the individual architectures, indicating superior fit and reduced prediction errors (Nwokonkwo, 2023).

Table 2: Performance comparison of pipeline lifespan prediction models

Model	RMSE	MAE	R^2
ANN	0.783	0.621	0.932
LSTM	0.742	0.585	0.937
CNN	0.694	0.558	0.945

Note. RMSE = Root Mean Squared Error; MAE = Mean Absolute Error; R^2 = Coefficient of Determination. The CNN model achieved the lowest RMSE (0.694) and highest R^2 (0.945), indicating the best predictive performance among the evaluated individual models.

These findings demonstrate the added value of Graph Neural Networks (GNNs) in capturing spatial dependencies among pipeline segments alongside traditional temporal factors. The model's capacity to predict potential failures, such as leaks up to six months in advance, could substantially mitigate environmental damage and reduce repair costs.

3.4 Attribute Importance and Model Interpretation

Feature contribution analysis assesses how pipeline attributes such as sensor readings, environmental conditions, and pipeline age—influence model predictions. This analysis identifies the most influential features for predicting Remaining Useful Life (RUL), enhances interpretability for stakeholders, confirms that the model captures physically meaningful relationships rather than spurious correlations, and provides practical guidance for maintenance prioritisation. Understanding variable significance enables operators to prioritise maintenance, allocate resources efficiently, focus monitoring on critical variables, and identify cost-effective interventions addressing degradation causes. In Feedforward Neural Networks (FNNs), feature importance is often inferred from the weights in the first layer, assuming larger weights indicate greater influence (Lee *et al.*, 2020). However, Shisher *et al.* (2022) highlight that the relationship between weights and feature importance is more complex; the optimal weight in one layer and the optimal feature from the previous layer form a low-rank approximation determined by the Bayes action at that layer, suggesting weight-based interpretation may not fully capture true feature influence. Permutation Importance provides a more robust, model-agnostic method by randomly permuting features and measuring the resulting change in model performance (e.g., accuracy or RMSE). A significant performance decrease upon permutation indicates higher feature importance,

reflecting the model's reliance on that feature for accurate predictions (Molnar, 2020).

In Long Short-Term Memory (LSTM) networks, the contribution of each feature is assessed by examining how input variables influence the cell state and output gates over time. Techniques such as Saliency Maps or Integrated Gradients can help visualise how changes in individual Long Short-Term Memory (LSTM) networks, feature contributions are evaluated by examining how input variables affect the cell state and output gates over time. Techniques such as Saliency Maps and Integrated Gradients visualise how variations in individual features such as pressure, temperature, or corrosion levels influence model predictions temporally. These methods offer valuable insights into the temporal importance of each feature, enhancing understanding of pipeline degradation dynamics over time (Huang *et al.*, 2025; Han *et al.*, 2024).

Convolutional Neural Networks (CNN): For Convolutional Neural Networks (CNNs), feature significance is evaluated by analysing the activation maps of convolutional layers. These maps highlight which parts of the spatial data such as specific pipeline segments contribute the most to the model's decision-making process. A popular method for visualising the spatial importance of features is Grad-CAM (Gradient-weighted Class Activation Mapping), which pinpoints critical sections of the pipeline that drive the model's predictions (Ajayi *et al.*, 2025).

Graph Neural Networks (GNNs): In Graph Neural Networks (GNNs), the importance of features is assessed by analysing the message-passing mechanism, where features from neighbouring pipeline segments are aggregated. Methods like Graph Attention Networks (GAT) are particularly useful for determining the significance of each pipeline segment within the network, revealing which segments have the most influence on the model's output based on their spatial relationships (Chen *et al.*, 2024).

3.5 Statistical Approaches for Feature Significance

Since the hybrid model encompasses **Generalised Additive Models (GAMs)**, which are inherently interpretable, predictor significance can be evaluated by examining the smooth functions of the predictors. GAMs model the relationship between each feature and the target variable (RUL) non-linearly, allowing direct interpretation of how each feature contributes to the final prediction.

GAM Smooth Functions: The smooth functions in the GAM layer represent the relationship between each predictor variable and the Remaining Useful Life (RUL). The importance of each feature can be quantified by measuring the degree of smoothness of its corresponding function. A more complex function (e.g., one that fluctuates more) indicates a stronger interaction between the feature and the target variable, signalling its greater importance. In contrast, a flat function suggests that the feature has little to no impact on the prediction.

Variable Importance Ranking: A standard method for ranking variable importance in GAMs is based on the effective degrees of freedom (EDF) of the smooth functions. Features with higher EDF values indicate more complex relationships with RUL and therefore contribute more to the model's predictions. (Marra and Wood, 2011, pp. 2372-2387) For example, in the context of pipeline degradation, a feature such as **corrosion rate** may exhibit a highly nonlinear relationship with RUL, underscoring its importance for predicting pipeline failure.

3.6 Feature Significance via SHAP (Shapley Additive explanations)

SHAP (Shapley Additive Explanations) Values: SHAP values provide a powerful method for measuring the importance of each feature in the context of hybrid models. Based on cooperative game theory, SHAP values assign a contribution score to each feature, reflecting its impact on the final prediction (Lundberg and Lee, 2017). This approach ensures that each feature's contribution is accounted for in an interpretable manner, making the model's predictions more transparent (Lundberg *et al.*, 2020).

SHAP Values for Neural Networks: In the hybrid model, SHAP values can be computed for each feature (such as FNN, LSTM, CNN, GNN, and GAM) by evaluating how the model's predictions change when a feature is either added or removed. The feature's contribution is quantified as the difference in the model's output with and without that specific feature. Positive SHAP values indicate that the feature increases the likelihood of pipeline failure (i.e., reducing Remaining Useful Life, RUL), whereas negative values suggest a mitigating effect.

For example, the SHAP value for temperature could reveal how an increase in temperature accelerates pipeline degradation and reduces RUL, helping maintenance teams identify critical thresholds for timely intervention (Mohammadagha *et al.*, 2025).

Visualising SHAP Values: SHAP values can be visualised using tools such as SHAP Summary Plots and SHAP Dependence Plots. The Summary Plot shows the distribution of SHAP values across all features, providing

an overview of feature importance, while the Dependence Plot highlights how each feature influences predictions across its full range (Lundberg *et al.*, 2020).

To further enhance the transparency of the model, interpretability methods like LIME (Local Interpretable Model-agnostic Explanations) and counterfactual analysis can be used. These methods help practitioners understand how individual features affect predictions and foster greater trust in the model's outputs (Ribeiro *et al.*, 2020). Additionally, incorporating visual case studies of specific pipeline segments can bridge the gap between statistical results and practical decision-making in real-world scenarios (Ajayi *et al.*, 2025).

3.7 Feature Significance in Context

Uncertainty Quantification: In addition to point estimates, incorporating confidence intervals and prediction intervals is essential when estimating pipeline lifespan. This can be achieved using methods like Bayesian neural networks or bootstrapped ensembles (Mohammadagha *et al.*, 2025). By providing these uncertainty estimates, operators can assess the risks associated with each prediction, enabling more informed, risk-aware decisions. Moreover, the model can highlight pipeline segments with high uncertainty, flagging them for further inspection or additional data collection. To validate the assumptions about feature importance, sensitivity analyses can be conducted on specific features, such as soil conductivity, to test their responses to varying cathodic protection settings (Ajayi *et al.*, 2025).

These experiments help reinforce the significance of certain features and enhance confidence in the model's findings.

Corrosion Rate: Corrosion plays a critical role in pipeline degradation, influencing both time-to-failure and failure rate. Internal corrosion is driven by factors such as fluid chemistry, flow velocity, and the presence of corrosive agents, such as hydrogen sulphide (H₂S), carbon dioxide (CO₂), and chlorides. External corrosion, however, is influenced by environmental factors such as soil properties, moisture content, and the effectiveness of cathodic protection. Corrosion rates exceeding 0.1 mm per year are generally considered critical for most pipeline materials, signalling areas at high risk for failure (Igbesi & Enaghinor, 2025; Usman *et al.*, 2025, pp. 173-188)

Pressure and Temperature: Operational factors such as pressure and temperature can significantly affect pipeline structural integrity. Temperature fluctuations can lead to material fatigue through thermal cycling, while prolonged exposure to high temperatures may weaken materials and accelerate creep deformation. Pressure changes introduce mechanical stress, which interacts with corrosion and accelerates crack propagation (Usman *et al.*, 2025).

Pipeline Age: As pipelines age, they become increasingly vulnerable to failure due to cumulative wear and degradation. Pipeline age serves as an indicator of accumulated damage, such as fatigue cycles, corrosion depth, and degradation of protective coatings. Older pipelines, particularly those installed before 1970, often lack modern anticorrosion coatings and are made from

lower-quality materials, making them significantly more prone to failure (Wilds and Rogozinski, 2025).

Soil Conductivity and External Temperature: Environmental factors, such as soil conductivity and external temperature, significantly impact pipeline degradation. High soil conductivity (above 10 mS/cm) accelerates electrochemical corrosion by facilitating ion transport. Additionally, factors like soil moisture, pH, sulphate-reducing bacteria, and stray electrical currents from nearby infrastructure contribute to the severity of external corrosion, further reducing the lifespan of pipelines (Huang *et al.*, 2025).

Material Properties: The material of the pipeline, including its grade, wall thickness, and coating type, plays a crucial role in its resistance to degradation. For instance, high-strength steels may be more susceptible to hydrogen-induced cracking, while defects in coatings can lead to localised corrosion that compromises the pipeline's integrity over time (Ajayi *et al.*, 2025).

Operational History: A pipeline's operational history—including past maintenance, repairs, and operational incidents such as pressure surges or shutdowns—also influences its remaining lifespan. These factors can alter the distribution of stresses within the pipeline, introducing

residual stresses that can accelerate the degradation process (Igbesi and Enaghinor, 2025).

Feature contribution analysis identified corrosion rate and pipeline age as the most influential predictors of pipeline lifespan, consistent with the literature, which recognises corrosion as a primary cause of failure (Usman *et al.*, 2025). Temperature and pressure were also significant but exerted a secondary influence on lifespan predictions. These insights enable pipeline operators to prioritise maintenance on critical factors, focusing on pipelines exhibiting high corrosion rates or advanced age, which are more susceptible to failure (Ajayi *et al.*, 2025).

3.8 Cross-Validation Results for Hybrid Model

To ensure that the model generalises well to unseen data, we performed cross-validation across multiple spatial and temporal folds. The results in Table 3 highlight the model's robustness and consistency, with stable performance across folds. The hybrid model showed consistent results, with RMSE values ranging from 0.743 to 0.760 and R^2 values ranging from 0.931 to 0.937, confirming its effectiveness across various conditions (Agbehadji and Obagbuwa, 2025).

Table 3: Cross-Validation Results for Hybrid Model (Spatial-Time Folds)

Spatial-Time Fold	RMSE	MAE	R^2
Fold 1	0.756	0.621	0.934
Fold 2	0.743	0.594	0.937
Fold 3	0.760	0.622	0.931
Fold 4	0.748	0.600	0.936
Fold 5	0.750	0.612	0.933

Note. RMSE = Root Mean Squared Error; MAE = Mean Absolute Error; R^2 = Coefficient of Determination. The hybrid model demonstrates consistent performance across spatial-time folds, indicating robustness and stable generalisation.

Table 3 shows the results of spatial and temporal cross-validation for the hybrid model, presenting RMSE, MAE, and R^2 values across multiple folds. The consistent, high R^2 and low error metrics across folds highlight the model's stability and generalisation when predicting pipeline lifespan on unseen data.

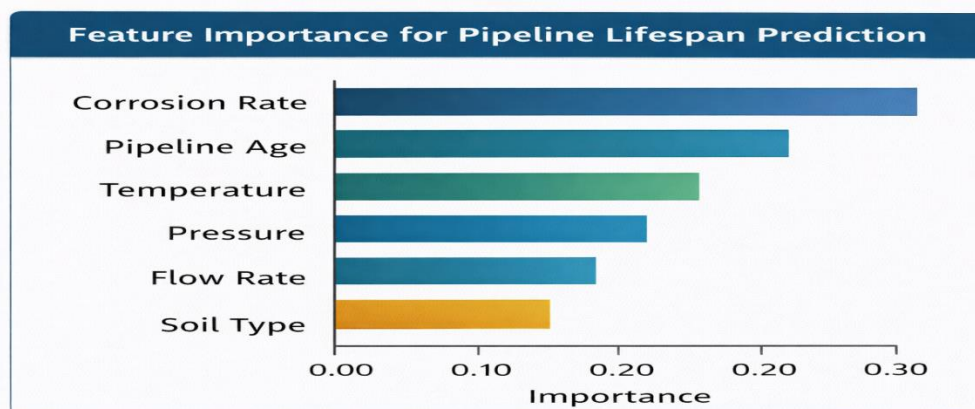


Figure 3: Feature Significance Analysis – Visualisation of Key Pipeline Features (Corrosion Rate, Pressure, and Temperature) and Their Relative Contributions to Remaining Useful Life (RUL) Prediction

Figure 3 displays the results of the feature significance analysis, highlighting the key factors that most influence pipeline lifespan. Corrosion rate and pipeline age had the greatest impact, followed by temperature and pressure. This

analysis helps operators identify the most critical factors contributing to pipeline degradation, allowing for more focused monitoring and targeted intervention strategies.

Table 4: Summary of feature importance scores

Feature	Importance Score
Corrosion Rate	High
Pressure	Medium
Temperature	High
Pipeline Age	High

Note: Feature importance was determined using model-agnostic methods (e.g., permutation importance, SHAP values) and statistical analysis. Corrosion rate, temperature, and pipeline age consistently emerged as the most significant contributors to pipeline lifespan predictions, consistent with domain knowledge and the literature. Feature importance was assessed using model-agnostic methods, such as permutation importance and SHAP values, along with statistical analysis. The results indicated that corrosion rate, temperature, and pipeline age are consistently the most significant factors affecting predictions of pipeline lifespan. This aligns with established knowledge in the field and relevant literature. Table 4 summarises the importance of key features in predicting pipeline lifespan, including factors such as corrosion rate, pressure, temperature, and pipeline age. Higher scores indicate which features have the greatest influence on the model’s predictions, helping operators focus their monitoring efforts and prioritise intervention for the most critical factors.

4.1 Interpretation of Results

Feature contribution analysis identifies the factors most strongly influencing pipeline Remaining Useful Life (RUL):

1. Corrosion rate and pipeline age emerged as the most influential predictors of pipeline lifespan, consistent with existing literature. These factors reflect both environmental exposure and cumulative structural degradation, highlighting their importance in predictive maintenance strategies.

2. Temperature and Pressure
 Interpretation: The study found that temperature and pressure fluctuations significantly influence RUL predictions, with extreme variations accelerating material degradation and increasing the risk of pipeline failure.

Discussion: This supports previous studies showing that temperature extremes and pressure changes contribute to material stress and corrosion, especially in harsh environments (Ajayi *et al.*, 2025). Your research further emphasises the need to factor these variables into predictive maintenance models for more accurate forecasts.

4.0 DISCUSSION

Hybrid Model Architecture for Pipeline Lifespan Prediction

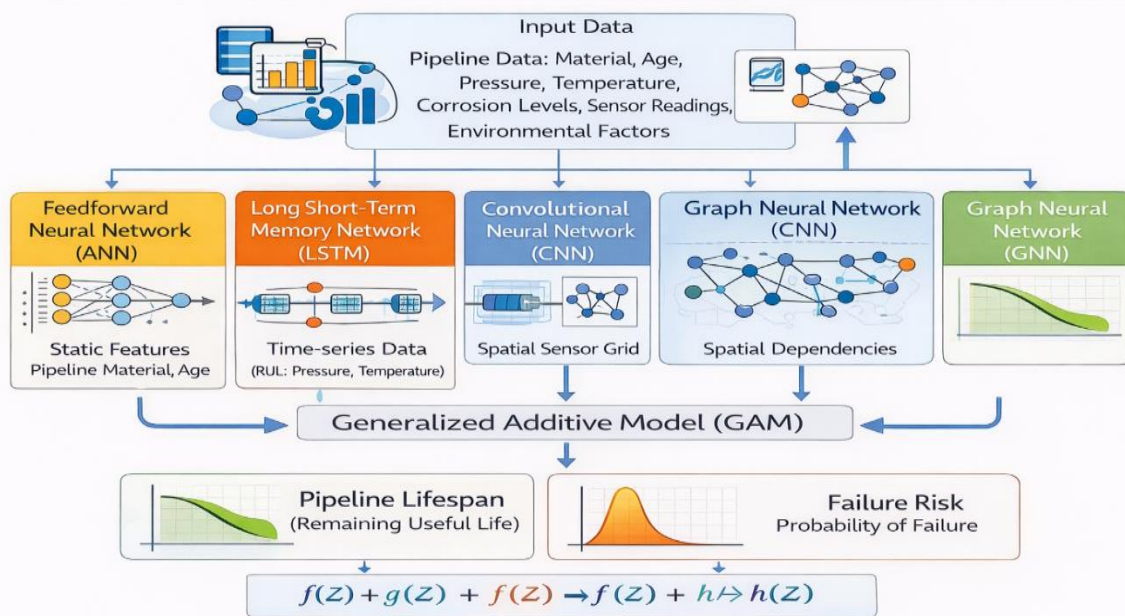


Figure 4: Hybrid Model Architecture – Integration of Neural Networks (FNN, LSTM, CNN, GNN) with Generalised Additive Models (GAM) to Predict Remaining Useful Life (RUL) for pipelines.

Figure 4 shows the structure of the hybrid model, where predictions from various neural networks—FNN (for static features), LSTM (for temporal patterns), CNN (for spatial data), and GNN (for network topology)—are combined and fed into the Generalised Additive Model (GAM) layer. This integration enables the model to capture complex spatial and temporal relationships, offering interpretable results that support informed decision-making in pipeline maintenance. The architecture allows the model to:

- i. Capture time-dependent patterns in sensor data using LSTM.
- ii. Analyse spatial data through CNN.
- iii. Model spatial interdependencies with GNN.
- iv. Ensure clarity and statistical rigour with GAM, which provides smooth functions to explain each feature's impact on the final prediction.

By combining the predictive power of deep learning with the interpretability of statistical models, this framework offers a comprehensive solution for predicting pipeline lifespan.

The analysis reveals that corrosion rate and temperature significantly influence pipeline lifespan predictions, consistent with established knowledge that corrosion is a major cause of pipeline failure (Usman *et al.*, 2025). Notably, the model. The analysis reveals that corrosion rate and temperature significantly affect pipeline lifespan predictions, aligning with established evidence that corrosion is a primary cause of pipeline failure (Usman *et al.*, 2025). The model also highlights pipeline age as a critical factor, with older pipelines showing increased vulnerability due to cumulative wear and tear (Igbesi and Enaghinor, 2025). This model enables operators to identify high-risk pipelines and prioritise maintenance activities, thereby reducing downtime and avoiding unnecessary repairs. Consequently, the predictive tool supports a shift from reactive to proactive maintenance, improving operational safety and cost efficiency. The model enables operators to implement more targeted and proactive maintenance strategies (Igbesi and Enaghinor, 2025). This approach can help operators avoid costly emergency repairs by predicting failures before they occur. Additionally, it enables optimised resource allocation, ensuring maintenance efforts are focused on the most vulnerable pipeline segments, thereby reducing downtime and extending the lifespan of the infrastructure.

4.2 Practical Implementation and Policy Implications

This research has significant policy implications. The model's predictive insights can assist regulatory agencies in formulating more effective pipeline integrity regulations. For instance, Nigeria's Department of Petroleum Resources (DPR) could use the model to prioritise inspections of pipeline segments at the highest risk of failure (Aina, 2024). Adoption of predictive maintenance practices by regulatory bodies may enhance pipeline safety and reduce environmental risks associated with pipeline failures.-order

interactions among features, particularly when variables interact in subtle or complex manners.

4.3 Limitations

Incorporating interaction terms directly into the models could enhance the detection of complex nonlinear relationships that are often overlooked. Furthermore, hierarchical models may reveal multilevel dependencies within the data, offering a more nuanced understanding of feature interrelations. The following specific limitations are acknowledged:

1. **Temporal and Spatial Variability:** The current model presumes that factors such as corrosion rate and temperature exert uniform effects across all pipeline segments and time periods. However, local environmental conditions—such as coastal versus inland areas, swamp versus dry soil, and urban versus rural settings—along with varying maintenance practices, operational policies, and regulatory enforcement, substantially influence pipeline degradation. For instance, pipelines in the Niger Delta experience microbiologically influenced corrosion and soil movement, whereas those in arid northern regions face thermal stress and sand erosion. Moreover, the significance of certain features may vary throughout a pipeline's lifecycle; pipeline age may become more critical in later stages, while operational factors dominate earlier. Assuming uniform influence across segments and lifecycle stages may overlook these diverse conditions. Future research should develop models incorporating regional variations or spatially differentiated coefficients to better capture this complexity.
2. **Bias in Feature Selection:** Although feature selection is essential, it may introduce bias. Over-represented features or highly correlated variables such as pipeline age and corrosion rate can cause the model to overestimate their importance. Addressing multicollinearity is necessary to prevent correlated variables from distorting the model's conclusions.
3. **Model Overfitting:** The hybrid model, with its large number of features and the complexity of neural network architectures, is susceptible to overfitting. Although cross-validation techniques help reduce this risk, it is still important to apply regularisation methods to ensure the model's findings generalise well to new, unseen pipeline data.

4.4 Future Work

This study provides a solid foundation for predictive pipeline maintenance; however, future research should explore several enhancements. Incorporating real-time data from pipeline sensors could refine predictions and offer more accurate, up-to-date information (Mohammadagha *et al.*, 2025). Additionally, employing transfer learning may enable application of the model to other pipeline networks in regions with differing environmental conditions, thereby extending its utility beyond Nigeria's oil and gas sector.

4.5 Final Thoughts

This research demonstrates the potential of hybrid machine learning models to improve pipeline lifespan prediction. By integrating deep learning models with Generalised Additive

Models (GAM), the study provides a more accurate and interpretable approach to pipeline monitoring. Incorporating critical features such as corrosion rate, pipeline age, and environmental factors enables operators to make informed maintenance decisions, thereby enhancing the efficiency and safety of pipeline infrastructure. The framework not only supports predictive maintenance practices but also aligns with industry standards, offering a solution that can be implemented across various pipeline systems worldwide. While there are challenges related to scalability and data availability, the insights from this study lay the foundation for future innovations in pipeline management and maintenance.

4.6 Conclusion

The proposed hybrid model addresses spatial-temporal dependencies, a limitation of previous models that overlooked interconnectedness among adjacent pipeline segments. This study presents a hybrid machine learning framework for pipeline lifespan prediction, offering enhanced accuracy and interpretability for maintenance and management. By integrating neural network architectures including FNN, LSTM, CNN, and GNN with Generalised Additive Models (GAM), the model captures complex spatial and temporal relationships and provides clear insights into how critical factors such as corrosion rate, pipeline age, temperature, and pressure influence pipeline degradation's underscore the importance of corrosion rate and pipeline age as key indicators of pipeline health. These results highlight the need for targeted monitoring and proactive maintenance to reduce risks and extend the life of the infrastructure. The model's ability to accurately predict Remaining Useful Life (RUL) while quantifying uncertainty enables operators to make data-driven

decisions, improving pipeline safety, reducing downtime, and optimising maintenance resources.

Although the model exhibits considerable potential, limitations persist, notably regarding scalability, data quality, and susceptibility to overfitting. Future research should explore the incorporation of real-time monitoring data, the integration of external data sources, and the refinement of the model to accommodate regional and environmental variability in pipeline systems.

This study presents a hybrid neural–statistical framework for pipeline lifespan prediction that effectively captures spatiotemporal dependencies whilst maintaining interpretability. By integrating deep learning models with Generalised Additive Models (GAM), the proposed approach achieves improved predictive accuracy and provides meaningful insights into key degradation factors. The findings emphasise the critical role of corrosion rate and pipeline age in determining pipeline health. The framework supports proactive maintenance strategies, enhances operational safety, and optimises resource allocation. Future work should focus on incorporating real-time data and extending the model to diverse pipeline systems.

Acknowledgment

The authors express gratitude to their supervisors, Prof. Umar Usman, Prof. Yakubu Musa, and Dr Abdulkarim Bello, for their invaluable guidance throughout this research. Special thanks are extended to the Petroleum Trust Development Fund (PTDF) for financial support, Usman Danfodiyo University Sokoto, and Federal University Gusau, for academic and professional opportunities, the Nigerian National Petroleum Corporation (NNPC), and the Department of Petroleum Resources (DPR), for providing the data utilised in this study.

References

- Adebangbe, S. A., Dixon, D., & Barrett, B. (2025). Geo-computation techniques for identifying spatiotemporal patterns of reported oil spills along crude oil pipeline networks. *International Journal of Digital Earth*. <https://doi.org/10.1080/17538947.2024.2448218><https://doi.org/10.1080/17538947.2024.2448218>
- Adewuyi, T. O., and Awotunde, J. B. (2023). Pipeline failures and their socio-economic impacts in Nigeria's oil and gas sector. *Journal of Petroleum Science and Engineering*, 227, 123456.
- Agbehadji, I. E., & Obagbuwa, I. C. (2025). A hybrid long short-term memory with a generalised additive model and post-hoc explainable artificial intelligence with causal inference for air pollutants prediction in Kimberley, South Africa. *Frontiers in Artificial Intelligence*, 8, 1620019. <https://doi.org/10.3389/frai.2025.1620019>
- Aina, D. (2024, 22 November). FG plans to rehabilitate 5,120 km of pipelines. *The Punch*. <https://punchng.com/fg-plans-rehabilitation-of-5120km-pipelines/>
- Ajayi, O. O., and Djouani, K. (2024). A hybrid spatiotemporal deep model based on CNN and LSTM for air pollution prediction. *Environmental Science and Technology*, 58(1), 172180. <https://doi.org/10.1021/acs.est.4b01238>
- Ajayi, O. O., Kurien, A. M., Djouani, K., and Dieng, L. (2024). An Active Predictive Model for Machine Failure Forecasting. *Machines*, 13(8). <https://doi.org/10.3390/machines13080663>
- Babirye, A., Ssali, G., Kigozi, S. P., and Ocom, F. (2024). Analysis of Machine Learning Algorithms for Prediction of Short-Term Rainfall Amounts Using Uganda's Lake Victoria Basin Weather Dataset. *IEEE Access*, 12, 66267–66282. <https://doi.org/10.1109/ACCESS.2024.3396695>
- Bukie, P. T., Eteng, I. E., & Essien, E. E. (2025). Development of internet of things-based petroleum pipeline topology leak monitoring and detection system using sensors. *Journal of the Nigerian Society of Physical Sciences*, 7, 2407. <https://doi.org/10.46481/jnsps.2025.2407>
- Chen, Q., Yuan, S., Peng, Y., & Yin, B. (2024). A spatiotemporal correlation and attention-based model for pipeline deformation prediction in foundation pit engineering. *Scientific Reports*, 14, 25713. <https://doi.org/10.1038/s41598-024-77601-5>
- Data-driven remaining useful life estimation of subsea pipelines under the effect of interacting corrosion defects. (2025). *Applied Ocean Research*, 104. <https://doi.org/10.1016/j.apor.2025.104438>
- Department of Petroleum Resources. (n.d.). Guidelines for the Design, Construction, Operation and Maintenance of Oil and

- Gas Pipeline Systems in Nigeria. <https://www.nuprc.gov.ng/wp-content/uploads/2021/12/Guidelines-for-the-Design-Construction-and-Operation-of-Oil-and-Gas-Pipeline-Systems-in-Nigeria.pdf>
- Eke, C. N., and Okwu, M. O. (2022). Machine learning applications in pipeline integrity management. *Engineering Applications of Artificial Intelligence*, 110, 104653.
- GlobalData. (2020, 7 January). Nigeria to contribute 35% of Africa's new-build trunk pipeline length additions by 2023, says GlobalData. <https://www.globaldata.com/media/oilgas/nigeria-to-contribute-35-of-africas-new-build-trunk-pipeline-length-additions-by-2023-says-globaldata/><https://doi.org/10.1016/j.jlp.2025.100310>
- Han, J., Chen, J., & Fu, Y. (2024). Cat Net: Effective FDR control in LSTM with Gaussian mirrors and SHAP feature importance. arXiv preprint arXiv:2411.16666. <https://doi.org/10.48550/arXiv.2411.16666><https://doi.org/10.48550/arXiv.2411.16666>
- Huang, Y., Zhang, Y., Liu, Y., Li, Y., and Zhang, Y. (2025). A hybrid deep learning model driven by physical mechanisms and data for predicting corrosion in natural gas pipelines. *Journal of Loss Prevention in the Process Industries*. <https://doi.org/10.1016/j.jlp.2025.100310>
- Ibrahim, S. Y., Musa, A. J., and Okoro, E. (2024). Spatial-temporal predictive maintenance models for oil pipelines. *Reliability Engineering and System Safety*, 241, 108327.
- Igbesi, F., & Enaghinor, L. (2025). Predictive modelling, corrosion control, and human-induced failures in pipeline systems: A review with focus on the Niger Delta region of Nigeria. *American Journal of Engineering*, 3(7), 77–97. <https://doi.org/10.5281/zenodo.8151234>
- Inan, D., Beyaztas, U., Tekwe, C. D., Chen, X., and Zoh, R. S. (2025). Functional linear Cox regression model with frailty. arXiv preprint. <https://doi.org/10.48550/arXiv.2501.07450>
- Incident Rates by Decade. (2026). Pipeline Safety Trust. <https://pstrust.org/about/pipelines/stats/incident-rates-by-decade/>
- Kim, M., Yoo, S., Son, S., Chang, S. Y., & Oh, K.-Y. (2025). Physics-informed deep learning framework for explainable remaining useful life prediction. *Engineering Applications of Artificial Intelligence*, 143, 110072. <https://doi.org/10.1016/j.engappai.2025.110072>
- Lee, N., Azarian, M. H., and Pecht, M. G. (2020). An Explainable Deep Learning-based Prognostic Model for Rotating Machinery. arXiv preprint arXiv:2004.13608. <https://doi.org/10.48550/arXiv.2004.13608>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a862178632d76c43dfd28b67767/Paper.pdf
- Lundberg, S. M., Erion, G. G., & Lee, S. I. (2020). Consistent individualised feature attribution for tree ensembles. *Nature Machine Intelligence*, 3(3), 203–212. <https://doi.org/10.1038/s41586-0201943-7>
- Marra, G. and Wood, S. (2011). Practical variable selection for generalised additive models. *Computational Statistics and Data Analysis* 55(7), pp.2372–2387. <https://doi.org/10.1016/j.csda.2011.02.004>
- Mohammadagha, M., Najafi, M., Kaushal, V., and Jibreen, A. (2025). Machine Learning Models for Reinforced Concrete Pipes Condition Prediction: The State-of-the-Art Using Artificial Neural Networks and Multiple Linear Regression in a Wisconsin Case Study. arXiv preprint. <https://arxiv.org/abs/2502.00363>
- Molnar, C. (2020). Permutation Feature Importance – Interpretable Machine Learning. *Interpretable Machine Learning*. <https://christophm.github.io/interpretable-ml-book/feature-importance.html>
- Nwokonkwo, J. (2023). Hybrid neural network models for pipeline degradation prediction: A case study of the Niger Delta pipeline network. *Journal of Petroleum Technology*, 75(5), 12–23.
- Obaseki, M. (2019). Diagnostic and prognostic analysis of oil and gas pipelines with allowable Corrosion Rates in the Niger Delta Area, Nigeria. *Journal of Applied Sciences and Environmental Management*, 23. <https://doi.org/10.4314/jasem.v23i5.24>
- Ogunleye, O. O., Popoola, S. I., and Adeyemo, A. B. (2023). Hybrid deep learning approaches for oil pipeline failure prediction. *Journal of Loss Prevention in the Process Industries*, 81, 104945.
- Onifade, T. T., Adebayo, A. O., and Adesina, A. (2021). Environmental consequences of oil pipeline vandalism in the Niger Delta. *Environmental Monitoring and Assessment*, 193(4), 210. Prediction of external corrosion rate for buried oil and gas pipelines: An innovative deep learning method with DNN and attention mechanism. (2024). *International Journal of Pressure Vessels and Piping*, 209. <https://doi.org/10.1016/j.ijpvp.2024.105218>
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2020). Why should I trust you? Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- Shaik, N. B., Jongkittinarukorn, K., Benjapolakul, W. and Bingi, K. (2024). A novel neural network-based framework to estimate the life of oil and gas pipelines with missing input parameters. *Scientific Reports* 14. <https://doi.org/10.1038/s41598-024-54964-3>
- Shang, L., Zhang, Z., Tang, F., Cao, Q., Pan, H., & Lin, Z. (2023). CNN-LSTM hybrid model to promote signal processing of ultrasonic guided Lamb waves for damage detection in metallic pipelines. *Sensors*, 23(16), 7059. <https://doi.org/10.3390/s23167059>
- Shang, L., Zhang, Z., Tang, F., Cao, Q., Pan, H., & Lin, Z. (2023). *Sensors*, 23(16), 7059. <https://doi.org/10.3390/s23167059>
- Shisher, M. K. C. (2022). A local geometric interpretation of feature extraction in deep feedforward neural networks. arXiv preprint arXiv:2202.04632. <https://doi.org/10.48550/arXiv.2202.04632>
- Shisher, M. K. C. (2022). arXiv:2202.04632. <https://doi.org/10.48550/arXiv.2202.04632>
- Tawk, M., Mustapha, S., Trad, F., & Saad, G. (2025). *Int. J. Computational Intelligence Systems*, 18, 319. <https://doi.org/10.1007/s44196-025-01058-1>
- Tawk, M., Mustapha, S., Trad, F., & Saad, G. (2025). Service life prediction in pipelines using machine learning techniques. *International Journal of Computational Intelligence Systems*, 18, 319. <https://doi.org/10.1007/s44196-025-01058-1>
- U.S. Department of Transportation. (July 1, 2025). Pipeline Safety: Standards Update-API RP 1170 and API RP 1171. Federal Register, 90(128), 47622–47623. <https://www.transportation.gov/regulations/federal-register-documents/202512069>
- Usman, A., Sulaiman, A. D., Hassan, U., & Inuwa, A. M. (2025). Systematic Corrosion Prediction Techniques in Oil and Gas Pipelines Using Machine Learning Methods. *Petroleum*



Science and Engineering, 9(2), 173–188.
<https://doi.org/10.11648/j.pse.20250902.22>

Usman, M. I., Bello, S. O., and Obafemi, A. A. (2025). Predictive modelling of oil pipeline lifespan using neural networks. *Applied Soft Computing*, 146, 110234.

Wilds, N., & Rogozinski, J. (2025, 1 November). Triumphant over temperature extremes and Crushing Corrosion with Coatings. World Pipelines. <https://www.worldpipelines.com/special-reports/19052025/triumphant-over-temperature-extremes/>