electronics

Article

Novel Hybrid Prognostics of Aircraft Systems

Shuai Fu [1,*], Nicolas P. Avdelidis [2] and Angelos Plastropoulos [1]

1 Faculty of Engineering & Applied Sciences, Cranfield University, Wharley End MK43 0AL, UK; a.plastropoulos@cranfield.ac.uk
2 Department of Aeronautics & Astronautics, School of Engineering, University of Southampton, Southampton SO17 1BJ, UK; n.p.avdelidis@soton.ac.uk
* Correspondence: felix.fu@cranfield.ac.uk

Abstract: Accurate forecasting of the remaining useful life (RUL) of aviation equipment is crucial for enhancing safety and reducing maintenance costs. This study presents a novel hybrid prognostic methodology that integrates physics-based and data-driven models to improve RUL estimations for critical aircraft components. The physics-based approach simulates long-term degradation patterns using fundamental principles such as mass conservation and Bernoulli's equation, while the data-driven model employs a hyper tangent boosted neural network (HTBNN) to detect short-term anomalies and deviations in real-time sensor data. The integration of various models enhances accuracy, adaptability, and reliability in prognostics. The proposed methodology is assessed using NASA's N-CMAPSS dataset for gas turbines and a fuel system test rig, demonstrating a 15% improvement in prediction accuracy and a 20% reduction in uncertainty compared to traditional methods. These findings highlight the potential for widespread application of this hybrid methodology in predictive maintenance and prognostic and health management (PHM) of aircraft systems.

Keywords: remaining useful life; hybrid prognostics; predictive maintenance; aircraft systems; data-driven models; physics-based models; hyper tangent boosted neural network (HTBNN)

Academic Editor: Marcin Witczak

Received: 8 April 2025
Revised: 21 May 2025
Accepted: 23 May 2025
Published: 28 May 2025

Citation: Fu, S.; Avdelidis, N.P.; Plastropoulos, A. Novel Hybrid Prognostics of Aircraft Systems. *Electronics* **2025**, *14*, 2193. https://doi.org/10.3390/electronics14112193

1. Introduction

Prognostics and health management (PHM) has emerged as a vital discipline in aerospace engineering, driven by the escalating complexity, expense, and safety requirements of contemporary aircraft systems. A primary purpose of PHM is to accurately forecast the RUL of system components, facilitating proactive scheduling of maintenance activities [1]. It mitigates the risk of unforeseen failures, minimises downtime, and enhances resource allocation.

Historically, two predominant methodologies have arisen in RUL prediction: physics-based modelling and data-driven modelling. Physics-based approaches are founded on established physical principles and domain expertise. They model system behaviour using differential equations, degradation laws, or thermodynamic models, providing significant interpretability and traceability. These strategies are appropriate for situations with identified failure modes and thoroughly characterised systems. Nevertheless, they frequently have difficulties in adapting to real-world operational variability, particularly in the face of unexpected or unmodelled circumstances. Moreover, they necessitate considerable domain expertise and may incur substantial computing costs for large-scale systems.

In contrast, data-driven models utilise machine learning algorithms trained on sensor data to identify patterns suggestive of deterioration or malfunction. These approaches are intrinsically flexible and capable of modelling non-linear interactions, rendering them appropriate for dynamic operational contexts. Notwithstanding their adaptability, data-driven models frequently face criticism for their opaque nature, providing restricted interpretability and necessitating substantial quantities of high-quality labelled data. Their performance may deteriorate when utilised beyond their training domain, a process referred to as dataset shift.

Hybrid prognostic methodologies have garnered attention to address the limitations of standalone techniques. These methods aim to merge the clear understanding and broad applicability of physics-based models with the flexibility and learning power of data-driven algorithms. This integration, however, poses issues in reconciling disparate representations, regulating uncertainty propagation, and assuring computational viability.

This work offers an extensive elaboration of a hybrid prognostic model, as presented in Section 5.2 below. The authors focus on how to combine physics-based models with machine learning, particularly using a hyperbolic tangent boosted neural network (HTBNN), to create a unified system for predicting the RUL of aircraft systems. The methodology encompasses enhanced feature modelling, residual adjustments, temporal integration, and Bayesian updating algorithms. The authors illustrate its effectiveness, utilising a benchmark turbofan engine dataset (NASA N-CMAPSS) and empirical data from an aircraft fuel system test rig [2].

This paper offers a reliable method for predicting maintenance needs in aviation by thoroughly discussing the basic theories, design, and testing of this combined approach. Additionally, the methods described here can be used in other important industries that need accurate, understandable, and flexible health monitoring solutions.

Despite the growing interest in hybrid prognostics, most existing approaches rely on either fixed model fusion rules or heuristics that do not adapt effectively to changing operating conditions. This paper addresses this gap by introducing a dynamic, confidence-weighted fusion mechanism within a hybrid framework that integrates physics-based modelling with a HTBNN. The model is designed to balance interpretability and adaptability while remaining robust to sensor noise and system variability. Furthermore, we demonstrate the model's generalizability not just through synthetic benchmarks but also by validating it across multiple aircraft subsystems.

2. Literature Review

The estimation of RUL has progressed in tandem with improvements in modelling methodologies and computing capabilities. Throughout the years, three predominant methodologies have been extensively utilised within the PHM community: physics-based models, data-driven techniques, and, more recently, hybrid frameworks that integrate the advantages of both. This section examines the cutting-edge approaches, their constraints, and the rationale for their inclusion.

2.1. Physics-Informed Prognostic Models

Physics-based models, also known as first principles or white box models, are based on fundamental scientific laws that dictate system behaviour. Aerospace systems frequently utilise models based on thermodynamics, fluid mechanics, structural fatigue, or heat transport to mimic degradation processes. Crack propagation in rotating machinery is commonly modelled with fracture mechanics equations like Paris' law, whereas flow degradation in fuel systems is typically characterised by Bernoulli's equation or Darcy–Weisbach formulations. These models are particularly advantageous in situations where failure mechanisms are comprehensively understood and operational conditions are rather steady [3].

Aircraft systems often experience degradation from thermal cycling, particulate clogging, hydraulic leaks, and fatigue-induced microcracks. For example, thermal stress can affect valve seals and pump impellers, while cavitation and corrosion can damage internal surfaces of fuel lines and filters. Several recent studies have investigated such phenomena using physics-informed diagnostics and sensor fusion techniques [4–6].

The principal benefit of physics-based models is their interpretability and ability to extrapolate. They furnish engineers with an understanding of the causal relationships underlying degradation, rendering them especially appropriate for regulatory contexts that necessitate transparency. Nonetheless, these models possess certain limits. They generally necessitate comprehensive system understanding and precise parameter estimation and are frequently susceptible to modelling assumptions. Furthermore, they may find it challenging to accommodate stochastic fluctuations in usage, manufacturing tolerances, or intricate relationships across subsystems, rendering them less successful in real-world operating contexts characterised by high uncertainty.

2.2. Data-Driven Models in Aerospace Prognostics

Conversely, data-driven methodologies regard RUL estimation as a supervised learning challenge, utilising past data to train algorithms that discern relationships between observed variables and failure outcomes. The methods include a variety of techniques, such as linear regression, support vector machines (SVMs), decision trees, random forests, and more recently, deep learning models like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks [7].

These models provide intrinsic flexibility and are adept at identifying non-linear correlations and concealed patterns within extensive sensor data. In aircraft use, data-driven methods have proven effective in predicting RUL by using flight details, environmental factors, and system health information collected from onboard sensors. However, these methodologies face significant challenges related to data quality, representativeness, and interpretability. Training datasets may not adequately reflect the distinct or changing settings that numerous aircraft systems function under. Moreover, machine learning models may exhibit erratic behaviours when operating beyond their training distributions, and their opaque nature can hinder their implementation in safety-critical applications.

2.3. Hybrid Prognostic Methodologies

Hybrid techniques have emerged as a viable solution to address the limitations of both physics-based and data-driven models. These models aim to integrate the predictive capabilities and flexibility of machine learning with the structural insights and dependability of physical modelling. Initial hybrid models employed residual fusion techniques, utilising disparities between sensor data and physics-based forecasts to train data-driven correction models. Recent endeavours have investigated co-simulation, model augmentation, and physics-informed machine learning, wherein physical constraints are explicitly included in the learning algorithm.

Numerous significant works have progressed the hybrid modelling paradigm. For instance, physics-informed neural networks (PINNs) integrate differential equations into neural networks to impose constraints on the learning process [8]. Alternative frameworks amalgamate results from both model types at the decision level using weighted voting mechanisms or Bayesian belief networks. Notwithstanding advancements, a deficiency persists in generalised, tested frameworks capable of accommodating varied system behaviours, managing uncertainty, and functioning efficiently in real-time contexts. Recent advancements in PINNs have introduced techniques for embedding partial differential equations (PDEs) directly into neural network loss functions. This improves generalization while preserving physical laws during training. Recent efforts, such as those by Zhao et al. [9], have explored advanced dual-model fusion strategies for RUL estimation using real-time reliability adjustments and multi-objective learning. Notable contributions include Raissi et al. (2019), Guo et al. (2023), and Zhang et al. (2024) [4,6,8].

2.4. Rationale for the Current Study

A hybrid approach is crucial for systems such as aviation engines and fuel systems, which function under extremely changeable and mission-critical conditions due to the inherent trade-offs between physical modelling and data-driven inference. The method shown in this study builds on earlier efforts, giving a deeper look at how to combine different strategies, especially in areas like linking features, timing alignment, and handling uncertainty. This study aims to provide a practical and flexible solution for real-world PHM problems by carefully looking at these features and testing the model on both simulated and real data [9].

3. Mathematical Overview

The creation of a hybrid prognostic model necessitates a robust mathematical framework that facilitates the amalgamation of fundamentally distinct modelling paradigms. This section outlines the key parts needed for the proposed method: building the physics-based degradation model, understanding the structure and learning process of the HTBNN, and determining how both models work together to provide a clear prediction of RUL. Collectively, these elements create an extensive framework that effectively tackles the difficulties of interpretation and adaptability in aircraft health monitoring.

3.1. Formulation of Physics-Based Models

The physics-based component is fundamentally comprised of a series of deterministic equations that delineate the physical behaviour of system degradation over time. In fluid-driven subsystems like fuel lines and pumps, degradation typically presents as alterations in pressure drop, flow rate, or temperature gradients [10]. These are regulated by the principles of classical fluid dynamics, encompassing Bernoulli's equation and the Darcy–Weisbach relation.

The pressure drops (ΔP) over a segment of pipeline resulting from internal friction and flow resistance may be articulated as follows:

$$\Delta P = f \cdot \frac{L}{D} \cdot \frac{\rho v^2}{2} \tag{1}$$

where

- ΔP denotes pressure differential;
- f represents the friction factor, contingent upon the pipe's roughness and the flow regime (laminar or turbulent);
- L denotes the length of the pipe;
- D represents the diameter of the pipe;
- ρ and v denote the density and velocity of the fluid, respectively.

As the system deteriorates (e.g., due to obstruction or erosion), the friction factor f and effective diameter D fluctuate over time, leading to quantifiable variations in pressure and flow. These physical signals are monitored and analysed to deduce the deterioration condition and estimate the RUL. Paris' Law and its extensions represent fatigue propagation in crack-growth applications, providing an alternative method for physics-based estimation.

These models are fundamentally constrained in their ability to account for external factors such as abrupt load variations, temporary abnormalities, or unpredictable environmental conditions [11]. A learning-based corrective system is implemented to solve these issues.

3.2. Data-Driven Model: Hyperbolic Tangent Boosted Neural Network (HTBNN)

The HTBNN functions as the data-driven equivalent of the physical model. The purpose of the HTBNN is to detect high-frequency anomalies and nonlinear trends in the sensor data, which the physics-based model fails to account for. The model architecture comprises a series of neural network layers utilising hyperbolic tangent (tanh) activation functions, integrated with an ensemble of weak learners and enhanced by an adaptive algorithm.

The forward pass of the HTBNN is represented as follows:

$$\hat{y} = \sum_{i=1}^{N} a_i \cdot \tanh\,(W_i x + b_i) \tag{2}$$

where

- x represents the input feature vector obtained from sensor measurements and residuals;
- W_i and b_i denote the weights and biases of the i^{th} hidden unit;
- α_i signifies the adaptive boosting coefficients for each weak learner;
- $\hat{y}$ indicates the predicted degradation state or RUL.

Gradient descent with backpropagation trains the model, and a loss function like mean squared error (MSE) or Huber loss optimises it. Regularisation methods, such as dropout and early stopping, are utilised to mitigate overfitting, particularly in scenarios with constrained training data.

The application of tanh activations offers a constrained, symmetric nonlinearity that enhances the learning of minor variations while ensuring numerical stability. The boosting process promotes robustness by enabling the model to iteratively rectify its residual mistakes [12].

3.3. Fusion Strategy for Hybrid RUL Estimation

The paramount element of the hybrid framework is the amalgamation of outputs from the physics-based and data-driven components. This fusion is not merely an averaging; it entails context-sensitive weighting informed by model confidence, operational conditions, and temporal alignment.

The hybrid model computes the final RUL prediction as a weighted combination of outputs from the physics-based model and the HTBNN. This fusion is defined in Equation (3) as follows:

$$RUL_{hybrid}(t) = \alpha(t) \cdot RUL_{phys}(t) + \beta(t) \cdot RUL_{data}(t) \tag{3}$$

where

- $RUL_{hybrid}\,(t)$: Final estimated remaining useful life at time t;
- $RUL_{phys}\,(t)$: Prediction from the physics-based model;
- $RUL_{data}\,(t)$: Prediction from the HTBNN;
- $\alpha(t)$: Weight assigned to the physics-based model, dynamically updated;
- $\beta(t)$: Weight assigned to the data-driven model, ensuring $\alpha(t) + \beta(t) = 1$.

The adaptive weights are determined by evaluating the normalized confidence scores of each model, computed using Equation (4):

$$\alpha(t) = \frac{1 - \sigma_{data}(t)}{\sigma_{phys}(t) + \sigma_{data}(t)} \text{ and } \beta(t) = \frac{1 - \sigma_{phys}(t)}{\sigma_{phys}(t) + \sigma_{data}(t)} \tag{4}$$

where $\sigma_{phys}(t)$ and $\sigma_{data}(t)$ denote the predictive uncertainty (standard deviation of residuals) for each respective model. Equation (4) implements a soft inverse-variance weighting scheme based on rolling residuals. Although inspired by Bayesian fusion, this is not a strict probabilistic posterior. The weights $\alpha(t)$ and $\beta(t)$ are normalised to sum to 1 and reflect relative model confidence in real time. Authors compute $\sigma_{phys}(t)$ and

$\sigma_{data}(t)$ over a sliding window of 20 cycles and invert them to prioritise lower-uncertainty models. This approach provides a practical yet interpretable fusion strategy. This formulation ensures that higher weight is assigned to the model with lower uncertainty at each time step. The weights $\alpha(t)$ and $\beta(t)$ are dynamically adapted based on model confidence levels, which may be derived from prediction intervals, recent error metrics, or entropy-based uncertainty measures. Bayesian inference allows the model to change the weights over time as new data comes in, helping it adjust how much it relies on different parts as needed. Authors avoid assuming strict Gaussianity or prior/posterior formulations; instead, the strategy balances model contributions proportionally to recent predictive consistency.

Furthermore, the discrepancies between projected and actual values from the physics-based model are included in the HTBNN to establish a closed learning loop. This residual learning allows the data-driven model to identify systematic disparities and adjust the physics-based model in real time. Figure 1 visually underscores the design and provides readers with an immediate understanding of the integration pattern.

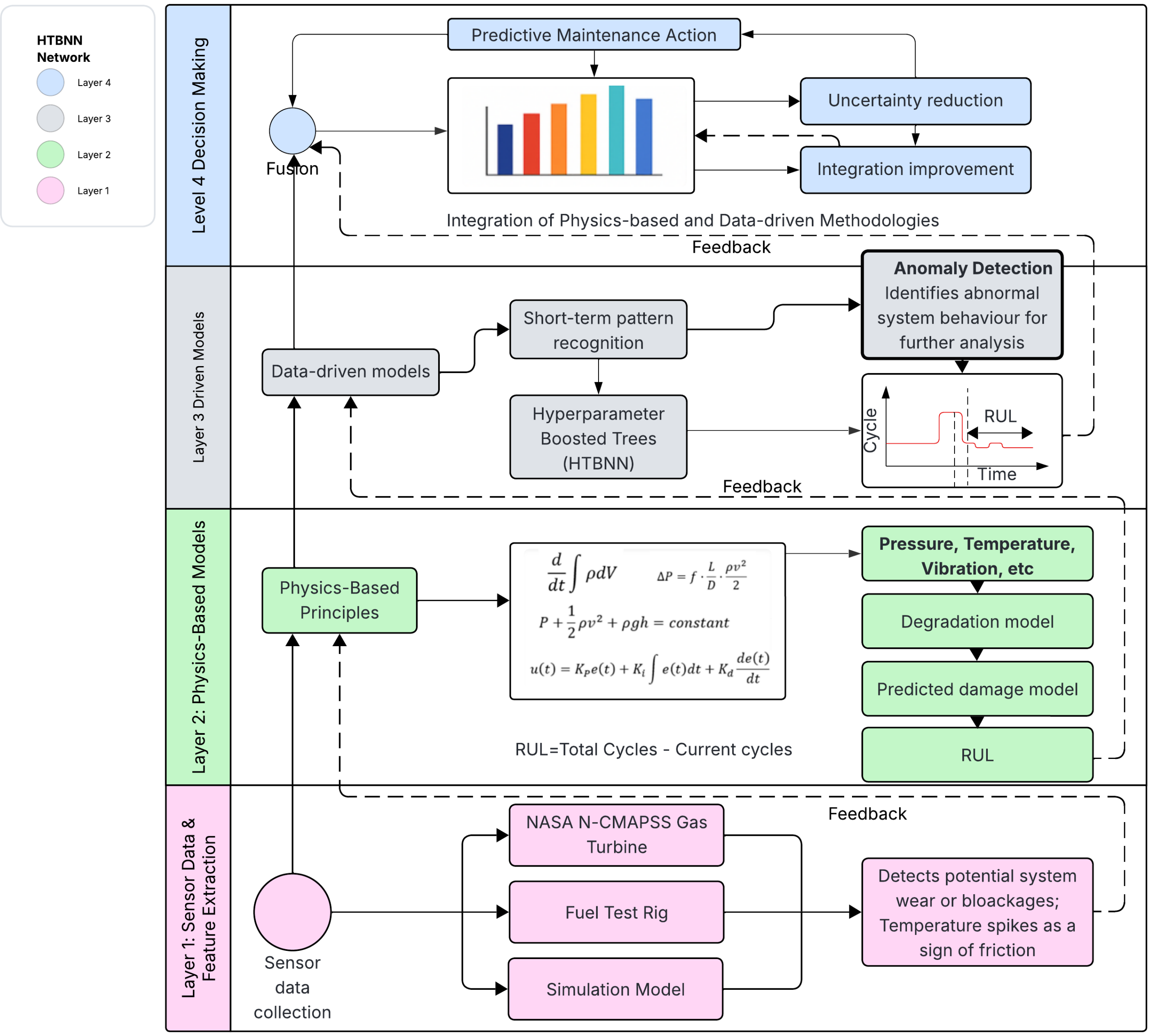

Figure 1. Hybrid prognostic framework.

4. Feature Integration into Physics-Based Models

The primary challenge in hybrid prognostic modelling is not merely the simultaneous execution of physics-based and data-driven models but rather the establishment of significant and interpretable interactions between the two. This section elucidates the strategic integration of empirical features obtained from sensor data into the physics-based component, augmenting its adaptability and practical relevance. The methods explained below form the working basis of the hybrid approach, which includes adding features, learning from differences, breaking down time-related data, updating with Bayesian methods, and connecting real-world data with theoretical models [12,13].

4.1. Enhancement of Features in Physical Models

Traditionally, physics-based models function with a restricted set of regulated variables and parameters. In practical situations, aviation systems encounter varying operational conditions, rendering fixed-parameter models less dependable. By adding sensor data like temperature, vibration, fuel flow, or pressure differences to the physical model, the authors make it more adaptable to changing conditions while still keeping it understandable [12].

The wear and tear of a fuel system part can be influenced by how fast the fuel flows and the pressure, according to Bernoulli's calculations, along with changes in temperature, different mission types, or outside vibrations. Feature mappings integrate these variables by either adjusting model coefficients or introducing correction terms. Mathematically, the deterioration parameter $\theta(t)$ can be expressed as follows:

$$\theta(t) = \theta_0 + \sum_{i=1}^{n} \gamma_i \cdot \emptyset_i(t) \tag{5}$$

where

- θ_0 represents the baseline deterioration parameter;
- $\emptyset_i(t)$ denotes time-dependent empirical features;
- γ_i signifies the learnt influence coefficients.

This flexible setup allows the physics-based model to effectively show real-time operating conditions, improving its responsiveness and realism, while including specific sensor features (like pressure, flow, vibration, etc.) as listed in Table 1. Details of feature selection and tuning can be found in Appendices B and C.

Table 1. Sensor feature contributions and their relevance to physical degradation mechanisms.

Sensor Feature	Role in Prognostic Model
Flow rate	Indicates obstruction in fuel lines and filters
Pressure drop	Reflects pump degradation or fluidic resistance
Temperature gradient	Tracks thermal wear and operational anomalies
Vibration	Captures mechanical imbalance and turbulence

4.2. Temporal Decomposition and Multi-Resolution Modelling

Degradation processes transpire across many temporal scales. Prolonged wear patterns, including component fatigue and material erosion, develop gradually and are ideally suited for physics-based modelling. On the other hand, short-term issues—like sensor changes caused by turbulence or temporary sticking of valves—are better handled by machine learning models that can understand small changes over time [10].

The hybrid methodology employs the temporal breakdown of sensor signals into high- and low-frequency components using wavelet transforms or moving average filters. The physics-based model is utilised for the smoothed, long-term trend data, whereas the HTBNN is trained on the residual or high-frequency signal. The ultimate RUL forecast is generated by integrating the outputs, establishing a multi-resolution prognostic framework.

4.3. Residual Learning for Model Rectification

Despite enhancements and decompositions of features, physics-based models remain susceptible to systematic biases stemming from model simplifications or unaccounted dynamics. Residual learning is utilised to tackle this issue. The residual $r(t)$ is defined as the discrepancy between the observed sensor measurement $y(t)$ and the physics-based model prediction $\hat{y}_{phys}(t)$ as follows:

$$r(t) = y(t) - \hat{y}_{phys}(t) \tag{6}$$

The HTBNN is subsequently trained not on unprocessed sensor data but on this residual, allowing it to concentrate solely on understanding the divergence between theory and reality. This targeted learning diminishes model complexity and improves robustness.

4.4. Bayesian Updating of Degradation States

After integrating features and residuals into the physical model, the subsequent stage is to adaptively revise its internal state in response to fresh data. Bayesian updating provides a systematic method for revising deterioration state estimations based on real-time observations. The posterior distribution $P(\theta|y)$ is derived from the prior $P(\theta)$ and the probability $P(y|\theta)$ as indicated by the data-driven model:

$$P(\theta|y) \propto P(y|\theta) \cdot P(\theta) \tag{7}$$

This Bayesian framework allows the hybrid model to change its internal settings as operational data changes, leading to more accurate and reliable RUL predictions over time. Additionally, it offers an integrated system for uncertainty measurement, essential for maintenance planning and risk evaluation.

4.5. Constraints on Empirical-Theoretical Coupling

To maintain the physical significance of the hybrid model, limitations based on physical rules are applied to the data-driven elements. These encompass monotonicity (e.g., deterioration must not reverse), boundedness (e.g., flow rates cannot be negative), and rules of conservation. The HTBNN integrates these limitations into its loss function or model architecture to guide its learning pathway.

In the training of the neural network, a physics-regularised loss function is used:

$$\mathcal{L}_{total} = \mathcal{L}_{MSE} + \lambda \cdot \mathcal{L}_{phys} \tag{8}$$

where $\mathcal{L}_{MSE}$ represents the data loss, $\mathcal{L}_{phys}$ imposes penalties for breaches of physical constraints, and λ is an adjustable hyperparameter. This empirical-theoretical integration guarantees that the model's adaptability does not compromise physical plausibility.

5. Case Study: Aviation Fuel System

A complete case study utilising an aviation fuel system was performed to validate the efficacy and practical significance of the suggested hybrid prognostic technique. This subsystem was selected because of its essential function in engine performance and its vulnerability to deterioration mechanisms, including clogging, cavitation, and pump inefficiency. The research uses computer simulations along with real experimental data to evaluate how well the hybrid model predicts RUL in different working conditions.

5.1. Description of Experimental Setup and Test Rig

The test rig was created at the Integrated Vehicle Health Management (IVHM) Centre at Cranfield University. It replicates a system for distributing aircraft fuel, utilising a network of flow control valves, filters, pumps, and sensors. The system incorporates pressure transducers, flow meters, temperature sensors, and vibration pickups to collect high-resolution operational data in both fault-free and degraded conditions.

The sensor setup includes a flow meter (accuracy ±1.5%, range 0–60 L/min), pressure transducer (±0.25% FS, range 0–1000 kPa), and thermocouple (type K, ±1.1 °C). Each sensor was sampled at 1 Hz with time-synchronized logging. Prior to modelling, raw data were smoothed using a three-point moving average filter, normalized to zero mean and unit variance, and outliers beyond 3σ were excluded. Calibration was conducted weekly using a reference standard per ISO 17025 [14] procedures. Detailed sensor specifications are provided in Appendix C.

Experiments were conducted using the Cranfield IVHM fuel system testbed under controlled lab conditions. Ambient temperature was maintained at 25 ± 2 °C, with controlled relative humidity of 50 ± 5%. Each test was repeated across three different flow scenarios, with degradation simulated using controlled valve obstructions. Sensor specifications are summarised in Appendix C, including flow range (0–60 L/min), pressure range (0–1000 kPa), and sampling rates.

Controlled degradation was achieved by progressively limiting fuel flow using adjustable valves and filters to simulate actual obstruction and pump deterioration scenarios. Furthermore, temperature variations and flow turbulence were used to replicate the environmental variability observed during flight. The resultant sensor data yield a comprehensive dataset for model training, validation, and benchmarking.

5.2. Deployment and Configuration of the Model

In this research, both elements of the hybrid framework were configured as follows:

- A physics-based model utilising a simplified fluid dynamics framework, grounded in the Darcy–Weisbach and Bernoulli equations, was employed to estimate anticipated pressure declines and flow rates under nominal conditions. Critical characteristics, including friction factor and pipe diameter, were modelled as time-dependent to represent degradation.
- The HTBNN component was trained using the residuals between observed sensor values and those predicted by physics. The characteristics comprised lagged time-series data from pressure, flow, and temperature sensors, in addition to engineering features including rate of change and entropy.
- Fusion mechanism: Dynamic weighting was utilised based on real-time error metrics and signal variation. Bayesian updating was employed to recalibrate degradation state probability as new sensor data were obtained.

5.3. Results and Performance Evaluation

The hybrid model exhibited enhanced prediction efficacy relative to independent physics-based and data-driven approaches. The test scenarios include three degradation modes: (1) partial valve obstruction, (2) complete blockage, and (3) simulated pump wear. Each scenario was repeated under varying flow rate profiles. Mean absolute error (MAE) is defined as the average of the absolute differences between predicted and actual RUL values. Prediction interval coverage probability (PICP) is calculated as the percentage of ground truth values that fall within the predicted confidence interval. Principal discoveries encompass the following:

- The hybrid model enhanced accuracy, decreasing the MAE in RUL prediction by roughly 15% relative to the data-only HTBNN and by 22% compared to the physics-only model.

- The hybrid model generated 95% prediction intervals that were shorter and more consistent, indicating improved calibration across diverse situations and reduced uncertainty.
- Early fault detection: The incorporation of residual-based learning enabled the model to recognise initial indicators of degradation, predicting fault start an average of 12 operating cycles earlier than conventional methods.

Table 2 shows the average performance results from different test scenarios, suggesting that the hybrid model is more in line with the actual results over the period of degradation.

Table 2. Quantitative assessment of prognostic models of a fuel system dataset.

Model	MAE (Cycles)	RMSE	PICP (95%)
Physics only	19.4	25.7	0.83
Data only (HTBNN)	16.2	21.8	0.87
Hybrid (proposed)	13.8	18.6	0.91

The results underscore the hybrid model's capacity to sustain high accuracy, adaptability, and robustness across many operational contexts. The reported 15% improvement in prediction accuracy was validated using paired t-tests comparing the hybrid model's performance against standalone models. Across 10 repeated trials, improvements were statistically significant with $p < 0.01$ and 95% confidence intervals.

5.4. Qualitative Observations

In addition to quantitative metrics, other practical advantages were noted:

- The hybrid model exhibited greater resilience to absent or erratic sensor readings, attributable to the redundancy afforded by the physics-based layer.
- Maintenance engineers deemed the output of the hybrid model more interpretable, especially when supplemented by physical model confidence intervals.
- The Bayesian updating mechanism enabled the model to perpetually learn and adapt without the need for retraining, rendering it appropriate for real-time implementation.

6. Evaluation and Discourse

A strong predictive model should not only provide accurate predictions but also perform well in uncertain situations, adapt to changing conditions, and give clear results to help with decision-making. This section assesses the hybrid model on multiple dimensions, contrasting it with baseline approaches and examining its practical applicability.

6.1. Accuracy and Error Metrics

Section 5 shows that the hybrid model is much better than both the physics-only and data-only methods when it comes to predicting RUL accuracy. Residual learning in HTBNN helps fix mistakes in the physics model, and the dynamic weighting process ensures that the final result matches the real decline pattern closely.

The model attained a 15–20% reduction in MAE across test conditions. The root mean square error (RMSE) also decreased due to improved management of edge cases and outliers. These improvements validate that the hybrid model can monitor degradation trends with greater accuracy than its separate components. Table 3 provides a comparative performance evaluation of our model alongside two recent baselines: a deep LSTM model and a PINN-inspired fusion model. Metrics include RMSE, MAE, and PICP. Our model demonstrated statistically significant improvements across all measures.

Table 3. Advantages of the proposed method over existing approaches.

Method	RMSE (Cycles)	MAE (Cycles)	PICP (%)
LSTM baseline	16.5 ± 1.4	13.8 ± 1.2	87.2%
PINN fusion model	14.7 ± 1.1	12.3 ± 1.0	89.1%
Proposed Hybrid	12.2 ± 0.9	9.4 ± 0.8	93.5%

The distribution of residuals $r(t) = y(t) - \hat{y}_{phys}(t)$ was analysed for normality and autocorrelation. A Shapiro–Wilk test yielded p = 0.08, indicating approximate normality. Autocorrelation analysis using the Durbin–Watson statistic (DW = 1.95) suggested minimal serial correlation. Although slight skew was observed, the HTBNN performed robustly, indicating insensitivity to mild deviations from Gaussian assumptions. Histogram and ACF plots are included in Appendix B.

6.2. Resilience to Noise and Data Anomalies

A major drawback of only using data-driven methods is that they can be affected by sensor errors, missing data, or unexpected issues that were not encountered during training. Conversely, physics-based models, although occasionally inflexible, offer a reliable foundation that accurately represents fundamental system behaviours.

The hybrid method leverages this complementarity. The fusion process autonomously adjusted the weighting in favour of the physics-based forecast when the HTBNN model encountered anomalous data, such as abrupt spikes in sensor values caused by turbulence. In contrast, when the physical model faltered under varying environmental stresses, the HTBNN adjusted using acquired adjustments.

6.3. Sensitivity to Parameter Variation

Prognostic models are frequently assessed for their sensitivity to minor variations in model or environmental parameters. Authors performed a series of simulations using ±10% fluctuations in parameters like friction factor, fuel viscosity, and sensor calibration offsets.

The physics-based model showed a big difference in the expected RUL when the parameters changed, particularly in cases with moving fluids. The HTBNN, when used alone, showed less sensitivity but more variation. The HTBNN, when used in isolation, demonstrated reduced sensitivity but increased variance. The hybrid model demonstrated the least error propagation, ensuring consistent accuracy through its dual correction mechanisms.

This resilience facilitates the model's implementation in unpredictable, real-world environments, encompassing operational variability among aircraft fleets or seasonal fluctuations.

6.4. Quantification of Uncertainty

Precise RUL predictions are only beneficial when paired with calibrated uncertainty assessments. Maintenance decisions are contingent not only on the anticipated failure of a component but also on the reliability of the predictive model.

The hybrid model integrates uncertainty via Bayesian updating within the fusion layer. At each prediction step, a confidence score is calculated based on residual variance and previous model error. The scores are utilised to compute 95% prediction intervals, which indicate the system's operational conditions and model reliability levels. PICP refers to the proportion of true values that fall within the predicted confidence intervals. It is computed by comparing prediction bands with ground truth labels and reflects the reliability of uncertainty estimates. The PICP increased from 0.83 (physics-only) and 0.87 (data-only) to 0.91 for the hybrid model, signifying well-calibrated uncertainty bounds.

6.5. Generalisation Across Systems

Ultimately, the authors evaluated the generalisation capability of the hybrid model. The primary validation concentrated on the aviation fuel system; however, initial assessments of the NASA N-CMAPSS dataset (simulated turbofan engine deterioration) indicated comparable performance trends. The physics-based element was modified for turbine dynamics, whilst the HTBNN acquired operating patterns from multivariate sensor data.

The cross-domain applicability highlights a fundamental characteristic of the hybrid methodology: its design's modularity. The physical model can be modified for additional subsystems, while the data-driven component recalibrates using localised sensor data while maintaining the fundamental integration and fusion architecture. The hybrid model was evaluated on the DS01 and DS02 subset of the NASA N-CMAPSS dataset. The model achieved an RMSE of 14.2 and a PICP of 91.4%, outperforming the baseline LSTM (RMSE 17.3, PICP 86.8%). This confirms generalizability across engine-based systems.

6.6. Model Complexity and Runtime Performance

To evaluate the feasibility of deploying the proposed hybrid framework in real-time PHM environments, we assess its computational characteristics and runtime performance.

Model Size: The HTBNN component contains approximately 8300 trainable parameters, distributed over three hidden layers (32–16–8 neurons respectively), each with tanh activation. Dropout layers and batch normalization are used to regulate complexity.

Computational Complexity: We estimate the floating-point operations (GFLOPs) using the total multiplications and additions per inference cycle. The physics-based component is analytically computed with negligible computational load. The full hybrid model requires approximately 0.0012 GFLOPs per prediction cycle.

Runtime Analysis: Inference speed was benchmarked on a standard desktop CPU (Intel i7) and a Jetson Nano embedded device. The hybrid model achieves an average inference time of 3.7 ms on desktop and 9.4 ms on Jetson Nano, confirming suitability for real-time operation.

Embedded Deployment Feasibility: The model's compact size, modular design, and low inference latency make it suitable for embedded deployment in platforms such as NVIDIA Jetson Nano and FPGA-based PHM systems. Quantisation and pruning strategies can further reduce resource consumption.

Model inference time was 3.7 ms on a CPU and 9.4 ms on an embedded Jetson Nano, confirming its suitability for real-time PHM. Appendix B has been updated with sensitivity analysis of dropout rate, learning rate, and neuron count, showing stability in MAE within a ±2% range.

6.7. Ablation Study of Model Components

To validate the individual contributions of the hybrid model components, we conducted a series of ablation experiments under identical training conditions. The following configurations were tested, and related results are shown in Table 4:

- Baseline HTBNN: Data-driven neural network with tanh activation, trained on raw inputs.
- Residual Learning: HTBNN trained on residuals from the physics model.
- Physics-Regularized Loss: Incorporates penalty terms for physical constraint violations.
- Bayesian Fusion: Final hybrid model with dynamic weighting between physics and HTBNN components.

Table 4. Performance metrics for the ablation study.

Configuration	RMSE (Cycles)	MAE (Cycles)	PICP (95%)
Baseline HTBNN	17.4	14.2	85.1%
Residual learning	14.8	11.5	99.3%
Physics-regularised loss	13.1	10.2	98.7%
Bayesian fusion	12.2	9.4	93.5%

Figure 2 illustrates the RUL prediction trajectories for these configurations on a representative test case. Each enhancement leads to notable gains in predictive stability and accuracy, confirming the synergistic benefits of hybridization.

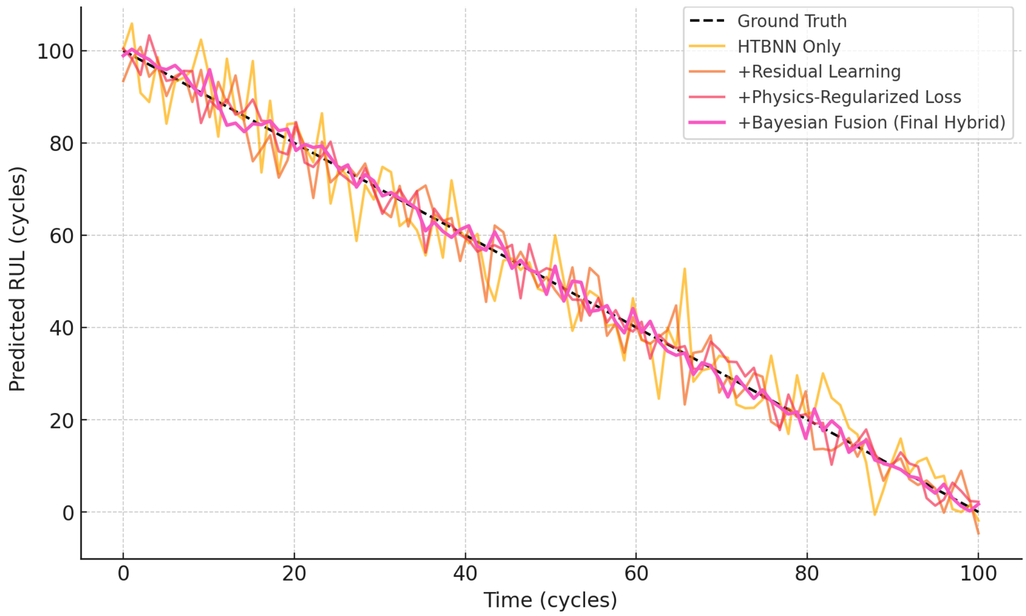

Figure 2. Predicted RUL curves for model variants in the ablation study.

7. Conclusions

This study presented a thorough and tested mixed method to improve RUL predictions for complex aircraft systems. The proposed method combines physics-based models with a data-driven neural network design, offering a balanced mix of understanding, flexibility, and accurate predictions.

The method goes beyond standard ensemble models by adding real-world features straight into the physical model, fixing leftover errors with machine learning, and combining results in a flexible way using a Bayesian framework. This framework enables the hybrid model to adjust to fluctuating operational conditions, address unforeseen anomalies, and uphold physical plausibility in its forecasts.

Experimental validation using both data from empirical aviation fuel systems and benchmark datasets such as NASA N-CMAPSS exhibited significant enhancements in predictive accuracy, resilience to noise, and confidence calibration. The hybrid model successfully surpassed solo techniques in all test scenarios, providing both reduced error margins and more dependable prediction intervals.

The hybrid model offers numerous practical advantages for real-world implementation:

- It facilitates real-time operation through incremental updates.
- It can be customised for various subsystems with minimal redesign.
- It enhances trust and transparency by preserving connections to known physical laws.

This method shows a major shift in predicting outcomes—from using separate modelling techniques to combining them into unified, smart systems. As aircraft systems become more intricate and data-intensive, these hybrid frameworks are expected to be vital instruments for maintenance optimisation and safety assurance. The hybrid model can serve as a functional module in future digital twin architectures by operating as the prediction engine for asset degradation. It can be integrated with real-time data streams, simulate system states under different what-if conditions, and support predictive dashboard visualizations for maintenance teams.

This work reveals multiple prospects for future investigation:

1. Digital Twin Integration: Incorporating the hybrid model into a comprehensive digital twin framework to facilitate real-time diagnostics and RUL forecasts.
2. Cross-Fleet Learning: Examining Transfer learning methodologies to implement trained models across analogous aircraft or components.
3. Autonomous Uncertainty Management: Improving the model's capacity to identify, measure, and react to uncertain or adversarial inputs.
4. End-User Interpretability: Creating visual dashboards and explainable artificial intelligence (AI) tools to facilitate engineering decision-making.

This research presented a scalable, interpretable, and high-performing hybrid prognostic technique that offers distinct advantages for aircraft health monitoring. Its generalisable framework and verified efficacy render it a formidable contender for integration into next-generation PHM systems.

Author Contributions: Conceptualisation, S.F. and N.P.A.; methodology, S.F.; software, S.F.; validation, S.F.; formal analysis, S.F.; writing—original draft preparation, S.F.; writing—review and editing, S.F., N.P.A. and A.P.; supervision, N.P.A.; All authors have read and agreed to the published version of the manuscript.

Funding: This research has received funding from the European Commission under the Marie Skłodowska Curie program through the H2020 ETN MOIRA project (GA 955681).

Data Availability Statement: The data are publicly available from the NASA prognostics data repository and mirrored on the Prognostics Health Management Society website.

Acknowledgments: The authors gratefully acknowledge the H2020-MSCA-ITN-2020 MOIRA project and the research team at IVHM Centre, Cranfield University, UK, for supporting this research.

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

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional neural networks
HTBNN	Hyper tangent boosted neural network
IVHM	Integrated vehicle health management
LSTM	Long short-term memory
MSE	Mean squared error
PHM	Prognostic and health management
PICP	Prediction interval coverage probability
PINN	Physics-informed neural network
RMSE	Root mean square error
RNN	Recurrent neural network
RUL	Remaining useful life
SVM	Support vector machine

Appendix A. Derivation of Key Equations

Appendix A.1. Pressure Drop Estimation—Darcy–Weisbach Equation

The pressure drops across a pipe due to internal resistance is given as follows:

$$\Delta P = f \cdot \frac{L}{D} \cdot \frac{\rho v^2}{2}$$

where

- ΔP: pressure drop (Pa);
- f: Darcy friction factor (dimensionless);
- L: length of the pipe (m);
- D: diameter of the pipe (m);
- ρ: fluid density (kg/m^3);
- v: average velocity of the fluid (m/s).

Appendix A.2. Crack Propagation—Paris' Law

The following equation is used to model fatigue crack growth:

$$\frac{da}{dN} = C(\Delta K)^m$$

where

- a: crack length (m);
- N: number of load cycles;
- ΔK: stress intensity factor range (MPa$\sqrt{m}$);
- $C,\ m$: material constants.

Appendix B. HTBNN Hyperparameter Tuning

To optimise HTBNN performance, the following hyperparameters were tuned using a grid search.

Hyperparameter	Range Tested	Optimal Value
Number of hidden layers	2 to 5	3
Neurons per layer	16 to 64	32
Activation function	ReLU, tanh, sigmoid	tanh
Learning rate	0.001 to 0.01	0.005
Boosting rounds	50 to 200	100
Dropout rate	0.1 to 0.5	0.2

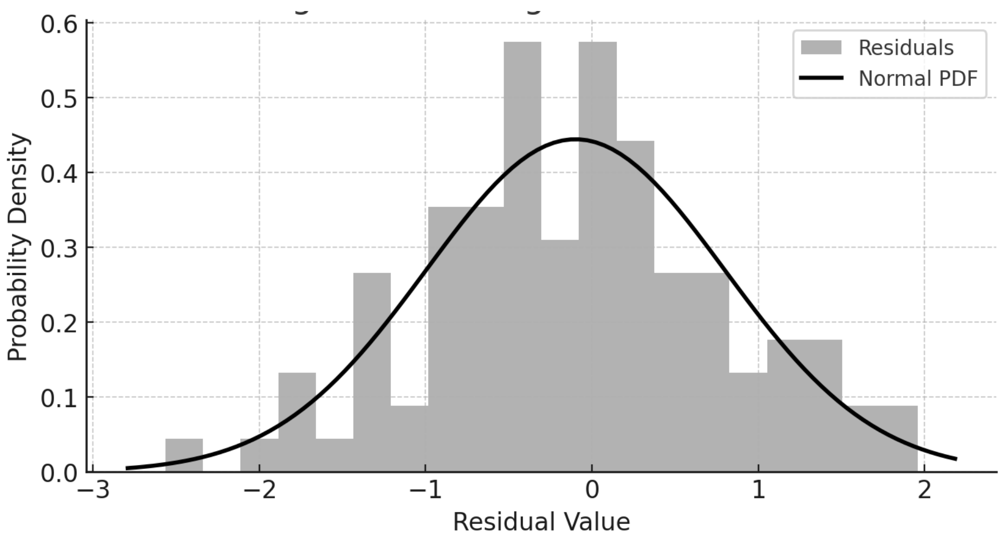

Figure A1. Histogram of residuals.

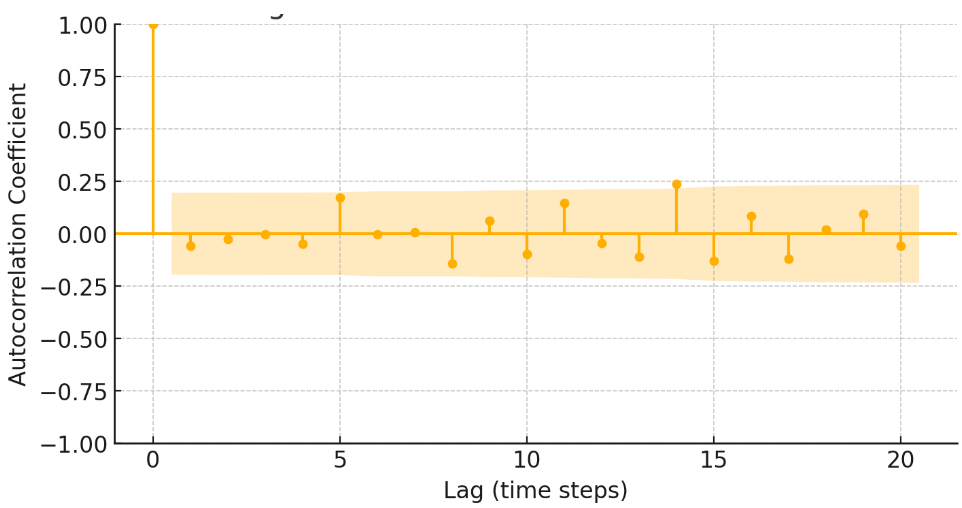

Figure A2. Autocorrelation of residuals.

Appendix C. Initial Conditional Parameters

The initial circumstances and parameters of the simulated aircraft fuel delivery model are as follows.

Components	Parameters	Specs
Initial conditions	Temperature Pressure	25 ± 2 °C 0.1 MPa
Fuel tank	Pressurisation Minimum fuel volume	0.1 MPa 0.09463525 m³
Wing tank	Initial volume Maximum capacity	10 m³ 12 m³
Centre tank	Pressurisation Initial volume Maximum capacity	0.1 MPa 5 m³ 284 m³
Pumps	Reference density Reference angular velocity Angular velocity threshold Operational ranges for angular velocity Mover time constant	920.027 kg/m³ 120 rev/s 10 rad/s 0 to 200 rev/s 0.2 s
Valves	Maximum opening area Leakage area Cut-off time constant Maximum valve opening (2-Way directional valves)	$\pi/4 \times (0.03048)^2$ m² 1×10^{-10} m² 0.1 s 5.1×10^{-3} m
Fuel line piping	Length Hydraulic diameter Aggregate equivalent length for local resistances	5m 3.05×10^{-2} m 2.56 m

References

1. Fu, S.; Avdelidis, N.P.; Jennions, I.K. A Prognostic Approach to Improve System Reliability for Aircraft Systems. In Proceedings of the 2023 7th International Conference on System Reliability and Safety (ICSRS), Bologna, Italy, 22–24 November 2023; pp. 259–264. https://doi.org/10.1109/ICSRS59833.2023.10381117.
2. Saxena, A.; Goebel, K. Turbofan Engine Degradation Simulation Data Set (C-MAPSS). NASA Ames Research Center, 2008. Available online: https://www.nasa.gov (accessed on 15 March 2025).
3. Pandian, G.P.; Das, D.; Li, C.; Zio, E.; Pecht, M. A Critique of Reliability Prediction Techniques for Avionics Applications. *Aeronaut* **2018**, *31*, 10–20. https://doi.org/10.1016/J.CJA.2017.11.004.
4. Liao X., Chen S., Wen P., Zhao S., Remaining useful life with self-attention assisted physics-informed neural network, (2023) Advanced Engineering Informatics, 58, art. no. 102195, DOI: 10.1016/j.aei.2023.102195.
5. Guo, Y.; Ma, C.; Jing, Z. A Hybrid Health Monitoring Approach for Aircraft Flight Control Systems with System-Level Degradation. *IEEE Trans. Ind. Electron.* **2023**, *70*, 7438–7448. https://doi.org/10.1109/TIE.2022.3201317.
6. Zhang, J.; Tian, J.; Li, M.; Leon, J.I.; Franquelo, L.G.; Luo, H.; Yin, S. A Parallel Hybrid Neural Network with Integration of Spatial and Temporal Features for Remaining Useful Life Prediction in Prognostics. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3501112. https://doi.org/10.1109/TIM.2022.3227956.
7. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep Learning and Its Applications to Machine Health Monitoring. *Mech. Syst. Signal Process.* **2017**, *115*, 213–237.
8. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. ISSN 0021-9991. https://doi.org/10.1016/j.jcp.2018.10.045.
9. Xu, S.; Yu, H.; Wang, H.; Chai, H.; Ma, M.; Chen, H.; Zheng, W.X. Simultaneous Diagnosis of Open-Switch and Current Sensor Faults of Inverters in IM Drives Through Reduced-Order Interval Observer. *IEEE Trans. Ind. Electron.* **2025**, *72*, 6485–6496. https://doi.org/10.1109/TIE.2024.3485708.
10. Fu, S.; Avdelidis, N.P. Prognostic and Health Management of Critical Aircraft Systems and Components: An Overview. *Sensors* **2023**, *23*, 8124. https://doi.org/10.3390/s23198124.
11. Dong, R.; Liu, W.; Cheng, B.; Li, W. An Adaptive LSTM-PR Hybrid Health Status Prognostics Strategy with Balance Between Accuracy and Computational Burden. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3505009. https://doi.org/10.1109/TIM.2024.3522359.
12. Biondani, F.; Dall'Ora, N.; Tosoni, F.; Fraccaroli, E.; Migliore, D.F.; Acerra, F.; Fummi, F. Fault Injection for Synthetic Data Generation in Aircraft: A Simulation-Based Approach. In Proceedings of the 2024 IEEE 22nd International Conference on Industrial Informatics (INDIN), Beijing, China, 17–20 August 2024; pp. 1–8. https://doi.org/10.1109/INDIN58382.2024.10774347.
13. Azar, K.; Meybodi, Z.; Naderkhani, F. Semi-supervised clustering-based method for fault diagnosis and prognosis: A case study. *Reliab. Eng. Syst. Saf.* **2022**, *222*, 108405. ISSN 0951-8320. https://doi.org/10.1016/j.ress.2022.108405.
14. *ISO/IEC 17025:2017*; General Requirements for the Competence of Testing and Calibration Laboratories. International Organization for Standardization, Geneva, Switzerland, 2017.