machines

MDPI

Review

Remaining Useful Life Prediction for Engineering Applications: A Critical Review of Methodologies, Capability Gaps, and System-Level Integration

Lin Wang [1,2,3], Yongmin Yang [1,2,*], Xu Luo [1,2] and Mengqiao Chen [1,2]

[1] College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China; wanglin8931@163.com (L.W.); luoxu2009@163.com (X.L.); chenmengqiao21@nudt.edu.cn (M.C.)
[2] National Key Laboratory of Equipment State Sensing and Smart Support, Changsha 410073, China
[3] Equipment Support and Remanufacturing Department, Army Arms University of PLA, Beijing 100072, China
* Correspondence: yangyongmin@163.com

Abstract

As one of the core technologies of predictive maintenance, the development of remaining useful life (RUL) prediction is gradually transitioning from early single-mechanism modeling to a new phase characterized by the deep integration of physics-based approaches, data-driven methods, and uncertainty awareness. This paper first analyzes the fundamental challenges facing this development, such as multi-stress coupling, sensor degradation, and non-stationary noise. By comparing the core advantages and applicability boundaries of statistical models, data-driven models, and hybrid models, it constructs a capability map for RUL prediction. It further points out that current RUL prediction still faces critical capability gaps in areas such as physical consistency and uncertainty decoupling. Finally, the paper distills a new paradigm for engineering implementation, including mechanism-guided neural architecture design and digital twin-driven online parameter adaptation. The research indicates that future RUL prediction studies must transcend the competition over accuracy metrics and shift toward the coordinated development of robustness, interpretability, and decision adaptability—a trinity guided by the principles of "trustworthy AI."

Keywords: remaining useful life prediction; predictive maintenance; hybrid modeling; uncertainty quantification; digital twin; edge intelligence

1. Introduction

Remaining Useful Life (RUL) prediction, as a core component of Predictive Maintenance (PdM), plays a crucial role in the operation and maintenance of modern industrial equipment. By accurately predicting the remaining lifespan of critical equipment, enterprises can transition from traditional scheduled maintenance or reactive repair models to proactive, condition-based maintenance strategies, thereby significantly reducing unplanned downtime and lowering maintenance costs. In high-safety-requirement sectors such as aviation, wind power, power electronics, and electric vehicles, RUL prediction not only enhances system reliability and availability but also effectively prevents accidents and operational disruptions, playing an irreplaceable role in ensuring personnel safety and asset integrity. Particularly in the field of lithium-ion battery management, precise RUL prediction is crucial for ensuring safe battery operation and preventing internal and external failures, thereby avoiding catastrophic events caused by complex operational characteristics and dynamic changes in model parameters during the aging process [1].

Academic Editor: Jihong Yan

Received: 22 May 2026
Revised: 20 June 2026
Accepted: 24 June 2026
Published: 26 June 2026

 https://doi.org/10.3390/machines14070724

With the deepening development of Industry 4.0 and smart manufacturing, equipment complexity continues to rise, making traditional maintenance strategies increasingly unable to meet modern industry's dual demands for high reliability and low-cost operations and maintenance. Planned maintenance often results in excessive costs due to neglect of the system's actual condition, while reactive maintenance may trigger unexpected failures due to a lack of foresight. Furthermore, when developing RUL prediction models for critical systems (such as aircraft engines), there are two major challenges: the scarcity of "run-to-failure" data samples and the reluctance of airlines to share data due to privacy concerns. These practical issues have driven the rapid advancement of RUL prediction theories and methods, prompting researchers to explore more accurate, reliable, and practical prediction techniques.

Existing RUL prediction methodologies can be broadly categorized into model-based approaches, data-driven techniques, and hybrid frameworks [2]. Model-based methods rely on physical modeling and typically exhibit good predictive accuracy when degradation parameters are estimated accurately, but they are often limited by uncertainties in model parameters. For example, as highlighted in recent reviews [3], in the Paris model for fatigue crack propagation in aluminum alloys, a variation in the exponent m between 3.6 and 4.2 can result in a range of up to 500% in life predictions. Data-driven methods, on the other hand, utilize machine learning and artificial intelligence techniques to process high-dimensional data. While they offer advantages in terms of generality and flexibility, they often face the challenge of a lack of physical interpretability. To provide a clear and structured overview of these methodologies, Table 1 summarizes the primary RUL prediction approaches, their typical engineering applications, and their respective benefits and limitations.

Table 1. Summary of RUL prediction approaches, applications, benefits, and limitations.

Approach Category	Specific Methods	Typical Engineering Applications	Key Benefits	Primary Limitations
Model-based	Paris Law; Archard's wear model; Equivalent circuit models	Fatigue crack propagation; Bearing wear; Battery degradation	High interpretability; Strong physical meaning; Good extrapolation capabilities.	Hard to derive exact physical equations for complex systems; Sensitive to parameter uncertainties.
Statistical	ARIMA; Kalman Filter; Particle Filter; Hidden Markov Models	Tool wear; Simple bearing degradation; Gearbox health monitoring	Mathematically rigorous; Handles uncertainty well; Low computational cost for inference.	Struggles with highly non-linear and complex degradation patterns; Requires strict statistical assumptions.
Data-driven	CNN; LSTM; Transformer; GAN; Random Forest	Turbofan engines (C-MAPSS); Complex rotating machinery; Power electronics	Excellent at capturing non-linear features; High predictive accuracy; Automated feature extraction.	"Black-box" nature lacking interpretability; Requires massive labeled datasets; Poor generalization to unseen conditions.
Hybrid/Physics-Informed	PINNs; Physics-guided RNNs; Digital Twin models	Wind turbine gearboxes; Aerospace structures; Li-ion batteries	Combines physical interpretability with data-driven accuracy; Better generalization; Requires less training data.	Complex model design; High training computational cost; Balancing physics and data loss functions is challenging.

In recent years, the rapid development of deep learning technologies has brought new breakthroughs to RUL prediction. Researchers have proposed various innovative frameworks. For example, the CRULP framework, which incorporates uncertainty quantification into the prediction process, transforms point predictions into interval estimates by integrating deep learning models with uncertainty-aware conformal quantile regression, thereby reducing the prediction interval width by 25.8–31.37% on an aircraft engine dataset [4]. To address the issue of error accumulation in iterative forecasting, the Attention-HTA Net was designed to internally handle the iterative forecasting process while significantly reducing the number of iterations required for forecasting through a hierarchical temporal aggregation strategy [5]. To address data sparsity and privacy concerns, a collaborative federated learning framework was applied to train RUL prediction models across airlines, resulting in a significant improvement in prediction accuracy for five airlines [6].

In terms of model architecture, the Multi-Dilation Causal Convolution Block-enhanced Squeeze-Excite Transformer (MDCCB-ISEformer) constructs temporal samples using a multi-layer sliding window approach and fuses window-level features to effectively capture cross-step dependencies. On the NASA C-MAPSS dataset, it reduced the RMSE by 1.5–11.5% compared to the best baseline model [7]. A hybrid framework combining CSI-EMD with the Dual-Channel Multi-Layer Feature Fusion Network (DCM-FFN) successfully mitigates the endpoint effects associated with traditional empirical mode decomposition, reducing RMSE by 15.33% on the NASA Aviation Turbine Engine dataset [8]. A GPT-based RUL prediction model (DMLP-GPT) leverages the few-shot learning capabilities of pre-trained language models. Using only 15% of the training data, it demonstrated that removing the GPT module results in a significant 24.6% increase in RMSE on complex datasets [9]. Deep reinforcement learning has been innovatively applied to the field of RUL estimation. By constructing an RUL estimation Markov Decision Process (RULeMDP) model, it achieved an improvement in the RMSE metric of 7.81–14.4% on the C-MAPSS dataset [10]. Additionally, the CONELPABO framework accelerates model training via parallel Bayesian optimization, thereby revealing a counter-intuitive phenomenon: architecturally streamlined combinations (e.g., dual-CNN cascades) can outperform numerous complex counterparts on long-horizon temporal series [11].

At the application level, researchers have developed specialized RUL prediction solutions for various domains. For wind turbines, a review study systematically analyzed RUL prediction techniques for key components (blades, gearboxes, bearings, and generators), with a particular focus on specific challenges such as uncertainty management and handling of variable operating conditions [12]. In the field of power electronic converters, FASSI et al. proposed Physical-Informed Machine Learning (PIML), which integrates physical knowledge into machine learning architectures, thereby addressing key challenges such as data resources, physical consistency, and generalization ability [13]. For tool wear prediction in high-speed CNC milling machines, a hybrid deep learning framework combining the advantages of LSTM and autoencoders achieved a prediction accuracy of nearly 98% on the PHM10 dataset [14]. An unsupervised framework integrating autoencoders, Gaussian mixture models, and self-attention LSTMs achieved effective early fault warning on real industrial data, with an R^2 as high as 0.986 [15].

However, existing research still faces numerous challenges. Many methods struggle to effectively capture degradation patterns due to the use of isolated time-series sampling strategies, and they rely solely on information from the current window for prediction, neglecting valuable insights from historical data. Regarding cross-domain adaptation, benchmark studies have found that many deep unsupervised domain adaptation methods may be unreliable under fair evaluation and exhibit significant performance variations across different datasets [16]. Research integrating predictive maintenance with calendar-based preventive maintenance and fleet operation scheduling remains insufficient, hindering the creation of commercial value [17]. Furthermore, the performance of existing data-driven algorithms in electric vehicle applications remains unsatisfactory, and health estimation methods must be dynamic to adapt to varying operating conditions [18].

While several review papers have explored RUL prediction techniques, a critical gap remains in systematically comparing these reviews and explicitly addressing the transition from algorithmic innovation to system-level engineering integration. Table 2 provides a comparative analysis of existing review studies on RUL prediction, highlighting their scope, key findings, and limitations.

Table 2. Comparative analysis of existing review studies on RUL prediction.

Reference	Scope/Objectives	Applications Covered	Conclusions	Coverage Period	Limitations of the Review
Reza et al. [1]	RUL mechanisms, network configs, and factors for EV batteries.	EV Li-ion batteries	ML/DL network configs; impact of data quality and aging mechanisms.	Up to 2023	Single-domain (EV); lacks multi-physics and system-level integration.
Cuesta et al. [12]	PHM techniques and challenges in wind energy.	Wind turbines (blades, gearboxes)	Variable operating conditions; data-driven and hybrid model trends	Up to 2024	Single-domain (Wind); limited PINNs and Digital Twin exploration.
Fassi et al. [13]	Physics-Informed ML (PIML) for power converters.	Power electronic converters	PIML addresses data scarcity and physical inconsistency effectively.	Up to 2023	Single-domain (Power electronics); lacks edge deployment and closed-loop decisions
Das and Kumar [18]	ML for EV battery capacity degradation and health.	EV battery SOH/RUL	ML excels in estimation but relies heavily on high-quality data.	Up to 2022	Focuses on traditional ML; lacks advanced DL, uncertainty and sensor degradation.
Patrizi et al. [19]	Degradation models and RUL for Li-ion battery testing.	Li-ion battery testing and modeling	Importance of empirical/electrochemical models in test design.	Up to 2023	Focuses on testing phase; lacks real-world multi-stress and system-level frameworks.
This Work	Methodologies, capability gaps, and system-level integration.	Cross-domain (Bearings, Batteries, Aero-engines, etc.)	Evolution to Digital Twins and closed-loop maintenance; computational and sensor analysis	2022–2026	N/A (Addresses gaps via cross-domain system-level integration and trustworthy AI).

To address the limitations identified in previous reviews and ensure a comprehensive, unbiased synthesis of the most recent literature, this study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A systematic search was performed across major databases (e.g., IEEE Xplore, Scopus, Web of Science, and ScienceDirect) using targeted keywords related to RUL prediction, physics-informed models, and system-level integration. After rigorous screening based on inclusion and exclusion criteria (focusing on peer-reviewed journal articles and state-of-the-art methodological innovations), core studies published in the recent five years (2022–2026) were selected for in-depth analysis. Foundational theories and background references cited in the introduction are not restricted by this timeframe to ensure theoretical completeness.

Based on this systematic review of recent advancements, the novel contributions of this work are fourfold: (1) It provides a critical comparative analysis of the predictive accuracy, computational efficiency, and applicability boundaries among statistical, data-driven, and physics-informed models based on the latest literature. (2) It systematically diagnoses the impact of real-world engineering constraints—such as sensor degradation and multi-stress coupling—on RUL prediction reliability. (3) It distills recent evolutionary trends and a future roadmap for algorithmic innovation specifically geared toward practical engineering implementation (e.g., Digital Twins and Foundation Models). (4) It constructs a comprehensive system-level integration framework that bridges the gap between point-in-time RUL prediction and closed-loop maintenance decision-making.

This review aims to systematically summarize the latest research progress in RUL prediction theories and methods, conduct an in-depth analysis of the strengths and weaknesses of different technical approaches, and explore future research directions. The structure of this paper is as follows: Section 2 summarizes the practical constraints faced in RUL prediction and their fundamental challenges to methodology; Section 3 outlines the capability maps and applicability boundaries of statistical, data-driven, and hybrid methods; Section 4 diagnoses the systemic limitations of existing methods in terms of robustness, interpretability, and uncertainty quantification; Section 5 distills an evolutionary path for algorithmic innovation geared toward engineering implementation; Section 6 constructs a system integration framework spanning from point-in-time prediction to closed-loop maintenance decision-making; and Section 7 summarizes the paper and looks ahead to future development directions.

2. Real-World Constraints and Methodological Challenges Facing RUL Prediction

RUL prediction faces multiple practical constraints in real-world industrial settings, which pose fundamental challenges to methodological design. This section systematically summarizes these constraints across five key areas—multi-stress coupling, sensor performance degradation, non-stationary noise and feature drift, resource limitations of edge devices, small-sample and cross-domain data scarcity, as well as sensor instrumentation and feature measurement considerations—and analyzes their profound impact on recent prediction methods.

2.1. Multi-Stress Coupling Effects

Multi-stress coupling effects significantly increase the complexity of degradation mechanisms. For example, in accelerated aging tests of optoelectronic couplers, the synergistic interaction between thermal stress (maintained at 130 °C) and electrical load (5 V input) leads to nonlinear degradation behavior, necessitating the use of physical models for extrapolation to address the difficulty in obtaining full-lifecycle data. Similarly, in battery systems, initial degradation model coefficients undergo drastic changes during current load profile transitions (e.g., from constant current to UDDS dynamic conditions); failure to track such changes will result in prediction failure [20]. In Industry 4.0 environments, changes in flexible manufacturing tasks trigger shifts in data distribution (i.e., changes in data mode). Ignoring this shift can cause a significant decline in the model's prediction accuracy in critical areas [21]. Furthermore, the degradation behavior of proton exchange membrane fuel cells is dynamically influenced by multiple operating variables; univariate models suffer from significantly reduced reliability due to the neglect of multivariate coupling [22].

2.2. Sensor Performance Degradation

Intrinsic sensor performance degradation directly leads to measurement errors. Recent research indicates that sensors degrade over time due to prolonged use; this degradation process should be modeled using the Wiener process and the Gamma process, rather than the traditional assumption of constant-mean Gaussian noise [23]. In optocoupler testing, transient voltage recovery effects have been observed. This phenomenon is attributed to the desorption of deposited ionic contaminants and requires the removal of non-monotonic segments via second-derivative detection to ensure physical consistency [24]. Due to sampling frequency limitations (e.g., approximately 5 kHz), low-cost sensors cannot cover the high-frequency natural frequencies (2–6 kHz) associated with the second stage of bearing degradation, resulting in inherent limitations in the early detection of microcrack failures [25].

2.3. Non-Stationary Noise and Feature Drift

Non-stationary noise and feature drift caused by operating condition changes severely interfere with feature extraction. Condition monitoring data often contains noise due to external environmental disturbances, data transmission issues, and variations among multi-source sensors, which directly reduces model accuracy. The data distribution changes dynamically with the progression of degradation; there are significant differences between the distribution of data acquired in the early stages and that in the later stages of severe vibration, making it extremely challenging to extrapolate global data using only a partial subset. Existing domain-adaptive methods often fail to distinguish between different phases of the operating profile (e.g., rising, steady-state, and falling), leading to alignment errors caused by under- or over-representation of specific phases. Notably, the Pearson correlation coefficient between bearing vibration data and time ranges between -0.3 and 0.3, indicating weak correlation. This highlights the practical constraint that traditional linear degradation models struggle to accurately describe complex degradation relationships. Experimental quantification shows that after adding 50% Gaussian noise, the prediction accuracy of traditional models (GRU, LSTM) decreases by 10–12.5%, and the RMSE increases by 62–81% [26].

2.4. Low-Latency and Low-Power Requirements for Edge Devices

The low-latency and low-power requirements of edge devices limit the direct deployment of deep learning models. Existing deep learning models are computationally intensive and thus difficult to adapt to edge devices with limited memory and computing power; at the same time, traditional physical models are complex to calibrate and computationally expensive. To address this constraint, the LiRUL framework employs a hybrid feature engineering pipeline combining PCA with domain-aware variables (temperature, C-rate). This approach minimizes model complexity (only 0.15 million parameters, memory footprint $<$ 5 MB) while preserving interpretability and achieving inference times under 1 s [27]. Similarly, for low-cost sensors, researchers have proposed a hybrid feature extraction method combining Short-Time Fourier Transform (STFT) and envelope analysis, using frequency-selective filters to remove non-fault frequency noise, making it suitable for low-sampling-rate scenarios.

2.5. Small Sample Sizes and Cross-Domain Data Scarcity

Small sample sizes and cross-domain data scarcity constitute common bottlenecks in the field of RUL prediction. Obtaining complete "run-to-failure" data is extremely difficult because safety-critical systems (such as aircraft engines) typically undergo preventive maintenance before failures occur, resulting in most condition monitoring data lacking failure labels. Airlines are reluctant to share data due to privacy concerns and conflicts of interest, making it difficult to effectively integrate cross-domain data even after anonymization. In wind power generation, acquiring samples of early-stage gearbox bearing failures is costly, and labeled samples of the same failure type are extremely scarce, making deep learning models highly prone to overfitting under small-sample conditions. Acquiring full lifecycle data for industrial systems is time-consuming and costly. Particularly in the aviation and aerospace sectors, where systems are not permitted to operate until a catastrophic failure occurs, only partial degradation data (i.e., truncated data) can be obtained [2]. Significant variability exists in degradation patterns between individual units (unit-to-unit variability), and the influence of uncertain resources limits the applicability of models based on historical data. Experiments show that as the number of training samples decreases, the increases in RMSE and MAPE for baseline models (such as MGRU) are significantly higher than those for robust methods. Furthermore, under small-sample training conditions (with only 20–25% of the initial data), the RMSE increase for some models reaches as high as 18.7%.

2.6. Impact of Sensors and Instrumentation on RUL Prediction Methods

The choice of sensors and their instrumentation characteristics profoundly impacts the performance and applicability of different RUL prediction methodologies. For model-based and statistical approaches (e.g., Kalman filters, Particle filters), sensor accuracy and measurement noise covariance directly dictate the precision of state estimation; excessive sensor noise or unmodeled sensor degradation can lead to filter divergence and prediction failure [22]. Conversely, deep data-driven models are highly sensitive to sensor sampling rates and signal fidelity. For instance, high-frequency vibration sensors capture rich fault signatures essential for Convolutional Neural Networks (CNNs), but they introduce massive data volumes that strain edge-computing resources, necessitating a trade-off between signal resolution and hardware limitations [24]. Furthermore, intrinsic sensor degradation (e.g., piezoelectric sensor sensitivity drift in high-temperature or harsh environments) introduces artificial feature shifts. If not properly decoupled or modeled, data-driven models may misinterpret these sensor-level drifts as system-level degradation, leading to false RUL predictions [28]. Therefore, instrumentation design must carefully balance sensor accuracy, sampling frequency, and environmental robustness with the specific computational and input requirements of the chosen predictive algorithm.

2.7. Input Feature Types and Measurement Considerations

The selection and reliable measurement of input features dictate the upper bound of RUL prediction accuracy. Common input features across engineering systems include: (1) vibration signals, widely used for rotating machinery (bearings, gearboxes), measured via accelerometers requiring high sampling rates (often >10 kHz) to capture high-frequency fault signatures; (2) thermal and electrical signals (voltage, current, temperature), crucial for Li-ion batteries and power electronics, measured via thermocouples and Hall-effect sensors [19]; and (3) Acoustic Emissions (AE), utilized for early-stage micro-crack detection, requiring high-sensitivity piezoelectric sensors. The reliability of these features is heavily dependent on sensor accuracy and rigorous noise control. In harsh industrial environments, electromagnetic interference and mechanical background noise severely degrade the Signal-to-Noise Ratio (SNR). To ensure high-quality inputs, hardware-level noise control (e.g., shielded cables, anti-aliasing filters) and software-level signal processing (e.g., wavelet denoising, empirical mode decomposition) are mandatory [8,26]. A low SNR not only obscures early degradation features but also forces data-driven models to learn noise patterns, resulting in catastrophic prediction failures under varying operating conditions.

2.8. Fundamental Challenges to Methodology

The aforementioned practical constraints pose multifaceted and fundamental challenges to RUL prediction methodologies.

First, it is necessary to integrate physical knowledge to enhance model robustness. One effective approach is to introduce physical constraint loss regularization, such as simultaneously incorporating monotonicity constraints L_{mono} and smoothness constraints L_{smooth}, which enforce that the predicted trajectories conform to the monotonically decreasing characteristics of electrochemical aging, thereby effectively suppressing prediction oscillations caused by non-stationary noise. Another approach is to combine single-particle models (SPMs) to extract physical degradation parameters (such as LAMPE, LAMNE, and LLI), thereby replacing traditional capacity measurement methods and more accurately quantifying degradation mechanisms.

Second, feature engineering requires innovative design. For example, the Noise-Conditional Convolutional Denoiser Autoencoder (NC-CDAE) adaptively suppresses heterogeneous noise types through an auxiliary noise estimation branch and Feature-Level Linear Modulation (FiLM) mechanism, avoiding issues of over-smoothing or insufficient denoising. The Dynamic Weighted Residual Network (DWRNet) utilizes the Optuna algorithm to dynamically adjust weighting factors, adaptively enhancing important features under low-sample-size conditions.

Third, data scarcity requires new learning strategies. In federated learning, decentralized validation procedures and novel parameter aggregation methods (such as full-sample/random validation strategies) can be adopted to provide accurate predictions even when the proportion of noisy data is high ($\alpha \geq 0.7$). The TimeGAN (Time Series Generative Adversarial Network) can perform segmented augmentation on data with varying degradation rates to address issues of insufficient training data and uneven distribution.

Fourth, cross-domain problems require refined domain adaptation. The OPS-DANN method sets up independent domain discriminators for each phase of the flight profile, achieving phase-level alignment of edge distributions, and achieves a 9% improvement in RMSE for short-to-long-range flight domain shift tasks. Training models separately after identifying data modalities can reduce the prediction MAE for minority-class data modalities by 60.50%.

Fifth, the hyperparameter tuning bottleneck requires efficient optimization algorithms. An algorithm based on Harris-Horn-O'Meara (HHO) optimization significantly improved the prediction accuracy of Random Forest and LightGBM (R^2 reached 0.971, MAPE as low as 0.23%) [29]. However, in the prediction of lithium-ion batteries for electric vehicles, hybrid methods face practical challenges due to high computational complexity and time-consuming hyperparameter tuning [1].

Finally, for constraints where tool degradation is difficult to measure directly, indirect evaluation strategies can be employed. For example, by using the wear state of the workpiece surface as a surrogate indicator, a hybrid data-driven framework (WearNet classification + BLSTM regression) kept prediction errors within ±5% [30], validating the method's effectiveness under complex contact-sliding conditions.

In summary, the practical constraints on RUL prediction span multiple dimensions, including data acquisition, signal quality, computational resources, and individual variability. These intertwined constraints pose significant challenges to existing methods in terms of robustness, interpretability, computational efficiency, and generalizability. Figure 1 summarizes the relationship between the identified constraints and methodological challenges. As shown in the figure, each constraint leads to one or more specific technical bottlenecks.

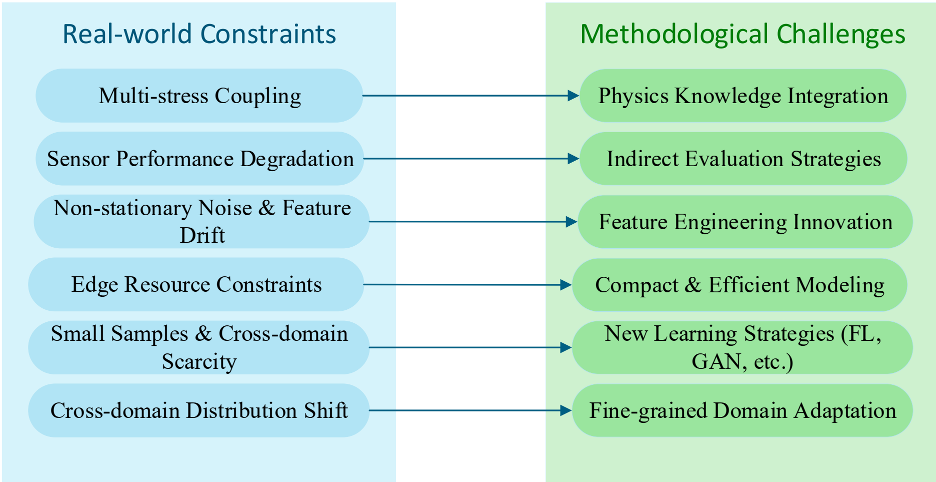


Figure 1. Mapping of real-world constraints and methodological challenges in RUL prediction.

To provide a consolidated overview of how recent studies have tackled these multi-faceted constraints, Table 3 summarizes the key literature discussed in this section, detailing the specific applications, the primary constraints addressed, and the corresponding methodological outcomes. As evidenced in the table, recent advancements have shifted from merely improving prediction accuracy under ideal conditions to developing robust, lightweight, and physically consistent frameworks capable of operating under severe real-world limitations, such as sensor degradation, edge-computing constraints, and indirect measurement scenarios.

Table 3. Summary of real-world constraints, addressed methods, and key outcomes in recent RUL prediction studies.

Reference	Application	Constraint	Key Outcomes	Prediction Efficiency *
[20]	Battery systems	Current load profile transitions (Multi-stress)	Tracked drastic changes in model coefficients to adapt to dynamic conditions.	Prevented prediction failure during dynamic load shifts; enhanced model robustness.
[21]	Industry 4.0 manufacturing	Data distribution shifts (Flexible tasks)	Highlighted significant accuracy decline when ignoring data mode shifts in critical areas.	Significant accuracy drop quantified when ignoring operational shifts.
[23]	General sensors	Intrinsic sensor degradation over time	Modeled sensor degradation using Wiener/Gamma processes instead of constant-mean Gaussian noise.	Improved state estimation robustness and reduced filter divergence risks.
[25]	Bearings	Low-cost sensor sampling limits	Identified inherent limitations in early microcrack detection due to high-frequency coverage gaps.	Highlighted detection blind spots caused by hardware sampling constraints.
[26]	Bearings	Non-stationary noise (50% Gaussian added)	Quantified the severe impact of noise on traditional deep learning models (GRU, LSTM).	RMSE increased by 62–81%; Accuracy decreased by 10–12.5%.

Table 3. *Cont.*

Reference	Application	Constraint	Key Outcomes	Prediction Efficiency *
[27]	Edge devices (LiRUL)	Low-latency and low-power constraints	Developed a hybrid feature engineering pipeline combining PCA with domain-aware variables.	Inference time < 1 s; Memory footprint < 5 MB; Model size 0.15 M parameters.
[29]	Li-ion batteries	Hyperparameter tuning bottleneck	Applied Harris Hawk Optimization (HHO) to tune Random Forest and LightGBM models.	R^2 reached 0.971; MAPE reduced to 0.23%.
[30]	Tool wear	Indirect measurement constraints	Utilized workpiece surface wear state as a surrogate indicator for tool degradation.	Prediction errors strictly maintained within ±5%.

* The "Prediction Efficiency" column includes quantitative metrics (e.g., RMSE, R^2, MAPE, inference time) where available or qualitative impact assessments for studies focusing on constraint identification rather than direct predictive modeling.

The surveyed studies underscore a paradigm shift: the research community is increasingly prioritizing the robustness and deployability of RUL prediction systems under real-world imperfections, rather than solely optimizing accuracy under idealized laboratory conditions.

3. Methodology Capability Map: Core Advantages and Application Boundaries of Statistical, Data-Driven, and Hybrid Models

3.1. Capabilities and Applications of Statistical Models

Statistical models demonstrate unique value in RUL prediction for specific equipment types and operating conditions due to their core advantages, including physical interpretability, stability under small sample sizes, and clear parameter meanings. Figure 2 presents a hierarchical classification of RUL prediction methods.

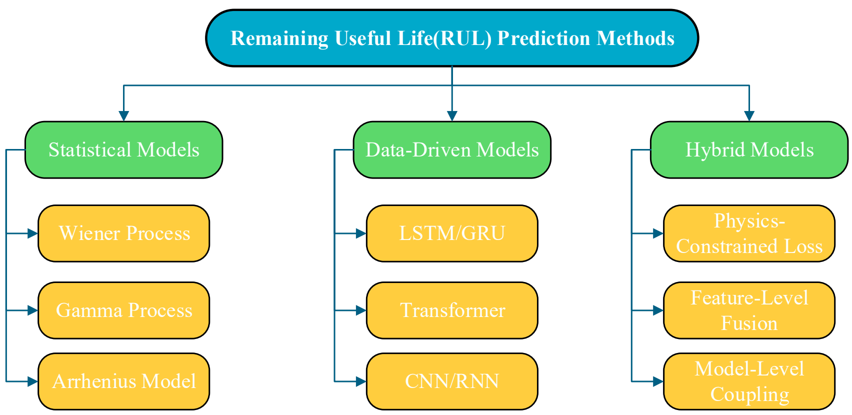


Figure 2. Classification System for RUL Prediction Methods.

In terms of physical interpretability, researchers introduced a novel PCA load interpretability method that maps the mathematically transformed principal components back to the real-world physical context. This method revealed the specific correlations between the RMS, mean, and frequency information of sensor readings and specific engine instability conditions. It also clarifies the specific physical significance of various statistical features in characterizing changes in core engine speed and pressure fluctuations, thereby significantly enhancing the interpretability of statistical features and the clarity of parameter meanings in data-driven models [31].

To address the issue of stability in small samples, statistical models employ innovative methods to ensure prediction reliability. For example, a dual-accelerated degradation test model based on the Wiener process was proposed, in which both the drift and diffusion parameters are influenced by the stress level. Generalized confidence intervals for model parameters were developed using the generalized pivot method. Monte Carlo simulations indicate that even with small sample sizes, the coverage of the proposed intervals remains very close to the nominal levels (90% and 95%) validated using accelerated degradation data from commercial white LEDs [32].

Regarding lithium-ion batteries, a cubic polynomial function is adopted as the nonlinear drift term of the Wiener process to capture the non-monotonic characteristics of capacity degradation. The proposed subjective parameter estimation method based on envelope extraction effectively avoids the overestimation of the diffusion coefficient, significantly improving prediction accuracy for small samples [33].

In terms of applicability, statistical models are particularly suitable for equipment types with relatively clear physical mechanisms and limited data. For equipment with non-monotonic degradation paths, models based on the Wiener process can provide accurate statistical inferences. Combined with online updates via Kalman filtering, this method is suitable for small-sample or single-device prediction scenarios affected by measurement noise.

3.2. Capabilities and Applications of Data-Driven Models

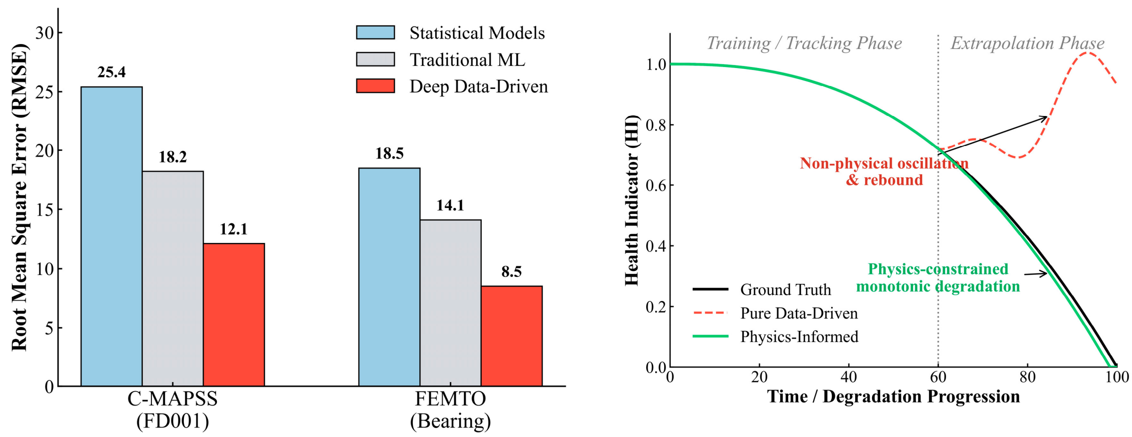
Data-driven models demonstrate exceptional ability to fit complex, nonlinear degradation trajectories in RUL prediction. For multi-rotor UAV propulsion systems, LSTM networks can achieve an end-to-end mapping from vibration signals to RUL by utilizing average peak frequency values [34].

In lithium-ion battery RUL prediction, hybrid deep learning architectures such as CNN-LSTM-DNN can effectively process multi-channel input data, achieving a significant 90.5% improvement in the RMSE metric on the NASA dataset [35]. By integrating deep learning with factorization machines, the XDFM model captures higher-order feature interactions, reducing the average minimum RMSE to 0.03 [36].

Furthermore, data-driven models significantly reduce the reliance on manual feature engineering. For instance, hybrid frameworks like MLP-LSTM can automatically perform feature extraction on state monitoring data vectors, contrasting with conventional models that require manual design [37]. Similarly, models like ForeNet and MDSCT automatically extract multiscale local subtle degradation features from raw SCADA or vibration signals [38,39]. The dual-channel framework combining TCN and Transformer networks achieved a maximum reduction of 48.98% in RMSE on the C-MAPSS dataset [40].

Advantages in Predictive Accuracy Compared to Statistical Models: Compared to traditional statistical models, data-driven approaches demonstrate overwhelming advantages in predictive accuracy under complex, non-linear degradation patterns. Statistical models (e.g., ARIMA, standard Wiener processes) rely on strict linear or Markovian assumptions, which fail to capture multi-stage degradation trajectories. Data-driven models automatically extract high-dimensional spatial-temporal features, typically reducing the RMSE by 20% to 40% on benchmark datasets like C-MAPSS and FEMTO. This superiority is visually demonstrated in Figure 3a, which compares the average RMSE across different model paradigms. As shown in the bar chart, deep data-driven models consistently achieve the lowest prediction errors compared to traditional machine learning and statistical baselines, validating their exceptional capability in fitting complex, non-linear degradation patterns.

Limitations Compared to Hybrid and Physics-Informed Models: Despite their high accuracy, purely data-driven models exhibit critical limitations when compared to hybrid models. They act as "black boxes" lacking physical interpretability, require massive amounts of "run-to-failure" data, and exhibit poor generalization when deployed in unseen operating conditions (out-of-distribution data). Purely data-driven models often produce non-physical predictions, such as fluctuating or non-monotonic RUL trajectories in the late stages of degradation. This limitation is schematically illustrated in Figure 3b. While pure data-driven models fit the training data well, they exhibit severe non-physical oscillations and rebounds during the extrapolation phase (red dashed line). In contrast, physics-informed models outperform conventional data-driven models by embedding physical degradation laws (e.g., Partial Differential Equations or monotonicity constraints) directly into the loss function ($Loss = Loss_{data} + Loss_{physics}$). As depicted by the green solid line in Figure 3b, this physics-constrained architecture forces the neural network to respect physical conservation laws, ensuring smooth and strictly monotonic degradation trajectories. This significantly reduces the reliance on large labeled datasets and drastically improves extrapolation capabilities in the late stages of degradation.



(**a**) (**b**)

Figure 3. Comparative analysis of predictive accuracy and physical consistency among RUL prediction paradigms: (**a**) Predictive Accuracy Comparison; (**b**) Physical Consistency and Extrapolation.

However, purely data-driven models have significant limitations in terms of generalization ability and data requirements. In fitting the wear evolution curves of tool rake faces, pure MLP models achieve a 100% prediction horizon when the training and validation sets share identical operating conditions. However, when conditions differ, the prediction horizon drops to 0%. Furthermore, these models typically require a large volume of "run-to-failure" experimental data to cover various process parameter combinations, resulting in insufficient robustness in small-batch or single-unit production scenarios [39,41].

Based on empirical research, data-driven models are particularly suitable for complex equipment with multivariate time-series sensor data, where converting one-dimensional signals into two-dimensional time-frequency maps or utilizing continuous wavelet convolution layers significantly enhances defect extraction [13,42,43].

3.3. Capabilities and Applications of Hybrid Models

Hybrid forecasting methods combine model-based and data-driven approaches, leveraging the data-driven component to handle complex nonlinear mappings while utilizing the mechanistic model to describe physical degradation laws [2]. The LiRUL framework integrates PCA with physics-aware domain variables and introduces physical constraint loss regularization (monotonicity and smoothness constraints), improving the monotonicity index from 0.82 to 0.97 and reducing long-term extrapolation errors by 40–60% [27].

For aluminum electrolytic capacitors, a library of Wiener degradation models was constructed to adaptively select the optimal model online, significantly reducing RMSE under limited data conditions [44]. Similarly, hybrid modeling for wet clutches integrates lifespan data with multi-source degradation data to update the drift parameter online, reducing the MAE to 0.31 h [45].

Hybrid models are applicable to a wide range of equipment types. For bearing fault prognosis, combining a dual-channel Transformer network with a nonlinear Wiener process effectively quantifies time-varying uncertainty [46]. In wind turbine main bearings, embedding a Physics-Informed Neural Network (PiNN) within RNN units integrates the ISO 281:2007 life formula to reduce prediction inaccuracies [12]. For high-reliability electronic devices, multi-stress-coupled accelerated degradation models effectively fit non-monotonic data under combined temperature, humidity, and electrical stress [47].

Furthermore, hybrid methods are highly suitable for the dynamic condition adaptation of batteries in uncontrolled environments, where integrating reinforcement learning with prior physical knowledge enables robust adaptation to dynamically changing usage conditions [19].

To visually compare the relative strengths and weaknesses of the three paradigms, Figure 4 presents a radar chart. The results indicate that the hybrid model achieves a substantial balance across all six dimensions.

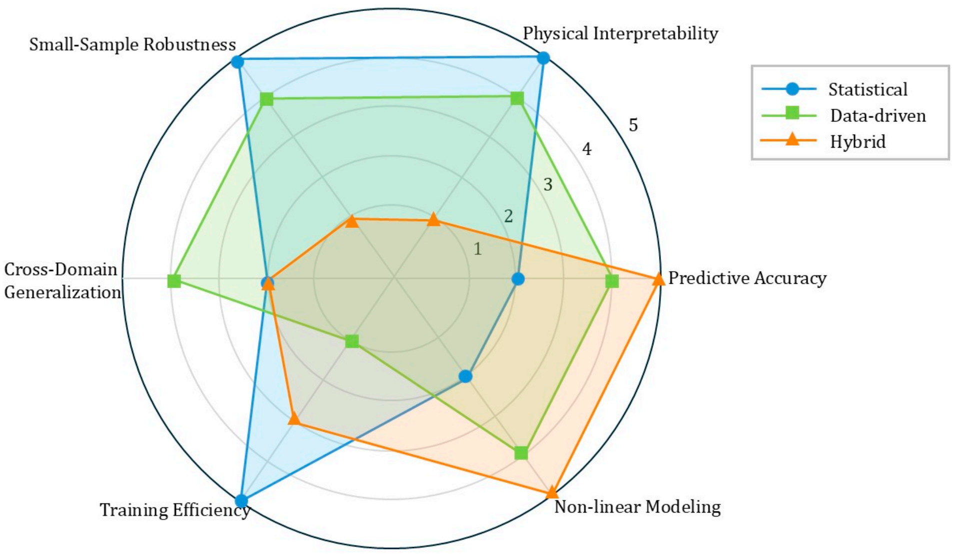


Figure 4. Capability radar chart for statistical models, data-driven models, and hybrid models across six key dimensions.

3.4. Computational Efficiency Comparison

While data-driven and hybrid models offer superior predictive accuracy, their computational efficiency remains a critical bottleneck for real-time engineering applications, particularly on edge devices. Conventional statistical models require minimal training time and possess ultra-low inference latency. In contrast, deep data-driven models incur massive computational costs during training, though their inference time can be optimized via model compression. Hybrid and physics-informed models often suffer from the highest training costs, as solving embedded physical equations during backpropagation is computationally intensive. Table 4 provides a critical comparison of the computational efficiency among these approaches.

Table 4. Computational efficiency comparison of statistical, data-driven, and hybrid models.

Model Category	Training Time	Inference Latency	Memory Footprint	Suitability for Edge Computing
Statistical (e.g., Wiener, PF)	Low (Seconds to Minutes)	Ultra-low (<1 ms)	Very Low (KB level)	Excellent
Data-driven (e.g., CNN, LSTM)	High (Hours to Days)	Low to Medium (1–10 ms)	High (MB to GB level)	Moderate (Requires quantization/pruning)
Hybrid/Physics-Informed	Very High (Days)	Medium (5–20 ms)	Very High (GB level)	Poor (Mostly restricted to cloud/server)

3.5. Experimental Evaluation Framework for RUL Prediction

The C-MAPSS dataset, provided by NASA, consists of four subsets—FD001 through FD004—and is widely used in research on the prediction of RUL for aircraft engines. This dataset is designed for aircraft engine simulation scenarios. The FD001 and FD003 subsets are conducted under single operating conditions, corresponding to single-fault and multi-fault modes, respectively, whereas the FD002 and FD004 subsets involve six different operating conditions, and FD004 includes two fault modes, significantly increasing the prediction difficulty of the latter two. These are suitable for evaluating a model's generalization capability under complex operating conditions and multi-fault modes [8]. In specific applications of the FD001 subset, studies typically adopt a single operating mode and a single failure mode configuration, selecting six sensors (T24, T30, T50, P30, Nf, Nc) strongly correlated with the degradation process from the original 21 sensors as input features to construct more accurate health indicators [48].

The PHM 2008 Challenge dataset contains run-to-failure data from 218 engines and is similarly oriented toward predictive scenarios for aero-turbofan engines; in this study, 10 sensors exhibiting degradation trends (2, 3, 4, 7, 11, 12, 15, 17, 20, 21) were selected from the original 21 sensors for health indicator construction, making it suitable for benchmarking large-scale engine data [49].

For predicting the RUL of rotating machinery such as bearings, the FEMTO-PRONOSTIA dataset provides accelerated bearing aging data. This dataset is based on dual-axis acceleration signal acquisition under constant speed and load conditions, with failure defined as a vibration amplitude reaching 20 g, making it suitable for short-term bearing degradation studies [16]. Similarly, the XJTU-SY dataset also targets bearing equipment but features a longer sampling window. It similarly uses an amplitude of 20 g as the failure threshold, which is suitable for bearing life prediction scenarios requiring longer observation periods, and its inter-individual variability is of significant value for evaluating model robustness on heterogeneous data.

3.6. Dataset Requirements and Training Dynamics

The success of data-driven and hybrid RUL prediction heavily relies on rigorous dataset curation and training dynamics management:

- Data Splitting: Unlike standard computer vision tasks, time-series RUL data requires chronological splitting rather than random shuffling to prevent data leakage. Sliding-window segmentation is employed to generate sequential samples while preserving temporal dependencies.
- Loss Convergence and Error Reduction: RUL prediction often utilizes asymmetric loss functions (e.g., the PHM08 scoring function) that penalize late predictions more heavily than early ones, aligning with conservative maintenance strategies. Gradient clipping and learning rate schedulers (e.g., Cosine Annealing) are essential to ensure smooth loss convergence and avoid local minima.

- Hyperparameter Tuning: The selection of hidden layers, sequence length, and batch size profoundly impacts performance. Automated tuning frameworks, such as Bayesian Optimization or Optuna, are increasingly used to replace exhaustive grid searches, significantly improving prediction accuracy [29].
- Overfitting Control: Given the scarcity of run-to-failure datasets, models are highly prone to overfitting. Techniques such as Dropout, Early Stopping, L2 regularization, and data augmentation (e.g., using TimeGAN to synthesize degradation trajectories) are strictly implemented to enhance model generalization.

3.7. Summary of Methodological Capabilities

To consolidate the findings of this section, Table 5 summarizes the key studies discussed, highlighting their applications, methodological innovations, and prediction efficiency.

Table 5. Summary of methodologies, key outcomes, and prediction efficiency in recent RUL prediction studies (Section 3).

Reference	Application	Methodology	Key Outcomes	Prediction Efficiency *
[31]	Aircraft engines	PCA load interpretability method	Mapped principal components to physical context; enhanced feature interpretability.	Clarified physical significance of statistical features for engine instability.
[32]	White LEDs	Dual-accelerated Wiener process model	Developed generalized confidence intervals for small-sample stability.	Coverage of confidence intervals remained close to nominal levels (90%/95%).
[35]	Li-ion batteries	CNN-LSTM-DNN hybrid architecture	Processed multi-channel input without deep electrochemical knowledge.	90.5% improvement in RMSE on NASA dataset.
[40]	Aero-engines (C-MAPSS)	Dual-channel TCN-Transformer	Extracted degraded features from multiple dimensions.	RMSE reduced by 48.98%; Score metric reduced by 92.49%.
[27]	Edge devices (LiRUL)	Physics-constrained lightweight learning	Introduced monotonicity and smoothness constraints to loss function.	Monotonicity index improved to 0.97; extrapolation errors reduced by 40–60%.
[44]	Aluminum electrolytic capacitors	Adaptive Wiener degradation model library	Adaptively selected optimal degradation model online based on historical similarity.	Significantly reduced RMSE and MRE under limited data conditions.
[46]	Rolling bearings	DCTN-CBAM + nonlinear Wiener process	Characterized degradation rates and quantified time-varying uncertainty.	Effectively processed multi-domain features with uncertainty quantification.

* The "Prediction Efficiency" column includes quantitative metrics (e.g., RMSE, R^2, MAPE, inference time) where available or qualitative impact assessments for studies focusing on methodological innovation rather than direct predictive modeling.

Despite the unique advantages of the three aforementioned paradigms in their respective scenarios, they collectively reveal systemic capability gaps when confronted with complex real-world constraints. The subsequent section will critically diagnose these systemic limitations, particularly focusing on robustness, physical interpretability, and uncertainty quantification.

4. Critical Capability Gaps: Systemic Limitations in Robustness, Interpretability, and Uncertainty Quantification

This section further identifies the systemic capability gaps exposed by existing technologies when faced with real-world constraints. These gaps primarily focus on five aspects: physical consistency, uncertainty quantification, the trade-off between physical consistency and lightweight solutions, differences in cross-domain adaptation phases, and data scarcity versus generalization capabilities. To convey this perspective intuitively, this paper employs the iceberg model for illustration, where the portion below the waterline represents those systemic gaps that are often overlooked but severely constrain the deployment of remaining life prediction in practical engineering, as shown in Figure 5.

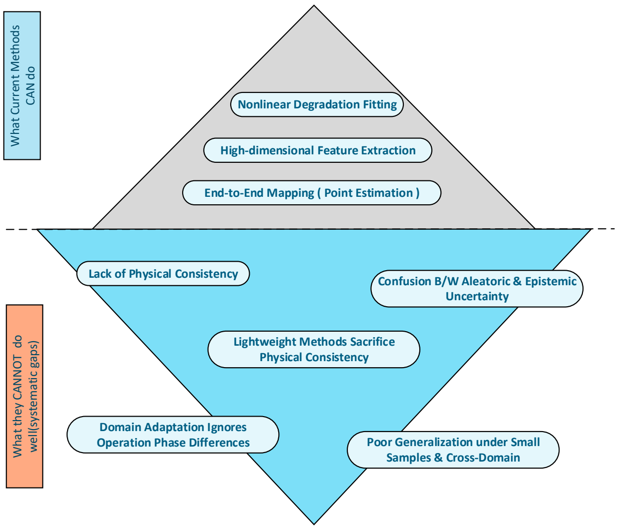


Figure 5. The Iceberg Model of the Gap in Remaining Life Prediction Capabilities.

4.1. Lack of Physical Consistency Constraints

Purely data-driven models (such as LSTM and GRU) typically operate as "black boxes," lacking physical interpretability and failing to provide a transparent link to electrochemical degradation processes, which can lead to predictions that violate physical laws. Experiments have shown that an unregularized LSTM baseline model violates monotonic degradation behavior in nearly 20% of cycles [27]. This physical inconsistency not only reduces the credibility of prediction results but also limits the model's engineering applications in safety-critical systems.

4.2. Confusion Between Epistemic and Aleatory Uncertainty in Uncertainty Quantification

Existing uncertainty quantification methods fail to effectively distinguish between these two types of uncertainty. The prediction intervals of Conformal Quantile Regression (CQR) and its variants fail to distinguish between the variability of the conditional distribution (i.e., aleatory uncertainty) and the uncertainty in estimating that distribution (i.e., epistemic uncertainty), resulting in overly narrow prediction intervals in regions with high epistemic uncertainty [50]. Furthermore, most deep learning-based methods can only perform point estimates, while the primary limitation of methods such as Probabilistic ResNet is that they capture only aleatory uncertainty and ignore epistemic uncertainty, causing the model to exhibit overconfidence when faced with out-of-distribution data [51].

4.3. The Trade-Off Between Computational Efficiency and Physical Consistency in Lightweight Solutions

In lightweight solutions designed for edge deployment, existing methods often sacrifice physical consistency in exchange for computational efficiency. Lightweight deep learning frameworks focus on structural optimization, such as model pruning and knowledge distillation, but neglect the integration of domain-aware features, interpretability, and physical consistency constraints. Additionally, the use of generic activation functions, such as ReLU, whose higher-order derivatives are zero, causes Physics-Informed Neural Networks (PINNs) to fail when computing the higher-order derivatives required for the physical loss function, thereby compromising the model's physical consistency and training effectiveness [13].

4.4. The Neglect of Operational Phase Differences in Cross-Domain Adaptation Methods

Existing cross-domain adaptation methods suffer from a systematic flaw in that they neglect differences between operational phases. Most domain adaptation methods fail to distinguish between different phases of the operating profile, leading to alignment errors caused by under- or over-representation of different operational phases. For example, in the prediction of RUL for aircraft engines, traditional adversarial domain adaptation methods fail to capture the unique differences in edge distributions across different operational phases. This leads to a decline in model performance in scenarios where varying flight durations result in different stress and degradation patterns [52].

4.5. Data Scarcity and Insufficient Cross-Domain Generalization Capability

Data scarcity severely limits a model's cross-domain generalization capability. In practical applications, real-time failure data from actual scenarios is often lacking. For example, hydraulic pump lifespan tests are lengthy, and failure may not occur during the testing process, making it difficult to obtain complete failure data to directly validate prediction results [53]. At the same time, deep learning methods are prone to overfitting when data from inverters, battery energy storage systems, or distributed energy sources is insufficient [54]. Furthermore, differences in degradation patterns between individual units (unit-to-unit variability) further exacerbate the difficulty of cross-domain generalization.

The capability gaps in the five aspects mentioned above are deeply intertwined, collectively constituting the primary obstacles preventing current RUL prediction technologies from transitioning from controlled laboratory environments to complex, real-world engineering applications.

To consolidate the analysis of these systemic limitations, Table 6 summarizes the key studies discussed in this section, highlighting the specific applications, the systemic gaps addressed, and their impact on prediction reliability.

Table 6. Summary of systemic capability gaps, addressed limitations, and key impacts in recent RUL prediction studies.

Reference	Application	Systemic Limitation	Key Outcomes *
[27]	Edge devices/Batteries	Lack of physical consistency constraints	Unregularized LSTM models violated monotonic degradation behavior in nearly 20% of cycles, reducing credibility.
[50]	General RUL prediction	Confusion between aleatory and epistemic uncertainty	CQR methods produced overly narrow prediction intervals in regions with high epistemic uncertainty.
[51]	General deep learning	Ignorance of epistemic uncertainty	Probabilistic ResNet exhibited severe overconfidence when confronted with out-of-distribution data.
[13]	Power electronics	Trade-off in lightweight solutions (ReLU activation)	Zero higher-order derivatives of ReLU caused PINNs to fail in computing physical loss functions, compromising consistency.

Table 6. *Cont.*

Reference	Application	Systemic Limitation	Key Outcomes *
[52]	Aircraft engines (C-MAPSS)	Neglect of operational phase differences in domain adaptation	Traditional adversarial methods failed to align edge distributions across varying flight phases, degrading performance.
[53]	Hydraulic pumps	Data scarcity and lack of complete failure data	Lengthy lifespan tests without observed failures hindered the direct validation of RUL prediction results.
[54]	Inverters/Energy storage	Insufficient data leading to overfitting	Deep learning models showed severe overfitting and poor cross-domain generalization under limited data conditions.

* The "Key Outcomes" column summarizes the qualitative impact of the identified systemic gaps on model reliability and engineering deployment, as these subsections focus on diagnosing limitations rather than proposing direct predictive improvements.

5. Methodological Evolution: A New Paradigm for RUL Prediction Aimed at Engineering Implementation

To address the aforementioned capability gaps, researchers have explored algorithmic innovations from multiple angles. This section distills the four most representative evolutionary paths: mechanism-guided neural architecture design, edge-oriented compact physics-AI fusion, multi-source degenerate joint representation and cross-domain adaptation, and digital twin-driven prediction methods. Figure 6 clearly illustrates the methodological evolution timeline, showing a gradual shift from purely data-driven black-box models toward architectures that integrate physical information and closed-loop mechanisms.

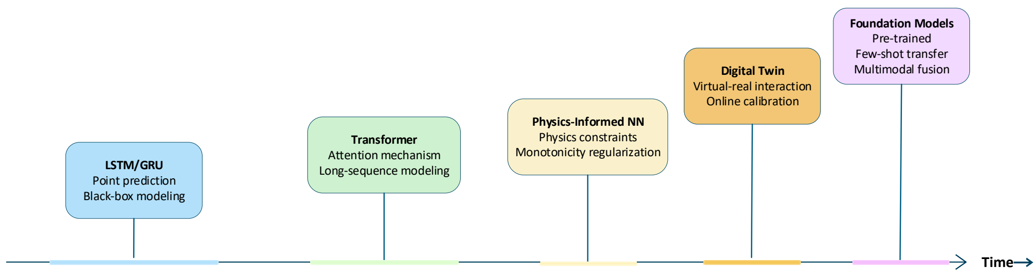


Figure 6. Roadmap for the Evolution of Residual Life Prediction Methodologies (2015–2025 and Beyond).

5.1. Mechanism-Driven Neural Architecture Design

In terms of mechanism-driven neural architecture design, the collaborative design of the Position-Sensitive Attention (PSA) mechanism and the Gated Hierarchical Long Short-Term Memory (GHLSTM) network [55] effectively addresses the issue of high-level features in traditional Transformer encoders being insensitive to local context. By calculating similarity scores that simultaneously consider input content and spatial encoding, this architecture achieves hierarchical regression prediction through multi-scale feature concatenation. Experimental validation on the C-MAPSS dataset demonstrates significant improvements, achieving a 4% enhancement in the Score metric on FD001 and up to 10% and 6% improvements in Score and RMSE metrics on FD004, respectively, outperforming current state-of-the-art methods.

5.2. Compact Physics-AI Fusion for Edge Computing

To address the practical implementation needs of edge computing scenarios with limited resources, the LiRUL framework [27] proposes a compact physics-AI fusion paradigm. By integrating a hybrid feature mapping strategy that combines PCA with physical domain variables, the framework concatenates explicit domain-aware feature statistics onto the principal components, achieving a balance between statistical compactness and physical interpretability. Experimental results indicate that LiRUL reduces RMSE by 45–60% and inference latency by 60% (below 1 s) compared to traditional models, utilizing merely 0.15 M parameters and <5 MB memory. Furthermore, the physical constraints elevate the monotonicity index to 0.97, effectively suppressing non-physical oscillations during long-term extrapolation.

5.3. Multi-Source Degradation Joint Representation and Cross-Domain Adaptation

In the area of multi-source degradation joint representation and cross-domain adaptation, the OPS-DANN method [52] overcomes the limitations of traditional domain adaptation frameworks by incorporating discrete phases of the operation profile into the fine-grained alignment process. By deploying dedicated domain discriminators for each phase (hard allocation) and utilizing probability distributions for smooth sample allocation (soft allocation), this method addresses alignment bias caused by multimodal feature distributions. In the short-to-long-haul (S $\rightarrow$ L) cross-domain task on the N-CMAPSS dataset, this phase-aware alignment mechanism reduces RMSE by at least 9% and improves the S-score by 2% compared to standard DANN methods, proving its superiority in capturing degradation-related features.

5.4. Digital Twin-Driven RUL Prediction Method

The new paradigm for RUL prediction driven by digital twins achieves a deep integration of physical models and data-driven approaches through a five-dimensional digital twin model [28]. This framework constructs a Wiener-Arrhenius accelerated degradation model to describe non-monotonic degradation trends under thermal stress. Concurrently, a parameter update mechanism based on Bayesian methods performs online adaptive adjustments to posterior distributions and dynamically selects the optimal error model via the Akaike Information Criterion (AIC). Experimental validation demonstrates that after three rounds of dynamic error correction via digital twin interaction, the RUL prediction error is reduced to 8 h. Furthermore, as interaction time increases, the prediction confidence interval narrows significantly, achieving effective decoupling and reduction of uncertainty.

To provide a consolidated overview of these evolutionary paths, Table 7 summarizes the key studies discussed in this section, highlighting their methodological innovations and prediction efficiency in engineering implementations.

Table 7. Summary of methodological evolutions, key outcomes, and prediction efficiency for engineering implementation.

Reference	Application	Methodological Evolution	Key Outcomes	Prediction Efficiency *
[55]	Aero-engines (C-MAPSS)	Mechanism-driven PSA and GHLSTM architecture	Addressed Transformer's insensitivity to local context via hierarchical regression.	4% Score improvement on FD001; 10% Score and 6% RMSE improvements on FD004.
[27]	Edge devices (LiRUL)	Compact physics-AI fusion with physical constraints	Balanced statistical compactness and physical interpretability for edge deployment.	RMSE reduced by 45–60%; Inference < 1 s; 0.15 M parameters; <5 MB memory.

Table 7. *Cont.*

Reference	Application	Methodological Evolution	Key Outcomes	Prediction Efficiency *
[52]	Aircraft engines (N-CMAPSS)	Phase-aware OPS-DANN for cross-domain adaptation	Aligned edge distributions across discrete operational phases (hard/soft allocation).	RMSE reduced by ≥9%; S-score improved by 2% in short-to-long-haul tasks.
[28]	Piezoelectric sensors	Digital Twin-driven Wiener-Arrhenius model	Achieved online Bayesian parameter updates and dynamic AIC error correction.	Prediction error reduced to 8 h; significant narrowing of confidence intervals.

* The "Prediction Efficiency" column highlights the quantitative performance gains and resource utilization metrics that demonstrate the feasibility of these evolved methodologies for real-world engineering deployment.

6. An Integrated RUL Prediction Framework for System-Level Applications: From Point-in-Time Predictions to Closed-Loop Maintenance Decisions

The core of the system-level RUL prediction integration framework lies in seamlessly embedding probabilistic prediction results into the closed-loop of O&M decision-making, thereby bridging the gap between point-in-time predictions and system-level maintenance optimization. This section reviews advancements in four areas: dynamic threshold setting, multi-objective maintenance optimization, multi-component coordinated scheduling, and online adaptive calibration. Figure 7 illustrates how probabilistic prediction, uncertainty quantification, and maintenance optimization are interconnected through online feedback, thereby transforming RUL prediction from an isolated algorithm into an actionable decision support system.

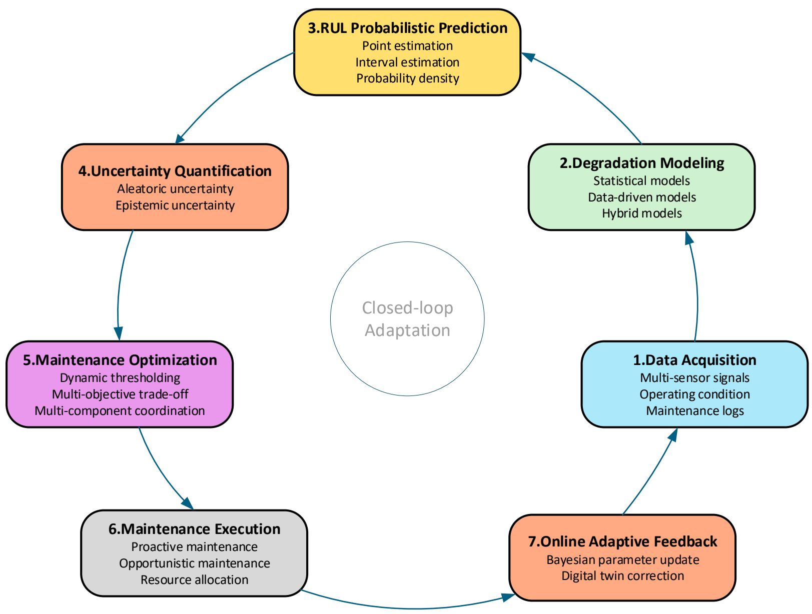


Figure 7. System-level closed-loop integration framework driven by RUL prediction for maintenance decision-making.

6.1. Dynamic Threshold Setting Based on the RUL Probability Distribution

Regarding dynamic threshold setting, Crespo del Castillo and Parlikad developed a mathematical framework integrating the RUL probability distribution (Weibull) with component criticality, revealing adaptive adjustment rules: high-criticality components undergo immediate intervention during the first available preventive maintenance window to mitigate risks, while low-criticality components maximize RUL utilization, intervening only near the end of life [17]. This mechanism provides a theoretical foundation for dynamic thresholds by quantifying the relationship between failure probability density and maintenance costs.

6.2. Multi-Objective Maintenance Optimization Integrating Uncertainty and Component Criticality

Multi-objective optimization models focus on explicitly handling RUL uncertainty and coordinating multi-dimensional objectives. Shulian Xie et al. innovatively combined an LSTM network with kernel density estimation (KDE-SGB) to generate a continuous probability distribution for RUL, and employed a weighted TOPSIS method to integrate maintenance cost rates, system availability, and reliability metrics, achieving closed-loop optimization from state prediction to decision-making [56]. In a bearing case study, this strategy significantly reduced maintenance costs by 24.49–70.02%, while simultaneously improving system availability (by 0.46–1.90%) and reliability (by up to 27.50%), ensuring the system remained failure-free under various operating conditions.

6.3. Multi-Component Opportunistic Maintenance and Collaborative Scheduling Strategies

Breakthroughs in multi-component collaborative scheduling have focused on system-level optimization under resource constraints. The end-to-end framework developed by Mitici et al. integrates CNNs with the Monte Carlo Dropout method into maintenance planning. By replacing traditional failure distribution assumptions with an update-reward process, it achieves multi-component collaborative scheduling under hangar and slot capacity constraints, reducing costs by 53% compared to time-based strategies [57]. Lv et al. introduced an opportunistic maintenance mechanism that schedules proactive actions based on predicted RUL, effectively reducing the total maintenance cost rate [58]. For mission-oriented systems, Ma et al. designed a dynamic decision-making mechanism allocating limited maintenance budgets during mission intervals, increasing the minimum mission success probability of an air defense system from 0.8538 to 0.9276 [59]. Furthermore, Bankati et al.'s RUL-driven "Daisy-Chain" energy management strategy employs a two-stage scheduling mechanism to move the most degraded components to the front of the queue, extending the lifespan of a multi-fuel cell stack system by 2.48 times [60].

6.4. Digital Twin-Assisted Online Adaptive Calibration of Degradation Models

Regarding online adaptive mechanisms, Dong et al. integrated an adaptive Wiener model with an incomplete maintenance strategy, proposing a parameter update method combining maximum likelihood estimation and Bayesian inference for real-time calibration of degradation model parameters [61]. Validated on gyroscope degradation data, this method significantly narrowed the RUL probability density function, reducing relative and mean squared errors. Yang et al. developed a correction mechanism based on real-time quality data, updating the posterior distribution of in-process acceptance rates via Bayesian estimation and utilizing Kolmogorov differential equations to update the machine state probability vector in real time, thereby driving a closed-loop maintenance decision-making process [62].

To consolidate the advancements in system-level integration, Table 8 summarizes the key studies discussed in this section, highlighting their integration strategies, application scenarios, and decision-making efficiency.

Table 8. Summary of system-level integration frameworks, closed-loop strategies, and decision-making efficiency.

Reference	Application	System-Level Integration Strategy	Decision-Making Efficiency *
[17]	General components	Dynamic threshold setting based on RUL PDF and criticality	High-criticality components undergo immediate intervention; low-criticality maximize RUL utilization.
[56]	Bearings	Multi-objective optimization via LSTM-KDE and weighted TOPSIS	Reduced maintenance costs by 24.49%–70.02%; improved availability and reliability.
[57]	Aircraft engines	End-to-end CNN + Monte Carlo Dropout for multi-component scheduling	Replaced traditional failure assumptions; reduced maintenance costs by 53% under capacity constraints.
[59]	Air defense systems	Dynamic budget allocation based on RUL PDF during mission intervals	Increased minimum mission success probability from 0.8538 to 0.9276.
[60]	Multi-fuel cell stacks	RUL-driven "Daisy-Chain" two-stage energy management scheduling	Extended system lifespan by 2.48 times by prioritizing degraded components.
[61]	Gyroscopes	Digital Twin-assisted adaptive Wiener model with Bayesian updates	Narrowed RUL probability density function; significantly reduced prediction errors via real-time calibration.

* The "Decision-Making Efficiency" column highlights the quantitative improvements in maintenance costs, system availability, mission success rates, and lifespan extension achieved by transitioning from isolated RUL predictions to closed-loop system-level decisions.

7. Conclusions and Outlook

7.1. Key Findings

As a core technology of predictive maintenance, the methodological evolution of RUL prediction has undergone systematic transitions: from a single paradigm to multidimensional collaboration, from static modeling to dynamic closed-loop systems, and from point estimation to probabilistic decision-making. The key findings of this study can be summarized as follows.

Statistical models possess irreplaceable advantages in terms of small-sample stability, parameter interpretability, and alignment with physical mechanisms, making them particularly suitable for high-reliability equipment with relatively clear degradation mechanisms and limited experimental data. Their core limitations lie in the flexibility required to model non-monotonic degradation paths and multi-stress coupling effects, necessitating enhancement strategies such as generalized confidence inference, nonlinear drift term design, and subjective parameter estimation to maintain statistical robustness.

Data-driven models demonstrate powerful high-dimensional nonlinear mapping capabilities and end-to-end feature learning potential, performing exceptionally well in multi-source sensor fusion, direct modeling of raw signals, and generalization to complex operating conditions. However, their empirical validity is highly dependent on the completeness, representativeness, and annotation quality of training data. Performance degradation is prone to occur in scenarios involving cross-condition transfer, small-sample extrapolation, and distribution shifts, and there are widespread issues of lacking physical consistency and insufficient characterization of uncertainty.

The hybrid modeling paradigm, by organically integrating mechanistic constraints with data adaptability, has emerged as the most engineering-feasible technical approach. On one hand, it leverages physical models to provide prior constraints on degradation, on the other hand, it utilizes data-driven modules to compensate for model errors, characterize stochastic effects, and capture unknown coupling mechanisms, thereby achieving a substantial balance between prediction accuracy, extrapolation robustness, and uncertainty decoupling capabilities.

Key methodological advancements are embodied in three structural breakthroughs. First, mechanism-guided neural architecture design, which embeds physical laws into network structures, ensures that the model possesses intrinsic physical consistency. Second, refined domain adaptation mechanisms break away from traditional global alignment paradigms. This is achieved by introducing operational profile segmentation, hard-soft allocation discriminators, and mode-aware training strategies, which significantly mitigate feature drift caused by operational heterogeneity. Third, a digital twin-driven closed-loop evolution framework—leveraging a five-dimensional digital twin to enable online Bayesian updates of physical model parameters, dynamic selection of error models, and real-time convergence of prediction confidence intervals, thereby shifting RUL prediction from offline evaluation to continuous calibration.

At the system-level application dimension, RUL prediction is rapidly integrating into core O&M decision-making workflows. Dynamic threshold setting is based on the joint optimization of component criticality and the RUL probability density function. Multi-objective maintenance planning explicitly models uncertainty to collaboratively balance cost, availability, and reliability. Multi-component collaborative scheduling implements energy reallocation and dynamic resource allocation based on RUL ranking. Empirical results demonstrate that this integrated framework can reduce maintenance costs by up to 53%, improve system availability, and increase mission success probability, validating the effectiveness of transitioning from single-point state awareness to system-level intelligent decision-making.

7.2. Open Questions

Although significant progress has been made in the field of RUL prediction, several structural gaps in capability remain to be addressed.

The challenge of balancing physical consistency with model lightweighting. Lightweight solutions such as model pruning often sacrifice physical consistency for computational efficiency. Maintaining fundamental physical properties (e.g., monotonicity) requires additional regularization, which inherently conflicts with parameter reduction. Designing compact yet physically credible models remains a key obstacle.

The decoupling dilemma in uncertainty quantification. Existing methods have not yet achieved effective decoupling between aleatory uncertainty (data-intrinsic randomness) and epistemic uncertainty (model knowledge gaps). Conformal prediction struggles to distinguish between uncertainty types, while Bayesian neural networks are computationally expensive and may produce overconfident intervals in out-of-distribution scenarios.

Cross-domain adaptation lacks theoretical modeling at the operational phase granularity. Traditional methods employ global alignment, ignoring distribution differences across operational phases. Although phase-based methods like OPS-DANN show promise, a systematic analytical framework regarding when phase-level adaptation is required and how to automatically identify phase boundaries is still lacking.

Generalization with limited data is constrained by data generation quality. Data augmentation methods (e.g., GANs, VAEs) can alleviate data scarcity, but the authenticity of generated degradation trajectories is difficult to guarantee, risking model overfitting. Furthermore, whether the inductive biases of deep learning models are theoretically suitable for RUL tasks requires further examination.

https://doi.org/10.3390/machines14070724

7.3. Outlook for the Future

Looking ahead, the RUL field is expected to achieve breakthrough progress in the following areas:

Theoretical unification driven by both physics and data. Developing hybrid neural-symbolic architectures capable of formal physical verification, embedding physical equations and conservation laws as hard constraints within neural networks. The maturation of differentiable physics simulators will enable end-to-end joint optimization, breaking away from the serial paradigm of "physical modeling first, followed by data calibration."

Edge-oriented, compact physics-AI fusion paradigms. Designing lightweight architectures that preserve fundamental physical mathematics while adhering to resource constraints. Techniques such as physics-informed knowledge distillation and hardware-friendly monotonic activation functions will enable edge devices to produce physically credible predictions.

Hierarchical uncertainty propagation models with runtime profile awareness. Establishing hierarchical Bayesian frameworks to distinguish between uncertainty sources (measurement noise, model error, environmental perturbations) and propagate them forward along the time axis. A hybrid paradigm combining conformal prediction and Bayesian methods is expected to balance computational efficiency and uncertainty decoupling.

Cross-domain generalization mechanisms based on causal inference. Moving beyond correlation-based learning, causal graph models can identify invariant causal outcomes during degradation, ensuring stability across operating conditions. Counterfactual reasoning can generate virtual intervention trajectories, fundamentally alleviating generalization challenges caused by data scarcity and distribution drift.

A predictive maintenance ecosystem empowered by foundation models. Pre-trained Large Language Models (LLMs) can extract prior knowledge from unstructured maintenance records, while time-series foundation models, pre-trained on massive multi-device data, can be fine-tuned with minimal target-domain samples. This paradigm will significantly lower the data threshold and reshape the PHM research roadmap.

In summary, the field of RUL prediction is at a critical juncture of transitioning from an "algorithm-driven" to a "system-driven" approach. The deep integration of physics and data, the refinement of uncertainty quantification, and decision-oriented closed-loop integration will be the core themes of future research.

Author Contributions: Conceptualization, L.W. and Y.Y.; methodology, L.W. and X.L.; formal analysis, L.W. and M.C.; investigation, L.W., X.L. and M.C.; resources, Y.Y.; data curation, X.L. and M.C.; writing—original draft preparation, L.W.; writing—review and editing, Y.Y., X.L. and M.C.; visualization, L.W. and M.C.; supervision, Y.Y.; project administration, Y.Y.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reza, M.S.; Mannan, M.; Mansor, M.; Ker, P.J. Recent advancement of remaining useful life prediction of lithium-ion battery in electric vehicle applications: A review of modelling mechanisms, network configurations, factors, and outstanding issues. *Energy Rep.* **2024**, *11*, 4824–4848. [CrossRef]
2. Office, J.E.; Gebraeel, N.; Lei, Y.; Li, N.; Si, X.; Zio, E. Prognostics and Remaining Useful Life Prediction of Machinery: Advances, Opportunities and Challenges. *J. Dyn. Monit. Diagn.* **2023**, *2*, 1–12.
3. Fu, S.; Avdelidis, N.P. Prognostic and Health Management of Critical Aircraft Systems and Components: An Overview. *Sensors* **2023**, *23*, 8124. [CrossRef] [PubMed]

4. Piao, S.; Huang, R.; Tsung, F. CRULP: Reliable RUL Estimation Inspired by Conformal Prediction. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3505411. [CrossRef]
5. Wang, Z.; Li, M.; Chen, Y.; Liu, Z. An interpretable hierarchical temporal aggregation network for RUL prediction. *Expert Syst. Appl.* **2025**, *289*, 128302. [CrossRef]
6. Landau, D.; De Pater, I.; Mitici, M.; Saurabh, N. Federated learning framework for collaborative remaining useful life prognostics: An aircraft engine case study. *Future Gener. Comput. Syst.* **2026**, *174*, 107945. [CrossRef]
7. Jiang, L.; Zhang, X.; Cao, H.; Zhang, Y. A transformer-based framework with historical data fusion for RUL prediction. *Meas. Sci. Technol.* **2025**, *36*, 106103. [CrossRef]
8. Zhou, H.; Wu, Q.; Peng, P.; Guo, Z. Turbofan Engine's RUL Prediction Based on the CSI-EMD and Double-Channel Multilayer Feature Fusion Network. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 6396–6405. [CrossRef]
9. Cui, H.; Guo, X.; Yu, L. Leveraging Pre-Trained GPT Models for Equipment Remaining Useful Life Prognostics. *Electronics* **2025**, *14*, 1265. [CrossRef]
10. Hu, Q.; Zhao, Y.P.; Wang, Y.; Peng, P.; Ren, L. Remaining Useful Life Estimation in Prognostics Using Deep Reinforcement Learning. *IEEE Access* **2023**, *11*, 32919–32934. [CrossRef]
11. David, S.M.; Juan, G.P.; Joaquín, B.D. CONELPABO: Composite networks learning via parallel Bayesian optimization to predict remaining useful life in predictive maintenance. *Neural Comput. Appl.* **2025**, *37*, 7423–7441. [CrossRef]
12. Cuesta, J.; Leturiondo, U.; Vidal, Y.; Pozo, F. A review of prognostics and health management techniques in wind energy. *Reliab. Eng. Syst. Saf.* **2025**, *260*, 111004. [CrossRef]
13. Fassi, Y.; Heiries, V.; Boutet, J.; Boisseau, S. Toward Physics-Informed Machine-Learning-Based Predictive Maintenance for Power Converters—A Review. *IEEE Trans. Power Electron.* **2024**, *39*, 2692–2720. [CrossRef]
14. Elminir, H.K.; El-Brawany, M.A.; Ibrahim, D.A.; Elattar, H.M.; Ramadan, E.A. An efficient deep learning prognostic model for remaining useful life estimation of high speed CNC milling machine cutters. *Results Eng.* **2024**, *24*, 103420. [CrossRef]
15. Lodygowski, T.; Szrama, S. Unsupervised Classification and Remaining Useful Life Prediction for Turbofan Engines Using Autoencoders and Gaussian Mixture Models: A Comprehensive Framework for Predictive Maintenance. *Appl. Sci.* **2025**, *15*, 7884. [CrossRef]
16. Krokotsch, T.; Ragab, M.; Wu, M.; Li, X.; Chen, Z.; Gühmann, C. From Inconsistency to Unity: Benchmarking Deep Learning-Based Unsupervised Domain Adaptation for RUL. *IEEE Trans. Autom. Sci. Eng.* **2025**, *23*, 3316–3328. [CrossRef]
17. del Castillo, A.C.; Parlikad, A.K. Dynamic fleet management: Integrating predictive and preventive maintenance with operation workload balance to minimise cost. *Reliab. Eng. Syst. Saf.* **2024**, *249*, 110243. [CrossRef]
18. Das, K.; Kumar, R. Electric vehicle battery capacity degradation and health estimation using machine-learning techniques: A review. *Clean Energy* **2023**, *7*, 1268–1281. [CrossRef]
19. Patrizi, G.; Martiri, L.; Pievatolo, A.; Meccariello, G.; Cristald, L.; Nikiforov, N.D. A Review of Degradation Models and Remaining Useful Life Prediction for Testing Design and Predictive Maintenance of Lithium-Ion Batteries. *Sensors* **2024**, *124*, 3382. [CrossRef] [PubMed]
20. El-Dalahmeh, M.; Al-Greer, M.; El-Dalahmeh, M.; Bashir, I. Physics-based model informed smooth particle filter for remaining useful life prediction of lithium-ion battery. *Measurement* **2023**, *214*, 112838. [CrossRef]
21. Fathi, K.; Sadurski, M.; Kleinert, T.; Venn, H.W. Source Component Shift Detection & Classification for Improved Remaining Useful Life Estimation in Alarm-Based Predictive Maintenance. In Proceedings of the 23rd International Conference on Control, Automation and Systems (ICCAS), Yeosu, Republic of Korea, 17–20 October 2023.
22. Zhang, T.; Hou, Z.; Li, X.; Chen, Q.; Wang, Q.; Lüddeke, C.; Wu, L.; Sun, W. A novel multivariable prognostic approach for PEMFC degradation and remaining useful life prediction using random forest and temporal convolutional network. *Appl. Energy* **2025**, *385*, 125569. [CrossRef]
23. Hachem, H.; Vu, H.C.; Fouladirad, M. Different methods for RUL prediction considering sensor degradation. *Reliab. Eng. Syst. Saf.* **2024**, *243*, 109897. [CrossRef]
24. Kim, J.H.; Kim, C.H.; Kim, D.; Shin, D.; Jang, I.; Jung, J.H.; Park, S. Hybrid Statistical–AI approach for RUL prediction of reactor protection system components. *Nucl. Eng. Technol.* **2026**, *58*, 103952. [CrossRef]
25. Schwendemann, S.; Rausch, A.; Sikora, A. A Hybrid Predictive Maintenance Solution for Fault Classification and Remaining Useful Life Estimation of Bearings Using Low-Cost Sensor Hardware. *Procedia Comput. Sci.* **2024**, *232*, 128–138. [CrossRef]
26. Yao, X.; Zhu, J.; Jiang, Q.; Yao, Q.; Shen, Y.; Zhu, Q. RUL prediction method for rolling bearing using convolutional denoising autoencoder and bidirectional LSTM. *Meas. Sci. Technol.* **2024**, *35*, 035111. [CrossRef]
27. Pournazari, J.; El-Dalahmeh, M.; Al-Dubai, A.; Liu, X.; Park, D.H.; Marco, J.; Dinh, T.; Ahn, J.H.; Niri, M. LiRUL: A physics-constrained and lightweight learning framework for edge-based RUL prediction. *Results Eng.* **2026**, *29*, 109537. [CrossRef]
28. Fu, C.; Gao, C.; Zhang, W. A Digital-Twin Framework for Predicting the Remaining Useful Life of Piezoelectric Vibration Sensors with Sensitivity Degradation Modeling. *Sensors* **2023**, *23*, 8173. [CrossRef] [PubMed]

29. Jafari, S.; Byun, Y.C. Optimizing Battery RUL Prediction of Lithium-Ion Batteries Based on Harris Hawk Optimization Approach Using Random Forest and LightGBM. *IEEE Access* **2023**, *11*, 1. [CrossRef]
30. Li, W.; Zhang, L.; Wu, C.; Wang, Y.; Cui, Z.; Niu, C. A data-driven approach to RUL prediction of tools. *Adv. Manuf.* **2023**, *12*, 6–18. [CrossRef]
31. Alomari, Y.; Andó, M.; Baptista, M.L. Advancing aircraft engine RUL predictions: An interpretable integrated approach of feature engineering and aggregated feature importance. *Sci. Rep.* **2023**, *13*, 13466. [CrossRef] [PubMed]
32. Jiang, P.; Yang, X. Reliability inference and remaining useful life prediction for the doubly accelerated degradation model based on Wiener process. *AIMS Math.* **2023**, *8*, 7560–7583. [CrossRef]
33. Su, K.; Deng, B.; Tang, S.; Sun, X.; Fang, P.; Si, X.; Han, X. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on a Cubic Polynomial Degradation Model and Envelope Extraction. *Batteries* **2023**, *9*, 441. [CrossRef]
34. Ozkat, E.C.; Bektaş, O.; Nielsen, M.J.; Anders, C.H. A data-driven predictive maintenance model to estimate RUL in a multi-rotor UAS. *Int. J. Micro Air Veh.* **2023**, *15*, 17568293221150171. [CrossRef]
35. Tiane, A.; Okar, C.; Alzayed, M.; Chaoui, H. Comparing Hybrid Approaches of Deep Learning for Remaining Useful Life Prognostic of Lithium-Ion Batteries. *IEEE Access* **2024**, *12*, 70334–70344. [CrossRef]
36. Nair, P.; Vakharia, V.; Borade, H.; Shah, M.; Wankhede, V. Predicting Li-Ion Battery Remaining Useful Life: An XDFM-Driven Approach with Explainable AI. *Energies* **2023**, *16*, 5725. [CrossRef]
37. Bharatheedasan, K.; Maity, T.; Kumaraswamidhas, L.A.; Durairaj, M. Enhanced fault diagnosis and remaining useful life prediction of rolling bearings using a hybrid multilayer perceptron and LSTM network model. *Alex. Eng. J.* **2025**, *115*, 355–369. [CrossRef]
38. Sun, N.; Tang, J.; Ye, X.; Zhang, C.; Zhu, X.; Wang, S.; Sun, Y. Remaining useful life prognostics of bearings based on convolution attention networks and enhanced transformer. *Heliyon* **2024**, *10*, e38317. [CrossRef] [PubMed]
39. Shah, S.S.; Daoliang, T.; Kumar, S.C. RUL forecasting for wind turbine predictive maintenance based on deep learning. *Heliyon* **2024**, *10*, e39268. [CrossRef] [PubMed]
40. Wang, R.; Zhang, Y.; Hu, C.; Yang, Z.; Li, H.; Liu, F.; Li, L.; Guo, J. A Parallel Prognostic Method Integrating Uncertainty Quantification for Probabilistic Remaining Useful Life Prediction of Aero-Engine. *Processes* **2024**, *12*, 2925. [CrossRef]
41. Bernini, L.; Malguzzi, U.; Albertelli, P.; Monno, M. Hybrid prognostics to estimate cutting inserts remaining useful life based on direct wear observation. *Mech. Syst. Signal Process.* **2024**, *210*, 111163. [CrossRef]
42. Li, W.; Deng, L. A hybrid model-based prognostics approach for estimating remaining useful life of rolling bearings. *Meas. Sci. Technol.* **2023**, *34*, 105012. [CrossRef]
43. Elsherif, S.M.; Hafiz, B.; Makhlouf, M.; Farouk, O. A deep learning-based prognostic approach for predicting turbofan engine degradation and remaining useful life. *Sci. Rep.* **2025**, *15*, 26251. [CrossRef] [PubMed]
44. Chen, J.; Miao, J.; Yin, Q. A remaining useful life prediction method of aluminum electrolytic capacitor with adaptive degradation model selection. *Microelectron. Reliab.* **2024**, *162*, 115509. [CrossRef]
45. Wu, J.; Li, P.; Wang, L.; Huang, X.; Yang, J.; Du, M. The RUL prediction based on improved Wiener degradation model for wet friction components. *Meas. Sci. Technol.* **2024**, *35*, 076126. [CrossRef]
46. Guo, J.; Wang, Z.; Li, H.; Yang, Y.; Huang, C.; Yazdi, M.; Kang, H. A hybrid prognosis scheme for rolling bearings based on a novel health indicator and nonlinear Wiener process. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 110014. [CrossRef]
47. Li, C.; Cheng, J.; Zhu, H.; Wen, B.; Zhao, X.; Kang, W. Degradation modeling and remaining useful life prediction for electronic device under multiple stress influences. *Sci. Rep.* **2025**, *15*, 19117. [CrossRef] [PubMed]
48. Huang, D.; Bai, R.; Zhao, S.; Wen, P.; Wang, S.; Chen, S. Bayesian Neural Network Based Method of Remaining Useful Life Prediction and Uncertainty Quantification for Aircraft Engine. In Proceedings of the IEEE International Conference on Prognostics and Health Management (ICPHM), Detroit, MI, USA, 8–10 June 2020.
49. Thakkar, U.; Chaoui, H. Remaining Useful Life Prediction of an Aircraft Turbofan Engine Using Deep Layer Recurrent Neural Networks. *Actuators* **2022**, *11*, 67. [CrossRef]
50. Rossellini, R.; Barber, R.F.; Willett, R. Integrating Uncertainty Awareness into Conformalized Quantile Regression. In Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS), Valencia, Spain, 2–4 May 2024.
51. Xuan, Q.L.; Adhisantoso, Y.G.; Munderloh, M.; Ostermann, J. Uncertainty-aware remaining useful life prediction for predictive maintenance using deep learning. In Proceedings of the 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering (CIRP ICME'22), Online, 13–15 July 2022.
52. Nejjar, I.; Geissmann, F.; Zhao, M.; Taal, C.; Fink, O. Domain adaptation via alignment of operation profile for Remaining Useful Lifetime prediction. *Reliab. Eng. Syst. Saf.* **2024**, *242*, 109718. [CrossRef]
53. Wu, F.; Tang, J.; Jiang, Z.; Sun, Y.; Chen, Z.; Guo, B. The Remaining Useful Life Prediction Method of a Hydraulic Pump under Unknown Degradation Model with Limited Data. *Sensors* **2023**, *23*, 5931. [CrossRef] [PubMed]
54. Arafat, M.Y.; Hossain, M.J.; Alam, M.M. Machine learning scopes on microgrid predictive maintenance: Potential frameworks, challenges, and prospects. *Renew. Sustain. Energy Rev.* **2024**, *190*, 114088. [CrossRef]

55. Chen, X. A novel transformer-based DL model enhanced by position-sensitive attention and gated hierarchical LSTM for aero-engine RUL prediction. *Sci. Rep.* **2024**, *14*, 10061. [CrossRef] [PubMed]
56. Xie, S.; Xue, F.; Zhang, W.; Zhu, Z. Data-Driven Predictive Maintenance Policy Based on Dynamic Probability Distribution Prediction of Remaining Useful Life. *Machines* **2023**, *11*, 923. [CrossRef]
57. Mitici, M.; De Pater, I.; Barros, A.; Zeng, Z. Dynamic predictive maintenance for multiple components using data-driven probabilistic RUL prognostics: The case of turbofan engines. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109199. [CrossRef]
58. Lv, Y.; Zheng, P.; Yuan, J.; Cao, X. A Predictive Maintenance Strategy for Multi-Component Systems Based on Components' Remaining Useful Life Prediction. *Mathematics* **2023**, *11*, 3884. [CrossRef]
59. Ma, W.; Zhang, Q.; Xiahou, T.; Liu, Y.; Jia, X. Integrated selective maintenance and task assignment optimization for multi-state systems executing multiple missions. *Reliab. Eng. Syst. Saf.* **2023**, *237*, 109330. [CrossRef]
60. Bankati, W.R.; Boulon, L.; Jemei, S. Remaining useful life prognostic-based energy management strategy for multi-fuel cell stack systems in automotive applications. *Int. J. Hydrogen Energy* **2024**, *82*, 374–383. [CrossRef]
61. Dong, Q.; Pei, H.; Hu, C.; Zheng, J.; Du, D. Remaining Useful Life Prediction Method for Stochastic Degrading Devices Considering Predictive Maintenance. *Sensors* **2025**, *25*, 1218. [CrossRef] [PubMed]
62. Yang, X.; He, Y.; Zhou, D.; Zhou, D.; Zheng, X. Mission reliability–centered maintenance approach based on quality stochastic flow network for multistate manufacturing systems. *Eksploat. Niezawodn.* **2022**, *24*, 455–467. [CrossRef]