PREDICTIVE INTEGRITY ASSESSMENT

Advanced Corrosion Monitoring and Predictive Integrity Assessment of Oil and Gas Pipelines in Saudi Arabia: Integrating Risk-Based Inspection, Degradation Modelling, and Preventive Maintenance

Azizur Rehman

Published November 2026

Abstract

Saudi Arabia's oil and gas pipeline systems operate across environments where saline soils, coastal exposure, variable water handling, and long asset lives can complicate corrosion control. Although inspection, corrosion prediction, and maintenance optimisation have each progressed, their integration into defensible operational decisions remains uneven. This critical review evaluates evidence published between 2020 and 2025 on pipeline monitoring, degradation modelling, risk-based inspection, and preventive maintenance, with emphasis on transferability to Saudi operating conditions. A structured thematic synthesis distinguishes direct measurements from inferred corrosion states, compares deterministic, probabilistic, and machine-learning approaches, and examines how uncertainty should influence intervention timing. The evidence indicates that high-resolution inspection alone cannot establish future integrity without sound defect matching, exposure histories, and calibrated growth models. Likewise, apparently accurate predictions can be misleading when validation ignores temporal dependence, geographic transfer, measurement error, or rare severe defects. This review develops an integrated evidence-to-decision framework linking sensing, data assurance, mechanism-informed forecasting, failure probability, consequences, and maintenance selection. A staged implementation pathway is proposed for operators managing heterogeneous pipeline networks and unequal data maturity. The central argument is that predictive integrity should be evaluated through decision reliability and risk reduction rather than predictive accuracy alone. The framework is conceptual; it does not report newly collected Saudi pipeline measurements or claim field-validated performance.

Keywords: Pipeline integrity; corrosion monitoring; Saudi Arabia; risk-based inspection; probabilistic degradation; preventive maintenance; predictive analytics

1 Introduction

Oil and gas pipelines are long-lived pressure-containing assets whose failure behaviour reflects interactions among material condition, operating history, surrounding environments, and human decisions. Corrosion remains especially difficult to manage because measurable metal loss is often localised, its growth rate varies over time, and available observations cover only parts of a network. External attack can begin beneath damaged coatings or shielding deposits, whereas internal attack depends on water availability, flow regime, dissolved gases, solids, microbial activity, and chemical treatment. The resulting mechanisms are not interchangeable. Wasim and Djukic [1] show why external corrosion assessment must consider several interacting damage processes; Sliem et al. [2] sim-

ilarly demonstrate that under-deposit corrosion requires evidence extending beyond a single thickness measurement.

Saudi Arabia provides a demanding application context rather than a uniform corrosion environment. Inland buried transmission lines, coastal and offshore crossings, industrial gathering networks, and facilities' interconnecting pipework encounter different combinations of soil chemistry, temperature, pressure cycling, and access constraints. These are relevant engineering considerations, not evidence that every Saudi route experiences the same severity. A national strategy must therefore support local exposure classification while retaining common rules for data quality and risk acceptance. The practical question is not simply which sensor measures the smallest defect, but whether operators can convert imperfect evidence into timely, proportionate action.

Recent studies span mechanistic corrosion models, statistical growth processes, magnetic inspection analytics, explainable learning, and maintenance optimisation [3–10]. Yet performance is frequently presented as an error score or a case-specific optimal inspection interval. Such quantities do not automatically establish reliability when sensor resolution changes between surveys, records are incomplete, or failure consequences differ sharply among locations. The literature requires synthesis across the full decision chain, from exposure to intervention. This review addresses that need without transferring numerical results from foreign case studies into unsupported Saudi forecasts.

A further distinction concerns scale. Local corrosion at a single weld, a long segment with repeated coating disbondment, and a network-wide inspection backlog present different decision problems. Segment-level averages may conceal the weakest location; anomaly-level predictions cannot alone explain how route access, shutdown constraints, and consequence concentration affect network priorities. The proper unit of analysis must therefore be selected before an algorithm is developed. A pipeline kilometre, corrosion anomaly, inspection joint, and pressure system cannot be treated as equivalent observations. Where these units are mixed, model evaluation may appear favourable without improving the actual integrity decision.

2 Aim, objectives and review questions

The aim is to establish a technically defensible, Saudi-relevant framework for converting corrosion observations into predictive integrity and preventive maintenance decisions. Unlike studies concerned chiefly with estimating corrosion rate, the review focuses on how evidence uncertainty, structural capacity, consequence severity, and intervention effectiveness interact across the asset life-cycle.

Four objectives guide the analysis: to distinguish complementary sensing methods and their blind spots; to appraise degradation models against physical plausibility, calibration demands, and transferability; to formulate a coherent connection between predicted damage, failure risk, and inspection intervals; and to identify implementation requirements for maintenance selection, assurance, and governance. Three questions follow. Which observations genuinely reduce uncertainty about the governing corrosion mechanism? Under what conditions do complex prediction methods improve decisions over transparent baselines? How should an operator select inspection, mitigation, repair, or replacement when both deterioration and intervention outcomes are uncertain?

3 Review design and methodological boundaries

This article uses a structured critical narrative review rather than a quantitative meta-analysis. The literature domain comprises peer-reviewed research dated 2020–2025 addressing oil or gas pipeline corrosion, condition monitoring, structural reliability, uncertainty analysis, inspection planning, and maintenance decisions. Searches were framed around combinations of "pipeline corrosion", "in-line inspection", "magnetic flux leakage", "ultrasonic monitoring", "Bayesian degradation", "risk-based inspection", "remaining strength", and "maintenance optimisation". Publisher records and DOI landing pages were used to check the identity and publication period of selected works. The final cited corpus contains 30 articles chosen for direct technical relevance and complementary methods; it is a purposive evidence set, not a statistically exhaustive census.

Eligibility required a discernible connection to pipeline integrity and a published research or review contribution within the specified period. Work concerned solely with wellbore casing, general machinery maintenance, or unrelated structural materials was not treated as direct pipeline evidence. Methods were compared by input requirements, target variable, validation design, uncertainty treatment, and operational relevance. Where articles reported distinct industrial examples, their findings were interpreted as bounded demonstrations rather than universal performance claims. No pooled model accuracy, incident frequency, or Saudi-specific rate was calculated because datasets and outcomes are not commensurable.

Evidence was organised into four analytical domains: monitoring and feature extraction; physical and statistical deterioration; risk and structural capacity; and preventive intervention. Contradictions were examined through likely differences in data generation, defect morphology, exposure regimes, or validation criteria. Conceptual integration then linked these domains through explicit feedback and decision gates. The approach strengthens traceability but has limitations: database coverage was not enumerated, unpublished operational records were unavailable, and reference selection may favour accessible studies. Accordingly, the paper offers a reasoned framework for subsequent validation, not a claim that a new predictive model has already been tested.

The synthesis also applied a qualitative appraisal rubric. Greater weight was assigned to studies that specified their observation source, corrosion mechanism, target variable, test separation, uncertainty assumptions, and practical decision context. Simulation-based evidence was useful for exploring difficult parameter combinations, but it was differentiated from prospective field validation. Studies reporting a correlation coefficient without a leakage-sensitive error analysis were not interpreted as proving readiness for deployment. This appraisal deliberately avoided numerical quality scores because reporting practices and study designs were too diverse for credible weighting. It also avoided equating citation frequency with engineering reliability.

4 Saudi operating exposures and corrosion mechanisms

A useful Saudi risk classification begins with mechanisms, not geography alone. Buried pipelines may face external corrosion where coating holidays expose steel to electrolytes, where cathodic protection is inadequate or shielded, or where differential aeration creates concentrated attack. Arid surface conditions do not eliminate subsurface electrolytes; conversely, a saline location does not prove active damage. Soil resistivity, moisture persistence, chloride concentration, pipe-to-soil potential, coating condition, and drainage should be evaluated jointly. The modelling study of Cui and Wang [6] identifies chloride, pH, potential, and age as influential variables, while its uncertainty analysis cautions against treating environmental covariates as fixed predictors.

Within a line, carbon dioxide and hydrogen sulphide chemistry, water wetting, solids deposition, local hydraulics, and inhibitor delivery can produce damage patterns [11]. Under-deposit conditions are particularly relevant where low velocity or stagnant sections retain water and particulate matter [2]. Sharma et al. [21] demonstrate the value of combined chemical, microbiological, and corrosion evidence when diagnosing a pipeline facility leak; Abilio et al. [22] underscore the classification difficulty of microbiologically influenced corrosion. These mechanisms require direct confirmation before chemical treatment is changed. Offshore exposure adds seawater chemistry, temperature, and coating or protection-system integrity; Chohan et al. [23] provide laboratory evidence of interacting environmental influences, not a universal field corrosion law.

Operating history must accompany exposure maps. Shutdowns, changes in water cut, pigging frequency, temperature excursions, pressure cycling, and fluid-source changes can alter both corrosion propensity and the relevance of previous inspections. A defensible segmentation scheme therefore combines stable characteristics, such as diameter, wall thickness, steel grade, and coating vintage, with dynamic conditions including water chemistry and inhibitor compliance. This creates segments that remain interpretable after operational changes and avoids assigning one average corrosion rate to dissimilar assets.

Environmental categorisation should be revisited after infrastructure modifications. New power supplies, crossings, nearby buried assets, drainage works, and coating repairs may change current distribution or moisture conditions. Internal risk can also shift after tie-ins, new production streams, commissioning chemicals, or revised pigging procedures. Exposure registers should record both the prevailing mechanism and competing hypotheses, especially when an unexpected defect is discovered. This prevents premature assignment of every anomaly to a familiar cause and encourages confirmatory excavation, sampling, or focused nondestructive testing where the failure consequences justify the cost.

5 Measurement architecture and evidence quality

Direct and indirect monitoring answer different questions. Magnetic flux leakage inspection detects disturbances associated with metal loss and can cover extensive accessible lines, but reporting depends on tool response, geometry, speed, sizing tolerances, and algorithm assumptions. Ultrasonic inspection offers wall-thickness information under suitable coupling conditions; neither modality automatically identifies the underlying chemistry. Shen and Zhou [25] demonstrate model-based interpretation of magnetic signals, while Cui and Wang [5] show that repeated in-line inspection can support defect analysis and forecasting. Both approaches depend on consistent feature registration between inspections.

Permanent or periodic ultrasonic thickness systems provide local time series, while electrical resistance probes, corrosion coupons, and electrochemical measurements provide information about local corrosivity. These measures should not be interpreted as instantaneous substitutes for the maximum pit depth elsewhere on a line. Chemical sampling can identify water, carbon dioxide, chlorides, solids, or inhibitor residuals but is vulnerable to sampling location and temporal bias. Cathodic-protection surveys describe an important control system rather than directly proving sound remaining wall. A well-designed programme uses complementary observations to distinguish a control failure from established structural degradation.

The most difficult measurement problem may be correspondence. A feature apparently growing between two surveys could reflect true wall loss, inaccurate axial matching, changed reference

datums, revised sizing algorithms, or different detection thresholds. Before estimating growth, operators should document tool identity, inspection date, calibration, stated accuracy, anomaly orientation, circumferential position, chainage transformation, and confidence in feature matching. An unmatched anomaly cannot be assigned a precise longitudinal growth rate merely because two inspection files exist. The same discipline is essential when combining laboratory corrosion data with operating observations.

Table 1 separates the principal evidence types by what they measure, what they may miss, and the validation needed before decisions. Its central lesson is that independent measurement channels are valuable because their errors and sensitivities differ. However, redundancy must be genuine: two models trained on the same flawed inspection labels do not supply independent confirmation. Monitoring architecture should be selected for decision relevance, failure consequences, tool access, and evidence gaps rather than sensor novelty alone.

For Saudi operators, digital integration is most useful when it preserves provenance. Each measurement should retain units, spatial position, acquisition conditions, quality flags, and the management decision it supports. A common asset identifier must link inspection runs, repairs, coating surveys, operational alarms, and incident investigations. The condition-monitoring digital-twin case of Wang et al. [24] illustrates the potential of combining physical, statistical, and signal-processing representations. Nevertheless, a digital model without reliable field synchronisation can simply accelerate the circulation of inconsistent information. Data assurance is therefore part of integrity engineering, not an administrative afterthought.

Sensor selection must additionally consider the scale at which an actionable change occurs. A high-frequency local probe may detect a rapid chemistry excursion but cannot map a remote pit; a whole-line inspection may resolve distributed metal loss yet miss brief process events between campaigns. Combining these outputs requires temporal alignment and an explicit statement of the spatial support of each variable. For example, an inhibitor residual sampled at one station should not be imputed without qualification to an entire heterogeneous network. Evidence fusion is credible only when extrapolation and representativeness are documented.

Table 1. Complementary evidence streams for corrosion monitoring and diagnostic assurance.

Evidence stream	Principal contribution	Important blind spot	Minimum validation expectation
In-line magnetic flux leakage	Broad metal-loss anomaly mapping	Sizing bias, feature matching, accessibility	Tool performance, calibration and anomaly reconciliation
Ultrasonic inspection	Wall-thickness profile or local trend	Coupling, coverage and geometry limitations	Repeatability and independent thickness checks
Coupons / resistance probes	Local corrosion activity over an interval	Not a remote maximum pit-depth measurement	Representativeness of location and fluid exposure
Cathodic-protection surveys	Condition of external electrochemical protection	Not direct proof of remaining wall thickness	Potentials, continuity, shielding and coating context
Chemistry / microbial testing	Mechanism evidence and treatment indicators	Sampling bias and changing flow regimes	Representative sampling and triangulation with damage

Note: Comparative appraisal derived from the reviewed literature; entries are not instrument acceptance specifications.

6 Degradation models: Prediction, uncertainty and transfer

Predictive models should distinguish corrosion initiation, metal-loss propagation, and structural failure. A corrosion-rate estimate describes change in a particular variable; remaining-life assessment additionally requires initial thickness, relevant defect geometry, uncertainty, operating pressure, and a failure criterion. Empirical and semi-empirical models are attractive when physical drivers and their validity limits are known. Cruz et al. [4] identify important limitations of uniform-corrosion formulations under operational conditions, and Fang et al. [3] use physics-derived scenarios to construct learning datasets when accessible field observations are scarce. Simulated labels can enlarge coverage, but they cannot independently establish field validity.

Simple deterministic extrapolation remains a necessary benchmark. Given two matched observations, an average depth growth rate can be expressed as the change in measured depth divided by elapsed time. However, this assumes compatible measurements and implicitly compresses acceleration, inhibition, and environmental switching into a single number. Stochastic processes acknowledge changing increments. Shao et al. [10] present a multistage Wiener-process formulation, while Cui and Wang [5] examine stochastic defect growth with inspection information. These methods can express dispersion and update expectations, although the chosen process and error distribution still require field checking.

Machine learning addresses nonlinear interactions, but its value depends on the sampling design. Guang et al. [16] propose an attention-based deep model for buried-pipeline corrosion; Hu [29] evaluates interpretable ensemble learning for internal corrosion rate; Xie et al. [30] combine a knowledge graph with a neural network. Their architectures should not be ranked by reported coefficients alone because sample sizes, response definitions, simulated inputs, and held-out data differ. Wang et al. [17] trace this shift from empirical to data-driven prediction, and Soomro et al. [18] document persistent difficulties in burst-pressure training data. A high fit on randomly divided defect records may overstate performance if adjacent defects or repeated readings from one pipeline appear in both training and testing sets.

A stronger evaluation reserves entire pipelines, time periods, or environmental regimes for testing. Models should be compared with transparent regression, physical baselines, and expert rules; report calibration as well as mean error; and test stability under missing covariates and measurement shifts. The ensemble Bayesian network of Cui and Wang [6] explicitly models predictive uncertainty, whereas the stacking approach of Wang and Lu [19] targets residual-strength prediction. These address different outputs. Predicting corrosion depth, burst pressure, and probability of failure are connected tasks, but none can be substituted for another without a structural limit-state calculation.

Uncertainty arises from at least four sources: measurement error, uncertain present condition, future exposure variability, and model inadequacy. Bayesian updating can revise forecasts after inspection, provided the likelihood honestly represents sizing and matching errors. Probability intervals that exclude plausible severe outcomes are dangerous even if central predictions are accurate. For highly consequential segments, operators should examine upper-tail metal loss and lower-tail strength rather than relying on the expected defect depth. Table 2 compares model families in these terms. The preferred strategy is often hybrid: physics constrains credible behaviour, statistical models quantify variation, and learning algorithms extract relationships where sufficient representative data exist.

A useful engineering test asks whether the forecast preserves plausible orderings under controlled changes. Holding other conditions constant, a more severe measured defect should not inexplicably imply greater remaining strength unless the model also accounts for compensating geometry or material properties. Sensitivity analysis should test realistic shifts in pressure, wall thickness, corrosive exposure, and instrument accuracy. Calibration plots, prediction-interval coverage, and errors in the deepest defect quantiles deserve particular attention. These diagnostics reveal failure modes hidden by a single overall coefficient of determination and provide clearer grounds for engineering acceptance.

Table 2. Comparative assessment of degradation and integrity-modelling approaches.

Model family	Required evidence	Critical limitation	Decision-relevant use
Empirical / mechanistic	Chemistry, exposure, material parameters	Validity range; omitted mechanisms	Transparent screening and physical baseline
Linear defect growth	Matched inspection dates and depths	Measurement error; non-stationarity	Simple benchmark with cautious extrapolation
Stochastic process	Repeated inspections; error distributions	Process choice; tail behaviour	Prediction intervals and risk progression
Machine learning	Representative labelled, diverse conditions	Leakage; shift; severe-defect scarcity	Nonlinear prediction after robust holdout
Bayesian / hybrid	Mechanistic priors and inspection updates	Prior sensitivity; likelihood misspecification	Updated uncertainty and decision analysis
Structural reliability	Defect geometry, pressure, material strength	Interacting defects; uncertain limit states	Leak/rupture probability and intervention gates

Note: Comparative appraisal derived from the reviewed literature; entries are not instrument acceptance specifications.

7 From degradation forecasts to risk-based inspection

Risk-based inspection connects deterioration to the timing and value of new evidence. A practical formulation estimates time-dependent risk from the probability of a defined failure scenario and the consequences of that scenario. Consequences may include injuries, environmental release, service interruption, repair cost, or strategic supply disruption. Their boundaries must be transparent: monetary aggregation does not remove ethical or regulatory constraints on unacceptable safety outcomes. Kumari et al. [12] combine neural and Bayesian analysis for corrosion-induced incident risk, while Mahmood et al. [20] demonstrate the role of mixed expert and empirical evidence in pipeline risk assessment.

The probability calculation requires an explicit failure mode. Leakage through a penetrating pit and rupture due to insufficient burst resistance have different limit states. Cluster interactions, surrounding stresses, material strength, and pressure variability influence structural capacity. Mensah and Sriramula [14,15] show that interacting corrosion defects and spatial uncertainty can alter reliability assessment. Accordingly, applying a single remaining-thickness threshold across every defect configuration can overlook geometry-driven effects. Failure modes must also consider dents and other coincident damage: Huang et al. [7] incorporate corrosion–dent interaction in in-spection planning and show why isolated-threat ranking can misrepresent an asset's condition.

Inspection priority must be distinguished from repair priority. An uncertain but potentially serious segment may benefit most from immediate inspection, whereas a well-characterised unacceptable

defect requires intervention rather than further modelling. A high-consequence route may justify more frequent evidence acquisition even when the estimated corrosion rate is modest. Conversely, repeated inspection of a stable low-consequence segment can consume resources better allocated to mitigation elsewhere. The decision variable is therefore not simply inspection interval but the expected benefit of information relative to inspection cost, exposure, and plausible loss.

A defensible inspection rule specifies three gates: a maximum tolerable risk, a minimum confidence level for the condition estimate, and a trigger for an earlier reassessment after process change or abnormal surveillance. Li et al. [8] connect pitting progression, risk, and maintenance decisions; Wang et al. [9] formulate maintenance objectives with multiple correlated failure modes. Their optimisation results should be treated as illustrations of possible frameworks, not transferable Saudi interval prescriptions. An operator must recalibrate risk limits, cost assumptions, and intervention effectiveness using local consequence maps and verified asset information.

Figure 1 presents the proposed evidence-to-decision architecture. The loop is intentionally bidirectional: intervention records and new inspections must update both the condition database and model credibility. Auditability demands that each risk ranking retain the governing defect, scenario, assumptions, uncertainty interval, last inspection date, and accountable reviewer. A single coloured risk score without these elements cannot support critical challenge or independent assurance.

The value of a new inspection can be framed as the improvement in expected decisions after its possible results are considered. Inspection is valuable when it could plausibly reverse a repair decision, reveal unacceptable damage, or prevent unnecessary isolation; it contributes less when every credible result leads to the same action. This perspective discourages mechanically short inspection cycles for assets already well understood. It also requires detection probability to be acknowledged. An inspection tool that cannot reliably size the controlling defect is not informative merely because it has produced another dataset. False negatives should remain explicit in the residual risk.

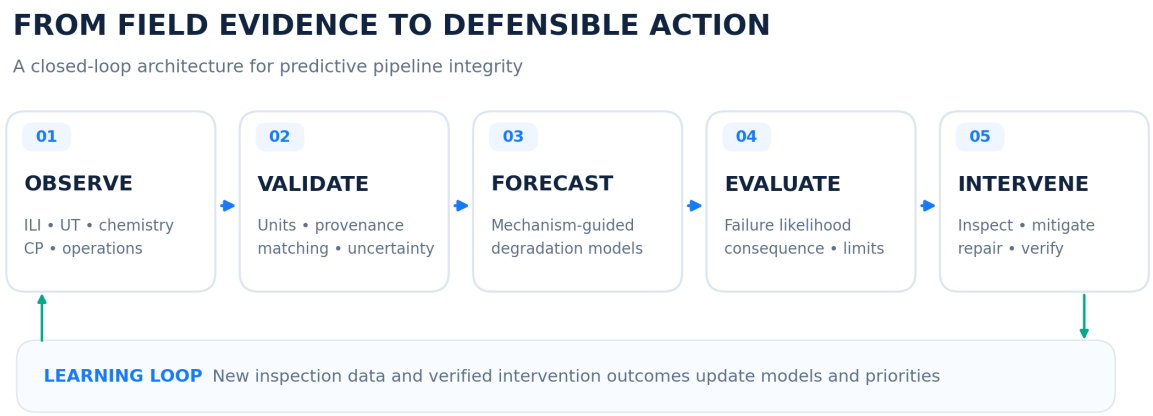


Figure 1. Conceptual evidence-to-decision architecture with feedback from verified field outcomes. Original schematic.

8 Preventive maintenance as a decision problem

Maintenance actions have different causal mechanisms. Coating refurbishment and cathodic-protection correction address external corrosion drivers; inhibitor optimisation, water removal, pigging, and targeted cleaning address particular internal mechanisms. Grinding, engineered repair sleeves,

composite repairs, pressure restriction, spool replacement, or full renewal address demonstrated loss of structural capacity. Actions should be chosen after mechanism diagnosis and fitness-for-service assessment, rather than from an automatically generated generic treatment list. In particular, a chemical remedy has little relevance to an external coating holiday, and increased inspection frequency does not itself reduce corrosion.

The choice among actions should account for their imperfect effectiveness. A repair may restore local strength but leave neighbouring exposure unchanged; chemical treatment may reduce average corrosion yet fail beneath persistent deposits. Wang et al. [9] evaluate competing cost and availability objectives under imperfect maintenance. Li et al. [8] connect maintenance selection with predicted pitting risk, while Zheng et al. [27] propose state-dependent, non-periodic decisions for natural gas lines. These contributions support adaptive planning, but the optimum is sensitive to assumptions about cost, inspection sensitivity, uncertainty, and the real duration of mitigation benefit.

A practical preventive policy combines immediate safety constraints with lifecycle economics. First, segments that breach structural or regulatory acceptance criteria require prompt engineered disposition. Second, remaining segments can be ranked by expected avoidable loss, considering how candidate actions change future degradation and failure probability. Third, the plan should incorporate work-window access, line shutdown implications, logistics, and availability of specialist inspection tools. Finally, effectiveness must be verified through post-action measurements and performance indicators, rather than assuming that completion of a work order proves risk reduction.

Lifecycle cost comparisons should distinguish expenditures incurred with certainty from uncertain consequences avoided. They should also include false alarms, undetected defects, delayed repairs, repeated mobilisation, and data-management overhead. A lower-cost option is not preferable when it breaches a hard safety constraint. Likewise, a highly sophisticated forecast delivers no value if action authorisation cannot respond before the risk threshold is reached. Figure 2 translates these principles into an operational sequence for selecting inspection, mitigation, or repair under uncertainty.

Maintenance packages should avoid treating different controls as statistically independent. Inhibitor performance may depend on cleaning and water removal, while external coating work can interact with cathodic-protection demand. Combining actions may create stronger risk reduction than isolated activities, but it may also increase operational disruption. Decision analysis should reflect these interactions, including uncertain implementation quality. An accountable engineer should retain authority to require additional verification whenever a proposed low-cost action conflicts with observed corrosion morphology or known failure history. The optimisation model supports engineering judgement; it cannot replace it.

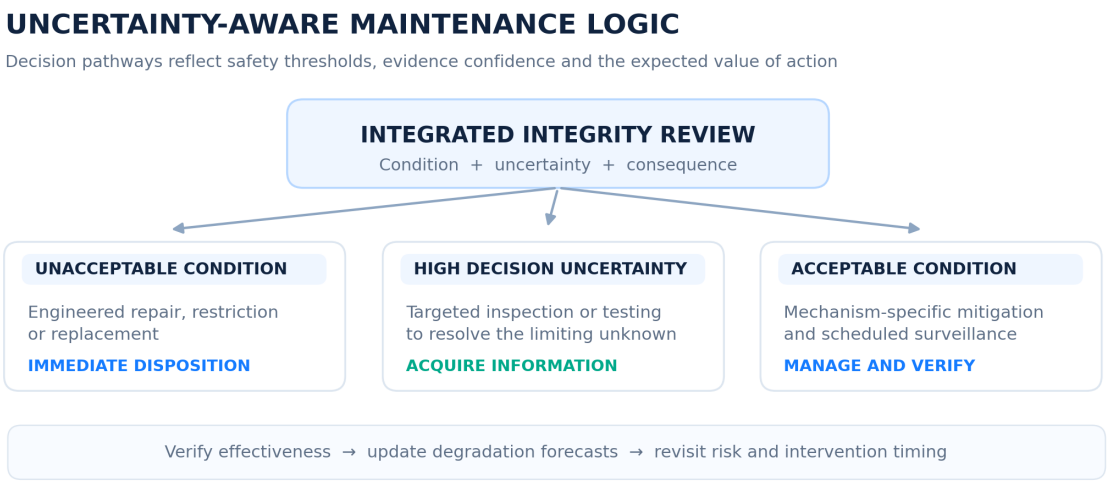


Figure 2. Conceptual logic for distinguishing immediate intervention, information acquisition, and surveillance. Original schematic.

9 Proposed implementation and assurance pathway

Implementation should begin with a minimum viable integrity dataset: verified pipeline geometry, design and operating limits, material and coating records, route exposure, inspection history, protection-system evidence, known failures, and intervention records. Each field should carry provenance and an explicit completeness flag. A route-level baseline should then partition the network by exposure mechanism, inspection access, and consequence setting. This is more defensible than deploying an undifferentiated predictive model across every asset from the outset.

A first implementation stage should select a small number of contrasting pilot segments, such as a mature inland route, a wet-service internal-corrosion concern, and a coastal or offshore exposure. The purpose is not to assume their severity but to test whether the framework handles materially different mechanisms. Historical inspection data should be reconciled manually for a sampled set of anomalies before automated growth estimates are accepted. Inspection uncertainty, censoring, and changes in tool specifications must be documented in the modelling dataset.

The second stage compares three forecasting levels: a transparent engineering baseline; a statistical growth model with uncertainty; and, only where justified, a hybrid or learned model. Prospective evaluation should measure severe-defect recall, forecast coverage, calibration, stability across routes, and the number of interventions changed by model outputs. Aryai et al. [13] address reliability with limited failure history, demonstrating why sparse evidence calls for explicit probabilistic treatment rather than unwarranted certainty. Wahono et al. [28] place risk-based approaches within wider asset integrity management, reinforcing the requirement to connect analytics with governance.

The third stage embeds controlled decisions within normal maintenance systems. Engineers should approve integrity thresholds, review exceptional predictions, and record reasons for overriding recommendations. Models should be versioned; changes in sensor hardware, chemistry, route conditions, or failure criteria should trigger revalidation. Leading indicators include unresolved critical anomalies, overdue protection-system defects, matched-ILI coverage, inhibitor non-compliance, and calibration drift. Lagging indicators such as leaks and repair expenditure remain necessary but cannot alone demonstrate improved control. Independent technical assurance should periodically test whether risk reductions attributed to the programme are supported by new measurements.

Procurement and contractor interfaces are important sources of hidden data loss. Inspection specifications should require raw and processed deliverables, anomaly confidence descriptors, calibration records, and an agreed method for reconciling stationing. Asset owners should preserve the ability to reproduce risk rankings independently of a proprietary dashboard. Cybersecurity controls must protect operational telemetry and model access without preventing traceable engineering review. Before a prediction is authorised for high-consequence decisions, its operational design domain, fallback rule, escalation pathway, and responsible technical authority should be documented in controlled procedures.

10 Research gaps and implications for Saudi networks

Three evidence gaps are especially consequential. First, many published models rely on narrow datasets, simulated conditions, or random train–test splits that do not establish transfer across operating systems [3,16,18]. Research should release de-identified, mechanism-labelled benchmark datasets with inspection uncertainty and longitudinal defect identifiers. Second, growth and structural strength are commonly assessed in separate studies, although their uncertainties propagate into the same maintenance decision [14,19]. Comparative trials should evaluate end-to-end decision quality, including missed severe defects and unnecessary intervention, not merely regression metrics.

Third, environmental drivers and maintenance effectiveness are often only partially observed. More credible evaluations would combine inspection runs with water chemistry, protection-system surveys, coating records, treatment history, and pressure data. Field studies should estimate whether information changes decisions and reduces verified risk under realistic resource constraints. The broad survey by Chen [26] supports the need to connect monitoring, inspection, and fitness-for-service assessment, but local validation remains indispensable.

For Saudi Arabia, future work could organise multi-operator pilots around common data definitions while retaining access controls for commercially sensitive asset records. Relevant performance measures include forecast-interval coverage, defect-detection effectiveness, intervention lead time, cost per risk reduction, and verified post-maintenance condition. Regulatory alignment, workforce competence, and contractor data interoperability should be evaluated alongside algorithmic performance. This research agenda is compatible with ambitions for safer, more reliable, and economically efficient energy infrastructure, but it does not imply that a particular national performance gain has already been measured.

11 Conclusions

Advanced corrosion monitoring becomes effective integrity management only when measurement, degradation forecasting, structural assessment, inspection planning, and preventive maintenance form an auditable feedback loop. The reviewed literature provides substantial progress in sensor interpretation, stochastic modelling, machine learning, and risk-informed intervention; it also exposes recurring weaknesses in transferability, error characterisation, and decision validation. No single prediction architecture is universally optimal, and no reported case-study accuracy should be adopted as a Saudi operating benchmark without independent evidence.

The proposed framework prioritises mechanism-specific exposure classification, provenance-preserving inspection data, physically credible forecasting, uncertainty-aware failure assessment, and

interventions whose effectiveness is subsequently verified. It treats inspection and repair as different decisions and places explicit acceptance thresholds above unconstrained cost optimisation. For Saudi pipeline networks, the most credible route forward is staged local validation across contrasting exposures, supported by disciplined data governance and accountable engineering review. Future research should demonstrate measurable risk reduction and calibrated performance in field settings, rather than relying solely on apparently improved prediction scores.

References

[1] Wasim M, Djukic MB. External corrosion of oil and gas pipelines: A review of failure mechanisms and predictive preventions. *Journal of Natural Gas Science and Engineering*. 2022;100:104467. https://doi.org/10.1016/j.jngse.2022.104467

[2] Sliem MH, Fayyad EM, Abdullah AM, Younan NA, Al-Qahtani N, et al. Monitoring of under deposit corrosion for the oil and gas industry: A review. *Journal of Petroleum Science and Engineering*. 2021;204:108752. https://doi.org/10.1016/j.petrol.2021.108752

[3] Fang J, Cheng X, Gai H, Lin S, Lou H. Development of machine learning algorithms for predicting internal corrosion of crude oil and natural gas pipelines. *Computers & Chemical Engineering*. 2023;177:108358. https://doi.org/10.1016/j.compchemeng.2023.108358

[4] Cruz JPB, Veruz EG, Aoki IV, Schleder AM, de Souza GFM, et al. Uniform corrosion assessment in oil and gas pipelines using corrosion prediction models – Part 1: Models performance and limitations for operational field cases. *Process Safety and Environmental Protection*. 2022;167:500–515. https://doi.org/10.1016/j.psep.2022.09.034

[5] Cui B, Wang H. Analysis and prediction of pipeline corrosion defects based on data analytics of in-line inspection. *Journal of Infrastructure Preservation and Resilience*. 2023;4:14. https://doi.org/10.1186/s43065-023-00081-w

[6] Cui B, Wang H. Pipeline corrosion prediction and uncertainty analysis with an ensemble Bayesian neural network approach. *Process Safety and Environmental Protection*. 2024;187:483–494. https://doi.org/10.1016/j.psep.2024.05.011

[7] Huang Y, Qin G, Yang M. A risk-based approach to inspection planning for pipelines considering the coupling effect of corrosion and dents. *Process Safety and Environmental Protection*. 2023;180:588–600. https://doi.org/10.1016/j.psep.2023.10.025

[8] Li X, Liu Y, Han Z, Chen G. A risk-based maintenance decision model for subsea pipeline considering pitting corrosion growth. *Process Safety and Environmental Protection*. 2024;184:1306–1317. https://doi.org/10.1016/j.psep.2024.02.072

[9] Wang Y, Xie M, Su C. Multi-objective maintenance strategy for corroded pipelines considering the correlation of different failure modes. *Reliability Engineering & System Safety*. 2024;243:109894. https://doi.org/10.1016/j.ress.2023.109894

[10] Shao X, Cai B, Ahmed S, Zhou X, Hu Z, Sui Z, et al. Towards proactive corrosion management: A predictive modeling approach in pipeline industrial applications. *Process Safety and Environmental Protection*. 2024;190:1471–1480. https://doi.org/10.1016/j.psep.2024.08.021

[11] Yang L, Zhang D, Fan H, Tan Z, Xing S, Guan X, et al. Investigating the micro-turbulent corrosion mechanism of pipeline defects based on a combined experimental and simulation approach. *Journal of Natural Gas Science and Engineering*. 2022;106:104745. https://doi.org/10.1016/j.jngse.2022.104745

[12] Kumari P, Halim SZ, Kwon JSI, Quddus N. An integrated risk prediction model for corrosion-induced pipeline incidents using artificial neural network and Bayesian analysis. *Process Safety and Environmental Protection*. 2022;167:34–44. https://doi.org/10.1016/j.psep.2022.07.053

[13] Aryai V, Baji H, Mahmoodian M. Failure assessment of corrosion affected pipeline networks with limited failure data availability. *Process Safety and Environmental Protection*. 2022;157:306–319. https://doi.org/10.1016/j.psep.2021.11.024

[14] Mensah A, Sriramula S. Probabilistic finite element-based reliability of corroded pipelines with interacting corrosion cluster defects. *International Journal of Pressure Vessels and Piping*. 2024;207:105086. https://doi.org/10.1016/j.ijpvp.2023.105086

[15] Mensah A, Sriramula S. Stochastic finite element-based reliability of corroded pipelines with interacting corrosion clusters. *Process Safety and Environmental Protection*. 2024;187:206–220. https://doi.org/10.1016/j.psep.2024.04.130

[16] Guang Y, Wang W, Song H, Mi H, Tang J, Zhao Z. Prediction of external corrosion rate for buried oil and gas pipelines: A novel deep learning method with DNN and attention mechanism. *International Journal of Pressure Vessels and Piping*. 2024;209:105218. https://doi.org/10.1016/j.ijpvp.2024.105218

[17] Wang Q, Song Y, Zhang X, Dong L, Xi Y, Zeng D, et al. Evolution of corrosion prediction models for oil and gas pipelines: From empirical-driven to data-driven. *Engineering Failure Analysis*. 2023;107097. https://doi.org/10.1016/j.engfailanal.2023.107097

[18] Soomro AA, Mokhtar AA, Hussin HB, Lashari N, Oladosu TL, Jameel SM, et al. Analysis of machine learning models and data sources to forecast burst pressure of petroleum corroded pipelines: A comprehensive review. *Engineering Failure Analysis*. 2024;155:107747. https://doi.org/10.1016/j.engfailanal.2023.107747

[19] Wang Q, Lu H. A novel stacking ensemble learner for predicting residual strength of corroded pipelines. *npj Materials Degradation*. 2024;8:87. https://doi.org/10.1038/s41529-024-00508-z

[20] Mahmood Y, Chen J, Yodo N, Huang Y. Optimizing natural gas pipeline risk assessment using hybrid fuzzy bayesian networks and expert elicitation for effective decision-making strategies. *Gas Science and Engineering*. 2024;125:205283. https://doi.org/10.1016/j.jgsce.2024.205283

[21] Sharma M, Liu H, Tsesmetzis N, Handy J, Place T, Gieg LM. Diagnosing microbiologically influenced corrosion at a crude oil pipeline facility leak site – A multiple lines of evidence approach. *International Biodeterioration & Biodegradation*. 2022;172:105438. https://doi.org/10.1016/j.ibiod.2022.105438

[22] Abilio ADA, Wolodko JD, Eckert RB, Skovhus TL. Development of an expert system for assessing failures in oil and gas pipelines due to microbiologically influenced corrosion (MIC). *Engineering Failure Analysis*. 2024;163:108426. https://doi.org/10.1016/j.engfailanal.2024.108426

[23] Chohan IM, Ahmad A, Sallih N, Bheel N, Salilew WM, et al. Effect of seawater salinity, pH, and temperature on external corrosion behavior and microhardness of offshore oil and gas pipeline: RSM modelling and optimization. *Scientific Reports*. 2024;14:16543. https://doi.org/10.1038/s41598-024-67463-2

[24] Wang T, Feng K, Ling J, Liao M, Yang C, Neubeck R, et al. Pipeline condition monitoring towards digital twin system: A case study. *Journal of Manufacturing Systems*. 2024;73:256–274. https://doi.org/10.1016/j.jmsy.2024.02.006

[25] Shen Y, Zhou W. Development of a convolutional neural network model to predict the size and location of corrosion defects on pipelines based on magnetic flux leakage signals. *International Journal of Pressure Vessels and Piping*. 2024;207:105123. https://doi.org/10.1016/j.ijpvp.2023.105123

[26] Chen P. Advancements and future outlook of safety monitoring, inspection and assessment technologies for oil and gas pipeline networks. *Journal of Pipeline Science and Engineering*. 2025;5:100267. https://doi.org/10.1016/j.jpse.2025.100267

[27] Zheng Y, et al. Random risk assessment model and risk-based maintenance decisions for natural gas pipelines. *Journal of Loss Prevention in the Process Industries*. 2025;95:105591. https://doi.org/10.1016/j.jlp.2025.105591

[28] Wahono T, Purniawan A, Mukhlash I, Putri ERM. Risk-based asset integrity management in the oil and gas industry from traditional to machine learning approaches: A systematic review. *Results in Engineering*. 2025;28:107287. https://doi.org/10.1016/j.rineng.2025.107287

[29] Hu J. Prediction of the internal corrosion rate for oil and gas pipelines and influence factor analysis with interpretable ensemble learning. *International Journal of Pressure Vessels and Piping*. 2024;212:105329. https://doi.org/10.1016/j.ijpvp.2024.105329

[30] Xie R, Fan Z, Hao X, Luo W, Li Y, Zhao Y, et al. Prediction model of corrosion rate for oil and gas pipelines based on knowledge graph and neural network. *Processes*. 2024;12:2367. https://doi.org/10.3390/pr12112367