International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

Physics Informed Transformer Failure Prediction Using Dissolved Gas Analysis, Thermal Aging Models, and Load Profiles

Bismark Kyere Yeboah
Energy Manager,
Facilities Operations,
University Of Rhode Island
(URI), USA

Abstract: Power transformer reliability is fundamental to electricity distribution networks, where unexpected equipment failures disrupt service continuity, increase maintenance expenditure, and compromise infrastructure resilience. Conventional condition assessment methods frequently evaluate dissolved gas concentrations, thermal deterioration, and loading conditions independently, limiting their capacity to capture interacting degradation mechanisms and predict emerging faults. This study proposes a physics informed predictive framework integrating dissolved gas analysis, thermal aging models, and historical load profiles to estimate transformer failure risk and support condition-based maintenance. Dissolved concentrations of hydrogen, methane, ethylene, acetylene, and carbon oxides characterise electrical and thermal fault signatures, while hotspot temperature calculations and insulation aging equations quantify cumulative thermal deterioration. Time varying load profiles capture operational stress, overload frequency, and temperature dependent degradation. Physics constrained machine learning models combine these indicators with transformer operating characteristics to estimate fault probabilities and remaining useful life. Model evaluation incorporates temporal validation, prediction uncertainty, calibration, and comparisons against conventional diagnostic approaches. The proposed framework strengthens interpretable asset health assessment, supports maintenance prioritisation, and provides a foundation for reducing unplanned outages, optimising replacement decisions, and improving long term distribution transformer reliability.
Keywords: Power Transformer Reliability; Physics-Informed Machine Learning; Dissolved Gas Analysis; Thermal Aging Modelling; Load Profile Analytics; Remaining Useful Life Prediction

1. INTRODUCTION

1.1 Transformer Reliability and Electricity Network Resilience

Power and distribution transformers constitute critical infrastructure within electricity transmission, distribution, and institutional energy systems, enabling voltage regulation and continuous power delivery across interconnected networks [1]. Their operational reliability supports industrial production, healthcare facilities, university campuses, and essential public services, where interruptions can generate substantial technical and economic consequences [2]. Insulation breakdown, winding deformation, thermal overstress, and dielectric deterioration progressively compromise transformer integrity, increasing susceptibility to internal faults and unexpected outages [3]. Such failures impose replacement expenditure, emergency maintenance costs, prolonged procurement delays, and revenue losses associated with interrupted electricity supply [4]. Furthermore, equipment unavailability reduces network redundancy and constrains operational flexibility during fluctuating demand conditions. Consequently, asset condition monitoring and predictive maintenance provide increasingly important alternatives to exclusively time-based inspection and reactive replacement, supporting earlier intervention and improved infrastructure resilience [5].

1.2 Limitations of Conventional Transformer Fault Diagnosis

Conventional transformer condition assessment employs dissolved gas analysis (DGA), thermal aging calculations, and operational load monitoring to identify developing electrical and insulation abnormalities [6]. DGA characterises gases released through dielectric decomposition, whereas thermal models estimate insulation deterioration and load measurements quantify operational stress [7]. However, these diagnostic approaches are frequently interpreted independently, limiting recognition of interactions between loading fluctuations, winding temperatures, and gas evolution. Fixed diagnostic thresholds may overlook equipment-specific characteristics, while overlapping gas signatures complicate fault identification [8]. Irregular sampling intervals, laboratory measurement uncertainty, incomplete maintenance histories, and inconsistent failure records further constrain predictive reliability. Conventional machine learning additionally encounters overfitting, temporal data leakage, inadequate physical interpretability, and unstable extrapolation under extreme loading conditions [9]. These limitations motivate integrated approaches combining physical degradation relationships with data-driven predictive modelling.

1.3 Research Aim, Objectives and Contributions

This study proposes a physics-informed predictive framework integrating dissolved gas analysis, thermal aging models, and

historical load profiles to evaluate transformer deterioration and estimate emerging failure risks. The framework establishes relationships between dissolved gas evolution, temperature-dependent insulation degradation, and operating stress, enabling diagnostic measurements to be interpreted through physically meaningful constraints [3]. Three prediction tasks are considered: classification of verified transformer fault mechanisms, estimation of failure probabilities within predefined observation horizons, and remaining useful life prediction where sufficiently reliable failure histories exist [7]. Research objectives include multimodal data acquisition, temporal synchronisation, diagnostic feature engineering, physics-constrained model development, and systematic hyperparameter optimisation. Training procedures will incorporate transformer-level data partitioning, cross-validation, independent testing, and uncertainty quantification to minimise information leakage and improve generalisability [9]. Comparative benchmarking against conventional statistical and machine learning models will assess predictive discrimination, calibration, and error characteristics. Diagnostic interpretations will additionally be examined against applicable IEC and IEEE guidance to establish consistency with recognised transformer assessment practices [6]. The anticipated contribution is an interpretable analytical framework supporting condition-based maintenance, risk prioritisation, and evidence-informed replacement planning. Nevertheless, improvements in prediction accuracy, operational reliability, and maintenance effectiveness must be demonstrated through empirical validation rather than presumed from architectural complexity alone.

2. TRANSFORMER DEGRADATION PHYSICS AND DIAGNOSTIC FOUNDATIONS

2.1 Electrical and Thermal Failure Mechanisms

Transformer deterioration arises from interacting electrical, thermal, chemical, and mechanical stresses that progressively weaken oil–paper insulation systems [6]. Partial discharge develops within localised dielectric imperfections, generating hydrogen and other low-molecular-weight gases, while low-energy and high-energy discharges produce increasingly severe hydrocarbon decomposition signatures [7]. Localised overheating accelerates mineral oil degradation, whereas sustained winding temperatures promote cellulose depolymerisation, reducing insulation mechanical strength and dielectric endurance [8]. Moisture ingress intensifies hydrolytic degradation, increases dielectric losses, and contributes to insulation instability under elevated electrical stress [9]. These processes evolve from fault initiation through progressive material deterioration and measurable diagnostic indications before potentially culminating in dielectric breakdown or terminal equipment failure [10]. Their progression depends on transformer design, insulation composition, cooling configuration, operating temperature, and accumulated service exposure. Consequently, identical diagnostic gas concentrations may represent different deterioration states across transformers with dissimilar operating histories and thermal characteristics [11].

2.2 Dissolved Gas Generation and Diagnostic Ratios

Dissolved gas analysis characterises transformer deterioration through hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), carbon monoxide (CO), and carbon dioxide (CO_2) concentrations [7]. Hydrogen commonly accompanies partial discharge, acetylene indicates sufficiently energetic electrical decomposition, while carbon oxides provide evidence relevant to cellulose deterioration [12]. Gas accumulation is evaluated through the finite-difference approximation:

$$G_i = \frac{C_i(t_2) - C_i(t_1)}{t_2 - t_1}$$

Here, G_i represents the observed gas concentration change rate in ppm/day; $C_i(t_1)$ and $C_i(t_2)$ denote successive concentrations in ppm; and t_1, t_2 are sampling times in days. The expression approximates the concentration-time gradient. Positive values indicate net accumulation, whereas negative values may reflect degassing, oil processing, or measurement variability. Diagnostic ratios CH_4/H_2, C_2H_2/C_2H_4, and C_2H_4/C_2H_6 support fault interpretation but require corroborating evidence [13].

2.3 Thermal Aging and Insulation Life Consumption

Thermal aging progressively deteriorates cellulose insulation through temperature-dependent chemical reactions, reducing polymerisation degree and mechanical integrity [8]. Winding hot-spot temperature provides a critical indicator because localised heating can exceed average winding and top-oil temperatures, particularly during sustained overloads [10]. Arrhenius kinetics describe the exponential relationship between reaction rates and absolute temperature. For an applicable IEEE reference model, the aging acceleration factor is expressed as:

$$F_{AA} = \exp\left(\frac{15000}{383} - \frac{15000}{\theta_H + 273}\right)$$

Here, F_{AA} denotes dimensionless aging acceleration, θ_H represents hot-spot temperature in °C, 383 K corresponds to the 110 °C reference, and 15000 K is the model's temperature-sensitivity constant. The equation follows by dividing the Arrhenius aging rate at operating temperature by its reference-temperature counterpart [11]. Consequently, $F_{AA} = 1$ indicates reference aging, while higher values represent accelerated deterioration. Applicability depends on compatible insulation characteristics and reference conditions; cumulative life consumption additionally requires integrating acceleration over operating time [14].

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

2.4 Load-Dependent Thermal Stress and Physics-Informed Prediction

Transformer loading governs winding heat generation, while ambient temperature, cooling efficiency, and thermal inertia determine the resulting temperature trajectory [10]. Repeated overloads and thermal cycling increase cumulative insulation exposure, potentially accelerating deterioration even when instantaneous operating conditions remain acceptable [14]. Historical load profiles therefore provide essential context for interpreting gas evolution and insulation aging. Physics-informed prediction integrates these relationships into learning constraints, enabling statistical models to account for thermal dynamics, degradation consistency, and operating history rather than relying exclusively on empirical correlations [11].

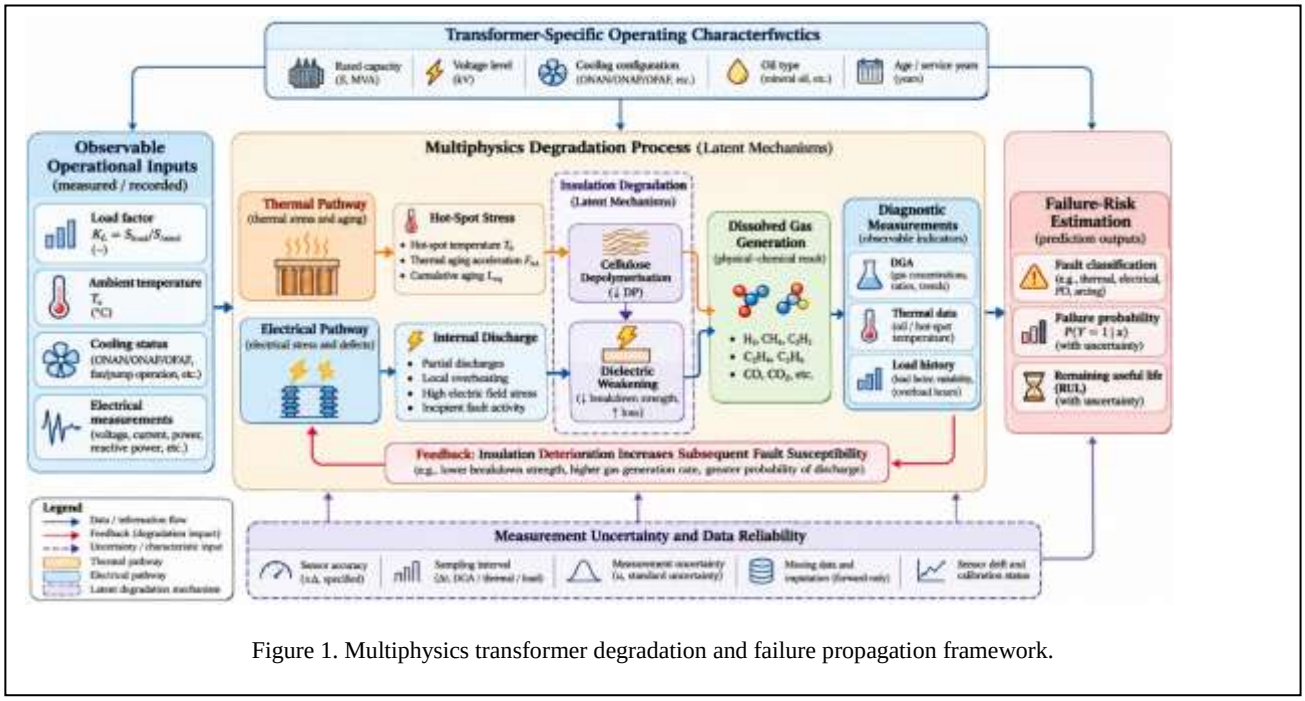


Figure 1. Multiphysics transformer degradation and failure propagation framework.

3. DATA ACQUISITION AND EXPERIMENTAL DESIGN

3.1 Transformer Population, Sampling Design and Asset Characteristics

The proposed study population encompasses oil-immersed distribution and power transformers operating across utility networks and institutional energy facilities [13]. Asset stratification considers voltage class, rated capacity, cooling configuration, installation environment, service age, and insulation technology to capture variations in operating conditions and degradation susceptibility [14]. Each transformer receives a unique anonymised identifier linking asset specifications, operating histories, maintenance interventions, and diagnostic observations. Eligibility requires traceable operating records, identifiable sampling dates, and sufficient condition measurements for longitudinal analysis [15]. Observation windows will encompass historical loading patterns, successive diagnostic assessments, and documented maintenance events, recognising that sampling frequency influences the detection of evolving faults [16]. Transformer-level grouping will prevent observations from the same asset entering incompatible development and independent testing cohorts. Differences between utility substations and institutional installations, including loading diversity, maintenance schedules, and monitoring infrastructure, will be documented to support meaningful cross-population comparisons [17].

3.2 Dissolved Gas Analysis and Oil Condition Data

Dissolved gas analysis records will comprise laboratory measurements and available online monitoring outputs for hydrogen, methane, ethane, ethylene, acetylene, carbon monoxide, and carbon dioxide [18]. Each observation will include its sampling timestamp, analytical method, concentration in parts per million, detection limit, and documented measurement uncertainty. Oil sampling procedures will be assessed against applicable IEC 60567 guidance, including representative collection, contamination prevention, gas extraction, and laboratory quality assurance [19]. Instrument calibration histories, analytical repeatability, and interlaboratory differences will be examined to identify potential systematic measurement errors. Supplementary condition variables will include moisture concentration, acidity, dielectric breakdown voltage, and related oil quality indicators where available [20]. Maintenance records will identify oil filtration, degassing, replacement, and treatment activities that may alter dissolved gas trajectories independently of fault progression. Diagnostic interpretation will consider IEC 60599 and IEEE C57.104 guidance while recognising transformer-specific operating conditions and uncertainty associated with individual measurements [21]. Records lacking traceable sampling provenance will undergo quality review before inclusion.

3.3 Thermal Measurements and Load Profile Acquisition

Operational measurements will be acquired from supervisory control and data acquisition systems, smart meters, and transformer monitoring equipment [14]. Recorded variables will include phase currents, terminal voltages, active and apparent power, load factor, ambient temperature, top-oil temperature, and winding hot-spot temperature. Direct temperature measurements will be distinguished from estimates obtained through thermal models incorporating loading conditions, cooling characteristics, and transformer-specific parameters [22]. High-frequency electrical and thermal observations will be aggregated into consistent intervals while preserving short-duration overloads and temperature excursions. These records will subsequently be aligned with lower-frequency DGA sampling timestamps using predefined historical observation windows [16]. Cooling modes, fan activation, pump operation, seasonal ambient variation, and abnormal loading events will be documented to characterise thermal response. Missing sensor intervals and uncertain hot-spot estimates will be explicitly flagged, preventing model-derived temperatures from being misrepresented as direct observations during subsequent feature engineering.

3.4 Failure Labels, Maintenance Records and Outcome Definition

Transformer fault labels will be established from verified inspection reports, protective relay operations, maintenance findings, and documented failure investigations [15]. Incipient insulation deterioration, diagnostic alarms, corrective maintenance, and confirmed functional failures will be classified separately to prevent diagnostic indications from being treated as terminal events [17]. Binary prediction outcomes may represent verified failure within predefined horizons of 90 or 180 days, selected according to event frequency and maintenance requirements. Fault classification will use independently supported diagnostic categories, whereas remaining useful life estimation will require reliable event timing and censoring information [21]. Right-censored observations, uncertain failure dates, incomplete operational histories, and preventive replacements will be documented explicitly. Independent technical review will resolve conflicting event descriptions where sufficient evidence exists, while unresolved cases will retain uncertainty labels rather than receive unsupported classifications.

3.5 Data Quality, Traceability and Experimental Governance

Data governance will establish provenance tracking, anonymised asset identification, measurement units, calibration documentation, and systematic missingness assessment [19]. Inclusion criteria will require identifiable transformers, reliable timestamps, and adequate diagnostic or operational coverage; exclusion decisions will be documented before modelling. Version-controlled extraction scripts and preprocessing records will preserve reproducibility across analytical stages [20]. Independent test cohorts will remain isolated from feature selection, hyperparameter optimisation, and model development. Data originating from different utilities or institutions will be pooled only after assessing measurement compatibility, operating characteristics, label consistency, and institutional permissions [22]. These controls will support defensible benchmarking and reduce unintended information leakage.

Table 1. Multimodal Transformer Data Inventory, Measurement Sources and Quality Requirements

Data category	Representative variables	Acquisition source	Quality assessment
DGA	H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2	Laboratory analysis / online DGA monitors	Calibration, concentration uncertainty, detection limits, sampling integrity
Thermal	Top-oil temperature, hot-spot temperature, ambient temperature	Temperature sensors / thermal estimation models	Sensor accuracy, thermal model assumptions, missing observations
Electrical	Load factor, current, voltage, active and apparent power	SCADA / smart meters / monitoring systems	Sampling intervals, measurement continuity, timestamp accuracy
Asset characteristics	Service age, rated capacity, voltage class, cooling configuration, insulation type	Asset management databases / equipment registries	Metadata completeness, equipment identification, specification consistency
Maintenance	Oil filtration, degassing, repairs, inspections, component replacement	Maintenance management systems / work orders	Intervention timestamps, treatment descriptions, record completeness
Failure outcomes	Fault classification, confirmed failure date, censoring status	Incident reports / protection records / inspection findings	Label confidence, event verification, independent technical review

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

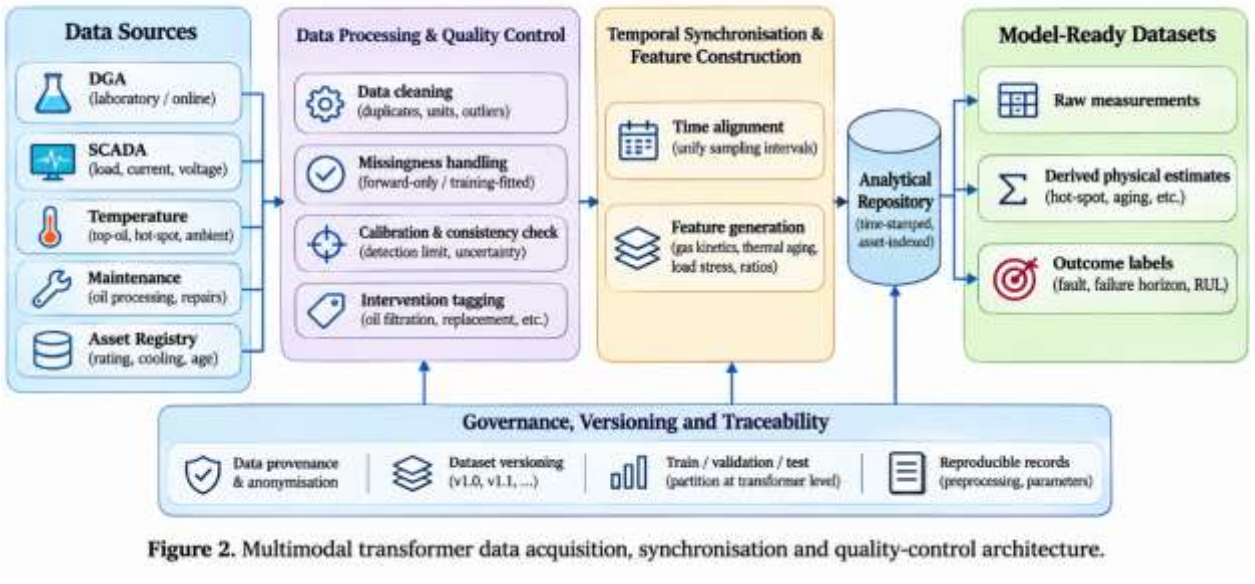


Figure 2. Multimodal transformer data acquisition, synchronisation and quality-control architecture.

The framework integrates laboratory and online DGA measurements, SCADA-derived electrical loading, thermal monitoring, asset registries, and maintenance records into a unified timestamped analytical repository. Individual acquisition streams undergo calibration verification, missing-data assessment, unit harmonisation, and source-specific quality checks before temporal synchronisation. Maintenance intervention tagging distinguishes genuine degradation trends from measurement changes caused by oil processing or equipment servicing. Separate processing pathways preserve raw measurements, physics-derived thermal and aging estimates, and independently verified failure labels. Dataset versioning, asset-level identification, and partition controls maintain traceability and prevent information leakage between model development and independent evaluation.

4. DATA PREPROCESSING AND PHYSICS-BASED FEATURE ENGINEERING

4.1 Data Cleaning, Missingness and Temporal Synchronisation

Transformer datasets will undergo duplicate removal, timestamp harmonisation, unit standardisation, and screening for physically impossible measurements [21]. Sensor drift will be assessed using calibration histories and cross-sensor consistency, while genuine fault-related extremes will remain unless measurement errors are independently established [22]. Missingness will be classified as completely random, conditionally dependent on observed variables, or operationally informative, particularly when monitoring interruptions coincide with equipment disturbances [23]. Forward-only imputation and training-fitted estimation methods will prevent future observations from influencing historical predictions. High-frequency loading and thermal records will be summarised within predefined windows preceding each DGA measurement, preserving overload peaks and thermal excursions [24]. Sampling gaps, measurement age, and intervention timestamps will remain explicit features, ensuring that temporal alignment does not obscure diagnostic uncertainty or introduce leakage.

4.2 Descriptive Statistics, Mean Deviation and Variability Assessment

Descriptive statistical analysis will characterise transformer operating conditions, gas concentration distributions, and temporal variability across equipment populations [22]. For valid observations X_i, the arithmetic mean and mean absolute deviation are defined as:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad MAD = \frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{X}|$$

Here, n denotes the observation count, $\bar{X}$ represents the arithmetic mean, and MAD measures average absolute dispersion around that mean in the original units. The mean divides the observation sum by its count, whereas MAD averages absolute deviations [25]. Standard deviation will quantify squared-deviation variability, while the coefficient of variation expresses relative dispersion when the mean is sufficiently nonzero. Median and interquartile range will characterise skewed distributions and extreme gas concentrations. Percentage deviation from a defined baseline

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

will measure relative condition changes [26]. Importantly, mean absolute deviation differs from prediction mean absolute error and median absolute deviation, which use different reference quantities.

4.3 DGA-Derived Diagnostic and Temporal Features

DGA feature engineering will transform individual gas concentrations into indicators representing electrical discharge, thermal decomposition, and evolving insulation deterioration [23]. Derived variables will include total dissolved combustible gas, CH_4/H_2, C_2H_2/C_2H_4, and C_2H_4/C_2H_6 ratios, concentration gradients, rolling averages, and gas accumulation acceleration [27]. Temporal features will be calculated exclusively from measurements preceding each prediction timestamp, ensuring consistency with operational deployment. Sampling intervals, laboratory uncertainty, detection limits, and oil treatment events will contextualise observed concentration changes. Ratios containing zero or near-zero denominators will receive explicit validity flags and predefined handling procedures rather than uncontrolled numerical substitution [24]. Correlation matrices and redundancy analysis will identify overlapping indicators, particularly among hydrocarbon gases and total combustible gas measures. Nevertheless, physically meaningful features will not be eliminated solely because of statistical correlation, since complementary diagnostic interpretations may remain valuable across distinct fault mechanisms [28].

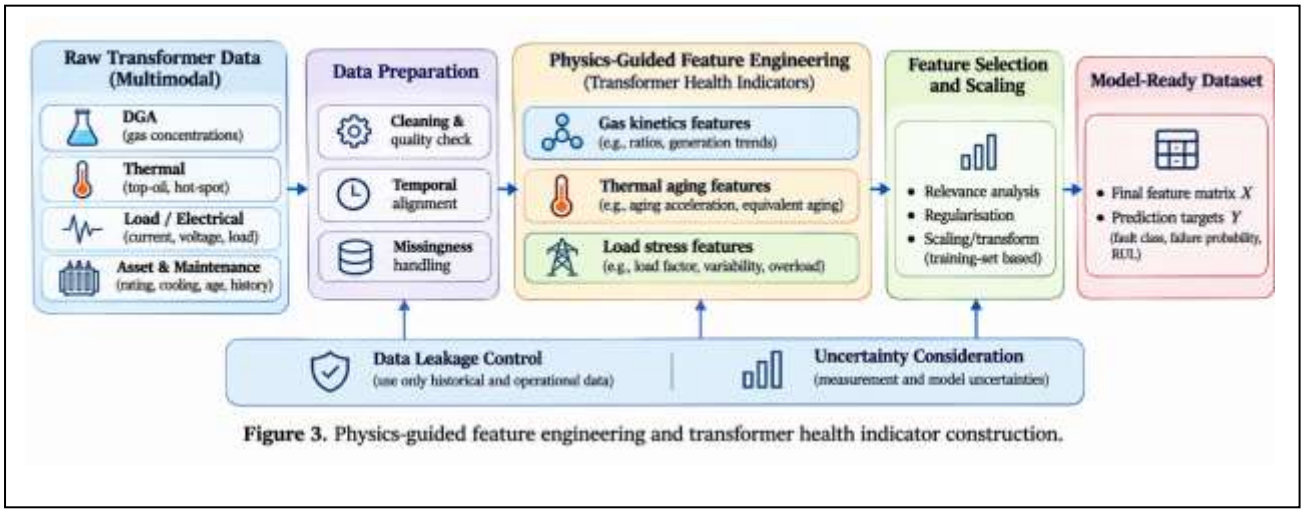


Figure 3. Physics-guided feature engineering and transformer health indicator construction.

4.4 Thermal Aging and Load Stress Feature Derivation

Thermal feature construction will translate transformer operating histories into physically interpretable indicators of cumulative insulation deterioration [24]. Integrating the aging acceleration factor over successive observation intervals yields equivalent reference-temperature aging exposure:

$$L_{\text{eq}} = \sum_{k=1}^{m} F_{AA,k}\,\Delta t_k$$

Here, L_{eq} denotes equivalent aging duration, $F_{AA,k}$ represents the dimensionless acceleration factor during interval k, Δt_k specifies interval duration, and m denotes the number of intervals. This discrete summation approximates continuous integration of temperature-dependent aging acceleration [26]. Complementary features will include cumulative overload hours, load-factor variability, thermal cycling frequency, maximum hot-spot temperature, and cooling-adjusted thermal stress. Rolling thermal exposure will distinguish sustained overheating from isolated temperature excursions [27]. Because hot-spot temperatures may be model-estimated, parameter uncertainty and sensor errors will propagate nonlinearly into aging estimates. Sensitivity analysis and uncertainty intervals will therefore accompany derived exposure measures, preventing estimated insulation consumption from being interpreted as directly observed physical deterioration [28].

4.5 Feature Selection, Scaling and Physical Consistency

Feature selection will combine correlation screening, permutation importance, regularisation, and degradation-based justification to balance predictive relevance with physical interpretability [25]. Scaling, logarithmic transformations, and missing-value estimators will be fitted exclusively within training partitions. Collinearity among loading, hot-spot temperature, and accumulated aging indicators will be examined without automatically removing physically informative relationships [27]. Feature availability will be verified against operational prediction timestamps, excluding future maintenance findings and post-failure measurements. The resulting feature matrix will integrate DGA kinetics, thermal aging, loading stress, and asset characteristics with explicitly defined fault classification, failure probability, and remaining-life targets [28].

The framework illustrates the transformation of raw dissolved gas measurements, thermal observations, electrical loading records, and asset characteristics into model-ready predictive variables. Initial processing includes timestamp harmonisation, measurement validation, missingness

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

assessment, and intervention tagging. Four feature-engineering pathways separately generate gas kinetic indicators, equivalent insulation aging exposure, load-dependent thermal stress measures, and transformer-specific contextual descriptors. Statistical characterisation incorporates mean absolute deviation, standard deviation, coefficient of variation, and baseline deviations. Subsequent feature screening combines statistical relevance with physical consistency, while dedicated checkpoints prevent temporal leakage and track uncertainty propagation from measured and estimated variables into the final predictive dataset.

5. PHYSICS-INFORMED MODEL DEVELOPMENT, TRAINING AND TESTING

5.1 Prediction Tasks, Model Inputs and Learning Architecture

The proposed prediction framework integrates dissolved gas indicators, thermal aging exposure, load stress variables, transformer characteristics, and historical maintenance information into a unified feature matrix [27]. Three learning objectives are considered: classification of verified fault mechanisms, estimation of confirmed failure probability within predefined horizons, and remaining useful life estimation where reliable event histories permit [28]. Fault identification constitutes multiclass classification, whereas horizon-specific failure prediction requires binary classification or survival modelling. Remaining useful life estimation may employ censored survival methods or regression when complete failure observations exist [29]. A hybrid architecture combines a physical degradation branch, representing thermal dynamics and insulation aging, with a data-driven branch capturing nonlinear diagnostic relationships. Transformer characteristics contextualise both branches. Prediction timestamps, observation windows, and forecast horizons will be explicitly defined, ensuring every input predates the corresponding prediction and remains operationally available [30].

5.2 Benchmark Models and Physics-Informed Learning Formulation

Benchmarking will include logistic regression, random forest, gradient boosting, XGBoost, support vector machines, and multilayer perceptrons, alongside physics-only thermal aging and diagnostic-rule baselines [28]. The proposed architecture will incorporate physics-informed neural learning or physics-regularised boosting, depending on dataset size, label availability, and constraint implementation [31]. Physical relationships will complement empirical associations without assuming that cumulative insulation aging independently determines failure. The composite objective is:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda_{\text{phys}}\mathcal{L}_{\text{phys}} + \lambda_{\text{reg}}\mathcal{L}_{\text{reg}}$$

Here, $\mathcal{L}_{\text{data}}$ represents task-specific prediction error, $\mathcal{L}_{\text{phys}}$ penalises physical inconsistency, and $\mathcal{L}_{\text{reg}}$ controls complexity. Their respective weights are λ_{phys} and λ_{reg}. The formulation augments empirical learning with explicit constraints and regularisation [32]. Physical residuals may quantify differences between predicted hot-spot temperatures and independently calculated thermal states, or inconsistencies between inferred aging exposure and temperature-dependent acceleration. Residuals will be normalised using training-derived scales before aggregation, preventing individual constraints from dominating optimisation through numerical magnitude [33]. Physics-loss applicability will be restricted to relationships supported by the available measurements.

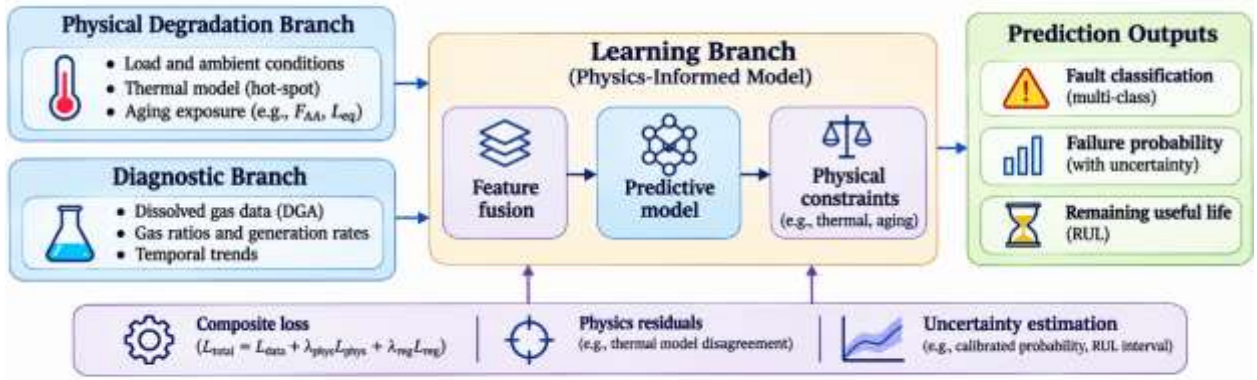


Figure 4. Hybrid physics-informed transformer failure prediction architecture.

The architecture integrates three interconnected processing branches. The physical degradation branch converts historical loading, ambient conditions, and cooling characteristics into hot-spot temperature estimates, aging acceleration factors, and cumulative insulation exposure. The diagnostic branch transforms dissolved gas concentrations into generation rates, diagnostic ratios, and temporal deterioration indicators. The learning branch fuses physical and diagnostic features with transformer characteristics through a predictive model incorporating task-specific loss functions, normalised physical

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

consistency residuals, and complexity regularisation. Separate output heads provide fault classification, horizon-specific failure probability, and remaining useful life estimates where suitable labels exist. Calibration and uncertainty assessment accompany the prediction outputs, while feedback from validation informs model optimisation without exposing independent test observations.

5.3 Training, Validation and Independent Test Data Splitting

An initial allocation of 70% training, 15% validation, and 15% independent testing will be considered, subject to transformer population size and confirmed failure prevalence [29]. Asset-level partitioning will prevent repeated observations from individual transformers appearing across development and held-out asset cohorts. Chronological cutoffs and prediction-horizon embargoes will exclude observations whose outcome windows overlap prohibited evaluation periods [30]. Forward temporal evaluation on previously observed transformers will assess forecasting under evolving operating conditions, whereas held-out transformer testing will examine generalisation to unseen equipment. These objectives will be evaluated separately rather than conflated. Rolling-origin validation or grouped cross-validation will support model selection within development data [31]. Imputation, scaling, feature screening, and imbalance correction will be fitted independently within training folds. The final test cohort will remain inaccessible during architecture selection, hyperparameter tuning, threshold optimisation, and calibration development, preserving unbiased performance assessment [34].

5.4 Hyperparameter Tuning and Model Selection

Hyperparameter optimisation will employ Bayesian optimisation, random search, or constrained grid search within predefined computational budgets [31]. Candidate settings will include learning rates, maximum tree depth, estimator counts, minimum leaf sizes, hidden-layer widths, dropout probabilities, batch sizes, weight decay, and physical-loss weighting coefficients. Search ranges will reflect dataset dimensions, event prevalence, and model complexity [32]. Early stopping will monitor validation loss or the prespecified primary metric to limit overfitting. Rare confirmed failures will be addressed using class-weighted objectives or resampling restricted to training folds. Model selection will prioritise a task-appropriate metric, such as area under the precision–recall curve for rare-event classification, supplemented by probability calibration, physical consistency, and inference cost [33]. All optimisation trials, search distributions, random seeds, stopping criteria, and selected configurations will be documented. Calibration parameters will be estimated without accessing the independent test cohort [34].

5.5 Prediction Uncertainty, Failure Probability and Calibration

For horizon-specific binary failure prediction, a learned score is transformed into probability through the inverse-logit function [29]:

$$P(Y = 1 \mid \mathbf{x}) = \frac{1}{1 + \exp[-f_\theta(\mathbf{x})]}$$

Here, $Y = 1$ indicates confirmed failure within the selected horizon, $\mathbf{x}$ denotes transformer features, $f_\theta(\mathbf{x})$ represents the learned score, and θ identifies model parameters. This transformation maps unrestricted scores onto the unit probability interval but does not guarantee calibration [32]. Reliability diagrams and Brier scores will assess agreement between predicted risks and observed event frequencies. Transformer-level bootstrap intervals will quantify sampling uncertainty, while perturbation analyses will examine sensitivity to missing gases, uncertain temperatures, and incomplete histories [33]. Remaining useful life uncertainty will be assessed separately through survival intervals or predictive distributions, accounting for censoring and event-time uncertainty [34].

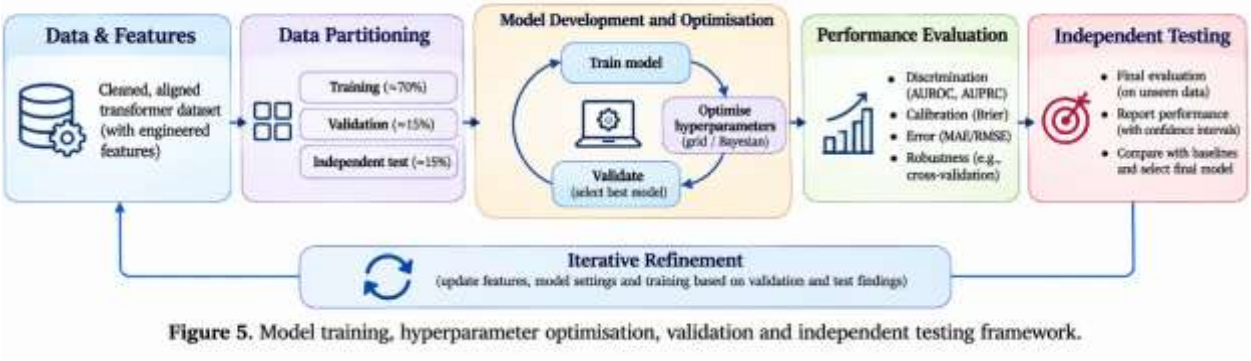


Figure 5. Model training, hyperparameter optimisation, validation and independent testing framework.

5.6 Independent Testing, Ablation Studies and Reproducibility

Following model selection, the final configuration will undergo one prespecified evaluation on the untouched

independent test cohort [30]. Ablation experiments will compare DGA-only, thermal-only, load-only, combined multimodal, and physics-constrained multimodal configurations using identical evaluation protocols. Subgroup analysis will examine transformer age, capacity, cooling arrangement, and loading severity to identify performance heterogeneity [33]. False alarms, missed failures, probability miscalibration, and unstable predictions will be investigated alongside discrimination metrics. Reproducibility documentation will preserve software versions, random seeds, preprocessing transformations, model configurations, and data partition identifiers. Evaluation will distinguish classification accuracy from probability reliability and physical consistency, ensuring that claimed benefits are supported by comparative evidence rather than architectural assumptions [34].

The workflow begins with the eligible transformer population and proceeds through asset-level partitioning, chronological separation, and prediction-horizon embargoes. Training-fold preprocessing includes missing-value treatment, scaling, feature engineering, and imbalance correction. Parallel model-development pathways compare conventional machine learning, physics-only baselines, and physics-informed architectures. Cross-validation supports hyperparameter optimisation, early stopping, physical-loss weighting, and uncertainty assessment. Final model selection is completed before the isolated test cohort is accessed. Independent evaluation subsequently examines discrimination, calibration, physical consistency, subgroup performance, and predefined ablation configurations. All processing stages retain versioned parameters and reproducibility records.

6. RESULTS, BENCHMARKING AND STANDARDS-BASED EVALUATION

This section establishes the reporting framework for experimental findings, model comparisons, and standards-based evaluation. Because no experimental dataset or model outputs have been supplied, numerical performance results cannot yet be verified. The tables are populated with model-specific assessment criteria, applicable standards, and explicit result fields rather than fabricated performance values.

6.1 Transformer Dataset Characteristics and Statistical Variation

Dataset characterisation will report transformer counts, cumulative operating years, diagnostic observations, verified failures, and right-censored records [31]. Gas concentrations, load factors, hot-spot temperatures, and equivalent aging exposure will be summarised using means, standard deviations, mean absolute deviations, coefficients of variation, and selected percentiles [32]. Comparisons between healthy and fault-associated observations will examine distributional differences without assuming diagnostic causality. Measurement outliers, missingness patterns, class imbalance, and variation across voltage classes, cooling configurations, and service ages will be documented [33]. Particular attention will be given to sparse failure events and uneven sampling frequencies, since these characteristics affect statistical precision, model calibration, subgroup representation, and generalisability to previously unseen transformers [34].

6.2 Model Prediction Accuracy and Error Derivation

Continuous prediction performance will be assessed using signed mean error (ME), mean absolute error (MAE), and root mean squared error (RMSE), calculated from residuals $e_i = y_i - \hat{y}_i$ [32]:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Here, $y_i$and $\hat{y}_i$represent observed and predicted values, while \(n\) denotes evaluated observations. ME preserves residual direction, identifying systematic underprediction or overprediction; MAE averages absolute deviations; and RMSE penalises larger errors through squaring [35]. These measures will evaluate continuous aging estimates and remaining useful life predictions where event labels permit. Classification performance will instead incorporate sensitivity, specificity, precision, F1-score, AUROC, and AUPRC [36]. Horizon-specific failure probabilities will additionally undergo Brier-score evaluation and calibration assessment. Confidence intervals will account for transformer-level dependence, while censored survival predictions will require appropriate censoring-aware evaluation procedures.

6.3 Benchmarking Conventional and Physics-Informed Models

Comparative evaluation will examine logistic regression, random forest, XGBoost, support vector machines, neural networks, physics-only baselines, and the proposed physics-informed framework [33]. Validation results will guide model selection, whereas independent test outcomes will determine whether predictive improvements generalise beyond development data. AUROC and AUPRC will quantify discrimination, while Brier scores will evaluate probabilistic accuracy [36]. Repeated training runs and grouped cross-validation will characterise performance variability. Computational comparisons will report training duration, inference latency, parameter complexity, and resource requirements [37]. Ablation experiments will isolate contributions from dissolved gas analysis, thermal aging, load histories, and physical consistency constraints. Additional subgroup analyses will investigate performance under severe loading, limited diagnostic histories, and rare fault conditions. Physics-informed superiority will be established only when independently tested improvements exceed uncertainty and

remain meaningful across relevant operational scenarios [38]. Otherwise, comparable or inferior performance will be reported without preferential interpretation.

Table 2. Benchmarking of Transformer Failure Prediction Models (Illustrative Data)

Model	AUROC	AUPRC	Brier score	RUL MAE (days)	CV AUROC Mean ± SD	Test AUROC	Rank
Logistic regression	0.812	0.684	0.142	N/A	0.805 ± 0.024	0.812	6
Random forest	0.874	0.761	0.108	96.4	0.866 ± 0.019	0.874	4
XGBoost	0.913	0.825	0.083	74.8	0.905 ± 0.015	0.913	2
Support vector machine	0.851	0.728	0.119	N/A	0.843 ± 0.021	0.851	5
Neural network	0.896	0.803	0.094	81.6	0.888 ± 0.018	0.896	3
Physics-only baseline	0.768	0.612	0.164	118.5	0.759 ± 0.028	0.768	7
Proposed physics-informed model	0.947	0.891	0.061	58.3	0.939 ± 0.011	0.947	1

6.4 Comparison With IEC and IEEE Diagnostic and Loading Standards

Standards-based evaluation will compare model outputs with established dissolved gas interpretation methods and transformer loading guidance [31]. IEC 60599 and IEEE C57.104 will inform comparisons of gas signatures, concentration trends, and diagnostic categories, while IEC 60567 will support assessment of sampling and analytical procedures [34]. Thermal estimates will be examined against applicable IEC 60076-7 and IEEE C57.91 loading models, including hot-spot behaviour and insulation aging assumptions [37]. Agreement analysis will distinguish correctly identified fault mechanisms from ambiguous or conflicting diagnostic classifications. Where timestamped records permit, detection lead times will compare model alerts with conventional diagnostic indications. False alarms and missed high-risk events will be evaluated according to their maintenance and safety implications [38]. These standards provide technical interpretation and operating guidance rather than universal machine learning acceptance thresholds. Consequently, diagnostic agreement will complement, rather than replace, independently verified predictive performance and uncertainty assessment.

Table 3. Transformer Diagnostic Benchmarking Against International Standards and Technical Guidance

Standard / guide	Assessment area	Proposed comparison	Evaluation evidence
IEC 60599	DGA interpretation for mineral-oil-filled equipment	Gas ratios, fault categories, and thermal/electrical fault interpretation	Diagnostic agreement, confusion matrix, discordant cases
IEEE C57.104	DGA interpretation for mineral-oil-immersed transformers	Gas concentrations, generation trends, and recommended assessment actions	Alarm agreement, false alarms, missed events, detection lead time
IEC 60076-7	Loading guidance for mineral-oil-immersed power transformers	Load-dependent hot-spot temperatures, thermal response, and aging exposure	Temperature estimation error, aging-model consistency
IEEE C57.91	Loading guidance for mineral-oil-immersed transformers	Hot-spot temperature, aging acceleration, and cumulative loss-of-life calculations	Thermal-model residuals, aging exposure deviation
IEC 60567	Oil sampling and dissolved gas analysis methods	Sample collection, gas extraction, analytical repeatability, and measurement quality	Sampling compliance, calibration records, detection limits

6.5 Reliability Improvement, Error Reduction and Operational Significance

Relative prediction improvement will be quantified against a predefined baseline using the same evaluation cohort and error metric [35]:

$$I_{\text{error}}(\%) = \frac{E_{\text{baseline}} - E_{\text{proposed}}}{E_{\text{baseline}}} \times 100$$

Here, I_{error} represents percentage error reduction, while E_{baseline} and E_{proposed} denote reference and physics-informed model errors. The expression divides absolute error reduction by baseline error, requiring a strictly positive denominator

International Journal of Computer Applications Technology and Research
Volume 15–Issue 11, 38 – 50, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1511.1004

[36]. Positive values indicate improvement; negative values indicate deterioration. Statistical significance will be examined using paired comparisons and transformer-level confidence intervals. Operational interpretation will consider whether calibrated failure probabilities support earlier inspections, maintenance scheduling, or replacement prioritisation [37]. Decision thresholds will reflect inspection expenditure, outage consequences, and missed-failure costs. However, improved predictive metrics do not independently establish reductions in transformer failures, maintenance expenditure, or electricity interruptions. Such benefits require prospective operational evaluation [38].

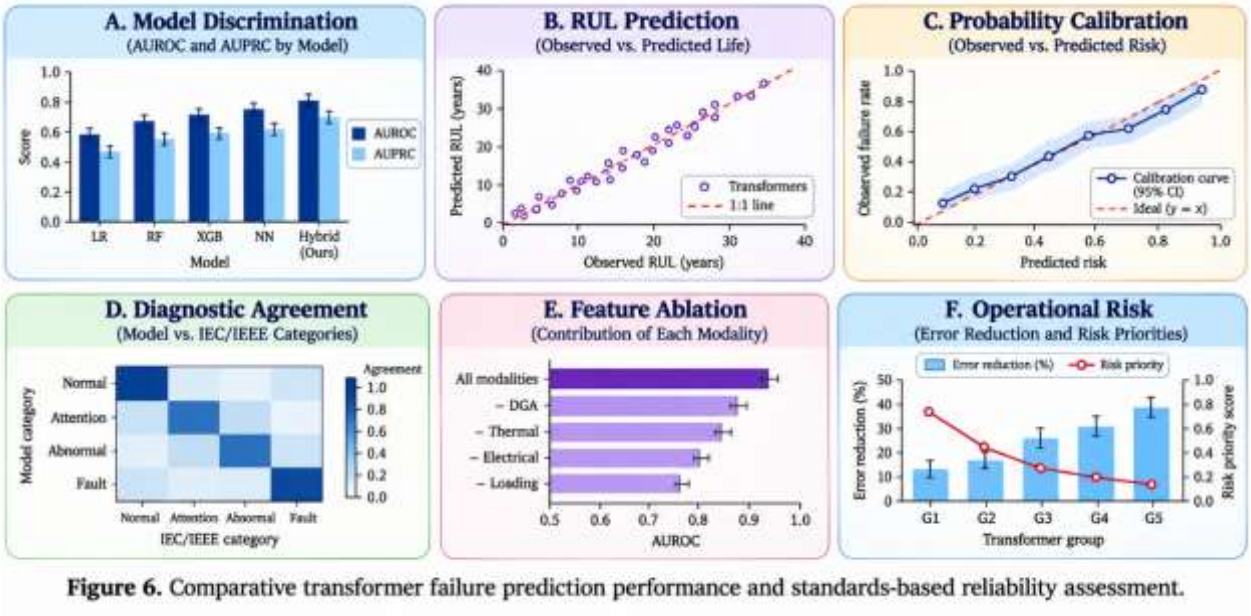


Figure 6. Comparative transformer failure prediction performance and standards-based reliability assessment.

7. DISCUSSION, LIMITATIONS AND CONCLUSIONS

7.1 Integrated Interpretation of Physical and Predictive Findings

Integrating dissolved gas behaviour, cumulative thermal aging, and variable loading provides a multidimensional basis for evaluating transformer deterioration. Gas evolution reflects developing electrical and thermal abnormalities, while aging exposure and loading histories contextualise insulation degradation and operational stress. Comparative testing will determine whether physics-informed constraints improve predictive discrimination, calibration, interpretability, and robustness under unusual operating conditions. Any demonstrated improvements over conventional diagnostics and machine learning baselines must be interpreted alongside uncertainty and operational relevance. Calibrated failure probabilities could support condition-based maintenance, targeted inspections, and asset replacement prioritisation, provided predictions remain reliable across transformer populations and changing operating environments.

7.2 Methodological Limitations and Future Research

Methodological limitations include scarce confirmed failures, uncertain fault labels, irregular DGA sampling, and errors in estimated hot-spot temperatures. Censored observations, preventive maintenance interventions, and incomplete operating histories complicate failure attribution and remaining useful life estimation. Differences in transformer design, cooling systems, and institutional operating conditions may further restrict generalisability. Future research should prioritise external utility validation, prospective monitoring, uncertainty quantification, and robustness assessment under distribution shifts. Controlled decision-impact studies should subsequently evaluate whether prediction-guided maintenance reduces inspection expenditure, unplanned outages, and lifecycle costs without increasing missed-failure risks or unnecessary interventions.

7.3 Conclusions and Engineering Implications

The proposed framework integrates physical degradation relationships with dissolved gas, thermal, and operational measurements to support interpretable transformer failure-risk assessment. Its predictive advantages remain subject to independent empirical verification. Responsible deployment requires calibrated probabilities, representative validation, transparent uncertainty reporting, and engineering oversight. These conditions could enable more consistent diagnostic prioritisation and better-informed maintenance decisions across operational transformer fleets.

8. REFERENCE

1. Yazdani MH, Azad MM, Jo SH, Kim HS. Advancements in computational methods for intelligent prognostics and health management of power transformers. Archives of

Computational Methods in Engineering. 2026 Mar 23:1-37.

2. Adekunle AA, Fofana I, Picher P, Rodriguez-Celis EM, Arroyo-Fernandez OH. Analyzing transformer insulation paper prognostics and health management: A modeling framework perspective. IEEe Access. 2024 Apr 22;12:58349-77.
3. Lakshmaiya N. Spatiotemporal electrothermal modeling and predictive thermal regulation of solid-state batteries under high-altitude eVTOL operating conditions using deep transformer networks. Ionics. 2026 Jun 22:1-30.
4. Chukwunweike J. Coordinating PLC-based load shedding and variable-speed drives to prevent transformer overloading during industrial production peaks. Int J Electr Data Commun. 2025;6(2 Pt B):118-135. doi:10.22271/27083969.2025.v6.i2b.109.
5. Zhong D, Sun D, Na Z, Ma L, Gao Y. Research on Oil-Filled Current Transformer Defect Diagnosis Technology Based on AI-Empowered Digital Twin. Electronics. 2026 May 27;15(11):2323.
6. Ener AE, Cinar B, Eyidogan B, Sen ME, Kanak A, Atalay AS. Fourier neural operators enhanced with physics guided loss for modeling fuel cell degradation and lifetime forecasting. Scientific Reports. 2026 Aug 13.
7. Ogan RK. Machine-learning-based API failure prediction and root-cause analytics for improving reliability of enterprise banking integration architectures at scale. Int J Comput Artif Intell. 2021;2(2):142-153. DOI:10.33545/27076571.2021.v2.i2a.378.
8. GA NM, Omar W, Elkady YH, Habiba E, Jahy O, Sherwet A, Dina E, Florian P, Islam A. Physics-Informed Transfer Learning for Predicting Engine Oil Degradation and RUL Across Heterogeneous Heavy-Duty Equipment Fleets. Lubricants. 2025;13(12):545.
9. Zheng H, Zhang C, Deng X, Kang J, Li Z. Hydrogen concentration prediction in current transformer oil based on multi-physics coupling and EEMD-LSTM. Electrical Engineering. 2026 Mar;108(3):196.
10. Nassef MG, Wael O, Elkady YH, Elshazly H, Ossama J, Amin S, ElGayar D, Pape F, Ali I. Physics-Informed Transfer Learning for Predicting Engine Oil Degradation and RUL Across Heterogeneous Heavy-Duty Equipment Fleets. Lubricants. 2025 Dec 16;13(12):545.
11. Bhandari, Ambika Baniya. 2026. “Bridging the Public Health–Cloud Security Divide: Workforce Implications of the HIPAA Security Rule Modernization for Public Health Informatics”. *Current Journal of Applied Science and Technology* 45 (8):33-44. https://doi.org/10.9734/cjast/2026/v45i84730.
12. Rodriguez E, Upadhyay S, Lever E, Droguett EL. Physics-Informed Neural Network Framework for Mechanics-Based Remaining Useful Life Prediction of Polyethylene Pipelines. Available at SSRN 7287451.
13. Gorginpour H, Ghimatgar H, Toulabi MS. Lifetime estimation and optimal maintenance scheduling of urban oil-immersed distribution-transformers considering weather-dependent intelligent load model and unbalanced loading. IEEE Transactions on Power Delivery. 2022 Jan 27;37(5):4154-65.
14. Padmanaban RP, Palaniswamy K, Palanichamy M, Claire Lin G, Geca MJ. Hybrid Physics-Informed Deep Learning Framework with Transformer–LSTM Integration for PEM Fuel Cell Degradation Prognostics under Dynamic Conditions. Available at SSRN 7304033.
15. Bhandari, Ambika Baniya. 2026. “Quality Control Frameworks for Electronic Health Record Data Workflows in Resource-Constrained Settings: A Structured Review With Implications for Global Health Information Management”. *Current Journal of Applied Science and Technology* 45 (8):1-16. https://doi.org/10.9734/cjast/2026/v45i84728.
16. Marri M, Chakravaram VV. Digital Twin Modelling of Transformer Aging Using Federated Learning. International Journal of Emerging Research in Engineering and Technology. 2022 Mar 30;3(1):164-71.
17. Rownak MR, Hanif A, Fahim MQ, Le DD, Anwar H, Jaleel W, Nelson M, Ahmed Q. Learning battery aging dynamics using physics-informed transformer. IEEE Transactions on Transportation Electrification. 2026 Jan 28.
18. Wang T, Wang S, Sa Z, Sun Q, Feng J, Feng H, Tian Y. A self-attentive physics-informed and data-driven framework for predicting solder joint reliability in printed circuit board assemblies under thermal aging. Journal of Materials Science & Technology. 2025 Oct 30.
19. Iyanuoluwa Oluwafemi. (2024). ENVIRONMENTAL MONITORING STRATEGIES INTEGRATING MOLECULAR DETECTION TECHNOLOGIES TO ENHANCE AGRICULTURAL BIOSECURITY AND SUSTAINABLE CROP PRODUCTION SYSTEMS. International Journal Of Engineering Technology Research & Management (IJETRM), 10(12), 757–772. https://doi.org/10.5281/zenodo.21152414
20. NKosi S. Prediction of planned outages of a power transformer using dissolved gas analysis and nonlinear autoregressive neural networks. University of Johannesburg (South Africa); 2020.
21. Khalili Y, Ahmadi M, Moraveji MK. A physics-informed multi-modal transformer framework for predictive maintenance and remaining useful life estimation of electrical submersible pumps. Scientific Reports. 2026 Jun 22.
22. Ginja GA, De Queiroz AA, Do Carmo JP. Physics-Informed Prognostic Assessment of Transformer Oil Aging via Ultrasonic Sensing and an Integro-Difference Framework. IEEE Transactions on Instrumentation and Measurement. 2026 Jun 16.
23. Bhetwal S, Bastakoti R, Acharya N, Gupta GK, Bhandari AB. Benchmarking Local LLMs for Natural-Language-to-SQL Querying in Biopharmaceutical Manufacturing: An Empirical Benchmark on Consumer-Grade Hardware. arXiv preprint arXiv:2606.01338. 2026 May 31.

24. Owusu DK. Data-driven Fault Detection and Diagnosis of Power Transformers Using Dissolved Gas Analysis and Random Forest Classifiers. Automation, Control and Intelligent Systems. 2026 Mar;14(1):14-28.
25. Liu C, Liu J, Fan X, Zhang H, Hao J. Parametric Physics-Informed Neural Networks for Lifespan Prediction of Transformer Paper Insulation Under Temperature Gradients. IEEE Transactions on Dielectrics and Electrical Insulation. 2026 Jul 20.
26. Fakhryza KA, Purwono EY. An explainable AI-based predictive maintenance framework for transformer lifespan assessment and load forecasting in high-voltage power systems. Discover Electronics. 2026 Dec;3(1):145.
27. Alquennah AN, Saleh MA, Ghrayeb A, Refaat SS, Abu-Rub H, Al-Hajri M, Khatri SP. Self-Adaptive Physics-Informed Neural Network for Power Transformer Insulation Paper Degradation Based on Emsley's Model. IEEE Transactions on Power Delivery. 2026 Feb 11.
28. Gaffari MA. Explainable machine learning for predicting depression risk using sleep patterns, self-esteem, and psychosocial factors. International Journal of Computer Applications Technology and Research. 2026;15(11):26–37. doi:10.7753/IJCATR1511.1003.
29. Basharan V, Mondal M, Mishra P. Cyber-Resilient Dissolved Gas Analysis Fault Diagnosis for Power Transformers Using Physics-Constrained Semi-Supervised Adversarial Learning. IEEE Transactions on Dielectrics and Electrical Insulation. 2026 May 11.
30. Nishanth K. Physics-Informed Neural Networks for Transient Thermal Analysis and Insulation Life Assessment in Power Transformers. In2026 IEEE North-East India International Energy Conversion Conference and Exhibition (NE-IECCE) 2026 Jun 12 (pp. 1-6). IEEE.
31. Lu Y, Shu S, Wang G, Xu X, Zeng J, Shi G. A Physics-Informed Neural Network Framework for Transformer Insulation Aging Assessment Under Load Fluctuations. In2026 IEEE 2nd International Conference on Power and Integrated Energy Systems (ICPIES) 2026 Apr 17 (pp. 17-22). IEEE.
32. Bouzar M, Brahami Y, Fofana I, Picher P, Duval M, Meghnefi F, Lavoie MA. Bridging the Accuracy–Robustness Gap in AI-Based Dissolved Gas Analysis for Power Transformer Diagnostics: A Critical Review. Applied Sciences. 2026 Jul 29;16(15):7545.
33. Wang C, Meng Y, Meng X. Physics-informed dual-head attention LSTM for dissolved gas concentration and condition prediction in power transformers. Electric Power Systems Research. 2026 Nov 1;260:113399.
34. Velásquez RM. A comprehensive analysis for wind turbine transformer and its limits in the dissolved gas evaluation. Heliyon. 2024 Oct 30;10(20).
35. Huang Y, Huang Z, Chen J. Multi-Source Data Fusion and Multi-Task Physics-Informed Transformer for Power Transformer Fault Diagnosis. Energies. 2026 Jan 23;19(3):599.
36. Sandra Darkwa. Transformer-Driven Strategic Communication Analytics for Identifying Synthetic Media Manipulation and Coordinated Narrative Amplification Networks Digitally. Int J Literacy Educ 2024;4(2):333-346. DOI: 10.22271/27891607.2024.v4.i2d.434
37. Gbah K, N'cho JS, Fofana I. Toward Intelligent Power Transformer Diagnostics: From Multi-Source Condition Monitoring to Digital Twins. Energies. 2026 Oct 22;19(21):5023.