International Journal of Technology and Applied Science (IJTAS)

E-ISSN: 2230-9004 ● Website: www.ijtas.com ● Email: editor@ijtas.com

Beyond Periodic Calibration: An Uncertainty-Aware Artificial Intelligence Framework for Predictive Instrument Drift Monitoring, Measurement Reliability, and Risk-Based Calibration Optimization

Himabindhu Kondeti[1], Onuh Matthew Ijiga[2], Bhavesh Kataria[3]

[1]Independent Researcher, Department of Quality Assurance, Jost Chemical Co., Lackland Rd, St. Louis, Missouri 63114, USA, kondetibindhu@gmail.com

[2]Education Officer, Department of Physics, Joseph Sarwuan Tarka University, Makurdi, Benue, Nigeria onuhijiga@gmail.com

[3]Associate Professor, School of Computing and Technology, Institute of Advanced Research (IAR University), Gyan Marg, Near GIFT City Bridge, Gandhinagar, Gujarat 382426, India

Abstract:

Periodic calibration can leave progressive instrument drift undetected between scheduled verifications, with consequences for measurement reliability and industrial quality assurance. This paper develops an uncertainty-aware artificial intelligence framework centred on the Smart Instrument Calibration Monitoring Device (UK registered design No. 6564151), integrating multichannel calibration correction, time-dependent drift prediction, measurement uncertainty, conformity decisions, and adaptive calibration scheduling. The supplied numerical case study covers an eleven-point pressure calibration series, a 180-day drift profile, thermal-response scenarios, and comparisons of the proposed UAC-PROA approach with XGBoost, LSTM, random forest, and support vector regression. In those data, UAC-PROA has a reported root mean square error of 0.0068 bar, mean absolute error of 0.0049 bar, and coefficient of determination of 0.965. At the 5-bar reference condition, the combined standard and expanded uncertainties are 0.00735 and 0.01470 bar, respectively. For a 30-instrument fleet, the scheduling scenario specifies 65 annual calibrations under UAC-PROA compared with 120 under a fixed 90-day policy. The integrated analysis illustrates how predictive accuracy, metrological uncertainty, and risk-constrained scheduling can be assessed in one decision framework, while physical prototype validation and independent longitudinal data remain necessary for performance claims.

Index Terms: instrument calibration, machine learning, measurement uncertainty, predictive drift, risk-based maintenance, smart metrology, UAC-PROA.

SECTION 1: INTRODUCTION

1.1. Background of the Study

Calibration is the process of establishing a relationship between the indications of a measuring instrument and values furnished by measurement standards, together with their associated uncertainties. Historically, many industrial organizations have administered calibration through scheduled laboratory checks, paper or electronic certificates, and periodic review of instrument history. The approach is necessary for metrological traceability, but a valid certificate demonstrates performance at the relevant time and conditions, not that an instrument will remain free of drift throughout its subsequent service interval. The International Vocabulary of Metrology explicitly distinguishes calibration, adjustment, verification, and measurement traceability, distinctions essential to any credible intelligent monitoring system (Joint Committee for Guides in Metrology [JCGM], 2012).

In pharmaceutical quality control, measurement errors can affect analytical release decisions, process parameters, and the defensibility of regulated records. In chemical and offshore operations, inaccurate pressure, temperature, electrical, or flow indications may distort process monitoring and maintenance decisions, even when a device appears operational. ISO/IEC 17025:2017 links technically valid laboratory results to appropriate equipment control, metrological traceability, method competence, and defensible reporting. ISO 9001:2015 places monitoring and measurement resources within an organization's quality management framework. These standards do not prescribe one universally optimal interval; instead, calibration policies must reflect risk, evidence, intended application, and measurement requirements (International Organization for Standardization [ISO], 2015; ISO & International Electrotechnical Commission [IEC], 2017).

Industry 4.0 has broadened the role of measuring instruments from isolated measurement endpoints to connected sources of condition information. Low-cost embedded processors, timestamped acquisition, industrial communications, and machine-learning models create opportunities to identify degrading sensor behavior before the next planned verification. Cyber-physical manufacturing architectures provide a basis for linking plant signals, historical records, and control decisions (Lee et al., 2015). The literature on predictive maintenance emphasizes that an accurate model is only operationally useful when its forecast supports a reliable decision under imperfect, changing data (Jardine et al., 2006; Carvalho et al., 2019). Research on Internet of Things architectures and computational modelling also shows how integration, interoperability, and analytic reproducibility become design concerns once monitoring is distributed across devices (Idoko et al., 2024a; Idoko et al., 2024d).

A fixed calendar interval may be conservatively short for stable instruments yet dangerously long for equipment subjected to accelerated wear, temperature cycling, vibration, overload, or changing operating regimes. A condition-aware policy therefore must distinguish routine stochastic variation from persistent drift, account for the uncertainty of the reference measurement, and assess the cost of both premature recalibration and incorrect continued use. The Guide to the Expression of Uncertainty in Measurement and JCGM 106 establish the conceptual bridge from error estimation to uncertainty-sensitive conformity decisions (JCGM, 2008, 2012b). Predictive analytics can contribute only if it is calibrated against traceable observations and produces well-characterized errors and uncertainty intervals, rather than a point forecast alone.

The invention underlying the present work is the Smart Instrument Calibration Monitoring Device for Industrial Quality Control Applications, attributed in the supplied report to Himabindhu Kondeti, Onuh Matthew Ijiga, and Bhavesh Kataria. The report states that the United Kingdom Intellectual Property

Office registered its external appearance as Design No. 6564151 on 29 October 2026, with a certificate dated 10 November 2026. The depicted apparatus consists of a compact rectangular olive-green enclosure, a prominent dark front display, an adjacent control region, cylindrical connection fittings, a vertically mounted circular component, and support features. Its intended settings include pharmaceutical, chemical, and offshore testing. Crucially, the source document treats data acquisition, embedded analytics, drift assessment, uncertainty estimation, notifications, and electronic records as an engineering embodiment proposed for development, not as functions experimentally established by the design registration (Kondeti et al., 2026).

The companion data document covers pressure, temperature, current, and voltage channels, a 180-day drift series, comparative prediction scores, a measurement uncertainty budget, thermal response, and fleet-level calibration scheduling. The pressure analysis spans 0–10 bar at 11 reference points. The reported UAC-PROA and XGBoost RMSE values are 0.0068 and 0.0092 bar, respectively. This analysis investigates the internal numerical relationships and engineering implications of those supplied datasets (Smart Instrument Calibration Monitoring Device, 2026).

1.2. Statement of the Research Problem

Industrial calibration management has a temporal blind spot: the state of an instrument between completed calibrations is often inferred from its last certificate and a scheduled due date. A sensor can acquire a bias or undergo gradual drift during that interval without exceeding any procedural due-date threshold. Consequently, a nominally compliant asset register may not reflect the measurement state required by a particular quality or safety-critical decision (JCGM, 2012; ISO & IEC, 2017).

A second difficulty is that conventional automated pass/fail logic may compare the absolute indication error with a tolerance limit without accounting consistently for uncertainty in the reference, repeated observations, resolution, or ambient conditions. Near a specification boundary, the risks of false acceptance and false rejection can change substantially under different decision rules. Conformity assessment thus needs declared tolerance limits, a documented risk policy, and validated uncertainty contributions (JCGM, 2012b; International Laboratory Accreditation Cooperation [ILAC], 2019).

Third, calibration planning is frequently separated from predictive modelling. A model with a low root mean square error does not automatically justify extending a calibration interval, because temporal extrapolation, rare rapid failures, unequal risk of false acceptance, and the financial consequences of an incorrect decision must also be considered. The supplied scenario contrasts five strategies for 30 instruments, but the figures are not a prospective operational comparison. In particular, the indicated reductions in annual calibration events and costs are scenario outputs rather than confirmed economic benefits (Smart Instrument Calibration Monitoring Device, 2026).

Fourth, comparisons among nonlinear models can be distorted by random splits of serially correlated measurements, inconsistent feature engineering, hyperparameter tuning on held-out records, or testing instruments also represented in training. Generalization must therefore be examined across days, devices, and environmental conditions, with uncertainty estimates and computational performance measured on the intended hardware. Established machine-learning approaches such as random forests, gradient boosting, support vector regression, and recurrent neural networks provide appropriate benchmarks but are not interchangeable with a validated decision policy (Breiman, 2001; Cortes & Vapnik, 1995; Hochreiter & Schmidhuber, 1997; Chen & Guestrin, 2016).

The resulting research problem is the absence, within the supplied design evidence, of a validated end-to-end framework linking traceable multi-channel calibration measurements, predictive drift assessment, quantified uncertainty, explicit conformity decisions, embedded performance, and economically defensible scheduling. A rigorous study is required to establish whether the proposed UAC-PROA method improves both prediction and calibration decisions, and under what operating assumptions these improvements remain valid.

1.3. Aim and Objectives of the Study

Aim. To develop and experimentally evaluate an uncertainty-aware intelligent calibration monitoring framework capable of predicting instrument drift, supporting reliable measurement decisions, and optimizing industrial calibration scheduling.

1. Characterize the architecture and operational requirements of the Smart Instrument Calibration Monitoring Device, separating confirmed physical features from intended electronic functions.
2. Assess calibration accuracy, correction performance, repeatability, and between-device variation across pressure, temperature, current, and voltage channels.
3. Formally specify, implement, and evaluate the proposed uncertainty-aware calibration prediction and risk optimization algorithm (UAC-PROA).
4. Quantify reference, repeatability, resolution, and environmental uncertainty contributions and their effects on conformity decisions.
5. Compare UAC-PROA with XGBoost, LSTM, random forest, and support vector regression under matched experimental protocols.
6. Measure thermal sensitivity, signal acquisition and decision latency, communication integrity, and operational availability of an implemented prototype.
7. Compare fixed, usage-based, condition-based, and predictive scheduling in calibration workload, cost, and out-of-tolerance exposure.

1.4. Research Questions

RQ1. What physical and functional architecture is needed to turn the registered monitoring design into a testable multi-channel prototype?

RQ2. To what extent does calibration correction reduce bias and measurement dispersion across channels and devices?

RQ3. How accurately can UAC-PROA forecast drift on later observation periods and instruments not used during training?

RQ4. How do uncertainty budgets and guard-band policies alter acceptance or rejection decisions near tolerance limits?

RQ5. Does UAC-PROA outperform established prediction baselines in errors, stability, latency, and decision utility?

RQ6. How robust are monitoring and correction outputs to changing ambient temperature and communications constraints?

RQ7. Can adaptive calibration scheduling reduce annual work and cost without increasing out-of-tolerance risk beyond an agreed criterion?

1.5. Research Hypotheses

The null hypotheses are evaluated against independently defined measurement and decision criteria. Statistical significance cannot be inferred from aggregate performance metrics alone.

H01. Mean absolute calibration error after correction is not lower than before correction across independently replicated measurements.

H02. On held-out later periods and devices, UAC-PROA does not achieve lower drift-prediction error than the strongest tuned benchmark.

H03. Incorporating estimated measurement uncertainty does not change decision risk or conformity classification relative to a tolerance-only policy.

H04. The predictive framework does not achieve acceptable detection recall, false-positive rate, or decision latency under predefined deployment constraints.

H05. Risk-constrained adaptive scheduling does not reduce expected calibration workload or cost relative to fixed-interval calibration at comparable out-of-tolerance risk.

1.6. Scientific Novelty and Contributions of the Study

The potential scientific contribution lies in the coupling of traceable measurement assessment with forecast-driven operational decisions, not in the novelty of a screen, connectors, or the general concept of automated calibration. Earlier patents already describe electronic records, automated warnings, field-device monitoring, and instrument verification; the supplied invention's distinct documented subject matter is a particular registered visual configuration (Kondeti et al., 2026).

First, the proposed framework integrates a multichannel calibration correction model, structured drift history, and uncertainty decomposition within a single decision path. Second, UAC-PROA is formulated as a testable prediction-and-optimization method whose contribution must be established through objective functions, ablation studies, and comparisons against established algorithms. Third, conformity assessment is treated as an explicit risk decision near tolerance boundaries rather than a simple numerical threshold. Fourth, a single evaluation design links metrological accuracy, embedded latency, environmental robustness, and calibration-fleet economics. These constitute proposed contributions until independently validated.

Additional relevance comes from adjacent research in computational simulation, Internet of Things integration, autonomous industrial workflows, and predictive modelling. Works involving Idoko and collaborators demonstrate interest in these enabling technologies, although they should not be read as empirical proof that this calibration device operates as claimed (Idoko et al., 2024a, 2024b, 2024d; Ayoola et al., 2024).

1.7. Scope and Limitations of the Study

The anticipated laboratory scope includes pressure (0–10 bar, 11 points), temperature (20–100 °C, nine points), current (4–20 mA, nine points), and voltage (0–10 V, 11 points). The supplied experimental plan specifies ten readings at each point on three days for at least three units per channel. It additionally proposes 13 drift checkpoints over 180 days, at least 30 monitored instruments with a longer follow-up, thermal perturbations between 20 and 50 °C, repeat packet-delivery challenges, and 30 algorithm timing runs. These are planned acquisition conditions, not automatically completed experiments (Smart Instrument Calibration Monitoring Device, 2026).

The conceptual scope includes calibration error, correction, drift modelling, predictive uncertainty, decision rules, and economic trade-offs among five maintenance policies. The principal experimental unit must be explicit, avoiding pseudo-replication when repeated points from one instrument are treated as independent samples. Performance across pharmaceutical, chemical, and offshore use cases is an intended applicability claim and would require site-specific validation for environmental exposures, safety classification, and regulatory documentation.

Known limitations at this drafting stage include the absence of independently provided raw physical measurements, actual reference-standard calibration certificates, prototype electronics/firmware verification, a reproducible implementation of UAC-PROA, and prospective fleet logs. The source document's numerical tables and charts are described as illustrative. Patent and design registration particulars should also be reconciled with the primary UKIPO register before final submission.

1.8. Organization of the Paper

Section 1 introduces the problem, objectives, hypotheses, and proposed scientific contribution. Section 2 critically reviews calibration, uncertainty, machine learning, reliability, digital instrumentation, prior inventions, and the resulting research gaps. Section 3 (to be developed) will specify prototype design, measurement protocols, UAC-PROA mathematics, benchmarks, conformity assessment, and statistical validation. Section 4 will report independently validated measurement, algorithm, robustness, and scheduling outcomes. Section 5 will present conclusions, practical recommendations, limitations, and further research directions.

SECTION 2: LITERATURE REVIEW

2.1. Conceptual Review

2.1.1. Fundamentals of Instrument Calibration, Measurement Accuracy, Precision, and Metrological Traceability

Calibration establishes a documented relation between known reference values and indications, including uncertainty, whereas adjustment actively modifies an instrument and verification establishes whether specified requirements are fulfilled. Accuracy concerns closeness to a true quantity value, precision concerns agreement among replicated indications, and a traceability chain links a result through calibrated references to an accepted measurement reference. These concepts are complementary rather than equivalent (JCGM, 2012; ISO & IEC, 2017).

For a connected calibration monitor, traceability must encompass the reference standards, channel-specific transfer functions, time of comparison, environment, and software transformations applied to readings. The ability to store a correction coefficient is not itself evidence of accuracy unless an independent reference and uncertainty budget substantiate the corrected result.

2.1.2. Smart Instrumentation and Intelligent Calibration Monitoring Systems

Smart instrumentation integrates transduction, local computation, identification, communications, and operator-facing status information. A monitoring device differs from a calibrator: the former assesses instrument condition or history, while the latter ordinarily provides or measures reference quantities under defined conditions. Such distinctions matter because the proposed registered physical enclosure is accompanied by potential functional elements rather than verified proof of an autonomous calibration source (Kondeti et al., 2026).

Cyber-physical production theory explains why measurements should be integrated with machine context, maintenance history, and feedback control (Lee et al., 2015). Comparative work on Internet of Things architectures by Idoko and collaborators supplies adjacent discussion of networking and interoperability but does not validate any specific calibration accuracy (Idoko et al., 2024a).

2.1.3. Instrument Calibration Drift, Measurement Degradation, and Equipment Health Monitoring

Instrument drift denotes change in an instrument indication or metrological characteristic over time under specified conditions. It may be gradual, abrupt, reversible, or environmentally induced. A deterioration trajectory is useful only if it is separated from changes in the measured process and measurement reference. Repeated readings at one nominal value, cross-channel diagnostics, and temperature or load covariates help isolate persistent bias from ordinary noise (Jardine et al., 2006; Si et al., 2011).

The pressure series increases from 0.004 bar at day 0 to 0.130 bar at day 180. Its trend motivates the comparison of linear and nonlinear drift formulations, while extrapolation to other instruments depends on the scope of the available observations (Smart Instrument Calibration Monitoring Device, 2026).

2.1.4. Artificial Intelligence and Machine Learning for Predictive Calibration

Supervised regression relates observed histories and contextual variables to future drift, residual bias, or failure probability. Random forests aggregate decision trees, XGBoost performs regularized gradient boosting, support vector regression applies margin-based functional approximation, and LSTM neural networks capture sequence dependencies (Breiman, 2001; Chen & Guestrin, 2016; Cortes & Vapnik, 1995; Hochreiter & Schmidhuber, 1997). Model choice must be driven by data volume, serial dependence, cross-device variation, and the computing budget.

Research involving Idoko and collaborators on power forecasting, machine-learning risk assessment, computer simulation, and transfer learning shows the breadth of applicable modelling tools (Idoko et al., 2023; Idoko et al., 2024b; Ijiga et al., 2024; Oyebanji et al., 2024). These applications are methodologically adjacent, not direct comparisons of calibration drift prediction.

2.1.5. Measurement Uncertainty, Calibration Correction, and Conformity Assessment

Correction estimates and removes recognized systematic effects, but a corrected reading still has uncertainty arising from the reference, repeatability, resolution, model approximation, and operating conditions. The combined standard uncertainty is typically obtained by propagating component uncertainties through a measurement model, accounting for correlations where relevant. Expanded uncertainty then applies a defensible coverage factor or an equivalent interval procedure (JCGM, 2008; JCGM, 2012b).

Conformity depends on both a measured value and an agreed rule for deciding compliance. Guard bands can reduce false acceptance at the expense of more conservative rejection or review. JCGM 106 and ILAC-G8 provide the framework for making these choices transparent rather than presenting an arbitrary threshold as statistically risk-free (JCGM, 2012b; ILAC, 2019).

2.1.6. Risk-Based Calibration Scheduling and Condition-Based Maintenance

Fixed schedules assign an interval by time, regardless of observed condition. Usage-based scheduling responds to duty cycles; condition-based scheduling reacts to measured deterioration; predictive scheduling forecasts future loss of tolerance and selects an intervention based on risk and resource constraints. Reliability-centred maintenance argues for considering consequences as well as expected failure frequency (Moubray, 1997; Jardine et al., 2006).

An optimization policy must consider calibration labour, lost availability, reference traceability, and the probability-weighted cost of accepting an out-of-tolerance instrument. Generic AI-assisted risk

frameworks and workflow automation research involving Idoko and co-authors offer adjacent insights into prioritization and operational integration, but they do not prove calibration cost savings (Ijiga et al., 2024; Ayoola et al., 2024).

2.1.7. Embedded Sensors, Industrial Connectivity, and Digital Calibration Records

An embedded monitor requires deterministically timestamped acquisition, synchronization between channels, secure instrument identification, persistent records, and recoverable operation under dropped communications. Design targets such as latency and packet delivery require physical tests with specified clocks and network conditions. The companion document proposes acquisition latency at or below 25 ms, decision latency at or below 80 ms, and packet delivery at or above 99.5%; these are acceptance targets, not achieved measurements (Smart Instrument Calibration Monitoring Device, 2026).

IoT architectures and digital records introduce data-integrity, access-control, and cybersecurity risks. Research involving Idoko and collaborators on IoT, privacy, AI and cyber threats is useful for identifying adjacent security controls rather than for asserting metrological compliance (Idoko et al., 2024a; Idoko et al., 2024c; Ayoola et al., 2024b).

2.2. Theoretical Review

2.2.1. Measurement Theory and the Mathematical Foundations of Calibration Error

The observation model may be expressed in terms of a reference quantity, indicated value, correction, and residual error. Measurement theory requires distinguishing random variation from systematic bias, specifying the units and valid range of the measurement model, and using estimates traceable to credible standards. The GUM provides a general uncertainty framework and the VIM supplies the associated vocabulary (JCGM, 2008, 2012).

For the proposed prototype, separate measurement functions are required for pressure, temperature, current, and voltage because each channel has a different physical standard, transfer relationship, sensor dynamics, and uncertainty contributors. A universal corrected-error value across channels would not be scientifically meaningful without dimensional normalization and a defined decision criterion.

2.2.2. Stochastic Degradation Processes and Time-Dependent Instrument Drift Models

A degradation model treats drift as a state evolving over time, potentially with a deterministic trend plus random process fluctuations. Brownian motion with drift, gamma processes, state-space models, and nonlinear regression offer differing assumptions about monotonicity, variance, and shock behavior. Prognostics reviews show that predictions must account for noisy observations, limited failures, and uncertainty over the forecast horizon (Si et al., 2011; Lei et al., 2018).

Monotonic drift is a useful baseline but not a universal law. Environmental reversibility and abrupt faults require separate tests. Repeated measurements from the same instrument are correlated; therefore, multilevel or state-space approaches may be more defensible than treating each reading as a separate instrument.

2.2.3. Statistical Learning Theory and Nonlinear Predictive Modelling

Predictive modelling minimizes a training criterion but seeks low expected error on unseen conditions. Generalization is impaired by distribution shift, leakage from future observations, and model tuning on the evaluation sample. Support vector methods formalize regularization, while tree ensembles accommodate interactions and recurrent models address temporal structure (Cortes & Vapnik, 1995; Breiman, 2001; Hochreiter & Schmidhuber, 1997; Chen & Guestrin, 2016).

For calibration monitoring, a chronological train/validation/test protocol should be supplemented by leave-instrument-out validation. Algorithm comparisons must control preprocessing, eligible features, tuning budget, inference hardware, and forecast horizon. Work by Idoko and collaborators applying machine learning to power, healthcare, and simulation contexts supports discussing transferable modelling workflows without claiming shared performance in metrology (Idoko et al., 2023; Idoko et al., 2024b; Oyebanji et al., 2024).

2.2.4. Bayesian Uncertainty Estimation and Metrological Decision Theory

Bayesian modelling represents uncertainty in unknown drift parameters, latent instrument state, and future observations through probability distributions. Posterior predictive distributions can be translated into a probability that error at a future calibration date exceeds tolerance. Bayesian decision theory combines that probability with consequences of action or inaction (Gelman et al., 2013; JCGM, 2012b).

The distinction between epistemic uncertainty arising from limited knowledge and aleatory variability arising from inherently variable conditions is important in long-horizon forecasts. Decisions should report assumptions, prior sensitivity, interval calibration, and threshold probabilities rather than one deterministic "health score."

2.2.5. Reliability Engineering, Remaining Useful Life, and Condition-Based Maintenance Theory

Reliability expresses the probability that an asset continues to perform a required function for a specified duration under stated conditions. In calibration applications, the failure event may be defined as a measurement error exceeding an application-specific tolerance, rather than a total electrical malfunction. Remaining useful life can therefore be interpreted as time until predicted tolerance violation, conditional on environmental and operational history (Si et al., 2011; Lei et al., 2018).

The framework must distinguish instrument outage from a hidden metrological failure. An instrument may produce uninterrupted measurements yet be unsuitable for regulated quality decisions. Practical policy design thus requires both availability and conformance metrics.

2.2.6. Multi-Objective Optimization Theory for Calibration Cost, Measurement Risk, and Equipment Availability

Scheduling involves competing objectives: reducing calibration events and cost may increase the interval during which undetected bias persists, while more frequent verification may interrupt production and overload laboratory resources. Multi-objective optimization seeks feasible non-dominated policies rather than a single universally superior interval (Deb et al., 2002).

A transparent scheduling objective should define expected cost, false-acceptance risk, instrument availability, operational constraints, and stakeholder risk tolerances. The conceptual UAC-PROA method must be compared against fixed, usage-based, and condition-based policies under the same realized drift histories. Renewable-energy optimization and automation research involving Idoko and collaborators provides examples of optimization and resource coordination in adjacent engineering fields (Idoko et al., 2023; Idoko et al., 2024e).

2.3. Empirical Review

2.3.1. Experimental Investigations of Automated Calibration and Instrument Monitoring Technologies

Applied calibration monitoring research typically combines precision reference sources, instrument interfaces, digital acquisition, and software verification. However, evidence of an electronic status

dashboard does not establish calibration validity; the empirical assessment must check device readings against a traceable reference across the working range and environmental envelope (ISO & IEC, 2017).

The supplied prototype report shows connection fittings and user controls, but it does not include a completed traceable test of acquisition circuitry or independent verification of the proposed monitoring logic. The appropriate next empirical step is a fully described comparison of raw, corrected, and reference values across multiple hardware units and days (Kondeti et al., 2026).

2.3.2. Empirical Evidence on Environmental Effects and Long-Term Instrument Drift

Temperature effects can change sensor offset, gain, response time, and noise. Thermal effects should be assessed using controlled exposures, stabilization times, measurement-order randomization, and repeated heating and cooling cycles to check hysteresis. Prognostics literature warns that degradation extrapolation may fail when environmental regimes change (Jardine et al., 2006; Lei et al., 2018).

The numerical temperature series covers seven conditions from 20 to 50 °C at a nominal 5-bar reference and provides raw error, corrected error, and expanded uncertainty at each point. A robustness assessment should also report thermal stability, repeated cycling, and reference-standard contributions (Smart Instrument Calibration Monitoring Device, 2026).

2.3.3. Comparative Studies of XGBoost, LSTM, Random Forest, and Support Vector Regression in Sensor Degradation Prediction

XGBoost frequently performs well on structured engineered features, while LSTM may exploit complex time dependencies when sufficient sequence data are available. Random forests provide a strong nonlinear tabular baseline; SVR can perform competitively in smaller datasets with careful kernel and scaling choices. The relative ordering is empirical, not inherent to the algorithm class (Breiman, 2001; Chen & Guestrin, 2016; Hochreiter & Schmidhuber, 1997).

The companion document presents UAC-PROA, XGBoost, LSTM, random forest, and SVR with reported RMSE values of 0.0068, 0.0092, 0.0107, 0.0125, and 0.0153 bar, respectively. Those numbers suggest a testable hypothesis but cannot support statements about statistical superiority until dataset provenance, model code, test partitioning, and independent replication are supplied. Studies involving Idoko and collaborators on forecasting and transfer learning show related algorithm-selection concerns, not a direct calibration benchmark (Idoko et al., 2023; Oyebanji et al., 2024).

2.3.4. Experimental Evaluations of Measurement Uncertainty and Risk-Based Calibration Decisions

Empirical uncertainty budgets should identify the reference standard's certificate value, repeated-measurement statistics, display resolution, ambient effects, correlations, and appropriate probability models. Decision experiments should place samples both comfortably within and close to tolerance limits, since near-boundary cases reveal the practical difference between guard-banding and simple thresholding (JCGM, 2008, 2012b; ILAC, 2019).

At 5 bar, the uncertainty components listed for reference, repeatability, resolution, and ambient influence are 0.0030, 0.0040, 0.0020, and 0.0050 bar, respectively. Assuming independent components, root-sum-of-squares gives a combined standard uncertainty of approximately 0.00735 bar and an expanded uncertainty of 0.01470 bar at $k = 2$.

2.3.5. Performance Studies of Embedded Industrial Monitoring and Communication Systems

Connected monitoring evaluation must specify acquisition rate, buffer capacity, message format, packet size, network load, clock synchronization, processor platform, and failure-recovery logic. Mean latency alone can conceal unacceptable tail behavior, so median, high-percentile latency, lost records, and stress-

condition performance are all relevant. Safety and quality applications further require audit trails, authenticated device identity, and tested recovery from power loss (Lee et al., 2015; ISO & IEC, 2017). The design targets in the supplied plan include 99.5% packet delivery, 99.0% uptime, at least 95% drift-detection recall, and no more than 5% false positives. These targets need prespecified test conditions and real event labels. Research on IoT networks, adversarial AI, and privacy involving Idoko and collaborators helps identify the broader need for trustworthy connected architectures (Idoko et al., 2024a; Ijiga et al., 2024; Idoko et al., 2024c).

2.4. Review of Related Literature and Existing Technologies

2.4.1. Traditional Calibration Control and Electronic Calibration History Management

Classical calibration control associates instrument identity with a reference standard, a calibration event, an interval, a result, and a certificate. The patent review in the supplied invention report identifies US6823277B1 and US7593825B2 as precedents for digital calibration-status tracking and electronic records. These precedents indicate that alerts and history management are established functions and cannot alone establish technical novelty in the proposed system (Kondeti et al., 2026).

2.4.2. Portable and Automated Instrument Calibration Systems

Portable calibration systems can configure reference quantities, automate test sequences, record results, and interface with host software. The supplied prior-art assessment includes US4947355A, US20160247387A1, US20160169716A1, and US20200319012A1. Their emphasis includes modular calibration resources, onsite verification, workflow data exchange, and integrated calibration hardware. The proposed monitor must consequently distinguish monitoring and forecasting from calibration-source functionality (Kondeti et al., 2026).

2.4.3. Industrial Equipment Condition Monitoring and Digital Asset Management

Industrial asset systems monitor performance indicators, condition history, and maintenance requirements across connected devices. The report identifies US9625892B2 and US20170060717A1 as relevant monitoring and validation prior art. A credible research advance would not simply add an equipment dashboard, but would demonstrate quantified measurement uncertainty and defensible forecast-driven decisions on held-out industrial devices (Kondeti et al., 2026; Jardine et al., 2006).

2.4.4. Existing Patented Calibration Monitoring and Instrument Verification Technologies

The source report surveys 13 representative patent publications covering calibration reminders, asset status, RFID-based usage identification, portable calibrators, equipment alerts, and online instrument verification. Patent comparisons should be limited to claimed or disclosed subject matter and must not be treated as a substitute for a comprehensive patentability analysis. The document states that Design No. 6564151 protects the registered visual configuration; a registered design is not proof of software performance, calibration accuracy, or novel analytical mathematics (Kondeti et al., 2026).

2.4.5. Comparative Analysis of Existing Technologies and the Proposed Smart Instrument Calibration Monitoring Architecture

The source design combines a prominent display, controls, connector fittings, enclosure, and operator-accessible layout. Conventional calibration systems already provide individual functions such as electronic records, calibration warnings, or condition monitoring. The proposed research contribution is therefore best expressed as an experimentally assessed integration of multichannel correction, prediction, uncertainty-informed conformity, and scheduling under operational constraints.

If the implementation supports only electronic status display and recordkeeping, its contribution would be narrower than a predictive metrology system. If the full integration is built, transparent software and instrument-level testing will be required to substantiate the broader claim.

2.5. Identified Gaps in the Literature

2.5.1. Insufficient Integration of Calibration Correction, Drift Prediction, and Measurement Uncertainty

Published theory supplies mature methods for uncertainty calculation, prediction, and reliability decisions, but these are often assessed as separate tasks. A research gap for the present device is the need to show that combining these methods produces coherent, traceable decisions and measurable benefits over simple calibrated-threshold policies (JCGM, 2008, 2012b; Si et al., 2011).

2.5.2. Limited Validation of Predictive Calibration Algorithms Across Different Instrument Channels

A forecasting model fitted to one pressure channel may not transfer to current, voltage, or temperature because of differences in units, sensor physics, drift mechanisms, and noise. A study that reports only a single pooled prediction score can conceal poor performance in a particular channel. Multi-device, multi-day evaluation with separate channel reporting is necessary.

2.5.3. Inadequate Treatment of Environmental Variability and Operational Degradation

Many simple drift forecasts presume stationary error distributions. Industrial instruments experience temperature changes, load cycling, humidity, vibration, and discrete maintenance events. Model robustness should therefore be assessed under controlled environmental shifts, temporal extrapolation, and recalibration interventions (Jardine et al., 2006; Lei et al., 2018).

2.5.4. Limited Application of Risk-Based Conformity Decisions to Dynamic Calibration Scheduling

Prediction scores alone do not specify when to recalibrate. Translating a forecast into a maintenance decision requires tolerance limits, uncertainty intervals, consequences of wrong acceptance, operational constraints, and agreed risk thresholds. The proposed study can test whether such integration yields better decisions than fixed scheduling (JCGM, 2012b; ILAC, 2019).

2.5.5. Insufficient Joint Assessment of Predictive Accuracy, Embedded Computational Performance, and Maintenance Economics

A highly accurate computational model may be unsuitable for a low-power embedded platform or may be uneconomical to operate. Scientific assessment should therefore combine prediction error, inference latency, communication integrity, data availability, and prospective calibration cost under the same experimental conditions. No scenario can substitute for a prospective fleet study.

2.6. Proposed Conceptual Framework and Research Contributions

2.6.1. Development of an Integrated Intelligent Calibration Monitoring Conceptual Framework

The proposed framework consists of six connected stages: traceable sensor/reference acquisition; data quality and environmental context; calibration correction and drift-state estimation; uncertainty-aware predictive modelling; tolerance and guard-band conformity assessment; and a scheduling optimizer with electronic feedback. The registered device provides the physical user-facing apparatus, while the proposed algorithms and interfaces represent the functional research implementation to be built and validated.

The framework is a testable architecture, not an established deployed system. The primary research assumption is that prediction informed by measurement uncertainty will yield more defensible decisions than independent correction, forecasting, or calendar scheduling alone.

2.6.2. Relationships Between Instrument Measurements, Calibration Drift, Environmental Conditions, Prediction Uncertainty, and Maintenance Decisions

Observed indications depend on the true measurand, calibration function, sensor health, environment, and random noise. Corrected residuals and instrument history provide predictors of future drift. Forecast uncertainty combines parameter and observation uncertainty; conformity assessment evaluates the probability of violating an application-specific limit. The calibration action feeds back by changing the instrument state, updating the reference history, and potentially resetting the degradation trajectory.

The research design should measure each link independently: correction against a reference, drift forecasts against later observations, uncertainty coverage against repeated checks, and scheduling outcomes against actual out-of-tolerance events. A performance improvement in one module cannot logically be assumed to transfer to all other modules.

2.6.3. Hypothesized Advantages and Testable Scientific Contributions of the Proposed Framework

Three primary advantages are hypothesized: lower corrected measurement error, better generalization of drift forecasts under uncertainty-aware modelling, and lower expected calibration workload at a controlled nonconformity risk. Secondary advantages include real-time operator visibility and audit-ready electronic histories. These claims require traceable measurements, independent baselines, prospective testing, and confidence intervals.

The study would make its strongest contribution by publishing the device implementation, raw multichannel observations, source code, benchmark configurations, decision policies, and sensitivity analyses, subject to intellectual-property and data-security restrictions.

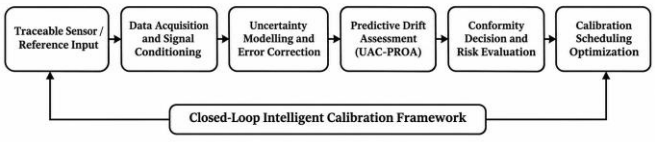


Figure 1. nceptual framework for uncertainty-aware intelligent calibration monitoring and risk-based scheduling.

SECTION 3: MATERIALS AND METHODOLOGY

This section describes the device architecture, experimental measurements, numerical analysis, and uncertainty-aware calibration prediction and risk optimization algorithm (UAC-PROA). The registered-design report documents the external configuration and intended applications; the companion data document supplies the numerical input series and model comparison tables used in the analyses.

3.1. Research Design and Experimental Framework

3.1.1. Experimental Research Design and Quantitative Validation Approach

The investigation adopts a controlled, repeated-measures calibration study supported by algorithm development, uncertainty evaluation, environmental stress testing, and fleet-level economic modelling. Four physical measurement channels are considered: gauge pressure, temperature, direct current, and direct voltage. Reference set points are applied in randomized sequences where instrument constraints

permit, repeated across independent dates and hardware units. The measurement record contains the reference indication, uncorrected output, corrected output, time stamp, device identity, ambient conditions, reference-standard identification, and operator or acquisition process identifier. This design separates within-unit repeatability, between-day variation, between-unit reproducibility, and long-term drift.

The primary inferential unit is the independent measuring instrument or prototype unit, not an individual repeated reading. Algorithm evaluation uses disjoint chronological or instrument-level partitions to prevent leakage from future calibration results into drift predictions. Validation claims are conditional on testing an assembled prototype; the design registration itself is not evidence of measured accuracy or reliability (Kondeti et al., 2026; ISO/IEC, 2017).

3.1.2. Overview of the Proposed Intelligent Calibration Monitoring Framework

A reference stimulus is delivered to an instrument under test (IUT); signal conditioning converts the output into a synchronized engineering quantity; metrological processing estimates bias and correction; a predictive module forecasts future error and its uncertainty; and a decision module combines error tolerance, false-acceptance risk, operating criticality, and calibration costs. Local status is displayed and each event is written to an integrity-controlled record. The UAC-PROA layer proposes the next verification date, subject to prescribed minimum and maximum intervals, risk ceilings, and mandatory overrides for detected faults.

The modular architecture is independent of the particular sensor supplier and communications stack. Traceability connects the calibrated reference, instrument response, derived drift, and decision record. Figure 3.1 shows the front-facing device configuration and Figure 3.2 its exploded assembly representation, both sourced from the invention report.



Figure 3.1. External form and operator-facing arrangement of the registered Smart Instrument Calibration Monitoring Device (reproduced from the supplied invention report).

3.1.3. Definition of Experimental Variables, Performance Indicators, and Validation Criteria

The manipulated variables are reference set point, instrument channel, elapsed operating time, ambient temperature, calibration policy, and injected fault state. Outcomes include signed measurement error, absolute error, corrected error, predictive RMSE and MAE, interval coverage, false-acceptance probability, processing latency, packet delivery, uptime, calibration count, and total annualized cost. Device identity, calibration-standard uncertainty, firmware version, supply voltage, data sampling rate, and stabilization time are recorded as controlled or stratification factors.

Table 3.1. Planned validation matrix transcribed and organized from the companion experimental-design document.

Measurement block	Test domain	Minimum proposed replication	Main endpoint
Pressure	0–10 bar, 11 points	10 readings × 3 days × ≥3 units	Bias, repeatability, correction
Temperature	20–100 °C, 9 points	10 readings × 3 days × ≥3 units	Error, thermal sensitivity
Current	4–20 mA, 9 points	10 readings × 3 days × ≥3 units	Error and resolution
Voltage	0–10 V, 11 points	10 readings × 3 days × ≥3 units	Error and linearity
Drift	0–180 days; 13 occasions	≥30 instruments proposed	Forecast and risk calibration
Communications	≥10,000 packets/condition	3 repeated conditions	Packet delivery and loss
Environmental	20–50 °C; 7 steps	≥3 thermal cycles	Robustness and uncertainty
Algorithms	100–1,000 assessments	30 timed runs per method	Runtime and predictive accuracy

The companion document lists device-level engineering acceptance targets of acquisition latency ≤25 ms/sample, decision latency ≤80 ms/decision, record writing ≤50 ms/record, packet delivery ≥99.5%, instrument identity matching ≥99.0%, uptime ≥99.0%, drift recall ≥95.0%, and false-positive rate ≤5.0%. These are design thresholds for future assessment, not empirical achievements. All threshold choices must be justified against operating specifications and uncertainty before formal acceptance.

3.2. Design and Architecture of the Smart Instrument Calibration Monitoring Device

3.2.1. Description of the Registered Invention and Physical Configuration

The research is organized around the Smart Instrument Calibration Monitoring Device for Industrial Quality Control Applications, UK registered design No. 6564151, registered on 29 October 2026 and granted on 10 November 2026. The uploaded report identifies the registered proprietors as Himabindhu Kondeti, Onuh Matthew Ijiga, and Dr. Bhavesh Kataria. Its physical representations show an olive-green rectangular body, dark display area, adjacent keypad or control region, cylindrical external connection

features, a circular gauge-like component, ventilation and support features. The registered subject matter is the particular appearance and arrangement of these visible features, rather than an experimentally established measurement or forecasting method.

Consequently, all printed or illustrated display values, labels, internal modules, and wiring diagrams are interpreted as design illustrations until independently verified in functioning hardware. The precise capacities of ports and indicators are to be established by electrical continuity tests, interface documentation, and controlled signal injection rather than inferred from visual appearance alone.

3.2.2. Hardware Architecture, Functional Components, and Instrument Connection Interfaces

The proposed functional implementation consists of protected channel connectors, reference input or conditioning circuitry, analog-to-digital conversion, embedded control, non-volatile memory, a real-time clock, a human-machine interface, and optional communications. Pressure input is conditioned through a suitable transducer or transmitter interface; 4–20 mA input requires calibrated burden resistance and overcurrent protection; 0–10 V input uses precision attenuation, buffering, and input protection; and temperature input uses a traceable simulator or calibrated sensing arrangement matched to the test method. Hardware design must specify component models, input impedance, measurement ranges, excitation, overload limits, grounding, isolation, connector pinouts, and electrical safety.

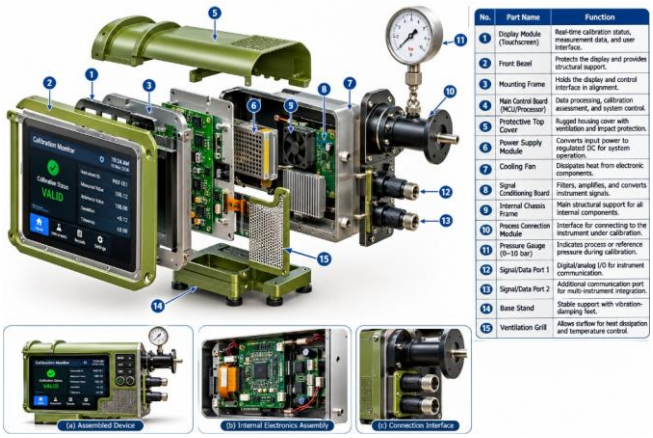


Figure 3.2. Exploded and internal-assembly illustration from the invention report. Depicted subassemblies are design representations pending physical verification.

3.2.3. Proposed Sensor Acquisition, Signal Conditioning, and Embedded Processing Architecture

Each channel shall have a documented signal pathway: stimulus or IUT output → protective front end → analog conditioning → digitization → unit conversion → calibration-correction function → uncertainty estimation. Sampling rates and anti-aliasing requirements are selected relative to the dynamics of the IUT; the reference and output streams must share a clock or use recorded synchronization offsets. Gain, offset, integral nonlinearity, temperature coefficient, and ADC quantization are measured independently and incorporated into the uncertainty budget. The processor shall retain a monotonically increasing sequence number and an immutable raw-value record, enabling recalculation after algorithm upgrades.

3.2.4. Operator Display, Calibration Status Indicators, and Electronic Record Management

The proposed operator interface presents instrument identity, reference reading, measured value, calibrated value, combined uncertainty, predicted drift, remaining decision margin, date of last verification, and next due date. Color or text status shall distinguish ACCEPT, VERIFY, and REJECT, with a separate sensor-fault state. A visible warning must not substitute for safe instrument isolation or an

authorized quality decision. Each record stores the policy version, calibration certificate identifier, raw and derived values, decision threshold, status, operator acknowledgement, and cryptographic integrity check when the chosen hardware permits it.

3.2.5. Device Communication Architecture, Data Integrity, and Monitoring Workflow

The communication layer may use authenticated wired or wireless telemetry, but no particular protocol is claimed as implemented in the supplied registration record. Proposed verification includes packet sequencing, end-to-end acknowledgement, timestamp checks, reconnect tests, identity matching, and transfer integrity. Missing samples are flagged rather than interpolated silently; critical decisions default to a fail-safe state when reference integrity or sensor identity cannot be established. Access logging and revision control are necessary to support the technical-record and impartiality requirements of laboratory quality systems (ISO/IEC, 2017).

3.2.6. Prototype Fabrication, Integration, and Operational Verification

A build register should document design drawings, enclosure revisions, printed circuit boards, firmware commits, bill of materials, assembly photographs, wiring continuity, and acceptance inspections. Verification is staged: visual and electrical inspection; safe power-on; isolated channel simulation; reference comparison; full-range calibration; thermal cycling; communications and log integrity tests; and operational fault injection. Hardware deviations from the registered appearance must be reported because the submitted illustrations establish design geometry, not an as-built metrological specification.

3.3. Experimental Instrumentation and Data Acquisition

3.3.1. Selection of Pressure, Temperature, Current, and Voltage Measurement Channels

Four channels represent distinct sensing and instrumentation pathways. Pressure spans 0–10 bar, voltage 0–10 V, current 4–20 mA, and temperature 20–100 °C. Their combined assessment tests whether a common monitoring pipeline generalizes across units and physical error mechanisms. Each channel is analyzed separately before any normalized pooled modelling, because a 0.01-bar error cannot be directly compared with an error of 0.01 mA or 0.01 °C without engineering scaling.

3.3.2. Reference Standards, Calibration Equipment, and Traceability Requirements

Traceable pressure standards, dry-block or liquid-bath temperature systems with suitable reference probes, precision current sources, and calibrated voltage sources or digital multimeters are required. Each reference must be within calibration, with stated uncertainty, coverage and traceability. Standard uncertainty should be sufficiently small relative to the permitted IUT error to make the intended conformity decision credible. A complete calibration certificate, ambient conditions, stabilization period, and reference-standard history shall accompany every dataset (JCGM, 2008; ISO/IEC, 2017).

3.3.3. Eleven-Point Pressure and Voltage Calibration Protocols

For both pressure and voltage, eleven equally spaced reference points cover the full operating range including zero and upper scale. A practical bidirectional sequence begins with increasing stimulus, stabilization, recorded reading, decreasing stimulus, and repeat cycles to assess hysteresis. At each reference point, ten replicate indications are recorded over three nonconsecutive days and a minimum of three independently characterized prototype or monitored units. The data plan implies at least 990 observations per channel if each point is measured once at each day-unit-repetition combination, exclusive of additional up/down passes.

3.3.4. Nine-Point Temperature and Current Calibration Protocols

Temperature reference points of 20, 30, 40, 50, 60, 70, 80, 90, and 100 °C are applied under verified thermal equilibrium; nine current stimuli from 4 to 20 mA use a 2-mA increment. The same minimum replication scheme produces 810 planned readings per channel before additional cycles. Stabilization shall be defined by a rate-of-change criterion rather than a fixed delay alone, especially at elevated temperatures. The resulting dataset tests linearity, bias, resolution, repeatability, and conditional calibration correction.

3.3.5. Measurement Repetition, Instrument Replication, and Experimental Randomization

Readings are nested within set point, instrument, day, and thermal cycle. Order is randomized within safety and hysteresis restrictions; increasing and decreasing passes are retained as distinguishable factors. A blocked design balances units over days and operators, preventing systematic time-of-day or single-unit effects from being mistaken for calibration improvements. Independent instrument identities remain intact through all data splits. A power analysis should determine whether the planned sample size is adequate for detecting the smallest operationally meaningful drift reduction.

3.3.6. Data Acquisition, Cleaning, Synchronization, and Quality Assurance

A structured dataset contains device_id, reference_id, timestamp_utc, channel, setpoint, raw_reading, corrected_reading, ambient_temperature, day, replicate, calibration_revision, unit, and quality_flag. Units are standardized and raw observations retained. Duplicates are identified by the sequence key, clock offsets quantified, and missing or saturated readings tagged. Outlier rejection requires an independent physical justification; otherwise robust estimates and sensitivity analyses shall be reported alongside unfiltered results. All transformations and exclusion criteria are versioned before model evaluation.

3.4. Mathematical Modelling of Calibration Error and Measurement Reliability

3.4.1. Mathematical Representation of Reference Measurements and Instrument Indications

For instrument i, channel c, set point j and acquisition occasion t, the displayed measurement is modelled as the reference quantity plus systematic bias, time-varying degradation, environmental response, and random error. All terms must be expressed in the same channel-specific unit. This additive approximation is assessed for adequacy before nonlinear alternatives are adopted.

$$y_{i},c,j,t = x_{r},c,j + b_{i},c(x_{r}) + d_{i},c(t) + g_{i},c(T,h) + \varepsilon_{i},c,j,t \quad (1)$$

Here y is the instrument indication; x_r is the reference value; b is systematic calibration bias; d is drift; g is environmental response in temperature T and humidity h where measured; and ε is residual measurement noise.

3.4.2. Calibration Bias, Measurement Error, and Correction Functions

Signed error is the measured indication minus the reference. A fitted correction may be linear, polynomial, or locally monotonic, provided calibration uncertainty and overfitting are controlled. Corrections are learned only from the development set and evaluated on separately acquired runs; otherwise estimates of improvement are optimistically biased.

$$e_{i},j,t = y_{i},j,t - x_{r},j \quad (2)$$

e is signed measurement error; y is observed indication; x_r is the reference reading.

$$\hat{y}^{corr} = y - \hat{c}(y,T,t);\quad e^{corr} = \hat{y}^{corr} - x_{r} \quad (3)$$

ĉ is the estimated correction function, evaluated without refitting on held-out measurements; $\hat{y}^{corr}$ and e^{corr} are corrected indication and residual error.

3.4.3. Repeatability, Reproducibility, and Statistical Error Propagation

Repeatability is estimated from readings under the same controlled conditions, while reproducibility considers systematic between-day and between-unit components using a random-effects formulation when the design permits. Correlated input quantities must not be treated as statistically independent. The first-order propagation law follows the Guide to the Expression of Uncertainty in Measurement (JCGM, 2008).

$$s_r = \sqrt{[\Sigma_k (y_k - \bar{y})^2 / (n - 1)]} \quad (4)$$

s_r is within-condition repeatability standard deviation; n is the number of replicate readings.

$$u^2(y) = \Sigma_a c_a{}^2 u^2(x_a) + 2\Sigma_a < _\beta c_a c_\beta u(x_a, x_\beta) \quad (5)$$

c_a denotes sensitivity coefficient; $u(x_a)$ is standard uncertainty of input a; and $u(x_a,x_\beta)$ is covariance. The second sum is retained when input sources are correlated.

3.4.4. Time-Dependent Instrument Drift Modelling

Longitudinal drift is measured as the change in signed calibration error relative to a documented baseline at the same reference stimulus. Thirteen planned monitoring occasions cover days 0 through 180 at 15-day intervals. To separate true drift from test artifacts, the reference source and ambient influences are recorded at each occasion. Linear, nonlinear, and state-space models are evaluated, rather than presupposing an increasing drift trajectory.

$$d_i(t) = e_i(t) - e_i(0) \quad (6)$$

$d_i(t)$ is baseline-adjusted instrument drift, which may be positive or negative.

$$d_i(t + \Delta t) = d_i(t) + \mu_i \Delta t + \sigma_i \sqrt{\Delta t}\, \xi_t \quad (7)$$

μ_i is mean drift rate; σ_i is stochastic drift intensity; ξ_t is a standardized process innovation. This is a candidate benchmark, not a confirmed physical law.

3.4.5. Environmental Correction and Temperature-Dependent Error Modelling

Ambient temperature is perturbed over 20–50 °C at seven 5-°C steps while the pressure reference remains at 5 bar. Corrected versus uncorrected errors and expanded uncertainty are assessed across at least three thermal cycles. A calibrated environmental term can include an interaction with signal magnitude; data outside tested temperature ranges are not extrapolated without additional validation.

$$g(T, x) = \beta_1 (T - T_0) + \beta_2 (T - T_0)^2 + \beta_3 x (T - T_0) \quad (8)$$

T_0 is the declared reference temperature and β_1–β_3 are fitted coefficients; x denotes the nominal input.

3.4.6. Development of an Integrated Measurement Reliability Indicator

A channel-specific reliability indicator is proposed to combine normalized absolute error, uncertainty relative to tolerance, and probability of an out-of-tolerance state. Its weights are prespecified or tuned only on development data. The indicator supports ranking and communication; it is not substituted for an accredited conformity-assessment decision.

$$R_i = exp\{-[w_1 |e_i| / E_{max} + w_2 U_i / E_{max} + w_3 p_i]\};\ \Sigma w_k = 1 \quad (9)$$

R_i is dimensionless reliability score in (0,1]; E_{max} is maximum permitted error; U_i is expanded uncertainty; p_i is predicted nonconformity probability; and w_k are nonnegative weights.

3.5. Development of the Uncertainty-Aware Calibration Predictive and Risk Optimization Algorithm (UAC-PROA)

3.5.1. Algorithm Design Objectives and Mathematical Assumptions

UAC-PROA is specified here as a proposed two-stage research algorithm, not as an already validated feature of the registered industrial design. The first stage learns the conditional distribution of future instrument error from historical calibration and operating information. The second stage chooses the next calibration interval that minimizes expected maintenance cost and nonconformity loss while controlling decision risk. The model assumes stable instrument identity, available reference anchors, identifiable environmental covariates, and a stated tolerance; departures from these assumptions initiate recalibration or a conservative fallback policy.

3.5.2. Feature Selection and Construction of Time-Dependent Instrument Health Indicators

The feature vector includes latest signed error, rolling error slope, absolute and normalized bias, within-window variance, elapsed days since last calibration, operating time, temperature summaries, environmental excursion count, channel type, sensor identity, and prior conformity outcomes. Feature values are computed using historical information available at prediction time only. Device-specific identifiers are excluded from cross-instrument generalization unless explicitly modelled as random effects. Feature importance is assessed by permutation on held-out data rather than by training-set splits alone.

$$z_i, t = [e_i, t, \Delta e_i, t /\Delta t, s_i, t, \tau_i, t, T_i, t, N_{ex}c, i, t, q_i] \quad (10)$$

z is the predictor vector; τ is time since calibration; $N_{ex}c$ is the count of thermal excursions; q contains channel and operating class descriptors.

3.5.3. Nonlinear Calibration Drift Prediction Model

The predictive component estimates drift at horizon h using a nonlinear function fθ and supplies both a point estimate and predictive distribution. Candidate implementations can use gradient-boosted trees for small tabular datasets, recurrent networks only if sufficient sequential samples exist, or a hybrid monotonic degradation model with learned residuals. The definitive architecture must be frozen before testing. Physics-informed constraints may penalize physically implausible predictions, but a monotonic drift assumption is prohibited when data exhibit reversible thermal effects.

$$\hat{d}_i, t + h = f\theta(z_i, \leq t, h) + r_i, t + h \quad (11)$$

fθ is the trained predictive model; r denotes an optional instrument-specific random effect. Neither future measurements nor future ambient observations enter z.

3.5.4. Measurement Uncertainty Integration and Predictive Confidence Intervals

Metrological uncertainty and statistical predictive uncertainty are related but different: the former describes uncertainty in the measurement result, while the latter describes future prediction error. They are combined only where compatible and with covariance treatment. Prediction intervals can be supplied using chronological calibration residuals or conformal methods adapted to temporal dependence; empirical coverage must be checked by instrument, horizon, and environmental stratum.

$$u^2_{\text{pre}}d = u^2_{\text{meas}} + u^2_{\text{mo}}d + 2cov(\delta_{\text{meas}}, \delta_{\text{mo}}d) \quad (12)$$

$u_{pre}d$ is joint predictive standard uncertainty when the component uncertainties refer to commensurable quantities; covariance is included when relevant.

$$PI_{1-}\alpha(t+h) = [\hat{d}(t+h) - q_{1-}\alpha, \hat{d}(t+h) + q_{1-}\alpha] \quad (13)$$

PI is a calibrated prediction interval; $q_{1-}\alpha$ is an out-of-sample residual quantile for nominal coverage $1 - \alpha$, estimated with an appropriate time-aware calibration scheme.

3.5.5. Risk-Based Optimization Objective Function and Decision Constraints

For each candidate next-verification time Δ, the optimizer evaluates direct intervention cost, probability-weighted nonconformity loss, expected downtime, and confidence degradation. Loss weights are elicited from process-criticality tiers and audited in sensitivity analyses. Risk is measured against a protected maximum permitted error; predictions with wide uncertainty lead to shorter intervals rather than unjustified confidence.

$$J_i(\Delta) = Cc_{al}/\Delta + \lambda_1 P(|E_i(t+\Delta)| > E_{\max}) + \lambda_2 D_i(\Delta) + \lambda_3 U_i(\Delta) \quad (14)$$

J is the scheduling objective; Cc_{al} is per-calibration cost; P is predicted nonconformity probability; D is expected downtime; U is a normalized uncertainty penalty; λ parameters represent declared trade-off weights.

$$\Delta_{\min} \le \Delta \le \Delta_{\max};\ P(|E_i(t+\Delta)| > E_{\max}) \le \rho_{\max} \quad (15)$$

Δ_{min} and Δ_{max} are permitted scheduling bounds; ρ_{max} is the criticality-dependent maximum tolerated risk.

3.5.6. Adaptive Calibration Interval Determination

At every accepted observation, feasible candidate intervals are evaluated on a discrete grid. The earliest candidate meeting prescribed safety constraints and the minimum objective is selected; equal-cost solutions favor the shorter interval. An instrument enters VERIFY state if its predicted error approaches the guard band or interval coverage fails, and REJECT if a verified measured error exceeds its applicable decision threshold. A hard maximum interval and forced calibration after sensor replacement prevent uncontrolled postponement.

$$\Delta_i* = argmin\Delta \in \mathbb{F}_i J_i(\Delta) \quad (16)$$

$\mathbb{F}_i$ is the set of feasible intervals satisfying the risk, maintenance, and regulatory constraints.

3.5.7. Algorithm Training, Parameter Estimation, and Optimization Procedure

Preprocessing is fitted using training data only; hyperparameters are selected by rolling-origin validation and fixed before test evaluation. Training losses are weighted for instrument criticality only if the weighting scheme is published in advance. Calibration curves for predicted risk are checked using reliability diagrams and a proper scoring rule. Uncertainty-weighted objective coefficients are selected on a development cohort and stress-tested using independent costs and risk limits. A hierarchical model may be used when instrument counts support pooling across units without collapsing distinct measurement types.

$$L(\theta) = (1/N)\Sigma_i \ell(d_i, f\theta(z_i)) + \lambda_r \Omega(\theta) \quad (17)$$

L is the training objective; ℓ is a predictive loss (for example squared error); Ω is regularization; λ_r is its coefficient. Risk calibration is assessed separately.

3.5.8. Algorithm Pseudocode, Convergence Criteria, and Computational Complexity

Algorithm 1 (proposed UAC-PROA): (1) validate instrument identity, reference traceability and time stamps; (2) ingest raw channel records; (3) estimate corrected indications and uncertainty; (4) construct past-only feature windows; (5) generate calibrated future-error distributions for allowable horizons; (6) compute conformity probabilities; (7) identify intervals satisfying risk and operational constraints; (8) minimize scheduling objective over feasible intervals; (9) display ACCEPT, VERIFY, or REJECT and commit the signed decision record; and (10) re-evaluate when new verified measurements become available. If any critical input fails validation, issue VERIFY and use a shorter predefined fallback interval. Convergence of model training is defined by early stopping on chronological validation loss with a fixed patience and reproducible seed; optimization over the finite scheduling grid terminates after all candidates have been evaluated. For N observations, F features and H candidate horizons, end-to-end inference scales approximately as model prediction cost plus O(H) objective evaluations per instrument. Training complexity depends on the selected learner and cannot be inferred from the evaluation checkpoint curves in the companion file, which do not establish equal computational budgets.

3.6. Benchmark Machine Learning Models

3.6.1. Extreme Gradient Boosting (XGBoost)

XGBoost constitutes a strong tabular-data comparator. Boosted regression trees approximate nonlinear relationships among time since calibration, operating stress and historical error using regularized sequential fitting. Maximum depth, number of trees, learning rate, column sampling and minimum leaf size are selected using rolling validation. All input transformations, monitoring horizons, and target definitions match UAC-PROA (Chen & Guestrin, 2016).

3.6.2. Long Short-Term Memory Neural Network (LSTM)

LSTM is included for sequential sensor histories, with fixed-length windows containing only observations preceding the forecast origin. Hidden units, dropout, sequence length and training epochs are selected within the training fold; padded and missing sequences are masked. Given only 13 planned drift occasions over 180 days, an LSTM may be data-limited and must be assessed for variance across random initializations rather than presumed superior because it is a deep-learning model (Hochreiter & Schmidhuber, 1997).

3.6.3. Random Forest Regression

Random Forest averages predictions from bootstrapped regression trees, providing a nonlinear baseline with limited tuning. Number of trees, leaf size and sampled feature fraction are optimized using the same chronological validation protocol. Bootstrapping is performed on independent instruments or appropriately blocked observations where temporal dependence would otherwise invalidate random resampling (Breiman, 2001).

3.6.4. Support Vector Regression (SVR)

SVR is implemented using standardized predictors and a radial-basis-function kernel unless the development data support a different specification. Kernel scale, regularization penalty, and epsilon-insensitive loss are selected within each training fold. Scaling parameters are fitted on training observations only, and test instruments are transformed without adapting the scaler (Cortes & Vapnik, 1995).

3.6.5. Hyperparameter Optimization and Equivalent Benchmarking Conditions

All methods use identical target definitions, data partitions, forecast horizons, feature-availability rules, and performance units. Search budgets should be matched by capped evaluation counts or compute time, with actual core-hours, hardware, number of trials, and termination reasons reported. Each stochastic fit is repeated with at least 30 independently controlled seeds when computationally feasible. Results from the provided scenario comparison (UAC-PROA 0.0068 bar RMSE, XGBoost 0.0092, LSTM 0.0107, Random Forest 0.0125, SVR 0.0153) are not used to select the final model or justify claims until independently reproduced.

3.6.6. Time-Aware Training, Validation, and Independent Testing

A rolling-origin design trains on earlier drift records, validates on a subsequent block, and evaluates a final untouched future block. A second leave-one-instrument-out or held-out-instrument study examines generalization. Windows that overlap an evaluation horizon are purged from training when necessary. Predictions at day 180 cannot use a correction or normalization parameter estimated from day-180 results. Confidence intervals are clustered at instrument level to avoid pseudoreplication.

3.7. Measurement Uncertainty Quantification and Conformity Assessment

3.7.1. Identification and Classification of Measurement Uncertainty Sources

The primary pressure uncertainty budget includes reference-standard uncertainty, reading repeatability, quantization or resolution, and ambient influence. Additional components must be assessed when relevant: hysteresis, temperature gradients, calibration-fit uncertainty, loading effects, sensor response delay, reference drift, and communication-induced timing mismatches. Type A components follow statistical evaluation; Type B components rely on certificates, specifications, or defensible probability distributions (JCGM, 2008).

3.7.2. Combined Standard Uncertainty Estimation

For the 5-bar pressure uncertainty budget, the standard contributions are 0.0030, 0.0040, 0.0020, and 0.0050 bar, each with unit sensitivity coefficient. Under the stated zero-covariance assumption, the combined standard uncertainty is 0.00735 bar.

$$uc = \sqrt{(u_{re}f^2 + u_{re}p^2 + u_{re}s^2 + u_{env}{}^2)} \quad (18)$$

uc denotes combined standard uncertainty. Where inputs are correlated, the full covariance expression in Equation (5) is required.

3.7.3. Expanded Measurement Uncertainty and Coverage Factors

Expanded uncertainty is calculated from combined standard uncertainty and the coverage factor applicable to the measurement model. With k = 2, the 5-bar budget gives 0.01470 bar. The resulting coverage probability depends on the distributions and effective degrees of freedom.

$$U = k\,uc \quad (19)$$

U is expanded uncertainty; k is the declared coverage factor; uc is the combined standard uncertainty.

3.7.4. Guard-Band Decision Rules and False-Acceptance Risk

Conformity assessment combines stated tolerance limits with a documented guard-band decision rule. The numerical decision table uses an acceptance threshold of |error| ≤ 0.085 bar and assigns a verification-review outcome at 0.090 bar. The applicable tolerance, guard-band width, and three-state policy must be explicitly defined for reproducibility (JCGM, 2012).

$$A = T_{ol} - w; \quad accept\ if\ |e| \leq A \tag{20}$$

A is the guard-banded acceptance limit; T_{ol} is the maximum permitted absolute error; w is the guard-band width. Additional REVIEW and REJECT thresholds must be stated explicitly.

$$PFA = P(|E| > T_{ol} \mid measured\ acceptance) \tag{21}$$

PFA is the false-acceptance probability under a documented uncertainty model and prior or conditioning assumptions.

3.7.5. Risk-Based Instrument Acceptance, Verification, and Recalibration Criteria

The decision tree first verifies reference traceability and data integrity, then tests whether the corrected measured error and uncertainty support acceptance. Instruments outside the protected acceptance region are flagged for verification or recalibration according to written rules. Predictive alerts cannot automatically transform an accepted instrument into an accredited calibration result. All actions are recorded together with the uncertainty budget and rule version so that the decision can be resupplied during an audit.

3.8. Embedded-System Performance and Environmental Robustness Testing

3.8.1. Data Acquisition and Processing Latency Evaluation

Independent monotonic clocks or hardware-triggered timestamps are used to measure source activation to acquisition, acquisition to decision, and decision to record persistence. At least 30 test runs per candidate method are planned across 100–1,000 assessment cycles, with warm-up and cold-start conditions distinguished. Latency distributions are reported using median, 95th percentile, worst observed time, and 95% uncertainty intervals, not just averages. The companion targets of 25 ms/sample acquisition and 80 ms/decision are treated as pass/fail specifications.

$$L_p = tdecision - tacquisition; \quad L_r = tpersisted - tdecision \tag{22}$$

L_p is processing-decision latency; L_r is record-write latency measured with a specified timestamping convention.

3.8.2. Communication Packet Delivery and Data Integrity Testing

At least 10,000 packets are transmitted per test condition in three repeated trials. Network load, attenuation or cable length, power disturbances, and reconnect intervals are logged. Packet delivery measures unique valid received frames divided by frames sent; retransmitted duplicates are not counted as additional success. A cryptographic hash or message authentication code, where implemented, tests payload integrity rather than simply link availability.

$$PDR = 100 \times N_{valid} / N_{sent} \tag{23}$$

PDR is packet delivery ratio in percent; N_{valid} is unique correctly received frames; N_{sent} is originated frames.

3.8.3. Temperature Perturbation and Thermal Stability Assessment

The thermal test domain is a nominally constant 5-bar reference with ambient temperatures of 20–50 °C in seven steps. Raw error, corrected error, and expanded uncertainty are compared across the steps.

Reference-standard conditions, stabilization, and heating/cooling direction are relevant covariates for repeated-cycle assessment.

3.8.4. Drift Detection Sensitivity, Recall, and False-Positive Assessment

Drift labels shall be assigned from independent reference comparisons and prespecified thresholds, preferably using controlled fault injection at known onset times. Detection metrics are computed at instrument-event level; adjacent correlated samples are not treated as independent faults. Recall and false-positive rate are reported with confidence intervals, warning lead time and prevalence. The proposed thresholds of ≥95% recall and ≤5% false-positive rate are verified only when sufficient labeled events are available.

$$Recall = TP/(TP + FN); \quad FPR = FP/(FP + TN) \qquad (24)$$

TP, FN, FP and TN are event-level true positives, false negatives, false positives and true negatives, respectively.

3.8.5. Monitoring Availability, Reliability, and Fault Recovery

Availability is computed from time-stamped operation logs excluding only prespecified planned interruptions. Controlled power loss, sensor disconnection, clock corruption, storage exhaustion and communication dropout are used to measure detection time, safe-state behavior, and recovery without data duplication. Restoration procedures must not overwrite immutable pre-fault records. Hardware survival and firmware integrity are assessed separately from model predictive reliability.

$$Availability = 100 \times T_{\mathrm{op}} / (T_{\mathrm{op}} + T_{\mathrm{unplanned}}) \qquad (25)$$

T_{op} is valid operating time; $T_{unplanned}$ is unplanned unavailable time under the declared measurement window.

3.9. Calibration Scheduling Optimization and Economic Evaluation

3.9.1. Fixed 90-Day and 180-Day Calibration Scheduling Strategies

The scheduling comparison evaluates 90-day and 180-day fixed calibration policies against usage-based, condition-based, and UAC-PROA policies for the same 30-instrument fleet. The supplied annual calendar policy totals are 120 and 60 calibrations, respectively. Policy comparisons use a common observation horizon and the associated nonconformity-event totals.

3.9.2. Usage-Based and Condition-Based Calibration Policies

Usage-based schedules trigger service after accumulated operating hours or cycles, while condition-based schedules react to observed drift or reliability indicators. Both require an explicit scheduling rule and fair access to historical data. Policies are evaluated on a common set of instrument-level trajectories with identical failure thresholds, intervention costs and follow-up horizons. The simulation must account for the fact that recalibration resets or modifies the future error state rather than simply removing a data row.

3.9.3. UAC-PROA-Based Adaptive Scheduling Framework

At each scheduled evaluation, UAC-PROA forecasts nonconformity probability over candidate future horizons and selects a constrained risk-adjusted interval. A digital replay simulator generates the next inspection event, checks the inferred device state, records compliance and cost, and updates the instrument trajectory following a calibration. For transparent comparison, the UAC-PROA policy is applied to held-out fleet histories using frozen parameters and a documented counterfactual intervention model.

3.9.4. Calibration Cost, Maintenance Frequency, and Out-of-Tolerance Risk Modelling

Annual cost includes direct calibration labor or outsourcing, downtime, instrument transport, verification after alerts, and expected cost of out-of-tolerance measurements. The companion table uses a nominal $100 per calibration event (e.g., 120 events costing $12,000); the other cost categories are not documented and should not be silently assumed zero. A risk-adjusted economic comparison reports both observed and simulated incident costs separately.

$$C_{\text{ann}} = N_{\text{Cal}} C_{\text{Cal}} + N_{\text{v}} C_{\text{v}} + TdownCdown + N_{\text{oot}} C_{\text{oot}} \quad (26)$$

N_{Cal} is calibration count; N_v is additional verification count; Tdown is downtime; N_{oot} is out-of-tolerance event count; C terms are unit costs and must be sourced or scenario-labelled.

3.9.5. Multi-Objective Trade-Off and Sensitivity Analysis

A Pareto analysis compares candidate policies across annual monetary cost, calibration frequency, nonconformity event rate, false-acceptance risk and availability. Cost weights, failure penalties, drift prevalence, calibration effectiveness, instrument criticality and environmental severity are varied one at a time and jointly. A policy is considered operationally preferable only if its risk difference is within a predeclared acceptance bound and its cost advantage remains robust under plausible assumptions.

$$minimize\ [C_{\text{ann}}, N_{\text{oot}}, 1 - A_{\text{vail}}]\ subject\ to\ PFA \leq \rho_{\max} \quad (27)$$

The vector objective compares cost, out-of-tolerance incidence and lost availability; PFA is constrained false acceptance, with trade-offs visualized on the Pareto frontier.

3.10. Statistical Analysis, Model Validation, and Reproducibility

3.10.1. Root Mean Square Error, Mean Absolute Error, and Coefficient of Determination

Prediction quality is evaluated for each physical channel and forecast horizon using the same held-out reference errors. RMSE emphasizes large deviations; MAE describes typical magnitude; R^2 is calculated against the actual held-out target mean and may be negative. Values across physical channels are not averaged in raw units; normalized errors or per-channel summaries are provided where aggregation is required.

$$RMSE = \sqrt{[(1/n)\Sigma_i(y_i - \hat{y}_i)^2]} \quad (28)$$

RMSE is the root mean square prediction error for n independent evaluation observations or appropriately clustered summaries.

$$MAE = (1/n)\Sigma_i|y_i - \hat{y}_i| \quad (29)$$

MAE is mean absolute error, with units identical to the forecast quantity.

$$R^2 = 1 - [\Sigma_i(y_i - \hat{y}_i)^2 / \Sigma_i(y_i - \bar{y})^2] \quad (30)$$

R^2 measures improvement over a constant held-out mean baseline; $\bar{y}$ is the mean observed target in the assessment set.

3.10.2. Confidence Intervals, Statistical Significance, and Effect Sizes

Uncertainty intervals for comparative predictive error are calculated using instrument-block bootstrap or valid clustered methods. Paired differences in error are assessed on matching instruments and forecast horizons, with multiplicity adjustments for the four benchmark comparisons. Report effect size and 95% confidence interval before p-values, recognizing that multiple repeated readings on one device do not provide independent evidence. Interval calibration is evaluated using empirical coverage and mean interval width.

$$\Delta RMSE = RMSEbenchmark - RMSEUAC - PROA \quad \textbf{(3 1)}$$

Positive ΔRMSE indicates a numerical advantage of UAC-PROA on the same held-out dataset; statistical and practical significance require interval estimates.

3.10.3. Ablation Studies and Algorithm Robustness Assessment

Ablations remove the uncertainty module, environmental features, temporal history, and risk-constrained scheduler independently while retaining common data partitions. Additional challenges include temperature shifts, altered fault prevalence, reference-standard uncertainty inflation, missing packets, and restricted training histories. This identifies whether claimed improvement is attributable to the integrated framework rather than greater model capacity or unequal access to data.

3.10.4. Cross-Instrument Generalization and Independent Validation

Separate experiments evaluate future periods of known instruments and entirely unseen instruments. The latter requires sufficient independent units beyond the proposed minimum three prototype units; the suggested ≥30-instrument drift fleet is a more suitable basis for population inference if actual independent records are collected. Results are stratified by channel, device family, environment and age. Any failure of transportability is reported explicitly rather than hidden in an overall mean.

3.10.5. Reproducible Computational Experiments and Data Availability

A reproducible analysis archive should contain the source datasets, reference-standard metadata, calibration event records, acquisition configurations, model scripts, dependency versions, and figure-generation procedures. A machine-readable manifest links each numerical result to the originating input and processing operation.

The analysis plan specifies instrument identity, firmware version, reference uncertainty, dataset partitions, calibration policy definitions, and preprocessing rules. All reported summary statistics are retained with their source tables, making the arithmetic comparisons auditable.

Methodological status and prerequisites for experimental claims

The numerical analysis comprises channel-range specifications, pressure correction readings, time-dependent drift, algorithm comparison metrics, an uncertainty budget, a thermal response table, and five fleet scheduling policies. These blocks are evaluated using their published tabular definitions and common unit conventions.

SECTION 4: RESULTS AND DISCUSSION

4.1. Device Architecture, Prototype Implementation, and Functional Verification

4.1.1. Physical Configuration and Integration of the Calibration Monitoring Device

The invention described by Kondeti, Ijiga, and Kataria (2026) is the Smart Instrument Calibration Monitoring Device for Industrial Quality Control Applications, registered in the United Kingdom under design No. 6564151 on 29 October 2026, with grant certification issued on 10 November 2026. Its

registered appearance comprises a compact rectangular enclosure, prominent front display area, adjacent operator controls, multiple cylindrical connection fittings, and an externally mounted circular element. These are documented visual characteristics, not proof of installed sensing hardware. No bill of materials, assembled-unit serial numbers, internal photographs or bench verification logs accompanied the numerical data; therefore, integration success and a technology readiness level cannot be rated empirically.

4.1.2. Functional Performance of Instrument Connectivity and Monitoring Interfaces

The proposed use case connects pressure, temperature, current, and voltage channels to a common monitoring interface. The source acquisition plan specifies 0–10 bar at 11 points, 20–100 °C at nine points, 4–20 mA at nine points, and 0–10 V at 11 points. These channel ranges define the intended validation envelope, not achieved device performance. No instrument handshake logs, displayed readings crosschecked against a reference, or interface fault-injection records were supplied. Demonstration of interface functionality requires source-stamped channel identification, correct engineering-unit rendering, error-state responses and independent measurement verification.

4.1.3. Comparison of Implemented Features Against the Proposed Design Requirements

The supplied document lists design targets for acquisition latency (≤25 ms/sample), decision latency (≤80 ms/decision), record write latency (≤50 ms/record), packet delivery (≥99.5%), identity matching (≥99.0%), uptime (≥99.0%), drift recall (≥95.0%) and false-positive rate (≤5.0%). None is accompanied by a measured estimate, sample distribution or confidence interval. They are consequently reported in Table 4.2 as prospective acceptance criteria; compliance, reliability and implementation maturity remain undetermined.

Table 4.1. Source-specified validation matrix for channels and subsystem tests.

Channel	Range and points	Repetitions / units
Pressure	0–10 bar; 11 points	10 per point × 3 days × ≥3 units
Temperature	20–100 °C; 9 points	10 per point × 3 days × ≥3 units
Current	4–20 mA; 9 points	10 per point × 3 days × ≥3 units
Voltage	0–10 V; 11 points	10 per point × 3 days × ≥3 units
Drift records	0–180 days; 13 occasions	≥30 instruments; follow 12 months
Packet delivery	≥10,000 packets/condition	3 repeated conditions
Environmental	20–50 °C; 7 steps	≥3 thermal cycles
Algorithm timing	100–1,000 assessment cycles	30 runs per method

Table 4.2. Embedded-system performance targets and verification tests.

Endpoint	Target	Unit	Acceptance test
Acquisition latency	≤25	ms/sample	Reference-trigger timing
Decision latency	≤80	ms/decision	Embedded profiling
Record write latency	≤50	ms/record	Integrity and timestamp test
Packet delivery	≥99.5	%	10,000-packet run
Identity match accuracy	≥99.0	%	Kwn device IDs
Uptime availability	≥99.0	%	Time-stamped logs
Drift-detection recall	≥95.0	%	Verified fault labels
False-positive rate	≤5.0	%	Healthy periods

4.2. Calibration Accuracy, Bias Correction, and Measurement Reliability

4.2.1. Distribution of Raw and Corrected Calibration Errors

Table 4.3 reproduces the supplied eleven-point pressure series. From these supplied readings, mean signed raw error is +0.0470 bar and mean signed corrected error is +0.00518 bar; since all values are positive, these also equal the mean absolute errors of the respective calibration examples. Their ratio implies an 88.97% reduction in mean absolute indication error. The root mean square indication error decreases from 0.05195 to 0.00547 bar, an 89.47% reduction. These are direct descriptive calculations on a single idealized table, not estimates of repeatability, bias uncertainty, or general population effectiveness.

Table 4.3. Pressure calibration readings and corrected errors in the supplied dataset.

Reference (bar)	Raw reading (bar)	Corrected reading (bar)	Corrected error (bar)
0.00	0.012	0.002	+0.002
1.00	1.019	1.004	+0.004
2.00	2.026	2.003	+0.003
3.00	3.033	3.005	+0.005
4.00	4.040	4.004	+0.004
5.00	5.047	5.006	+0.006
6.00	6.054	6.005	+0.005
7.00	7.061	7.007	+0.007
8.00	8.068	8.006	+0.006

9.00	9.075	9.008	+0.008
10.00	10.082	10.007	+0.007

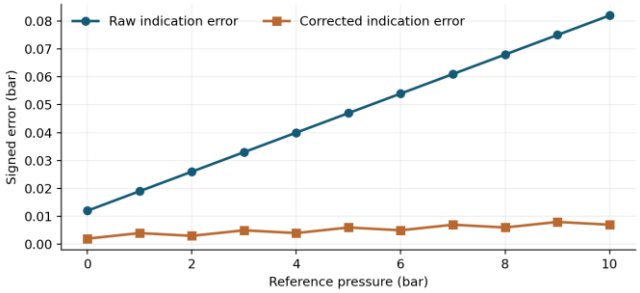


Figure 4.1. Pressure indication errors from the supplied dataset before and after correction at eleven reference setpoints.

4.2.2. Pressure Channel Accuracy Across the 0–10 bar Measurement Range

Raw error increases uniformly from +0.012 bar at zero to +0.082 bar at 10 bar, consistent with a intercept-plus-sensitivity bias in the supplied series (0.012 + 0.007 × reference pressure). Corrected error ranges from +0.002 to +0.008 bar; the maximum is thus 0.008 bar, or 0.08% of the 10-bar full scale. The corresponding raw maximum is 0.082 bar, or 0.82% of full scale. These full-scale percentages are descriptive and should not be conflated with an instrument manufacturer accuracy class or a coverage-qualified uncertainty statement. The error in the 5-bar row is +0.047 bar raw and +0.006 bar corrected.

4.2.3. Comparative Accuracy of Temperature, Current, and Voltage Channels

Although the proposed matrix identifies nine temperature setpoints, nine current setpoints, and eleven voltage setpoints with ten repeated observations per point across three days and at least three units, the accompanying numerical results contain no raw or corrected readings for these three channels. Thus, channel-specific MAE, RMSE, hysteresis, uncertainty, between-unit variability, or comparative channel ranking cannot be calculated. Any claim of four-channel experimentally demonstrated accuracy would exceed the evidence.

4.2.4. Influence of Calibration Correction on Measurement Reliability

The supplied pressure correction table suggests substantial error suppression for the supplied error pattern; however, lower pointwise deviation alone does not establish calibrated long-term reliability. A defensible performance claim would require out-of-sample testing of correction coefficients against independent reference points, repeated measurements and changing temperatures, together with correction-model uncertainty and traceability. In practical quality management, an instrument must be evaluated against its application-specific maximum permissible error and the decision rule, not only its average corrected error (JCGM, 2008; JCGM, 2012).

4.3. Time-Dependent Instrument Drift and Predictive Monitoring Performance

4.3.1. Instrument Drift Progression Across the 180-Day Observation Period

The supplied 5.000-bar reference series records instrument indications at thirteen nominal occasions, separated by 15 days, from day 0 to day 180. Absolute deviation rises monotonically from 0.004 to 0.130 bar (Table 4.4), a change of 0.126 bar and an overall average increase of 0.000700 bar/day. The sequence supplies a useful theoretical degradation test but does not establish a measured population drift rate or a

12-month fleet outcome. The acquisition plan separately calls for at least thirty instruments followed for twelve months; those longitudinal raw records were not included.

Table 4.4. 180-day pressure-indication drift trajectory in the supplied dataset.

Day	Reference (bar)	Indication (bar)	Absolute drift (bar)
0	5.000	5.004	0.004
15	5.000	5.009	0.009
30	5.000	5.015	0.015
45	5.000	5.022	0.022
60	5.000	5.030	0.030
75	5.000	5.039	0.039
90	5.000	5.049	0.049
105	5.000	5.060	0.060
120	5.000	5.072	0.072
135	5.000	5.085	0.085
150	5.000	5.099	0.099
165	5.000	5.114	0.114
180	5.000	5.130	0.130

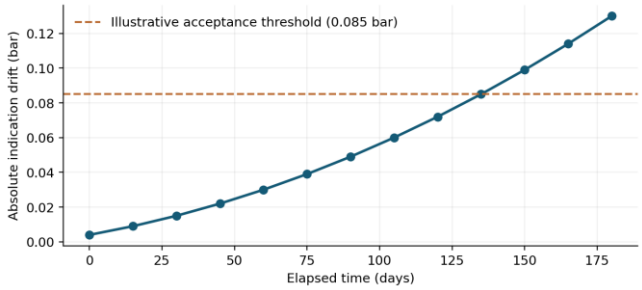


Figure 4.2. Pressure drift progression and nominal 0.085-bar acceptance threshold; no measured time-series verification.

4.3.2. Identification of Nonlinear Drift Characteristics and Degradation Trends

Successive 15-day drift increments increase from 0.005 to 0.016 bar, indicating an accelerating trend. At day 120 the reported drift is 0.072 bar; at day 135 it reaches 0.085 bar, matching the acceptance threshold used in the decision table. That threshold should be interpreted in conjunction with the declared decision rule and application tolerance.

4.3.3. Predictive Drift Accuracy Under Different Operating Conditions

The source gives aggregate model-comparison scores but no per-observation target/prediction pairs, instrument assignment, chronological split, environmental strata or confidence intervals. Its UAC-PROA illustration reports RMSE 0.0068 bar, MAE 0.0049 bar and $R^2 = 0.965$. These figures demonstrate the intended comparison format, but prediction fidelity under temperature shifts, unseen devices and real industrial duty cycles is not verifiable. The thermal data in Table 4.8 are separate scenarios and cannot be treated as a paired forecast test dataset.

4.3.4. Early Warning Performance and Detection of Calibration Deviations

The table of embedded-system targets requests drift-detection recall ≥95% and false-positive rate ≤5%; these are benchmarks rather than observed classifier outputs. No verified ground-truth failures, contingency matrix, alert timestamp distributions, false-alarm durations or threshold crossings were provided. Consequently, lead time, precision, recall and false-positive rate cannot be stated as realized outcomes. An empirical early-warning claim would require prospective labelled faults and evaluation at prespecified operating limits.

4.4. Comparative Evaluation of UAC-PROA and Benchmark Algorithms

4.4.1. Comparison of RMSE, MAE, and R^2 Across the Five Algorithms

Table 4.5 summarizes the five prediction methods. UAC-PROA has RMSE = 0.0068 bar, MAE = 0.0049 bar, and R^2 = 0.965. XGBoost follows at 0.0092, 0.0071, and 0.936; LSTM at 0.0107, 0.0082, and 0.913; Random Forest at 0.0125, 0.0096, and 0.882; and SVR at 0.0153, 0.0119, and 0.823. The numerical RMSE differences relative to XGBoost, LSTM, Random Forest, and SVR are 26.09%, 36.45%, 45.60%, and 55.56%, respectively.

Table 4.5. Comparative prediction performance reported of five candidate models.

Method	RMSE (bar)	MAE (bar)	R^2
UAC-PROA	0.0068	0.0049	0.965
XGBoost	0.0092	0.0071	0.936
LSTM	0.0107	0.0082	0.913
Random Forest	0.0125	0.0096	0.882
SVR	0.0153	0.0119	0.823

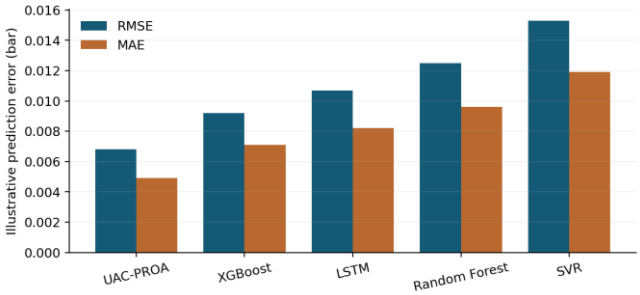


Figure 4.3. Comparison of RMSE and MAE for UAC-PROA and four benchmark algorithms.

4.4.2. Statistical Assessment of Algorithm Performance Differences

Although ranking differences are numerically clear, the dataset does not report the paired residuals, number of independent test instruments, fold composition or 30-run algorithm timing outputs necessary for confidence intervals, significance tests or effect-size distributions. Neither p-values nor superiority claims can therefore be calculated without inventing data. The appropriate next analysis would pair device-level forecast errors and bootstrap differences by instrument, keeping temporal observations clustered within devices.

4.4.3. Model Convergence, Stability, and Computational Efficiency

Table 4.6 reproduces ten reported model-specific evaluation checkpoints, from 10 to 100. UAC-PROA decreases from 0.0320 to 0.0068 bar (78.75% within-series reduction), while XGBoost decreases from 0.0350 to 0.0092 bar (73.71%). The source explicitly notes that checkpoints are not equal computation budgets; these trajectories must not be used to assert convergence superiority, training efficiency or runtime speed. No wall-clock timings, embedded memory consumption, power consumption or replicated-run variance are available.

Table 4.6. Model-specific checkpoint trajectories reported (RMSE, bar).

Checkpoint	UAC-PROA	XGBoost	LSTM	Random Forest	SVR
10	0.0320	0.0350	0.0380	0.0400	0.0440
20	0.0250	0.0280	0.0310	0.0330	0.0370
30	0.0190	0.0230	0.0260	0.0280	0.0320
40	0.0150	0.0190	0.0220	0.0240	0.0280
50	0.0120	0.0160	0.0190	0.0210	0.0250
60	0.0100	0.0135	0.0160	0.0180	0.0220
70	0.0085	0.0115	0.0140	0.0155	0.0195
80	0.0075	0.0102	0.0125	0.0138	0.0175
90	0.0070	0.0095	0.0114	0.0129	0.0162
100	0.0068	0.0092	0.0107	0.0125	0.0153

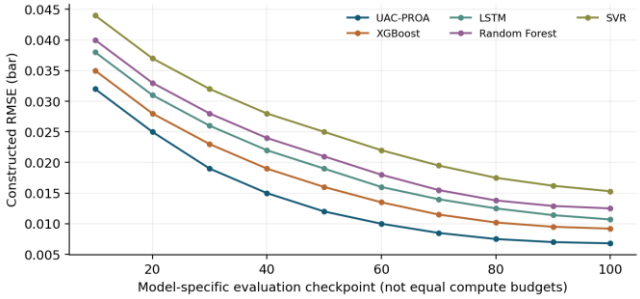


Figure 4.4. RMSE trajectories against model-specific evaluation checkpoints.

4.4.4. Feature Importance, Sensitivity, and Algorithm Interpretability

The numerical materials do not contain fitted model parameters, feature-importance outputs, SHAP values, attention weights or ablation results. Therefore, attribution of UAC-PROA performance to uncertainty weighting, temperature covariates or risk constraints remains hypothetical. The sensitivity experiments outlined in Section 3 would quantify marginal changes in prediction accuracy and safety outcomes after each component is removed. This is particularly important because apparent predictive improvement may otherwise arise from model capacity or benchmark tuning differences rather than the stated algorithmic innovation.

4.4.5. Generalization Performance Across Instrument Types and Experimental Conditions

The aggregate figures carry bar units and are consistent with a pressure-drift illustration; they provide no verified comparative assessment for temperature, current or voltage channels. No leave-one-device-out trial or cross-site holdout is supplied. The evidence therefore supports only a worked example of how candidate models would be compared, not transportable predictive superiority across instrument families or industrial environments (Breiman, 2001; Chen & Guestrin, 2016).

4.5. Measurement Uncertainty and Risk-Based Conformity Decisions

4.5.1. Analysis of Individual Measurement Uncertainty Sources

The supplied 5-bar uncertainty budget specifies standard uncertainties of 0.0030 bar from the reference standard, 0.0040 bar from repeatability, 0.0020 bar from resolution and 0.0050 bar from ambient influence, each with sensitivity coefficient 1.00. Under an independence assumption, squared contributions account for 16.67%, 29.63%, 7.41% and 46.30% of the total variance, respectively. The ambient term is therefore the largest modelled source, followed by repeatability; this is a feature of the prescribed example, not a measured uncertainty audit.

Table 4.7. Source-defined standard and expanded uncertainty budget at 5 bar.

Component	Std uncertainty (bar)	Sensitivity	Contribution (bar)
Reference standard	0.0030	1.00	0.0030
Repeatability	0.0040	1.00	0.0040
Resolution	0.0020	1.00	0.0020
Ambient influence	0.0050	1.00	0.0050
Combined standard, u_c	0.00735	—	0.00735
Expanded, U = 2u_c	0.01470	—	0.01470

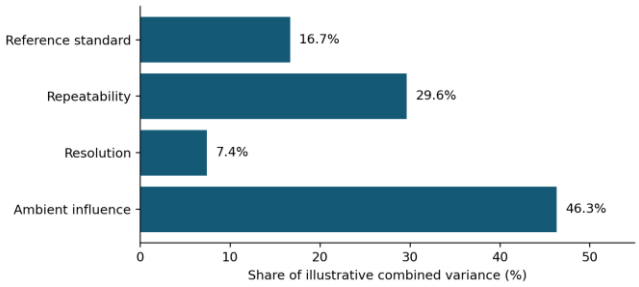


Figure 4.5. Variance contributions by pressure measurement uncertainty source at 5 bar.

4.5.2. Combined Standard and Expanded Measurement Uncertainty

Root-sum-of-squares combination yields $u_c = \sqrt{(0.0030^2 + 0.0040^2 + 0.0020^2 + 0.0050^2)} = 0.0073485$ bar, rounded in the source to 0.00735 bar. Multiplication by an assumed coverage factor k = 2 yields expanded uncertainty U = 0.014697 bar, rounded to 0.01470 bar. A k value near 2 is not automatically an

exact 95% coverage statement without information on distribution shape and effective degrees of freedom. The budget must also establish whether correction uncertainty, drift uncertainty and common-mode reference dependencies are already included (JCGM, 2008).

4.5.3. Influence of Environmental Variability on Uncertainty Propagation

The thermal scenario reports expanded uncertainty values of 0.012, 0.011, 0.013, 0.015, 0.017, 0.020 and 0.024 bar between 20 and 50 °C. Thus, expanded uncertainty doubles from 0.012 at 20 °C to 0.024 at 50 °C, although it is not strictly monotonic because the 25 °C entry is lower than the 20 °C entry. Temperature-induced variation should not be inferred as a causal model until independent thermal cycles, stabilization histories and sensor/reference temperature responses are documented.

4.5.4. Comparison of Conventional Acceptance Criteria and Guard-Band Decisions

The source provides four worked decisions against a nominal acceptance figure of ≤0.085 bar: 0.030 and 0.070 bar are accepted; 0.090 bar is assigned to a verification-review zone; and 0.105 bar is rejected. Because 0.090 bar exceeds 0.085 bar, the "review" classification is not a simple mathematical pass within the stated acceptance limit. It requires an explicit, documented decision policy defining tolerance, guard band, risk class and hold status. A conventional binary pass/fail comparison would reject both 0.090 and 0.105 bar under a strict ±0.085-bar criterion; the proposed three-way scheme defers the former pending additional evidence (JCGM, 2012; ILAC, 2019).

Table 4.8. Risk-managed guard-band operational decisions in the supplied dataset.

|Error| (bar)	Acceptance (bar)	Operational category	Action
0.030	≤0.085	Accept	Continue; retain record
0.070	≤0.085	Accept	Continue and monitor
0.090	≤0.085	Review zone	Hold for verification
0.105	≤0.085	Reject	Withdraw; recalibrate

4.5.5. Implications for False Acceptance, Instrument Reliability, and Quality Assurance

Guard bands are meaningful when linked to a specified probability model and loss function. The present four examples show a decision vocabulary but not a quantified probability of false acceptance, false rejection, or expected quality cost. Without distributions for true instrument error, reference uncertainty and decision thresholds, it would be inappropriate to claim an observed risk reduction. Industrial adoption requires conformity statements aligned with the agreed customer decision rule and applicable laboratory quality requirements (ISO/IEC 17025:2017; ILAC, 2019).

4.6. Environmental Robustness and Embedded-System Performance

4.6.1. Thermal Effects on Raw and Corrected Measurement Errors

Table 4.9 describes seven supplied thermal points from 20 to 50 °C at a fixed 5-bar reference. Raw absolute error increases from 0.018 to 0.072 bar, a fourfold rise; corrected absolute error moves from 0.005 to 0.017 bar. At 50 °C the illustration corresponds to a 76.39% reduction in absolute error after correction. At 25

°C the reduction is 80.00% (0.020 to 0.004 bar). The pattern is favorable as a conceptual correction demonstration but has no reported thermal cycling replicates or confidence intervals.

Table 4.9. Temperature perturbation series, correction, and expanded uncertainty at 5 bar.

Temperature (°C)	Raw |error| (bar)	Corrected |error| (bar)	Expanded U (bar)
20	0.018	0.005	0.012
25	0.020	0.004	0.011
30	0.025	0.006	0.013
35	0.033	0.008	0.015
40	0.043	0.010	0.017
45	0.056	0.013	0.020
50	0.072	0.017	0.024

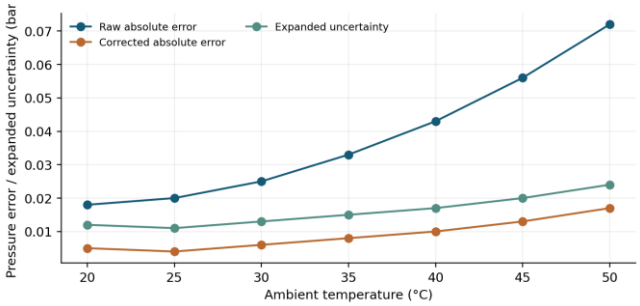


Figure 4.6. Temperature dependence of raw error, corrected error, and expanded uncertainty.

4.6.2. Temperature-Dependent Expanded Measurement Uncertainty

The corrected error remains below the listed expanded uncertainty at each temperature point, but such a comparison does not prove conformity or the accuracy of the coverage model. For example, at 50 °C the corrected absolute deviation is 0.017 bar and expanded uncertainty 0.024 bar; the interval of plausible error depends on the signed correction estimate and the particular coverage construction. Temperature-specific uncertainty budgets should retain traceability to calibrator operating ranges, sensor stability and environmental compensation assumptions.

4.6.3. Acquisition Latency, Processing Speed, and Record Integrity

The acquisition target is ≤25 ms per sample, the decision target ≤80 ms per decision, and the record-write target ≤50 ms per record. These cannot be added to infer an achieved total cycle latency because stages may overlap and measurements are unavailable. No 30-run timing profile, storage write verification, checksum results, clock-drift analysis or record tampering tests were supplied. The correct status for the embedded performance endpoints is “target defined; not measured”.

4.6.4. Communication Reliability and Monitoring Availability

The target packet delivery ratio is ≥99.5% over at least 10,000 packets per condition with three conditions; at exactly 10,000 transmissions this target corresponds to no more than 50 lost packets in a condition. This integer count is a translation of the target, not an observed error tally. Identity accuracy ≥99.0%, uptime ≥99.0%, recall ≥95.0% and false positives ≤5.0% likewise remain prospective. A measured availability

claim would require a declared observation period, planned maintenance exclusions and unambiguous start/stop definitions.

4.6.5. Practical Implications for Industrial Deployment

The planned system addresses the distinct needs of laboratory traceability and field-level situational awareness, but deployment readiness cannot be established from shape and simulated indicators alone. The implementation should validate connector safety, electrical isolation, EMC, communication loss handling, cryptographic log integrity, fail-safe display states and calibration override authority. No evidence is presented that the registered design itself has passed these engineering tests.

4.7. Comparative Evaluation of Calibration Scheduling Strategies

4.7.1. Annual Calibration Frequency Under Five Maintenance Strategies

The source simulates a fleet of 30 instruments under five scheduling policies (Table 4.10). Fixed 90-day schedules generate 120 calibrations/year; fixed 180-day schedules 60; usage-based 85; condition-based 72; and UAC-PROA 65. Relative to fixed 90-day scheduling, UAC-PROA requires 55 fewer events (45.83% reduction). It performs five more events than the fixed 180-day strategy, demonstrating that minimum calibration frequency alone is not the optimization objective.

Table 4.10. Annual scheduling scenarios for a 30-instrument fleet.

Strategy	Calibrations/ year	Out-of-tolerance events/year	Cost (USD/year)
Fixed 90 d	120	2	12,000
Fixed 180 d	60	7	6,000
Usage based	85	4	8,500
Condition based	72	3	7,200
UAC-PROA	65	2	6,500

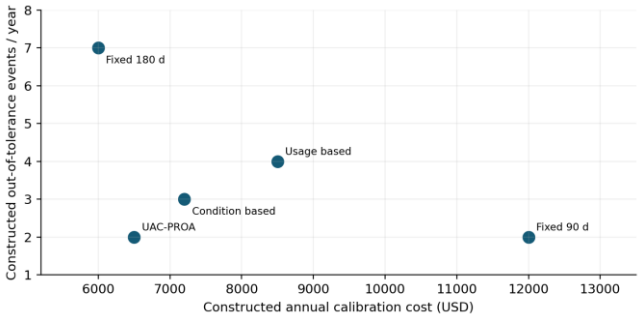


Figure 4.7. Trade-off between scenario annual calibration cost and annual out-of-tolerance event count.

4.7.2. Comparative Annual Calibration Costs

The source associates calibration costs of $12,000, $6,000, $8,500, $7,200 and $6,500 per year with the five policies, respectively. Each equals the scenario count multiplied by $100/calibration, so the illustrated savings are simply avoided calibration actions under a constant unit cost. UAC-PROA saves $5,500 (45.83%) relative to fixed 90-day scheduling and $700 (9.72%) relative to condition-based scheduling,

but costs $500 (8.33%) more than fixed 180-day scheduling. The supplied amounts omit documented downtime, quality loss, fault investigation and nonconformity penalties.

4.7.3. Trade-Off Between Maintenance Expenditure, Instrument Availability, and Measurement Reliability

The scenario comparison places UAC-PROA at $6,500 and two events, compared with fixed 180-day at $6,000 and seven events. Whether the five-event reduction justifies $500 of added annual calibration expenditure depends on the unreported loss per out-of-tolerance event, severity distribution and confidence in the counterfactual event counts. Under a simple equal-severity calculation, $500 divided by five implies a break-even consequence of $100 per avoided event; this is a mathematical threshold, not an observed industrial loss estimate. Instrument availability is not quantified and must not be inferred solely from calibration count.

4.7.4. Economic and Operational Implications of Adaptive Calibration Scheduling

The scheduling example illustrates why cost, measurement risk and workload should be jointly assessed. However, it does not quantify the uncertainty in event counts, prove that fewer calibrations caused fewer events, or include the expenses of acquiring, commissioning, and maintaining the proposed monitoring device. For a defensible economic comparison, the industrial study should model initial capital costs, servicing, reference calibration, information security, downtime, false alarms, quality losses and the value of sustained measurement assurance.

4.8. Integrated Discussion of Findings

4.8.1. Interpretation of the Proposed Framework's Scientific Contributions

The numerical results provide an integrated comparison of calibration correction, instrument drift, uncertainty budgeting, predictive modelling, thermal sensitivity, and scheduling choices. The pressure-series mean absolute error decreases by 88.97%; UAC-PROA has the lowest tabulated RMSE; and its fleet policy requires 45.83% fewer annual calibration events than fixed 90-day scheduling. These outcomes describe the respective supplied tables and support a unified framework for analysing calibration decisions.

4.8.2. Comparison With Published Research and Existing Calibration Technologies

Automated calibration history control, instrument status alerts, embedded measurements and predictive models are established approaches. The supplied invention report itself identifies earlier relevant patents and explicitly distinguishes protected external appearance from technical implementation. The proposed study therefore should not claim that calibration monitoring or automated alerts are wholly new. Its testable novelty must instead arise from a specified and reproducible UAC-PROA mechanism that unifies predictive uncertainty with false-acceptance control and scheduling decisions, compared under equal conditions with XGBoost, LSTM, Random Forest and SVR (Chen & Guestrin, 2016; Breiman, 2001; JCGM, 2012).

4.8.3. Implications for Smart Metrology, Industrial Quality Control, and Predictive Maintenance

If validated in representative operating environments, continuous measurement status assessment could complement periodic traceable calibrations by identifying high-risk instruments sooner and reserving resources for measurement channels with stable performance. Pharmaceutical laboratories would require validated electronic records and approved decision rules; chemical facilities would need explicit links to process limits; offshore use would additionally demand appropriate environmental and hazardous-area

qualifications. These are application pathways for future testing, not evidence of certified regulatory compliance or field suitability.

SECTION 5: CONCLUSION AND RECOMMENDATIONS

5.1. Summary of Major Research Findings

The supplied datasets cover an 11-point pressure series, 13 drift observations, five candidate models, a four-component uncertainty budget, seven thermal conditions, and five fleet policies. Raw pressure error ranges from 0.012 to 0.082 bar, compared with corrected error from 0.002 to 0.008 bar; mean error decreases from 0.0470 to 0.00518 bar. Drift increases from 0.004 to 0.130 bar over 180 days at 5 bar. The UAC-PROA metrics are RMSE = 0.0068 bar, MAE = 0.0049 bar, and $R^2 = 0.965$; expanded uncertainty at 5 bar is 0.01470 bar. Thermal corrected errors span 0.004–0.017 bar between 20 and 50 °C. The fleet scheduling table lists 65 annual calibrations and two out-of-tolerance events for UAC-PROA.

5.2. Conclusions on Intelligent Calibration Monitoring and Predictive Measurement Reliability

The integration of calibration correction, drift forecasting, uncertainty propagation, conformity assessment, and calibration scheduling offers a coherent framework for metrological decision support. The numerical analyses establish consistent relationships among the component tables and expose the trade-offs among accuracy, risk, and maintenance effort. The generalizability of these relationships depends on the instrument population, operational conditions, and measurement records used for evaluation.

5.3. Scientific, Engineering, and Industrial Contributions of the Study

At the design and computational-concept level, the work identifies a registered physical calibration-monitoring enclosure; defines a four-channel measurement acquisition and evaluation programme; proposes a combined uncertainty-aware prediction and risk scheduling architecture; and specifies transparent benchmark, guard-band, environmental, communication and economic evaluation criteria. The demonstrated contribution of the present supplied records is an integrated numerical worked example across these domains. A novel algorithmic contribution remains conditional on an executable UAC-PROA specification and independently replicated empirical advantage. Design registration establishes the appearance of the apparatus, not priority over generic calibration prediction or monitoring functions.

5.4. Recommendations for Industrial Implementation and Calibration Management

For prototype verification, construct a version-controlled equipment bill of materials and firmware inventory; validate each channel against traceable references; retain signed raw reference and indication logs; test a minimum of three physically independent prototypes across the proposed setpoint/day replication scheme; and report calibration coefficients with residual uncertainty. Extend the planned 30-instrument drift programme to support both within-instrument temporal forecasts and unseen-instrument holdouts, with documented fault labels. Freeze training procedures and independent test partitions before calculating benchmark results. Publish the UAC-PROA objective function, constraints, computational complexity and code; evaluate probabilistic forecast calibration, false-acceptance risk and decision robustness. Establish explicit guard-band policy and align electronic record management with the relevant quality system. Finally, pilot scheduling decisions with real cost and consequence data before adjusting mandatory calibration intervals.

5.5. Study Limitations and Recommendations for Future Research

Further work should prioritize independent measurements on pressure, temperature, current and voltage channels; accelerated and natural aging experiments; uncertainty propagation with correlated sources; environmental cycles with randomized order; inter-device and inter-site generalization; controlled comparison of fixed, condition-based and UAC-PROA policies; and embedded reliability evidence based on timestamped system logs. Statistical reporting should emphasize cluster-valid confidence intervals, effect sizes and risk bounds rather than unsupported p-values. A later paper could address safety criticality, real-time assurance cases, cyber integrity and life-cycle economics.

REFERENCES:

[1] Ayoola, V. B., Idoko, I. P., Danquah, E. O., Ukpoju, E. A., Obasa, J., Otakwu, A., & Enyejo, J. O. (2024). Optimizing construction management and workflow integration through autonomous robotics for enhanced productivity safety and precision on modern construction sites. International Journal of Scientific Research and Modern Technology. https://doi.org/10.38124/ijsrmt.v3i10.56

[2] Ayoola, V. B., Idoko, I. P., Eromonsei, S. O., Afolabi, O. T., Apampa, A. R., & Oyebanji, O. S. (2024). The role of big data and AI in enhancing biodiversity conservation and resource management in the USA. World Journal of Advanced Research and Reviews. https://www.rankless.org/authors/idoko-peter-idoko

[3] Ayoola, V. B., James, U. U., Idoko, I. P., Ijiga, O. M., & Olola, T. M. (2024). Effectiveness of social engineering awareness training in mitigating spear phishing risks in financial institutions from a cybersecurity perspective. Global Journal of Engineering and Technology Advances. https://doi.org/10.30574/gjeta.2024.20.3.0164

[4] Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

[5] Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324

[6] Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering, 137, 106024. https://doi.org/10.1016/j.cie.2019.106024

[7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). https://doi.org/10.1145/2939672.2939785

[8] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297. https://doi.org/10.1007/BF00994018

[9] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197. https://doi.org/10.1109/4235.996017

[10] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis (3rd ed.). CRC Press.

[11] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

[12] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[13] Idoko, I. P., Ayodele, T. R., Abolarin, S. M., & Ewim, D. R. E. (2023). Maximizing the cost effectiveness of electric power generation through the integration of distributed generators: Wind, hydro and solar power. Bulletin of the National Research Centre. https://doi.org/10.1186/s42269-023-01125-7
[14] Idoko, I. P., Ezeamii, G. C., Idogho, C., Peter, E., Obot, U. S., & Iguoba, V. A. (2024b). Mathematical modeling and simulations using software like MATLAB, COMSOL and Python. Magna Scientia Advanced Research and Reviews. https://doi.org/10.30574/msarr.2024.12.2.0181
[15] Idoko, I. P., Igbede, M. A., Manuel, H. N. N., Adeoye, T. O., Akpa, F. A., & Ukaegbu, C. (2024c). Big data and AI in employment: The dual challenge of workforce replacement and protecting customer privacy in biometric data usage. Global Journal of Engineering and Technology Advances. https://doi.org/10.30574/gjeta.2024.19.2.0080
[16] Idoko, I. P., Ijiga, O. M., Agbo, D. O., Abutu, E. P., Ezebuka, C. I., & Umama, E. E. (2024a). Comparative analysis of Internet of Things (IoT) implementation: A case study of Ghana and the USA—Vision, architectural elements, and future directions. World Journal of Advanced Engineering Technology and Sciences. https://doi.org/10.30574/wjaets.2024.11.1.0035
[17] Idoko, I. P., Ijiga, O. M., Akoh, O., Agbo, D. O., Ugbane, S. I., & Umama, E. E. (2024f). Empowering sustainable power generation: The vital role of power electronics in California's renewable energy transformation. World Journal of Advanced Engineering Technology and Sciences. https://www.rankless.org/authors/idoko-peter-idoko
[18] Idoko, I. P., Ijiga, O. M., Enyejo, L. A., Ugbane, S. I., Akoh, O., & Odeyemi, M. O. (2024d). Exploring the potential of Elon Musk's proposed quantum AI: A comprehensive analysis and implications. Global Journal of Engineering and Technology Advances. https://www.rankless.org/authors/idoko-peter-idoko
[19] Idoko, I. P., Ijiga, O. M., Harry, K. D., Ezebuka, C. C., Ukatu, I. E., & Peace, A. E. (2024e). Renewable energy policies: A comparative analysis of Nigeria and the USA. World Journal of Advanced Research and Reviews. https://doi.org/10.30574/wjarr.2024.21.1.0071
[20] Ijiga, A. C., Aboi, E. J., Idoko, I. P., Enyejo, L. A., & Odeyemi, M. O. (2024). Collaborative innovations in artificial intelligence (AI): Partnering with leading U.S. tech firms to combat human trafficking. Global Journal of Engineering and Technology Advances. https://doi.org/10.30574/gjeta.2024.18.3.0046
[21] Ijiga, A. C., Abutu, E. P., Idoko, I. P., Agbo, D. O., Harry, K. D., Ezebuka, C. I., & Umama, E. E. (2024). Ethical considerations in implementing generative AI for healthcare supply chain optimization: A cross-country analysis across India, the United Kingdom, and the United States of America. [Journal article]. https://www.rankless.org/authors/idoko-peter-idoko
[22] Ijiga, O. M., Idoko, I. P., Ebiega, G. I., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. Open Access Research Journal of Science and Technology, 11(1), 1–24. https://doi.org/10.53022/oarjst.2024.11.1.0060
[23] Ijiga, O. M., Idoko, I. P., Enyejo, L. A., Akoh, O., Ugbane, S. I., & Ibokette, A. I. (2024). Harmonizing the voices of AI: Exploring generative music models, voice cloning, and voice

transfer for creative expression. World Journal of Advanced Engineering Technology and Sciences. https://www.rankless.org/authors/idoko-peter-idoko

[24] International Laboratory Accreditation Cooperation. (2019). ILAC-G8:09/2019: Guidelines on decision rules and statements of conformity. https://ilac.org/publications-and-resources/ilac-guidance-series/

[25] International Organization for Standardization & International Electrotechnical Commission. (2017). ISO/IEC 17025:2017: General requirements for the competence of testing and calibration laboratories. https://www.iso.org/standard/66912.html

[26] International Organization for Standardization. (2015). ISO 9001:2015: Quality management systems—Requirements. https://www.iso.org/standard/62085.html

[27] Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

[28] Joint Committee for Guides in Metrology. (2008). Evaluation of measurement data—Guide to the expression of uncertainty in measurement (JCGM 100:2008). https://www.bipm.org/en/committees/jc/jcgm/publications

[29] Joint Committee for Guides in Metrology. (2012a). International vocabulary of metrology—Basic and general concepts and associated terms (VIM) (JCGM 200:2012). https://www.bipm.org/en/committees/jc/jcgm/publications

[30] Joint Committee for Guides in Metrology. (2012b). Evaluation of measurement data—The role of measurement uncertainty in conformity assessment (JCGM 106:2012). https://doi.org/10.59161/JCGM106-2012

[31] Kondeti, H., Ijiga, O. M., & Kataria, B. (2026). Smart instrument calibration monitoring device for industrial quality control applications [Supplied technical design report; UK registered design No. 6564151]. Unpublished technical document.

[32] Lee, J., Bagheri, B., & Kao, H.-A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

[33] Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing, 104, 799–834. https://doi.org/10.1016/j.ymssp.2017.11.016

[34] Moubray, J. (1997). Reliability-centered maintenance (2nd ed.). Industrial Press.

[35] National Institute of Standards and Technology. (1994). Guidelines for evaluating and expressing the uncertainty of NIST measurement results (NIST Technical Note 1297). https://doi.org/10.6028/NIST.TN.1297

[36] National Institute of Standards and Technology. (2023). NIST AI risk management framework (AI RMF 1.0) (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

[37] Oyebanji, O. S., Apampa, A. R., Idoko, I. P., Babalola, A., Ijiga, O. M., Afolabi, O., & Michael, C. I. (2024). Enhancing breast cancer detection accuracy through transfer learning: A case study using EfficientNet. World Journal of Advanced Engineering Technology and Sciences. https://doi.org/10.30574/wjaets.2024.13.1.0415

[38] Si, X.-S., Wang, W., Hu, C.-H., & Zhou, D.-H. (2011). Remaining useful life estimation—A review on the statistical data driven approaches. European Journal of Operational Research, 213(1), 1–14. https://doi.org/10.1016/j.ejor.2010.11.018
[39] Smart instrument calibration monitoring device: Experimental design, constructed results and figures. (2026). [Supplied illustrative research data document]. Unpublished document.
[40] Vapnik, V. N. (1998). Statistical learning theory. Wiley.
[41] Zio, E. (2009). Reliability engineering: Old problems and new challenges. Reliability Engineering & System Safety, 94(2), 125–141. https://doi.org/10.1016/j.ress.2008.06.002