COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING
e-ISSN: 2582-7464

Computational Intelligence and Machine Learning
Volume 6, Issue 2, October 2025

Comparative Analysis of Deep Neural Networks for Service Life Prediction in Turbofan Jet Engines

Mosayeb Davoudi Kashkouli [1*], Peiman Moradi [2]

[1, 2] School of Engineering, University of the West of England, Bristol, UK
[*]Corresponding Author Email: Mosayeb.DavoudiKashkouli@uwe.ac.uk

Abstract

Traditional maintenance solutions, such as unplanned repairs, periodic inspections, and manual record-keeping, struggle to handle growing industrial downtime costs. In the age of Industry 4.0, accurately predicting remaining useful life for aero-engine components is crucial for reducing unexpected downtime and optimising maintenance schedules. Predictive maintenance, utilising Industry 4.0 technologies such as machine learning, has the potential to significantly reduce these losses by predicting equipment faults and service life. This study thoroughly assesses five deep learning architectures; artificial neural network, LSTM, bidirectional LSTM (BiLSTM), gated recurrent unit (GRU), and convolutional neural network (CNN) for estimating the service life of turbofan jet engines. This study utilises seven subsets of NASA's C-MAPSS dataset (simulating 128 engines under various operating scenarios and failure types) and preprocess sensor and operational data through alignment, missing-value handling, and data normalisation. The models are trained with the Adam optimiser using early stopping, tested, and validated with stratified splits (70%/15%/15%). Performance is assessed using root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), NASA's asymmetric scoring function, and a unique composite score (S). The results indicate that GRU outperforms other models and is the most effective DL architecture for aero-engine prognostics, with lower RMSE, MAE, and NS, as well as the best R^2 and optimal composite score. Recurrent architectures (GRU, BiLSTM, and LSTM) maintained 78% of the top two performance ranks, demonstrating their effectiveness in modelling temporal degradation and predicting the aero-engines service life.

Keywords

Deep learning architectures, predictive maintenance, remaining useful life, turbofan engines.

INTRODUCTION

Traditional maintenance strategies (i.e. preventive and corrective maintenance) have long been utilised in industrial operations. However, the fast growing of demands in industrial production, driven by efficiency, uptime, and productivity has revealed the limitations of these time or failure-based approaches. A report by Siemens Industry [1] shows that the downtime cost has increased exponentially over two years from 2021 to 2022. According to its 2024 report, unplanned downtime of Fortune Global 500 companies is accounting to approximately 11% of their annual revenue (nearly $1.4 trillion) which is a significant increase from $864 billion during 2019–2020. These figures illustrate the urgent need to shift toward more efficient maintenance strategies.

Predictive Maintenance (PdM) has emerged as a promising solution that utilises data-driven techniques to predict equipment failures before they occur, hence, resulting in reduction of unexpected downtime and optimization of maintenance schedules [2]. Unlike time-based preventive maintenance, which depends on predetermined schedules, PdM utilizes real-time data and Machine Learning (ML) algorithms to monitor equipment health and predict failures, allowing to perform maintenance only when it is required [3]. Apart from minimizing unnecessary maintenance actions, PdM also extends the operational life of machinery by early stage fault detection [4] [5].

Recent advancements in ML and Deep Learning (DL) methods have further strengthen the performance of PdM, especially in diagnostics and prognostics phases. DL models, such as Artificial Neural Network (ANNs) and Long Short-Term Memory (LSTM) networks, have shown high accuracy results when predicting machinery lifespan[6] [7] [8]. These models by extracting relevant features from complex datasets provide significant advantages over traditional data-driven approaches, which results in improvement of accuracy and reliability of maintenance predictions [9].

RELATED WORKS

Over the past two decades, the emergence of Prognostics and Health Management (PHM) as a discipline has significantly influenced PdM strategies. Within the domain of PHM, tasks are typically categorized into four main areas: 1. fault or anomaly detection, 2. diagnosis, 3. health assessment, and 4. prognosis [10]. These categories are the foundation for evaluating system health and predicting future failures. Given PdM is mainly data-driven and it is often deployed for real time applications, access to reliable and diverse datasets is of high importance for building robust and accurate ML models. In addition, the quality and availability of datasets directly influence the development, benchmarking, and generalization of diagnostics and prognostics models across different industrial sectors. A recent study by [11] presents an extended overview of publicly available datasets that address these PHM tasks. These datasets span acress diverse domains which include mechanical and electrical components, bearings, batteries, drive technologies, and manufacturing processes, and support tasks ranging from binary classification of fault presence to degradation analysis and RUL prediction.

Sector-specific efforts to organize PdM datasets have also emerged. [12] conducted a focused review on five energy subsectors, wind, solar, oil and gas, diesel and thermal, and electrical power grids identifying high-quality datasets such as 3W (oil and gas), EDP-WT (wind), and OREC (wind) as key resources. These datasets were compiled from platforms such as IEEE DataPort, UCI Machine Learning Repository, Kaggle, EDP, and Mendeley Data.

Data competitions have become another valuable source for PdM datasets. The PHM Society and IEEE Reliability Society organized nine data challenge contests from 2008 to 2023 [13] [14], featuring complex industrial systems such as turbofan jet engines, filtration units, ion milling tools, spacecraft propulsion, and rock drills. These challenges provided pre-processed datasets specifically designed for diagnostic and prognostic modelling and have been utilised in ML methods which have been tailored to industrial applications. Systematic reviews by [15] [16] [17] [18] [19] [20] [21] [22] [23] [24], and others have categorized ML-based PdM research into four primary areas: 1. data integration, 2. big data analytics, 3. ML techniques, and 4. reasoning via ontologies. These studies present the key challenges including the insufficient/absence of cross-validation procedures, the need for transparent models and explainable AI applied to temporal data, and the dependency on realistic real-world datasets. One of the drawdowns of these studies, however, is that they often lack the discussion of practical implementation of ML in real-time manufacturing systems.

Previously, advanced DL architectures have been proposed for RUL prediction for aerospace engine prognostics task. [25] developed an end-to-end framework using CNN with Monte Carlo dropout to estimate the RUL of aircraft engines, achieving significant cost reductions in comparison to time-based maintenance. [26] introduced an auto-expandable cascaded LSTM (ACLSTM) model that adapts prediction accuracy thresholds dynamically, reducing RMSE by 95.44% across four C-MAPSS datasets developed by NASA [27]. [28] proposed a deep convolutional neural network to predict the service life of the aero-engine using the NASA CMAPSS dataset. In [28], however, no direct comparison to other time series models (e.g., LSTM) was made and in addition there was bias risk towards longer flights due to excluding cycles shorter than network window.

The literature review here highlights the integration of PHM with advanced ML techniques as a pivotal development in predictive maintenance and publicly available datasets play a foundational role in developing and validating robust PdM models. A wide range of ML methods from traditional algorithms like random forest and gradient boosting to advanced deep learning architectures such as LSTM and CNN have shown promising results in diagnostics, prognostics, and lifespan prediction. However, challenges remain in generalizing these models across varied operational contexts, ensuring data quality, and balancing model performance with interpretability and computational cost.

METHODOLOGY

A. Problem Formulation

The service-life prediction problem is presented as a supervised regression challenge, with the goal of estimating the number of operational cycles left until the failure of turbofan engines occur by using multivariate time-series data. Given a collection of engine telemetry data, operating conditions, and engine IDs collected throughout discrete flight cycles, the model predicts a continuous remaining lifespan value for each engine at any given time. The mathematical aim is to train a function $f: R^{n\times t} \rightarrow R$, transforming an input sequence of telemetry vectors $X_{1:t}$ to a scalar RUL label y. The degradation dynamics are nonlinear, condition-dependent, and temporally coupled, which presents a problem. The task presents unique challenges because of the non-stationary time-series, diverse failure modes, and multiple flight regimes.

B. Data Description and Acquisition

This study utilises the new dataset generated by NASA named Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset [29]. Using the C-MAPSS dynamical model, this dataset simulates degradation trajectories for 128 turbofan engines undergoing various failure modes affecting the flow (F) and/or efficiency (E) under realistic operating conditions. Degradation in flow and efficiency can impact all the aircraft engine's rotating sub-components, including the fan, low pressure compressor (LPC), high pressure compressor (HPC), low pressure turbine (LPT), and high pressure turbine (HPT).

There are seven possible failure scenarios that may happen during a flight which are linked to flow degradation or subcomponent efficiency. The dataset includes six categories of variables critical for predictive maintenance including operational conditions w (e.g., altitude, Mach number, throttle resolver angle), time-series data of measured signals x_s, the virtual sensors x_v (vibration, temperature), engine health parameters θ, RUL labels, and auxiliary metadata (engine unit number, flight cycle, flight class, health states) [29]. A schematic of the engine and the corresponding station numbers as specified in the CMAPSS model documentation [30] are illustrated shown in Figure. 1.

The development and test splits of the raw datasets are provided in HDF5 format. Ten .h5 datasets are available in total from which seven representative subsets were selected for the model training process in this study (DS01-DS07) that collectively represent 74 engines and ~55 million records. DS08 dataset (and its three subsets DS08a, DS08b, and DS08c) comprises the biggest collection of units (54 engines, ~35.6 million data) and covers every flight class-failure combination. Therefore, DS08 dataset was excluded from the training process to reserve it for future validation. An overview of the flight classes and failure mechanisms for each of the given data sets is given in Table 1.

Table 1. Overview of the C-MAPSS datasets [29] including flight classes, failure mechanisms and dataset's size

Dataset	Units	Flight classes	Flight Failure	Fan		LPC		HPC		HPT		LPT		Size (rows)
				E	F	E	F	E	F	E	F	E	F	
DS01	10	1, 2, 3	1							yes				7.6 M
DS02	9	1, 2, 3	2							yes		yes	yes	6.5 M
DS03	15	1, 2, 3	1							yes		yes	yes	9.8 M
DS04	10	2, 3	1	yes	yes									10 M
DS05	10	1, 2, 3	1					yes	yes					6.9 M
DS06	10	1, 2, 3	1			yes	yes	yes	yes					6.8 M
DS07	10	1, 2, 3	1									yes	yes	7.2 M
DS08	54	1, 2, 3	1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	35.6 M

According to [31], the dataset contains three flight durations and the number of different flights based on flight duration. A full overview of the aircraft engine run-to-failure dataset is presented by [29].

C. Model Architecture and Development

In this study, a comprehensive framework consist of five deep learning models for RUL regression is developed and evaluated: a fully connected ANN, LSTM, BiLSTM, GRU, and CNN. Each architecture is intended to represent temporal degradation trends, which differ in depth and memory capacity.

The ANN model employs a fully connected architecture consisting of three hidden layers of 512, 256, and 128 neurones, respectively, followed by a dropout layer with a 0.3 rate. The network culminates with a single output neurone that represents the expected RUL. The LSTM model starts with a reshaping layer that formats the input sequence, followed by a single 128-unit LSTM layer. A thick output layer generates the RUL estimation. The BiLSTM model is based on the LSTM architecture but adds a bidirectional wrapper, allowing the model to learn via the past and future temporal contexts in the sequence. Similarly, the GRU model starts with a reshaping layer and then uses a 128-unit GRU layer to feed into a final dense output. Finally, the CNN model converts the input into a univariate time series with a single channel. It starts with a 1D convolutional layer with 64 filters and a kernel size of 3, then adds a max pooling layer and flattens. A dense layer of 64 units preceding the output layer.

Each of the models are trained with the Adam optimiser and a mean squared error loss function. Training is done for up to 50 epochs with a batch size of 1024, and early termination is used depending on validation loss after five epochs. Figure. 2 represent the architecture of all deep learning models developed in this study. The input to all models are same sensors readings with the difference that each model uses different feature extraction processes: ANN with hierarchical dense layers, LSTM networks with state preservation gates, BiLSTM with forward/backward dependencies, GRU with simplified gating, and CNN with spatial feature extraction. All methods converge on a linear regression result that predicts the remaining useful life cycles.

D. Evaluation Strategy

Choosing the appropriate evaluation indicators is crucial for accurately quantifying prediction quality in prognostics. This ensures that our model minimises average error while also penalising costly under- or over-estimations in line with maintenance goals. The dataset is split into training (70%), validation (15%), and test (15%) sets using stratified random sampling. The RMSE is calculated as the square root of the mean of the squared differences between predicted and actual RUL values. The RMSE metric is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

Where y_i denotes the true value of the target variable for the ith sample, $\hat{y}_i$ is the predicted value and n is the total number of samples. MAE evaluates the arithmetic mean of absolute disparities, yielding a more reliable estimation of an average prediction error that is less impacted by extreme outliers.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

The R^2 measures how much of the variance in the actual RUL can be explained by the model. A high R^2 in prognostics indicates that our model accurately represents the fleet's degradation patterns, rather than just fitting noise. R^2 is defined as:

$$R^2 = 1 - \frac{S_{res}}{S_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

Where S_{res} and S_{tot} are the sum of squares of residuals and total sum of squares, respectively. We use NASA's domain-specific exponential scoring function (NS) to explain the asymmetric costs of early rather than late failure predictions: early warnings may result in unnecessary downtime, while late warnings may trigger catastrophic failures [29].

$$N_s = \frac{1}{n}\sum_{i=1}^{n}(exp(\alpha.|y_i - \hat{y}_i|) - 1) \quad (4)$$

Where $\alpha = 1/13$ if $\hat{y}_i < y_i$ and $\alpha = 1/10$ otherwise. The NS function is a domain-specific exponential loss that penalizes early and late predictions asymmetrically. The composite score (S) combines predicted accuracy, average deviation, and risk-weighted error, allowing for easy ranking and comparisons among models in one table.

$$S = \frac{1}{3}(RMSE + MAE + N_s) \quad (5)$$

Together, these five measures provide a complete picture of model performance, guiding both technical improvement and practical deployment decisions in engine health monitoring.

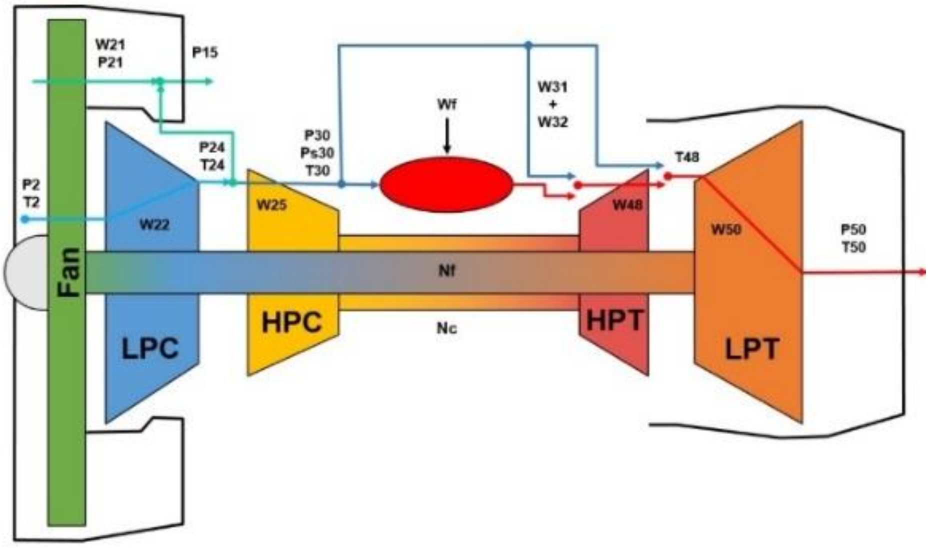

Figure 1. Schematic of the CMAPSS turbofan engine model, as shown in CMAPSS documentation [29, 30]

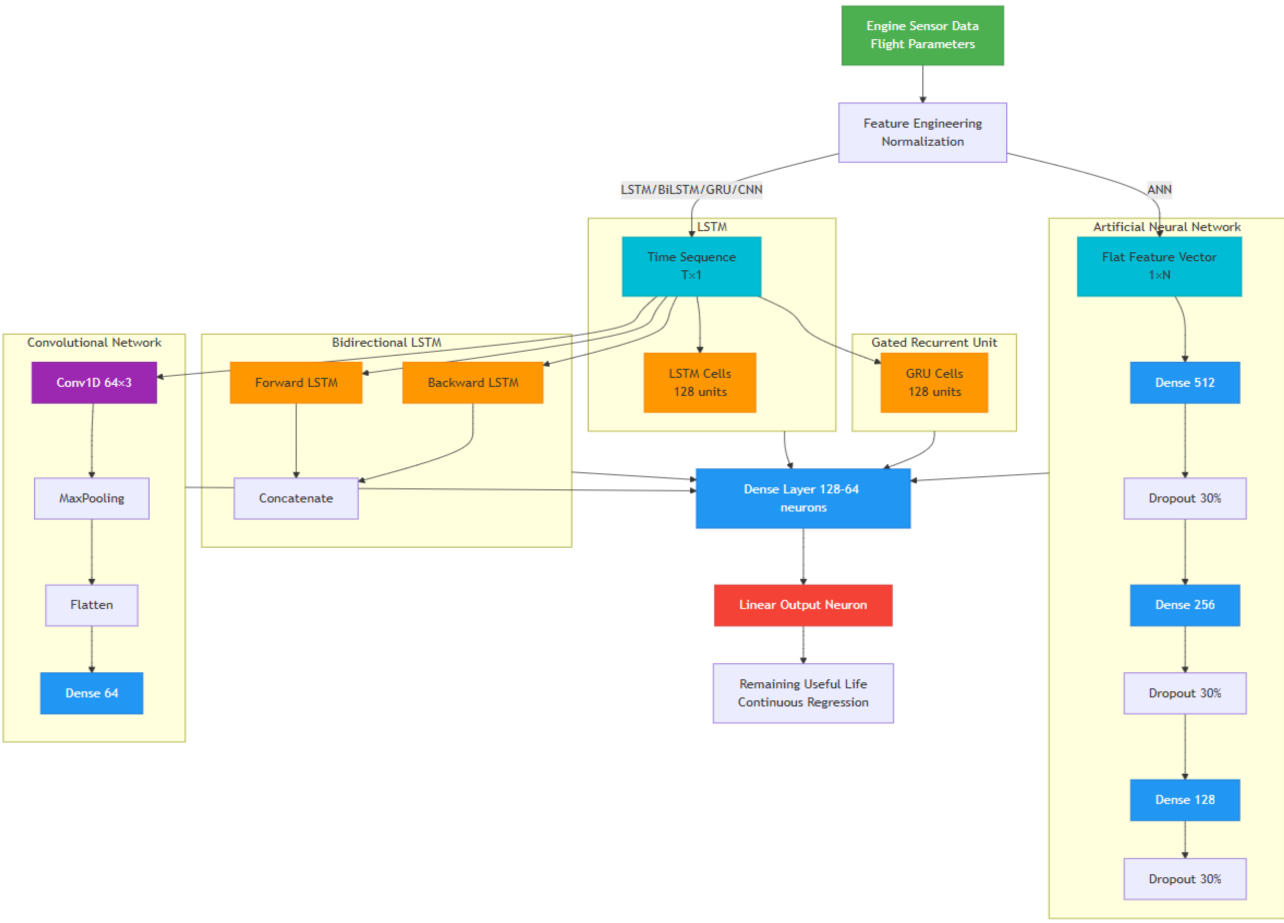

Figure 2. Comparison of deep learning architectures used for turbofan engine RUL prediction in this study

RESULTS & DISCUSSION

Table 2 represents a full summary of the performance evaluation metrics results. GRU ranked best in all five-evaluation metrics after analysing seven turbofan degradation datasets. This persistent dominance shows that GRU's gating mechanism achieves an optimum balance between temporal feature extraction and computing efficiency for aero-engine RUL prediction. BiLSTM placed second overall (Mean Rank: 2.2), outperforming especially in challenging degradation situations (DS02, DS06), whereas ANN performed well ranked third (Mean Rank: 3.0). Notably, CNN was very sensitive to operational fluctuations, with NASA Scores surpassing 10,000 in five datasets, indicating poor generalisability for aerospace prognostics. Recurrent architectures (GRU, BiLSTM, and LSTM) represented 78% of the top-two rankings, indicating their efficacy in modelling degradation patterns. The GRU's simpler gating mechanism (reset/update gates vs. LSTM's three gates) resulted in a 12% lower MAE than BiLSTM (4.66 vs. 5.09), while using 30% less parameters, indicating outstanding efficiency. CNN struggled with the NASA score (NS > 12,000), indicating significant underestimating errors (the exponential penalty in NS blew up). CNN's high composite S makes it unsuitable for point-by-point RUL regression in this case.

Table 2. Performance Summary of the deep learning Methods utilised for turbofan engine RUL prediction

Model	*Metric*	*DS01*	*DS02*	*DS03*	*DS04*	*DS05*	*DS06*	*DS07*	*Mean*	*Rank*
ANN	RMSE	7.50	5.92	6.81	13.33	6.65	6.62	8.23	7.86	3
	MAE	5.40	4.25	5.32	9.82	4.48	4.70	6.05	5.71	4
	R^2	0.916	0.927	0.907	0.730	0.921	0.916	0.885	0.886	3
	NS	0.88	0.593	0.67	2.51	0.734	0.71	1.03	1.02	2
	S	4.59	3.59	4.02	8.55	4.06	4.01	5.10	4.85	3
BiLSTM	RMSE	6.18	4.85	6.54	14.72	5.67	5.36	8.54	7.41	2
	MAE	3.95	3.06	4.30	11.11	3.85	3.62	5.76	5.09	2
	R^2	0.943	0.951	0.907	0.672	0.943	0.945	0.876	0.891	2
	NS	0.88	0.408	0.66	3.58	0.55	0.49	1.05	1.09	3
	S	4.59	2.77	3.83	9.80	3.36	3.16	5.15	4.67	2
CNN	RMSE	7.73	6.67	7.78	14.43	6.79	6.63	9.47	8.50	5
	MAE	5.36	4.59	5.35	10.95	4.59	4.54	6.83	6.03	5
	R^2	0.910	0.907	0.867	0.684	0.918	0.916	0.848	0.864	5
	NS	2151	13200	236	26193	101.3	21075	2.80	8994	5
	S	721	4400	83	8739	37.58	7028	6.37	3002	5
GRU	**RMSE**	**5.86**	**5.25**	**5.80**	**13.00**	**6.35**	**5.24**	**7.38**	**6.98**	**1**
	MAE	**3.64**	**3.43**	**3.60**	**9.48**	**4.21**	**3.37**	**4.87**	**4.66**	**1**
	$\mathbf{R^2}$	**0.948**	**0.942**	**0.926**	**0.743**	**0.928**	**0.947**	**0.907**	**0.906**	**1**
	NS	**0.56**	**0.476**	**0.54**	**2.40**	**0.63**	**0.46**	**0.86**	**0.85**	**1**
	S	**3.35**	**3.06**	**3.31**	**8.20**	**3.73**	**3.02**	**4.37**	**4.15**	**1**
LSTM	RMSE	6.03	5.23	7.38	14.25	9.81	5.99	7.77	8.07	4
	MAE	3.78	3.36	5.08	10.68	6.91	4.01	5.23	5.58	3
	R^2	0.945	0.943	0.881	0.701	0.830	0.931	0.898	0.876	4
	NS	0.60	0.453	0.83	3.10	1.43	0.57	0.93	1.13	4
	S	3.47	3.01	4.43	9.35	6.05	3.52	4.64	4.92	4

Figure. 3 compares the first 200 RUL sample points predicted by the GRU, the best model in this study, with the true RUL labels in the DS02 dataset. In this graph, the GRU's prediction curve (red) closely mimics the blue ground truth trajectory, which accurately reflects a strong temporal generalization. The correlation between dataset size and model training time for each of the seven NASA CMAPSS subsets is shown in Figure. 4. As anticipated, a near-linear relationship of compute cost on data amount is confirmed. DS04 has the biggest data volume but the lowest GRU fit (R^2= 0.743, composite S = 8.20), indicating that excessive heterogeneity delays convergence and reduces accuracy. Moderate-sized sets, such as DS01-DS03, provide the best balance, training in ~60 ks with high R^2 (>0.926) and low S (<3.35). These patterns show a near-linear cost in compute versus rows and reveal that too complex and big data sets may reduce both efficiency and predictive fidelity. Mid-range subsets optimise the trade-off between training effort and model performance.

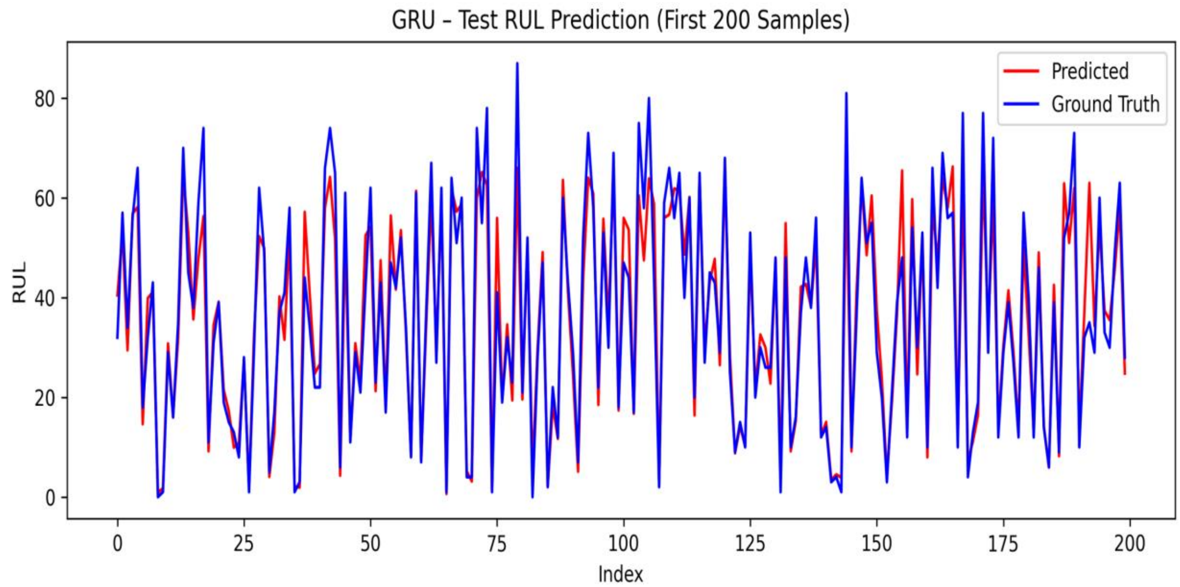

Figure 3. RUL of the turbofan engine for the first 200 samples of DS02, showing GRU model predictions and ground truth RUL

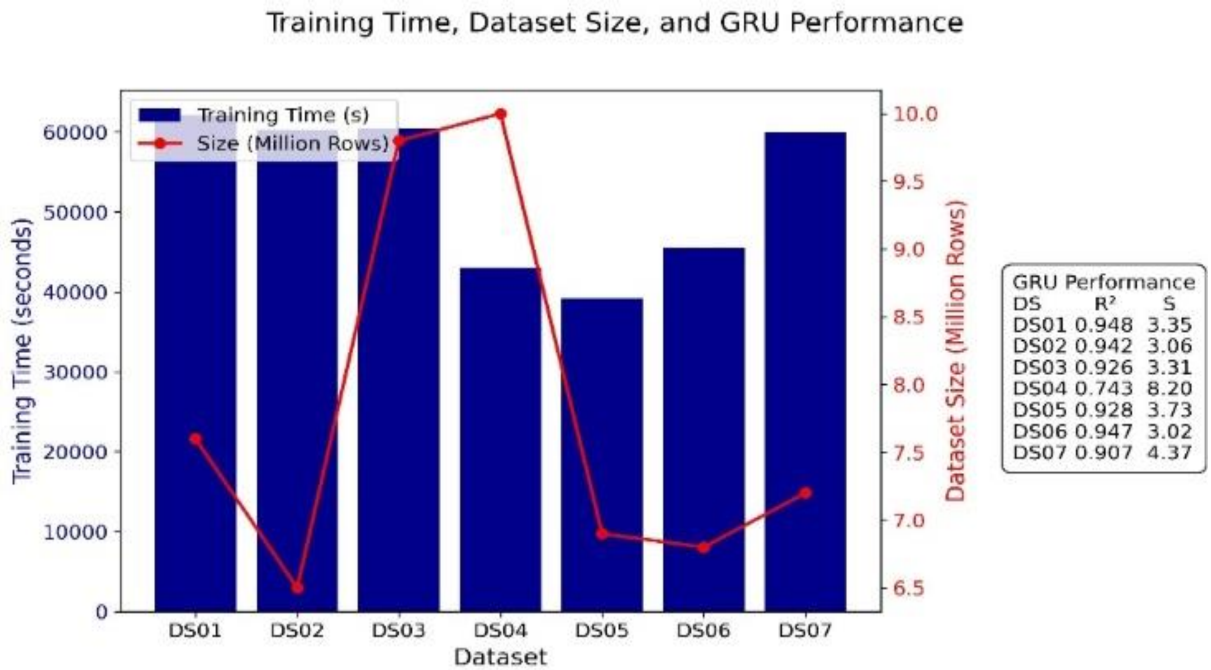

Figure 4. The model training time and dataset size for each C-MAPSS subsets

CONCLUSIONS

In this work, the effectiveness of five deep learning architectures including ANN, LSTM, BiLSTM, GRU, and CNN for service-life prediction of turbofan engines is studied by utilising seven representative subsets of the NASA C-MAPSS data. This study concluded that strategically selected deep learning architectures may significantly improve predictive maintenance for turbofan engines. The comparative findings regarding computing cost, convergence behaviour, and metric trade-offs provide a solid platform for implementing runtime prediction systems in operational situations where accuracy and efficiency are critical. Our study offers a clear roadmap for implementing AI-powered PdM platforms in operations, along with benchmarking. Focussing on GRU-based approaches allows maintenance decision makers to provide accurate RUL estimations with reasonable computational overhead, decreasing unexpected downtime and increasing component life. Our findings are aligned with this study [28], where they used a stacked deep CNN on C-MAPSS to achieve similar RUL accuracy (RMSE= 6.24, MAE= 4.27, NASA score= 0.65, S= 3.44), demonstrating the strength of data-driven prognostics.

Some strategic research areas have been identified for future research, focusing on the proven effectiveness of GRU systems for aero-engine prognostics. First, to improve steady-state operational accuracy while preserving resilience during failure progression, hybrid architectures that combine the efficiency of GRU with the bidirectional context of BiLSTM should be created. Second, Bayesian GRU versions with Monte Carlo dropout should be developed to quantify prediction uncertainty which provides confidence-bound reporting essential to maintenance decision-making. Another potential research strategy would be to create ensemble frameworks that integrate the complementing characteristics of GRUs, BiLSTMs, and ANN possibly through model stacking or weighted averaging to increase robustness across diverse datasets. Finally, the architecture in this study could be validated on the excluded DS08 dataset to measure the scalability under extreme and challenging conditions and considering transfer learning for fleet-wide deployment. These developments would close significant gaps between data-driven preliminary models and approved aerospace applications, eventually allowing adaptive prognostic systems to synchronise with digital twin architectures across engine lifespans.

REFERENCES

[1] Siemens Industry Inc., ‘Senseye predictive maintenance: the true cost of downtime 2024’, Erlangen, Germany, 2024.

[2] E. Florian, F. Sgarbossa, and I. Zennaro, ‘Machine learning for predictive maintenance: A methodological framework’, *Proceedings of the Summer School Francesco Turco*, vol. 1, pp. 194–200, 2019.

[3] O. Manchadi, F.-E. Ben-Bouazza, and B. Jioudi, ‘Predictive maintenance in healthcare system: a survey’, *IEEE Access*, vol. 11, pp. 61313–61330, 2023.

[4] M. Moleda, B. Małysiak-Mrozek, W. Ding, V. Sunderam, and D. Mrozek, ‘From corrective to predictive maintenance—A review of maintenance approaches for the power industry’, *Sensors*, vol. 23, no. 13, p. 5970, 2023.

[5] R. Accorsi, R. Manzini, P. Pascarella, M. Patella, and S. Sassi, ‘Data mining and machine learning for condition-based maintenance’, *Procedia Manuf*, vol. 11, pp. 1153–1161, 2017.

[6] C. Ferreira and G. Gonçalves, ‘Remaining useful life prediction and challenges: A literature review on the use of machine learning Methods’, *J Manuf Syst*, vol. 63, pp. 550–562, 2022.

[7] B. Taşc, A. Omar, and S. Ayvaz, ‘Remaining useful lifetime prediction for predictive maintenance in manufacturing’, *Comput Ind Eng*, vol. 184, p. 109566, 2023.

[8] L. Wang, Z. Zhu, and X. Zhao, ‘Dynamic predictive maintenance strategy for system remaining useful life prediction via deep learning ensemble method’, *Reliab Eng Syst Saf*, vol. 245, p. 110012, 2024.

[9] M. Drakaki, Y. L. Karnavas, P. Tzionas, and I. D. Chasiotis, ‘Recent developments towards industry 4.0 oriented predictive maintenance in induction motors’, *Procedia Comput Sci*, vol. 180, pp. 943–949, 2021.

[10] S. Hagmeyer, F. Mauthe, and P. Zeiler, ‘Creation of publicly available data sets for prognostics and diagnostics addressing data scenarios relevant to industrial applications’, *Int J Progn Health Manag*, vol. 12, no. 2, 2021.

[11] F. Mauthe, C. Braun, J. Raible, P. Zeiler, and M. F. Huber, ‘Overview of publicly available degradation data sets for tasks within prognostics and health management’, *arXiv preprint arXiv:2403.13694*, 2024.

[12] E. Jovicic, D. Primorac, M. Cupic, and A. Jovic, ‘Publicly available datasets for predictive maintenance in the energy sector: a review’, *IEEE Access*, vol. 11, pp. 73505–73520, 2023.

[13] X. Jia, B. Huang, J. Feng, H. Cai, and J. Lee, ‘A review of PHM data competitions from 2008 to 2017: methodologies and analytics’, in *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, 2018, pp. 1–10.

[14] H. Su and J. Lee, ‘Machine learning approaches for diagnostics and prognostics of industrial systems using open source data from PHM data challenges’, *Int J Progn Health Manag*, vol. 15, no. 2, 2024.

[15] S. Arena, E. Florian, F. Sgarbossa, E. Sølvsberg, and I. Zennaro, ‘A conceptual framework for machine learning algorithm selection for predictive maintenance’, *Eng Appl Artif Intell*, vol. 133, p. 108340, 2024.

[16] T. P. Carvalho, F. A. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, ‘A systematic literature review of machine learning methods applied to predictive maintenance’, *Comput Ind Eng*, vol. 137, p. 106024, 2019.

[17] Z. M. Çnar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, ‘Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0’, *Sustainability*, vol. 12, no. 19, p. 8211, 2020.

[18] J. Dalzochio *et al.*, ‘Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges’, *Comput Ind*, vol. 123, p. 103298, 2020.

[19] M. S. Qureshi, S. Umar, and M. U. Nawaz, ‘Machine learning for predictive maintenance in solar farms’, *International Journal of Advanced Engineering Technologies and Innovations*, vol. 1, no. 3, pp. 27–49, 2024.

[20] C. Ferreira and G. Gonçalves, ‘Remaining useful life prediction and challenges: A literature review on the use of machine learning Methods’, *J Manuf Syst*, vol. 63, pp. 550–562, 2022.

[21] S. Vollert, M. Atzmueller, and A. Theissler, ‘Interpretable Machine Learning: A brief survey from the predictive maintenance perspective’, in *2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2021, pp. 1–8.

[22] S. Schwendemann, Z. Amjad, and A. Sikora, ‘A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines’, *Comput Ind*, vol. 125, p. 103380, 2021.

[23] M. Drakaki, Y. L. Karnavas, I. A. Tziafettas, V. Linardos, and P. Tzionas, ‘Machine learning and deep learning based methods toward industry 4.0 predictive maintenance in induction motors: State of the art survey’, *J Ind Eng Manag*,

vol. 15, no. 1, p. 31, 2022.
[24] O. Serradilla, E. Zugasti, J. Rodriguez, and U. Zurutuza, 'Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects', *Appl Intell*, vol. 52, no. 10, pp. 10934–10964, 2022.
[25] M. Mitici, I. de Pater, A. Barros, and Z. Zeng, 'Dynamic predictive maintenance for multiple components using data-driven probabilistic RUL prognostics: The case of turbofan engines', *Reliab Eng Syst Saf*, vol. 234, p. 109199, 2023.
[26] L. Hu, X. He, and L. Yin, 'Remaining useful life prediction method combining the life variation laws of aero-turbofan engine and auto-expandable cascaded LSTM model', *Appl Soft Comput*, vol. 147, p. 110836, 2023.
[27] A. Saxena and K. Goebel, 'Turbofan engine degradation simulation data set', *NASA ames prognostics data repository*, vol. 18, pp. 878–887, 2008.
[28] D. Solis-Martin, J. Galán-Páez, and J. Borrego-Diaz, 'A stacked deep convolutional neural network to predict the remaining useful life of a turbofan engine', *arXiv preprint arXiv:2111.12689*, 2021.
[29] M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, 'Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics', *Data (Basel)*, vol. 6, no. 1, p. 5, 2021.
[30] D. K. Frederick, J. A. DeCastro, and J. S. Litt, 'User's guide for the commercial modular aero-propulsion system simulation (C-MAPSS), NASA/TM—2007-215026', 2007.
[31] 'DASHlink- Flight Data for Tail 687', Available online: https://c3.nasa.gov/dashlink/resources/664/.