



Journal of Frontiers in Multidisciplinary Research

A Machine Learning–Enhanced Model for Predicting Pipeline Integrity in Offshore Oil and Gas Fields

Dulo Chukwuemeka Wegner ^{1*}, Adumike Kenechukwu Nicholas ², Ojoh Odoh ³, Kehinde Ayansiji ⁴

¹⁻³ University of Hull, United Kingdom

⁴ Hydrography, University of Plymouth, United Kingdom

* Corresponding Author: **Dulo Chukwuemeka Wegner**

Article Info

E-ISSN: 3050-9726

P-ISSN: 3050-9718

Volume: 02

Issue: 02

July – December 2021

Received: 04-09-2021

Accepted: 06-10-2021

Published: 07-11-2021

Page No: 331-342

Abstract

There are several difficulties that come with maintaining the structure and integrity of subsea pipelines. Some of these difficulties are caused by the extreme weather conditions and other issues during operations, which greatly speed up the erosion of the pipeline. These pipelines are of extreme importance when it comes to extracting oil and gas, yet the breakdown of the pipes is often overlooked. While traditional risk assessments or erosion control measures are important, they tend to focus on reactive measures to slow given periods of erosion, instead of predictive measures. This navigates predicting pipeline evaluations with the help of an advanced machine learning structure. The model uses Remote Operated Vehicle Inspection Reports, Cathodic Protection Survey Reports, and Maintenance Registries to produce an informed estimate with regard to disintegration, such as loss of armor, coating, and corrosion fatigue. Supervised learning Support Vector Machines and Random Forest techniques are applied to provide a more informed revolutionary learning experience, as opposed to traditional learning methods. The predictive model relies on empirical assessments that are measured in vast quantities in sea pipelines and are compared with ground truths to be validated. The design of the model uses sets of features that are relevant to sea pipes. The results demonstrate the novel model adds value to pre-existing methods used by risk, as it relies on the machine learning process, predicting more accurately sooner and with less inaccuracies. In addition, the analysis of the importance of individual features associated with the seabed topography, cath protection potential, and operating pressure for their value in proactive maintenance presents actionable insights and contextual value. This shows the value of machine learning in moving the management of pipeline integrity from a reactive to a more proactive approach, which in turn, improves reliability, minimizes downtime, and reduces inspection costs. This framework enhances the ability for decision-making based on evidence while fulfilling the core aim of deep-water energy production, which is the utmost safety and ecological integrity. The focus of the next stage of the project will be on the deployment of the model to varying offshore conditions and incorporating real-time data from sensors for interval-based assessment.

DOI: <https://doi.org/10.54660/IJFMR.2021.2.2.331-342>

Keywords: Pipeline integrity, Offshore oil and gas, Machine learning, Predictive modeling, Structural health monitoring, ROV inspection

1. Introduction

Offshore oil and gas activities are among the most technologically advanced and most capital-demanding activities in the energy sector (Ericson *et al.*, 2019; Castiblanco, 2021). These activities depend on a network of subsea pipelines, the primary system for transporting gas and oil from production wells to processing and storage facilities, as well as export terminals. These pipelines are usually in remote, ecologically sensitive areas. Their uninterrupted, reliable operation is necessary for profitability and conservation (Hacquard *et al.*, 2019; Mäkitie, 2020). It is of national importance and needs to stress the importance of subsea pipeline infrastructure: any disruption in the integrity of the pipeline not only threatens energy security, but also increases operation costs and has the potential for serious environmental damage (Bordoff *et al.*, 2018; Lakshminarayanan, 2019).

It follows that safeguarding the subsea pipeline infrastructure has emerged as a central focus in the engineering and operation of offshore oil and gas projects (Omoya Papadopoulou, 2019; Cherepovitsyn *et al.*, 2020).

Even though offshore pipelines are crucial, they are frequently subjected to conditions that speed up deterioration and increase the risk of failures (Drumond *et al.*, 2018; Khalid *et al.*, 2020). Corrosion, particularly when it comes to marine and saline environments, remains the greatest risk to both internal and external pipelines as it severely hinders the efficacy of electrochemical reactions. 'Passive' strains from shifting seabeds, vessel activities, and even installation defects have even more mechanical stress. Additionally, low temperatures and extreme sub-sea pressure add to the more extreme operational conditions, which work against materials and exacerbate their fatigue (Zhang *et al.*, 2020; Ho *et al.*, 2020). Over time, the amplification of these factors increases the risk of ruptures, leaks, or structural collapse, which all have grave safety and environmental consequences (Qin *et al.*, 2020; Adamo *et al.*, 2020). The inability to monitor advanced systems and the offshore operations, which elevated their visibility due to highly publicized pipeline incidents, has rallied focus to the advanced methods of predictive integrity management, and the sublime importance of it (Stark, 2018; Nilsson and Christensen, 2019).

Negative proof supporting current monitoring and inspection procedures, such as remotely operated vehicle (ROV) inspections and ultrasonic testing, are time-consuming and unrewarding inspection procedures that document pipe conditions while ignoring degradation evolution (Shen *et al.*, 2018; Mattioli *et al.*, 2019). Moreover, damage accumulation is often only reported after extensive damage has occurred, as there are no proactive measures in place to counter damage, a situation extremely insufficient due to few unmonitored inspection procedures. Interpretation subjectivity in Degnan *et al.* 2019; Gerry *et al.* 2020 is a common limitation in most inspection reports, as humans corresponding, responsible for finding minor deterioration signs, lean towards automation matrix, ignoring maximal processes. This lag to Revica, reluctance to utilize current business processes, is associated with composition gaps in predicted failures and uncalculated operational risks that are to be accounted for. Parker *et al.* 2018 and Biggio and Kastanis 2020 shed light regard business critical gaps in unmonitored validated offshore systems. Decision-making systems are absent in place and are predicted for predicted future, thus accurate deviation measurements that can track changes in erosion surfaces are critical.

Addressing this gap demands the application of different technologies that can interpret inspection results and provide insight in predictive form. Machine learning can take deep learning models and utilize them to solve the advanced analytics process of corrosion and defect reporting (Ahmad *et al.* 2018; Goldsmith *et al.* 2018). Complex models can be designed and trained to analyze inspection data, historical records of failures, and the environmental surroundings to predict failures before they escalate and critically damage the pipeline. Machine learning models, unlike rule-based systems, grow in predictive accuracy the more data they are trained with (Banda *et al.* 2018; Islam *et al.* 2020). This means that Machine learning models will refine their predictions over time through repeated training, leading to a far more accurate model. This detachment from defined, inspection-based methods machine learning model is

predictive and offers a means of adaptive integrity monitoring.

This machine learning-enhanced model will utilize ROV inspection data along with cathodic protection survey data to predict offshore pipeline degradation. ROVs equipped with video and sensor data collection capabilities are routinely used in subsea inspection operations and provide direct visual and quantitative evidence of pipeline condition with high-resolution videos, images, and other datasets. Cathodic protection surveys, on the other hand, provide electrochemical data depicting the effectiveness of corrosion-prevention systems along with crucial external corrosion indicators. The goal of the model is to provide predictive integrity assessments, as opposed to just predictive reporting, by combining the aforementioned data streams within the machine learning framework. In other words, the machine learning framework will determine the external corrosion risk condition and, in turn, forecast pipe degradation hotspots. Moreover, the model will predict failure potential and recommend optimum maintenance windows to minimize operational and maintenance intervention downtimes.

This model advances this response to both technical and environmental issues. Predictive integrity assessment, from the technical side, minimizes outages, resource wastage, and prolongs the lifespan of a pipeline. On the environmental side, it minimizes the chances of leaks and spills that can ruin marine life and attract regulatory involvement (Olawuyi, 2018; Byrnes and Dunn, 2020). At the crossroads of offshore engineering, data science, and environmental protection, this piece of research enriches the desire to improve the safety and sustainability of offshore oil and gas structures.

2. Methodology

To ensure transparency, reproducibility, and rigor in the synthesis of evidence regarding the use of machine learning-enhanced models for predicting pipeline integrity in offshore oil and gas fields, the methodology adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach. A thorough literature search was carried out across several academic and technical databases, including Scopus, Web of Science, IEEE Xplore, ScienceDirect, and Google Scholar, as part of the identification step of the review process. Boolean operators were used to filter and widen the results of the search by combining keywords like "machine learning," "pipeline integrity," "offshore oil and gas," "predictive modeling," and "structural health monitoring." Additional sources were found by tracking the backward and forward citations of pertinent papers in order to lessen bias.

During the screening phase, duplicates were eliminated, and titles and abstracts were assessed for relevance in relation to predetermined inclusion and exclusion criteria. Studies that offered hybrid computational techniques or machine learning models applied to corrosion prediction, leak detection, pipeline integrity management, or failure risk assessment in offshore environments were accepted. Studies unrelated to oil and gas pipelines, those lacking a predictive modeling component, works that were solely theoretical and lacked implementation data, and publications written in languages other than English were also excluded.

Full-text evaluations of a few chosen papers were necessary for the eligibility step. Studies that did not give evaluation measures, did not offer insights relevant to offshore pipeline systems, or lacked adequate methodological detail were

further disqualified. Model types (e.g., neural networks, support vector machines, random forests, ensemble learning), input parameters (e.g., operational data, sensor readings, environmental conditions, historical failure data), datasets used, validation methods, and performance indicators (e.g., accuracy, precision, recall, and area under the ROC curve) were all recorded during the methodical data extraction process.

Studies that satisfied all requirements were combined into the synthesis during the last inclusion step. To determine trends in model development, the function of feature engineering, the combination of sensor-based monitoring and predictive analytics, and the efficacy of models in actual offshore environments, a qualitative and quantitative evaluation was conducted. Limitations mentioned in the research, including a lack of data, problems with sensor reliability, and difficulties extrapolating models to a variety of offshore situations, were also taken into account in the synthesis. The PRISMA flow reduced selection bias and improved methodological transparency by guaranteeing an organized process from identification to synthesis. By enabling proactive maintenance, lowering failure risks, and improving operational safety, the results offered a solid body of evidence showing how machine learning-enhanced models may improve pipeline integrity management in offshore oil and gas sectors.

2.1. Literature Review

Pipeline integrity management emphasizes the importance of offshore operations, as the impact of structural breakdowns can be dire. The pipelines function as the main avenue for the transportation of hydrocarbons from the subsea reservoirs to the surface facilities. Their failure can lead to catastrophic environmental damage, overwhelming costs, and safety issues (Kenny, 2018; Chernov and Sornette, 2019). In the longer term, practitioners and scholars have developed several frameworks for analyzing, monitoring, and protecting the integrity of subsea pipelines. These frameworks have transformed from simplistic approaches to a technologically advanced machine learning predictive modeling paradigm that radically shifts the enhancement of reliability and structural integrity.

From the beginning, maintaining pipeline integrity has been about checking the pipeline. ROVs and AUVs are used primarily for visual checks of the external pipeline and coating surveys, and for probing protective coating for marine growth. Along with ultrasonic testing and magnetic flux leakage, they are also used in subsea surveys for wall corroded sections, checking for internal corrosion, and corrosion bursts. Maybe important, but they still come with an expectation of exclusivity, with a high price, and offshore support being shallow, which in turn necessitates the support of operator intelligence for interpretation.

Another important aspect that remains is integrity. For example, offshore pipelines are highly susceptible to internal corrosion from multiphase flow conditions and external corrosion due to seawater exposure. Having corrosion coupons, electric resistance probes, and protective surveys to guard and assess corrosion rate and protective performance is one class of techniques out of many that have been guiding the way (after Groysman and Wright, 2019). But with the use of static frames of frameworks also comes limitations, especially in regards to capturing locals and the variation endurance across long pipeline stretches.

Moreover, there are empirical and semi-empirical models that attempt to predict degradation and failure probabilities. The corrosion rate equation and probabilistic fracture mechanics provide an underpinning theoretical basis to estimate integrity risk with respect to operational and material parameters (of) corrosion (of the) material(s) (Larrosa *et al.*, 2018; Wu *et al.*, 2019). These models are advantageous in risk-based inspection planning; however, the effectiveness of the models is limited due to assumptions that, due to oversights of nonlinear and dynamic interactions, are common in real subsea environment interactions.

In the past 20 years, there has been an increase in predictive modeling as industries are shifting from reactive to proactive and preventive maintenance. Infrastructure reliance has focused on the usage of stochastic modeling, Bayesian updating, and finite element analysis (FEA) to simulate degradation and estimate remaining useful life (RUL). In the case of pipeline systems, stochastic processes are used to estimate the uncertainty of corrosion rates, while FEA allows the engineer to model the stress distribution to fatigue within dynamic loading systems from seabed movement and internal pressure fluctuations (Teply *et al.*, 2018; Razvarz *et al.*, 2020).

Moving past deterministic methods, probabilistic risk assessment (PRA) remains a pivotal aspect of evaluating reliability. PRA frameworks formulate risk-based maintenance plans by combining inspection records with statistical models of degradation rates. While such frameworks excel at allocating inspection resources, their dependence on historical information, coupled with expert-derived probability functions, stifles performance in intricate, information-dense settings (Asghar *et al.*, 2018; Marrec *et al.*, 2020).

Recent advances have focused on data-driven predictive frameworks, where sensor networks and digital twins are employed to approximate operational conditions in real time. In particular, digital twin technology has enabled continuous monitoring by virtually replicating pipeline systems equipped with the capacity to simulate degradation under a myriad of conditions (Konchenko *et al.*, 2020; Therrien *et al.*, 2020). However, the computational costs associated with the scalability of these models still remain an ongoing issue.

The emergence of machine learning (ML) in the oil and gas industry has signified a digital transformation by facilitating the utilization of diversified and extensive datasets. In the context of pipeline integrity management, ML models are capable of anticipating outcomes by synthesizing ROV inspection imagery, cathodic protection data, flow information, and environmental variables.

Supervised learning techniques have demonstrated particular promise among machine learning approaches. Support Vector Machines (SVMs), Random Forests (RFs), and Artificial Neural Networks (ANNs) are used in oil and gas operations for classification of failure modes, prediction of degradation rates, and forecast of equipment failures (Ao *et al.*, 2018; Yahşi *et al.*, 2019). For example, SVMs have been used to recognize normal and abnormal operating conditions of subsea pipelines. Random Forests make strong predictions by combining many decision trees, thereby reducing overfitting and improving prediction interpretability. ANNs are very useful in learning complex relationships in a dataset that involves operational variables and degradation outcomes, but are also very expensive computationally.

Machine learning applications in the oil and gas industry are

not limited to pipeline integrity. In reservoir engineering, ML has been applied to predict production rates, optimize drilling parameters, and model fluid dynamics. In health, safety, and environment (HSE) management, predictive models estimate the probabilities of accidents and improve emergency response plans (Ajayi *et al.*, 2019; Jensen, 2019). These successes demonstrate the wider applicability of ML techniques and support their use for pipeline integrity management.

Over the years, the predictive performance of supervised learning algorithms has been enhanced through feature importance ranking and SHAP (Shapley Additive explanations) values, which help engineers pinpoint the most crucial components of integrity risks (Elshawi *et al.*, 2019; Linardatos *et al.*, 2020). This, as a result, augments the precision in forecasting and assists in the strategic formulation of predictive maintenance plans through evidence-based recommendations.

There is a linear progression in the management of pipeline integrity in the literature, beginning with inspection-centric approaches of measurement and physical observation, extending to empirical and probabilistic models formulated to synthesize and generalize degradation processes, and now to the contemporary focus on machine learning–driven models that maximize predictive accuracy from large-scale datasets. Although traditional methods are still essential for legal obligations and ground truth validation, their value is enhanced by predictive models as the need for foresight in resource management and allocation grows. At this juncture, machine learning provides the most transformative advancement in offshore oil and gas maintenance planning, driving down uncertainty and optimizing the maintenance schedule to enhance operational safety (Atri, 2018; Angelopoulos *et al.*, 2019).

2.2. Data Sources and Feature Engineering

The ghastly conditions offshore have not aided the development of a machine learning enhanced model for predicting the pipeline integrity of the oil and gas offshore operations. The machine learning model remains undeveloped because of the most critical aspect, the diversity and quality of the input data. Offshore pipelines exist within a multilayered and complex environment. Within which pipelines undergo a series of degrading processes as shown in Figure 1 (Cai *et al.*, 2019; Adumene *et al.*, 2020). Thus, the models still require a myriad of complex datasets that capture the primary attributes regarding the pipeline and the overarching operational environment. Some such fundamental sources of data include inspection outputs from the ROV, results from cathodic protection surveys, and other operational and environmental data. However, the ROV inspection remains arguably the most important and insightful component of data processing. Data processing requires the ROV inspection data to stream the raw data outputs into pertinent model inputs, which involves advanced feature engineering. Unprocessed measurements need machine learning algorithms in order to draw: unmeasured outputs and predictions. The ROV inspection data features main components utilized in other data and unprocessed streams, which emerge from different features to explain model predictions.

Among these ROVs, the ones that are used in the ROV integrations multiplayer are more than ROVs as they are equipped with high-definition cameras, monitors, and

structural/sonar ROV inspection systems located within and external to a survey area (Forsyth, 2018). ROVs are skilled in providing invisible structural and systems that traditional means of inspection cannot access (Primeau, 2019). The frameworks of ROVs with monitors that can capture them during operations have provided circumstantial evidence to the predictive model and structural ROVs. Many of these predictions relate to the predictive modeling of systems.

First, visual imaging provides continuous footage and stills that capture the external state of the pipeline surfaces. These images can show surface discoloration, protective coating degradation, and localized fouling by marine organisms. Image-processing methods, such as convolutional neural networks (CNNs), can be trained to find features associated with the onset of corrosion, such as surface roughness, texture, and pixel intensity (Vijayalakshmi, 2020; Rifai *et al.*, 2020).

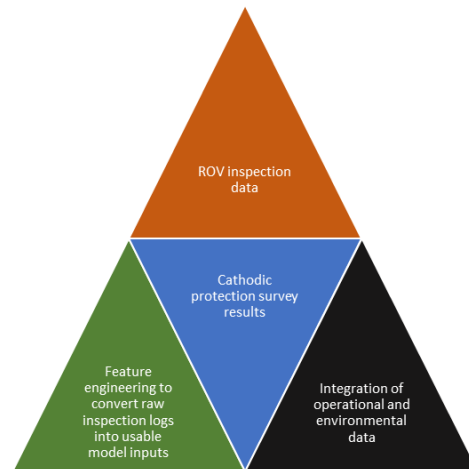


Fig 1: Data Sources and Feature Engineering

Second, ROVs find structural dents, ovalities, welds misalignment and local deformations acquired from the installation process, anchor strikes, and seabed movement. The changes caused by these factors produce significant predictors of zones where unnecessary stress is concentrated and cracks or leaks are likely to form. Structural anomaly data can be represented in geometric features, which can then be quantified as anomaly type, depth, and spatial distribution along the pipeline (Golan and El-Yaniv, 2018; Bao *et al.*, 2019).

Third, the detection of cracks is considered the most critical for predicting failure, since this data can be used to build failure prediction models. ROVs equipped with imaging and ultrasonic or laser scanning devices can find surface and subsurface cracks. Important engineered features such as crack length, orientation, and propagation rate over time can inform models on the likelihood of fatigue progression and failure.

Finally, ROV surveys capture marine growth, like barnacles, algae, and other biofouling species. Biofouling and marine growth, in general, augment external loading, alter drag around the structure, and accelerate the deterioration of the coating, which, in turn, affects the rate of corrosion (Lau, 2019; Stringer and Polagye, 2020). Growth thickness, coverage area, and biomass density are some of the many features that help quantify the biofouling effects.

The main form of defense against external corrosion of subsea pipelines is cathodic protection (CP), and its

monitoring provides fundamental electrochemical parameters for integrity assessment. CP surveys usually along the pipeline assess the distribution of both the voltage and the current, which, for the pipeline, provides direct information on the integrity of the protective coating and the likelihood of corrosion.

In this context, voltage measurements are used as indicators of the protective potential. If the protective potential is high enough to counteract corrosion, a sufficiently negative voltage can be recorded. On the other hand, if the potential is less negative, more positive, or fluctuates, then there is a possibility of coating breakdown and lack of CP current, as well as the possible shielding due to the marine growth. Time series of voltage readings can be engineered into features such as mean potential, minimum values, or deviation from the designated values (Pinzón and Colomé, 2019; El Hariri *et al.*, 2019).

The workload and efficiency CP system are confirmed by evidence from concurrent measurements. A high current demand usually indicates a current coating disbandment or exposure of bare steel. Stable current values, on the other hand, tend to indicate an intact coating. Engineered features may include current density per unit area, variability of the current with time, and other anomalies with respect to baseline conditions (Leigh *et al.* 2019; Amanambu *et al.* 2019).

Furthermore, the CP surveys offer a secondary method of assessing the condition of a coating. Differences in the response of the system to electrochemical perturbation are associated with the breakdown of the coating, blistering, or even detaching. Features like coating resistance and effective coverage ratios can be automated by integrating the CP datasets with the pipeline specifications in hand (Azarmi and Huang 2018; Eltai *et al.* 2019). Together, these features assist in estimating the likelihood of corrosion, as it remains one of the most determining factors to the risk of degradation a pipeline may face.

The degradation of a pipeline is not an independent happening. It takes place in conjunction and in consideration of the operational and environmental factors surrounding it. In the case to increase precision of forecast, constructed predictive models will have to integrate operational and environmental stressor data alongside the datasets derived from ROV and CP surveys.

Metrics such as internal fluid temperature, pressure, and flow rates of a system have a key influence on its internal corrosion rates, wax deposition, and hydrate formation as well: some of the processes and phenomena of which, operational and maintenance expenditure are greatest. For example, extreme temperatures and pressures may accelerate the kinetics of chemical reactions, resulting in the exacerbation of crack propagation of the materials. Features of which may include average operational ranges, frequency and duration of temperatures and pressures above and below design limits, and patterns of cyclic overloading.

Equally important are the environmental parameters. Conditions on the seabed, like its soil, its burial depth, and its slope, have much to say about the external stresses on a structure, the geohazards of which are landslides and the trawling of some. Corrosion and biofouling of which are classified as the oceanic parameters, salinity, dissolved oxygen, and current. These parameters have been studied and integrated as the spatial and temporal features by Nevshupa *et al.* (2018) and Azam *et al.* (2019), with the aim to adopt

machine learning techniques in modelling the environmental variability which pipelines experience over their lifespan.

The preparation of the datasets in their remoteness requires feature engineering. This strategy pertains the structure of machine learning datasets that minimizes slope of learning in the algorithm. In accomplishing such a strategy, two objectives are realized: the first being the proportion of the dataset which is being retained without change, while the second being the most useful features extracted which correlate with maximizing model performance on degradation prediction.

In the case of image data from ROV inspections, raw pixel information is often too high-dimensional to be used directly in models (Langenkämper *et al.*, 2020). Through preprocessing techniques such as edge detection, segmentation, and texture analysis, features such as fractures, cracks, surface roughness, and the extent of surface marine growth can be extracted and represented. Compact high-level feature embeddings can also be constructed directly from CNNs.

For the time-series CP data, engineered features include statistical descriptors (mean, variance, skewness), threshold indicators (time exceeded potential ranges), and frequency-domain features that expose recurrent patterns of degradation.

For the operational and environmental variables, engineered features can include interaction terms (temperature \times pressure as e.g. stress multiplier), cumulative exposure (total time in corrosive environments), and classified encodings of seabed soil types.

Methods of feature selection, such as recursive feature elimination, principal component analysis (PCA), or mutual information ranking, can, for the most part, streamline the input set to the most predictive features (Nie *et al.*, 2018; Doreswamy *et al.*, 2020). This minimizes the chances that models will be burdened with unnecessary and duplicative features that can contribute to overfitting and decreased generalizability.

The last stage of the feature engineering process deals with the mapping of a greatly varied and unstructured set of inputs, everything from images to chemical and physical recordings of the surroundings, into a clean, feature-rich, and multi-faceted data structure suitable for predictive modeling. This data structure serves as the basis for supervised learning approaches to be applied, where the pipeline integrity prediction models can be built using support vector machines, random forests, and other neural networks (Jiang *et al.* 2020; Sheykhmousa *et al.* 2020).

2.3. Model Architecture and Workflow

Constructing an architecture and workflow designed to facilitate machine learning enhancement on a pipeline integrity predictive model for offshore oil and gas deposits is a highly nuanced task as shown in figure 2. The myriad predictive results from a model are only partially the result of the algorithms selected. Much more is to be taken into account to make sense of the entire framework for collecting data, preprocessing data, training the data, optimizing the data, and interpreting the results as shown in table 1. The framework proposed seeks to combine ROV inspection data with results from cathodic protection surveys to provide the necessary data architecture for making intentional and well-grounded forecasts of the risk of degradation (Kowalczyk *et al.*, 2019; O'Byrne *et al.*, 2020).

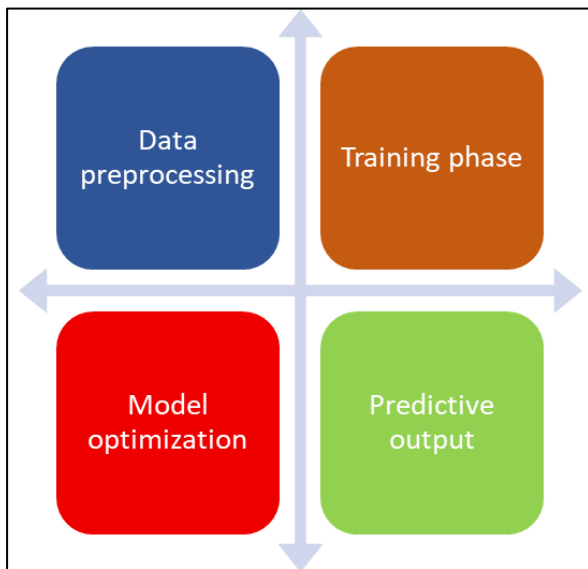


Fig 2: Model Architecture and Workflow

The first stage of the workflow is the preprocessing stage. The predictive outcomes of the model are closely tied to the input data, which is the model's foundational premise.

Offshore pipeline datasets are typically a hodgepodge of a wide variety of types of data, such as cathodic protection surveys, which generate structured numeric datasets, and ROV inspections, which produce unstructured or semi-structured images. The architecture is designed to clean, normalize, and fill missing value estimates to be harmonized. Data cleaning is the removal of data that compromises the integrity of the predictive model by inserting copies, correcting sensor errors, as well as removing features that do not add sufficient information to degradation prediction. In the case of depth data, where ROV passes are overlapped, recordings of depth readings are consolidated. Readings that are outliers of sensor drift are also excluded from the dataset, where no value is found with no elements present. Normalization solves the issue of transforming dissimilar units and scales—cathodic potential (often expressed in volts), depth of corrosion (measured in millimeters), and operational pressure (measured in bar) (Heidersbach, 2018; Groyzman, 2019). This, in turn, preserves the integrity of the other features by preventing the model from learning from features with greater magnitudes of numbers (Model). Common normalization techniques are z-score standardization and min-max normalization.

Table 1: Model architecture and workflow key activities and contributions

Workflow Stage	Key Activities	Purpose / Contribution
Data Preprocessing	- Cleaning noisy or corrupted sensor readings. - Normalization of heterogeneous datasets (ROV imagery, CP survey values). - Handling missing or incomplete data through imputation or interpolation.	- Ensures data quality and consistency. - Reduces bias and improves model reliability. - Prepares datasets for effective training.
Training Phase	- Combining ROV inspection imagery with cathodic protection (CP) survey datasets. - Extracting features related to corrosion, coating degradation, and structural anomalies. - Splitting datasets into training, validation, and testing subsets.	- Leverages multi-source data for robust learning. - Enables the model to identify patterns across visual and electrochemical signals. - Improves predictive accuracy of degradation trends.
Model Optimization	- Hyperparameter tuning using cross-validation. - Ranking feature importance to identify the most critical predictors (e.g., CP potential drop, crack density). - Avoiding overfitting through regularization and dropout strategies.	- Enhances model performance and generalizability. - Prioritizes features that most influence degradation risk. - Improves confidence in model outputs.
Predictive Output	- Generating probability scores of pipeline degradation risk. - Estimating timelines for required intervention or maintenance. - Visualizing predictions through dashboards or digital twins.	- Provides actionable insights for operators and engineers. - Supports risk-based maintenance planning. - Reduces the likelihood of catastrophic failures and unplanned downtime.

Handling missing values is another fundamental aspect of the model is the model's accuracy and reliability. Sensors might stop working, or a survey might not be extensive enough, and therefore, missing data might be present. For small gaps mean or median might be used for estimation; however, more sophisticated techniques like the k-nearest neighbor or model-based estimation are used for larger missing spans. Lack of data might be missing with a purpose, which may contain useful data, like a sequence of missing data might mean a breakage of the system. Therefore, missing data are not just considered as such; rather, the features are generated with which they might be present.

The focus of this stage of the training process is on the integration of two primary datasets: the ROV inspection data and the results of the cathodic protection survey. The ROV data vertically and horizontally visualize the disruption of coatings, the growth of marine flora and fauna, and the

occurrence of physical changes (Chen *et al.*, 2019; Nocerino *et al.*, 2020). These data, often in 'wild' or unstructured forms and images, are often in need of organization, which can be done through the use of edge detection and mapping, or other feature extraction and analysis techniques. The cathodic protection survey data, in stark contrast, consists of 'crisply' defined numbers that identify the potential of a pipeline, giving one an understanding of the electrochemical and corrosion phenomena of a given pipeline.

The process of 'joining' datasets is delicate. The alignment ensures that, in the case of images of the cathodic pipelines taken by an ROV, there is a match which the respective pipeline segments and survey periods of the other pipeline. The result is a covariate dataset that contains data regarding both physical observations and cathodic protection measurements simultaneously.

The dataset constructed captures historical degradation and

failure outcomes and uses them as labels for training supervised models such as Support Vector Machines (SVM) and Random Forests (RF) (Mohammed *et al.*, 2019; Musumeci *et al.*, 2020). SVMs are proficient at manipulating and delineating high-dimensional feature spaces and determining optimal margins across domains, while Random Forests enhance SVMs by treating them as weak classifiers—collapsing the outcomes of a multitude of SVM classifiers into a single, stronger classifier to minimize overfitting. Training SVMs concerns splitting a dataset into a model-fitting (training) and an evaluation (validation) set to check internal model validity and generalization to unseen datasets. Accurate and dependable predictions rely heavily on optimization. In the case of hyperparameter tuning, SVMs are noted for their kernels (linear, radial, and base functions), regularization strategies (C), and gammas to compute efficiency, while Random Forests are recognized for their component trees, maximum depths, and parameter tuning to mitigate overfitting (Probst *et al.*, 2019; Nadi and Moradi, 2019). Configuration of grid and very coarse randomized searches, in combination with cross-validation for hyperparameter tuning, is founded on the desire to minimize model error. Of equal importance in optimization is the ranking of features importance. In Random Forests, importance scores can be computed from the number and the quality of the splits involving certain variables. For instance, cathodic potential may come out as one of the most important predictors of corrosion risk, while bathymetric slope or variation of the coating thickness may be considered important but less significant. The studies of feature importance enhance the interpretability of models and help in formulating data acquisition plans by identifying which features should be the primary focus of future data collection (Carvalho *et al.*, 2019; Mi *et al.*, 2020).

Moreover, techniques for dimension reduction, such as Principal Component Analysis (PCA), can be employed to transform a data set of very large dimensions into one of fewer dimensions and yet retain the significant variation present in the data. This simplifies the model, which helps in reducing computational burden while also addressing the risk of overfitting.

In the end, the model attempts to give outputs that are both easy to understand and can be used to predict actions that will be helpful in managing the integrity of the system. The model does not simply assign pipelines into a set of broken or unbroken, but instead gives the chances of a certain pipeline sustaining damage over a given period. The model can predict which section of a pipeline is most likely to sustain damage, and thus lose structural integrity, along with the time period in which this will happen. These outputs are expressed as probabilities and make it possible to prioritize the most critical actions. This allows operators to decide where inspection and maintenance will be most effective.

Predicting the remaining useful life (RUL) of pipeline sections is made possible by the model's estimated solicited timelines for intervention (Kerwin and Adey, 2020; Meda, 2020). Its temporal scope is measured by the historical progression patterns of degradation and custom-fit operational and environmental variables. For example, a pipeline segment located in an area of strong ocean currents and with low cathodic protection potential tends to have a high degradation probability and a short RUL, indicating that prompt action is necessary.

Through dashboards, the model's predictive outputs can be

visualized by mapping degradation risk over networked pipelines, permitting targeted deployment of resources to the most critical undertaking. These outputs, combined with integration to digital twin platforms, enable inspection and sensor data to be continuously incorporated, facilitating dynamic predictive maintenance and management of pipeline integrity.

The integration of sophisticated data science with offshore engineering practice is evident in the model structure and the operational workflow. The model ensures cleansing and harmonization of the heterogeneous data through extensive data preprocessing phases. A synergistic ROV inspection and cathodic protection survey data model allows for the training phase to derive a solid multivariate dataset for supervised learning. Predictive accuracy is maximized through hyperparameter tuning and ranked feature importance while simultaneously allowing for model clarity and explanation (Probst *et al.*, 2019; Freitas, 2019). Proactive pipeline maintenance and risk-based decision-making are aided by predictive outputs in the form of intervention timelines and probability scores.

Incorporating this workflow into offshore integrity management systems will enable companies to shift from a model of reactive inspection and maintenance to predictive and preventative strategies that reduce operational downtime, mitigate subsea environmental risks, and improve safety for offshore operations (Golightly *et al.*, 2018; Raj and Surianarayanan, 2020).

2.4. Implications for Offshore Pipeline Integrity Management

Offshore pipeline integrity assessment moving to machine learning enhanced models is a shift from reactive to predictive management practices. Offshore oil and gas operators had relied on static inspections, visual checks, and manual reviews to capture and address any issues, structural or corrosion-related (Skovhus *et al.*, 2019; Tan, 2020). Such techniques, even though insightful, are incapable of forecasting degradation and preventing failures before they take place. Integrating machine learning with remote-operated vehicle (ROV) inspection, cathodic protection (CP) survey, and other operational parameters opens up new value streams with far-reaching consequences for pipeline integrity management, as shown in Figure 3. Such consequences include maintenance strategies, operational safety, regulatory compliance, and environmental compliance, which show a positive value add from a technical perspective and socio-environmental standpoint. One of the most immediate and practical implications of models built on machine learning is the possibility of proactive maintenance scheduling. The offshore oil and gas sector is used to performing inspections on pipelines at scheduled points in time set by regulatory and company policy, with little regard to risk. This is a recipe for overspending on inspections. The reality is that some inspections show little degradation, while other inspections are alarmingly neglected and show high levels of deterioration (Weise *et al.*, 2018).

Machine learning algorithms can utilize data from past inspections, CP surveys, operational data, and other information to predict degradation curves and segment time to failure for individual pipeline components. With these predictive model outputs, operators can better time maintenance activities to real requirements within the specified timelines, rather than the arbitrary and time-based

schedules. In addition, as a risk-based maintenance paradigm, this approach focuses resource allocations to the most valuable and critical pipeline sections, thereby improving cost-effectiveness.

Predictive models can further improve the economic value of maintenance by allowing operators to focus maintenance repairs on a prioritized basis. Suppose the system determines that certain segments of a pipeline can be classified as higher-priority for coating repairs or CP system additions, while other sections are adequately structurally sound, then repairs can be streamlined. Within this optimized framework, the highly specialized active divers and ROVs, as well as the more general repair assets, can be most economically distributed to address higher risk repairs (Thompson and Guihen, 2019; Will, 2019). Consequently, this methodology drives unscheduled repair actions to a minimum by lowering pipeline downtimes, enhancing overall pipeline service life, and controlling associated operational costs.

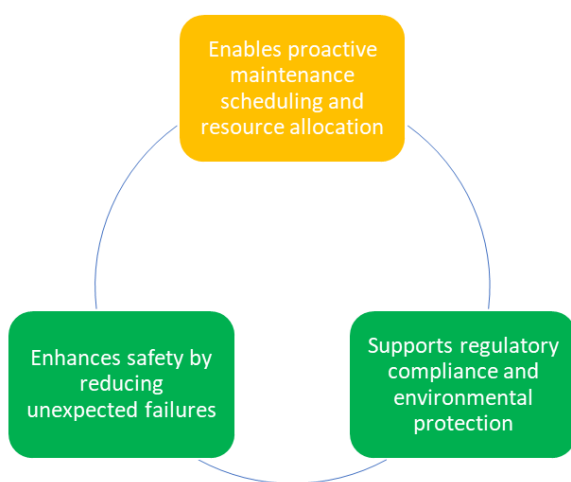


Fig 3: Implications for Offshore Pipeline Integrity Management

Safeguarding operational safety is another major advantage of predictive integrity assessment. Offshore pipelines are exposed to harsh conditions, and when things go wrong, the consequences for people and nature can be catastrophic. Leaks and ruptures that discharge hydrocarbons at high pressure pose a danger of fire, explosion, and toxic inhalation. Furthermore, casualties and injuries to offshore workers are a secondary risk, as failures may also impact the integrity of platforms and floating production systems.

This is exactly the insight provided by machine learning; by predicting the probability of failure in certain operational or environmental conditions, warning systems can be designed to give sufficient lead-time to deal with the hot spots that are most likely to fail. Predictive outcomes lead to proactive measures where CP failures, cracks, or coating failures that can lead to emergencies are managed much earlier in the protective zone rather than the reactive zone. This is critical in offshore emergencies where the unpredictable weather and limited mobility can restrict timely response. Predictive integrity management protects offshore personnel as well as adjacent shipping lanes, fishing activities, and coastal populations that might suffer from the failure of offshore pipelines. Therefore, the safety concerns go beyond the oil and gas operators and include other maritime stakeholders, which emphasizes the importance of predictive modeling.

These offshore pipeline operators are bound by legal requirements that include a series of inspections, reporting,

and maintenance activities. These activities are intended to systematically reduce the chances of an oil spill, gas leak, or adjacent environment. These activities also shield the operators from penalties for not being able to reduce the chances of an oil spill, gas leak, or adjacent environment. Compliance, however, with these regulations is very difficult to achieve due to the constraints that are imposed by legacy monitoring solutions. Those who are incapable of defining emerging risks or cannot provide evidence of risk to the satisfaction of audits (Yang and Li, 2018; Hubbard, 2020).

ML predictive systems enhance regulatory compliance capabilities by integrating risk-level assessments with data-driven evaluation of pipeline conditions. These systems have predictive integrity assessments, which not only quantify a pipeline's health but also create auditable data actions. This data provides evidence of compliance with frameworks set by the American Petroleum Institute (API) and other regulatory bodies of the International Maritime Organization (IMO) and International Organization for Standardization (ISO).

Predictive models support timely maintenance by identifying areas of necessary work before issues arise. This substantially decreases the risk of hydrocarbons being released into alien ecosystems, which helps in the reduction of maintenance costs and timely draws compliance fines from authorities. This also aids in mitigating the violation of Canadian and American waters, which serves large penalties, avoids fishery and tourism reputation loss, and avoids considerable conservative actions taken for corporate social responsibility. Integrating predictive insights into machine learning outputs allows for the inclusion of environmental predictive risk assessments within the broader ecological models and allows for the simulation of possible leaks or failures under various scenarios. This added functionality improves preparedness and resilience offshore operators can respond to an incident in a timely and effective with added effectiveness.

Predictive integrity management in offshore oil and gas does hold w strategic also has predictive integrity management. Systems augmented with machine learning technology increase planned productivity and subsequently planned downtime in the process, resulting in more productivity aligned with the global energy supply target (Lee *et al.* 2020, Kalusivalingam *et al.* 2020). This results in a first-mover advantage of digital transformation aligned with the company's focus on policy entrepreneurship relative to the leadership role in technology.

Access to the stakeholders improves the value added to investors and governments with the added value of the machine learning technology. offshore operators supervising human offshore activities provide added value with assurance on risk management and on environmental resource protection. This process enhances the trust in the system, streamlining the processes regarding permits and licensing, with the results of diminished operational risk perceived as insurance.

The implementation of systems that utilize machine learning technology for the management of offshore pipelines comes with new horizons. With the use of machine learning technology, predictive maintenance becomes operational, along with better management of resources. Further, it improves safety by minimizing the chances of failure and helps in abiding by regulatory requirements that relate to the environment. All these factors change the paradigm of pipeline management from reactive and expensive to

predictive, efficient, and environmentally conscious. Machine learning not only serves as an improvement to practices of inspection, but also embodies a change toward sustainability and resilience in the management of infrastructure for offshore oil and gas.

3. Conclusion and Future Work

This presentation underscores the abilities and provides the design of a machine learning model for forecasting pipeline integrity within offshore oil and gas fields. The model utilizes diverse data sets, including ROV inspection images, maintenance logs, and cathodic protection survey data. The model also illustrates the best practices of advanced analytics within integrity management. Predictive approaches outperform the traditional inspection-driven models, which are largely reactive, inefficient, and expensive. The model applies supervised learning techniques, including Support Vector Machines and Random Forests, to pinpoint nonlinear degradation patterns.

This model reaps many profits. Proactive maintenance minimizes disruptive downtimes and reduces operational expenses. The remaining useful life calculation and risk score estimation also improve the efficiency of decision-making. Operators can focus on more urgent tasks that need immediate attention. The model also improves the efficiency of answer prediction and provides primary drivers of degradation, which target mitigation and strategic planning. The model's features position it as a scalable, practical tool to enhance the safety, reliability, and sustainability of offshore pipeline operations.

The findings validate the potential of machine learning, but there is still more to be done. Of primary importance is the widening of datasets. Models still rely on very specific datasets, which is a problem for situational generalization in versatile offshore circumstances. Datasets that are more diverse in region, climate, and other situational variables would help in generalization, as well as robustness and adaptability.

The application of digital twin technology also offers considerable potential. By constructing a digital twin of a physical pipeline, for instance, it is possible to monitor performance in real time and simulate degradation. The use of machine learning to analyze performance—and the application of digital twins to simulate the conditions—could streamline performance forecasting and scenario testing.

The increasing predictive performance offered by the use of deep learning approaches is also relevant. Convolutional Neural Networks (CNNs), for instance, could facilitate the automated identification of coating defects or cracks in ROV footage. Furthermore, temporal deterioration estimation could be improved through the use of Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) systems, which offer improved estimates for deterioration progression.

Finally, future models should incorporate active monitoring fusion. With the offshore industry adopting more IoT sensors, enormous volumes of data regarding pressure, temperature, and electrochemistry can be continuously streamed and incorporated into predictive models. Changing the predictive models to incorporate real-time data would fundamentally change how pipeline integrity management is done, turning predictive interval revisions into a continuously modified process and enabling operators to act in real-time as risks intensify.

The discourse around offshore pipeline integrity is advanced by the ease of implementation of machine learning. It opens new data sources to merge with digital twins, real-time info, and incorporates deep learning approaches on the predictive models to enhance the value of the models. Acting on these steps will boost the industry's reputation for protecting the environment and critical infrastructures, as well as sustaining energy production, in the rapidly growing difficulties of the offshore projects.

4. References

1. Adamo N, Al-Ansari N, Sissakian V, Laue J, Knutsson S. Dam safety problems related to seepage. *J Earth Sci Geotech Eng.* 2020;10(6):191-239.
2. Adumene S, Adedigba S, Khan F, Zendejboudi S. An integrated dynamic failure assessment model for offshore components under microbiologically influenced corrosion. *Ocean Eng.* 2020;218:108082.
3. Ahmad T, Lund LH, Rao P, Ghosh R, Warier P, Vaccaro B, *et al.* Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc.* 2018;7(8):e008081.
4. Ajayi A, Oyedele L, Davila Delgado JM, Akanbi L, Bilal M, Akinade O, *et al.* Big data platform for health and safety accident prediction. *World J Sci Technol Sustain Dev.* 2019;16(1):2-21.
5. Amanambu AC, Li L, Egbinola CN, Obarein OA, Mupenzi C, Chen D. Spatio-temporal variation in rainfall-runoff erosivity due to climate change in the Lower Niger Basin, West Africa. *Catena.* 2019;172:324-34.
6. Angelopoulos A, Michailidis ET, Nomikos N, Trakadas P, Hatziefremidis A, Voliotis S, *et al.* Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. *Sensors.* 2020;20(1):109.
7. Ao Y, Li H, Zhu L, Ali S, Yang Z. Logging lithology discrimination in the prototype similarity space with random forest. *IEEE Geosci Remote Sens Lett.* 2019;16(5):687-91.
8. Asghar A, Farooq H, Imran A. Self-healing in emerging cellular networks: review, challenges, and research directions. *IEEE Commun Surv Tutor.* 2018;20(3):1682-709.
9. Atri P. Optimizing financial services through advanced data engineering: a framework for enhanced efficiency and customer satisfaction. *Int J Sci Res.* 2018;7(12):1593-6.
10. Azam HM, Alam ST, Hasan M, Yameogo DDS, Kannan AD, Rahman A, *et al.* Phosphorous in the environment: characteristics with distribution and effects, removal mechanisms, treatment technologies, and factors affecting recovery as minerals in natural and engineered systems. *Environ Sci Pollut Res Int.* 2019;26(20):20183-207.
11. Azarmi F, Huang Y. Mitigating pipeline corrosion using a smart thermal spraying coating system. Washington, DC: United States Department of Transportation, Pipeline and Hazardous Materials Safety Administration; 2018. Report No.: NDSU-12-20-2018.
12. Banda JM, Seneviratne M, Hernandez-Boussard T, Shah NH. Advances in electronic phenotyping: from rule-based definitions to machine learning models. *Annu Rev*

- Biomed Data Sci. 2018;1(1):53-68.
13. Bao Y, Tang Z, Li H, Zhang Y. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Struct Health Monit.* 2019;18(2):401-21.
 14. Biggio L, Kastanis I. Prognostics and health management of industrial assets: current progress and road ahead. *Front Artif Intell.* 2020;3:578613.
 15. Bordoff JE, Half AM, Losz A. New realities, new risks: rethinking the strategic petroleum reserve. 2018.
 16. Byrnes TA, Dunn RJ. Boating-and shipping-related environmental impacts and example management measures: a review. *J Mar Sci Eng.* 2020;8(11):908.
 17. Cai B, Shao X, Liu Y, Kong X, Wang H, Xu H, *et al.* Remaining useful life estimation of structure systems under the influence of multiple causes: subsea pipelines as a case study. *IEEE Trans Ind Electron.* 2020;67(7):5737-47.
 18. Carvalho DV, Pereira EM, Cardoso JS. Machine learning interpretability: a survey on methods and metrics. *Electronics.* 2019;8(8):832.
 19. Castiblanco S. The role of oil and gas companies in a low-carbon transition: strategic and technological point of view. 2021.
 20. Chen X, Wu G, Hou S, Fan J, Dang J, Chen Z. Development of tactile imaging for underwater structural damage detection. *Sensors.* 2019;19(18):3925.
 21. Cherepovitsyn A, Tsvetkova A, Komendantova N. Approaches to assessing the strategic sustainability of high-risk offshore oil and gas projects. *J Mar Sci Eng.* 2020;8(12):995.
 22. Chernov D, Sornette D. Specific features of risk management in the industrial and agricultural sectors. In: *Critical risks of different economic sectors: based on the analysis of more than 500 incidents, accidents and disasters.* Cham: Springer International Publishing; 2019. p. 13-145.
 23. Degan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, *et al.* Perceptual and interpretive error in diagnostic radiology—causes and potential solutions. *Acad Radiol.* 2019;26(6):833-45.
 24. Doreswamy, Hooshmand MK, Gad I. Feature selection approach using ensemble learning for network anomaly detection. *CAAI Trans Intell Technol.* 2020;5(4):283-93.
 25. Drummond GP, Pasqualino IP, Pinheiro BC, Estefen SF. Pipelines, risers and umbilicals failures: a literature review. *Ocean Eng.* 2018;148:412-25.
 26. El Hariri M, Harmon E, Youssef T, Saleh M, Habib H, Mohammed O. The IEC 61850 sampled measured values protocol: analysis, threat identification, and feasibility of using NN forecasters to detect spoofed packets. *Energies.* 2019;12(19):3731.
 27. Elshawi R, Al-Mallah MH, Sakr S. On the interpretability of machine learning-based model for predicting hypertension. *BMC Med Inform Decis Mak.* 2019;19(1):146.
 28. Eltai EO, Musharavati F, Mahdi ES. Severity of corrosion under insulation (CUI) to structures and strategies to detect it. *Corros Rev.* 2019;37(6):553-64.
 29. Ericson S, Engel-Cox J, Arent D. Approaches for integrating renewable energy technologies in oil and gas operations. Colorado: Joint Institute for Strategic Energy Analysis; 2019.
 30. Forsyth RA, Hartzell CS, Chader SA, Browne TM. Underwater inspection and imaging technologies for pipelines. In: *Pipelines 2018.* Reston, VA: American Society of Civil Engineers; 2018. p. 608-17.
 31. Freitas AA. Automated machine learning for studying the trade-off between predictive accuracy and interpretability. In: *International cross-domain conference for machine learning and knowledge extraction.* Cham: Springer International Publishing; 2019. p. 48-66.
 32. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, *et al.* Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ.* 2020;369:m1501.
 33. Golan I, El-Yaniv R. Deep anomaly detection using geometric transformations. *Adv Neural Inf Process Syst.* 2018;31.
 34. Goldsmith BR, Esterhuizen J, Liu JX, Bartel CJ, Sutton C. Machine learning for heterogeneous catalyst design and discovery. 2018.
 35. Golightly D, Kefalidou G, Sharples S. A cross-sector analysis of human and organisational factors in the deployment of data-driven predictive maintenance. *Inf Syst E-Bus Manag.* 2018;16(3):627-48.
 36. Groyzman A. Nondestructive testing and corrosion monitoring. In: *Non-destructive evaluation of corrosion and corrosion-assisted cracking.* Hoboken: Wiley; 2019. p. 261-409.
 37. Groyzman A. Nondestructive testing and corrosion monitoring. In: *Non-destructive evaluation of corrosion and corrosion-assisted cracking.* Hoboken: Wiley; 2019. p. 261-409.
 38. Hacquard P, Simoën M, Hache E. Is the oil industry able to support a world that consumes 105 million barrels of oil per day in 2025? *Oil Gas Sci Technol.* 2019;74:88.
 39. Heidersbach R. Metallurgy and corrosion control in oil and gas production. Hoboken: John Wiley & Sons; 2018.
 40. Ho M, El-Borgi S, Patil D, Song G. Inspection and monitoring systems subsea pipelines: a review paper. *Struct Health Monit.* 2020;19(2):606-45.
 41. Hubbard DW. The failure of risk management: why it's broken and how to fix it. Hoboken: John Wiley & Sons; 2020.
 42. Islam RU, Hossain MS, Andersson K. A deep learning inspired belief rule-based expert system. *IEEE Access.* 2020;8:190637-51.
 43. Jensen RC. Risk-reduction methods for occupational safety and health. Hoboken: John Wiley & Sons; 2019.
 44. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther.* 2020;51(5):675-87.
 45. Kalusivalingam AK, Sharma A, Patel N, Singh V. Enhancing predictive maintenance in manufacturing using machine learning algorithms and IoT-driven data analytics. *Int J AI ML.* 2020;1(3).
 46. Kenny S. Offshore pipelines—elements of managing risk. In: *Methods in chemical process safety.* Vol. 2. Amsterdam: Elsevier; 2018. p. 289-325.
 47. Kerwin S, Adey BT. Optimal intervention planning: a bottom-up approach to renewing aging water infrastructure. *J Water Resour Plan Manag.* 2020;146(7):04020044.
 48. Khalid HU, Ismail MC, Nosbi N. Permeation damage of polymer liner in oil and gas pipelines: a review.

- Polymers. 2020;12(10):2307.
49. Konchenko A, Chatar C, Doronichev S, Saetern D, Bruns J. Oilfield virtual twin. In: SPE Annual Technical Conference and Exhibition; 2020 Oct. p. D031S022R006.
 50. Kowalczyk M, Claus B, Donald C. AUV integrated cathodic protection ICP inspection system—results from a North Sea survey. In: Offshore Technology Conference; 2019 Apr. p. D011S002R004.
 51. Lakshminarayanan R. Small state security dilemma: Kuwait after 1991. Chennai: Notion Press; 2019.
 52. Langenkämper D, Van Kevelaer R, Purser A, Nattkemper TW. Gear-induced concept drift in marine images and its effect on deep learning classification. *Front Mar Sci.* 2020;7:506.
 53. Larrosa NO, Akid R, Ainsworth RA. Corrosion-fatigue: a review of damage tolerance models. *Int Mater Rev.* 2018;63(5):283-308.
 54. Lau K. Susceptibility of bridge steel and concrete components to microbiological influenced corrosion (MIC) and microbiological influenced deterioration (MID) in Florida. 2019.
 55. Lee J, Ni J, Singh J, Jiang B, Azamfar M, Feng J. Intelligent maintenance systems and predictive manufacturing. *J Manuf Sci Eng.* 2020;142(11):110805.
 56. Leigh C, Alsibai O, Hyndman RJ, Kandanaarachchi S, King OC, McGree JM, *et al.* A framework for automated anomaly detection in high frequency water-quality data from in situ sensors. *Sci Total Environ.* 2019;664:885-98.
 57. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. *Entropy.* 2021;23(1):18.
 58. Mäkitie T. Corporate entrepreneurship and sustainability transitions: resource redeployment of oil and gas industry firms in floating wind power. *Technol Anal Strateg Manag.* 2020;32(4):474-88.
 59. Marrec R, Abdel Moniem HE, Irvani M, Hricko B, Kariyeva J, Wagner HH. Conceptual framework and uncertainty analysis for large-scale, species-agnostic modelling of landscape connectivity across Alberta, Canada. *Sci Rep.* 2020;10(1):6798.
 60. Mattioli M, Cherubini P, Baldoni A. New frontiers for the pipeline integrity management. In: CORROSION 2019. Houston: Association for Materials Protection and Performance; 2019. p. 1-16.
 61. Meda R. Real-time data pipelines for demand forecasting in retail paint distribution networks. *Glob Res Dev.* 2020;5(12).
 62. Mi JX, Li AD, Zhou LF. Review study of interpretation methods for future interpretable machine learning. *IEEE Access.* 2020;8:191969-85.
 63. Mohammed B, Awan I, Ugail H, Younas M. Failure prediction using machine learning in a virtualised HPC system and application. *Cluster Comput.* 2019;22(2):471-85.
 64. Musumeci F, Magni L, Ayoub O, Rubino R, Capacchione M, Rigamonti G, *et al.* Supervised and semi-supervised learning for failure identification in microwave networks. *IEEE Trans Netw Serv Manag.* 2021;18(2):1934-45.
 65. Nadi A, Moradi H. Increasing the views and reducing the depth in random forest. *Expert Syst Appl.* 2019;138:112801.
 66. Nevshupa R, Martinez I, Ramos S, Arredondo A. The effect of environmental variables on early corrosion of high-strength low-alloy mooring steel immersed in seawater. *Mar Struct.* 2018;60:226-40.
 67. Nie F, Yang S, Zhang R, Li X. A general framework for auto-weighted feature selection via global redundancy minimization. *IEEE Trans Image Process.* 2019;28(5):2428-38.
 68. Nilsson EA, Christensen M. Arctic geopolitics, media and power. London: Taylor & Francis; 2019. p. 142.
 69. Nocerino E, Menna F, Gruen A, Troyer M, Capra A, Castagnetti C, *et al.* Coral reef monitoring by scuba divers using underwater photogrammetry and geodetic surveying. *Remote Sens.* 2020;12(18):3036.
 70. O'Byrne M, Ghosh B, Schoefs F, Pakrashi V. Applications of virtual data in subsea inspections. *J Mar Sci Eng.* 2020;8(5):328.
 71. Olawuyi DS. Environmental impacts and risks of extractive operations. In: Extractives industry law in Africa. Cham: Springer International Publishing; 2018. p. 269-322.
 72. Omoya OA, Papadopoulou KA, Lou E. Reliability engineering application to pipeline design. *Int J Qual Reliab Manag.* 2019;36(9):1644-62.
 73. Peng X, Anyaoha U, Liu Z, Tsukada K. Analysis of magnetic-flux leakage (MFL) data for pipeline corrosion assessment. *IEEE Trans Magn.* 2020;56(6):1-15.
 74. Pinzón JD, Colomé DG. Real-time multi-state classification of short-term voltage stability based on multivariate time series machine learning. *Int J Electr Power Energy Syst.* 2019;108:402-14.
 75. Primeau E. Innovations in inspection processes for marine pipelines. In: SPE Offshore Europe Conference and Exhibition; 2019 Sep. p. D011S002R006.
 76. Probst P, Boulesteix AL, Bischl B. Tunability: importance of hyperparameters of machine learning algorithms. *J Mach Learn Res.* 2019;20(53):1-32.
 77. Probst P, Wright MN, Boulesteix AL. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2019;9(3):e1301.
 78. Qin G, Zhang P, Hou X, Wu S, Wang Y. Risk assessment for oil leakage under the common threat of multiple natural hazards. *Environ Sci Pollut Res Int.* 2020;27(14):16507-20.
 79. Raj P, Surianarayanan C. Digital twin: the industry use cases. In: *Advances in computers.* Vol. 117, No. 1. Amsterdam: Elsevier; 2020. p. 285-320.
 80. Razvarz S, Jafari R, Gegov A. Flow modelling and control in pipeline systems: a formal systematic approach. Vol. 321. Cham: Springer Nature; 2020.
 81. Rifai AP, Aoyama H, Tho NH, Dawal SZM, Masruroh NA. Evaluation of turned and milled surfaces roughness using convolutional neural network. *Measurement.* 2020;161:107860.
 82. Sarwar A, Khan F, Abimbola M, James L. Resilience analysis of a remote offshore oil and gas facility for a potential hydrocarbon release. *Risk Anal.* 2018;38(8):1601-17.
 83. Shen J, Forsyth R, Kilgore R. Underwater inspection of bridge substructures using imaging technology. Washington, DC: Federal Highway Administration (US); 2018. Report No.: FHWA-HIF-18-049.
 84. Sheykhoumousa M, Mahdianpari M, Ghanbari H,

- Mohammadimanesh F, Ghamisi P, Homayouni S. Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2020;13:6308-25.
85. Skovhus TL, Taylor C, Eckert RB. Modeling of microbiologically influenced corrosion—limitations and perspectives. In: *Oilfield microbiology*. Boca Raton: CRC Press; 2019. p. 15-34.
86. Song H, Yang L, Liu G, Tian G, Ona D, Song Y, *et al.* Comparative analysis of in-line inspection equipments and technologies. In: *IOP Conference Series: Materials Science and Engineering*. Vol. 382, No. 3; 2018 Jul. p. 032021.
87. Stark A. *Public inquiries, policy learning, and the threat of future crises*. Oxford: Oxford University Press; 2018.
88. Stringer CC, Polagye BL. Implications of biofouling on cross-flow turbine performance. *SN Appl Sci.* 2020;2(3):464.
89. Tan MY. A critical overview of monitoring infrastructural health using corrosion probes. *Corros Eng Sci Technol.* 2020;55(2):103-17.
90. Teplý B, Rovnaníková M, Řoutil L, Schejbal R. Time-variant performance of concrete sewer pipes undergoing biogenic sulfuric acid degradation. *J Pipeline Syst Eng Pract.* 2018;9(4):04018013.
91. Therrien JD, Nicolai N, Vanrolleghem PA. A critical review of the data pipeline: how wastewater system operation flows from data to intelligence. *Water Sci Technol.* 2020;82(12):2613-34.
92. Thompson F, Guihen D. Review of mission planning for autonomous marine vehicle fleets. *J Field Robot.* 2019;36(2):333-54.
93. Vijayalakshmi A. Deep learning approach to detect malaria from microscopic images. *Multimed Tools Appl.* 2020;79(21):15297-317.
94. Weise FJ, Hayward MW, Aguirre RC, Tomeletso M, Gadimang P, Somers MJ, *et al.* Size, shape and maintenance matter: a critical appraisal of a global carnivore conflict mitigation strategy—livestock protection kraals in northern Botswana. *Biol Conserv.* 2018;225:88-97.
95. Will M. *An operations guide to safety and environmental management systems (SEMS): making sense of BSEE SEMS regulations*. Cambridge: Gulf Professional Publishing; 2019.
96. Wright RF, Lu P, Devkota J, Lu F, Ziomek-Moroz M, Ohodnicki Jr PR. Corrosion sensors for structural health monitoring of oil and natural gas infrastructure: a review. *Sensors.* 2019;19(18):3964.
97. Wu W, Li Y, Cheng G, Zhang H, Kang J. Dynamic safety assessment of oil and gas pipeline containing internal corrosion defect using probability theory and possibility theory. *Eng Fail Anal.* 2019;98:156-66.
98. Yahşi M, Çanakoğlu E, Ağralı S. Carbon price forecasting models based on big data analytics. *Carbon Manag.* 2019;10(2):175-87.
99. Yang D, Li M. Evolutionary approaches and the construction of technology-driven regulations. *Emerg Mark Financ Trade.* 2018;54(14):3256-71.
100. Zhang B, Gong R, Wang T, Wang Z. Causes and treatment measures of submarine pipeline free-spanning. *J Mar Sci Eng.* 2020;8(5):329.