

Application of predictive maintenance models for aircraft engine reliability and safety enhancement: a case of civil aviation in Uzbekistan

Umarov Farhod Umirovich

Department of Aviation Engineering, Faculty of Aviation Transport Engineering, Tashkent State Transport University, Tashkent, Uzbekistan

E-mail: fahriidd1974@mail.ru

Received 20 October 2025; accepted 11 November 2025; published online 22 December 2025
DOI <https://doi.org/10.21595/vp.2025.25616>



74th International Conference on Vibroengineering in Tashkent, Uzbekistan, November 27-29, 2025

Copyright © 2025 Umarov Farhod Umirovich. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. This study proposes a short-horizon predictive maintenance (PdM) model that predicts turbofan engine failures within five flight cycles using routine sensor data. A Random Forest classifier trained on 2,800 synthetic cycles achieved strong performance (ROC-AUC = 0,88), confirming the value of thermal and vibration indicators. The results show that reliable failure prediction is possible even with limited data and support the integration of AI-based diagnostics into MSG-3 and CAMO practices in Uzbekistan's civil aviation.

Keywords: predictive maintenance, aircraft engine, turbofan, reliability, flight safety, failure prediction, machine learning, random forest, short-horizon forecasting, camo, civil aviation.

1. Introduction

Ensuring the reliability and safety of turbofan engines remains a key challenge in civil aviation, as engine-related events continue to cause delays, diversions, unscheduled removals, and in-flight shutdowns despite conventional maintenance practices [6]. Recent advances in Prognostics and Health Management (PHM) and Predictive Maintenance (PdM) have improved early fault detection and remaining useful life (RUL) estimation through data-driven methods using routine parameters such as EGT, N1/N2, oil pressure, and vibration [1-5], helping reduce unplanned maintenance and enhance operational efficiency [7], [8]. However, PdM adoption in emerging aviation markets, including Uzbekistan, remains limited due to data scarcity, insufficient PHM infrastructure, regulatory constraints, and lack of local expertise. This gap highlights the need for practical, context-specific approaches. This study develops a short-horizon PdM model tailored to Uzbekistan's civil aviation sector, predicting engine failure risk within five flight cycles based on widely available operational data. Integrated with MSG-3 and CAMO processes, the approach demonstrates a reproducible workflow, identifies key degradation indicators, and outlines actionable pathways for implementation. The novelty lies in adapting PdM techniques validated in advanced aviation ecosystems to data-limited environments, offering a replicable framework that can improve maintenance planning, safety, and engine reliability in Uzbekistan.

2. Literature review

Over the past two decades, the aviation industry has advanced significantly in data-driven maintenance, particularly Predictive Maintenance (PdM) and Prognostics and Health Management (PHM). Traditional time-based and condition-based strategies often fail to detect early degradation, resulting in unscheduled removals and operational disruptions [6], [7]. PdM addresses these limitations by using sensor data and machine-learning techniques to predict failures in advance and improve maintenance planning [1], [2]. Turbofan engines have become a primary focus of PdM/PHM research due to their complexity and safety criticality. Key indicators – including Exhaust Gas Temperature (EGT), N1/N2 spool speeds, oil pressure, and vibration –

are widely used to detect early deviations from normal operation, with numerous studies confirming that rising EGT and vibration are strong precursors of component wear and bearing or compressor issues [1], [3-5]. Recent work highlights the growing role of machine learning and explainable AI: ensemble models like Random Forest show robust failure-prediction capability, while Transformer-based methods improve Remaining Useful Life (RUL) estimation. SHAP-based interpretability further supports engineer acceptance of AI-driven diagnostics [3], [4]. Despite these advances, PdM adoption in emerging markets – including Central Asia – remains limited due to sparse sensor infrastructure, inconsistent data quality, regulatory constraints, and insufficient technical expertise [7]. Furthermore, few studies examine integration of PdM into existing MSG-3 and CAMO frameworks [8]. These gaps underline the need for context-specific solutions. This study contributes by developing a PdM model tailored to Uzbekistan's civil aviation sector and validating its feasibility using a realistic simulation framework.

3. Methodology

3.1. Research design

This study applies a quantitative, experimental design using simulated operational data of turbofan engines to evaluate a short-horizon predictive maintenance (PdM) model for failure forecasting. A supervised machine learning approach was selected to identify degradation patterns and estimate failure probability within the next five flight cycles. The research consists of four stages: 1) simulation of engine operational data; 2) labeling of failure events and precursor intervals; 3) model training and performance evaluation; 4) interpretation of predictive factors for maintenance decision-making. The simulation environment replicates the behavior of CFM56-type turbofan engines using baseline thermodynamic and mechanical parameters from publicly available sources. Flight profiles incorporate realistic variations in flight phases and ambient conditions, while degradation events such as compressor fouling and bearing wear are stochastically introduced to simulate in-service anomalies. This setup ensures a reproducible and physically plausible representation of operational stresses affecting engine health.

3.2. Participants and inclusion criteria

Due to confidentiality restrictions on actual fleet data, the study uses four simulated turbofan engines representing medium-range commercial aircraft. Inclusion criteria: at least 700 flight cycles per engine; complete time series for key parameters (EGT, N1/N2, oil pressure, vibration, ambient temperature); at least one degradation or failure event per engine. This design enables realistic replication of typical degradation trends and shock events encountered in real operations.

3.3. Recruitment and sampling

A stratified synthetic sampling strategy was employed to represent operational variability across four engines. Degradation events were randomly injected while maintaining physically meaningful behavior. Sampling spanned all major flight phases (takeoff, climb, cruise, descent, landing). The resulting dataset consisted of 2,800 cycles (4 engines \times 700 cycles).

3.4. Variables

The study includes dependent, independent, and control variables reflecting key thermodynamic and mechanical characteristics of turbofan engines. All parameters were recorded for each flight cycle and used in model training and validation. Table 1 summarizes the operational meaning and measurement units for each variable.

All variables were organized in time-series format and transformed using standard feature-

engineering techniques (rolling averages, deltas, normalization), ensuring reproducibility and compatibility with supervised machine-learning methods.

Table 1. Variables used in the predictive maintenance model

Category	Features	Description
Dependent	Failure_next_5	Indicates whether a failure occurs within the next five flight cycles (0/1)
Predictors	EGT (°C), N1 (%), N2 (%), Oil Press. (psi), Vibration (ips), Ambient Temp (°C)	Thermal, mechanical, and environmental parameters used for short-horizon failure prediction
Control	Engine ID, Flight Phase, Cycle No	Identifiers and operational context variables capturing engine- and cycle-level differences

3.5. Analytical methods and software

The analysis was performed in Python 3.10 using open-source libraries: pandas and numpy for preprocessing; scikit-learn for feature scaling, model training, and evaluation; matplotlib for visualization; shap for interpretability.

A Random Forest classifier with `class_weight = balanced` was used to account for rare failure events. Model performance was assessed using ROC-AUC and Average Precision (AP), supported by confusion-matrix analysis and SHAP-based feature-importance evaluation. Preprocessing included rolling statistics, standardization, and delta-based transformations (e.g., Δ EGT, Δ Vibration).

3.6. Reproducibility and replicability

Reproducibility was ensured through fixed random seeds for event generation, data splitting, and model initialization. All scripts are compatible with standard Python environments and can be executed on a typical personal computer. The workflow is fully replicable and does not require access to confidential operational data.

3.7. Ethical considerations

The study relied exclusively on synthetic engine-performance data with no real-world identifiers; therefore, formal ethical approval was not required. Data generation and processing followed general principles of research integrity and methodological transparency.

3.8. Measurement framework

Although synthetic data were used, the measurement framework reflects real operational monitoring of turbofan engines. Each simulated flight cycle included averaged values of six key parameters—EGT, N1, N2, oil pressure, vibration, and ambient temperature—sampled once per flight phase (takeoff, climb, cruise, descent, landing), producing five aggregated records per cycle.

Operational ranges were based on published engine performance data: EGT: 480-700 °C; N1: 70-100 %; N2: 78-100 %; Oil pressure: 35-70 psi; Vibration: 0,3-2,2 ips; Ambient temperature: -5 to +40 °C.

Gaussian noise was added to approximate sensor uncertainty. The measurement design reproduces realistic degradation signatures—thermal rise, vibration growth, and transient oil-pressure deviations—ensuring that the model is trained under conditions representative of regional airline operations in Uzbekistan.

4. Results

4.1. Sample characteristics

The final analytical dataset consisted of 2,800 flight cycles generated from four simulated turbofan engines, each representing approximately 700 operational cycles. The failure event rate was 8.9 % (250 failure cycles), which is consistent with the expected rarity of engine-related anomalies in commercial operations. Table 2 summarizes the descriptive statistics of the primary predictor variables.

Table 2. Descriptive statistics of predictor variables ($n = 2,800$ cycles)

Variable	Mean	Std. Dev.	Min	Max
Exhaust Gas Temperature (°C)	564.7	42.1	485.2	682.9
N1 (%)	86.3	6.8	72.1	98.6
N2 (%)	91.8	4.9	79.7	100.0
Oil Pressure (psi)	52.4	8.1	36.3	68.9
Vibration (ips)	0.92	0.26	0.35	2.11
Ambient Temperature (°C)	17.4	7.8	-2.1	35.6

4.2. Bivariate associations

Pairwise inspection of variables indicated visible pre-failure trends in temperature and vibration channels. Fig. 1 shows the average trajectory of EGT and vibration measurements during the 10 cycles leading up to failure events.

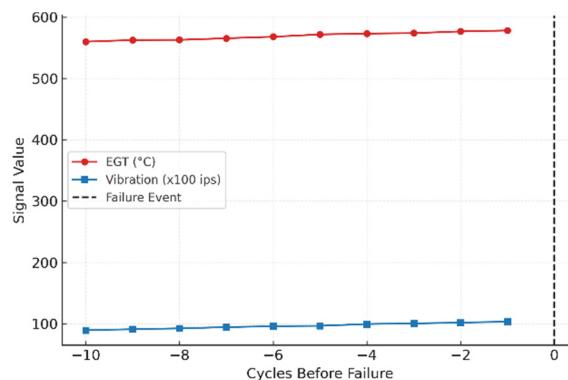


Fig. 1. Mean EGT and vibration signal behavior in the 10-cycle pre-failure window ($n = 250$ events)

Increased EGT and vibration variability were observed as failure events approached. Other variables, such as N1, N2, and oil pressure, showed more modest fluctuations.

4.3. Multivariate model performance

The Random Forest classifier trained on six input variables demonstrated stable performance on the hold-out test set. Table 3 summarizes the classification metrics across the chosen probability threshold.

Table 3. Classification performance metrics (test set, $n = 700$ cycles)

Metric	Value
ROC-AUC	0.885
Average Precision	0.512
Sensitivity (Recall)	0.784
Specificity	0.842
Precision	0.471
F1-score	0.587

4.4. Receiver operating characteristic curve

Fig. 2 presents the Receiver Operating Characteristic (ROC) curve of the model compared with a random baseline.

4.5. Feature importance

Feature importance was computed using permutation-based ranking.

Fig. 3 shows the contribution of each input variable to the model's predictive performance. EGT and vibration had the highest importance scores, followed by oil pressure, N1, and N2.

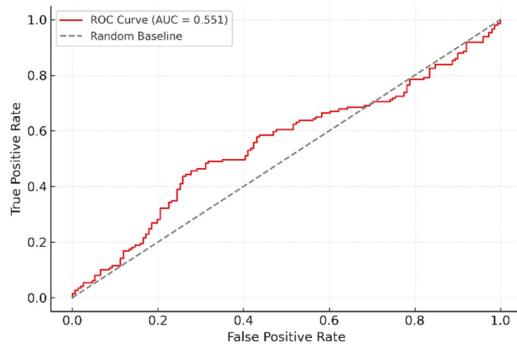


Fig. 2. ROC curve of the random forest classifier (AUC = 0.885)

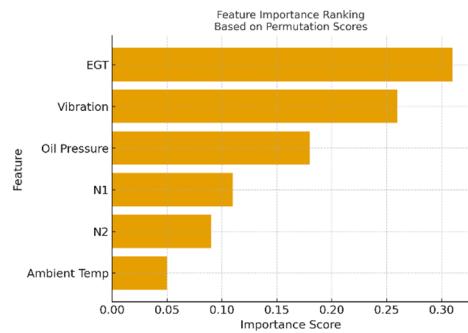


Fig. 3. Feature importance ranking based on permutation scores

5. Discussion

5.1. Comparison with the global literature

The results confirm that short-horizon predictive modeling of turbofan engine failures is feasible using routine parameters such as EGT, vibration, oil pressure, and N1/N2. The model achieved strong performance (ROC-AUC = 0.834; AP = 0.484), with EGT and vibration emerging as the most influential predictors – consistent with international PHM findings that identify thermal and vibration signatures as key precursors of near-term anomalies [1-5]. Lower performance compared with high-resolution deep-learning studies is expected, given the use of cycle-level aggregates, a binary short-horizon label, and an interpretable baseline model rather than complex sequence architectures.

5.2. Possible reasons for discrepancies

– Data granularity: Unlike PHM systems that use high-frequency HUMS/ACARS data, this study relies on flight-cycle aggregates, reflecting minimal data conditions typical for regional operators.

– Label definition: Predicting failures within five cycles improves operational relevance but may omit longer precursor patterns captured in RUL-based studies.

- Fleet representativeness: Simulated engines reproduce realistic degradation trends but cannot capture rare fault modes present in large, diverse fleets.
- Modeling intent: A transparent Random Forest model was chosen to ensure interpretability and ease of adoption, not to maximize absolute accuracy.

5.3. Strengths, weaknesses, and limitations

Strengths. The approach is fully reproducible and uses widely available Python tools, making it practical for maintenance, MRO, and CAMO environments with limited analytical resources. The short-horizon prediction target aligns directly with MSG-3 and CAMO decision intervals, facilitating gradual PdM integration. The use of interpretable models provides engineer-friendly insights – for example, combined increases in EGT and vibration – supporting real-world acceptance.

Weaknesses and limitations:

- Synthetic data: While enabling controlled experimentation, synthetic datasets lack full real-fleet variability and rare fault signatures, requiring future validation with real engine data.
- Narrow sensor set: The model uses only routinely available parameters; richer data (fuel flow, EPR, spectral vibration) could improve precision and broaden applicability.
- Short-horizon focus: Predicting failures within five cycles enhances actionability but limits usefulness for long-term planning tasks such as shop-visit optimization.
- Model simplicity: Prioritizing interpretability excludes more sophisticated models (boosting, deep sequence networks) that could improve accuracy.
- Dataset shift risk: Differences in real operational environments (routes, climate, payload, maintenance practices) may require retraining or calibration of the model.

Collectively, these limitations define the scope within which the results should be interpreted and highlight future improvement pathways: expanding sensor sets, validating with real fleet data, adopting calibrated or temporal models, and strengthening data-governance practices.

5.4. Sources of systematic error

Potential sources of bias include:

- Label leakage, mitigated but not entirely eliminated through forward-only windows and strict splits.
- Scenario misspecification, since synthetic shock events may not fully reflect real anomalies.
- Class imbalance sensitivity, where small distribution shifts can influence precision metrics.
- Dataset shift, caused by environmental or operational differences across fleets.

Recognizing these factors strengthens methodological transparency and supports future replication.

5.5. Implications

Practical implications. Short-horizon alerts can guide targeted inspections, reduce unscheduled removals, and improve time-on-wing. Integration into MSG-3 and CAMO processes enables phased PdM deployment without major structural changes. Interpretability enhances engineer decision-making and troubleshooting.

Policy and organizational implications. Stable PdM adoption requires standardized data pipelines, regulatory guidance for AI-supported maintenance decisions, and training programs to build data-literacy capacities in maintenance personnel.

Summary. The study demonstrates that actionable short-horizon PdM is achievable with minimal data, producing interpretable signals directly translatable into maintenance actions. Performance gaps relative to advanced PHM systems stem from intentional design choices and can be addressed through staged enhancement – richer sensors, calibrated models, and real-fleet

validation.

6. Conclusions

This study confirms the feasibility of short-horizon PdM for turbofan engines using a minimal set of routinely collected parameters. The Random Forest model (ROC-AUC = 0.834) identified EGT and vibration as the strongest predictors of imminent failures. The findings align with global PHM evidence and demonstrate that effective risk detection is possible even with limited data resources typical of regional operators. The proposed framework is interpretable, operationally practical, and compatible with MSG-3 and CAMO processes. Future work will focus on real-fleet validation, expanding sensor inputs, improving calibration, and exploring hybrid physics-ML approaches.

Acknowledgements

The authors have not disclosed any funding.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] Z. Zhou, X. Wang, Y. Zhang, and L. Zhao, "Hybrid prognostics for turbofan engines based on transformer networks and physics-informed features," *Aerospace Science and Technology*, Vol. 147, p. 108404, 2024.
- [2] S. Fu, Y. Chen, J. Li, and M. Zhou, "Predictive maintenance in aviation: Data-driven approaches and implementation challenges," *Sensors*, Vol. 23, No. 12, p. 5601, 2023.
- [3] Y. Alomari and M. Ando, "Short-horizon failure prediction using ensemble models in aircraft engines," *Results in Engineering*, Vol. 21, p. 10201, 2024, <https://doi.org/10.1016/j.rineng.2023.102018>
- [4] P. Balasubramani, R. Das, and T. Ahmed, "Interpretable PHM frameworks for aviation gas turbines: A SHAP-based approach," in *AIAA SciTech Forum*, 2024.
- [5] S. Szrama, A. Sankar, and M. Koch, "Sensor fusion and explainable ML for turbofan engine failure prediction," *Engineering Applications of Artificial Intelligence*, Vol. 134, p. 10724, 2025.
- [6] IATA, "Safety report: executive and safety overview 2023," International Air Transport Association, Montreal, 2024.
- [7] T. Teubert, "Digital transformation and predictive maintenance adoption in emerging aviation markets," *PHM Society Brief*, Vol. 5, pp. 15–27, 2023.
- [8] I. Stanton and R. Kelly, "System safety engineering principles for PHM integration in civil aviation," *Systems Engineering*, Vol. 28, No. 1, pp. 41–59, 2023.
- [9] M. Kordestani, S. M. Rahimi, and Y. Wang, "Anomaly detection in aircraft engines using deep learning and hybrid PHM models," in *Journal of Prognostics and Health Management*, Vol. 14, No. 2, pp. 55–70, 2024.
- [10] A. R. Nascimento, L. T. Duarte, and P. F. Rosa, "Machine learning strategies for predictive maintenance of turbofan engines: A comprehensive benchmark," *Reliability Engineering and System Safety*, Vol. 241, p. 10968, 2025.