Predicting the Remaining Useful Life (RUL) of a Turbofan Jet Engine

Tareq Alaa Alsahaf
Maintenance and Reliability, Industrial Systems Engineering
King Fahd University of Petroleum and Minerals
Dhahran - Saudi Arabia
g202036800@kfupm.edu.sa

I. Literature Review & Critical Analysis

Remaining Useful Life (RUL) prediction of turbofan jet engines is a core task in Prognostics & Health Management (PHM). Accurate RUL estimation enables optimized maintenance scheduling, reduces unexpected failures, and improves reliability. In recent years, the NASA C-MAPSS simulation dataset has become the standard benchmark for data-driven RUL research due to its multi-sensor, multi-condition degradation trajectories. This survey briefly reviews several trusted and highly cited approaches that align closely with the methodology used in this project. The focus is on simple, practical comparisons of machine-learning and heuristic-based strategies, highlighting their advantages, disadvantages, and performance under common criteria such as data source, scalability, benchmarking, and real-time suitability.

❖ ***LSTM-Based RUL Prediction***

Ellefsen et al. (2019) applied Long Short-Term Memory (LSTM) networks to model non-linear temporal degradation patterns in turbofan engines using the C-MAPSS dataset. LSTM demonstrated strong baseline accuracy due to its ability to capture long-term dependencies in sensor sequences.

➢ ***Advantages:***

1. Learns temporal degradation effectively.
2. Outperforms traditional ML (e.g., SVR, Random Forest).
3. Widely accepted benchmark.

➢ ***Disadvantages:***

1. Weak scalability on multi-condition datasets (FD002–FD004).
2. High training time.
3. No real-time evaluation.

❖ ***CNN + Attention Networks***

Muneer et al. (2021) proposed a 1-D CNN with an attention mechanism to prioritize critical time steps in sensor data. This method reduces training cost compared to LSTM while achieving competitive accuracy on FD001.

➢ ***Advantages:***

1. Faster training and simpler structure.
2. Improved interpretability via attention.
3. Good accuracy on single-condition data.

➢ ***Disadvantages:***

1. Performance decreases in multi-regime datasets.
2. Still fully dependent on C-MAPSS simulation.
3. Real-time execution not addressed.

❖ ***Hybrid CNN-LSTM-Attention Models***

Deng et al. (2024) introduced a hybrid architecture combining CNN, LSTM, and attention layers. This framework showed improved robustness across all four C-MAPSS subsets, making it stronger under varied environmental conditions.

➢ ***Advantages:***

4. High accuracy across FD001–FD004.
5. Captures local features + long-term trends.
6. Strong scalability to noise and varying conditions.

➢ ***Disadvantages:***

4. Increased computational load.
5. Complex architecture; difficult to deploy.
6. No real-time performance study.

❖ ***Domain-Adaptation-Based RUL Prediction***

Nejjar et al. (2023) employed domain adaptation on the N-CMAPSS dataset to address realistic variations in engine operating regimes. Their approach aligns training and testing distributions using adversarial learning.

➢ ***Advantages:***

7. Best scalability under domain shift.
8. More realistic multi-flight-profile modeling.
9. Strong robustness to varying environments.

➢ ***Disadvantages:***

7. Complex and computationally heavy.
8. Still based on simulated data.
9. Not suitable for real-time embedded systems.

A. Comparative Evaluation

Method	Data Source	Scalability	Benchmarking	Real-Time Suitability
LSTM	C-MAPSS	Weak for multi-condition	Compared with ML baselines	Low
CNN + Attention	C-MAPSS	Moderate	Compared with CNN/LSTM	Moderate
CNN-LSTM-Attention	C-MAPSS	Strong across FD001–FD004	Ablation studies	Low
Domain Adaptation	N-CMAPSS	Excellent	Compared with DA methods	Low

Across all studies, deep learning remains the dominant approach for turbofan RUL prediction. LSTM models provide simple and strong baselines but lack robustness under varying operational conditions. CNN-based attention models are faster and more interpretable but limited to simpler datasets. Hybrid CNN-LSTM-attention networks deliver the best accuracy and scalability, though at the cost of higher computational complexity. Domain-adaptation methods offer the best generalization performance but require sophisticated training procedures and remain unsuitable for real-time deployment. A common limitation across all works is their dependence on simulated C-MAPSS or N-CMAPSS data, with very limited validation on real-world engine fleets.

B. Industry Practices and Practical Approaches to Jet-Engine RUL Prediction

In the aviation industry, the estimation of Remaining Useful Life (RUL) for turbofan engines is approached very differently from academic studies that rely heavily on the NASA C-MAPSS dataset. Engine manufacturers such as GE Aerospace, Rolls-Royce, Pratt & Whitney, and Safran deploy large-scale Health & Usage Monitoring Systems (HUMS) that collect real-time operational data—including exhaust gas temperature (EGT) margin, fuel flow, vibration signatures, oil debris analysis, and performance deterioration trends. Instead of depending on a single deep-learning model, industrial RUL estimation relies on a combination of physics-based degradation models, engineering domain knowledge, and fleet-wide statistical analytics. For example, Rolls-Royce's *Power-by-the-Hour* model uses thermodynamic performance margin tracking (especially EGT margin decay) combined with historical maintenance records to evaluate when components such as compressor blades, turbine blades, bearings, or combustor liners will reach their end of life. GE's Prognostics Engine Health Management (EHM) systems use a blended approach of simplified physics models (e.g., compressor efficiency decay, turbine creep life estimation) and anomaly detection algorithms fed from vibration and performance data logged during each flight cycle.

A key difference from academic methods is that industry rarely performs direct RUL regression. Instead, they focus on condition indicators (CIs)—such as corrected core speed, pressure ratios, EGT margin, or vibration amplitudes—and track their rate of change over time. These CIs feed into physics-informed rules or statistical thresholds that trigger maintenance actions. In high-risk components like turbine blades, companies rely on creep and fatigue damage accumulation models, often based on finite-element thermal analysis and stress modeling, rather than deep learning. Machine learning is used, but typically in supportive roles such as anomaly detection, clustering of abnormal behaviors, sensor fusion, or trend extrapolation, rather than full end-to-end RUL prediction. Moreover, industry must ensure certification, safety, and explainability, which makes black-box deep networks difficult to deploy without extensive validation. Thus, interpretability, reliability, and regulatory compliance often outweigh raw accuracy. In summary, while academia explores increasingly complex deep learning architectures for RUL, real aviation practice favors hybrid frameworks that combine physics models, engineering knowledge, and conservative statistical monitoring to ensure safety and regulatory acceptance.

C. Limitations of Current Research vs. Real-World Needs

Although academic research in RUL prediction—particularly using the NASA C-MAPSS dataset—has advanced rapidly through deep neural networks, significant gaps remain when these models are compared with the requirements of real aviation operations. First, the vast majority of published work relies solely on simulated datasets, meaning the learned models do not capture the full complexity of real turbine engines, such as sensor drift, maintenance actions, variable fuel quality, environmental effects, or operational irregularities. Real engines operate under diverse mission profiles, and their degradation patterns seldom follow the clean trends seen in C-MAPSS. This creates a major generalization gap, making most academic models unsuitable for direct deployment in safety-critical environments.

Second, RUL models in the literature often pursue high prediction accuracy, but aviation industry standards demand interpretability, traceability, and certification. Regulatory bodies (FAA, EASA) require maintenance-related decision-making tools to provide transparent reasoning and to adhere to strict safety margins. However, deep learning architectures such as LSTM, CNN-LSTM, and transformer-based models remain largely black-box systems, making them difficult to certify. Industry therefore prefers hybrid approaches—combining physics-based degradation modeling with statistical indicators—because these methods align better with engineering intuition and allow engineers to justify decisions such as early removals or on-condition maintenance.

A third limitation is that academic studies rarely consider real-time constraints. Engine health monitoring systems used by Rolls-Royce, GE Aerospace, and Pratt & Whitney require fast inference, low power consumption, and reliable operation on embedded avionics or ground-based decision platforms. In contrast, many research models require large GPUs, long training durations, and heavy preprocessing pipelines that cannot be integrated into operational maintenance systems. Furthermore, academic work often overlooks fleet-level analytics, historical maintenance logs, and operational metadata—resources heavily used by industry. OEMs track parameters such as exhaust gas temperature (EGT) margin decay, core speed deterioration, vibration anomalies, and oil debris signatures over thousands of engines worldwide, whereas academic models usually analyze only isolated sensor sequences.

In summary, while current research pushes the boundaries of algorithmic performance, it does not yet address the broader ecosystem of aviation maintenance: regulatory compliance, real-world noise, mixed data sources, interpretability, and scalable deployment across global fleets. Bridging these gaps requires not only improved machine learning techniques but also the integration of physics knowledge, domain expertise, and system-level engineering considerations used by industry leaders.

References

[1] T. Ellefsen, M. B. Dwyer, K. S. Kahl, and D. M. Nicholson, “Remaining useful life predictions for turbofan engines using LSTM networks,” Reliability Engineering & System Safety, vol. 183, pp. 248–264, 2019. doi: 10.1016/j.ress.2018.11.011.

[2] M. Muneer, A. Athar, M. U. Ghani, and M. A. Jaffar, “Data-driven deep learning-based attention mechanism for remaining useful life prediction: Case study application to turbofan engine analysis,” Electronics, vol. 10, no. 20, p. 2453, 2021. doi: 10.3390/electronics10202453.

[3] Z. Deng, X. Yan, S. Li, and Z. Zhao, “Prediction of remaining useful life of aero-engines based on CNN–LSTM–attention,” Discover Mechanical Engineering, vol. 2, no. 1, 2024. doi: 10.1007/s44196-024-00639-w.

[4] A. Nejjar, B. Chebel-Morello, and N. Zerhouni,

“Domain adaptation via alignment of operation profile for remaining useful lifetime prediction,” arXiv preprint arXiv:2302.01704, 2023.

[5] M. Arias-Chao, C. Kulkarni, K. Goebel, and O. Fink, “Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics,” Data, vol. 6, no. 1, 2021. doi: 10.3390/data6010005.

[6] M. Salinas-Camus, J. Chacón-Rivas, and A. Pérez, “Uncertainty in aircraft turbofan engine prognostics on the C-MAPSS dataset,” Proceedings of the PHM Society Conference, vol. 16, no. 1, 2024.

[7] A. D. Kwakye, “Platform health management for aircraft maintenance,” Proc. Inst. Mech. Eng. Part G: J. Aerospace Engineering, 2024. doi: 10.1177/09544100231219736.

[8] M. A. Chao, O. Fink, and K. Goebel, “Fusing physics-based and deep learning models for prognostics of complex systems,” Reliability Engineering & System Safety, vol. 217, 2022.

doi: 10.1016/j.ress.2021.107722.

[9] I. de Pater, P. van Houtum, and G. J. van Houtum, “Alarm-based predictive maintenance scheduling for aircraft engines with imperfect remaining useful life prognostics,” Reliability Engineering & System Safety, vol. 222, 2022. doi: 10.1016/j.ress.2022.108341.

[10] C. Peng, X. Liu, and Z. Hu, “Remaining useful life prognosis of turbofan engines based on the C-MAPSS dataset,” Scientific Reports, vol. 12, Article 7072, 2022. doi: 10.1038/s41598-022-10191-2.

[11] Rolls-Royce plc, “Intelligent Engine Health Monitoring (EHM),” 2019. Available: https://www.rolls-royce.com/media/ourstories/discover/2019/intelligent-engine-health-monitoring.aspx

[12] K. Hemmerdinger, “The power of engine health information,” Aviation International News (AINonline), Apr. 26, 2024. Available: https://www.ainonline.com/aviation-news/business-aviation/2024-04-26/power-engine-health-information

[13] GE Aerospace, “Health and Usage Monitoring System (HUMS),”2023. Available: https://connectedaircraftsupport.geaviation.com/support/solutions/articles/43000537206-health-and-usage-monitoring-system-hums-

[14] Rolls-Royce,“Engine Health Management,” Ingenia, Issue 39, June 2009. Available: https://www.ingenia.org.uk/articles/engine-health-managment/

[15] Gabelli Funds, “Power-by-the-Hour: The Rolls-Royce business model,” 2019. Available: https://gabelli.com/research/airplanes-power-by-the-hour/