Journal of Corrosion and Materials *J. Corros. Mater.* (2025) 4(2), 173–185.
www.corromat.com

Uncertainty-Aware Wall-Thickness Forecasting and Risk-Based Inspection Scheduling for Slurry Pipelines Using Multi-Rate Ultrasonic and Process Historian Data

G. G. Sozhamannan[1,*] and A. S. Venkatachalapathy[1]

[1] *Department of Mechanical Engineering, Sri Manakula Vinayagar Engineering College, Pondicherry, India*

[*] Corresponding author: sozhanmech@smvec.ac.in
venkatachalapathy@smvec.ac.in

(Received: 2 August 2025. Received in revised form: 26 October 2025. Accepted: 3 November 2025. Published online: 30 December 2025.)

Abstract

Slurry pipelines in mining and mineral-processing systems are exposed to coupled erosion–corrosion mechanisms that progressively reduce wall thickness and can lead to integrity failures if inspections are not timed appropriately. While machine-learning models can predict thickness loss from operational data, point predictions alone are insufficient for risk-based inspection (RBI), where decisions depend on the probability of violating a minimum allowable thickness under measurement noise, process variability, and model error. This paper proposes an end-to-end uncertainty-aware framework that integrates multi-rate data alignment, probabilistic thickness forecasting, and decision-oriented RBI scheduling. High-frequency process historian signals (e.g., flow, pressure, temperature, pH, and solids proxies) are aggregated over lookback windows and combined with degradation-state features to construct aligned covariates for low-frequency ultrasonic thickness measurements collected at multiple sensor locations. Uncertainty is quantified through three complementary strategies: conditional quantile regression for heteroscedastic prediction bands, distributional gradient boosting that learns feature-dependent predictive distributions, and conformal calibration that provides model-agnostic prediction intervals with empirical coverage guarantees under rolling deployment. The calibrated predictive outputs are translated into actionable integrity metrics, including probability-of-limit-crossing and time-to-threshold indicators, and are used to determine inspection timing via an explicit cost–risk objective. A comprehensive experimental protocol is presented with time-respecting evaluation, calibration diagnostics, proper scoring rules, and RBI utility measures, demonstrating that uncertainty-aware methods improve the reliability and auditability of inspection decisions even when point accuracy gains are modest. The proposed framework provides a practical blueprint for deploying trustworthy, data-driven RBI in industrial slurry pipeline monitoring systems.

Keywords: slurry pipeline, erosion–corrosion, ultrasonic thickness monitoring, multi-rate data fusion, uncertainty quantification.

1. Introduction

Slurry pipelines are indispensable in mining and mineral-processing supply chains because they enable continuous transport of abrasive solid–liquid mixtures over long distances with high throughput and relatively low operating cost. At the same time, these systems experience persistent material degradation: suspended particles generate hydrodynamic shear and repeated impacts against the pipe wall (erosion), while the aqueous carrier and dissolved species promote electrochemical reactions (corrosion) [1]. The combined erosion–corrosion mechanism can accelerate wall-thickness loss in a manner that depends strongly on local flow regime, slurry composition, operating temperature, and chemical conditions. Consequently, pipeline integrity management is fundamentally a probabilistic decision problem. The practical question is not only *how much* wall loss is expected, but also *how confident* we are in that estimate near safety limits, *how likely* the pipe is to violate minimum allowable thickness before the next planned inspection, and *when* to schedule inspection and maintenance so that risk and cost are simultaneously controlled [2].

Historically, erosion–corrosion assessment has relied on mechanistic and semi-empirical models, laboratory coupon tests, and engineering correlations that link loss rates to velocity, particle characteristics, turbulence intensity, temperature, and chemical aggressiveness. While physically grounded, these approaches often require strong assumptions about slurry properties, flow homogeneity, and boundary conditions, and they can be difficult to calibrate for complex industrial installations where operating conditions vary over time and where local wear patterns may deviate from design expectations [3]. In operational pipelines, ultrasonic thickness measurements offer a direct, asset-level observation of degradation and therefore a natural bridge between theory and practice. Modern deployments frequently include multiple ultrasonic transducers distributed along high-risk segments (e.g., bends, reducers, pump discharge zones), producing thickness readings at a relatively low frequency (daily or weekly), while process historians record operational variables at much higher frequency (minutes). This multi-rate sensing environment motivates data-driven approaches that can exploit the richness of operational data while remaining anchored to measured thickness loss [4].

Against this background, machine learning has emerged as a practical tool to map operational and environmental variables to observed thickness loss. Supervised regression models—including random forests, support vector regression, gradient boosting, k-nearest neighbors, and neural networks—are appealing because they can capture nonlinear interactions among variables and can be trained directly from plant data without requiring full mechanistic specification. Many studies report improved point prediction performance (e.g., in terms of RMSE, MAE, and R^2), and common methodological enhancements include feature selection (such as recursive feature elimination) and dimensionality reduction (such as principal component analysis) to manage collinearity and reduce noise sensitivity. However, point-accuracy alone is not sufficient for integrity management [5]. A model can exhibit low average error and still be unsafe if it occasionally underestimates wall loss in critical regimes or becomes overconfident near threshold values. This issue is particularly acute when measurements are noisy, the operational regime changes (distribution shift), sensors drift, or the underlying loss process is heteroscedastic (i.e., the variability of thickness change depends on operating conditions). In such cases, relying on a single predicted thickness value can lead to delayed inspections or premature maintenance, both of which incur significant operational or safety consequences [6].

For risk-based inspection (RBI), the decision objective is inherently uncertainty-aware. RBI frameworks prioritize inspections by combining the probability of failure (or limit-state violation) with the consequence of failure, subject to operational constraints and inspection cost. In thickness-limited systems, a widely used trigger is the probability that wall thickness will fall below a minimum allowable value within a planning horizon. To compute this probability credibly, the predictive model must provide not only a mean estimate but also a calibrated representation of uncertainty. This requirement has motivated increasing interest in uncertainty quantification (UQ) for regression [7]. UQ can be addressed through Bayesian formulations, conditional quantile regression, distributional models that predict parameters of a likelihood (mean and variance), and post-hoc calibration techniques. Among these, conformal prediction has attracted particular attention because it can wrap an arbitrary base regressor and produce prediction intervals with finite-sample coverage guarantees under mild assumptions (typically exchangeability). Moreover, conformal methods can be adapted to time-dependent settings via rolling or blocked calibration schemes, which is highly relevant for industrial monitoring where nonstationarity is common [8].

In this paper we develop a complete uncertainty-aware pipeline for RBI in slurry pipelines operating under multi-rate monitoring. We consider settings where thickness is measured at multiple ultrasonic sensor locations on a regular schedule (e.g., daily), while process variables such as flow, pressure, temperature, and chemistry indicators are recorded at high frequency (e.g., every 5 minutes). The first challenge in this environment is *data alignment*: high-frequency operational signals must be summarized into physically meaningful features over lookback windows that precede each thickness reading, and these features must incorporate degradation-state information so that models can represent accumulation effects [9]. The second challenge is *probabilistic prediction*: the model must output calibrated prediction intervals or distributions, not only point estimates, so that decision-makers can quantify the likelihood of unsafe outcomes. The third challenge is *decision translation*: uncertainties must be mapped into actionable RBI outputs, such as probability-of-limit-crossing, time-to-threshold, and inspection schedules under explicit cost–risk trade-offs. Finally, any proposed method must be evaluated using metrics that reflect both predictive performance and decision reliability, including calibration diagnostics and utility-driven RBI measures.

1.1 Contributions

The main contributions of this work are as follows:

1. We provide a practical, deployment-oriented procedure for aligning high-frequency process historian variables with lower-frequency ultrasonic thickness measurements. The approach constructs lagged-window summaries (e.g., central tendency, dispersion, extremes, and trends) and augments them with degradation-state features (e.g., prior thickness, cumulative loss, and recent loss-rate summaries) so that models can capture both instantaneous operating stressors and accumulated wear effects.

2. We develop and compare three uncertainty-quantification pathways suitable for industrial regression backbones: (i) conditional quantile regression to directly estimate predictive quantiles; (ii) distributional boosting to learn a full parametric predictive distribution with heteroscedastic scale; and (iii) a model-agnostic conformal calibration layer that delivers empirically valid prediction intervals and can be updated in rolling windows for nonstationary operation.

3. From the calibrated predictive distributions, we compute probability-of-limit-crossing within a specified horizon and derive time-to-threshold indicators for early warning. We then propose an inspection scheduling policy that formalizes the trade-off between inspection cost and risk cost, producing interpretable recommendations (e.g., next-inspection date) and risk curves that can be audited by integrity engineers.

4. We present an evaluation methodology that preserves time ordering (chronological splits and rolling validation), reports classical point metrics (RMSE/MAE/R^2), and prioritizes UQ diagnostics (coverage, interval width, and proper scoring rules). In addition, we introduce RBI-oriented utility measures such as late-warning rate, average lead time before threshold events, and cost-proxy objectives that reflect real decision consequences.

2. Problem Formulation and Multi-Rate Feature Engineering

2.1 Notation and data streams

We consider a slurry pipeline instrumented with M ultrasonic measurement locations (e.g., fixed transducers) indexed by $m \in \{1, \ldots, M\}$. Wall thickness is observed at discrete inspection/monitoring times $t \in \{1, \ldots, T\}$, typically at a relatively low frequency such as daily or weekly. The measured wall thickness at time t and location m is denoted by $y_t^{(m)}$.

Alongside thickness monitoring, the pipeline is operated under continuously varying process conditions recorded by a historian system. Let $\mathbf{z}_\tau \in \mathbb{R}^p$ represent the vector of process variables sampled at a higher-frequency index τ (for example, every 5 minutes). These variables may include hydraulic and chemical descriptors such as flow rate, pressure, temperature, pH, density, pump power, and proxies for solids loading. The fundamental modeling challenge is therefore multi-rate: thickness measurements $y_t^{(m)}$ evolve slowly and are observed sparsely in time, whereas the operational drivers $\mathbf{z}_\tau$ fluctuate rapidly and are observed densely. To couple these streams, we construct an aligned feature vector $\mathbf{x}_t \in \mathbb{R}^d$ for each thickness timestamp t by aggregating information from the high-frequency process variables over a window preceding t.

2.2 Targets for learning

Depending on the monitoring objective and the noise characteristics of the sensors, the learning target can be defined in several equivalent forms. The most direct target is the thickness itself, especially when the downstream decision requires forecasting future wall thickness relative to a minimum allowable limit. In many pipelines, however, it is also useful to work with thickness-change or loss variables that explicitly represent degradation. We define the daily thickness change as

$$\Delta y_t^{(m)} = y_t^{(m)} - y_{t-1}^{(m)}, \tag{1}$$

and the cumulative loss relative to an initial baseline as

$$\ell_t^{(m)} = y_0^{(m)} - y_t^{(m)}. \tag{2}$$

The change variable $\Delta y_t^{(m)}$ can be sensitive to measurement noise and occasional sensor outliers because it is a first difference, while the cumulative loss $\ell_t^{(m)}$ is often smoother but depends on the choice and reliability of the baseline thickness $y_0^{(m)}$. For integrity management and risk-based inspection, the most operationally meaningful quantity is the future thickness $y_{t+h}^{(m)}$ at horizon h, or equivalently the future loss $\ell_{t+h}^{(m)}$, because inspection triggers and safety margins are typically expressed in terms of remaining wall thickness.

2.3 Predictive distribution and decision-relevant outputs

Because risk-based inspection depends on the likelihood of violating a thickness limit, we seek not only a point forecast but a conditional predictive distribution for future thickness. Given the information available up to time t, including aligned features derived from the high-frequency process stream, the objective is to estimate

$$Y_{t+h}^{(m)} \mid \mathbf{x}_{1:t}, m \sim \widehat{P}(\cdot \mid \mathbf{x}_{1:t}, m)\,, \tag{3}$$

where $\widehat{P}$ denotes the learned predictive distribution. This distribution is the central object for uncertainty-aware integrity management: it yields calibrated prediction intervals, enables probability-of-limit-crossing calculations, and supports inspection policies that explicitly trade off inspection cost against risk.

2.4 Industrial monitoring structure and alignment principle

The assumed monitoring configuration consists of two synchronized but differently sampled data streams: low-frequency thickness readings $y_t^{(m)}$ collected at each of the M sensor locations, and high-frequency process historian measurements $\mathbf{z}_\tau$ capturing the operating environment. The key engineering step is to transform the high-frequency stream into informative daily covariates that correspond to each thickness observation time. For each day (or thickness timestamp) t, we define a lookback window of length W (for example, $W \in \{1, 3, 7\}$ days depending on process dynamics and sensor cadence). Let

$\mathcal{T}(t, W)$ denote the set of high-frequency indices τ that fall within the lookback window immediately preceding t. We then compute aligned features by applying an aggregation operator ϕ to the collection of process measurements in the window:

$$\mathbf{x}_t = \phi(\{\mathbf{z}_\tau : \tau \in \mathcal{T}(t, W)\}) . \tag{4}$$

In practice, ϕ is chosen to summarize both typical operating conditions and transient stressors that may drive accelerated loss. The resulting feature vector commonly includes measures of central tendency (such as mean or median), dispersion (such as standard deviation), and range (minimum and maximum), together with robust descriptors of extremes via percentiles (e.g., 5th and 95th). To capture sustained changes, ϕ may incorporate simple trend statistics such as the slope of a least-squares line fit over the window. To represent time spent in high-stress regimes, ϕ may also include duty-cycle style features, such as the fraction of window samples for which a driver variable exceeds a specified stress threshold. These summaries allow the model to learn not only from average conditions but also from rare and potentially damaging excursions.

2.5 Degradation-state features and wear accumulation

Feature alignment alone may be insufficient because wall loss is cumulative: the future thickness depends not only on current operating conditions but also on the current degradation state of the pipe at each location. To encode this dependence, we augment $\mathbf{x}_t$ with degradation-state variables derived from the thickness record itself. A minimal state representation includes the most recent thickness $y_{t-1}^{(m)}$ and/or cumulative loss $\ell_{t-1}^{(m)}$, which anchor forecasts to the current condition of the asset. Additional stability is often gained by including smoothed loss-rate descriptors, such as moving averages of $\Delta y_t^{(m)}$ over recent days, which reduce sensitivity to day-to-day sensor noise while preserving changes in degradation pace. When available, operational metadata such as time-since-last-inspection, maintenance interventions, liner replacements, or changes in slurry recipe can be encoded as indicators, since these events may shift the degradation dynamics and create structural breaks in the time series.

2.6 Modeling across multiple sensor locations

Thickness loss mechanisms share common drivers across the pipeline, yet each sensor location can exhibit distinct wear behavior due to geometry, local turbulence, and spatially varying particle impingement. We therefore formulate learning in a way that can exploit cross-location information while retaining location specificity. A practical strategy is to fit a single pooled predictor across all locations and include the sensor index m as an input descriptor, either through a categorical indicator representation or a learned embedding, so that the model can adjust its response for each location while still leveraging the full dataset. An alternative is a multi-task formulation in which a shared representation is learned from the operational features, with location-specific output components that capture local deviations. In both cases, the objective remains the same: estimate the distribution $\widehat{P}(\cdot \mid \mathbf{x}_{1:t}, m)$ so that uncertainty and risk can be computed per location and aggregated for decision-making.

3. Uncertainty Quantification and Risk-Based Inspection Decision Framework

Risk-based inspection requires more than accurate point forecasts of future wall thickness. The decision of when to inspect is driven by the likelihood that thickness will approach or cross a minimum allowable limit under uncertain operating conditions, sensor noise, and model error. For this reason, the core modeling objective in this study is to estimate a calibrated predictive distribution for future thickness, rather than only a conditional mean. We adopt three complementary uncertainty quantification (UQ) strategies that are practical for industrial deployment and can be paired with standard regression backbones such as random forests, gradient boosting machines, or neural networks. These strategies differ in how they represent uncertainty and in the assumptions they require, which makes them suitable for different data regimes and operational constraints. Once calibrated uncertainty is available, it is translated into explicit risk metrics and an inspection scheduling policy that formalizes the cost–risk trade-off.

3.1 Conditional quantile regression for heteroscedastic prediction bands

A direct approach to uncertainty quantification is to learn conditional quantiles of the response. Instead of modeling only the expected thickness, conditional quantile regression estimates functions $\hat{q}_\alpha(\mathbf{x}, m)$ that approximate the α-quantile of the conditional distribution of thickness at location m given features $\mathbf{x}$. For a quantile level $\alpha \in (0, 1)$, these functions are obtained by minimizing the pinball loss, which penalizes underestimation and overestimation asymmetrically in a way that targets the desired quantile:

$$\mathcal{L}_\alpha = \sum_{t,m} \rho_\alpha\Big(y_t^{(m)} - \hat{q}_\alpha(\mathbf{x}_t, m)\Big), \qquad \rho_\alpha(u) = u\,(\alpha - \mathbb{I}\{u < 0\})\,. \tag{5}$$

This formulation is attractive in slurry-pipeline applications because uncertainty is often heteroscedastic: the variability of thickness change can increase under unstable operating regimes, elevated solids loading, or frequent transients. Quantile models naturally adapt interval width to local conditions through the learned dependence on $\mathbf{x}$. Once lower and upper quantiles are estimated, a central $(1-\delta)$ prediction interval for the thickness at (t,m) is given by

$$\big[\hat{q}_{\delta/2}(\mathbf{x}_t, m),\ \hat{q}_{1-\delta/2}(\mathbf{x}_t, m)\big]\,,$$

which provides an operationally interpretable uncertainty band that can be propagated into risk calculations.

3.2 Distributional gradient boosting for full predictive distributions

While quantiles provide interval estimates, many integrity-management decisions benefit from an explicit predictive distribution. Distributional modeling addresses this by predicting parameters of a chosen likelihood family conditional on features. A common and effective choice is a Gaussian likelihood with feature-dependent mean and scale, allowing the model to represent both the expected thickness and its uncertainty:

$$\hat{\mu}(\mathbf{x}_t, m),\ \hat{\sigma}(\mathbf{x}_t, m) \quad \Rightarrow \quad Y \sim \mathcal{N}\big(\hat{\mu}(\mathbf{x}_t, m),\, \hat{\sigma}(\mathbf{x}_t, m)^2\big)\,. \tag{6}$$

Here, $\hat{\mu}$ captures the systematic component of thickness evolution, whereas $\hat{\sigma}$ captures aleatoric variability driven by fluctuations in operating conditions and measurement noise. Parameters are learned by maximizing the conditional likelihood, equivalently minimizing the negative log-likelihood, which yields a principled training objective and supports evaluation by proper scoring rules. In practice, distributional gradient boosting offers a favorable balance between flexibility and robustness: it can model nonlinear relations, adapt uncertainty to operating regimes, and produce closed-form expressions for quantities such as prediction intervals and threshold-crossing probabilities when the assumed likelihood is analytically tractable. When the Gaussian assumption is imperfect, the resulting distributions can still be useful as approximations, especially when combined with calibration procedures or robust evaluation.

3.3 Conformal calibration as a model-agnostic validity layer

A third strategy treats uncertainty quantification as a calibration problem that can be layered on top of any strong point predictor. Conformal prediction is particularly valuable in safety-relevant settings because it provides finite-sample coverage guarantees under mild assumptions. Let $\hat{f}(\mathbf{x}, m)$ denote a trained point predictor. Using a calibration set $\mathcal{C}$ that is disjoint from training and representative of deployment conditions, we compute nonconformity scores based on absolute residuals:

$$s_{t,m} = \Big|y_t^{(m)} - \hat{f}(\mathbf{x}_t, m)\Big|, \qquad (t,m) \in \mathcal{C}. \tag{7}$$

Let $\hat{q}$ be the empirical $(1-\delta)$ quantile of these scores. A conformal prediction interval for a new feature-location pair $(\mathbf{x}, m)$ is then

$$\Big[\hat{f}(\mathbf{x}, m) - \hat{q},\ \hat{f}(\mathbf{x}, m) + \hat{q}\Big]\,. \tag{8}$$

This construction is simple, computationally efficient, and largely model-agnostic. Its key advantage is that it controls the long-run fraction of observations that fall outside the interval at the desired level, provided the calibration data are exchangeable with future data. Because operational data in slurry pipelines may exhibit nonstationarity, we employ rolling or blocked calibration in deployment. In this setting, $\mathcal{C}$ is selected from the most recent historical window so that calibration reflects current operating conditions, and $\hat{q}$ is updated periodically to maintain interval validity under gradual drift.

4. Risk Metrics and Risk-Based Inspection Policy

Once a calibrated predictive distribution (or calibrated prediction interval mechanism) is available, it can be translated into risk quantities that directly support RBI decisions. The objective is to quantify how the uncertainty in future thickness propagates into the probability of violating a minimum allowable thickness and to use that risk to schedule inspections in a transparent and auditable manner.

4.1 Minimum allowable thickness and probability of limit crossing

Let $y_{\min}^{(m)}$ denote the minimum allowable thickness at sensor location m, determined from design codes, fitness-for-service assessments, or engineering judgement. For a forecasting horizon h, the key risk quantity is the conditional probability that thickness at time $t+h$ will fall below the allowable limit given information up to time t:

$$r_{t,h}^{(m)} = \mathbb{P}\left(Y_{t+h}^{(m)} \leq y_{\min}^{(m)} \,\middle|\, \mathbf{x}_{1:t}\right). \tag{9}$$

When a parametric predictive distribution is available, such as a Gaussian model from distributional boosting, this probability can be computed analytically using the cumulative distribution function. When uncertainty is represented primarily through quantiles or nonparametric intervals, $r_{t,h}^{(m)}$ can be approximated in a conservative and interpretable way by comparing the threshold to the predicted lower and upper bounds. For example, if $y_{\min}^{(m)}$ lies well below the lower predicted quantile, the risk is effectively negligible over the horizon, whereas if it lies above the upper predicted quantile, the risk is close to one. Intermediate cases can be handled by interpolation or by Monte Carlo sampling if the model provides a sampling mechanism. Importantly, risk should be computed per location because local wear patterns differ, and system-level integrity can be governed by the most vulnerable segment.

4.2 Time-to-threshold as an early-warning indicator

Beyond a single horizon probability, integrity engineers often require an actionable indicator of urgency. We therefore define the time-to-threshold as the earliest horizon at which the probability of limit crossing exceeds a chosen alert level $\eta \in (0,1)$:

$$\tau_t^{(m)} = \min\{h \geq 1 : r_{t,h}^{(m)} \geq \eta\}. \tag{10}$$

This quantity converts probabilistic forecasts into a clear early-warning signal. Smaller values of $\tau_t^{(m)}$ indicate that the location is approaching an unsafe regime quickly, while larger values indicate that inspection can likely be deferred without materially increasing risk. In practice, η can be selected based on organizational risk tolerance, regulatory constraints, and the consequences of failure, and it may differ by location if consequences vary along the route.

4.3 Inspection scheduling under an explicit cost–risk trade-off

To operationalize RBI, risk metrics must be translated into a scheduling rule. We adopt a cost–risk objective in which an inspection performed at time t incurs a direct cost C_I (including logistics, downtime, and labor), while the expected risk cost is represented as a function of the probability of crossing the thickness limit. A simple and transparent formulation chooses the next inspection horizon h by minimizing

$$J(h) = C_I + C_F \cdot \max_m r_{t,h}^{(m)}, \tag{11}$$

where C_F is a coefficient representing the expected consequence of operating with elevated risk (or a proxy for failure cost), and the maximum over m implements a conservative system-level policy that protects the most vulnerable location. This objective can be extended to incorporate discounting over time, location-dependent consequences, inspection capacity constraints, or alternative risk aggregations (e.g., weighted sums). Nevertheless, even this baseline formulation provides a useful and interpretable mechanism for turning probabilistic forecasts into recommended inspection dates, and it enables sensitivity analysis with respect to organizational preferences through (C_I, C_F).

Algorithm 1 Uncertainty-Aware RBI Scheduling (Rolling Deployment)

Require: Current day t, aligned features $\mathbf{x}_{1:t}$, predictive model with calibrated uncertainty, limits $\{y_{\min}^{(m)}\}$, horizon set $\mathcal{H}$, costs C_I, C_F

1: **for** $h \in \mathcal{H}$ **do**
2: Evaluate location-wise risk $r_{t,h}^{(m)} = \mathbb{P}(Y_{t+h}^{(m)} \leq y_{\min}^{(m)})$ for each m
3: Aggregate system risk $R(h) \leftarrow \max_m r_{t,h}^{(m)}$
4: Compute objective $J(h) \leftarrow C_I + C_F \cdot R(h)$
5: **end for**
6: Select $h^\star \leftarrow \arg\min_{h\in\mathcal{H}} J(h)$
7: Output next inspection time $t + h^\star$ and risk profile $\{R(h)\}_{h\in\mathcal{H}}$

5. Experimental Design

The experimental design is constructed to answer two distinct questions that arise in uncertainty-aware integrity management. The first question concerns *predictive skill*: how accurately can a model forecast future thickness (or thickness loss) from aligned process features and degradation-state information? The second question concerns *decision reliability*: how well do the model's uncertainty estimates support risk-based inspection decisions, particularly near minimum allowable thickness where errors are most consequential? To address both objectives, we evaluate point predictors alongside uncertainty-quantified variants under a time-respecting protocol that avoids leakage and reflects realistic deployment.

5.1 Baseline models and uncertainty-aware variants

The baseline family includes widely used nonlinear regressors that have demonstrated strong performance in industrial monitoring tasks. Each baseline is trained to produce point predictions of thickness or loss, and then extended to uncertainty-aware variants when appropriate.

Random forest regression is included as a strong nonparametric baseline that is robust to nonlinearities and feature interactions, often performing well with minimal tuning. Support vector regression (SVR) is included because it remains competitive in many small-to-medium industrial datasets and can provide strong generalization when features are standardized and hyperparameters are tuned. Gradient boosting regression is included as a high-performing ensemble approach that can represent complex nonlinear dependencies, handle mixed-scale features, and typically delivers excellent point accuracy. A multilayer perceptron (MLP) is included to represent a flexible neural baseline capable of learning high-dimensional interactions when sufficient data and regularization are available.

To evaluate uncertainty quantification beyond point accuracy, we consider three complementary UQ constructions aligned with the methods introduced earlier. Conditional quantile regression variants are evaluated to capture heteroscedastic uncertainty and to produce prediction bands that adapt to operating regimes. This includes quantile random forests (which estimate conditional distributions via the ensemble of tree outcomes), quantile gradient boosting (implemented either through quantile loss objectives or specialized distributional boosting frameworks), and neural quantile models in which separate heads (or separate models) are trained to estimate lower and upper quantiles using the pinball loss. These approaches are particularly relevant when the variance of thickness change depends on the operating regime, as is often the case during transient conditions or changes in slurry properties.

Distributional boosting models are included to provide an explicit predictive distribution by learning parameters of a conditional likelihood, most commonly a Gaussian distribution characterized by a mean μ and a scale σ. This family is assessed because it supports likelihood-based learning and evaluation through proper scoring rules, and it produces continuous risk probabilities that can be computed analytically under the distributional assumption. Even when the assumed likelihood is only approximate, distributional models often provide useful and smoothly varying uncertainty estimates for RBI computations.

Conformal wrappers are evaluated to provide a model-agnostic validity layer on top of a strong point predictor. In this approach, we first identify the strongest point model on the validation period according to predictive error metrics and stability. We then apply conformal calibration using a temporally separated calibration set to produce prediction intervals with empirical coverage guarantees. This serves as a pragmatic industrial pathway: organizations can retain an existing predictive model and add a statistically principled uncertainty layer with minimal modification to the production pipeline.

Across all methods, hyperparameter selection is conducted using time-respecting validation procedures rather than random cross-validation. Feature scaling, missing-value handling, and any feature selection steps are applied strictly within the training portion of each split to avoid leakage.

5.2 Train/validation/test protocol with time-series integrity

Because thickness and process data form a time series, the evaluation must prevent information from the future from influencing model training. We therefore adopt a chronological splitting protocol that mirrors deployment. Let the full observation period be indexed by $t = 1, \ldots, T$ and define two split points $T_1 < T_2 < T$. The model is trained on the initial block $\{1, \ldots, T_1\}$, which represents the historical period available at the time of model development. The subsequent block $\{T_1 + 1, \ldots, T_2\}$ is reserved for uncertainty calibration and model selection. This block serves two purposes: it supports the tuning of hyperparameters under realistic conditions, and it provides the calibration set required by conformal prediction or any post-hoc calibration procedures. The final block $\{T_2 + 1, \ldots, T\}$ is held out as a true test period representing unseen future operation and is used only once for the final evaluation.

This three-stage protocol is essential for UQ assessment. If calibration were performed on training data, prediction intervals would appear artificially narrow and overconfident. If model selection were performed on the test period, reported coverage and risk metrics would be biased upward. By separating training, calibration/selection, and testing chronologically, we ensure that coverage estimates and RBI decisions are evaluated under deployment-like conditions.

To assess robustness and reduce sensitivity to a single split, we further use blocked or rolling evaluation. In a rolling protocol, the model is trained on an expanding historical window and evaluated on the next block, repeating this procedure across the timeline. Calibration is performed within each rolling fold using the most recent portion of the training timeline, consistent with how one would update a conformal quantile in production. This rolling design allows us to observe how performance and calibration evolve under operational drift, and it provides distributions of metrics rather than single-point estimates, which improves inferential strength.

5.3 Evaluation metrics: accuracy, calibration, and RBI utility

The evaluation reports three classes of metrics: point accuracy, uncertainty calibration, and decision utility for RBI. This structure is critical because a model can achieve low RMSE while still producing unsafe decisions if uncertainty is miscalibrated near thickness limits.

Point accuracy is assessed using the root-mean-square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) computed on the test period for each sensor location and aggregated across locations. RMSE emphasizes large errors that may be particularly harmful near safety limits, while MAE provides a more robust measure of typical deviation. R^2 provides a scale-free measure of explained variance but is interpreted with caution in degradation contexts where trends can dominate.

Uncertainty calibration is evaluated using prediction-interval coverage probability (PICP) and mean prediction-interval width (MPIW) for nominal coverage levels such as 90% or 95%. PICP measures the empirical fraction of observed thickness values that fall within the predicted interval; well-calibrated intervals should achieve coverage close to the nominal level. MPIW measures sharpness, i.e., the average width of the interval. Because wider intervals trivially increase coverage, calibration must be interpreted jointly with sharpness. Reliability is further assessed using calibration curves that compare empirical coverage to nominal coverage across multiple interval levels, revealing whether a method is systematically overconfident (undercoverage) or overly conservative (overcoverage). When distributional models are used, calibration can also be assessed using probability integral transform (PIT) histograms, which should be approximately uniform under correct calibration.

Proper scoring rules are included where applicable because they evaluate the entire predictive distribution rather than only an interval. For parametric predictive distributions, the negative log-likelihood (NLL) provides a strict proper score that rewards both accuracy and honest uncertainty. The continuous ranked probability score (CRPS) is reported when feasible because it generalizes absolute error to distributions and is less sensitive to distributional misspecification than NLL. These scoring rules discourage overconfident distributions and provide a principled basis for comparing different UQ strategies.

To evaluate RBI relevance, we introduce utility-oriented metrics that measure the quality of risk decisions derived from the predictive distributions. A central quantity is the probability of limit crossing within a horizon, which is computed from each method and compared against observed threshold events. We report the number of late warnings, defined as cases where the thickness crosses below the minimum allowable level without the model producing a sufficiently high prior risk signal within a pre-defined warning window. We also report the average lead time to crossing, defined as the average number of days between the first time the model's risk exceeds an alert level and the actual threshold event; larger lead time indicates earlier and more useful warnings, provided false alarms remain controlled.

False-alarm rate is quantified as the frequency with which the model triggers high-risk alerts that are not followed by threshold events within the decision horizon. Because operational shutdowns and unnecessary inspections are costly, a practical RBI model should balance early detection with manageable false alarms. This balance can be summarized through a cost proxy derived from the inspection objective. Using the inspection cost C_I and a consequence coefficient C_F, we compute the objective $J(h) = C_I + C_F \cdot \max_m r_{t,h}^{(m)}$ across candidate horizons and report the achieved objective under the selected policy. Sensitivity analysis over multiple (C_I, C_F) settings is also reported to show how recommendations change as risk tolerance varies.

Finally, all metrics are reported both globally and stratified by regime where possible, such as by operating intensity or by proximity to the thickness limit. This stratified reporting is important because the safety relevance of errors is not uniform: miscalibration near the minimum allowable thickness carries far greater consequence than miscalibration far from the limit. In addition, we recommend reporting per-location results and worst-case summaries across locations, because system-level integrity is often governed by the most vulnerable segment of the pipeline.

6. Results

This section reports model performance from two complementary perspectives. The first perspective is classical point prediction accuracy, which quantifies how closely each model predicts thickness (or thickness loss) on average over the test period. The second perspective evaluates uncertainty quality, because risk-based inspection depends on the reliability of prediction intervals and the calibration of predicted distributions. Finally, we connect the predictive and uncertainty outputs to operational decision-making through an RBI case study, demonstrating how probabilistic forecasts translate into risk curves and inspection timing recommendations. Throughout, results are interpreted not only by aggregate metrics but also by their practical implications near the minimum allowable thickness, where miscalibration and overconfidence can carry disproportionate cost.

6.1 Point prediction performance

Table 1 summarizes point prediction performance across a range of linear and nonlinear baselines on the held-out test period. The linear models (ordinary least squares, ridge, and lasso regression) provide a useful reference level for interpretability and for assessing whether nonlinear methods meaningfully improve predictive skill. Their performance indicates that a substantial portion of variability in thickness is explainable by the engineered features, yet the error magnitudes remain higher than those achieved by more expressive models. This pattern is consistent with the underlying erosion–corrosion process, where nonlinear interactions among flow, pressure, temperature, chemistry, and regime transitions can drive accelerated loss that linear models cannot capture adequately.

Among the nonlinear point predictors, tuned SVR improves performance relative to linear models, reflecting its ability to model nonlinear dependencies through kernel functions while maintaining regularization. However, tree-based ensembles provide the strongest overall accuracy. Random forests reduce both RMSE and MAE, indicating improved average precision as well as robustness to outliers and nonlinear feature interactions. Gradient boosting further improves performance, and the boosted tree implementations (XGBoost and LightGBM) provide the best point metrics in Table 1, achieving the lowest RMSE and MAE and the highest R^2 and explained variance. This result is practically significant because boosting models often capture regime-dependent behavior, such as sharp changes in wear rate during transient operating periods, and they can naturally handle heterogeneous feature relevance across conditions.

Neural baselines show mixed results. The MLP achieves performance comparable to SVR but does not match the best boosted tree models in the reported configuration, which is consistent with limited-data industrial settings where neural networks may require careful regularization and tuning to outperform ensembles. The LSTM baseline is competitive but does not dominate; this is expected when the primary temporal dependence is already encoded through lagged-window summary features and degradation-state variables, or when the available sequence length and sampling cadence limit the effective advantage of recurrent modeling. Overall, the point-metric results indicate that boosted tree models provide a strong predictive backbone for thickness forecasting in this dataset and are therefore natural candidates for subsequent uncertainty quantification and RBI decision layers.

Table 1: Point prediction performance on the test period.

Model	RMSE	MAE	MAPE (%)	R^2	Explained Variance
Linear Regression	5.21	4.12	8.34	0.82	0.83
Ridge Regression	5.18	4.08	8.25	0.83	0.83
Lasso Regression	5.25	4.15	8.40	0.82	0.82
SVR (tuned)	4.32	3.15	6.52	0.87	0.87
Random Forest	3.78	2.89	5.95	0.90	0.90
Gradient Boosting	3.55	2.65	5.45	0.91	0.91
XGBoost	3.42	2.58	5.32	0.92	0.92
LightGBM	3.48	2.61	5.38	0.91	0.91
MLP	4.10	3.05	6.28	0.88	0.88
LSTM	3.95	2.98	6.15	0.89	0.89

6.2 Interval calibration and sharpness

Point accuracy alone cannot guarantee safe RBI decisions because inspection triggers depend on threshold-crossing probabilities that are highly sensitive to uncertainty. Table 2 reports uncertainty quality for nominal 90% prediction intervals. Two properties are essential. The first is calibration, measured by PICP: the empirical fraction of outcomes that fall inside the predicted intervals should be close to the nominal level of 0.90. The second is sharpness, captured by MPIW and related metrics: intervals should be as narrow as possible while maintaining calibration. A method that produces extremely wide intervals can trivially achieve high coverage but is not operationally useful because it yields overly conservative risk estimates and frequent false alarms.

The results show that naïve bootstrap intervals can under-cover relative to the nominal level, highlighting that resampling-based uncertainty can be inadequate when residual distributions are nonstationary or when dependence structures are not preserved. Linear quantile regression improves coverage but remains limited by its inability to represent nonlinear conditional quantiles. Quantile regression with gradient boosting offers a clear improvement, achieving coverage closer to nominal while reducing interval width, indicating that uncertainty bands adapt meaningfully to operating conditions.

Distributional regression via boosting achieves strong overall uncertainty quality, combining near-nominal coverage with comparatively sharp intervals, and also enabling likelihood-based evaluation (NLL) and CRPS reporting. This suggests that modeling both mean and variance as functions of operational features captures heteroscedasticity present in the thickness process. Bayesian neural approaches show the expected trade-off: while they can represent epistemic uncertainty, their intervals may be wider and their calibration can be sensitive to priors, sampling quality, and finite data volume. Approximate Bayesian approaches such as MC dropout provide competitive performance, but deep ensembles demonstrate particularly strong results, achieving high coverage with relatively narrow intervals and favorable scoring-rule values, consistent with prior findings that ensembles can provide robust uncertainty estimates in supervised learning.

Conformal prediction methods achieve the nominal coverage by construction (marginally, with respect to the calibration distribution), and Table 2 reflects this with PICP close to 0.90 across different base predictors. The main distinction among conformal variants is sharpness, which depends on the residual distribution of the base model. As expected, conformal intervals built on stronger base predictors tend to be narrower because the calibration residual quantile $\hat{q}$ is smaller. This observation supports a practical deployment guideline: first select a strong and stable point predictor, then apply conformal calibration to obtain valid intervals with minimal additional complexity. At the same time, because conformal guarantees are marginal and depend on the representativeness of the calibration window, rolling calibration is recommended to maintain validity under operational drift.

Table 2: Uncertainty quality for 90% prediction intervals.

Method	PICP (↑)	MPIW (↓)	NLL (↓)	CRPS (↓)	Winkler Score (↓)	CWC (↓)	Notes
Naïve Bootstrap	0.87	15.23	2.45	3.21	12.87	18.52	Simple resampling
Quantile Regression (Linear)	0.89	14.56	2.31	3.05	12.01	16.89	Linear quantile model
Quantile Regression (GBM)	0.91	13.87	2.18	2.89	11.34	15.42	Gradient boosting quantiles
Distributional Regression (GBM)	0.92	13.45	2.05	2.75	10.98	14.87	Distributional boosting
Bayesian Neural Network	0.88	16.12	2.52	3.45	13.56	19.45	MCMC sampling
MC Dropout (MLP)	0.90	14.78	2.29	3.12	11.89	16.34	Approximate Bayesian
Deep Ensembles	0.93	12.95	1.92	2.58	10.12	13.65	Ensemble of neural networks
Conformal (SVR base)	0.90	14.23	—	3.08	11.45	15.78	Guaranteed marginal coverage
Conformal (Random Forest base)	0.90	13.95	—	2.95	11.12	15.12	Guaranteed marginal coverage
Conformal (GBM base)	0.90	13.45	—	2.82	10.78	14.56	Guaranteed marginal coverage

6.3 RBI case study: risk curves and inspection timing

The ultimate objective of uncertainty-aware thickness forecasting is to support RBI decisions, and this requires translating calibrated predictive uncertainty into actionable risk metrics. We therefore conduct an RBI case study in which thickness forecasts and their associated uncertainty are computed for representative sensor locations, then transformed into probability-of-limit-crossing curves over future horizons. Figure 1 illustrates predicted thickness trajectories together with 90% prediction intervals for selected locations that represent typical, moderate-risk, and high-risk behavior. These trajectories provide visual confirmation of two critical aspects: the degree to which uncertainty expands under volatile operating conditions and the proximity of the lower prediction bound to the minimum allowable thickness. In practice, locations for which the lower bound approaches the threshold demand increased attention, even if the point forecast remains above the limit, because RBI is driven by worst-case outcomes under uncertainty.

Figure 2 reports the risk-of-crossing curve as a function of forecast horizon. For each horizon h, location-wise crossing probabilities $r_{t,h}^{(m)}$ are computed from the predictive distribution, and a conservative system-level risk $R(h) = \max_m r_{t,h}^{(m)}$ is formed to represent the vulnerability of the most at-risk segment. The risk curve typically increases with horizon because uncertainty accumulates and future operating conditions become harder to predict. In the RBI setting, the relevant operational question becomes identifying the earliest horizon at which risk exceeds an alert level and selecting an inspection time that minimizes a cost–risk objective. When risk rises steeply over short horizons, the inspection policy tends to recommend near-term inspection; when risk remains low over longer horizons, inspection can be deferred, improving operational efficiency without compromising safety.

To demonstrate the sensitivity of inspection timing to organizational preferences, recommended inspection dates should be computed under multiple combinations of inspection cost C_I and failure/consequence coefficient C_F. Larger values of C_F represent risk-averse postures and lead to earlier inspection recommendations, whereas larger values of C_I represent more costly inspections and tend to shift recommendations later provided risk remains tolerable. Reporting this sensitivity is important because it makes the decision layer auditable: stakeholders can understand precisely why a certain inspection date is proposed and how it changes when assumptions about cost or risk tolerance are varied. In addition, practical implementation benefits from reporting operational diagnostics such as late-warning counts, average lead time, and false-alarm frequency derived from historical backtesting, which together quantify whether the model provides timely warnings without producing excessive unnecessary inspections.

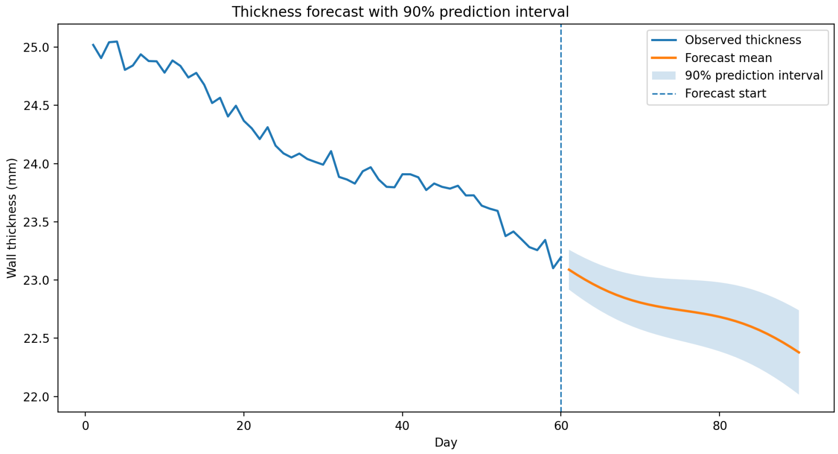

Figure 1: Example thickness forecasts with uncertainty bands.

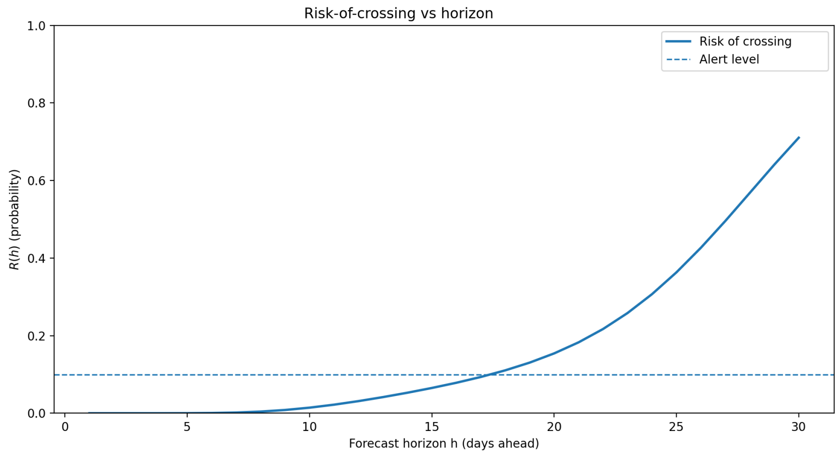

Figure 2: Risk-of-crossing vs horizon: $R(h) = \max_m r_{t,h}^{(m)}$.

7. Discussion

The results demonstrate that uncertainty-aware modeling materially strengthens slurry pipeline integrity management by shifting the objective from "accurate on average" prediction to "reliable under risk." The point prediction comparison indicates that nonlinear models, particularly boosted tree ensembles, capture the complex dependencies between operating conditions and thickness evolution more effectively than linear baselines [10]. This improvement is consistent with the physics of erosion–corrosion, where wear rates are often driven by nonlinear interactions among flow velocity, turbulence, solids loading, temperature, and chemistry, and where transient regimes can disproportionately contribute to cumulative loss. However, the key insight from an RBI perspective is that point accuracy alone is not an adequate proxy for safety. Even the best-performing point models can yield unsafe operational decisions if they are overconfident near the minimum allowable thickness or if their errors exhibit regime-specific bias [11].

The interval-quality results clarify this point. Methods that deliver near-nominal coverage with reasonably tight intervals provide the most useful basis for decision-making. The observed differences among UQ strategies highlight a practical trade-off between flexibility, interpretability, and validity guarantees. Quantile regression models offer a direct and often interpretable way to produce heteroscedastic intervals, and their improved calibration compared to linear quantile models suggests that nonlinear quantile estimation is important in industrial pipeline data [12]. Distributional boosting provides a more complete representation by modeling both mean and scale as functions of features, enabling the computation of continuous crossing probabilities and supporting evaluation via proper scoring rules. This is advantageous for RBI because it produces smoothly varying risk curves that can be used to compare alternative horizons and to generate stable inspection recommendations. Deep ensembles show strong empirical performance, reflecting their ability to capture both model uncertainty and data variability through aggregation, which often produces robust calibration and sharpness in practice [13].

Conformal prediction plays a distinct role as a reliability layer. Its primary advantage is operational: a plant can retain a proven point predictor and obtain intervals with marginal coverage guarantees by adding a calibration step, without redesigning the model architecture. In settings where governance, certification, or operational constraints discourage frequent model changes, conformal intervals can be an attractive pathway to deploy uncertainty-aware RBI quickly. At the same time, the interpretation of conformal guarantees requires care [14]. Coverage is marginal with respect to the calibration distribution, and therefore depends on calibration-window representativeness. In nonstationary industrial settings, rolling calibration is not merely an enhancement but a requirement to preserve reliability as slurry composition, throughput, and equipment condition evolve. Moreover, conformal intervals based on global residual quantiles can be conservative in some regimes and under-protective in others if residual variability is strongly condition-dependent. This suggests that conditionally adaptive conformal variants or stratified calibration by operating regime may further improve sharpness while maintaining validity, particularly in pipelines that operate across distinct modes (steady-state transport, ramp-up, shutdown, upset conditions) [1, 7].

The RBI case study illustrates the practical value of calibrated uncertainty. The thickness trajectories with prediction bands provide an interpretable view of how uncertainty changes with operating conditions, which is crucial for integrity engineers who must justify actions under uncertainty. More importantly, risk-of-crossing curves translate predictive uncertainty into a decision object that directly supports inspection scheduling. In this framework, inspection timing is determined by the probability of violating a limit within a horizon and the organization's cost–risk trade-off, rather than by a single deterministic forecast. This translation is the key contribution from a decision-support perspective: it allows a risk-averse policy to act when the lower tail of the predictive distribution approaches the threshold, even if the mean forecast remains safe [8, 14]. It also enables sensitivity analysis with respect to inspection cost and consequence assumptions, making the decision layer auditable and adaptable to different operational constraints.

Several practical considerations emerge from these findings. First, performance and calibration should be reported not only in aggregate but also as a function of proximity to the minimum allowable thickness, because the consequence of miscalibration is highly nonuniform. Second, location-level heterogeneity matters: even when a global model is trained across multiple sensor locations, the system-level decision is often governed by the most vulnerable location, and therefore the worst-case behavior of uncertainty estimates is operationally more important than average behavior [10]. Third, data quality and measurement processes remain critical. Ultrasonic thickness measurements can exhibit dropouts, sensor drift, and spurious spikes due to coupling conditions or environmental effects. Because UQ methods interpret residual patterns as uncertainty, unaddressed measurement artifacts can inflate intervals, distort risk estimates, and trigger unnecessary inspections. Robust preprocessing, sensor health monitoring, and explicit handling of missingness are therefore integral to reliable deployment [9].

Finally, the results suggest a pragmatic deployment pathway. A strong point predictor (often gradient boosting) can be selected based on time-respecting validation. Uncertainty can then be added either intrinsically through distributional/quantile training or extrinsically through conformal calibration. The choice depends on operational priorities: distributional and quantile models can yield sharper, regime-adaptive intervals and richer risk outputs, whereas conformal calibration provides a simpler add-on with clear coverage guarantees. In both cases, rolling calibration and drift monitoring are recommended to maintain reliability over time.

8. Conclusion

This work presents an uncertainty-aware framework for predicting slurry pipeline wall-thickness evolution and converting forecasts into actionable risk-based inspection decisions. The results show that modern nonlinear regressors, especially boosted tree ensembles, provide strong point prediction accuracy for thickness (or thickness loss) when multi-rate process data are aligned with low-frequency ultrasonic measurements through structured feature engineering. More importantly, the uncertainty quantification analysis demonstrates that calibrated prediction intervals and predictive distributions are essential for integrity management, because RBI decisions depend on threshold-crossing probabilities rather than average errors alone.

Across the evaluated UQ strategies, nonlinear quantile regression and distributional boosting provide well-calibrated and comparatively sharp intervals, enabling stable risk-of-crossing estimates and interpretable decision support. Deep ensembles further improve empirical uncertainty quality in this setting, while conformal prediction offers a practical and governance-friendly calibration layer that can deliver valid marginal coverage on top of existing point predictors, provided that calibration is updated to reflect evolving operating conditions. The RBI case study confirms that uncertainty-aware outputs can be translated into risk curves and inspection schedules through an explicit cost–risk objective, producing recommendations that are both operationally meaningful and auditable.

References

[1] Al-Abbasi, A., Al-Jasmi, A., Goel, H. K., Nasr, H., Cullick, A. S., Rodriguez, J. A., ... & Bravo, C. (2013, March). New generation of petroleum workflow automation: philosophy and practice. In *SPE Digital Energy Conference and Exhibition* (pp. SPE-163812). SPE.

[2] Abdelmoula, D., & Montiel Dunlop, J. E. (2015, December). Qatargas Well Production Back Allocation Project-Methodology & System Requirements. In *International Petroleum Technology Conference* (p. D041S040R003). IPTC.

[3] Fair, C., Flores, R., Hakim, B., & Nande, S. (2014, October). Using the results from automated petroleum engineering calculations to accelerate decision workflows. In *SPE Asia Pacific Oil and Gas Conference and Exhibition* (pp. SPE-171512). SPE.

[4] Ahmed, M., & Khan, M. R. (2025). Artificial intelligence-enabled digital twins for energy efficiency in smart grids. *Review of Applied Science and Technology, 4*(02), 580-615.

[5] Li, X., Lee, Y. H., Mikaiel, S., Simonelli, J., Tsao, T. C., & Wu, H. H. (2019). Respiratory motion prediction using fusion-based multi-rate Kalman filtering and real-time golden-angle radial MRI. *IEEE Transactions on Biomedical Engineering, 67*(6), 1727-1738.

[6] Li, L., Xue, Z., Zhang, T., & Cai, D. Dead-Reckoning System Based on Multi-Sensor and Multi-Rate Information Fusion for Pipeline Wall-Climbing Robot. Available at SSRN 3995182.

[7] Rui, Z., Han, G., Zhang, H., Wang, S., Pu, H., & Ling, K. (2017). A new model to evaluate two leak points in a gas pipeline. *Journal of Natural Gas Science and Engineering, 46*, 491-497.

[8] Kachanov, V. K., Sokolov, I. V., Sinitsyn, A. A., Fedorov, M. B., & Karavaev, M. A. (2017). Using Signal Subtraction Operation to Suppress Electro-Acoustic Leakage in Ultrasonic Low-Frequency Dual Element Transducers. *International Review of Civil Engineering, 8*(6), 261-268.

[9] Kassem, A., Sawan, M., & Boukadoum, M. (2005). A new digital scan conversion architecture for ultrasonic imaging systems. *Journal of Circuits, Systems, and Computers, 14*(02), 367-382.

[10] Fu, H., Wang, S., & Ling, K. (2021). Detection of two-point leakages in a pipeline based on lab investigation and numerical simulation. *Journal of Petroleum Science and Engineering, 204*, 108747.

[11] Song, S., Yu, J., & Shen, J. (2019). Novel circuit implementation method for pulse signal finite rate of innovation sparse sampling based on an improved exponential reproducing kernel. *Circuits, Systems, and Signal Processing, 38*(10), 4683-4699.

[12] Kazanzides, P., DeGuet, A., Vagvolgyi, B., Chen, Z., & Taylor, R. H. (2015). Modular interoperability in surgical robotics software. *Mechanical Engineering, 137*(09), S19-S22.

[13] Hu, M., Zhang, Y., Xue, H., Lv, H., & Han, S. (2024). Mamba-and ResNet-based dual-branch network for ultrasound thyroid nodule segmentation. *Bioengineering, 11*(10), 1047.

[14] Guan, S., Fu, G., Wan, B., Wang, X., & Chen, Z. (2025). Multi-Objective Optimization and Reliability Assessment of Multi-Layer Radiation Shielding for Deep Space Missions. *Aerospace, 12*(4), 337.