mathematics

MDPI

Article

Integration of Foundation Models and Federated Learning in AIoT-Based Aircraft Health Monitoring Systems

Igor Kabashkin

Engineering Faculty, Transport and Telecommunication Institute, Lauvas 2, LV-1019 Riga, Latvia; kiv@tsi.lv

Abstract: The study presents a comprehensive framework for integrating foundation models (FMs), federated learning (FL), and Artificial Intelligence of Things (AIoT) technologies to enhance aircraft health monitoring systems (AHMSs). The proposed architecture uses the strengths of both centralized and decentralized learning approaches, combining the broad knowledge capture of foundation models with the privacy-preserving and adaptive nature of federated learning. Through extensive simulations on a representative aircraft fleet, the integrated FM + FL approach demonstrated consistently superior performance compared to standalone implementations across multiple key metrics, including prediction accuracy, model size efficiency, and convergence speed. The framework establishes a robust digital twin ecosystem for real-time monitoring, predictive maintenance, and fleet-wide optimization. Comparative analysis reveals significant improvements in anomaly detection capabilities and reduced false alarm rates compared to traditional methods. The study conducts a systematic evaluation of the benefits and limitations of FM, FL, and integrated approaches in AHMS, examining their implications for system robustness, scalability, and security. Statistical analysis confirms that the integrated approach substantially enhances precision and recall in identifying potential failures while optimizing computational resources and training time. This paper outlines a detailed aviation ecosystem architecture integrating these advanced AI technologies across centralized processing, client, and communication domains. Future research directions are identified, focusing on improving model efficiency, ensuring generalization across diverse operational conditions, and addressing regulatory and ethical considerations.

Keywords: aviation health monitoring systems; foundation models; federated learning; Artificial Intelligence of Things; predictive maintenance; aircraft safety; machine learning in aviation; cognitive computing

MSC: 93A15

check for updates

Citation: Kabashkin, I. Integration of Foundation Models and Federated Learning in AIoT-Based Aircraft Health Monitoring Systems. *Mathematics* **2024**, *12*, 3428. https://doi.org/10.3390/math12213428

Academic Editor: Daniel-Ioan Curiac

Received: 13 October 2024
Revised: 27 October 2024
Accepted: 31 October 2024
Published: 31 October 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Background and Motivation

The aviation industry has always been at the forefront of technological innovation, driven by the imperative to ensure safety, reliability, and efficiency in one of the most complex and demanding operational environments. In recent years, the rise of artificial intelligence (AI) with advanced machine learning (ML) techniques, particularly foundation models (FMs) and federated learning (FL), has offered new opportunities to enhance aviation health monitoring systems (AHMSs). These systems play a crucial role in monitoring the health of aircraft, predicting potential failures, and optimizing maintenance schedules, thereby ensuring that aircraft remain safe and operational.

Foundation models, known for their ability to generalize across a wide range of tasks due to their training on massive and diverse datasets, hold great potential for AHMSs. They can analyze vast amounts of data to detect patterns and predict failures that might not be apparent through traditional methods. However, these models are typically trained

in centralized environments, which can pose significant risks to data privacy and security, especially when dealing with sensitive aircraft operational data.

On the other hand, federated learning offers a decentralized approach that mitigates some of these privacy concerns by keeping data localized on individual aircraft. This method allows each aircraft to learn from its own data and share only model updates with a central server, which then aggregates these updates to refine the global model. While this approach preserves privacy, it introduces new challenges, such as ensuring that the global model generalizes well across different aircraft and operating conditions, managing communication overhead, and protecting against potential security vulnerabilities during the federated learning process.

The article explores a synergistic approach to aircraft health management systems, harmonizing the strengths of foundation models and federated learning. This integrated methodology seeks to use the comprehensive knowledge base of foundation models while capitalizing on the distributed, privacy-preserving nature of federated learning, thus offering a nuanced solution that addresses the complex demands of modern aviation maintenance and safety protocols.

1.2. Related Works

From a technical standpoint, FMs are not entirely new, as they build upon long-established deep neural networks and standard transfer learning techniques. However, the combination of their immense scale and the large datasets used for training has led to novel emergent capabilities that can be transferred across applications and further fine-tuned to create a wide array of AI-driven solutions.

Foundation models, as defined in [1], refer to "any model that is trained on broad data and can be adapted to a wide range of downstream tasks." These models are characterized by task-agnostic pre-training followed by fine-tuning for specific applications, representing a new approach to the construction and deployment of AI systems. This shift, often referred to as a paradigm shift in AI, presents both significant opportunities and challenges [2], a central theme explored in this manuscript. Foundation models may be trained on a single modality, such as language, vision, robotics, or reasoning, or demonstrate multi-modal capabilities [3] with some prominent examples include OpenAI's GPT series [4]. While foundation models are applied across various domains, their pre-training paradigm has been particularly influential in advancing benchmarks in natural language processing, where applications like ChatGPT, based on the GPT foundation model, are now capable of generating text that closely mimics human writing [5].

Report [6] provides a comprehensive overview of the FMs and their implications for competition and consumer protection. It explores the rapid expansion of foundation models, highlighting key trends such as the increasing number of small but capable models, advancements in on-device deployment, and the growing influence of open-source FMs. The report identifies critical inputs for FM development, including data, compute, and expertise, emphasizing the strategic role of partnerships and collaborations in accessing these resources. The document also delves into the integration of FMs across various platforms, particularly by major technology firms, and discusses early adoption trends in digital markets.

Monitoring aircraft structural health under varying loads is essential for aviation and aerospace engineering, yet traditional load calibration methods through ground testing are expensive and inefficient. Study [7] demonstrates the application of foundation models in aviation by introducing a deep learning-based aircraft load model for strain prediction and load calibration. The proposed model follows a two-phase approach; first, it establishes the causal relationship between key flight parameters and strain data, before integrating the prediction equation into the monitoring process to develop a general load model for real-time calibration of load coefficients.

The use of datasets plays a significant role in advancing the development and application of FMs in aviation, particularly in tasks such as aircraft recognition, classification,

and automated analysis of aircraft features. The FGVC-Aircraft dataset [8] is one of the examples of such datasets. It is a specialized benchmark designed for fine-grained visual categorization of aircraft. This dataset consists of over 10,000 images, representing 102 distinct aircraft variants, categorized across a four-level hierarchy: model, variant, family, and manufacturer.

NASA and IBM have jointly released a groundbreaking open-source Geospatial AI Foundation Model based on NASA's Harmonized Landsat Sentinel-2 dataset [9]. This foundation model represents a significant advancement in the use of AI for Earth observation and analysis, using the power of deep learning to handle vast geospatial datasets for monitor natural disasters and predict environmental outcomes.

Paper [10] explores the application of FMs, particularly generative AI models, in aviation safety analysis to improve efficiency and accelerate the processing of incident reports. An FM was employed to generate concise incident synopses from aviation safety narratives.

Federated learning is a technology in which multiple local clients and a central server cooperate to learn a global model in a situation where data are decentralized, and local clients are IoT devices [11].

Paper [12] presents a systematic review of FL, where multiple consumers collectively solve machine learning problems while keeping training data distributed and private. The review categorizes FL, explores its application areas, and examines its relationship with blockchain technology.

Paper [13] introduces a novel clustering FL method incorporating a self-attention mechanism for bearing fault diagnosis in aerospace applications, where traditional fault diagnosis techniques are hindered by data scarcity and security issues. By developing a deep neural network that captures both local and global information from raw input for feature extraction and clustering data from different equipment with similar distributions, the method effectively utilizes distributed data while ensuring privacy.

Paper [14] discusses the crucial role of condition monitoring in enhancing industrial productivity through intelligent fault diagnosis systems, which help in accurately scheduling condition-based maintenance. Highlighting the challenges of data privacy, security risks, and competitive barriers often encountered due to data distribution among different regions or institutions, it proposes FL as a solution to collaboratively train reliable fault diagnosis models without compromising data security.

Paper [15] proposes an FL-based, privacy-preserving intrusion detection system to safeguard against uplink and downlink cyber-attacks on controller–pilot data link communication systems. To validate the efficacy of this solution, a unique training dataset was created using air–ground communication data near Arlanda airport in Sweden. The resulting anomaly detection model, trained through FL, outperformed centrally and locally trained models in accuracy and precision, thus demonstrating its suitability for securing sensitive aviation communications.

Study [16] evaluates distributed collaborative learning techniques, specifically FL, for predictive maintenance in the aircraft industry's prognostic and health management (PHM) systems. Traditional centralized degradation estimation methods have faced scalability and data privacy issues, which FL addresses by enabling privacy-preserving, decentralized ML directly at the network edges. Through experiments comparing federated average (FedAvg) and federated proximal term (FedProx) algorithms in feed-forward neural networks, it was found that while FedAvg achieves prediction errors comparable to centralized models, it shows variability in accuracy across iterations. Conversely, FedProx demonstrates a steady decrease in prediction error when using specific parameter settings, suggesting its potential for more stable and accurate PHM implementations in aviation.

Article [17] discusses enhancing unmanned aerial vehicles (UAVs)' network intelligence using ML to address traditional cloud-centric ML issues like privacy, latency, and resource burdens in both civil and military applications. A decentralized FL model is proposed to allow UAVs to collaboratively train an ML model without centralizing data,

mitigating risks such as single points of failure and adapting to unreliable UAV network nodes and links.

Maintenance is one of the most prominent areas for the application of AI, and it represents a vast domain for AI's practical use. The study in [18] examines the use of convolutional neural networks alongside autonomous drones for automating visual inspections during aircraft maintenance. Building on earlier research, the paper introduces techniques to enhance defect detection, achieving improved accuracy in identifying issues such as dents by employing specific image augmentations and pre-classification strategies.

In [19], a novel ML and Internet of Things (IoT)-based approach is presented to predict the thermal performance of aircraft wing anti-icing systems. The approach, which utilizes an artificial neural network, is demonstrated to be more efficient and timesaving than traditional computational fluid dynamics methods, highlighting its potential application within aviation.

Paper [20] surveys various statistical and ML methods aimed at providing more efficient, accurate, and data-driven analyses of aircraft environmental impacts, including fuel burn, emissions, and noise. It summarizes key research themes and identifies opportunities for further integration of these techniques to enhance the sustainability of aviation operations.

In [21], the use of deep neural networks and transfer learning on images of aircraft lap joints is proposed for automatically detecting corrosion. The method achieves precision comparable to that of trained human operators, offering support for maintenance personnel and enabling more automated condition-based maintenance procedures.

A study focused on helicopters in [22] explores the potential for active vibration control, specifically using individual blade control to reduce vibration loads at the hub. By integrating various models and methods, including a fuzzy neural network, the research indicates that this method can effectively reduce hub vibration loads, offering valuable insights for designing helicopter vibration control laws.

Four data-driven frameworks for predicting the exhaust gas temperature baseline of aeroengines, which is essential for engine health analysis and flight safety, are discussed in [23]. Using real engine data, ML methods were trained, with the generalized regression neural network model demonstrating the highest accuracy and efficiency, making it highly suitable for practical airline applications.

Article [24] provides a comprehensive review of the application of ML to lithium-ion battery research, particularly in areas such as material research, battery health estimation, and fault diagnosis, with a specific focus on aviation batteries and green aviation technology. The study outlines the strengths and weaknesses of various ML techniques and highlights the potential for future advancements in the field.

In [25], the use of ML, specifically a multilayer perceptron neural network, is explored to model the transient performance of aero engines, with a focus on heat transfer during transient operations. This model, trained on finite element simulation data and refined with actual measurements from aero engines, accurately replicates thermal transients, proving useful for aviation applications.

A data-driven approach to predicting base pressure in suddenly expanded flows, which influences base drag in aerodynamic vehicles, is introduced in [26]. ML models trained on data from response equations accurately predict base pressure, aiding in optimizing base drag for rockets and missiles.

The study in [27] introduces ML techniques, specifically a deep neural network (DNN) and a random forest classifier (RFC), to predict null motions in a four-control moment gyroscope used for satellite attitude control. The RFC method is shown to outperform the DNN in accuracy, allowing for reliable predictions of null motions even for maneuvers not present in the training data.

1.3. Research Gaps, Contributions, and Paper Structure

In recent years, advancements in AI have significantly impacted health monitoring systems across various sectors. However, in aviation, the application of AI-based health monitoring still faces considerable challenges. Existing research primarily focuses on either foundation models or federated learning independently. Foundation models offer powerful generalization capabilities but are usually implemented in centralized architectures, which can impose significant privacy and data security risks. Conversely, federated learning presents a decentralized approach, allowing data to remain local while only model updates are aggregated. Despite the promise of these approaches, limited research has explored their combined application in aviation health monitoring, leaving an opportunity to enhance system robustness, scalability, and adaptability by leveraging the strengths of both.

One of the primary gaps in current research is the scalability of AI-based health monitoring systems in aviation. The sheer size and complexity of modern aviation fleets generate massive, diverse datasets that require efficient processing for real-time decision-making. Traditional centralized systems can become bottlenecks due to the extensive data aggregation required, often resulting in delayed insights that are not viable for dynamic, safety-critical environments like aviation. This research addresses this gap by proposing a scalable architecture that integrates both centralized and decentralized AI models, using the generalization capabilities of FMs to manage broad patterns across the fleet while allowing local adaptations through FL to capture aircraft-specific conditions.

Another key gap pertains to privacy and security concerns inherent in centralized data processing. Aircraft health monitoring data are sensitive, with substantial implications for both security and competitive advantage. While centralized foundation models can utilize vast datasets, they also introduce vulnerabilities related to data breaches and unauthorized access, as raw operational data need to be transmitted to a central server. This research bridges this gap by incorporating federated learning principles, which mitigate privacy risks by keeping data localized on each aircraft while sharing only model updates. This decentralized approach not only enhances data privacy but also aligns with emerging data protection regulations and best practices in secure data handling, thus ensuring a robust privacy-preserving framework for AI-driven health monitoring.

Another critical challenge in applying AI to aviation health monitoring is achieving compliance with strict regulatory standards. The "black box" nature of many AI models limits transparency, making it difficult for aviation professionals and regulators to interpret and validate these systems' decisions. This lack of interpretability hampers adoption in safety-critical domains where every decision must be justifiable and auditable. To address this, the proposed architecture incorporates explainable AI techniques that improve model interpretability, offering insights into the decision-making process. By providing clear and interpretable outputs, this study enhances the explainability of the integrated model, ensuring that the system's predictions can be scrutinized and verified, which is crucial for regulatory compliance in the aviation industry.

The aviation environment is highly variable, with diverse operational conditions across fleets and individual aircraft. Many existing AI approaches in aviation struggle to adapt effectively to this variability, often leading to suboptimal performance in non-standard or highly specific conditions. By integrating FL, this study presents a model that dynamically adjusts to local operational contexts while benefiting from the broader knowledge encapsulated by FMs. This hybrid approach allows the health monitoring system to capture fleet-wide patterns through a generalized model, which can then be fine-tuned for each aircraft, thus achieving a balance between general applicability and local relevance.

The purpose of this article is to address these gaps by proposing a comprehensive framework that integrates foundation models, federated learning, and AIoT technologies for aircraft health monitoring. Specifically, this study aims to present an integrated architecture that uses the strengths of both centralized and decentralized learning approaches, explore scalable solutions for fleet-wide deployment of AI-driven health monitoring sys-

tems, discuss strategies for ensuring data privacy and model security in aviation contexts, examine methods for enabling real-time processing and decision-making using edge computing and efficient model design, consider the regulatory implications of AI in aviation and propose approaches for compliance, and investigate techniques for improving the interpretability and explainability of AI models in aircraft health monitoring. By addressing these gaps, this article presents a new approach to aviation health monitoring through the possibilities of AI and IoT technologies.

The remainder of this paper is structured as follows. Section 2 presents the framework for collaborative intelligence in aircraft health monitoring, detailing its main components, including AIoT technologies, FMs, FL, and the integrated model approach. Section 3 discusses the results of implementing this framework, addressing the architecture of the aviation ecosystem and providing a comparative analysis of FMs, FL, and the integrated approach in the context of big data and cognitive computing for AHMS. Section 4 provides a discussion on the comparative analysis of the three approaches, examining their benefits and limitations, and explores future directions of research in AI-driven AHMS. Section 5 concludes the paper, summarizing the key contributions and implications of this work for the future of aircraft health monitoring and aviation safety.

2. Framework for Collaborative Intelligence in Aircraft Health Monitoring

The study framework for collaborative intelligence (SFCI) in aircraft health monitoring (AHM) represents a comprehensive approach to using advanced data processing and cognitive computing technologies in the aviation industry. This framework (Figure 1) outlines a progression from foundational Artificial Intelligence of Things (AIoT) technologies to a sophisticated, integrated architecture for aircraft health monitoring.

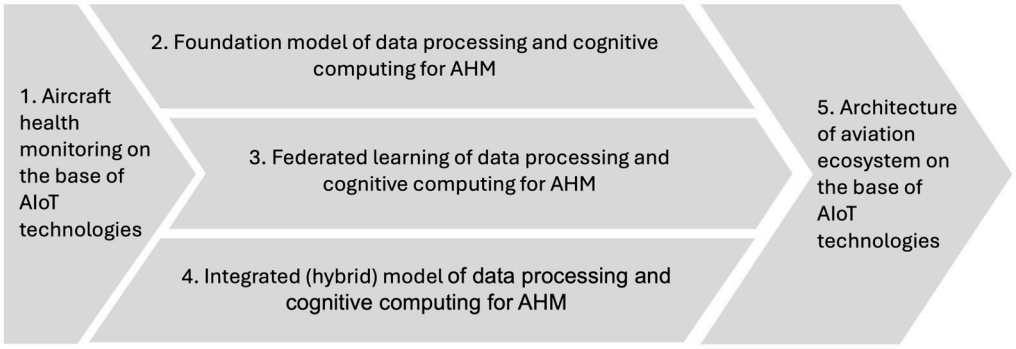


Figure 1. Study framework for collaborative intelligence in AHM.

The framework begins with the fundamental application of AIoT technologies in aircraft health monitoring. This stage sets the groundwork for data collection and initial processing. AIoT in this context refers to the integration of IoT devices with AI capabilities directly on aircraft. These technologies enable real-time data collection from various aircraft systems and components, creating a rich dataset for analysis and decision making.

Building upon the AIoT base, the framework introduces foundation models for data processing and cognitive computing. These models represent a centralized approach to ML in the aviation ecosystem. Foundation models are trained on large, diverse datasets aggregated from multiple aircraft and flights. They capture broad patterns and insights applicable across various aircraft types and operational scenarios. This centralized approach allows for the development of robust, generalizable models that can serve as a baseline for aircraft health monitoring across a fleet.

A foundation model for AHM integrates data from multiple sources, such as sensors, maintenance logs, and flight operations data, to build a generalized understanding of aircraft health and performance. By training on comprehensive datasets, the FM captures patterns and trends across different aircraft and operational conditions, enabling accurate predictions and insights into system health. In this essay, we will explore the structure,

functionality, and benefits of the FM as applied to AHM, highlighting its role in data processing and cognitive computing.

The framework then incorporates federated learning, which represents a paradigm shift towards decentralized machine learning. In this approach, individual aircraft or groups of aircraft train models on their local data without sharing raw information. This method preserves data privacy and reduces the need for massive data transfers. The locally trained models are then aggregated to create a global model that benefits from the collective learning of the entire fleet. Federated learning is particularly valuable in aviation due to the sensitive nature of operational data and the diverse conditions each aircraft experiences.

FL fundamentally changes how data are processed in AHMS by decentralizing the model training process. Rather than aggregating all the data from multiple aircraft into a central server for model training, FL allows each aircraft to train a model locally on its own operational and maintenance data. The locally trained models are then shared with a central server, which aggregates the model parameters (such as weights) rather than the raw data.

The framework culminates in an integrated or hybrid model that combines the strengths of both FMs and FL. This approach uses the broad insights from FMs while allowing for personalization and adaptation through FL techniques. In practice, this could involve starting with an FM trained on centralized data, which are then fine-tuned using FL based on the specific operational data of each aircraft. This hybrid approach optimizes the balance between generalized knowledge and aircraft-specific insights.

The integrated model for AHM combines the strengths of both FMs and FL into a hybrid framework that enhances data processing and cognitive computing across the aviation ecosystem. This model ensures that the broad generalization capabilities of the FM are augmented by the localized learning and fine-tuning provided by FL, creating a system that is both scalable and adaptable to the needs of individual aircraft.

The final component of the framework is the overarching architecture of the aviation ecosystem built on AIoT technologies. This architecture integrates all the previous elements into a cohesive system. It encompasses not only the aircraft themselves but also ground systems, communication networks, data centers, and various stakeholders in the aviation industry. This comprehensive ecosystem approach ensures that the insights generated from aircraft health monitoring can be effectively utilized across the entire aviation value chain.

3. Results

3.1. Aircraft Health Monitoring on the Base of AIoT Technologies

The integration of AIoT technologies in AHM has significantly transformed the way modern aircraft are monitored, maintained, and operated. AIoT-based AHM systems continuously monitor aircraft systems, process large quantities of data, and provide actionable insights in real time. This enables early fault detection and predictive decision making, thus improving operational safety, reducing unplanned downtime, and optimizing aircraft performance. The architecture of AIoT-based AHM is composed of several interrelated components:

- Sensors installed throughout the aircraft monitor critical subsystems such as engines, avionics, hydraulics, and environmental systems. These sensors collect data on parameters like temperature, pressure, vibration, fuel consumption, and more, which are then transmitted to edge computing devices or cloud systems for further analysis. By gathering continuous data on the health of various systems, these sensors provide real-time insights into potential malfunctions or areas of concern.
- Edge computing plays a pivotal role in real-time data processing within the aircraft. With the vast amounts of data generated by sensors, processing some of these data locally through onboard computing devices reduces the need to transmit large datasets to the cloud. AI models embedded in edge devices perform initial analysis, such as predicting component failure or detecting anomalies in the system. This local

processing improves the speed of decision making, ensuring critical actions are taken without delay.

- AIoT has the ability to analyze massive datasets and generate insights with AI-driven analytics. AI algorithms process both historical and real-time data to make predictions about system health, maintenance needs, and operational optimizations. ML models are trained on historical maintenance logs, operational data, and sensor readings to identify patterns of failure and predict future issues. Additionally, cognitive AI systems can interpret unstructured data such as pilot reports or technician notes and correlate this information with sensor data to generate actionable insights.
- Cloud platforms serve as central repositories for processing and storing the large volumes of data generated by aircraft sensors. These platforms provide the computational power required for training AI models and running advanced data analytics. Data lakes store the data from multiple aircraft, allowing for aggregated analysis and continuous improvement of AI models. With centralized data, aviation stakeholders can collaborate to enhance aircraft performance, monitor fleet health, and implement long-term improvements.
- Communication infrastructure is vital for the successful operation of AIoT-based AHMS. Data must be transmitted securely between aircraft and ground stations, and insights must be relayed in real time to relevant stakeholders. Communication networks include satellite communication (SATCOM), which allows data transmission during flights over remote regions, and 5G/6G networks, which provide high-speed connections for data uploads when the aircraft is on the ground. These networks enable seamless data flow between the aircraft, ground systems, and cloud platforms, ensuring that information is always up-to-date and available.

3.2. Analysis of FM, FL, and Integrated Model Approaches in AHMS

In the modern aviation industry, the convergence of big data and cognitive computing has paved the way for advanced AHMS. This section explores the FM, FL, and integrated model approaches, comparing how they interact with big data and cognitive computing to enhance aviation systems.

3.2.1. Foundation Model Approach in AHMS

The FM approach relies heavily on centralized data management and processing. In this model, data from various sources across the aviation ecosystem such as aircraft sensors, maintenance logs, and flight operations are aggregated into a centralized data lake. This central repository enables comprehensive analysis and model training, utilizing large-scale datasets. The centralized nature of this approach allows for scalability, as it can manage the volume, velocity, and variety of data characteristic of big data environments.

From a cognitive computing perspective, the FM is a large, pre-trained model that captures general patterns and knowledge from the aggregated data. This pre-training phase uses cognitive computing techniques to simulate human-like understanding and reasoning, allowing the model to generalize across multiple tasks such as anomaly detection, predictive maintenance, and flight optimization. Once trained, the FM can be deployed across the fleet, where it can process new data, make predictions, and support decision making in real time.

In the FM approach, data are centralized and processed in large batches. Let D represent the total dataset collected from various sources, where

$$D = \{d_1, d_2, \ldots, d_n\}$$

Each d_i represents a multidimensional vector of sensor readings, environmental conditions, and historical maintenance logs associated with different aircraft subsystems. The dataset D captures both operational and environmental variables, which are necessary for accurate health monitoring.

Data preprocessing is crucial for reducing noise and ensuring the model is trained on useful features. The preprocessing function P is applied to the dataset to extract meaningful features:

$$D' = P(D)$$

where $D\prime$ is the preprocessed dataset ready for training.

The foundation model M is trained using machine learning techniques, such as deep learning, to identify correlations between input data and aircraft system health outcomes. Training is performed on the preprocessed dataset D':

$$M = T(D')$$

where T represents the training function, typically involving gradient descent optimization over a large neural network architecture. The model learns to minimize a loss function L that measures the discrepancy between predicted outcomes and actual outcomes (e.g., component failures, wear, and tear).

The training process can be represented as the minimization of the loss function L over the parameters θ of the model:

$$\min_{\theta} L(D'; \theta)$$

Once trained, the foundation model is deployed for real-time aircraft monitoring. The model receives continuous data streams from the aircraft's onboard sensors during flight operations. Let x_t represent the real-time input vector of sensor readings at time t. The foundation model processes this input and generates a health prediction $\hat{y}_t$, which could represent probabilities of failure, wear rates, or system anomalies:

$$\hat{y}_t = M(x_t)$$

This output can trigger maintenance alerts or further inspections based on predefined thresholds.

A key function of FM in AHMS is detecting anomalies in system behavior. The model continuously evaluates new data inputs x_t and compares them with the expected system behavior. An anomaly score $S(x_t)$ is computed based on the deviation of real-time sensor data from typical operating conditions:

$$S(x_t) = \|M(x_t) - \overline{y}\|$$

where $\overline{y}$ is the expected system behavior under normal operating conditions. If the anomaly score exceeds a threshold τ, an alert is generated:

$$\text{Alert if } S(x_t) > \tau$$

The FM can also predict future failures by analyzing trends in historical and real-time data. Given the current health state $\hat{y}_t$ and the system's operational history, the FM can predict the time to failure t_f:

$$t_f = f(M(x_t), H_t)$$

where H_t represents the historical data, and f is a predictive function that estimates the time remaining before a component reaches a critical state.

The FM can be continuously fine-tuned using new data from multiple aircraft. Let D_{new} represent new incoming data, as follows:

$$M\prime = T(D' \cup D_{new})$$

This allows the model to adapt to changing operational conditions and maintain accuracy over time.

Cognitive capabilities are developed through the model M, which can generalize across various tasks:

$$O = I(M, D_{new})$$

Foundation models provide a powerful approach to aircraft health monitoring by processing large-scale data and enabling predictive maintenance. The mathematical framework of FM focuses on data aggregation, preprocessing, training, real-time monitoring, anomaly detection, and predictive maintenance scheduling. By using continuous learning, FMs can maintain high performance and adaptability in diverse operational environments.

Figure 2 represents the aviation ecosystem using the foundation model approach. This diagram illustrates how the various layers interact within the three-domain infrastructure model.

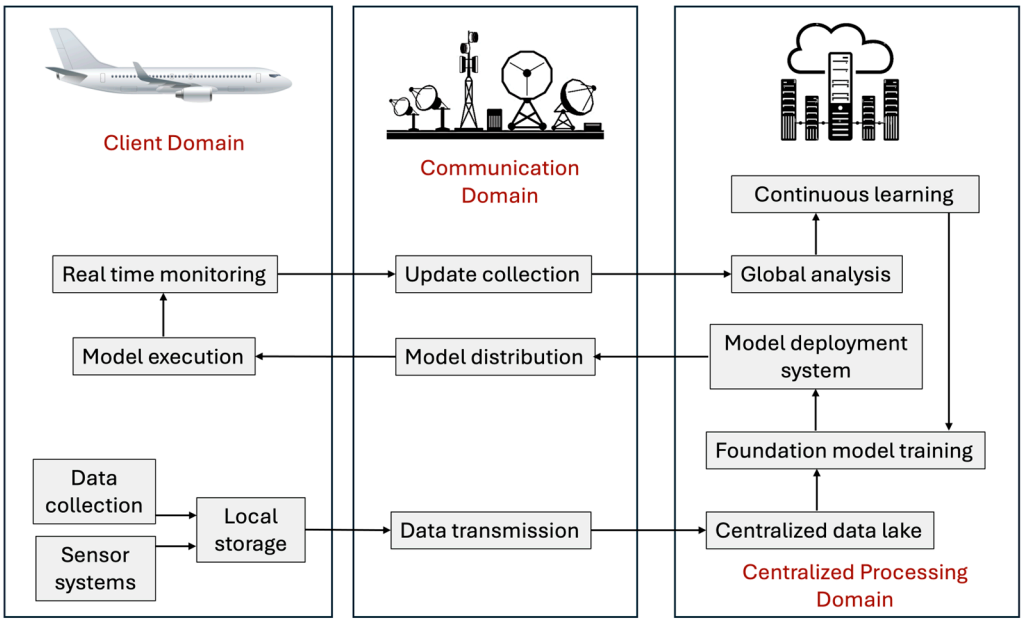


Figure 2. Aviation ecosystem with foundation model approach.

The above diagram shows how the FM approach relies heavily on the centralized component for model training and improvement, while using the decentralized components (aircraft and aviation services providers) for data collection and model execution. The communication channels play a crucial role in facilitating the exchange of data and models between these components.

One of the defining features of the FM for AHM is its ability to support cognitive computing, i.e., the simulation of human-like reasoning and decision-making processes. Cognitive computing allows the model to analyze data, generate insights, and make predictions that go beyond basic pattern recognition, incorporating elements of understanding, reasoning, and learning.

Anomaly detection is a key function of the FM. By analyzing historical data and real-time sensor readings, the model can detect deviations from normal operating conditions, indicating potential system failures. For example, an engine sensor may show an unusual increase in vibration, which could be an early sign of mechanical wear. The foundation model flags this anomaly and generates an alert, prompting further inspection and potential maintenance actions.

In addition to anomaly detection, the FM excels at predictive maintenance. By analyzing patterns in sensor data over time, the model can predict when a component is likely to fail and recommend maintenance before the failure occurs. This proactive approach minimizes unexpected downtime, reduces the risk of in-flight issues, and optimizes maintenance schedules by ensuring that parts are only replaced when necessary.

The FM provides real-time decision support to pilots, ground crews, and maintenance teams. During flight operations, the model continuously processes incoming data from aircraft systems, generating insights and recommendations in real time. For example, if a sensor detects abnormal behavior in the hydraulic system, the model can assess the severity of the issue and suggest corrective actions, such as adjusting flight parameters or preparing for maintenance after landing.

This real-time capability is powered by edge computing, where parts of the model are deployed directly on the aircraft to enable immediate data processing and decision-making. The foundation model's cognitive computing capabilities allow it to prioritize actions based on the severity of issues and the current operational context, ensuring that critical decisions are made quickly and accurately.

3.2.2. Federated Learning Approach in AHMS

In contrast to the centralized FM approach, the federated learning approach operates on a decentralized model. Data are kept locally on each client, such as individual aircraft, addressing significant concerns related to data privacy and security. This decentralized approach fundamentally alters the interaction with big data. Instead of aggregating data in a central repository, each client processes and analyzes its data locally, requiring robust local storage and computing capabilities.

FL reduces the need for centralized data storage by focusing on local data processing and sharing only model updates, such as gradients or weights, with a central server. This approach mitigates communication overhead and preserves privacy, as raw data never leave the local environment. However, managing these decentralized data flows and ensuring effective communication across the network introduces new complexities, especially when dealing with the diverse and heterogeneous data typical in aviation systems.

From a cognitive computing standpoint, FL empowers each client to develop cognitive capabilities that are highly specialized to its local environment. These localized cognitive models are then aggregated into a global model, representing a collective intelligence that benefits from the diverse experiences across the fleet. This approach aligns well with privacy-preserving cognitive computing principles, allowing the development of sophisticated cognitive models without compromising data privacy.

In the FL approach, data D_i

$$D_i = \left\{ d_1^i, d_2^i, \ldots, d_{n_i}^i \right\}$$

represent the local dataset on the i-th aircraft, where n_i is the number of data points available on aircraft i, and N is the total number of participating aircraft. Each local dataset D_i consists of sensor readings, operational logs, and historical maintenance records.

Each aircraft independently trains a local model on its dataset. Let $M_i(\theta_i)$ represent the local model on the i-th aircraft, parameterized by θ_i. The loss function $L_i(\theta_i, D_i)$ on the i-th aircraft is minimized locally using gradient descent:

$$\theta_i^{(t+1)} = \theta_i^{(t)} - \eta \nabla_{\theta_i} L_i(\theta_i, D_i)$$

where $\theta_i^{(t)}$ represents the model parameters at iteration t on the i-th aircraft, η is the learning rate, and $\nabla_{\theta_i} L_i$ is the gradient of the local loss function with respect to the model parameters.

This process allows each aircraft to fine-tune its model based on local data without sharing the raw data themselves.

After each aircraft i trains its local model $M_i(\theta_i)$ t communicates only the model updates (gradients or parameter changes) to a central server, without transmitting the raw data. The central server aggregates the model updates to form a global model M_{global} that represents the knowledge gained from all participating aircraft.

The most used aggregation algorithm in FL is federated averaging. The server computes the global model by taking a weighted average of the local model parameters:

$$\theta_{global}^{(t+1)} = \frac{1}{N}\sum_{i=1}^{N} w_i \theta_i^{(t)} \tag{1}$$

where
$\theta_{global}^{(t+1)}$ represents the global model parameters after aggregation at iteration $t+1$,
$\theta_i^{(t)}$ represents the local model parameters on aircraft i at iteration t,
$w_i = \frac{n_i}{\sum_{i=1}^{N} n_i}$ is the weight assigned to each aircraft, proportional to the size of its local dataset n_i,
N is the total number of participating aircraft.

After each global model update, the new global parameters $\theta_{global}^{(t+1)}$ are sent back to each aircraft, where they are used to update the local models. Aircraft can proceed with either of the following:

- Directly replacing their local models with the global model:

$$\theta_i^{(t+1)} = \theta_{global}^{(t+1)}$$

- Fine-tuning the global model using their local data:

$$\theta_i^{(t+1)} = \theta_{global}^{(t+1)} - \eta \nabla_{\theta_i} L_i\left(\theta_{global}^{(t+1)}, D_i\right)$$

This fine-tuning process allows for personalization, where the global model serves as a starting point but is further adapted to local operational conditions and aircraft-specific data.

Since transmitting large models frequently can be bandwidth-intensive, especially in the context of aviation where satellite communication may be required, several strategies are employed to reduce communication overhead.

- Model compression, where only significant changes in the model parameters (i.e., those above a certain threshold) are transmitted, thus reducing the amount of data shared.
- Sparse updates, where instead of sharing all parameters, only a fraction of parameters that have undergone the most significant updates are transmitted.
- Periodic communication, where instead of transmitting updates after every local iteration, updates are shared after a fixed number of local training steps.

Mathematically, if $\theta_i^{(t)}$ is the local model at iteration t, the model update is computed as follows:

$$\Delta\theta_i^{(t)} = \theta_i^{(t)} - \theta_{global}^{(t)}$$

Then, only a compressed or sparse version of $\Delta\theta_i^{(t)}$ is transmitted to the server.

In the context of AHMS, the global objective is to minimize the aggregate loss across all participating aircraft. The central server's role is to iteratively minimize this global loss by aggregating local updates from the aircraft and updating the global model accordingly.

In AHMS, different aircraft may experience vastly different operating conditions, leading to non-identically distributed data across the fleet. This poses a challenge for the FL model, as global models may struggle to generalize across all aircraft when local data distributions differ significantly.

Local model training on each client is represented by T_i:

$$M_i = T_i\left(D'_i\right)$$

where M_i is the locally trained model on client i.

The local models M_i are aggregated to form a global model M_{global}. The aggregation function A can be represented as follows:

$$M_{global} = A(M_1, M_2, \ldots, M_n)$$

where n is the number of clients (aircraft).

The global model M_{global} represents a collective cognitive capability, using the diverse experiences across all clients.

Real-time inference on each client uses the locally trained model:

$$O_{rt,i} = I_i(M_i, x_i)$$

where x_i is the real-time data input on client i, and $O_{rt,i}$.

Federated learning in AHMS provides a privacy-preserving, scalable, and efficient framework for predictive maintenance and anomaly detection across a fleet of aircraft. By combining local model training on each aircraft with global aggregation, FL ensures that sensitive operational data remain localized while still benefiting from the collective knowledge of the entire fleet. The mathematical framework of FL focuses on local training, global model aggregation, efficient communication, and strategies for handling non-IID data, making it a robust solution for modern AHMS.

Figure 3 represents the aviation ecosystem using the federated learning approach.

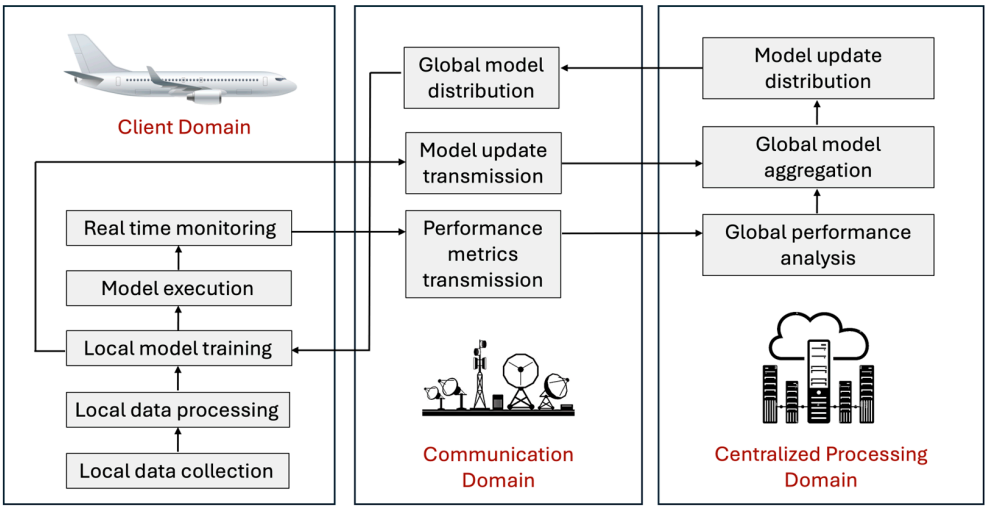


Figure 3. Aviation ecosystem with FL model approach.

There are some key differences from the FM approach:

- More processes occur in the decentralized component (on aircraft), including model training.
- The centralized component is lighter, focused on aggregation rather than direct model training.
- Communication channels primarily transmit model updates and aggregated models, not raw data.
- There is a more prominent feedback loop between local training and global aggregation.

The Figure 3 shows how the FL approach distributes the computational load and preserves data privacy by keeping most operations on individual aircraft, while still benefiting from fleet-wide learning through the aggregation process.

FL in AHM is not just about decentralized data processing; it also integrates cognitive computing to enable higher-level decision making and reasoning. Cognitive computing allows the system to go beyond simple data analysis by mimicking human-like

thought processes, such as understanding context, learning from new data, and making informed decisions.

With FL, each aircraft can generate cognitive insights that are tailored to its specific operational environment. Cognitive computing enables the local model to interpret complex patterns in data, learn from past experiences, and provide actionable insights for maintenance and operations.

For example, if an aircraft consistently experiences high levels of turbulence during certain flight paths, the local model may learn to correlate these conditions with increased wear on certain components, such as the landing gear or hydraulic systems. The cognitive system can then recommend specific maintenance actions based on these insights, ensuring that the aircraft remains in optimal condition even under challenging operational conditions.

These localized insights are valuable because they reflect the unique experiences of each aircraft. By processing data on the edge, the cognitive system can provide real-time recommendations that are specific to the aircraft's current flight and maintenance history.

While each aircraft benefits from localized cognitive insights, FL allows these insights to be shared with the broader fleet through the aggregation of model updates. After local models are trained on individual aircraft, the model updates are sent to a central server, where they are aggregated to create a global model.

This global model benefits from the collective intelligence of all participating aircraft. Each aircraft contributes to the overall knowledge of the system, enabling the global model to learn from the diverse operational experiences of the entire fleet. The global model is then redistributed to each aircraft, where it can be further fine-tuned based on local data. This cycle of learning ensures that the model continuously improves and adapts over time, both at the local and global levels.

The collaborative nature of FL enables the AHMS to provide cognitive insights that are not only relevant to individual aircraft but also reflective of the broader operational environment. This enhances the overall decision-making capabilities of the system, enabling fleet-wide optimizations in maintenance planning, operational efficiency, and safety.

3.2.3. Integrated (FM with FL) Model Approach

The integrated model approach combines the strengths of both the FM and FL, creating a hybrid system that uses both centralized and decentralized big data processing. Initially, large-scale data are centralized to train an FM, which provides a general-purpose cognitive baseline. This centralized training allows the system to exploit the benefits of big data processing, such as deep learning on extensive datasets, resulting in a robust FL that captures broad patterns across the aviation ecosystem.

Once the FL is deployed to individual aircraft, the system shifts to a decentralized mode, where each client uses its local data to fine-tune the model. This localized fine-tuning ensures that the cognitive model is not only generalized but also adapted to the specific operational environment of each aircraft. FL is then employed to periodically aggregate these locally fine-tuned models into an updated global model. This process allows the global model to benefit from the collective knowledge across the fleet, while still maintaining the privacy and security advantages of decentralized data processing.

The integrated model approach represents a balance between the centralized and decentralized paradigms, combining the scalability and generalization capabilities of the FM with the privacy-preserving and context-specific learning capabilities of FL. In terms of cognitive computing, this approach enables the development of both general and specialized cognitive capabilities. The FM provides baseline cognitive intelligence, which is then enhanced and adapted through localized fine-tuning and continuous learning.

The integrated model approach combines centralized and decentralized data management.

The foundation model serves as the baseline model, pre-trained on a large, diverse, and centralized dataset D_{global}, which aggregates data from various sources, including

historical maintenance records, sensor data, operational logs, and external factors such as environmental conditions.

Let θ_{FM} represent the parameters of the FM. The model is trained using a machine learning algorithm that minimizes a global loss function L_{global} based on the centralized dataset D_{global}:

$$\theta_{FM} = \arg\min_{\theta} L_{global}\left(\theta,\ D_{global}\right)$$

where θ_{FM} represents the optimized parameters of the foundation model after training, $L_{global}\left(\theta, D_{global}\right)$ is the loss function that measures the discrepancy between the model predictions and actual outcomes in the global dataset.

The FM can capture generalized patterns across diverse aircraft and operational environments, providing a robust baseline for further personalization using FL.

After the centralized FM is trained, it is deployed to individual aircraft. Each aircraft receives the global model θ_{FM} and uses its local dataset D_i to fine-tune the model to better reflect its specific operational conditions. This is achieved through federated learning, where each aircraft updates the model based on local data.

Let θ_i represent the local model on the i-th aircraft. The local training process on aircraft i involves minimizing a local loss function L_i, which is adapted to the local dataset D_i:

$$\theta_i^{(t+1)} = \theta_{FM}^{(t)} - \eta \nabla_{\theta_i} L_i\left(\theta_{FM}^{(t+1)}, D_i\right)$$

where η is the learning rate, $L_i(\theta, D_i)$ represents the local loss function for the i-th aircraft, and $\theta_i^{(t+1)}$ are the updated model parameters after fine-tuning on the local dataset D_i.

This process allows each aircraft to personalize the global FM model based on its own operational history, capturing aircraft-specific nuances that the global model may not fully reflect.

After the local models have been fine-tuned using the local data, each aircraft sends only the model updates (not the raw data) back to the central server. The central server aggregates these updates using a weighted averaging method, commonly referred to as federated averaging.

The aggregated global model is computed in accordance with (1).

The global aggregation combines the knowledge from all participating aircraft, ensuring that the updated model reflects both the general patterns learned by the FM and the specific adaptations from local datasets.

The integrated approach operates in a continuous learning cycle. After each round of global aggregation, the updated global model $\theta_{global}^{(t+1)}$ is redistributed to all aircraft for further fine-tuning on their local data. This iterative process allows the model to continuously improve over time, adapting to new data, changing operational conditions, and evolving system health patterns across the fleet.

At each iteration $t+1$, the updated global model is as follows:

$$\theta_{global}^{(t+1)} = \frac{1}{N}\sum_{i=1}^{N} w_i\left[\theta_{global}^{(t)} - \eta \nabla_{\theta_i} L_i\left(\theta_{global}^{(t)}, D_i\right)\right]$$

The iteration continues until the global loss function L_{global} converges or reaches a desired threshold.

In the integrated approach, the overall objective is to minimize both the global loss on the centralized dataset and the local loss on each aircraft. The total loss function L_{total} can be expressed as a combination of the global loss L_{global} and the local losses L_i from the individual aircraft. This combined loss function ensures that the global model captures both general and specific patterns, optimizing for fleet-wide performance as well as aircraft-specific needs.

In AHMS, aircraft may experience significantly different operating conditions, leading to non-identically distributed (non-IID) data. The integrated FM + FL approach is well suited to handling such variations, as it allows each aircraft to fine-tune the FM using its specific data.

A key aspect of handling non-IID data is to ensure that the local models do not diverge excessively from the global model. To achieve this, a regularization term can be added to the local loss function, encouraging the local models to stay close to the global model:

$$L_i^{regularized}(\theta_i, D_i) = L_i(\theta_i, D_i) + \lambda \left\| \theta_i - \theta_{global} \right\|^2$$

where λ is the regularization parameter, which controls the trade-off between local adaptation and global consistency.

This regularization term helps prevent overfitting to local data and ensures that the local models maintain alignment with the global model, while still allowing for personalization based on local operating conditions.

The integrated FM + FL approach in AHMS provides a powerful solution that balances the need for generalization across a fleet with the necessity for personalized adaptations to individual aircraft. By combining the global knowledge captured by the foundation model with the local adaptations made possible by federated learning, this approach ensures robust and accurate aircraft health monitoring, predictive maintenance, and anomaly detection across diverse operational environments. The mathematical framework includes global model training, local fine-tuning, iterative aggregation, and communication efficiency strategies, making it highly suitable for complex, real-time systems like AHMS.

Figure 4 represents the aviation ecosystem using the integrated model approach, which combines elements of both the FM and FL approaches.

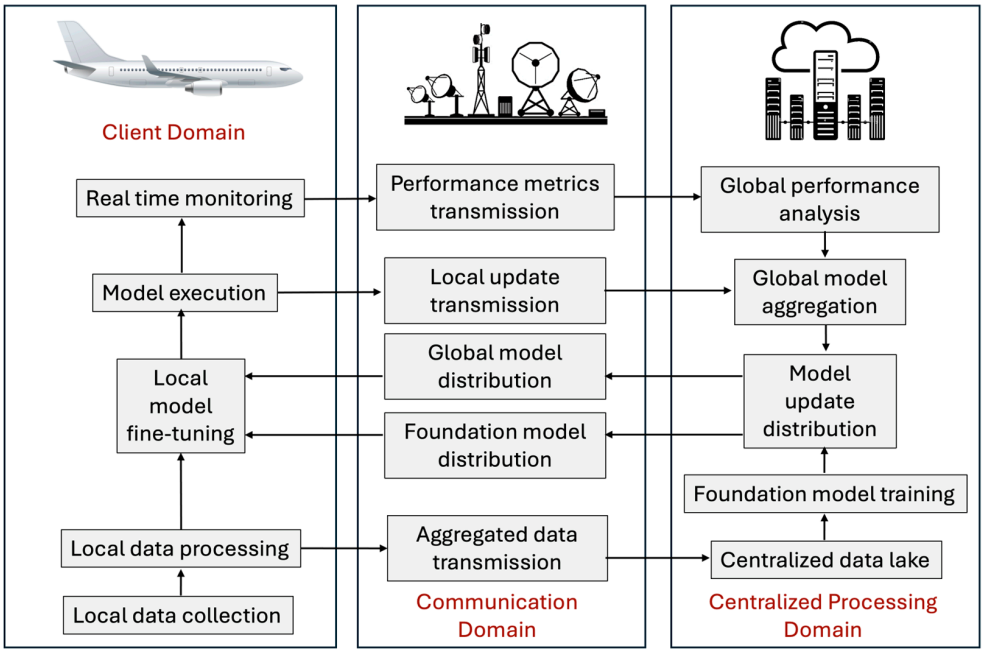


Figure 4. Aviation ecosystem with integrated model approach.

Figure 4 highlights several key aspects of the integrated model approach.

- The hybrid nature of the system, combining centralized foundation model training with decentralized federated learning.
- The dual role of the central component in both FM training and FL aggregation.
- The enhanced role of communication channels, facilitating various types of data and model exchanges.

- The continuous cycle of centralized training, local fine-tuning, global aggregation, and model redistribution.

There are key differences of integrated model from both FM and FL pure approaches, detailed as follows:

- It includes both a centralized data lake for FM training and local processing for FL.
- The central component performs both FM training and federated aggregation.
- Communication channels are more diverse, handling raw data, foundation models, local updates, and aggregated models.
- There is a more complex feedback loop involving both centralized and decentralized learning processes.

Integrated approach allows for a robust global model trained on comprehensive data while still enabling personalization and privacy-preserving learning on individual aircraft. This hybrid approach enables aircraft to benefit from a global understanding of fleet-wide trends while allowing for personalized model adaptation based on local data.

The integrated model not only enhances data processing but also enables advanced cognitive computing, which involves mimicking human-like reasoning, understanding, and decision-making capabilities. The integrated model enables real-time anomaly detection by continuously processing sensor data from the aircraft. Cognitive computing algorithms are embedded within the model to recognize deviations from normal operating conditions. For instance, if an engine's temperature begins to rise beyond normal thresholds, the model can detect this anomaly and alert the maintenance team before the situation escalates into a critical failure.

The integrated model excels in predictive maintenance. By analyzing historical data and current operational conditions, the model can predict when a specific component is likely to fail. Maintenance teams can use this information to proactively schedule repairs or part replacements, reducing unplanned downtime and improving the overall reliability of the aircraft.

One of the key features of cognitive computing in the integrated model is its ability to provide context-aware decision support. This means that the system not only detects anomalies or maintenance needs but also considers the operational context in which these issues occur. For example, the system might consider factors such as the aircraft's current flight route, altitude, and weather conditions when making recommendations.

This context-aware decision-making capability ensures that the actions recommended by the model are relevant and timely. In some cases, the model might suggest adjustments to the flight path to avoid excessive turbulence, which could impact sensitive components such as landing gear or hydraulic systems.

The integrated model benefits from adaptive learning, meaning that it continuously improves its performance based on new data from both individual aircraft and the entire fleet. The model is constantly updated through the FL process, ensuring that it evolves in response to changing operational conditions and emerging trends in aircraft health.

This continuous learning capability is crucial for maintaining the accuracy and relevance of the model over time. As new aircraft are introduced into the fleet or as environmental conditions change, the model can adapt to these changes without requiring a complete retraining from scratch. This adaptive nature makes the integrated model both efficient and scalable.

3.2.4. Loss Function Structure

It is possible to define a universal form of the loss function for all three approaches (FM, FL, and the integrated FM + FL) in the context of AHMS. However, the specific adaptation and application of the loss function will vary depending on the architecture and characteristics of each approach. The core structure of the loss function can remain consistent across all three methods, while modifications are made based on how data are processed and models are trained (centrally, locally, or in combination).

1. Universal Loss Function Structure

At its core, the loss function in AHMS aims to minimize prediction errors (e.g., anomalies, failures, or continuous health metrics). A universal loss function can be expressed as follows:

$$L_{total}(\theta) = \sum_{i=1}^{N} w_i L_i(\theta,\ D_i)$$

where $L_i(\theta,\ D_i)$ is the local loss function on the i-th dataset D_i, w_i are the weights that control the contribution of each aircraft's local loss to the total loss, and θ represents the model parameters being optimized (a global model for FM and integrated approaches, and local models for FL).

This structure can accommodate the following:

- Classification tasks (binary for anomaly detection or multi-class for failure modes).
- Regression tasks (continuous health metrics like remaining useful life).

Now, let us look at how this universal loss function is applied in each approach.

2. Foundation Models

In the FM approach, the loss function is primarily focused on centralized data processing. The model is trained on a large dataset D_{global}, and the loss is computed globally across the entire dataset:

$$L_{FM}\left(\theta_{global}^{(t+1)}, D_{global}\right) = \sum_{i=1}^{N} w_i L_i\left(\theta_{global}, D_i\right)$$

This global loss function allows the model to capture general patterns across the fleet of aircraft. Depending on the task, $L_i(\theta, D_i)$ could be one of the following:

- Binary cross-entropy for anomaly detection (binary classification).
- Categorical cross-entropy for multi-class failure mode classification.
- Mean squared error (MSE) or mean absolute error (MAE) for predicting continuous metrics.

3. Federated Learning

In the FL approach, the loss function is computed locally on each aircraft and optimized independently. Each aircraft i has its local dataset D_i, and the model parameters θ_i are updated based on local data. The global loss is then a weighted average of these local loss functions:

$$L_{FL}\left(\theta_{global}\right) = \sum_{i=1}^{N} w_i L_i(\theta_i, D_i)$$

After local optimization, the model updates are aggregated centrally. Each local loss function L_i will be similar to the loss used in the FM approach, but applied to local data, as follows:

- Binary cross-entropy for anomaly detection.
- Categorical cross-entropy for multi-class classification.
- MSE/MAE for regression tasks.

FL allows each aircraft to optimize its own loss function independently based on local conditions while contributing to the global model.

4. Integrated FM + FL Approach

In the integrated FM + FL approach, the loss function has both global and local components. The global FM is first trained on a large dataset and then fine-tuned on each

aircraft using FL. The total loss can be represented as a combination of the global loss (from the FM) and the local losses (from FL):

$$L_{FM+FL}(\theta) = \alpha L_{global}\left(\theta_{global}, D_{global}\right) + \sum_{i=1}^{N} \beta_i L_i(\theta_i, D_i)$$

where

$L_{global}\left(\theta_{global}, D_{global}\right)$ is the global loss from the FM trained on the centralized dataset;

$L_i(\theta_i, D_i)$ is the local loss for each aircraft, fine-tuning the model using FL; and

a and β_i are weights that control the contribution of global vs. local losses.

Like the previous two approaches, L_i can take the form of binary cross-entropy, categorical cross-entropy, or MSE/MAE, depending on the specific task being performed on each aircraft.

While the core structure of the loss function is universal across FM, FL, and FM + FL approaches, it can be adapted to specific tasks, as follows:

- Anomaly detection—binary cross-entropy for detecting abnormal states.
- Failure mode classification—categorical cross-entropy for classifying types of failures.
- Health metric prediction—MSE or MAE for continuous predictions like remaining useful life or wear levels.

These specific adaptations affect how the loss is computed but follow the same overall structure across all approaches.

While the core mathematical structure of the loss function is universal across FM, FL, and FM + FL approaches, the way it is applied and optimized differs. FM focuses on global optimization, FL emphasizes local optimization with subsequent global aggregation, and FM + FL combines both global and local optimization strategies. The specific form of the loss function (binary, multi-class, or regression) is determined by the type of task but can be integrated into a common framework for all three approaches.

3.3. Architecture of an AIoT-Based Aviation Ecosystem Integrating FM and FL Approaches

The modern aviation industry requires sophisticated systems to ensure the safety, reliability, and efficiency of aircraft operations. To achieve this, an integrated approach combining various advanced technologies such as FMs, FL and the AIoT platform has been proposed. This approach harnesses the strengths of these technologies, offering a robust and scalable architecture capable of addressing the complex challenges faced by AHMS.

The Figure 5 represents the architecture of comprehensive solution for modern AHM, integrating the capabilities of modern technologies within a unified framework.

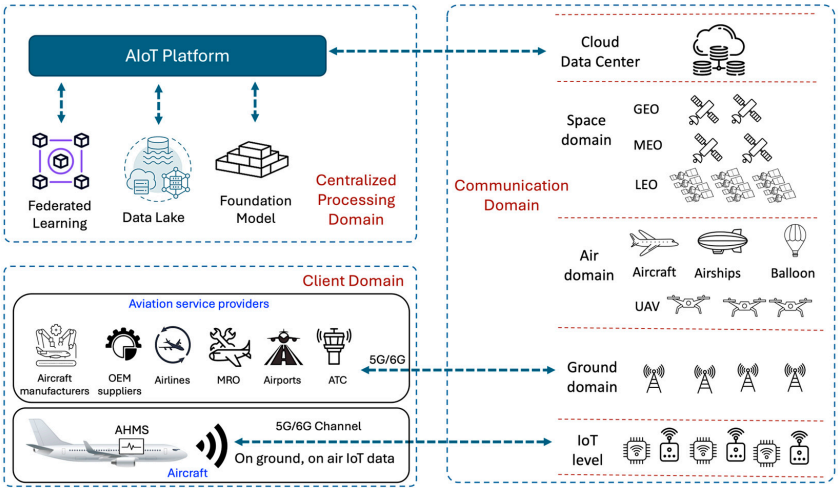


Figure 5. Architecture of aviation ecosystem for AHM.

The architecture is divided into three main domains:

1. Centralized processing domain;
2. Client domain;
3. Communication domain.

Each domain plays a crucial role in the operation of the aviation ecosystem, ensuring that data are collected, processed, and utilized effectively across the entire system.

1. Centralized processing domain

This domain is the heart of the architecture, where large-scale data processing and model training take place. It comprises three key components, detailed as follows:

- The FM serves as the core of the system, providing a general-purpose cognitive baseline that is trained on a comprehensive dataset aggregated from various sources within the aviation ecosystem. This model captures broad patterns and relationships that are applicable across different scenarios and aircraft, making it a crucial component for generalized decision-making and predictive analytics.
- The data lake acts as the central repository where all collected data are stored. These data include aircraft sensor readings, maintenance logs, flight operations data, and external data such as weather reports. The data lake enables extensive data analysis and model training, leveraging big data technologies to manage the volume, velocity, and variety of data.
- While the FM provides a strong starting point, the FL component allows for the decentralized fine-tuning of models on individual aircraft. This ensures that the global model benefits from local adaptations, reflecting the unique conditions experienced by different aircraft. FL aggregates the locally fine-tuned models to update the global model periodically, preserving data privacy and reducing communication overhead.
- The AIoT platform integrates these components, enabling seamless interaction between centralized data processing, FM training, and FL. The platform manages the orchestration of these processes, ensuring that data flow smoothly between components and that the system remains responsive and adaptable.

2. Client domain

The client domain consists of various stakeholders within the aviation ecosystem, including aircraft manufacturers; original equipment manufacturer (OEM) suppliers; airlines; maintenance, repair, and overhaul (MRO) providers; airports; and air traffic control (ATC).

Each stakeholder generates and processes data relevant to their operations. Aircraft, equipped with AHMS, collect real-time data on both the ground and in-flight. These data are processed locally using edge computing devices, where the foundation model is fine-tuned to reflect the specific conditions of each aircraft.

The AHMS on each aircraft operates in real time, using the fine-tuned model to provide insights such as predictive maintenance recommendations, anomaly detection, and operational optimization. The results are shared with relevant stakeholders, ensuring that decisions are informed by the latest data and model predictions.

3. Communication domain

The communication domain facilitates data transfer between the client domain and the centralized processing domain. It encompasses several levels, detailed as follows:

- The cloud data center acts as the central hub for storing and processing data, as well as hosting the foundation model and federated learning algorithms.
- The space domain includes communication satellites in GEO, MEO, and LEO orbits, which provide global connectivity for data transmission, especially in remote or airborne operations.
- The air domain encompasses the various aircraft types, including traditional aircraft, UAVs, airships, and balloons. These platforms generate data and require communication links for real-time data transfer and model updates.

- The ground domain consists of ground stations and cellular networks (e.g., 5G/6G) that facilitate communication between airborne platforms and the central data center.
- The IoT level represents the IoT devices embedded within aircraft and ground infrastructure, which collect and transmit data to the AIoT platform for processing.

The communication layer is crucial for maintaining the integrity of the system. It ensures that data collected by aircraft are securely transmitted to the centralized processing domain and that updates to the global model are efficiently distributed back to the clients.

4. Integration of models within the architecture

This architecture uses the strengths of FMs, FL and AIoT to create a resilient and adaptive aviation ecosystem. The FM provides a robust baseline for decision making across the fleet, while FL ensures that the model is continuously updated based on the latest operational data from individual aircraft. The AIoT platform orchestrates these processes, managing the data flows and ensuring that the system remains scalable and secure.

By integrating these technologies, the architecture can handle the diverse and dynamic environments encountered in aviation, from the unique conditions faced by individual aircraft to the broader trends captured in the centralized data lake. This multi-layered approach ensures that the AHMS can provide personalized, real-time insights for individual aircraft while using fleet-wide data for continuous improvement.

4. Discussion

4.1. Simulation-Based Case Study

This case study explores three approaches to AHMS: foundation models, federated learning, and an integrated approach that combines elements of both FM and FL. The FM approach uses large, pre-trained models that capture broad patterns across entire fleets. FL, on the other hand, enables personalized model training on individual aircraft while preserving data privacy. The integrated approach aims to harness the strengths of both methods, potentially offering a more robust and adaptable solution.

This study simulates a fleet of 100 aircraft, each equipped with sensors monitoring critical engine parameters. The performance of these three approaches in predicting engine failures is then compared. Through this comparison, this study aims to shed light on the relative strengths and weaknesses of each method, particularly in terms of prediction accuracy, adaptability to individual aircraft characteristics, computational efficiency, and privacy preservation.

By analyzing these approaches in a simulated environment, this research seeks to provide insights that can guide the development and implementation of more effective AHMS in real-world aviation contexts. This case study serves as a practical exploration of the theoretical concepts discussed in the broader research on integrating foundation models and federated learning in AIoT-based aircraft health monitoring systems.

The simulation results and subsequent analysis offer a comprehensive evaluation of each approach, highlighting their respective advantages and limitations. These findings contribute to the ongoing discourse on optimizing AHMS and may inform future directions in aviation safety and maintenance strategies.

4.1.1. Simulation Methodology for AHMS Experiment

The objective of this simulation is to compare the performance of three approaches to AHMS: FM, FL and integrated FM + FL. The evaluation focuses on prediction accuracy, error rates, training efficiency, and adaptability to individual aircraft characteristics.

To evaluate the proposed health monitoring models, we generated synthetic data for 100 aircraft over 1000 flight hours, incorporating a wide range of operational scenarios and simulating realistic aircraft behavior. Each aircraft was assigned unique characteristics, such as engine age, historical maintenance data, and typical flight routes, to introduce variability across the dataset. For each flight hour, synthetic sensor readings were generated, covering

parameters such as engine temperature, vibration levels, fuel consumption, oil pressure, and exhaust gas temperature.

The data generation process involves setting normal operating ranges for each parameter, adding aircraft-specific bias to simulate individual variations, incorporating random noise to mimic sensor fluctuations, introducing gradual drift to represent wear and tear over time, and injecting failure patterns.

In our synthetic dataset, we introduced failure patterns to simulate realistic degradation and fault scenarios for select aircraft, creating nuanced data reflective of potential failure conditions. For five randomly selected aircraft, we emulated gradual component degradation by progressively altering key sensor readings. These alterations were based on known patterns in aircraft maintenance data, simulating conditions such as increased vibration levels, rising engine temperatures, fluctuating oil pressures, and excessive fuel consumption that typically signal early-stage wear and potential failure.

Each failure pattern was designed to exhibit a distinct progression, where deviations from normal sensor ranges began subtly, increasing in severity over the simulated flight hours. For example, minor increases in engine vibration were introduced initially, followed by further increments to reflect progressive mechanical imbalance. Similarly, for oil pressure anomalies, slight pressure reductions over time were embedded to mimic issues like lubrication deficiencies. This staged approach aimed to train the model on early warning signs, offering insights into gradual wear before reaching a critical failure threshold. Such an approach ensures that the model can detect both immediate anomalies and gradual trends, enhancing its predictive power for real-world applications.

For the foundation model implementation, a deep neural network with three hidden layers is employed. The model is trained on 80% of the entire fleet's data and validated on the remaining 20%, using a centralized training process on a simulated high-performance server.

The federated learning approach utilizes a lightweight neural network with two hidden layers. Each aircraft uses its own data for training, with 100 rounds of federated averaging and five epochs of local training per round. The FedAvg algorithm is used for aggregation.

The integrated FM + FL approach begins by training an initial foundation model on 10% of the fleet's anonymized and aggregated data. This model then serves as a starting point for federated learning, with the last layer allowed to be personalized for each aircraft. The process involves 50 rounds of federated averaging, with three epochs of local training per round.

Evaluation metrics for each approach include accuracy, precision, recall, F1 Score, Area Under the ROC Curve (AUC-ROC), training time, model size, and communication overhead (for FL and integrated approaches).

The simulation procedure involves several steps:

1. Generating synthetic data for 100 aircraft over 1000 flight hours, including failure patterns for five randomly selected aircraft.
2. Training and validating the FM on the aggregated data, then deploying it to all aircraft in the test set.
3. Initializing models on each aircraft for FL, performing 100 rounds of federated learning and evaluating the personalized models.
4. Training the initial FM for the integrated approach, performing 50 rounds of federated fine-tuning, and evaluating the personalized models.
5. Using the trained models from each approach to predict engine failures on a held-out test set and calculating all evaluation metrics.
6. Conducting statistical tests to determine the significance of differences between approaches.
7. Performing sensitivity analysis by varying key parameters such as the number of aircraft, failure rate, and noise levels to assess the robustness of each approach under different conditions.

Data analysis and visualization includes creating comparative tables and charts, plotting learning curves, generating receiver operating characteristic (ROC) curves, creating confusion matrices, analyzing feature importance, and examining specific case studies of individual aircraft.

The simulation environment utilizes Python 3.13.0 as the programming language [28] and TensorFlow 2.17 [29] as the deep learning framework. TensorFlow Federated (Release 0.88.0) [30] is used for federated learning implementation. NumPy 2.1 [31] is employed for synthetic data creation, while Pandas 2.2.2 [32], Matplotlib 3.9.0 [33], and Seaborn 0.13 [34] are used for data analysis and visualization.

Potential extensions to the simulation include multi-component health monitoring, simulating dynamic fleet composition, testing transfer learning scenarios, implementing privacy analysis with differential privacy, and simulating varying network conditions to test the robustness of FL and integrated approaches.

This comprehensive simulation methodology provides a rigorous framework for comparing the three AHMS approaches, generating meaningful, quantitative results that demonstrate the strengths and weaknesses of each approach in a realistic, simulated environment.

4.1.2. Results of Simulation

The results of the simulation provide a detailed comparative analysis of the three AHMS approaches: FM, FL and the integrated FM + FL approach. These visualizations are derived from extensive simulations conducted on a fleet of 100 aircraft over a period of 1000 flight hours, with various sensor data points collected to predict potential engine failures. Figures 6–13 collectively illustrate the performance differences and critical evaluation metrics for each approach.

Figure 6 presents the model size comparison, showing that the FM approach requires significantly larger storage and computational resources than FL and the integrated approach. This demonstrates that FL and FM + FL are more suitable for resource-constrained environments in which minimizing model size is a priority.

Figure 7 depicts the training time across the different approaches. The integrated FM + FL approach strikes a balance, reducing the computational burden associated with centralized FM training while maintaining a robust model performance similar to FL. FL exhibits lower overall training time due to its decentralized nature but at the cost of reduced generalization capabilities.

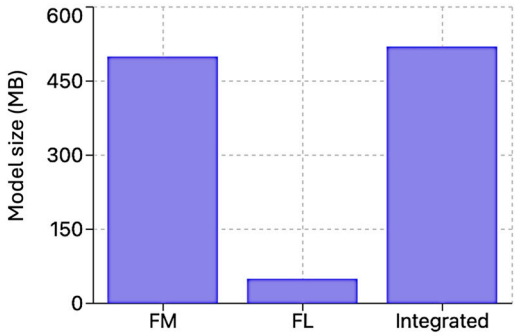


Figure 6. Model size chart.

Figure 8 summarizes the key performance metrics, including accuracy, precision, recall, and F1 score, highlighting that while the FM approach performs best in terms of overall accuracy, the integrated approach offers a comparable performance with better adaptability to individual aircraft variations, as reflected in its improved precision and recall scores.

Figure 9 illustrates the learning curves over multiple training epochs, with the integrated approach demonstrating faster convergence compared to both FM and FL. This indicates that the integrated method efficiently combines the strengths of global model learning with localized fine-tuning, leading to faster and more stable learning.

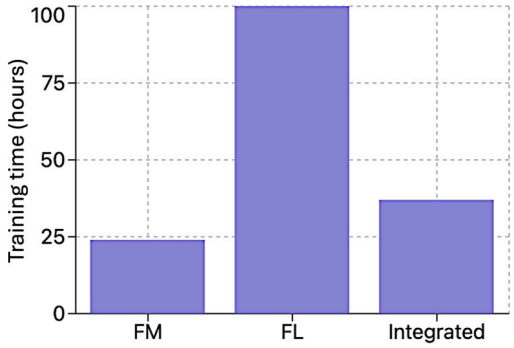


Figure 7. Training time chart.

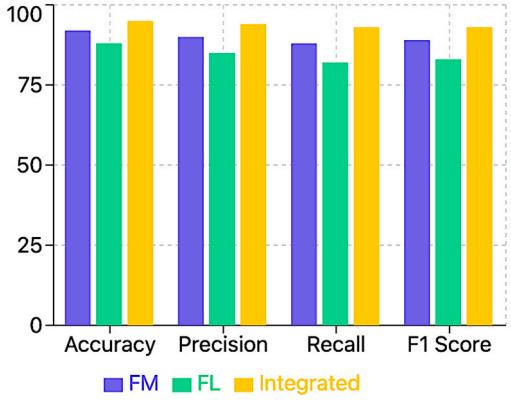


Figure 8. Comparative performance chart displaying key metrics (in percent).

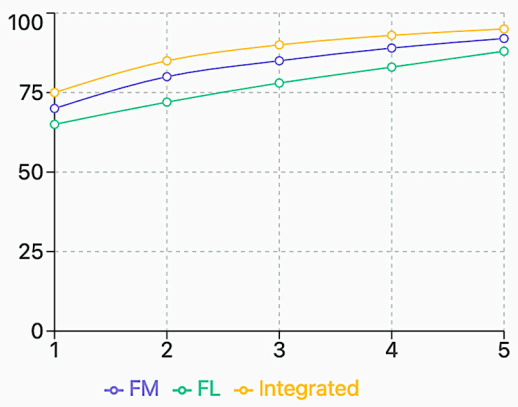


Figure 9. Learning curves that illustrate how model performance improves over training epochs (in percent).

Figure 10 displays the receiver operating characteristic (ROC) curves, where the area under the curve (AUC) is highest for the integrated FM + FL approach, indicating superior performance in distinguishing between normal and anomalous aircraft behavior, crucial for accurate failure predictions.

Figure 11 focuses on error rates, showing that while the FM approach generally maintains a lower overall error rate, the integrated approach balances error minimization across multiple aircraft, resulting in more consistent performance across the fleet.

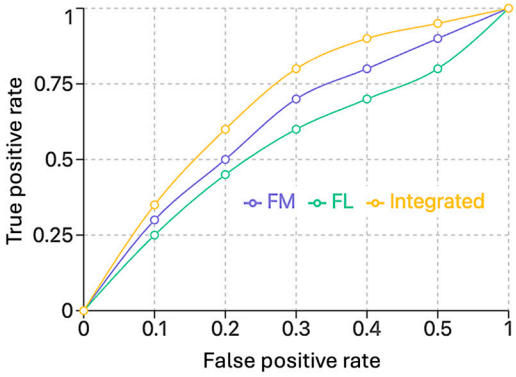


Figure 10. ROC curves.

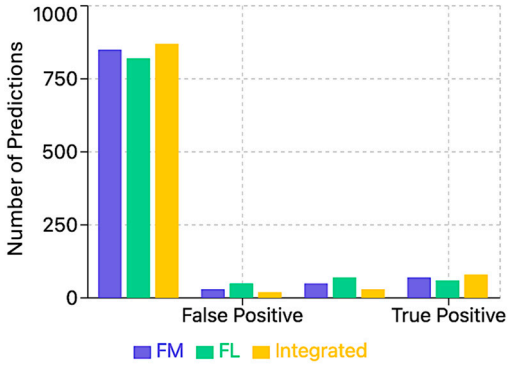


Figure 11. Error rate chart.

Figures 12 and 13 present case studies for a specific aircraft (Aircraft A). Figure 12 ranks the relative significance of various sensor readings in predicting engine failure, with the integrated approach offering more accurate and contextually aware predictions. Figure 13 compares normal operational ranges with actual readings across different engine parameters, highlighting the integrated approach's ability to adjust for individual aircraft characteristics while maintaining fleet-wide accuracy.

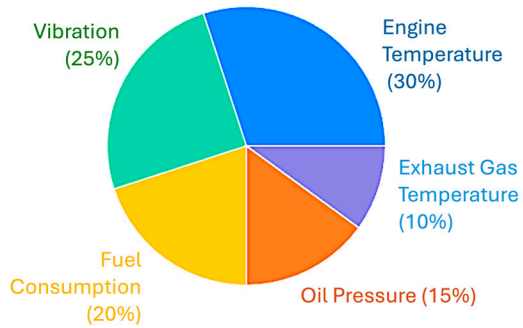


Figure 12. The relative significance of different sensor readings in predicting engine failure (case study for aircraft A).

The analysis of these visual representations provides insights into the relative merits of each approach and the potential benefits of integrating foundation models with federated learning in the context of aircraft health monitoring.

The integrated approach provides a scalable solution that retains high accuracy while optimizing resources, model size, and adaptability to individual aircraft characteristics. This approach demonstrates potential for real-world implementation in predictive aircraft health monitoring systems.

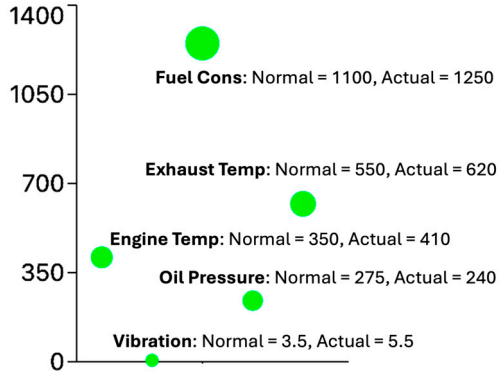


Figure 13. Comparison of normal operating ranges with actual readings across various engine parameters (case study for aircraft A).

4.2. Comparative Analysis of FM, FL and Integrated Approaches in AHMS

The FM approach relies on centralized, large-scale AI models trained on vast datasets encompassing various aspects of aircraft health and performance. This centralization offers significant advantages in terms of global consistency and comprehensive knowledge integration. All users of the system benefit from the same, extensively trained model, ensuring uniformity in health monitoring and predictive maintenance across an entire fleet or even multiple airlines. The ability to quickly incorporate new data and simultaneously update all users is a notable strength, particularly in responding to industry-wide issues or newly discovered fault patterns.

However, the centralized nature of foundation models also introduces vulnerabilities that impact system robustness. The most glaring issue is the risk of a single point of failure. If the central server experiences downtime or network connectivity issues, it could potentially disrupt AHMS operations across the entire user base. This centralization also presents a high-value target for cyber-attacks, with the potential for large-scale data breaches that could compromise sensitive operational data from multiple airlines or fleet operators.

Moreover, while FMs excel at capturing general patterns and common scenarios, they may struggle with local variations or rare conditions not well represented in the central dataset. This limitation could lead to suboptimal performance in unique operational environments or failure to detect anomalies specific to aircraft or routes.

In contrast, the FL approach distributes the learning process across multiple nodes, typically individual aircraft or airline operations centers. This decentralized architecture inherently enhances several aspects of system robustness. Fault tolerance is significantly improved, as the system can continue functioning even if some nodes fail or lose connectivity. This graceful degradation ensures that AHMS capabilities persist, albeit potentially at a reduced capacity, in the face of partial system failures.

FL also shines in its adaptability to local conditions. Each node trains on its local data, allowing the model to capture and respond to specific operational environments, aircraft configurations, or regional factors. This localization can lead to more accurate and relevant health monitoring for individual aircraft or fleets.

From a security and privacy standpoint, FL offers robust protection of sensitive data. Raw operational data remain local, with only model updates being shared. This approach significantly reduces the risk of large-scale data breaches and helps in complying with data protection regulations.

However, the decentralized nature of FL introduces its own set of challenges to system robustness. Maintaining global consistency across all nodes can be difficult, potentially leading to divergent model behaviors in different parts of the system. This inconsistency could result in varying levels of monitoring quality or conflicting maintenance recommendations across a fleet. Additionally, FL systems may struggle with rare event detection, as unusual occurrences might not propagate effectively through the network of locally trained models.

The integrated approach, combining elements of both FL and FL, attempts to harness the strengths of both while mitigating their weaknesses. This hybrid system can start with a foundation model that encapsulates broad industry knowledge, which is then continuously refined and adapted through federated learning techniques.

This integration enhances system robustness in several ways. It reduces the single point of failure risk associated with pure FMs by allowing for continued operation based on local models if connection to the central system is lost. The periodic synchronization between local and global models helps to maintain a balance between global consistency and local optimization. This approach is particularly adept at handling both common and rare scenarios, using the comprehensive knowledge of the FM while remaining adaptable to local conditions.

In terms of security and privacy, the integrated approach maintains many of the advantages of FL. Sensitive operational data primarily remain local, reducing the risk of large-scale data breaches. At the same time, the FM component allows for the incorporation of industry-wide insights and patterns that might not be captured in any single node's local data.

However, the integrated approach is not without its challenges. The complexity of managing both global and local model components can introduce new failure modes and increase the overall system's complexity. Ensuring smooth integration and resolving conflicts between global and local models require sophisticated algorithms and careful system design.

Table 1 compares the benefits and limitations of using an integrated model versus using an FM or FL independently.

Table 1. Comparison of FM, FL, and integrated model (FM + FL) approaches.

Approach	Strengths	Limitations	Suitable Use Cases
Foundation Model	- High generalization across fleet - Consistent model performance across systems - Capable of handling large datasets centrally	- Limited adaptation to individual aircraft - Requires centralized data, raising privacy concerns	Fleets with similar operating conditions where centralized data processing is feasible
Federated Learning	- Preserves privacy by keeping data local - Adapts to specific conditions of individual aircraft - Reduces risk of single point of failure	- Potential inconsistencies due to local learning - Limited fleet-wide generalization for rare events	Scenarios requiring privacy protection and adaptive learning to diverse environments
Integrated (FM + FL)	- Combines generalization of FM with FL's local adaptation - Balances global insights with fleet-wide personalization - Flexible in diverse operational settings	- Added complexity in managing centralized and decentralized components - Requires sophisticated synchronization and computational resources	Large fleets with varied operational conditions and a need for privacy and adaptability

Each approach is analyzed based on its strengths, limitations, and suitable use cases. This table aims to highlight key distinctions and the specific advantages and trade-offs of each method. Notably, the FM approach excels in providing a broad, generalized model but may lack local adaptation and privacy. In contrast, FL is advantageous for privacy-preserving, locally adaptable learning but can encounter challenges with consistency across the fleet. The integrated FM + FL approach offers a hybrid solution, combining global insights with local flexibility, though it introduces added complexity. This comparative analysis supports informed decision making on model selection based on an organization's operational priorities, privacy needs, and computational resources.

To enhance the robustness of AHMS, regardless of the chosen approach, several strategies can be employed. Implementing hybrid connectivity solutions that allow for both online and offline operation can ensure continuity during network disruptions. Multi-level redundancy, whether in server infrastructure for FMs or in node participation for federated systems, can significantly improve fault tolerance.

Advanced anomaly detection systems that monitor not just aircraft health but also model behavior itself can provide an additional layer of robustness, alerting to unexpected model outputs or performance degradation. Employing secure enclave technologies for sensitive computations can enhance security across all approaches.

For federated and integrated systems, adaptive learning rates can help balance between global consistency and local adaptations. Blockchain-based versioning for model updates can ensure integrity and traceability of model evolution. Regular fuzzing and adversarial testing can identify and address potential vulnerabilities, while carefully designed graceful degradation protocols can maintain core functionalities even when the system operates with reduced capabilities.

As aviation continues to embrace AI technologies, the robustness of AHMSs will remain a critical concern. Ongoing research and development in this field should focus on further enhancing the strengths and mitigating the weaknesses of each approach. The future of AHMS likely lies in increasingly sophisticated integrated systems that can provide the global insights of FMs, the adaptability and privacy of FL, and the high-level reliability demanded by aviation safety standards.

4.3. Challenges in Integrating FM and FL

Integrating foundation models (FMs) with federated learning (FL) within an aircraft health monitoring system presents several technical and operational challenges that need to be carefully addressed to ensure a functional and scalable solution.

One of the primary challenges in combining FM and FL is maintaining synchronization across decentralized learning nodes (i.e., individual aircraft) while updating a central, foundation-based model. The decentralized nature of FL relies on local updates that need to be effectively aggregated with the FM to avoid inconsistencies that could reduce model accuracy or robustness. This requires advanced aggregation algorithms and synchronization protocols that can balance local adaptations with global model coherence.

Integrating FM and FL increases the computational demands on both centralized servers and local aircraft devices. For FL, models must be trained on distributed nodes with limited computing capacity. Balancing this demand with the FM's requirement for centralized processing necessitates optimized resource management and, where feasible, the use of lightweight model architectures. Resource constraints are particularly pronounced in real-time applications, where model updates need to be processed without interrupting critical onboard systems.

While FL inherently offers privacy advantages by keeping data localized, integrating it with a centralized FM framework introduces new privacy and security considerations. Protecting sensitive operational data requires secure aggregation protocols and encryption techniques to prevent unauthorized access during model updates and synchronization. This is particularly important in aviation, where data breaches could impact safety and regulatory compliance.

Aircraft operate under diverse conditions, resulting in data distributions that are often non-identically distributed across the fleet. An integrated FM + FL approach must address these variations to ensure the model's predictive accuracy across different operating environments. This is challenging because an FM trained on generalized data may not always capture localized patterns effectively, while FL's decentralized approach might overly adapt to specific conditions. Balancing global generalization with local adaptation requires carefully designed algorithms to handle data heterogeneity without compromising model performance.

These challenges highlight the complexity of designing a robust, scalable FM + FL-based health monitoring system for aviation. Addressing these issues is essential to ensure that the integrated model can effectively meet the diverse operational, privacy, and performance requirements of a real-world aviation context.

4.4. Future Directions of Research in AI-Driven AHMS

As the aviation industry continues to embrace digital transformation, the integration of advanced ML techniques into AHMS represents a critical area of innovation. The combination of FMs and FL has shown significant potential in enhancing the safety, reliability, and efficiency of aircraft operations. However, to fully realize the benefits of these technologies, ongoing research is necessary to address the challenges and limitations inherent in their application. This section explores future directions of research that could advance the field of AHMS, focusing on scalability, robustness, security, and the seamless integration of AI technologies into the complex ecosystem of aviation.

One of the primary challenges in deploying AI-driven AHMS is ensuring scalability across a global fleet of aircraft. As the size and diversity of data increase, so do the demands on computational resources and communication infrastructure. Future research should focus on developing more efficient algorithms for FL that reduce communication overhead and optimize the aggregation of model updates. Techniques such as federated optimization, hierarchical FL and gradient compression could play crucial roles in making these systems more scalable and responsive.

As FMs continue to grow and complexity, there is a need for more efficient model architectures that can deliver high performance without requiring prohibitively large computational resources. Research into lightweight models, such as distilled or quantized versions of FMs, could enable their deployment on edge devices, such as those onboard aircraft, without compromising accuracy or functionality. These advances would facilitate real-time health monitoring and decision making, even in resource-constrained environments.

Ensuring that AHMS can generalize across diverse operational conditions and aircraft types is another key area for future research. While FMs provide a strong starting point, they must be continually adapted and fine-tuned to reflect the specific environments in which they operate. Research into more sophisticated transfer learning techniques, which allow models to transfer knowledge from one domain to another with minimal retraining, could help to improve the robustness and adaptability of AHMS.

Additionally, developing more robust FL algorithms that can handle non-independent and identically distributed data distributions is essential. Aircraft in different regions may experience vastly different operating conditions, and a one-size-fits-all model may not perform well across all scenarios. Future research should explore personalized FL approaches that enable each aircraft to benefit from global knowledge while also retaining highly customized models that reflect their unique operational contexts.

Security and privacy are paramount in the aviation industry, where the consequences of data breaches or system failures can be catastrophic. While FL offers inherent privacy advantages by keeping data localized, it also introduces new vulnerabilities, such as the risk of adversarial attacks or model poisoning. Future research should focus on enhancing the security of FL frameworks through techniques such as differential privacy, secure multi-party computation, and homomorphic encryption.

As AI systems become more integrated into critical aviation infrastructure, ensuring their resilience against adversarial attacks is crucial. Research into adversarial robustness (developing models that can withstand and correctly respond to manipulated inputs) will be essential for maintaining the integrity and reliability of AHMS. This includes exploring adversarial training techniques and anomaly detection systems that can identify and mitigate potential threats in real time.

The future of AHMS research must also address the challenge of integrating AI-driven models with existing aviation systems and workflows. The aviation industry is heavily regulated, with stringent requirements for safety, reliability, and compliance. Any AI

system deployed in this context must not only meet these standards but also be compatible with the existing technological infrastructure.

Research into explainable AI is critical in this regard. As AI models become more complex, their decision-making processes often become opaque, making it difficult for human operators and regulators to understand how specific conclusions were reached. Developing AI systems that can provide clear, interpretable explanations for their decisions will be crucial for gaining trust and ensuring that these systems can be safely integrated into aviation operations. Research should focus on creating interoperable AI systems that can easily integrate with existing maintenance management systems, flight control systems, and other aviation software. This includes developing standardized interfaces and protocols that allow seamless communication between AI-driven AHMS and other critical systems in the aviation ecosystem.

As AI becomes more prevalent in aviation, ethical and regulatory considerations will increasingly come to the forefront. Future research must address the ethical implications of AI-driven decision-making in aviation, particularly in areas such as predictive maintenance, where the stakes are high. This includes exploring the potential biases that might be embedded in AI models and ensuring that these models are developed and deployed in ways that are fair, transparent, and accountable.

Additionally, as regulatory bodies begin to develop guidelines and standards for AI in aviation, research should support these efforts by providing evidence-based insights into the best practices for AI development, deployment, and monitoring. This includes collaborating with regulators to develop frameworks that ensure AI systems meet the highest safety and reliability standards without stifling innovation.

By addressing these challenges through focused research and development, the aviation industry can harness the full potential of AI to enhance safety, efficiency, and reliability in aircraft operations.

5. Conclusions

This paper has explored the integration of foundation models, federated learning, and AIoT technologies in aviation health monitoring systems, presenting a novel framework that promises to dramatically change aircraft safety, efficiency, and reliability. The proposed architecture uses the strengths of both centralized and decentralized approaches, offering a robust solution to the complex challenges faced in modern aviation.

The integration of FMs provides a strong baseline for generalized knowledge across the aviation ecosystem, capturing broad patterns and insights from diverse data sources. This is complemented by the FL approach, which enables personalized model adaptation while preserving data privacy and reducing communication overhead. The incorporation of AIoT technologies further enhances the system's capabilities, allowing for real-time data collection, edge computing, and seamless communication across the aviation network.

The comparative analysis of FMs, FL and the integrated approach has highlighted the unique advantages and limitations of each method. The integrated model emerges as a promising solution, combining the global consistency and comprehensive knowledge of FMs with the adaptability and privacy benefits of FL. However, this approach also presents challenges in terms of complexity, resource requirements, and the need for sophisticated orchestration between centralized and decentralized processes.

The proposed architecture for the aviation ecosystem, encompassing centralized processing, client, and communication domains, provides a roadmap for implementing these advanced technologies in real-world aviation environments. This multi-layered approach ensures that the health monitoring system can provide personalized, real-time insights for individual aircraft while using fleet-wide data for continuous improvement.

Key challenges identified in this study include ensuring scalability across large, diverse fleets of aircraft, maintaining model robustness in the face of heterogeneous data and operational conditions, safeguarding against security threats and privacy breaches, and addressing the ethical and regulatory implications of AI-driven decision-making in aviation.

Future research directions outlined in this paper point towards developing more efficient algorithms for FL, exploring lightweight versions of FMs for edge deployment, enhancing the security and privacy of FL frameworks, and improving the explainability and interpretability of AI models in aviation contexts. Additionally, the integration of these advanced AI systems with existing aviation infrastructure and workflows presents both a challenge and an opportunity for future innovation.

The collaborative intelligence framework proposed in this study represents a new step forward in aviation health monitoring. By combining the strengths of FMs, FL and AIoT technologies, this approach has the potential to dramatically enhance predictive maintenance capabilities, optimize operational efficiency, and improve overall safety in the aviation industry.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bommasani, R.; Hudson, D.A.; Adeli, E.; Altman, R.; Arora, S.; von Arx, S.; Bernstein, M.S.; Bohg, J.; Bosselut, A.; Brunskill, E.; et al. On the Opportunities and Risks of Foundation Models. 2021. Available online: https://arxiv.org/abs/2108.07258 (accessed on 5 September 2024).
2. Li, X.; Tian, Y.; Ye, P.; Duan, H.; Wang, F.Y. A novel scenarios engineering methodology for foundation models in metaverse. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *53*, 2148–2159. [CrossRef]
3. Reed, S.; Zolna, K.; Parisotto, E.; Colmenarejo, S.G.; Novikov, A.; Barth-Maron, G.; Gimenez, M.; Sulsky, Y.; Kay, J.; Springenberg, J.T.; et al. A Generalist Agent. 2022. Available online: https://arxiv.org/abs/2205.06175 (accessed on 5 September 2024).
4. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are fewshot learners. In Proceedings of the 2020 Advances in Neural Information Processing Systems, Online, 6–12 December 2020; Volume 33, pp. 1877–1901.
5. Bender, E.M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Toronto, ON, Canada, 3–10 March 2021; pp. 610–623.
6. Competition and Markets Authority (CMA). *AI Foundation Models: Technical Update Report*; CMA: London, UK, 2024. Available online: https://www.gov.uk/government/publications/ai-foundation-models-update-paper (accessed on 5 September 2024).
7. Sun, C.; Li, H.; Dui, H.; Hong, S.; Sun, Y.; Song, M.; Cai, D.; Zhang, B.; Wang, Q.; Wang, Y.; et al. A multi-model architecture based on deep learning for aircraft load prediction. *Commun Eng.* **2023**, *2*, 47. [CrossRef]
8. FGVC-Aircraft Benchmark. Available online: https://www.robots.ox.ac.uk/~vgg/data/fgvc-aircraft/ (accessed on 5 September 2024).
9. Blumenfeld, J. *NASA and IBM Openly Release Geospatial AI Foundation Model for NASA Earth Observation Data*; NASA Earthdata: Huntsville, AL, USA, 2023. Available online: https://www.earthdata.nasa.gov/news/impact-ibm-hls-foundation-model (accessed on 5 September 2024).
10. Ray, T.; Bhat, A.P.; White, R.T.; Nguyen, V.M.; Pinon Fischer, O.J.; Mavris, D.N. Examining the Potential of Generative Language Models for Aviation Safety Analysis: Case Study and Insights Using the Aviation Safety Reporting System (ASRS). *Aerospace* **2023**, *10*, 770. [CrossRef]
11. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, PMLR, Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
12. Ogundokun, R.O.; Misra, S.; Maskeliunas, R.; Damasevicius, R. A Review on Federated Learning and Machine Learning Approaches: Categorization, Application Areas, and Blockchain Technology. *Information* **2022**, *13*, 263. [CrossRef]
13. Li, W.; Yang, W.; Jin, G.; Chen, J.; Li, J.; Huang, R.; Chen, Z. Clustering Federated Learning for Bearing Fault Diagnosis in Aerospace Applications with a Self-Attention Mechanism. *Aerospace* **2022**, *9*, 516. [CrossRef]
14. Berghout, T.; Benbouzid, M.; Bentrcia, T.; Lim, W.H.; Amirat, Y. Federated Learning for Condition Monitoring of Industrial Processes: A Review on Fault Diagnosis Methods, Challenges, and Prospects. *Electronics* **2023**, *12*, 158. [CrossRef]
15. Khan, S.; Gaba, G.S.; Gurtov, A. *A Federated Learning Based Security for Controller-Pilot Data Link Communication*; Linköping University: Linköping, Sweden, 2022; Available online: https://www.icas.org/ICAS_ARCHIVE/ICAS2022/data/papers/ICAS2022_0704_paper.pdf (accessed on 5 September 2024).

16. Llasag Rosero, R.H.; Silva, C.; Ribeiro, B. Remaining Useful Life Estimation in Aircraft Components with Federated Learning. *PHM Soc. Eur. Conf.* **2020**, *5*, 9. [CrossRef]
17. Qu, Y.; Dai, H.; Zhuang, Y.; Chen, J.; Dong, C.; Wu, F.; Guo, S. Decentralized Federated Learning for UAV Networks: Architecture, Challenges, and Opportunities. *IEEE Netw.* **2021**, *35*, 156–162. [CrossRef]
18. Doğru, A.; Bouarfa, S.; Arizar, R.; Aydoğan, R. Using Convolutional Neural Networks to Automate Aircraft Maintenance Visual Inspection. *Aerospace* **2020**, *7*, 171. [CrossRef]
19. Abdelghany, E.S.; Farghaly, M.B.; Almalki, M.M.; Sarhan, H.H.; Essa, M.E.-S.M. Machine Learning and IoT Trends for Intelligent Prediction of Aircraft Wing Anti-Icing System Temperature. *Aerospace* **2023**, *10*, 676. [CrossRef]
20. Gao, Z.; Mavris, D.N. Statistics and Machine Learning in Aviation Environmental Impact Analysis: A Survey of Recent Progress. *Aerospace* **2022**, *9*, 750. [CrossRef]
21. Brandoli, B.; de Geus, A.R.; Souza, J.R.; Spadon, G.; Soares, A.; Rodrigues, J.F., Jr.; Komorowski, J.; Matwin, S. Aircraft Fuselage Corrosion Detection Using Artificial Intelligence. *Sensors* **2021**, *21*, 4026. [CrossRef] [PubMed]
22. Yang, R.; Gao, Y.; Wang, H.; Ni, X. Fuzzy Neural Network PID Control Used in Individual Blade Control. *Aerospace* **2023**, *10*, 623. [CrossRef]
23. Wang, Z.; Zhao, Y. Data-Driven Exhaust Gas Temperature Baseline Predictions for Aeroengine Based on Machine Learning Algorithms. *Aerospace* **2023**, *10*, 17. [CrossRef]
24. Chen, J.; Qi, G.; Wang, K. Synergizing Machine Learning and the Aviation Sector in Lithium-Ion Battery Applications: A Review. *Energies* **2023**, *16*, 6318. [CrossRef]
25. Baumann, M.; Koch, C.; Staudacher, S. Application of Neural Networks and Transfer Learning to Turbomachinery Heat Transfer. *Aerospace* **2022**, *9*, 49. [CrossRef]
26. Quadros, J.D.; Khan, S.A.; Aabid, A.; Alam, M.S.; Baig, M. Machine Learning Applications in Modelling and Analysis of Base Pressure in Suddenly Expanded Flows. *Aerospace* **2021**, *8*, 318. [CrossRef]
27. Papakonstantinou, C.; Daramouskas, I.; Lappas, V.; Moulianitis, V.C.; Kostopoulos, V. A Machine Learning Approach for Global Steering Control Moment Gyroscope Clusters. *Aerospace* **2022**, *9*, 164. [CrossRef]
28. Python. Available online: https://www.python.org/downloads/ (accessed on 7 October 2024).
29. TensorFlow. An End-to-End Platform for Machine Learning. Available online: https://www.tensorflow.org/ (accessed on 7 October 2024).
30. TensorFlow Federated: Machine Learning on Decentralized Data. Available online: https://www.tensorflow.org/federated (accessed on 7 October 2024).
31. NumPy. The Fundamental Package for Scientific Computing with Python. Available online: https://numpy.org/ (accessed on 7 October 2024).
32. Pandas. Available online: https://pandas.pydata.org/ (accessed on 7 October 2024).
33. Matplotlib: Visualization with Python. Available online: https://matplotlib.org/ (accessed on 7 October 2024).
34. Seaborn: Statistical Data Visualization. Available online: https://seaborn.pydata.org/ (accessed on 7 October 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.