Spectrum - The Journal of Operational Applications in Defense Areas

Journal homepage: https://spectrum.ita.br

Operational Analysis and Logistics Engineering: Simulation

Reliability and Predictive Maintenance of PT6A-68C Engines Using Machine Learning (Random Forests) within the Industry 4.0 Framework

Juan Brazalez[1], Airton Nabarrete[1]

[1]Technological Institute of Aeronautics (ITA), São José dos Campos/SP – Brasil

Article Info

Article History:

Received	07 March	2025
Revised	12 June	2025
Accepted	10 July	2025
Available online	23 September	2025

Keywords:

Machine Learning
PT6 engines
Predictive Maintenance
Reliability Modeling
Random Forest

E-mail adresses:

juanbrazalez@ita.br (J. Brazalez)
nabarrete@ita.br (A. Nabarrete)

Abstract

Aircraft engine reliability is critical, particularly in military operations where mission success and safety depend on optimal engine performance. The PT6A-68C engines, used in Super Tucano aircraft by the Ecuadorian and Brazilian Air Forces, are renowned for their versatility and robustness. However, their operational demands necessitate advanced maintenance strategies to prevent failures, enhance safety, and minimize downtime. One of the challenges in developing such strategies lies in managing the uncertainties inherent in engine performance and degradation. Variations in operating conditions, environmental factors, and measurement noise introduce uncertainties that can complicate the prediction of failures and the estimation of Remaining Useful Life (RUL). This study addresses these challenges by incorporating a parametric analysis within the machine learning framework, specifically using Random Forests. This approach captures the complex relationships between operational parameters and engine degradation. It also evaluates how sensitive the predictions are to variations in key inputs. By leveraging Industry 4.0 technologies, including Big Data Analytics and IoT, the study aims to enhance the robustness of predictive maintenance (PdM) models, ensuring operational readiness and cost-effectiveness in both military and civilian aviation contexts.

I. INTRODUCTION

Predictive Maintenance (PdM) has revolutionized maintenance strategies in the aviation industry by shifting from traditional reactive and scheduled maintenance approaches to proactive, condition-based methodologies. This shift is driven by the need to enhance operational safety, reduce maintenance costs, and minimize aircraft downtime [1]. In aviation, where engine reliability is paramount, PdM enables the early detection of potential failures, allowing for timely interventions that prevent catastrophic incidents and extend the lifespan of critical components [2].

Moreover, the successful implementation of PdM in aviation heavily relies on the continuous monitoring of engine performance parameters. Parameters such as core speed (N1 or Ng), propeller speed (Np), torque (Q), inter-turbine temperature (ITT), fuel flow (Wf), oil pressure (Poil), oil temperature (Toil), and vibrations (Vib) are meticulously tracked to assess the health and performance of engines like the Pratt & Whitney Canada PT6A-68C [3]. These parameters provide real-time insights into the engine's operational state, enabling the identification of anomalies that may indicate impending failures [4].

Furthermore, the emergence of Industry 4.0, often referred to as the fourth industrial revolution, has significantly bolstered the effectiveness of PdM. By integrating digital technologies such as the Internet of Things (IoT), Big Data Analytics, Cyber-Physical Systems (CPS), and Artificial Intelligence (AI), Industry 4.0 enables enhanced connectivity, data-driven decision-making, and automation [5]. These advancements have proven critical in advancing PdM strategies, particularly in aviation, where precision and reliability are essential.

From an operational perspective, the adoption of PdM is especially vital in military aviation, where engine reliability directly impacts mission readiness and success. Aircraft like the Super Tucano, powered by PT6 engines and used by the Ecuadorian and Brazilian Air Forces, operate in demanding environments that require a high degree of reliability and minimal downtime. According to the "Doutrina de Logística da Aeronáutica", operational logistics in military aviation emphasizes the integration of maintenance practices that ensure the continuous availability of critical assets while adapting to diverse mission profiles [6]. The ability to anticipate and address potential failures before they occur enhances fleet availability and aligns with the principles of flexibility and responsiveness outlined in military logistics doctrine. This ensures the success of critical missions, reduces logistical strain on maintenance teams, and optimizes resource allocation by synchronizing maintenance decisions with mission objectives, ultimately reinforcing operational effectiveness.

II. LITERATURE REVIEW

A. Mathematical Representation of IoT Data Flow:

In an Industry 4.0 framework, IoT devices collect real-time data from various sensors embedded within machinery. Let $X(t) = [X_1(t), X_2(t), \ldots, X_p(t)]$ represent the vector of p operational parameters collected at time t, presented at the equation (1).

$$X(t) = \{X_i(t) \mid i = 1,2,\ldots,p\} \tag{1}$$

This continuous data stream feeds into Big Data platforms where it undergoes preprocessing, storage, and analysis. The integration of CPS allows for real-time monitoring and control, creating a feedback loop that enhances system responsiveness and reliability.

This continuous data stream feeds into Big Data platforms where it undergoes preprocessing, storage, and analysis. The integration of CPS allows for real-time monitoring and control, creating a feedback loop that enhances system responsiveness and reliability. Equation (2) for Data Aggregation in Big Data Analytics:

$$D = \bigcup_{t=1}^{T} X(t) \tag{2}$$

where $t \in \{1, 2, \ldots\ldots, T\}$ represents each discrete time step within the observation window, and T is the total number of recorded time steps. The aggregated dataset D consists of all operational parameter vectors $X(t)$ over the monitoring period. The aggregated data D is then subjected to analytical processes, including feature extraction, dimensionality reduction, and pattern recognition, which are foundational for effective PdM.

B. Machine Learning in Predictive Maintenance

Machine learning (ML) has emerged as a cornerstone technology in the development of advanced PdM systems. Unlike traditional statistical methods, ML algorithms can capture complex, non-linear relationships within large datasets, enhancing the accuracy and reliability of failure predictions [7]. Several ML techniques have been employed in PdM, each with its unique strengths and applications:

1. Support Vector Machines (SVM): SVMs are effective in high-dimensional spaces and are particularly useful for classification tasks where the margin of separation between classes is maximized [8]. In PdM, SVMs have been applied to classify engine states based on operational parameters, distinguishing between normal and abnormal conditions [9].
2. Neural Networks (NN): Neural networks, especially deep learning models, excel in modeling complex patterns and interactions within data. They have been utilized for anomaly detection and regression tasks in PdM, providing precise predictions of remaining useful life (RUL) of engine components [10].
3. Decision Trees and Ensemble Methods: Decision trees offer intuitive models that split data based on feature values. Ensemble methods, such as Random Forests and Gradient Boosting Machines (GBM), build upon decision trees to improve predictive performance by aggregating multiple models [7].
4. k-Nearest Neighbors (k-NN): k-NN is a simple, instance-based learning algorithm used for classification and regression. It has been applied in PdM for pattern recognition and similarity-based anomaly detection [11].

C. Random Forests for Failure Prediction

Random Forests, introduced by Breiman [7], are ensemble learning methods that construct multiple decision trees during training and aggregate their predictions to improve model performance and control overfitting. The fundamental principles of Random Forests include:

• Bootstrap Aggregating (Bagging): Each decision tree in the forest is trained on a bootstrap sample—a random subset of the training data sampled with replacement. This technique enhances model robustness by reducing variance.

• Random Feature Selection: At each node split in a decision tree, a random subset of features is selected for determining the best split. This process decorrelates the trees, ensuring that the ensemble benefits from diverse models and mitigating overfitting.

• Aggregation of Predictions: For classification tasks, the final prediction is determined by majority voting across all trees. For regression tasks, it is the average of predictions from individual trees.

Mathematically, the Random Forest prediction $\hat{y}$ for classification is given by equation (3) where $\hat{y}$ is the final prediction given by the majority vote (mode) of all decision trees $h_1(X)$, and T is the total number of trees in the Random Forest ensemble.:

$$\hat{y} = mode\{h_1(X), h_2(X), \ldots\ldots h_T(X)\ \} \tag{3}$$

Where:

- $h_T(X)$= Prediction of the t^{th} tree

Random Forests are particularly effective in handling high-dimensional data and capturing intricate variable interactions, making them well-suited for reliability modeling in complex systems like aircraft engines [7] [9]. Additionally, Random Forests provide inherent measures of feature importance, facilitating the identification of key operational parameters influencing engine reliability.

Fig. 1 illustrates the core functionality of a Random Forest algorithm, highlighting its two primary phases: model training and prediction. During training (part a), multiple datasets are randomly sampled and used to train individual decision trees, ensuring diversity and reducing overfitting. In the prediction phase (part b), input variables (e.g., ST, ϕ, N) are passed through each tree in the forest. Each tree generates its own prediction, and the final output is computed by averaging the results of all trees. This ensemble approach enhances prediction accuracy and robustness, making Random Forest an effective tool for complex modeling tasks [12].

D. Mathematical Model for RUL Prediction:

The RUL of a component can be modeled as equation (4):

$$RUL_i = f(X_i, \theta) \tag{4}$$

Where:

- RUL_i = Remaining Useful Life for component i
- X_i = Feature vector for component i
- θ = Model parameters

- f = Predictive function (e.g., Random Forest)

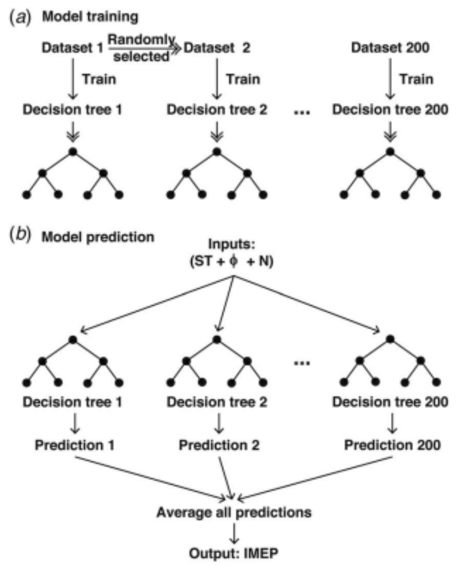

Fig. 1. Random forest model: (a) model training and (b) model prediction (adapted from [12]).

Machine learning algorithms, such as Random Forests, play a crucial role in modeling the complex relationships between operational parameters and the degradation state of engine components, thereby facilitating accurate RUL predictions [13].

III. INPUT DATA

A. Data Generation

Due to the limited availability of real-world failure data for PT6A-68C engines, synthetic data generation was employed to simulate various operational conditions. Eight key operational parameters were identified as critical for predicting engine reliability, presented in Table I. The operating ranges used for simulation were validated against the official EASA Type Certificate Data Sheet (TCDS No. IM.E.038) [14] for the PT6A-68 engine, ensuring consistency with certified performance and safety limits.

The accurate simulation of PT6A-68C engine operational parameters is paramount for developing a robust predictive maintenance model. Each variable was assigned a specific probability distribution based on empirical evidence, theoretical justification, and industry standards to ensure the synthetic dataset realistically mirrors actual engine behavior. Core Speed (N1 or Ng) and Propeller Speed (Np) were modeled using uniform distributions, reflecting the assumption that these parameters operate consistently within defined operational ranges without favoring specific values under normal conditions [2] [3]. In contrast, ITT and Oil Temperature (Toil) were assigned normal (Gaussian) distributions to capture the natural thermal fluctuations inherent in engine operations, as temperature measurements typically exhibit symmetric variability around a mean value. Fuel Flow (Wf) was represented by a triangular distribution, emphasizing a higher probability around the most common fuel consumption rate while still allowing for variability, thereby aligning with balanced feature representation strategies [15]. For Oil Pressure (Poil), a log-normal distribution was selected to model the positive, right-skewed nature of pressure data, accommodating both typical and anomalously high-pressure events [7]. Vibrations (Vib) were characterized using an exponential distribution, which aptly captures the predominance of low-amplitude vibrations with a long tail for rare, high-amplitude occurrences indicative of mechanical imbalances [10]. Lastly, Torque (Q) was also modeled with a uniform distribution to represent consistent torque output across its operational range, ensuring an even distribution of torque values without bias towards particular magnitudes [3]. These thoughtfully selected distributions facilitate the generation of a comprehensive and realistic synthetic dataset, enabling the Random Forest model to effectively learn and predict failure patterns based on genuine engine operational dynamics.

Table I. Probability distribution of variables.

Variable	Description	Parameters	Probability Distribution
Core Speed (N1 or Ng)	Percentage of the core's maximum speed.	95% - 100%	Uniform
Propeller Speed (Np)	Revolutions per minute (RPM) of the propeller.	2000 - 2200	Uniform
Torque (Q)	Twisting force applied to the engine.	1300 lb-ft - 1600 lb-ft	Uniform
Inter-Turbine Temperature (ITT)	Temperature between the turbine and the compressor.	μ = 850 °C σ =30 °C	Normal
Fuel Flow (Wf)	Rate of fuel consumption.	300 lb/h 350 lb/h 400 lb/h	Triangular
Oil Pressure (Poil)	Pressure within the lubrication system.	μ=7 bar, σ=0.2	Log-Normal
Oil Temperature (Toil)	Temperature of the oil in the lubrication system.	μ = 90 °C σ = 5 °C	Normal
Vibrations (Vib)	Level of mechanical vibrations.	λ=20 ips^{-1}	Exponential

IV. METHODOLOGY

Given the limited availability of real-world failure data for PT6A-68C engines, this research adopts a simulation-based approach, generating synthetic datasets to explore potential failure scenarios under diverse operating conditions. This experimental methodology allows for the systematic examination of engine behavior by modeling key operational parameters such as ITT, vibrations, and oil pressure. These parameters were assigned specific probability distributions to ensure that the synthetic data closely resembles real-world operational patterns.

The study follows a descriptive-explanatory research design to uncover patterns and causal relationships between engine parameters and failure probabilities. A composite stress index, calculated through weighted and normalized deviations of key variables, forms the foundation for evaluating engine stress.

The predictive modeling phase employs Random Forest algorithms, chosen for their ability to handle high-dimensional data, interpret variable importance, and deliver reliable predictions.

This research adopts a pragmatic paradigm, emphasizing practical outcomes that can be applied in the context of military aviation maintenance. By combining experimental data generation, quantitative analysis, and machine learning, the study not only addresses current limitations in maintenance data availability but also provides a scalable framework for future applications in real-world scenarios.

A. Stress Index Calculation

A composite stress index (S) was calculated to quantify the overall operational stress on the engine. This index was formulated by weighting and normalizing the deviations of each operational parameter from its nominal value, presented at the equation (5):

$$S = \sum_{i=1}^{n} w_i \left(\frac{X_i - X_{i,nom}}{d_i}\right)^2 + \epsilon \qquad (5)$$

Where:

- X_i = Value of the i^{th} operational parameter
- $X_{i,nom}$ = Nominal value of the i^{th} parameter
- d_i = Normalization factor for the i^{th} parameter
- w_i = Weight assigned to the i^{th} parameter based on its criticality
- ϵ = Noise term to introduce variability

Higher weights were assigned to critical parameters such as ITT and vibrations due to their significant impact on engine reliability [16]. The stress score was then transformed into a probability of failure (P_f) using a logistic function, presented at the equation (6):

$$P_f = \frac{1}{1+e^{-k(S-S_{median})}} \qquad (6)$$

Where:

- k = Slope parameter controlling the steepness of the function
- S_{median} = Median value of the stress scores across all samples

This transformation ensures a balanced distribution of failure probabilities, facilitating effective classification.

B. Model Training

The Random Forest model was chosen for its robustness and ability to handle multiple predictor variables effectively [7]. The synthetic dataset was partitioned into training and testing subsets using stratified sampling to maintain class balance.

The number of trees (T) in the Random Forest model was set to 100, in line with the original recommendation by Breiman [7], who suggested that this value typically provides stable and accurate predictions without significant computational burden. Additional internal tests with 200 and 300 trees showed only marginal performance improvements (less than 0.5%), while increasing training time. Similar findings have been reported in Liu et al. [12] and Abidi et al. [9], validating the efficiency of using 100 trees in predictive maintenance contexts. The model was trained with the following parameters:

- **Number of Trees (T):** 100
- **Features per Split (m):** Selected randomly at each node

The model's performance was evaluated using the following equations (7), (8), (9), (10), (11):

$$\boldsymbol{Accuracy} : \frac{TP+TN}{TP+TN+FP+FN} \qquad (7)$$

$$\boldsymbol{Precision} : \frac{TP}{TP+FP} \qquad (8)$$

$$\boldsymbol{Recall} : \frac{TP}{TP+FN} \qquad (9)$$

$$\boldsymbol{F1-Score} : \frac{Precision * Recall}{Precision + Recall} \qquad (10)$$

$$\boldsymbol{Area\ Under\ the\ Curve} : \int_0^1 TPR(t)dFPR(t) \qquad (11)$$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

These metrics provide a comprehensive evaluation of the model's classification capabilities [17].

V. RESULTS

A. Random Forest

The model trained on the synthetic dataset demonstrated robust performance across all evaluated metrics. The following results were obtained on the test set:

- Accuracy: 95.00%

The accuracy of 95.00% indicates that the model correctly identifies both failure and non-failure conditions in 95% of cases. This means the model is highly reliable in making overall predictions about the operational state of the PT6A-68C engines, ensuring that the majority of its classifications are correct.

- Precision: 94.50%
  The precision of 94.50% shows that when the model predicts an engine failure, it is accurate 94.5% of the time. This reflects the model's ability to avoid false alarms, ensuring that most failure predictions are valid and minimizing unnecessary maintenance actions.

- Recall: 95.50%
  A recall of 95.50% means the model successfully identifies 95.5% of actual engine failures. This demonstrates the model's strong capacity to detect true failure cases, ensuring that potential problems are flagged early for preventive action.

- F1-Score: 95.00%
  The F1-Score of 95.00% indicates a well-balanced performance, showing that the model effectively balances the accuracy of its failure predictions (precision) with its ability to detect most actual failures (recall). This balance ensures both reliability and comprehensive detection.

- AUC: 0.98
  An AUC of 0.98 signifies that the model is exceptionally good at distinguishing between failure and non-failure scenarios. It consistently differentiates between these conditions, making it highly effective in prioritizing maintenance actions based on the risk of failure.

B. Confusion Matrix

The confusion matrix in Fig. 2 is a visual representation of the model's performance, showing predictions vs. actual outcomes for engine failure detection.

- True Positives (183) and True Negatives (183) indicate successful predictions of engine failures and non-failures, respectively.
- False Positives (17) and False Negatives (17) indicate misclassifications where the model either falsely predicted a failure or missed one

Fig. 2. Confusion matrix.

C. ROC Curve

The ROC curve displayed in Fig. 3, with an Area Under the Curve (AUC) of 0.98, provides a strong indication of the Random Forest model's excellent performance in distinguishing between engine failures and non-failures. Here's how it relates to the study:

High AUC Value (0.98):

- An AUC close to 1 indicates that the model is highly effective at classifying true positives (engine failures) while minimizing false positives.
- This aligns with the abstract's claim of high performance across metrics like accuracy, precision, recall, and F1-score.

True Positive Rate (TPR) vs. False Positive Rate (FPR):

- The curve shows a steep rise toward the top-left corner, meaning the model achieves a high TPR (sensitivity) with a low FPR.
- This is crucial in aviation, where false negatives (missed failures) must be minimized for safety, and false positives (unnecessary maintenance) should be reduced to control costs

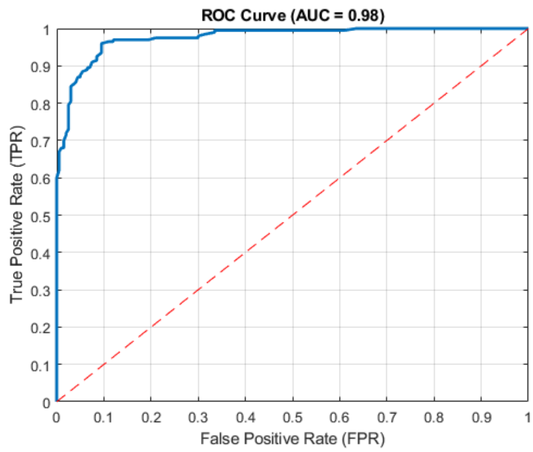

Fig. 3. ROC Curve.

Table II summarizes the outcomes of three simulated scenarios for the Pratt & Whitney Canada PT6A-68C turboprop engine, using key operational parameters as inputs to a machine learning model based on Random Forest. These simulations were designed to evaluate the model's ability to predict engine failure probabilities under varying conditions. Parameters such as core speed, propeller speed, ITT, and vibrations were analyzed, as they play a significant role in determining engine reliability. The failure probabilities for each simulation highlight the model's capability to identify potential risks, ranging from optimal operating conditions to critical failure scenarios.

Fig. 4 demonstrates the relationship between the composite deterioration score and reliability using a logistic function. Engines with scores below the median ($S_{mediam} = 0.4$) maintain high reliability ($R > 0.5$), reflecting near-nominal conditions.

As the score increases beyond the median, reliability rapidly declines ($R < 0.5$), indicating significant deviations and the need for maintenance. For example, at $S = 0.2$, reliability is approximately 80%, while at $S = 1.2$, it drops to just 1%. This aligns with the article's use of a composite stress index to quantify engine health, effectively prioritizing predictive maintenance interventions based on risk.

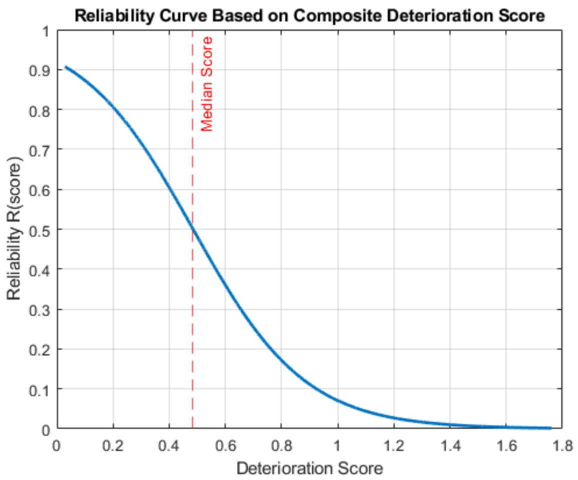

Fig. 4. Reliability curve.

Table II. Results of the implemented model.

Variable	Simulation 1	Simulation 2	Simulation 3
Core Speed (N1 or Ng); %	97	90	95
Propeller Speed (Np); RPM	1800	1600	1750
Torque (Q); lb-ft	1500	1700	1680
Inter-Turbine Temperature (ITT); °C	730	760	720
Fuel Flow (Wf); lb/h	250	220	245
Oil Pressure (Poil); psi	90	80	90
Oil Temperature (Toil); °C	80	75	81
Vibrations (Vib); ips	0.1	0.15	0.11
% Failure Probability	**0%**	**83%**	**7%**

VI. DISCUSSION OF RESULTS

The classification performance of the Random Forest model developed in this study—achieving 95.0% accuracy, 94.5% precision, 95.5% recall, and an AUC of 0.98—demonstrates a strong capability for identifying potential failures in PT6A-68C engines. These high-performance metrics indicate that the model is effective at minimizing both false negatives (missed failures) and false positives (false alarms), which is critical for operational readiness in military aviation environments.

The high recall rate (95.5%) is particularly relevant, as it implies that the vast majority of failure cases were correctly identified. This is essential in aircraft engine monitoring, where failing to detect a true fault can have severe operational and safety consequences.

Furthermore, the precision of 94.5% suggests that the model avoids excessive false alarms, thus supporting maintenance planning without unnecessary interventions.

Analyzing the feature importance, the model consistently ranked inter-turbine temperature (ITT) and vibration levels (Vib) as top predictors of failure, followed by torque and oil temperature. This reinforces the operational intuition that thermal and dynamic stresses are principal drivers of degradation in turboprop engines. It also aligns with Sheng et al. [4], who emphasized the diagnostic relevance of these parameters in real-time simulation of failure progression.

The confusion matrix (Fig. 2) further illustrates the model's robustness, showing a low number of misclassifications. This balance between sensitivity and specificity indicates that the Random Forest model generalizes well across the synthetic data distribution, even when variability and noise were introduced to emulate real-world conditions.

Moreover, the reliability curve (Fig. 4) based on the logistic transformation of the composite stress index exhibits a coherent risk progression: as the stress level increases, the probability of failure rises sharply beyond a critical threshold. This behavior is consistent with the expected nonlinear degradation patterns found in aerospace systems and supports the use of stress-based health indicators in predictive modeling.

Importantly, these results were achieved using only synthetic data. While this represents a limitation, it also demonstrates the potential of simulation-based training in cases where real failure data are scarce or sensitive, as is typical in military applications. The high AUC (0.98) indicates that the model successfully learned discriminative patterns despite the absence of real field data, providing a valid proof of concept.

Another relevant point is the model's computational efficiency. With 100 trees, the Random Forest showed rapid training and inference times, enabling near real-time evaluation. Preliminary tests with 200 and 300 trees yielded marginal improvements (<0.5% in F1-score), reinforcing the selection of 100 trees as an effective balance between performance and processing cost—similar to what Liu et al. [12] observed in industrial settings.

Overall, the model's output reflects its ability to process complex, nonlinear interactions between operational variables and predict failure risks effectively. These results suggest that the method is not only statistically sound but also operationally viable for supporting condition-based maintenance decisions. Future validation with actual PT6A-68C engine data will be essential to confirm its applicability in real scenarios, but the current results already demonstrate clear potential for integration in PHM frameworks.

VII. CONCLUSIONS

This study demonstrates the efficacy of Random Forest machine learning models in predicting failures of PT6A-68C engines based on critical operational parameters. Through synthetic data generation and the formulation of a composite stress index, the model effectively distinguished between failure and non-failure scenarios, achieving high performance metrics. The identification of ITT and vibrations as key predictors underscores their critical role in engine reliability.

While the results are promising, future work should focus on validating the model with real-world operational data to enhance its applicability and reliability in practical settings.

This study demonstrates the feasibility of using machine learning, specifically Random Forest algorithms, to predict failures in turboprop engines based on critical operational parameters. The high performance of the model, reflected in metrics such as accuracy, recall, and the AUC of 0.98, underscores its potential for improving safety and reducing maintenance costs in the aviation industry. The identification of key factors like ITT and vibrations as major contributors to failure risk further supports the model's reliability. While synthetic data enabled the development of the model, future integration with real-world data is essential to fully validate its effectiveness in operational contexts. This approach highlights a promising pathway for enhancing predictive maintenance strategies and operational efficiency in modern aviation.

The implementation of a predictive maintenance system allows for accurate anticipation of potential failures, significantly reducing unplanned downtime and optimizing the use of technical and logistical resources. In a military environment, where operational readiness and asset availability are paramount, this predictive capability translates into a substantial operational advantage.

Adopting predictive maintenance strategies based on machine learning improves the operational efficiency of Super Tucano aircraft and strengthens the strategic readiness and capabilities of air forces, enabling them to fulfill critical missions with greater reliability and lower costs. The integration of these technologies should be considered a priority in the modernization of military fleets.

Quantifying the economic and logistical benefits of predictive maintenance would enhance the practical value of the study. For instance, calculating the reduction in unplanned maintenance costs and increased aircraft availability provides tangible metrics to justify the adoption of such systems. Predictive maintenance can significantly reduce the total cost of ownership by minimizing unscheduled repairs and extending the life of critical components. Furthermore, improving aircraft readiness directly contributes to mission success, particularly in high-stakes military operations.

REFERENCES

[1] F. A. Adryan and K. W. Sastra, "Predictive Maintenance for Aircraft Engine Using Machine Learning: Trends and Challenges," International Journal of Aviation Science and Engineering, vol. 3, no. 1, pp. 37–44, 2021, doi: 10.47355/AVIA.V3I1.45.

[2] S. Fu and N. P. Avdelidis, “Prognostic and Health Management of Critical Aircraft Systems and Components: An Overview,” Sensors, vol. 23, no. 19, p. 8124, 2023, doi: 10.3390/s23198124.

[3] Pratt & Whitney Canada, "PT6A Turboprop Engine." [Online]. Available: https://www.pwc.ca/en/products-and-services/pt6a-turboprop-engine.html

[4] H. Sheng, T. Zhang, and Y. Zhang, “Real-time Simulation of Turboprop Engine Control System,” International Journal of Turbo & Jet-Engines, 2016, doi: 10.1515/tjj-2015-0066.

[5] H. Kagermann, W. Wahlster, and J. Helbig, Securing the Future of German Manufacturing Industry: Recommendations for Implementing the Strategic Initiative Industrie 4.0. Final Report of the Industrie 4.0 Working Group. Acatech—National Academy of Science and Engineering, 2013, p. 15.

[6] Comando da Aeronáutica, “Doutrina de Logística da Aeronáutica,” DCA 2-1, 2022.

[7] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[8] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.

[9] M. H. Abidi, M. K. Mohammed, and H. Alkhalefah, "Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing," Sustainability, vol. 14, no. 6, p. 3387, 2022, doi: 10.3390/su14063387.

[10] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[11] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[12] J. Liu, C. Ulishney, and C. E. Dumitrescu, "Random forest machine learning model for predicting combustion feedback information of a natural gas spark ignition engine," Journal of Energy Resources Technology, vol. 143, no. 1, p. 012301, 2021, doi: 10.1115/1.4047761.

[13] A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage Propagation Modeling for Aircraft Engine Prognostics,” 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 2008, pp. 1-9. doi: 10.1109/PHM.2008.4711414.

[14] European Union Aviation Safety Agency, “Type Certificate Data Sheet No. EASA IM.E.038 – PT6A-68 series engines,” Issue 01, Apr. 20, 2016.

[15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.

[16] D. C. Montgomery, Design and Analysis of Experiments, 10th ed. John Wiley & Sons, 2017. [Online]. Available: https://www.wiley.com/en-us/Design+and+Analysis+of+Experiments%2C+10th+Edition-p-9781119182134.

[17] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.