Research Article

Hybrid Paris Physics-Informed Neural Network for Predicting CFM56-7B Engine Fatigue Failure

Yebeen Hwang

Hankuk Academy of Foreign Studies, 50, Oedae-ro 54beon-gil, Mohyeon-eup, Cheoin-gu, Yongin-si, Gyeonggi-do, Republic of Korea

ABSTRACT

Existing methodologies for predictive maintenance in aviation engines have largely diverged into data-centric and physics-centric models, each constrained by their unchangeable limitations. In this study, a hybrid framework integrating both perspectives was developed to address fatigue-induced failures in the CFM56-7B engine. Specifically, a Bayesian Physics-Informed Neural Network (B-PINN) was constructed, embedding Paris' Law within a deep learning structure and modeling key fatigue parameters as probabilistic distributions. Selected sensor data from NASA's CMAPSS FD004 dataset was employed to assume latent stress signals and simulate fatigue crack propagation. The results show that the proposed model has advantages on interpretability and reliability of fatigue predictions but also quantifies uncertainty through variational inference and Monte Carlo dropout.

Keywords: Physics-Informed Neural Network (PINN); Fatigue Crack Growth Prediction; Bayesian Uncertainty Quantification; CMAPSS FD004 Dataset

INTRODUCTION

The Boeing 737 is one of the most widely used aircraft models. According to Boeing Company's Boeing Commercial and Deliveries data through 12/31/2024 (1), a total of 19,579,737 models were purchased by commercial airlines. At the same time, Boeing 737 models have experienced many accidents (Figure 1). Although the Boeing 737-600/700/800/900 is one of the most widely used aircraft models globally, its hull loss accident rate (0.17%) is not exceptionally low compared to other frequently operated models. For example, the A320/321/319 NEO family has a lower accident rate of 0.08%, and the aircraft models such as the A380, 737 Max and EMB-170/175/190 series report even lower or zero accident rates.

One notable accident involving a Boeing 737-800 occurred in Muan, South Korea, on December 29, 2024. The accident resulted in the deaths of 179 passengers and left two others severely injured. Such aircraft accidents can be caused by multiple factors, such as runway excursions (RE), loss of control in flight (LOC-I), controlled flight into terrain (CFIT), abnormal runway contact (ARC), undershoot/overshoot (USOS), system/component failure or malfunction (SCF) (3). Among those flight dangers, SCF is the most fatal and frequently arising factor of aircraft accidents. The average hull loss accident rate of SCF was 0.05%, accounting for a 10-year moving average

Corresponding author: Yebeen Hwang, E-mail: 19370@hafs.hs.kr.

Received May 27, 2025**; Accepted** July 28, 2025
https://doi.org/10.70251/HYJR2348.34156168

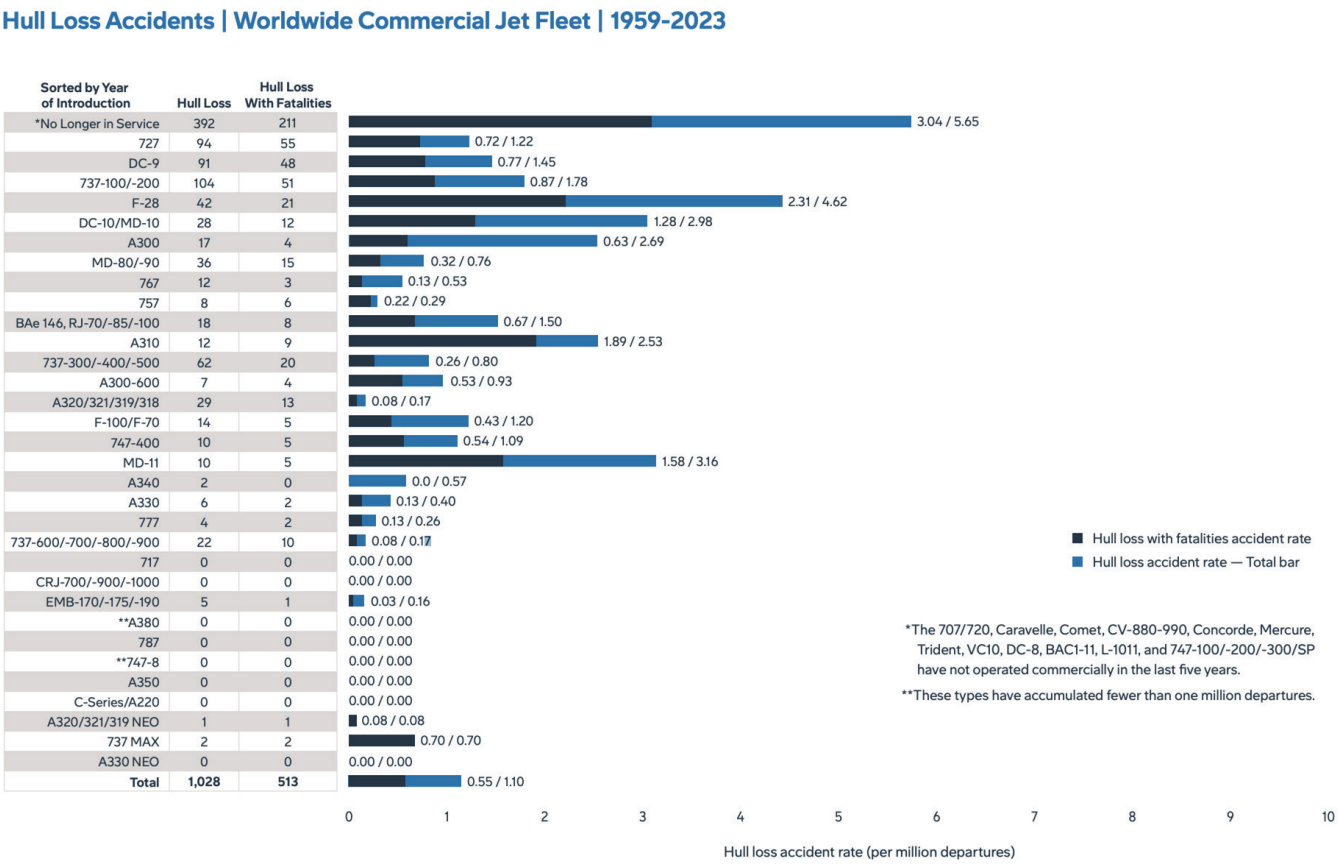

Figure 1. Accident Rates by Airplane Type.

(4). The Boeing aircraft models 737-600, 700, 800, and 900 all use CFM56-7B engines. CFM56 is a turbofan jet engine series made by CFM International. Among all of the CFM56 engine series, the CFM56-7B is the most advanced model and requires a low cost of operation.

As mentioned, one of the most fatal causes of airplane accidents is due to SCF. Many types of SCFs exist, including engine fatigue, thermal stress, foreign object debris/damage (FOD), electrical failures. This research primarily focuses on fatalities due to engine fatigue. This study provides Physics-Informed Machine Learning (PIML) to prevent aircraft accidents caused by engine fatigue of the jet engine CFM56-7B. The research investigates how the PIML model can detect early signs of fatigue, providing information for the early replacement of fatigued engines.

Problem Statement

Existing Remaining Useful Life (RUL) prediction methods for jet engines are usually divided into two types: purely data-driven models and purely physics-based models. Data-driven models are flexible, but they don't really explain what's happening inside the engine, and they also need a lot of failure data — which is pretty much impossible in aviation, where failures are supposed to be rare. On the other hand, physics-based models rely too much on complicated formulas and material constants, and getting those constants right without a sufficient number of experiments is not realistic. Hence, this study focuses on using a physics-informed machine learning (PIML) approach and data-driven prediction with physical laws like Paris' Law to make the model both accurate and interpretable.

Research Objectives and Contributions

The Physics-Informed Neural Network (PINN) developed in this study is structured around the classical Paris' Law formulation of fatigue crack growth, expressed as da/dN=C(ΔK)^m, where C and m are material-dependent parameters (Figure 2) (5). These parameters are regularized using empirical values drawn from titanium alloy fatigue datasets for Ti-6Al-4V, per ASTM E647 and

related studies (6, 7). Sensor features from the FD004 subset of the C-MAPSS dataset (7) are used as model inputs, with domain-guided selection identifying low-pressure turbine temperature (T24), pressure ratio (P30), vibration, and fan/core rotational speeds (Nf, Nc) as principal fatigue indicators. The hybrid PINN loss function incorporates both data fidelity by RNN loss function and physics-informed terms, minimizing deviation from observed crack growth trajectories while ensuring compliance with Paris' Law. Following (7), a bias correction component is added via an RNN-informed residual to accommodate deviations under complex loading (Figure 3). To model uncertainty in the parameters C and m, the framework is extended to a Bayesian PINN (B-PINN) using variational inference techniques as proposed by (9, 10), enabling credible intervals over predicted crack lengths. Input feature relevance is indirectly assessed via sensor selection informed by heat map correlation patterns and stress-prediction sensitivity analysis.

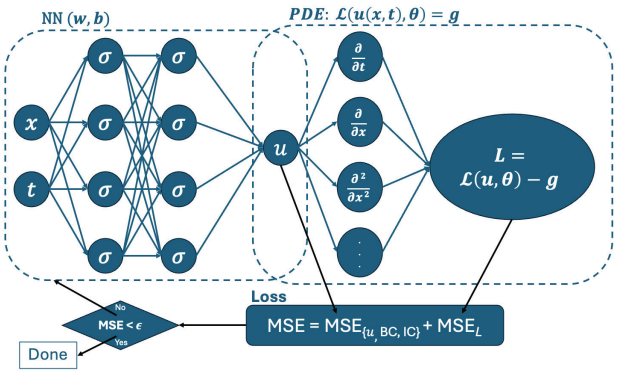

Figure 2. Schematic of a Physics-Informed Neural Network (PINN) (5).

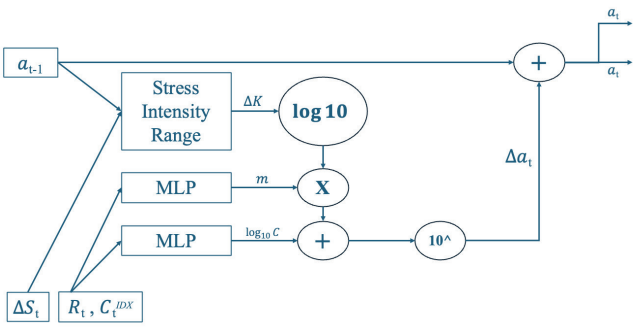

Figure 3. Proposed RNN cell for corrosion-fatigue crack propagation. MLP stands for multilayer perceptron, ΔS_t represents stress range, R_t is stress ratio, and C_t^{IDX} means corrosion environment index (7).

THEORETICAL BACKGROUND

Turbojet Engine and Engine Failure

How Jet Engines Work

A jet engine is an aircraft engine that provides propulsion by expelling a reaction mass. The jet engine makes the aircraft go forward and lift the aircraft into the air. To prevent a decrease in engine power due to the flight level, a large amount of air is sucked into the engine, compressed, and continuously combusted with fuel. Depending on the type of jet engine, a fan may be attached at the front, and the compression stages may vary. However, the basic principle remains the same: compressing a large amount of air and then controlling combustion to proceed (11). The most widely used jet engine model in commercial aircraft that people use is the turbofan engine. The turbofan engine is a type of turbo engine with a fan attached, developed with the goal of improving fuel efficiency and reducing noise. The thrust of a turbofan is the combined thrust created by the fan and the thrust generated by the energy of the gases combusted inside the engine.

CFM56-7B Engine

CFM56-7B engine The CFM56-7B engine is one of the most widely used turbofan jet engines developed by CFM International. The engine powers several high-usage aircraft models, notably the Boeing 737 Next Generation and 737-600/-700/-800/-900 series, making it one of the most produced and extensively operated jet engines in commercial aviation. The CFM56-7B engine has six variants, which differ slightly in specifications by each model. All of the variants have the same basic hardware, and the thrust rating is only changed through the programming of the engine's FADEC (Full Authority Digital Engine Control) unit. Its thrust rate differs between 19,500lbs and 27,300lbs. The engine uses a two-shaft design with a 61-inch fan, followed by a three-stage LPC (Low Pressure Compressor), a nine-stage HPC (High Pressure Compressor), a one-stage HPT (High Pressure Turbine), and a four-stage LPT (Low Pressure Turbine).

The engine is flat-rated, considering that flat rating occurs due to the pressure limit, and every jet engine has limitations on the maximum thrust available, up to ISA + 15°C (30°C) for most models, but CFM56-7B engines extend this limit to ISA + 50°C (for high-altitude operations). ISA represents the International Standard Atmosphere, whose standard is 15°C, 1013.25 hPa, and 1.225 kg/m³.

In terms of operating altitude, typical cruise conditions for aircraft equipped with the CFM56-7B occur between 35,000 and 39,000 ft, corresponding to atmospheric pressures below 30 kPa. These flight settings show low ambient pressures, high Mach numbers (typically Mach 0.78–0.82), and maximum thermal stress conditions on the turbine blades. The EGT redline is 950°C, while actual EGT (Exhaust Gas Temperature: the temperature of the turbine exhausting gases when leaving the turbine unit) values range between 857°C and 940°C, depending on the model. EGT margins are higher in lower-thrust models and decrease as thrust increases (12). There was a technical update in 2007, called the TI version, which improved engine durability and reduced maintenance costs by 12%. By this, the CFM56-7B engine model's fuel consumption lowered by 1%, and NOx emissions reduced by 20-30%. This update also included new 3D-Aero compressor blades, an optimized single-annular combustor, and redesigned turbine blades and low-pressure nozzles. These improvements reduced interaction losses between the high-pressure and low-pressure turbines, changing LLP (Life Limited Parts) from the original engine components (12).

The fan module has a life limit of 30,000 EFC (Engine Flight Cycles), where common jet engines' EFC is 20,000. The fan module includes the fan disc, booster spool, and fan shaft. HPC (High-Pressure Compressor) is set at 20,000 EFC, covering the front shaft, components from stages 1-9, and the CDP (Compressor Discharge Pressure) seal. HPT (High-Pressure Turbine) also has a life limit of 20,000 EFC, which includes the front shaft, air seal, disc, and rear shaft. LPT (Low-Pressure Turbine) has a life limit of 25,000 EFC, covering the shaft, conical support, and stages 1-4 discs (12).

Crack Growth and Paris' Law

Paris Law and Fatigue Studies

Standard formula of Paris' law is defined as

$$\frac{da}{dN} = C \cdot (\Delta K)^m \qquad \text{Eq. [1]}$$

where a is crack length, N is load cycles, ΔK is stress intensity factor range, and C is the crack growth coefficient; m is the Paris exponent. Usually in metallic blade condition, constant C and m vary as shown below (5).

$$C \in [10^{-12}, 10^{-8}],\ m \in [2.5,\ 4] \qquad \text{Eq. [2]}$$

Paris' law is applied for measuring fatigue crack growth rates under constant amplitude loading. It shows that crack length per load cycle is proportional to the stress intensity factor range. This formulation works as the fundamental concept in this paper's Physics-Informed Neural Network (PINN) model.

Not only Paris' law, but there are also many equations explaining crack growth of the fatigue material, such as Miner's rule. Miner's rule assumes linear accumulation of fatigue damage, but it fails to accurately represent real-world damage under variable amplitude loading. In contrast, Paris' law models incremental crack growth per cycle, which makes the model better capture the nonlinear behaviors. According to (13), nonlinear summation methods estimate failure up to 23% earlier than Miner's prediction.

$$\Delta K = Y \cdot \sigma \cdot \sqrt{\pi a} \qquad \text{Eq. [3]}$$

In this equation, ΔK is a stress intensity factor range, and a is the crack length. σ is the effective stress applied to the material, and C and m are material constants that are specific to the engine material (e.g., Ti-6Al-4V used in turbine blades). Y is a geometry factor. According to (14), the model treats mechanical stress σ not as a direct measurement but as a latent variable inferred from sensor inputs, especially under operating conditions like altitude, shaft speed, and total temperature. This study derives the approach of (14) applying the formula of the stress intensity factor range.

Fatigue Fracture Mechanics In Turbo Jet Engine

A 2016 National Transportation Safety Board (NTSB) report details a fatigue crack in the fan blade. The report was about an accident of a Boeing 737 aircraft with two CFM56-7B engines equipped; the engine experienced engine fan blade separation, loss of inlet cowling, and cabin depressurization while climbing through flight level 310.

M. Fox (15) found that the engine failure was initiated from a subsurface crack and propagated via high-cycle fatigue (HCF) under operating conditions where engine speed was more than 10,000 rpm. The fan blades became more susceptible to fatigue crack initiation under the combined effects of loading/friction changes over time and engine operational differences. Not only the primary causes of the crack, the report points out the influence of the abnormal residual stress affecting the unusual fatigue of the metal blade of the jet engine. Moreover, the report indicates a possible correlation between stress anomalies and fatigue crack initiation accidents, highlighting the need for improved predictive maintenance strategies.

Machine Learning Model

Predictive Maintenance

Predictive maintenance (PdM) is a strategy that predicts when a machine might fail so that maintenance can be done before an actual breakdown occurs. Instead of fixing something after it fails (reactive maintenance) or checking it on a fixed schedule (preventive maintenance), PdM continuously tracks the condition of a system in real-time. By analyzing sensor data and operational history, PdM can reduce unexpected failures, cut maintenance costs, and improve overall safety.

Before machine learning was widely used, PdM of jet engines relied on conventional condition monitoring techniques.

- Vibration Analysis: Detects early signs of mechanical problems like imbalance, misalignment, and bearing wear (16).
- Oil and Debris Analysis: Checks engine oil for metal particles to assess component wear and tear (17).
- Temperature Monitoring: Tracks heat buildup from friction, poor lubrication, or inefficient cooling (18).
- Non-Destructive Testing (NDT): Techniques like ultrasonic testing (UT) and eddy current inspection (ECI) to find surface and internal cracks (19).

While these techniques work, they require frequent manual inspections and have limitations in detecting complex failure modes. With machine learning models, PdM has become much more accurate and less dependent on human inspections.

Physics-Informed Neural Network (PINN)

PINN is deep learning models that incorporate physics-based loss terms (e.g., PDEs or ODEs) in the training objective. It takes predictive maintenance a step further by combining physics-based models with machine learning. Unlike traditional machine learning models that rely only on historical data, PIML incorporates physical principles and engineering fields like fatigue models, thermodynamics, and material stress analysis to make better predictions (20). Using the PIML model, machine learning models' interpretability, robustness, and generalization with limited data increase outstandingly. Physics-Informed Machine Learning (PIML) For the CFM56-7B engine, PIML for the CFM56-7B engine can detect fatigue cracks. By using a combination of finite element analysis (FEA) and machine learning, a machine learning model can identify early fatigue cracks. Also, by using computational fluid dynamics (CFD) along with neural networks, it can analyze how temperature changes impact turbine blade durability due to thermal stress of metal blades.

In fatigue modeling under corrosive environments, a hybrid PINN structure can be implemented with a recurrent neural network (RNN) learning residual dynamics (e.g., oxidation effects) (6).

The physics loss from Paris' Law and data mismatch were balanced using dynamic weighting:

$$L_{total} = \alpha \cdot L_{physics} + (1 - \alpha) \cdot L_{RNN} \quad \text{Eq. [4]}$$

Where L_{total} is the overall loss function used to train the model; $L_{physics}$ is a loss function that shows how well the model follows Paris' Law; and L_{RNN} is a loss function that represents how the model matches with the actual sensor data. α is a weighting coefficient that controls the weighted value between the physics-based loss function and data-driven loss function. By weighting between two loss functions, the hybrid model encourages adopting bias correction modules to capture unmodeled fatigue phenomena (like thermal fatigue or corrosion).

According to a study by A. Yucesan (9), the hybrid PINN predicts fatigue with higher accuracy by applying physics. The study by Yucesan showed a 25% RMSE reduction compared to black-box DNNs.

Limitations of Prior Fatigue Models and Novelty of the Proposed Framework

Previous research in fatigue modeling has made notable progress by embedding Paris' Law into neural network architectures to enhance the physical interpretability of data-driven models. For instance, neural fatigue cohesive (NFCOH) models have been augmented with Paris-law-derived regularization to predict crack propagation in composite laminates, significantly reducing training sample requirements (24). Similarly, particle filtering frameworks have been applied to multiple-crack scenarios in metallic specimens, coupling modified Paris models with Lamb wave-informed deep autoencoders for improved state estimation (25). These approaches demonstrate the potential of physics-informed neural models in reducing data dependency and increasing reliability. However, they have been primarily validated in controlled laboratory environments with explicit measurements of stress or crack length, not under the complex, sensor-limited conditions of operational jet engines.

In contrast, diagnostic frameworks applied to the CFM56-7B engine—such as Gas Path Analysis (GPA)

and GM(1,1)-based statistical degradation models—rely heavily on observable parameters like exhaust gas temperature (EGT), shaft speeds, or pressure ratios (26, 27). While these methods can track performance degradation trends, they do not infer unmeasurable internal stress states, which are central to fatigue damage initiation and propagation. Moreover, the practical environment of in-flight engines presents additional challenges: fewer available sensor channels, high ambient noise, varying operating conditions, and an absence of direct stress measurements. As a result, current maintenance diagnostics remain limited in their ability to detect early-stage fatigue risk driven by latent mechanical variables. This study addresses this critical gap by introducing a hybrid Bayesian Physics-Informed Neural Network designed to estimate latent internal stress directly from sensor data and simulate fatigue crack growth via embedded Paris Law dynamics. Unlike traditional methods that treat stress as a known or externally measured quantity, this framework models stress as a hidden variable, inferred from selected features in the C-MAPSS dataset filtered to resemble CFM56-7B operating conditions (28). This inference is then integrated into a physics-based crack growth calculation using ΔK and Paris' Law, while Bayesian uncertainty modeling accounts for sensor noise and parameter ambiguity. By jointly learning both physical and statistical relationships, the proposed method offers a novel approach to fatigue life prediction—uniquely suited for real-world aviation scenarios where stress cannot be directly measured yet is the primary driver of structural failure.

Challenges in Parameter Estimation of C, m

Constants C and m in Eq. [1] are easily affected by the material microstructure, heat treatment, and surface design. Even within the same alloy (e.g., Ti-6AI-4V), the value can fluctuate a lot. For an accurate estimation of the constant, stress intensity factors must be calculated using Digital Image Correlation (DIC) and the Extended Finite Element Method (XFEM). However, due to the limitations of these methodologies, direct estimation of C and m using techniques like DIC or XFEM has challenges in practical fatigue analysis. Specifically, accurate estimation of the stress intensity factor ΔK which is defined by:

$$\Delta K = Y \cdot \sigma \cdot \sqrt{\pi a} \qquad \text{Eq. [3]}$$

where Y is the geometry correction factor, σ is the applied stress, and a is the crack length. However, in various influences such as turbine blades, Y may vary spatially and cannot be expressed analytically, necessitating FEM-based estimation. Even when ΔK is known, fitting the Paris' Law Eq. (1) requires accurate measurement of both da/dN and ΔK, which are often corrupted by noise in the dataset. The equation below shows when we take logarithms to linearize the relationship, based on the variables given in Eq. [1].

$$log\left(\frac{da}{dN}\right) = log\, C + m\, log(\Delta K) \qquad \text{Eq. [5]}$$

Even small errors in ΔK can cause large deviations in the slope and intercept $log\ C$, making least-squares fitting highly unstable.

Moreover, under variable amplitude loading and non-isothermal conditions, the assumption of constant C and m becomes invalid, and fatigue crack growth becomes path-dependent. This renders deterministic estimation methods insufficient.

To overcome these challenges, a Bayesian Physics-Informed Neural Network (B-PINN) is used as a framework (8). The parameters $\theta = (C, m)$ are treated as latent variables with a posterior distribution:

$$P(\theta \mid D) \propto P(D \mid \theta) \cdot P(\theta) \qquad \text{Eq. [6]}$$

with the likelihood term computed from sensor-observed crack lengths $a^{(i)}$ as:

$$P(D \mid \theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(a_{model}{}^{(i)}(\theta) - a_{observed}{}^{(i)})^2}{2\sigma^2}\right) \qquad \text{Eq. [7]}$$

Sampling from this posterior is calculated by Hamiltonian Monte Carlo (HMC), which defines a Hamiltonian as:

$$H(\theta \mid r) = -\, log P(D \mid \theta) - log P(\theta) + \frac{1}{2} r^T M^{-1} r \qquad \text{Eq. [8]}$$

and this equation is extended by:

$$\frac{d\theta}{dt} = M^{-1} r,\ \frac{dr}{dt} = -\nabla_\theta U(\theta) \qquad \text{Eq. [9]}$$

Alternatively, Variational Inference approximates the posterior by a parametric density :

$$Q(\theta;\zeta) = \prod_{i=1}^{2} q(\theta_i;\ \zeta_{u,i},\ \zeta_{\rho,i}) \qquad \text{Eq. [10]}$$

and optimizes the parameters by minimizing the equation:

$$D_{KL}(Q(\theta) \parallel P(\theta \mid D)) \approx E_{\theta \sim Q}[log\, Q(\theta) - log\, P(\theta) - log\, P(D \mid \theta)] \qquad \text{Eq. [11]}$$

Finally, the overall training objective incorporates Paris' Law via a soft penalty:

$$L_{total} = \alpha \cdot L_{physics} + (1-\alpha) \cdot L_{RNN} \quad \text{Eq. [4]}$$

where

$$L_{physics} = \sum_{i=1}^{N} \left(\frac{da}{dN}^{(i)} - C(\Delta K^{(i)})^m \right)^2 \quad \text{Eq. [12]}$$

enforces the tendency of applying the fatigue law, and accounts for observational constancy.

METHODS AND MATERIALS

Model Architecture

Dataset Selection and Preprocessing

In this study, the FD004 subset of the NASA CMAPSS dataset was selected for training and evaluation of the proposed predictive model. The CMAPSS (Commercial Modular Aero-Propulsion System Simulation) dataset, developed by NASA, is widely used for research in aircraft engine performance prediction and health monitoring. It contains high-fidelity simulated sensor data generated under realistic flight and degradation conditions, rather than real-world flight data.

The FD004 subdataset best approximates the long-cycle, full-condition degradation of turbofan engines like the CFM56-7B, with high variability in operating conditions and complete sensor readings relevant to fatigue dynamics. The CMAPSS dataset contains 21 sensor measurements per cycle, along with three operational settings, generated using NASA's high-fidelity turbofan engine simulator. These sensors include key parameters such as pressure, temperature, and shaft rotational speeds (RPM), making the dataset highly suitable for fatigue modeling and failure prediction research.

Initially, the goal of this research was to focus specifically on CFM56-7B engine degradation, given its widespread use in commercial aviation and its importance in understanding fatigue-related failures. However, there was a limitation of datasets due to the lack of publicly available failures of certain engine types. According to the previous study by Saxena et al (7), the CMAPSS dataset is a widely recognized resource for simulating jet engine degradation and fatigue crack propagation.

While the dataset is not engine-specific, to further align the training input space with the operational characteristics of the CFM56-7B, we performed data preprocessing.

To preprocess the raw CMAPSS dataset with the actual flight conditions of the CFM56-7B mentioned in Section 2.1.2, we applied an additional sorting step before normalization and modeling. In NASA's CMAPSS dataset, the three operating settings (op_1, op_2, op_3) correspond respectively to altitude (per thousands of feet), Mach number (Ma), and ambient pressure (kPa). The following criteria were made to identify, among the highly diverse and extensive CMAPSS engine datasets, those operational segments most closely aligned with the characteristics of the CFM56-7B engine. Further details are provided in Section 2.1.2.

- op_1 > 35 selects engine states at altitudes above 35,000 ft, which corresponds to the typical cruise ceiling of the 737NG equipped with CFM56-7B engines. At this altitude range, the engine is flat-rated to maintain maximum thrust even under ISA +50°C conditions.
- op_2 > 0.75 captures high-speed phases where the Mach number exceeds 0.75. This aligns with the Mach 0.78–0.84 cruise regime of narrow-body commercial aircraft, ensuring that aerodynamic loads are consistent with real-world use.
- op_3 < 100 enforces ambient pressure levels below 100 kPa, reflecting the low-pressure environment experienced at cruising altitudes (typically <30 kPa at 35,000 ft), which amplifies thermal stress on turbine components.

This set of three criteria narrows the dataset down to a smaller and more consistent group of engine cycles. After applying the filters, about 37.12% of the original FD004 data remains. This filtered subset more accurately reflects the high-stress operating conditions where fatigue cracks are likely to form in CFM56-7B engine components. For PIML, 12 sensor channels were selected based on their relevance to modeling engine stress and fatigue according to a previous study (7). These selected 12 sensors—sensors 2, 3, 4, 7, 11, 12, 13, 14, 15, 17, and 20—are critical values to determine fatigue cracks that are likely to be formed. To ensure uniformity and eliminate scale differences between sensors, the sensor data was normalized using MinMax scaling. This scaling method maps all sensor values between 0 and 1. Table 1 is sourced from reference (21).

Among the listed sensors, sensor2, 4, 7, 12, 13, and 14 were used as final model inputs based on correlation analysis (Fig. 4). According to Fig. 4, sensors 2, 4, 7, 12, 13, and 14 were chosen based on the correlation heatmap

analysis, which revealed both high intra-group consistency and complementary diversity. Sensors 2, 7, and 12 show near-perfect correlations (≥ 0.99), capturing critical operational trends that remain consistent across engine cycles. Sensor 4 demonstrates moderate correlation with other variables, offering additional dynamic information. Meanwhile, sensors 13 and 14 exhibit strong mutual correlation (≈ 0.79) and contribute distinct non-linear and transient characteristics, making the combination of these sensors suitable and comprehensive for modeling fatigue-related behavior.

Table 1. Sensor Name and Number used in NASA CMAPSS FD004 dataset

sensor #	sensor name
2	Total pressure at fan inlet (Ps30)
3	Total pressure in bypass duct (Ps50)
4	HPC outlet pressure (P30)
7	High-pressure turbine outlet temperature (T50)
11	Static pressure at HPC outlet (P40)
12	Ratio of fuel flow to Ps30
13	Corrected fan speed (Nc)
14	Corrected core speed (Nc)
15	Bypass Ratio (BPR)
17	Bleed Enthalpy
20	HPT coolant bleed
21	LPT coolant bleed

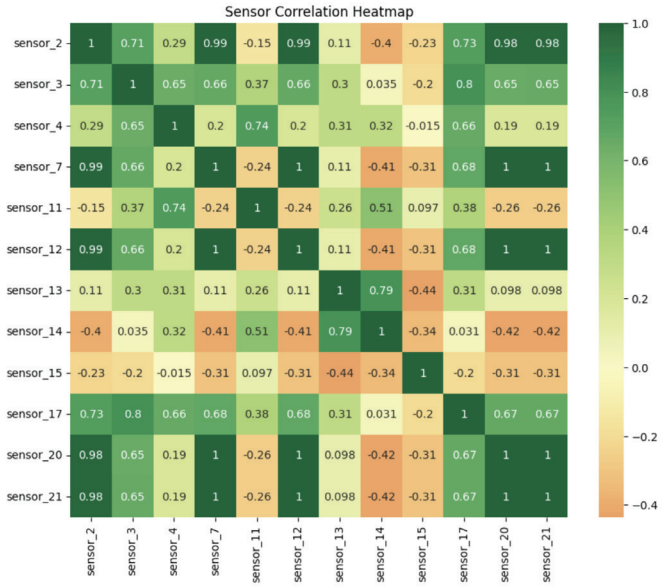

Figure 4. Sensor Correlation Heatmap.

Crack Growth Modeling

As mentioned above, this study presents a Hybrid PINN designed to model fatigue crack growth in jet engine components using sensor data. The framework integrates physical laws—specifically Paris' Law—with neural representations to infer latent stress and crack growth rates (Figure 5). Since the direct stress measurements are unavailable in the C-MAPSS dataset, the model estimates stress as a latent variable from six selected sensors. The estimated stress is then used to compute the stress intensity factor ΔK using Eq. [4], where crack length evolves over time. The hybrid model couples this physics-informed estimate with a residual neural network to learn non-ideal effects not captured by theory. To ensure the model adheres to both data-driven patterns and physical laws, a dual-component loss function is implemented. The prediction loss minimizes the mean squared error between predicted and theoretical crack growth values, while the physics loss enforces consistency with Paris' Law by comparing predicted da/dN against physically

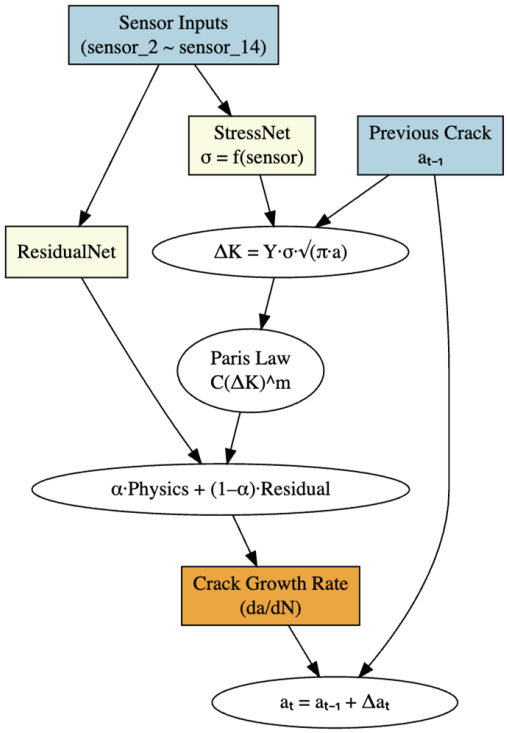

Figure 5. Architecture of the crack growth modeling using Hybrid-PINN.

expected rates. The model is trained on cycle-level data across all engine units, with sensor inputs scaled and batch-fed using a custom PyTorch dataset. The resulting framework offers a generalizable approach to estimate crack growth rates under realistic operating conditions, enabling both interpretability through physical grounding and flexibility through learned residuals.

Uncertainty Modeling (Bayesian PINN)

To calculate the epistemic uncertainty in crack growth prediction, a Bayesian PINN (B-PINN) model was employed by modeling the Paris' Law parameters C and m as probabilistic distributions (Figure 6). Utilizing Monte Carlo dropout during inference, the model makes multiple forward passes to approximate the posterior distribution of predicted crack growth rates. This enables the calculation of both the mean and standard deviation of uncertainty of the prediction at each cycle. This approach allows the PINN model to convey uncertainty-aware predictions for fatigue behavior under data-limited and noisy sensor conditions.

Baselines for Comparison Using RNN and Pure Paris' Law

Referring to the performance of the proposed hybrid PINN model (8), a recurrent neural network (RNN) and a pure Paris' Law model with fixed literature-based parameters $C = 3 \times 10^{-12}$ and $m = 3.1$ were implemented. The pure Paris baseline serves as a physics-only reference, applying uniform crack growth prediction across all engine units without learning from sensor data. This comparison highlights the hybrid PINN model of data-driven learning and adaptive stress estimation in capturing unit-specific degradation patterns.

Dataset and Experimental Setup

The experimental setup uses the FD004 subset of the C-MAPSS dataset. Six key sensors—sensor_2, sensor_4, sensor_7, sensor_12, sensor_13, and sensor_14—were selected based on physical interpretability and statistical significance from the correlation heatmap shown in Fig. 4. These sensors were chosen due to their mutual correlation and relevance to fatigue-related behavior, with sensor_2 (fan speed) and sensor_12 (static pressure) being particularly aligned with fatigue propagation indicators reported in prior studies (20).

Sensor data was scaled using MinMax normalization, and training was performed over 10 epochs using a Hybrid Physics-Informed Neural Network (Hybrid PINN). The model consists of three components shown below.

- a StressNet sub-network that maps sensor input to latent stress values σ,
- a physics-based Paris' Law branch that computes fatigue crack growth using inferred stress values, and
- a residual data-driven branch that learns nonlinearities not captured by the physics model.

The Paris' Law branch estimates the crack growth rate mentioned in Eq. [1] where the constants used were $C = 1 \times 10^{-12}$, $m = 3.2$, and geometry correction factor $Y = 1.1$, embedded as domain knowledge (22, 23). An α-value of 0.5 was used to equally weight the physics-based and data-driven branches during training. For uncertainty modeling, Monte Carlo Dropout was applied at inference time. Specifically, 30 stochastic forward passes were performed per prediction cycle, allowing estimation of epistemic uncertainty through the standard deviation of predicted crack growth rates. This approach enables cycle-wise confidence intervals under data-limited and noisy sensor conditions. The full Python code used for model training, testing, and visualization is provided in the *Appendix*.

Metrics of the PINN Model

To evaluate the interpretability of the Hybrid PINN model, we follow the analysis framework suggested by Raissi et al. [2017] (20). Their study showed that sensor_2 (analogous to Nf, or fan speed) and sensor_12 (related to Ps30, or static pressure) contributed significantly to crack growth prediction—both of which are among our selected

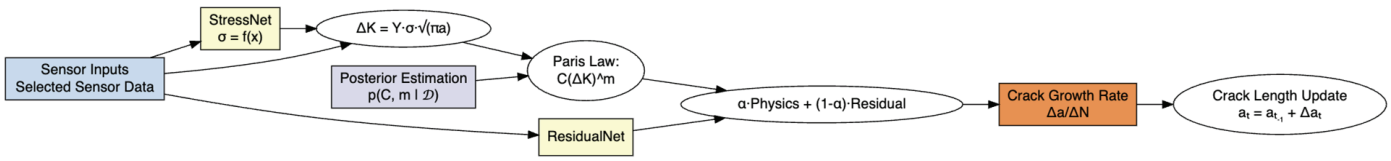

Figure 6. Architecture of the Bayesian Physics-Informed Neural Network (B-PINN) for Fatigue Crack Growth Modeling.

features. In our model, the StressNet sub-network plays a key role in this process. It takes multivariate sensor readings as input and estimates a latent internal stress value \sigma. This stress is then used to compute the stress intensity factor $\Delta K = Y \cdot \sigma \cdot \sqrt{\pi a}$, which is fed into Paris' Law to calculate crack growth rate. The residual_ net branch, in parallel, learns any remaining data-driven discrepancies. Both sub-networks are implemented using PyTorch's nn.Sequential modules.

RESULTS

Model Performance and Training Stability

At Epoch 0, average loss is 537.81, which is a high value. This loss value sharply drops to 0.0843 by epoch 9, indicating that the model is well-adjusting the iterations. From epoch 0 to 2, loss decreases drastically from 537.81 to 7.5516 to 1.7319. Average loss At epoch 5, average loss is 0.0413, and at epoch 9, average loss reaches 0.7795. The training loss decreased rapidly from 537.81 to 0.084 over 10 epochs, indicating fast and stable convergence of the Hybrid PINN model. Overall, we can assure that the model can learn the underlying physics-based loss function, but some instability remains in later stages of training.

Crack Growth Rate Trajectory Analysis

To interpret the result of the graph intuitively, the top 3 units with the lowest average uncertainty were selected—Unit 1, Unit 2, and Unit 3—and visualized in detail (see Figs. 7–9). Each figure includes three plots: predicted crack growth rate, Bayesian da/dN with uncertainty, and inferred stress trajectory with latent uncertainty. The predicted crack growth rate graph's x-axis represents engine cycles, and the y-axis (blue) shows the predicted crack growth rate da/dN in mm/cycle. Bayesian PINN Prediction + Uncertainty Graph's x-axis indicates engine cycles, the left y-axis (blue) shows the mean predicted crack growth rate da/dN in mm/cycle, and the right y-axis (red) presents the epistemic uncertainty as the standard deviation. Stress Prediction and Uncertainty Graph's x- axis represents engine cycles, the left y-axis (blue) displays the inferred internal stress in MPa estimated by the stress_net, and the right y-axis (red) shows the standard deviation of predicted stress as uncertainty.

For Unit 1, the predicted crack growth rate shows increasing trends beyond 100 cycles, with fluctuations intensifying toward the end. The corresponding Bayesian uncertainty remains low and stable throughout, reinforcing the model's confidence. Stress prediction for Unit 1 exhibits moderate oscillations, and uncertainty gradually stabilizes across cycles, indicating improved internal stress estimation.

Unit 2 shows a more variable crack growth profile with alternating peaks and troughs, especially after 60 cycles. However, the epistemic uncertainty remains consistently low, validating the model's robustness even under fluctuating inputs. Stress estimation in Unit 2 is more volatile, but the uncertainty envelope remains within acceptable bounds, suggesting successful learning of latent stress patterns.

Unit 3 exhibits a crack growth pattern similar to Unit 1, with a strong upward trend after 80 cycles. The Bayesian uncertainty again remains minimal, highlighting the predictive confidence of the model in this unit. The latent stress plot shows consistent high-frequency variations, but the uncertainty is controlled, suggesting that the model remains stable despite sensor-driven disturbances.

Overall, the visualization of these top-performing

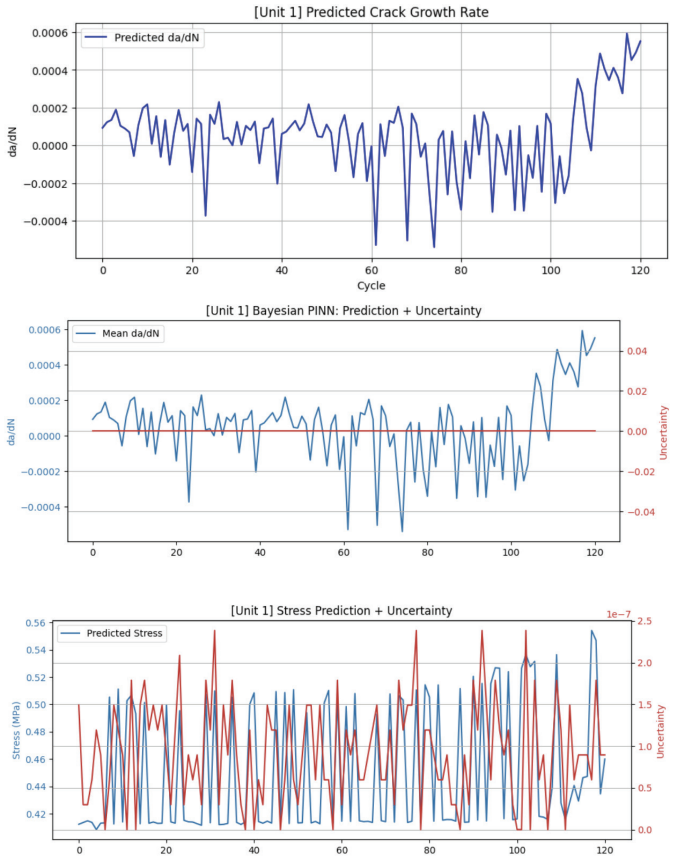

Figure 7. Plots for Unit 1 (predicted crack growth rate, Bayesian da/dN with uncertainty, and inferred stress trajectory with latent uncertainty).

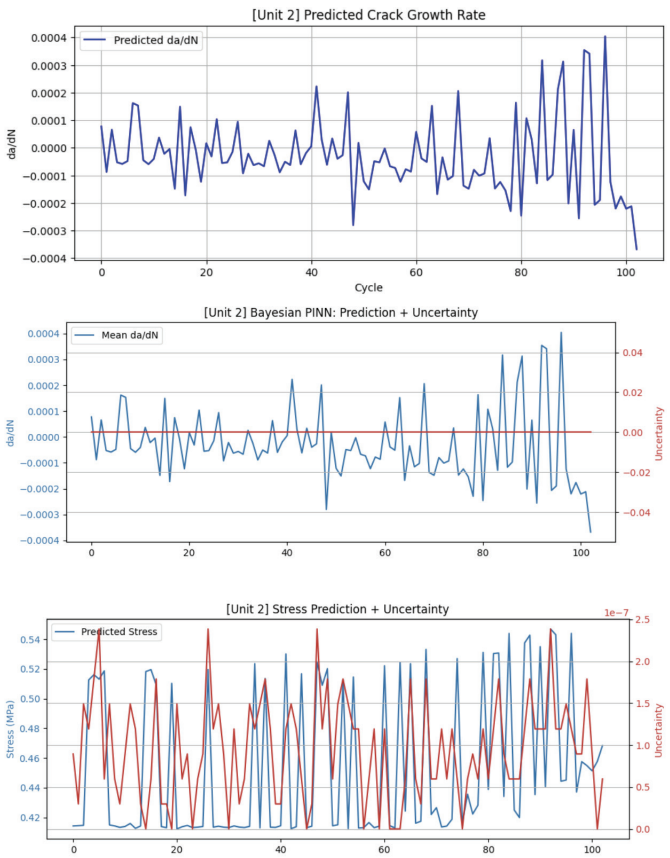

Figure 8. Plots for Unit 2 (predicted crack growth rate, Bayesian da/dN with uncertainty, and inferred stress trajectory with latent uncertainty).

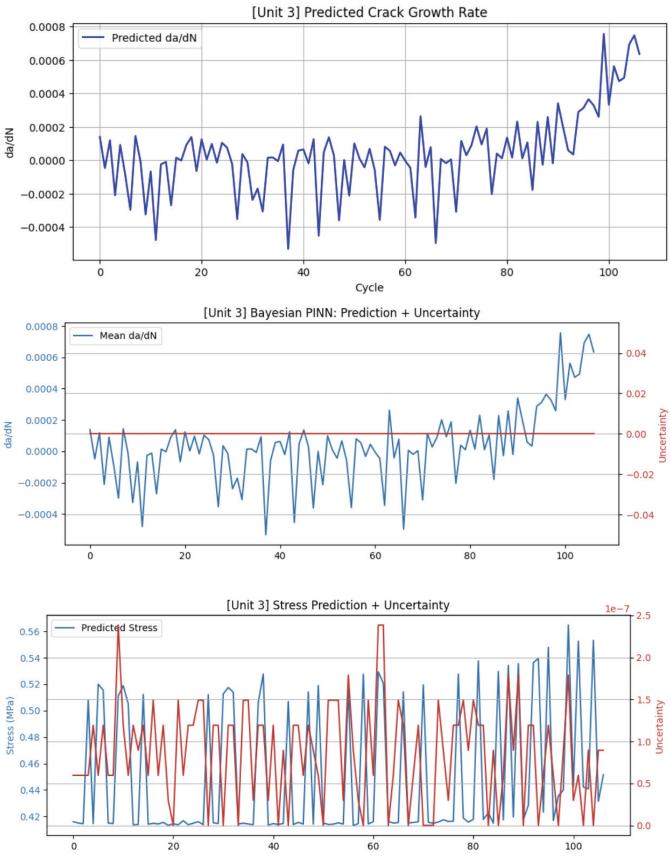

Figure 9. Plots for Unit 3 (predicted crack growth rate, Bayesian da/dN with uncertainty, and inferred stress trajectory with latent uncertainty).

units confirms the B-PINN's ability to not only provide accurate crack growth predictions but also offer interpretable uncertainty quantification. This dual capability is essential for real-world deployment in fatigue diagnostics of CFM56-7B engines, especially under data-limited and sensor-noisy conditions.

DISCUSSION

The alpha value 0.5 reflects an equal weighting between the physics-based Paris' Law loss and the data-driven residual loss. While this allows the model to benefit from both physical regularization and empirical learning, it may not be an optimal value for the perfect machine learning model. In the future using a dynamic α schedule or Bayesian optimization to tune α during training, especially since overfitting to noisy sensor data or underfitting physical laws may occur at different stages of training, will be beneficial for finding the exact value of the alpha.

While Monte Carlo dropout provides epistemic uncertainty, certain cycle regions showed high predictive uncertainty, as shown in Fig. 9, potentially due to underrepresented stress regimes in training. Applying Bayesian ensembling or using HMC sampling might result in better posterior approximation.

Another significant limitation lies in the latent stress inference mechanism. Since true stress measurements are not available in the CMAPSS dataset, the model relies solely on sensor proxies to estimate internal stresses. This assumption introduces a modeling gap—especially under complex, multiaxial loading conditions or dynamic engine phases (e.g., takeoff, descent)—where sensor-stress correlation may be nonlinear or weak. Without direct stress calibration from test-rig data or finite element validation, the inferred stress remains a best-guess approximation, which can propagate uncertainty into the fatigue crack predictions.

A threshold value of 0.00025 inch/cycle is applied in this study to identify the starting point of high-risk fatigue crack growth behavior. This value is referenced from the standard shown in ASTM E647 (5). In practical applications, this threshold serves as a critical limit beyond which crack propagation is considered potentially unstable or dangerous. Although the threshold provides a useful benchmark for risk classification, it varies a lot depending on specific operating conditions, material microstructures, or environmental influences such as temperature and load spectrum. Therefore, in future studies, specific calibration or experimental validation might be required for specific engine types or service scenarios. While the proposed framework demonstrates strong predictive capabilities on simulated datasets, its applicability to real-world scenarios remains unverified due to the absence of actual failure data from CFM56-7B engines. The lack of flight records and post-maintenance inspection reports limits our ability to validate the model against observed crack propagation or component failure. This constraint affects the generalizability of the results and should be considered when interpreting the findings. Future work should incorporate operational data from real engines to fully evaluate the model's robustness and diagnostic utility. Ultimately, these refinements would enhance the model's applicability in real-world Structural Health Monitoring (SHM) systems and contribute to safer, more reliable predictive maintenance frameworks in commercial aviation.

CONCLUSION

This study presents hybrid Physics-Informed Neural Network (PINN) and Bayesian PINN (B-PINN) modeling for predicting fatigue-induced crack growth in the CFM56-7B jet engine and uncertainty of the predicted crack growth. By applying Paris' Law into the neural network loss function and integrating sensor-derived stress estimates as residual neural networks (RNN), the model successfully integrates physical theory with data-driven learning. The results show that the model can accurately track crack growth trajectories, estimate latent stress behavior, and provide reliable uncertainty calculation using Bayesian inference. The fixed threshold value (0.00025 inch/cycle) allowed the model to clearly identify the potential risk of the crack growth.

Looking forward, future improvements mentioned in the Discussion part are required. These improvements include refining the α-balancing parameter, enhancing training stability across epochs, and validating the model against actual post-maintenance data. Deploying this PINN model in real-world SHM (Structural Health Monitoring) systems may enable predictive maintenance (PdM) scheduling and help prevent catastrophic engine failures in jet engines such as CFM56-7B engines.

Moreover, real-world validation using post-flight CFM56-7B engine data remains a crucial step to confirm the model's generalizability and operational accuracy. Adding such real-flight data would provide essential insights on model performance under actual service conditions and support further development beyond simulation-based testing.

FUNDING SOURCES

The author has no funding sources that supported the research and preparation of this article.

DECLARATION OF CONFLICT OF INTERESTS

The author declares that there are no conflicts of interest regarding the publication of this article.

APPENDIX

https://colab.research.google.com/drive/1b6KRZnEUChynndWwGmYEq5pUDlecUb-E?usp=sharing

REFERENCES

1. Tableau. *Sales Operations Dashboard*. Available from: https://public.tableau.com/app/profile/salesoperations/viz/shared/DRRPC5YX8 (accessed on 2025-02-06)
2. Boeing. *Statistical Summary of Commercial Jet Airplane Accidents*. Available from: https://www.boeing.com/content/dam/boeing/boeingdotcom/company/about_bca/pdf/statsum.pdf (accessed on 2025-02-06)
3. Airbus. *A Statistical Analysis of Commercial Aviation Accidents*, 2024. Available from: https://accidentstats.airbus.com/wp-content/uploads/2024/02/20230873_A-Statistical-analysis-of-commercial-aviation-accidents-2024-version.pdf (accessed on 2025-02-06)
4. SKYbrary. *Uncontained Engine Failure*. Available from: https://skybrary.aero/sites/default/files/bookshelf/4342.pdf (accessed on 2025-02-06)
5. Meng X, Li Z, Zhang D, Karniadakis GE. PPINN: Parareal physics-informed neural network for time-dependent PDEs. Computer Methods in Applied Mechanics and Engineering. 2020 Oct 1; 370: 113250. https://doi.org/10.1016/j.cma.2020.113250
6. ASTM International. *ASTM E647 – Standard Test Method for*

Measurement of Fatigue Crack Growth Rates, 2023. Available from: https://www.astm.org/e0647-23.html (accessed on 2025-03-22)

7. Dourado A, and Viana F. Physics-informed neural networks for corrosion-fatigue prognosis. *AIAA SciTech Forum*, 2019. Available from: https://arc.aiaa.org/doi/10.2514/6.2019-0113
8. Saxena A, and Goebel K. C-MAPSS dataset benchmark for prognostics. In: *Proceedings of the International Conference on Prognostics and Health Management (PHM)*, 2008. Available from: https://www.phmsociety.org/events/conference/phm/08/data-challenge
9. Yang Y, Meng X, and Karniadakis G. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *Journal of Computational Physics*. 2021; 425: 109913. Available from: https://doi.org/10.1016/j.jcp.2020.109913
10. Yucesan A. A physics-informed neural network for wind turbine main bearing fatigue. *Energies*. 2020; 13 (24): 6591. Available from: https://doi.org/10.3390/en13246591
11. Lundberg S, and Lee S-I. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2017. Available from: https://arxiv.org/abs/1705.07874 (accessed on 2025-03-22)
12. Rolls-Royce. *The Jet Engine*, 5th ed. John Wiley & Sons, 2015. ISBN: 978-0-470-01641-9.
13. Aircraft Commerce. *Owner's & Operator's Guide: CFM56-7B*, 2008. Available from: https://www.aircraft-commerce.com/wp-content/uploads/aircraft-commerce-docs1/Aircraft%20guides/CFM56-7B/ISSUE58_CFM56_7B_GUIDE.pdf (accessed on 2025-03-22)
14. Fatemi A. Cumulative fatigue damage and life prediction theories: A survey of the state of the art for homogeneous materials. *International Journal of Fatigue*. 1998; 20(1): 9–34. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0142112397000819. https://doi.org/10.1016/S0142-1123(97)00081-9
15. Li M, Wang X, Sun Y, and Chen H. Fatigue life prediction of aircraft engine blades using physics-informed neural networks. *Materials & Design*. 2024; 245: 111596. Available from: https://www.sciencedirect.com/science/article/pii/S0264127524006427
16. Fox M, Rossey W, and Raguet M. Fan blade fatigue fractures in CFM56-7B engines. *Journal of Failure Analysis and Prevention*. 2023; 23: 1438–1451. Available from: https://doi.org/10.1007/s11668-023-01702-y
17. ZEISS Group. Vibration analysis: Predictive maintenance. Available from: https://www.zeiss.com/metrology/en/explore/topics/vibration-analysis.html (accessed on 2025-03-22)
18. ASM International. Condition Monitoring. *ASM Digital Library*, 2017. Available from: https://dl.asminternational.org/handbooks/edited-volume/50/chapter-abstract/616388/Condition-Monitoring (accessed on 2025-03-22)
19. SKF Group. Bearing damage and failure analysis. Available from: https://cdn.skfmediahub.skf.com/api/public/0901d1968064c148/pdf_preview_medium/0901d1968064c148_pdf_preview_medium.pdf (accessed on 2025-03-22)
20. Wikipedia Contributors. Condition monitoring. *Wikipedia*, Sep. 17, 2019. Available from: https://en.wikipedia.org/wiki/Condition_monitoring (accessed on 2025-03-22)
21. Raissi M, Perdikaris P, and Karniadakis G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *arXiv preprint*, arXiv:1711.10561, 2017. Available from: https://arxiv.org/abs/1711.10561 (accessed on 2025-03-22)
22. Wassimderbel. NASA Predictive Maintenance (RUL). *Kaggle*, 2023. Available from: https://www.kaggle.com/code/wassimderbel/nasa-predictive-maintenance-rul (accessed on 2025-03-22)
23. Heimes, H. Recurrent neural networks for remaining useful life estimation. In: *Proceedings of the 2008 International Conference on Prognostics and Health Management*, Denver, CO, USA. 2008; p. 1–6. https://doi.org/10.1109/PHM.2008.4711422
24. Zhang A, Yucesan Y, and Shen H. Bayesian physics-informed neural networks for fatigue crack growth modeling. *arXiv preprint*, arXiv:2003.06097, 2020. Available from: https://arxiv.org/abs/2003.06097 (accessed on 2025-03-22)
25. Tao C, Zhang C, Ji H, and Qiu J. A Paris-law-informed neural fatigue cohesive model and its application to open-hole composite laminates. *International Journal of Solids and Structures*. 2023; 267: 112158. https://doi.org/10.1016/j.ijsolstr.2023.112158
26. Wang L, Zhang C, Tao C, Ji H, et al. Prediction of multiple fatigue crack growth based on modified Paris model with particle filtering framework. *Mechanical Systems and Signal Processing*. 2023; 190: 110124. https://doi.org/10.1016/j.ymssp.2023.110124
27. Verbist ML, Visser WPJ, van Buijtenen JP, and Duivis R. Gas Path Analysis on KLM In-Flight Engine Data. ASME Turbo Expo 2011: Power for Land, Sea, and Air, Vancouver, Canada. 2011; pp. 149–157. https://doi.org/10.1115/GT2011-45625
28. A GM(1,1) Markov Chain-Based Aeroengine Performance Degradation Forecast Approach Using Exhaust Gas Temperature – Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/EGTM-sequence-and-trend-analysis-for-CFM56-aeroengine-31_fig1_274918409 (accessed on 2025-07-15)
29. Li M, Wang X, Sun Y, and Chen H. Fatigue life prediction of aircraft engine blades using physics-informed neural networks. *Materials & Design*. 2024; 245: 111596. https://doi.org/10.1016/j.matdes.2024.111596ons of Prior Fatigue Models and Novelty of the Proposed Framework