© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1720220

Artificial Intelligence for Predictive Maintenance and Structural Health Monitoring in Aerospace Systems: A State-of-the-Art Review

ABE OLAOLUWA OJO[1], FALODUN JOSHUA MAYOWA[2], BABALOLA EMMANUEL AYOOLUWA[3], EKEBAFE GODSFAVOUR ONOFUEH[4], OMIDIJI IYANUOLUWA ISAAC[5]
[1, 2, 3, 4] Department of Mechanical and Aerospace Engineering, Afe Babalola University
[5]School of Aeronautic Science and Engineering, Beihang University

Abstract - AI is moving the aerospace industry toward intelligent and data-driven maintenance practices and the aerospace industry is undergoing a paradigm shift to intelligent and data-driven maintenance systems with AI. This paper provides a comprehensive review of the state-of-the-art of AI methods for predictive maintenance (PdM) and structural health monitoring (SHM) in aerospace systems. We systematically classify and study machine learning (ML) and deep learning (DL) solutions for aircraft engine prognostics, composite damage detection, fatigue life prediction, and impact characterization for supervised, unsupervised and hybrid learning frameworks. A new taxonomy is suggested that correlates the AI techniques to the typical aerospace SHM functions: damage detection, localization, classification and quantification, remaining useful life (RUL) estimation. The review also explores new trends such as physics-informed neural networks (PINNs), digital twins, federated learning, and Explainable AI (XAI) in the context of enabling reliable aerospace SHM. The critical analysis of the TRLs shows that supervised and hybrid methods are at the TRL 5–6 (operational stages), while federated learning and physics-informed deep learning are still in the early experimental stages (TRL 2–4). Data scarcity, environmental and operational variability are identified as key challenges along with mitigation measures, interpretability for certification and real-time onboard processing constraints are identified with mitigation measures. A roadmap to the future of self-learning, intelligent, and certifiable SHM systems that are adaptive and able to diagnose damages autonomously and integrate digital twins without hassle is presented as the conclusion of the paper. This review will be useful to researchers and practitioners in the field of AI-assisted aerospace maintenance and safety.

Index Terms - Artificial Intelligence, Predictive Maintenance, Structural Health Monitoring, Aerospace Systems, Aircraft Engine Prognostics, Composite Damage Detection, Physics-Informed Neural Networks, Digital Twins, Federated Learning, Explainable AI.

I. INTRODUCTION

The aerospace industry worldwide is increasingly under pressure to improve aircraft safety, to cut down on operating expenses and aircraft downtime, and to cope with the growing complexity of the vehicles. With the increasing requirements of modern aviation, the traditional aircraft maintenance strategies, namely reactive (corrective) and preventive (scheduled) maintenance are no longer effective [1] [2]. Reactive maintenance is reactive and helps address the subsequent failures leading to unplanned downtime, flight cancellations and possible safety issues. Fixed interval preventive maintenance frequently results in the need to replace components unnecessarily and inappropriately uses resources [2].

A new paradigm based on condition-based monitoring, prognostics and health management (PHM) has come into existence, which is known as predictive maintenance (PdM) [3]. By using real-time sensor information, sophisticated analysis, and machine learning, PdM can anticipate component degradation and determine the remaining useful life (RUL), allowing maintenance to be performed just in time for maximum safety and cost savings. At the same time, Structural Health Monitoring (SHM) systems have developed from offline (non-destructive testing (NDT) techniques to continuous monitoring of aerospace structures via embedded sensor networks [4].

Incorporating artificial intelligence (AI) and machine learning (ML) into aerospace PHM and SHM has ushered in unprecedented capabilities. The deep learning models have shown the excellent performance in capturing the complex spatiotemporal dependency of sensor data, especially convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer architectures [4]. The aerospace domain, however, has quite a few challenges: high operational variability (humidity, temperature, vibration), limited availability of labelled damage data, strict certification requirements for safety-critical systems, and the need for real-time onboard processing [2].

This paper gives a thorough survey of the state-of-the-art methodologies for aerospace PdM and SHM based on AI. We systematically classify the ML/DL approaches, discuss new methods such as physics-informed neural networks (PINNs), digital twins and federated learning, and point to challenges and future directions. The review collates the results of research carried out between 2019 and 2026, but focuses on more recent advances (2024–2026) for use as a reference point for researchers and practitioners.

II. REQUIREMENT FOR AEROSPACE SYSTEM DEGRADATION AND MONITORING

A. Aerospace System Critical Failure Modes

There are multiple ways of system degradation in an aerospace system, for instance, in propulsion, airframe and avionics.

Aircraft Engine Degradation: Turbofan engines are subjected to severe mechanical and thermal stress, causing the gradual degradation of engine components such as erosion of high-pressure compressor (HPC), turbine blade creep and bearing wear. NASA C-MAPSS is an established dataset of multivariate sensor data from engines under different fault modes and operating conditions used for benchmarking RUL prediction models [5].

Damage on Composite Structure: The damage to the composites, commonly referred to as composite structure damage, refers to damage to composite materials such as carbon fiber reinforced polymer (CFRP) composites, which are used in modern aircraft airframes (Boeing 787, Airbus A350) and are vulnerable to barely visible impact damage (BVID), delamination and matrix cracking, as well as fiber breakage [6]. These damage modes can have a substantial impact on the load capacity and are not apparent by visual means.

Fatigue and Corrosion: Metallic structures undergo fatigue crack growth and corrosion from cyclic loading and environmental exposure. The vibration analysis, strain measurement, and electrochemical sensing are the methods for monitoring [4].

Sensor Faults: Voltage, current, temperature, and vibration sensor degradation or drift may result in incorrect PHM decisions, thereby resulting in missed detections and/or false alarms.

B. The Requirements and Regulatory Environment for SHM.

Aerospace SHM systems have a number of requirements that are very strict and which must be adhere to.

Real-Time Operation: Damage detection and RUL estimation have to be performed in milliseconds so as to allow taking action timely and to prevent catastrophic propagation.

High Reliability: Nothing is more important than low false negative rates for critical faults while these should be virtually 0%, and false positive rates must be reduced to a minimum to prevent unnecessary grounding and maintenance costs.

Environmental Robustness: Algorithms need to be able to perform under the temperature ranges between −55°C and +85°C, and humidity, vibrations, and electromagnetic interference conditions.

Certification and Interpretability: FAA, EASA, and EU AI Act require safety-critical AI systems to give back transparent, traceable decision-making.

Certification and Interpretability: The FAA, EASA, and EU AI Act require safety-critical AI systems to

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1720220

give back transparent and traceable decision-making. There are major hurdles in the way of certification in the case of “black-box” models.

Scalability: It should be possible to implement solutions that are platform-independent in terms of aircraft types, profiles, and structures, and that don't require a lot of retraining per platform.

III. TAXONOMY OF AI METHODS FOR AEROSPACE SHM

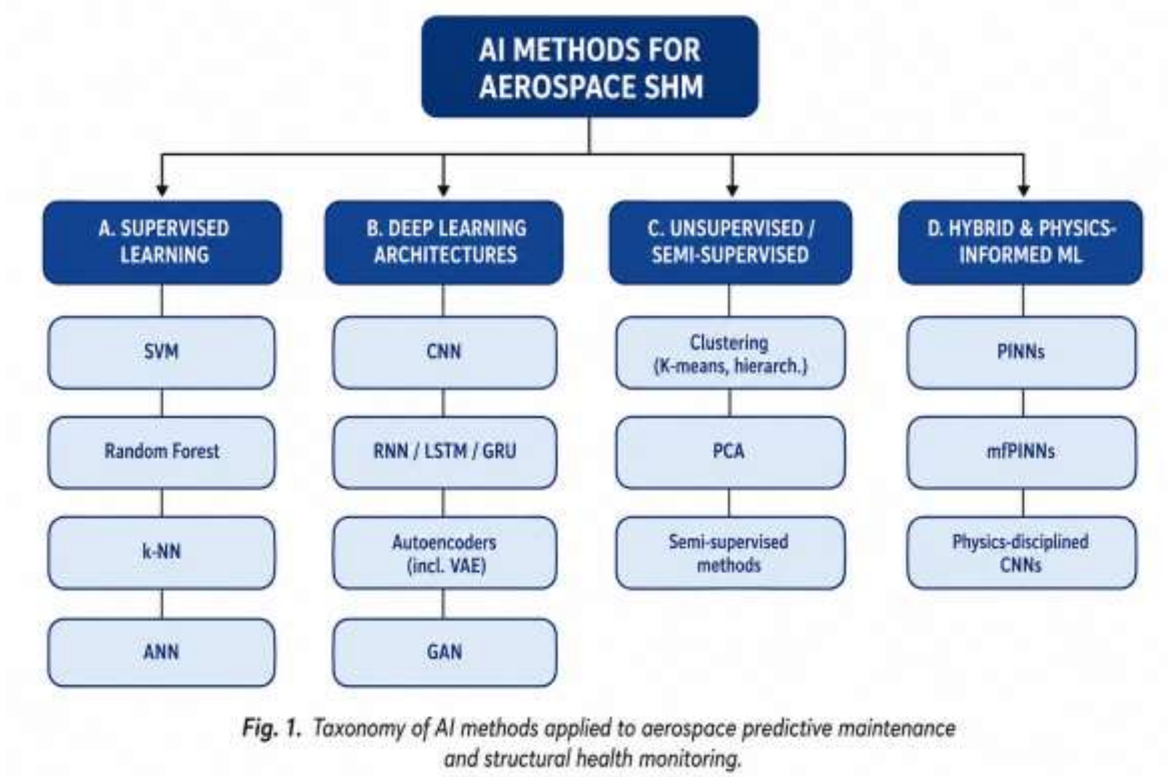

Fig. 1. Taxonomy of AI methods applied to aerospace predictive maintenance and structural health monitoring.

A. Unsupervised Learning Approaches

Existing SHM applications in the aerospace industry are mostly dominated by supervised learning algorithms which are trained using labeled databases of healthy and damaged structures.

Support Vector Machines (SVMs): Works well for binary classifications (healthy vs. damaged) and for small sample size. The SVMs have been widely used for Lamb-wave damage detection in composite plates with good accuracy achieved with a suitable feature extraction. But performance deteriorates in the presence of noisy and/or nonlinear damage states.

Random Forests (RFs): Ensemble methods that are robust to sensor noise, and can give interpretable feature importance rankings. The performance of RFs in impact energy classification and crack detection based on strains has been shown to be good, while the speed of inference is suitable for application onboard.

k-Nearest Neighbors (k-NN): Simple, intuitive algorithms without training, used for impact localization of CFRP panels, k-Nearest Neighbors (k-NN). Some limitations are sensitivity to feature scaling and the computational cost for large datasets.

Artificial Neural Networks (ANNs): Multi-layer perceptron’s are able to perform flexible nonlinear regression and classification. For composite beams, vibration-based ANN has been able to reach >95% accuracy with the assistance of optimization algorithms for damage quantification.

B. Deep Learning Architectures

Deep learning builds upon the ANNs that have multiple layers in between and they are able to extract features automatically from the high dimensional inputs.

Convolutional Neural Networks (CNNs): CNNs are good at capturing the spatial information in imaging and spectrogram data [8]. For aerospace SHM, CNNs are used in combination with the guided-wave signals converted to 2D images by the Gramian Angular Fields to detect and localize damages in CFRP panels, with a classification accuracy of 99.1%. for damage detection and 99.3% for damage localization [9]. Up to 95% of the training parameters can be saved with the reduced training time using transfer learning [10].

Recurrent Neural Networks (RNNs/LSTMs/GRUs): They are well-suited for handling temporal data, such as sensor data, and are suitable for fatigue life prediction and RUL estimation. LSTM networks have been successful in degradation modelling from the sensors of an aircraft engine.

Autoencoders: Unsupervised deep learning models which learn compact representations of healthy-state data. Reconstruction errors can be used as indicators of anomalies without the aid of labelled data to detect delamination [11]. Variational autoencoders (VAEs) also provide a way of conducting probabilistic anomaly detection and uncertainty quantification.

Generative Adversarial Networks (GANs): Limited labeled damage data in aerospace applications pose the challenge of generating realistic damage scenarios to supplement limited training data sets, which GANs can solve.

C. *Unsupervised and Semi-Supervised Learning*

Unsupervised methods tackle the problem of limited availability of labeled damage data, learning from the unlabeled or by far healthy data.

Clustering Algorithms: In the AE and guided-wave case, new damage states can be discovered with clustering algorithms (K-means and hierarchical). These techniques are more useful in the beginning when labeled data are not available [13].

Principal Component Analysis (PCA): Dimensionality reduction algorithms are used to obtain lower-dimensional subspaces from high dimensional data from several sensors, which are then used to detect novelty when operating under different conditions.

Semi-Supervised Approaches: those methods that take advantage of a large amount of unlabeled data and some labeled data. Improved sensitivity to rare damage events in composite laminates has been shown with autoencoders and graph-based methods. Autoencoders and graph-based methods have been shown to be more sensitive to rare damage events in composite laminates.

D. Hybrid and Physics-informed Machine Learning

Hybrid approaches use a combination of physics-based modelling and data-driven learning to follow the principle of leveraging the domain knowledge by boosting the generalization capacity.

Physics-Informed Neural Networks (PINNs): PINNs: equations governing the physical phenomenon (wave propagation, heat transfer, constitutive laws) are directly embedded in the loss function of the neural network. For aerospace SHM, the accuracy of PINNs is found to be the improvement from <80% to >90% in damage detection in various CFRP layups, surpassing the accuracy of purely data-based methods [14]. The PINNs achieve up to 40% reduction of computational costs compared to FE methods and do not require meshing.

Multi-Fidelity PINNs (mfPINNs): These are used to incorporate low fidelity physics-based models and high-fidelity experimental data. The mfPINNs reduced the average localization error by 95% over the conventional ANNs in the case of anisotropic composites for acoustic emission source localization [15].

Physics-Disciplined CNNs: The convolutional filter is restricted to only learn physical operations leading to meaningful feature extraction which is in line with domain knowledge. This is especially useful when working with certified filters, in which case the learned filters can be compared and contrasted to the real-world principles.

IV. APPLICATIONS OF AI IN AEROSPACE SHM TASKS.

A. Virtual Field Testing and Simulation

The basic SHM tasks are to detect damage (Level 1) and to locate the damage (Level 2).

Guided-Wave Based Methods: Lamb waves generated by piezoelectric transducers (PZTs) are sensitive to delamination, cracks and corrosion. Localization error of 7.58% in CFRP plates have been obtained by CNNs processing on the basis of guiding-wave spectrograms [9]. The GLFCFN features a multi-head attention layer for fusing features from local and global levels, an 1D-CNN for extracting local features and a gated MLP for extracting global information.

Vibration Based Methods: Modal Analysis with Acceleration and Strain sensors allows damage to be detected by changes in natural frequencies and mode shapes. ANNs have been used to classify cracks in composite laminates and have been trained on modal parameters where 95% accuracy has been obtained.

Imaging Methods: Using computer vision techniques such as ResNet and prototypical networks, the impact damage from surface images is classified with >94% accuracy, without any embedded sensors being used for the barely visible impact damage (BVID) recognition. The U-Net architectures segment CT images with >88% accuracy to identify warp, weft and interfacial cracks in 3D woven composites in 3D.

B. Damage Classification and Quantification

Damage classification (Level 3) and quantification (Level 4) give actionable information for maintenance planning.

Acoustic Emission (AE) Analysis: K means clustering and SVM classification of AE signals are used to identify matrix cracking, delamination, fiber/matrix debonding and fiber breakage in carbon/epoxy composites. Digital image correlation (DIC) is used to confirm cluster assignments, linking displacements and strains to the damage modes.

Probabilistic Method: Damage size and severity are quantified using probabilistic methods (Bayesian framework and sparse reconstruction algorithms). Probability imaging algorithms for Bayesian learning have been proven to consistently visualize the location and size of fatigue damage in CFRP laminates.

Deep Learning for Quantification: Defect classification and detection of defect type (hole, crack, intact) in CFRP plates are performed with high accuracy using Deep Learning for Quantification based on DenseNet and denoising diffusion probabilistic models (DDPM). Under noisy conditions, DDPM + DenseNet combination is superior to damage classification based on GANs.

C. Fatigue Life Prediction and RUL Estimation

One of the most important SHM functions is fatigue prognosis, which is essential for maintaining structural integrity for the operational life.

Deep Learning for RUL: In particular, deep learning algorithms such as LSTM networks and hybrid CNN-LSTM models have been shown to be effective in learning complex temporal degradation patterns from aircraft engine sensor data for the prediction of RUL. The Health-Aware Transformer (HAT) combines statistically derived Mahalanobis Distance metrics with attention mechanisms, and has demonstrated interpretable RUL prediction with RMSE of 11.27 on C-MAPSS FD001.

Transformer Architectures: The STAR (Two-Stage Attention-based Hierarchical Transformer) architecture is proposed to learn variable temporal and sensor-wise attention, achieving the best performance compared to state-of-the-art methods in terms of RMSE and Score for C-MAPSS FD002 [16] with a margin of 12% and 15%, respectively. The drawback of the quadratic sequence length is that it does not apply to very long time series.

Physics-Informed Approaches: PINNs can insert the Paris-Erdogan equations into constraints of the neural network to make physically consistent predictions of RUL while being applicable to different loading scenarios.

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1720220

Highlighting the effects of the impact event.

D. Impact Detection and Characterization

The impact events are a serious threat to composite structures which must be rapidly detected with energy evaluation.

CNN-Based Impact Localization: Training CNNs on 2D images which are based on PZT sensor signals, this method achieves 94.3% accuracy in predicting the impact location and >98.3% accuracy in predicting the energy classification. It is a method that can be generalized to untrained impact positions with >95% accuracy.

Prototypical Networks: Few-shot learning approaches with prototypical networks based on ResNet allow for damage detection with limited training data and, when using embedded glass/carbon composite sensors, outperform other methods.

Real-Time Monitoring: Lightweight CNN architectures designed for edge deployment can be used to detect impacts in real-time during flight operations, and can have inference times that are suitable for onboard processing requirements.

V. EMERGING PARADIGMS IN AEROSPACE SHM

A. Digital Twins for Predictive Maintenances

Aerospace maintenance is changing with the help of digital twin, they are virtual copies of physical assets that are up-to-date and in sync.

Physics Based Digital Twins: High fidelity finite element models (FEM) simulate behaviour of the structure to operational loads, providing synthetic training data to ML algorithms and "what-if" scenario analysis to support maintenance planning [17].

Data-Driven Digital Twins: ML models that are connected to an IoT sensor network to form a living digital copy of the asset that evolves as the asset changes. In pilot projects, the use of a digital twin has increased the efficiency of the predictive planning process by 32% and given the asset a 24% longer lifespan.

Hybrid Digital Twins: Physics-based and data-driven Digital Twins are combined to take advantage of physics' extrapolation capability and adaptability of data-driven learning. The hybrid methods have been shown to be 25% more accurate at predicting damages than purely data-driven methods.

B. Fleet-Wide SHM: Federated Learning

Federated learning allows for distributed aircraft fleets to collaborate on model training without having to share sensitive aircraft operational data.

Collaborative Improvement of Aircraft Models without Breaching Aircraft Operator Propriety and Competitive Information: Aircraft operators can work together to improve their SHM models without compromising their proprietary flight data and competitive information. The centralized data requirement reduction by 40-60% with the model accuracy has been shown using federated learning [18].

Physics-Informed Federated Learning: Integrating physics constraints into the federated learning framework helps to minimize the need for large-scale labeled data sets. DLR's federated PINN framework for ultrasonic SHM includes the acoustic wave equation in their decentralized learning, whereby they can learn together in aircraft fleets.

Challenges: Communication overhead, Heterogeneity in models across fleet types, Byzantine fault tolerance (malicious participants) are ongoing research areas.

C. Explainable AI for Trustworthy SHM

Explainability is key for aerospace certification and operator trust.

SHAP and LIME: Post-hoc explanation methods were used to explain feature importance for vibration and wave-based SHM, and found that certain frequency bands and wave propagation paths are more important in predicting damage than others [19].

Attention Mechanisms: Saliency maps and attention visualizations show saliency maps and attention visualizations of structural images and spectrograms

to explain the prediction results of the CNN and transformer networks in an intuitive manner.
Inherently Interpretable Models: The inherent interpretation of physics-disciplined CNNs with reversed mapping (CNN-RM) enables explainable decision making with feedback connections, resolving certification roadblocks, and achieves 95.1% testing accuracy.

Regulatory Alignment: XAI is also becoming crucial for commercialization, especially in light of the EU AI Act and FAA guidelines that are increasingly demanding AI systems to be interpretable for safety-critical applications. In addition, the EU AI Act and FAA guidelines are putting increasing pressure on the interpretable nature of AI systems in safety-critical applications, making XAI integration crucial for commercialization.

VI. THE CRITICAL CHALLENGES AND TECHNOLOGY READINESS.

A. Data Scarcity and Quality

Since there are few examples of labeled flight data and damage, it is still the biggest challenge for aerospace ML. Solutions include:

i. Simulations with synthetic data using physics-based simulators and GANs
ii. The following week, the researchers transferred their learning to the other platforms and operational domains and observed the results.
iii. Application of machine learning to unlabeled data exploitation, namely semi-supervised and self-supervised learning approaches
iv. A few-shot learning using prototypical networks and meta-learning. A few shot learning using prototypical networks and meta-learning.

B. Environmental and Operational Variability

The environmental and operational conditions (EOCs) to which aerospace structures are exposed are extreme and pose challenges to the robustness of the models:

i. Temperature changes can cause changes in the calibration of these sensors and the properties of the materials being measured
ii. Humidity & Salt Exposure Cause Corrosion and Delamination to accelerate.
iii. The different vibration and acoustic noise create disturbances in the sensor signals.
iv. Dynamic responses of structures vary with the loading variations.

Some mitigation methods are adversarial training, domain adaptation, and physics-informed constraints representing physical laws which do not change. Denoising and compression of guided-wave data with manifold learning approaches (such as Diffusion Maps combine with autoencoders) enhance the capability of anomaly detection in different operating conditions.

C. Certification and Interpretability

Deep learning models are “black box” models, which cause large difficulties for aerospace certification:

i. All regulatory bodies need to have clear and accountable decision-making processes.
ii. The traditional verification and validation (V&V) methods are not suitable to be applied to neural networks
iii. Uncertainty quantification must be done for risk-aware decisions

Physics-informed and hybrid methods can also lead to certification by enforcing physical limits on predictions. For regulatory compliance, XAI techniques are used for providing transparency; for uncertainty-aware frameworks, such as Bayesian neural networks, ensemble methods, they facilitate confidence-based decision thresholds.

D. Real-Time Onboard Processing

For the deployment of AI systems on the edge of aircrafts, the following is required:

i. Accurate and precise results, even under early warning conditions.
ii. Low latency inference (<10ms for critical faults)
iii. Completely safe for aircraft electrical system use small sizes for embedded processors (FPGAs, microcontrollers)
iv. The ability to withstand the effects of radiation and electromagnetic interference.

Lightweight architectures like MobileNet, EfficientNet and model compression methods such as pruning and quantization are vital for an on-board

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1720220

deployment. A summary of the computational cost comparison of all ML methods is given in Table 1.

TABLE 1. Computational Cost Comparison of ML/DL Methods for Aerospace SHM

ML/DL Method	Training Complexity	Inference Latency	Memory Footprint	Onboard/Edge Suitability
SVM	Low	< 1 ms	Low	High
Random Forest	Low-Medium	< 1 ms	Low-Medium	High
k-Nearest Neighbors	None (lazy learner)	Medium-High (scales with data)	Medium	Medium
ANN (MLP)	Medium	1-5 ms	Low-Medium	High
CNN	High	5-15 ms	Medium-High	Medium (edge-optimized variants only)
RNN / LSTM / GRU	High	10-20 ms	Medium-High	Medium
Autoencoder / VAE	Medium-High	5-10 ms	Medium	Medium
GAN	Very High	Not used at inference (offline data generation)	High	Low
PINN	High	5-10 ms	Medium	Medium
mfPINN	Very High	10-15 ms	Medium-High	Low-Medium
Transformer (STAR / HAT)	Very High	15-30 ms	High	Low
Lightweight CNN (MobileNet/EfficientNet, pruned/quantized)	Medium (with compression)	< 5 ms	Low	High

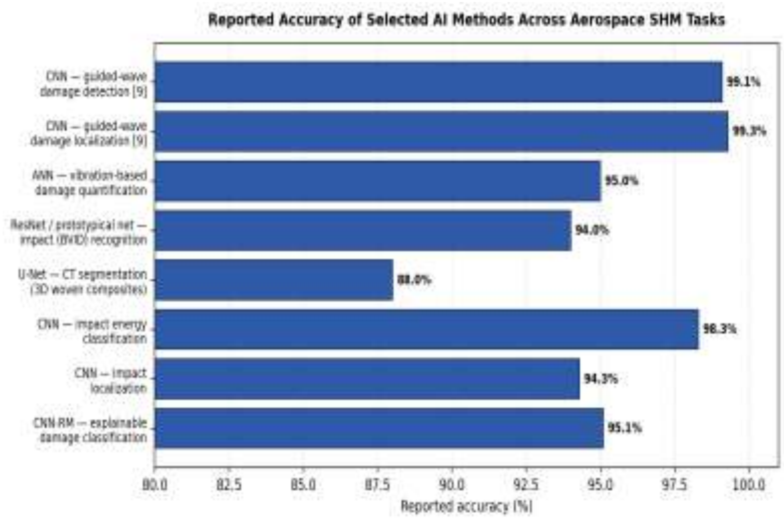

Fig. 3. Reported accuracy figures for representative AI methods, as cited in Sections III-IV of this review.

Fig. 3. Reported accuracy figures for representative AI methods, as cited in Sections III-IV of this review.

E. Technology Readiness Assessment

The TRLs as seen in Table 2 indicate the readiness level of AI techniques for aerospace SHM. Table 2 lists the technology readiness level (TRL) of AI techniques for aerospace SHM.

TABLE 2. Technology Readiness Levels (TRL) of AI Techniques for Aerospace SHM

TRL Range	Readiness Status	AI Techniques	Description
TRL 5-6	Near-Operational	Supervised learning (SVM, RF, CNN); hybrid model-based damage detection and localization	Validated under controlled/relevant conditions; undergoing flight testing toward full operational use
TRL 4-5	Experimental	Digital twins, federated learning, physics-informed deep learning	Component/subsystem validated in laboratory settings; promising robustness but unproven at fleet scale
TRL 2-3	Concept	Quantum computing and neuromorphic computing for SHM	Basic principles observed and technology concept formulated; early exploratory research stage

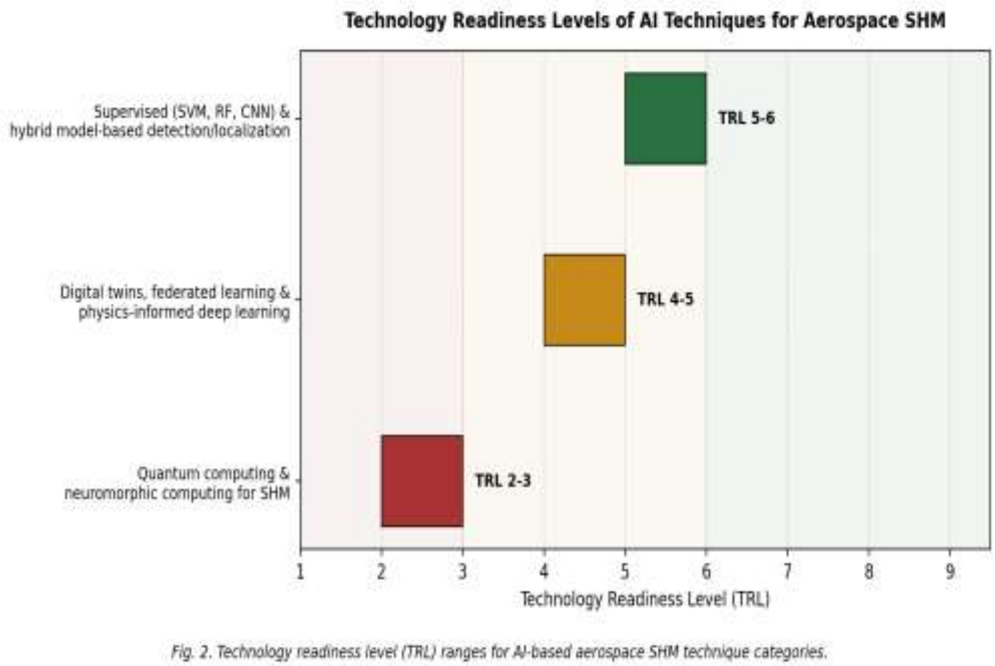

Fig. 2. Technology readiness level (TRL) ranges for AI-based aerospace SHM technique categories.

Under controlled conditions, supervised learning (SVM, RF, CNN) and hybrid model-based damage detection and localization (TRL 5 – 6: Near Operational). These are still being flight tested for complete operational use.

- TRL 4–5 (Experimental): Digital twins, federated learning and physics-informed deep learning. They offer good potential to ensure both robustness and operation, but are still in early stages of experimentation.

TRL 2-3 (Concept): Optimization and Neuromorphic Computing with Quantum Devices for SHM.

Some of the common obstacles in all the TRLs are the lack of data, sensitivity to EOCs, the problems of interpretability and technology being available for real time processing.

VII. FUTURE RESEARCH DIRECTIONS

A. Edge Intelligence & Autonomous SHM

Future SHM systems will be based on the direct use of AI at the sensor level to allow autonomous damage diagnosis without accessing the cloud. The energy efficiency and latency of the neuromorphic sensor and processing-in-memory architectures will be orders of magnitude better.

B. Digital Twins at Multi-Scale and Multi-Physics

Multi-scale PINNs will connect the microstructural damage mechanics (fiber/matrix debonding) with macrostructural performance (wing deflection) in future. Higher-order multi-scale PINNs (HOMS-PINNs) have been shown to be highly computational efficient to connect the microscopic unit cell functions with homogenization solution at the macroscopic level.

C. Standardization and Benchmarking

Benchmarking isn't easy because of a lack of uniformity. We suggest using Aerospace SHM Benchmark (ASB), consisting of:

i. A multi-modal dataset with annotations (guided-wave, vibration, thermal and imaging).
ii. Contractual obligations and commitments (terms and conditions, warranties, representations, conditions, guarantees, and the like)
iii. Certification-ready evaluation protocols

D. Collaboration between Humans and AI in Maintenance Decision-Making.

Future SHM systems will focus on human-AI collaboration instead of automatic systems. XAI interfaces will enable operators to interact with the system in a natural way, allowing them to understand the reasoning behind the decisions made and to make better-informed choices. The trust and quality of the decision made by an operator will be improved with XAI interfaces that offer natural language explanations, interactive visualization and uncertainty-aware recommendations.

VIII. CONCLUSION

In this paper, we have given an in-depth review of the state of the art of AIRPMS and SHM in the field of aerospace systems. A comprehensive survey has been conducted and the machine learning (ML) and deep learning (DL) techniques (from the traditional supervised techniques to the emerging physics-informed, federated and explainable AI techniques) have been organized and linked to various aerospace SHM tasks such as damage detection, damage localization, damage classification and quantification, and RUL estimation.

The review finds that although there are some operational readiness technologies, such as supervised and hybrid approaches (TRL 5–6), there are transformative technologies, such as digital twins, federated learning, and physics-informed deep learning, that are still at early experimental stages (TRL 2–4). The challenges of scarcity of data, variability of the environment, certification hurdles and real-time processing restrictions have been identified and specific mitigation options have been proposed.

In the future, AI-powered SHM systems will play a crucial role in making aerospace systems safer, more efficient, and more environmentally friendly by enabling autonomous, intelligent, and sustainable aviation.

In conclusion, the aerospace industry's future is looking towards autonomous, intelligent, and sustainable aviation, and AI-powered SHM systems are going to be essential for making this a reality. The roadmap outlined here will serve as a framework for the future generation of certifiable, trustworthy, and adaptive aerospace maintenance systems, both with edge intelligence and multi-scale digital twins driving the innovation, as well as standardization and human-AI collaboration.

REFERENCES

[1] W. J. C. Verhagen, B. F. Santos, F. Freeman, P. van Kessel, D. Zarouchas, T. Loutas, D. Yeun, N. Hodzic, C. Zamboni, and I. Heiets, "Condition-based maintenance in aviation: Challenges and opportunities," Aerospace, vol. 10, no. 9, p. 762, 2023. doi: 10.3390/aerospace10090762.

[2] M. J. Scott, W. J. C. Verhagen, M. T. Bieber, and P. Marzocca, "A systematic literature review of predictive maintenance for defence fixed-wing aircraft sustainment and operations," Sensors, vol. 22, no. 18, p. 7070, 2022. doi: 10.3390/s22187070.

[3] S. Fu and N. P. Avdelidis, "Prognostic and health management of critical aircraft systems and components: An overview," Sensors, vol. 23, no. 19, p. 8124, 2023. doi: 10.3390/s23198124.

[4] G. Scarselli and F. Nicassio, "Machine learning for structural health monitoring of aerospace structures: A review," Sensors, vol. 25, no. 19, p. 6136, 2025. doi: 10.3390/s25196136.

[5] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in Proc. 1st Int. Conf. Prognostics and Health Management (PHM08), Denver, CO, USA, Oct. 2008.

[6] A. Wronkowicz-Katunin, A. Katunin, and K. Dragan, "Reconstruction of barely visible impact damage in composite structures based on non-destructive evaluation results," Sensors, vol. 19, no. 21, p. 4629, 2019. doi: 10.3390/s19214629.

[7] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," Off. J. Eur. Union, OJ L, 2024/1689, 12.7.2024. Available: http://data.europa.eu/eli/reg/2024/1689/oj.

[8] M. Rautela and S. Gopalakrishnan, "Ultrasonic guided wave based structural damage detection and localization using model assisted convolutional and recurrent neural networks," Expert Syst. Appl., vol. 167, p. 114189, 2021. doi: 10.1016/j.eswa.2020.114189.

[9] Y. Liao, X. Qing, Y. Wang, and F. Zhang, "Damage localization for composite structure using guided wave signals with Gramian angular field image coding and convolutional neural networks," Compos. Struct., vol. 312, p. 116871, 2023. doi: 10.1016/j.compstruct.2023.116871.

[10] S. Sawant, A. Sethi, S. Banerjee, and S. Tallur, "Unsupervised learning framework for temperature compensated damage identification and localization in ultrasonic guided wave SHM with transfer learning," Ultrasonics, vol. 130, p. 106931, 2023. doi: 10.1016/j.ultras.2023.106931.

[11] H. Lee, H. J. Lim, T. Skinner, A. Chattopadhyay, and A. Hall, "Automated fatigue damage detection and classification technique for composite structures using Lamb waves and deep autoencoder," Mech. Syst. Signal Process., vol. 163, p. 108148, 2022. doi: 10.1016/j.ymssp.2021.108148.

[12] K. K. Prajapati, A. Ghosh, and M. Mitra, "Semi-supervised generative adversarial network (SGAN) for damage detection in a composite plate using guided wave responses," Mech. Syst. Signal Process., vol. 232, p. 112686, 2025. doi: 10.1016/j.ymssp.2025.112686.

[13] L. B. Andraju and G. Raju, "Damage characterization of CFRP laminates using acoustic emission and digital image correlation: Clustering, damage identification and classification," Eng. Fract. Mech., vol. 277, p. 108993, 2023. doi: 10.1016/j.engfracmech.2022.108993.

[14] Y. Song, S. Shan, Y. Zhang, and L. Cheng, "Physics-guided neural network for structural health monitoring with Lamb waves through boundary reflection elimination," Struct. Health Monit., vol. 25, no. 2, pp. 1219–1236, 2026. doi: 10.1177/14759217241305050.

[15] N. M. M. Kalimullah, A. Shelke, and A. Habib, "A probabilistic framework for source localization in anisotropic composite using transfer learning based multi-fidelity physics informed neural network (mfPINN)," Mech. Syst. Signal Process., vol. 197, p. 110360, 2023. doi: 10.1016/j.ymssp.2023.110360.

[16] Z. Fan, W. Li, and K.-C. Chang, "A two-stage attention-based hierarchical transformer for turbofan engine remaining useful life prediction," Sensors, vol. 24, no. 3, p. 824, 2024. doi: 10.3390/s24030824.

[17] J. W. Y. Chia, W. J. C. Verhagen, J. M. Silva, and I. S. Cole, "A review and outlook of airframe digital twins for structural prognostics and health management in the aviation industry," J. Manuf. Syst., vol. 77, pp. 398–417, 2024. doi: 10.1016/j.jmsy.2024.09.024.

[18] D. Landau, I. de Pater, M. Mitici, and N. Saurabh, "Federated learning framework for collaborative remaining useful life prognostics: An aircraft engine case study," arXiv:2506.00499, 2025. doi: 10.48550/arXiv.2506.00499.

[19] X. Wang, Z. Wei, Z. Wang, S. Wei, Y. Li, and M. M. Shahzad, "Explainable AI-driven optimal feature selection for the identification of structural damage," Struct. Control Health Monit., 2025, Art. no. 7253150. doi: 10.1155/stc/7253150.