A Digital Twin-Enabled Deep Learning Framework for Remaining Useful Life Prediction of Turbofan Engines

N. Ismayil Knai[1,*], B. Arun[2], B. V. Manikandan[3],
S. Sheik Abdullah[4], D. Lakshmi[5]

[1,5]Dept. of Electrical and Electronics Engineering, AMET University, Chennai, India

[2]Dept. of Electrical and Electronics Engineering, RVS College of Engineering & Technology, Dindugal, India

[3]Dept. of Electrical and Electronics Engineering, Mepco schlenk Engineering College, Sivakasi, India

[4]School of Electronics, Electrical and Biomedical Technology, Kalasalingam Academy of Research and Education, India

ABSTRACT

Accurate prediction of the Remaining Useful Life (RUL) of industrial machinery is essential for developing intelligent predictive maintenance and digital twin systems. This study proposes a Long Short-Term Memory (LSTM) neural network model to estimate the RUL of turbofan engines by analyzing multivariate time-series sensor data obtained from the NASA Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset. The model was designed to capture temporal dependencies within the sensor readings in order to learn the complex patterns of degradation that occur over time. Four datasets, namely FD001, FD002, FD003, and FD004, were examined, and the FD001 dataset was selected as the baseline because it represents a single operational condition with a clearly defined degradation trend. The trained LSTM model achieved a Mean Absolute Error (MAE) of 38.533 and a Root Mean Square Error (RMSE) of 51.069, showing that it can closely follow the actual degradation trajectory with a high degree of accuracy. Correlation analysis identified several key sensors, including sensor_7, sensor_12, sensor_20, and sensor_21, as the most influential variables for predicting RUL. The findings indicate that deep learning models can effectively represent mechanical degradation and can be integrated into digital twin frameworks to enable real-time health monitoring, proactive maintenance scheduling, and data-driven decision-making in industrial environments.

Keywords Remaining Useful Life, Long Short-Term Memory, Predictive Maintenance, Digital Twin, Turbofan Engine

Submitted: 22 June 2025
Accepted: 5 September 2025
Published: 24 May 2026

Corresponding author
N. Ismayil Knai,
ismayilkani@gmail.com.

Additional Information and Declarations can be found on page 177

DOI: 10.47738/ijrm.v3i2.51

INTRODUCTION

The continuous evolution of industrial automation and digital transformation has introduced new paradigms in asset management and maintenance. In modern industries such as aerospace, energy, and manufacturing, equipment operates under highly dynamic conditions, generating large volumes of sensor data in real time. These data contain critical information about system performance, operational stress, and component degradation. Effectively interpreting this data can significantly improve system reliability and safety while reducing maintenance costs [1]. Predictive maintenance, which utilizes data-driven models to anticipate failures before they occur, has become an essential

How to cite this article: N. I. Knai, B. Arun, B. V. Manikandan, S. S. Abdullah and D. D. Lakshmi, "A Digital Twin-Enabled Deep Learning Framework for Remaining Useful Life Prediction of Turbofan Engines," *Int. J. Res. Metav.*, vol. 3, no. 2, pp. 164-179, 2026.

component of Industry 4.0. Unlike corrective maintenance, which reacts to failures, or preventive maintenance, which follows fixed schedules, predictive maintenance aims to optimize maintenance activities based on the actual health condition of the equipment. This approach enables organizations to maximize equipment availability and minimize unexpected breakdowns, supporting both operational efficiency and sustainability.

Accurate prediction of the RUL of machinery is one of the core challenges in predictive maintenance. RUL refers to the estimated time or number of operational cycles remaining before a component or system reaches its failure threshold. Determining RUL requires analyzing complex temporal patterns and nonlinear relationships within multivariate sensor data. Early approaches to RUL prediction employed statistical models such as regression analysis, proportional hazards models, and Kalman filters [2]. Although these methods provided useful insights, they often assumed linear degradation behavior and required strong prior knowledge about the system's physical model. As industrial systems have become more complex, these assumptions have proven insufficient, leading to the adoption of machine learning and deep learning techniques that can learn degradation patterns directly from data without explicit modeling of physical processes.

In recent years, deep learning models have demonstrated remarkable success in addressing the challenges of RUL prediction. Architectures such as Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and LSTM networks have shown superior capability in extracting meaningful features from high-dimensional sensor data [3]. Among these, the LSTM network has received particular attention due to its ability to capture long-term temporal dependencies and mitigate the vanishing gradient problem encountered in traditional Recurrent Neural Networks (RNN). Several studies have reported that LSTM-based models outperform conventional machine learning algorithms such as Support Vector Machines (SVM) and Random Forests when applied to the NASA C-MAPSS dataset [4]. State-of-the-art approaches have extended LSTM architectures with attention mechanisms, bidirectional structures, and hybrid CNN-LSTM combinations to further enhance predictive accuracy and robustness. However, despite these advancements, practical implementation remains limited due to challenges in model interpretability, computational efficiency, and adaptability to real-world industrial environments.

While recent studies have improved prediction accuracy, important research gaps remain in the integration of predictive models into operational decision-making frameworks. Many existing works focus on optimizing numerical performance metrics such as MAE and RMSE, but often overlook the interpretability of the models and their applicability in complex maintenance systems. Additionally, most research on the C-MAPSS dataset concentrates on single-condition subsets, such as FD001, which do not fully represent the diversity of real-world operational conditions. As a result, current models may perform well in controlled environments but fail to generalize across multi-condition and multi-fault scenarios such as those in FD002 and FD004 [5]. Furthermore, there is limited exploration of how deep learning-based RUL predictions can be integrated with emerging digital twin technologies to create dynamic, interactive representations of physical systems that evolve alongside their real-world counterparts. Addressing this gap requires developing models that not only predict degradation accurately but also enhance system transparency and support real-time decision-making.

To address these challenges, this study proposes an LSTM-based framework for predicting the Remaining Useful Life of turbofan engines as part of a digital twin system. The model utilizes multivariate time-series data from the NASA C-MAPSS dataset to capture the temporal dynamics of sensor readings and identify patterns associated with mechanical degradation. The FD001 subset is selected as the baseline for model training and evaluation because it represents a single operating condition with clear degradation behavior, allowing the model to establish a strong foundation before extension to more complex datasets. The research not only focuses on achieving high prediction accuracy but also on understanding sensor relevance through correlation analysis, which provides physical interpretability of the model's outputs. By integrating deep learning-based RUL prediction with digital twin concepts, this study contributes to the development of intelligent maintenance systems capable of simulating degradation processes, visualizing component health in real time, and supporting proactive maintenance scheduling.

Ultimately, the proposed approach aims to bridge the gap between data-driven prognostics and real-world industrial implementation. The integration of LSTM modeling with digital twin technology creates a foundation for adaptive and self-learning maintenance frameworks that can evolve alongside physical assets. Such systems can continuously update predictions as new data become available, enabling a closed-loop feedback mechanism between the digital and physical environments. The outcomes of this research are expected to advance the state of predictive maintenance by providing a scalable and interpretable methodology that supports intelligent decision-making, improves asset reliability, and enhances operational efficiency in complex industrial ecosystems.

Literature Review and Related Works

The prediction of RUL has become a critical focus area in prognostics and health management, as accurate estimation enables timely maintenance interventions and improved operational reliability. Early approaches primarily relied on statistical and physics-based models such as linear regression, Wiener processes, and proportional hazard models [6], [7]. These techniques provided interpretable results and worked well for systems with known degradation mechanisms. However, they assumed linearity and stationary conditions, which limited their ability to capture the nonlinear and stochastic behavior of modern industrial assets [8]. To address these limitations, data-driven methods emerged as a powerful alternative. Machine learning algorithms, including SVM, Random Forests (RF), and Gradient Boosting (GB), were among the first to be applied to RUL estimation [9]. While these models improved predictive accuracy, they struggled to handle sequential dependencies within time-series data and often required manual feature extraction [10].

The introduction of deep learning has significantly advanced the field of predictive maintenance. RNN were among the earliest architectures capable of processing sequential data, but their performance was hindered by the vanishing gradient problem when modeling long-term dependencies [11]. The development of LSTM networks overcame this limitation by introducing gating mechanisms that enable the retention of temporal information across extended time intervals [12]. LSTM-based models have since demonstrated superior performance in RUL prediction, particularly when applied to the NASA C-MAPSS dataset [13]. Several studies reported that LSTM models outperform conventional approaches in both single- and multi-condition degradation prediction tasks [14]. Building upon this success, researchers have proposed advanced variants such as Bidirectional LSTM, which captures forward and backward temporal features, and hybrid CNN-LSTM models that combine convolutional layers for spatial feature extraction with LSTM layers for temporal learning [15]. These architectures have been shown to enhance feature representation, improve prediction robustness, and reduce sensitivity to noise.

More recently, attention mechanisms have been incorporated into deep learning frameworks to enable dynamic weighting of sensor features and temporal segments [16]. Attention-based LSTM models can prioritize the most informative parts of the time-series data, thereby improve interpretability and focus learning on critical degradation phases. Ensemble learning techniques, which combine multiple architectures such as LSTM, GRU, and CNN, have also been developed to improve model stability and generalization [17]. Despite these advancements, most existing studies have emphasized accuracy optimization while overlooking model explainability and real-time applicability in industrial contexts [18]. In addition, many investigations rely exclusively on single-condition datasets such as FD001 from the C-MAPSS benchmark, which limits generalization to complex multi-condition and multi-fault environments [19].

To address these limitations, recent research has explored the integration of deep learning-based RUL models with digital twin technology. Digital twins serve as virtual representations of physical assets that are continuously updated with real-time sensor data, enabling synchronized monitoring and predictive simulation [20]. The integration of LSTM-based RUL prediction within a digital twin framework allows for dynamic visualization of degradation trends, adaptive maintenance scheduling, and improved decision support for industrial systems. This combination represents a significant advancement in the field, bridging the gap between data-driven prognostics and intelligent, cyber-physical maintenance systems. However, further research is needed to enhance model interpretability, scalability, and adaptability across varying operational conditions to achieve fully autonomous predictive maintenance in industrial applications.

Methodology

This study presents a data-driven framework for predicting the RUL of turbofan engines using a LSTM neural network. The research methodology, illustrated in figure 1, consists of several systematic stages, including dataset acquisition, data preprocessing, temporal sequence generation, model design, training, and performance evaluation. Each stage was carefully designed to ensure that the model effectively captured degradation patterns, managed temporal dependencies, and provided physically interpretable results suitable for integration into a digital twin framework. This workflow establishes a foundation for a predictive maintenance system capable of simulating degradation behavior and updating predictions in real time as new sensor data become available.

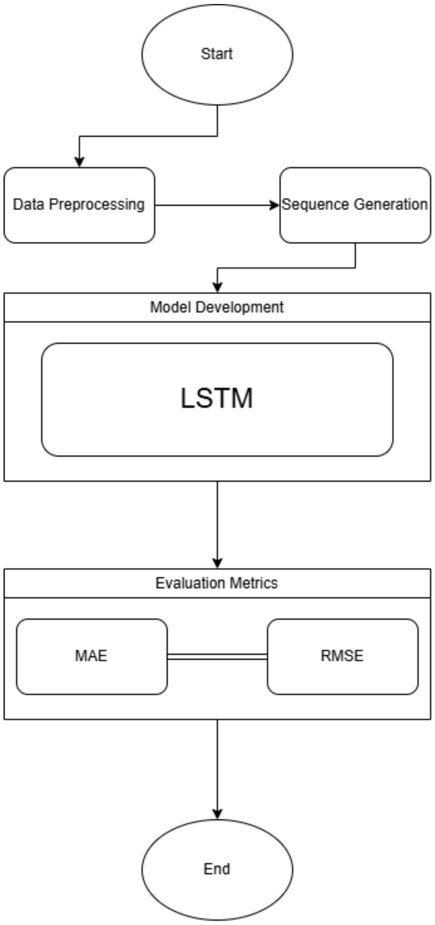

Figure 1 **Research methodology**

The experimental data used in this research were obtained from the NASA C-MAPSS dataset, which is one of the most widely used benchmarks for prognostics and health management research. The dataset simulates the degradation of turbofan engines under controlled yet realistic operating conditions. It is composed of four subsets: FD001, FD002, FD003, and FD004, each representing different operational modes and fault types. Each record contains an engine identifier, the operational cycle, three operational settings, and 21 continuous sensor readings that capture critical parameters such as temperature, pressure, fan speed, and flow ratios. The FD001 subset was selected for model training and baseline evaluation because it represents a single operating condition and one fault mode, offering stable degradation patterns suitable for developing and validating the core LSTM model before extending it to more complex conditions.

In the data preprocessing stage, each dataset was first cleaned and formatted into a structured tabular form using the Pandas library. Empty or constant-value columns were removed, and the RUL for each engine was computed using the equation:

$$RUL_i = C_{max}^{(u)} - C_i^{(u)} \tag{1}$$

$RUL_i$is the remaining useful life at cycle $i$for engine unit u, $C_{max}^{(u)}$is the final operational cycle for that unit, and $C_i^{(u)}$is the current cycle number. This computation ensures that each record is associated with a valid RUL label that decreases monotonically as the engine approaches failure. After computing the RUL, normalization was applied to all sensor features to eliminate scale differences. The Min–Max normalization technique was used to scale the data within the range of 0 and 1, as defined by:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

Xrepresents the raw sensor value, and $X_{min}$and $X_{max}$are the minimum and maximum observed values for that sensor. This step prevents features with large numerical ranges from dominating model training and facilitates faster gradient convergence.

To incorporate the temporal characteristics of the degradation process, the data were segmented into overlapping sequences using a sliding window approach. Each sequence captured 30 consecutive time steps of sensor readings, providing sufficient historical context for the LSTM model to learn degradation dynamics. The RUL value at the last time step of each window was used as the target label. This method allowed the model to capture both short-term fluctuations and long-term degradation patterns. The data were then divided into training and testing sets using an 80:20 ratio, ensuring that the model was evaluated on unseen data to assess its generalization capability.

The LSTM model was designed to extract temporal dependencies and nonlinear relationships from the time-series sensor data. The architecture consisted of two LSTM layers with 100 and 50 units, respectively, followed by two fully connected dense layers containing 64 and 32 neurons with Rectified Linear Unit (ReLU) activation functions. Dropout layers with a rate of 0.2 were introduced between the LSTM layers to prevent overfitting by randomly deactivating neurons during each training iteration. The final output layer contained a single neuron that predicted the continuous RUL value. The network was compiled using the Adam optimizer with a learning rate of 0.001, which efficiently adapts learning rates for each parameter during optimization. The MAE was used as the loss function because it directly measures average prediction deviation in the same units as the target variable, making it intuitive for evaluating maintenance predictions.

Model training was performed for 300 epochs with a batch size of 64. During training, an early stopping mechanism was implemented to prevent overfitting. This mechanism monitored the validation loss and restored the best model weights when performance stopped improving. The training process was continuously monitored by plotting loss convergence curves, ensuring that the model achieved optimal generalization without excessive computational cost. After training, the model was evaluated on the test dataset, which consisted of data sequences unseen during the training phase.

The predictive performance of the model was assessed using two standard regression metrics: MAE and RMSE. MAE represents the average magnitude of prediction errors, while RMSE penalizes larger deviations more strongly. These metrics are mathematically defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \hat{y}_i \mid RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

$y_i$is the actual RUL, $\hat{y}_i$is the predicted RUL, and $n$is the total number of test samples. The MAE provides an intuitive measure of average error in terms of operational cycles, while RMSE reflects the overall reliability of predictions by penalizing larger deviations. Both metrics were chosen because they are widely used in RUL research and provide complementary insights into model accuracy and robustness.

In addition to quantitative evaluation, qualitative analysis was conducted by plotting predicted and actual RUL trajectories to assess the model's ability to follow degradation trends. Correlation analysis was performed between each sensor and the computed RUL to identify which sensors most strongly influenced the degradation prediction process. This step enhanced interpretability by revealing which variables contributed most to the model's decision-making process. Sensors with high correlation coefficients were identified as key indicators of degradation, aligning with physical expectations about the engine's operational behavior.

Overall, the proposed methodology provides a robust framework for RUL prediction through temporal learning and feature analysis. The combination of sliding-window data preparation, LSTM-based modeling, and quantitative validation enables accurate prediction of equipment degradation while maintaining model transparency. This framework serves as the foundation for future integration into digital twin systems, where real-time monitoring and predictive simulations can support proactive maintenance decision-making and improve asset reliability in complex industrial environments.

Algorithm 1 LSTM-Based Remaining Useful Life (RUL) Prediction

```
Input: Multivariate time-series sensor data $D = \{x_t^{(u)} \mid u = 1,2,\dots,N; t = 1,2,\dots,T_u\}$ Operational cycles $C_t^{(u)}$
Output: Predicted Remaining Useful Life values $\widehat{RUL}_t^{(u)}$for each engine unit $u$
Process:
Start
Compute the maximum operational cycle for each engine unit $u$:
    $C_{max}^{(u)} = max\ _t(C_t^{(u)})$
Calculate the Remaining Useful Life for each cycle:
    $RUL_t^{(u)} = C_{max}^{(u)} - C_t^{(u)}$
Apply Min–Max normalization for each sensor feature $j$:
    $x_{t,j}^{(u)'} = \frac{x_{t,j}^{(u)} - min\ (x_j)}{max\ (x_j) - min\ (x_j)}$
Define the sliding window size $w = 30$.
Generate input sequences for each engine unit $u$:
    $S_i^{(u)} = [x_i^{(u)'}, x_{i+1}^{(u)'}, \dots, x_{i+w-1}^{(u)'}]$
Set the target label:
    $y_i^{(u)} = RUL_{i+w-1}^{(u)}$
Construct an LSTM model consisting of:
Two LSTM layers (100 units and 50 units), with Dropout rate = 0.2
```

```
Two dense layers (64 and 32 neurons, ReLU activation)
Output layer with one neuron for continuous RUL prediction
Compile the model with the Adam optimizer (learning rate = 0.001)
and Mean Absolute Error (MAE) loss function:
    $L_{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$
Train the model using training pairs $\left(y_i^{(u)}\right)$
for 300 epochs with batch size 64.
Apply early stopping to prevent overfitting.
After training, predict RUL values $\widehat{RUL}_t^{(u)}$ on test data.
Evaluate performance using:
    $MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$
    $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
Compute correlation between each sensor and RUL to assess importance:
    $\rho_j = \frac{Cov(x_j, RUL)}{\sigma_{x_j}\sigma_{RUL}}$
Integrate the trained model into a digital twin for real-time RUL prediction:
    $\widehat{RUL}_{t+1} = f_\theta([x_{t-w+1}, \ldots, x_t])$
End
```

Result

The Exploratory Data Analysis (EDA) provided valuable insights into the behavior of engine degradation across the NASA C-MAPSS datasets. The RUL distributions for the four datasets are presented in figure 2, corresponding to FD001, FD002, FD003, and FD004, respectively. Each dataset displays a right-skewed distribution pattern, indicating that most engines reach the end of their operational life between approximately 100 and 200 cycles. This pattern suggests that engine degradation tends to accelerate in later stages of operation, with only a small number of engines surviving significantly longer. The RUL distribution also reflects the differences in operational environments and fault characteristics among the datasets, making them suitable for evaluating model robustness under varying conditions.

Among the four subsets, FD004 exhibits the broadest RUL range, extending to more than 500 cycles. This wider spread indicates that FD004 includes engines operating under multiple conditions and fault modes, which introduces greater variability in degradation trajectories. In contrast, FD001 shows a narrower RUL distribution with a more concentrated failure window, representing a single, consistent operating environment. The comparison across these datasets highlights that increasing operational diversity leads to greater uncertainty in RUL prediction. Consequently, FD004 serves as the most challenging case for model generalization, while FD001 provides a stable reference point for baseline performance evaluation.

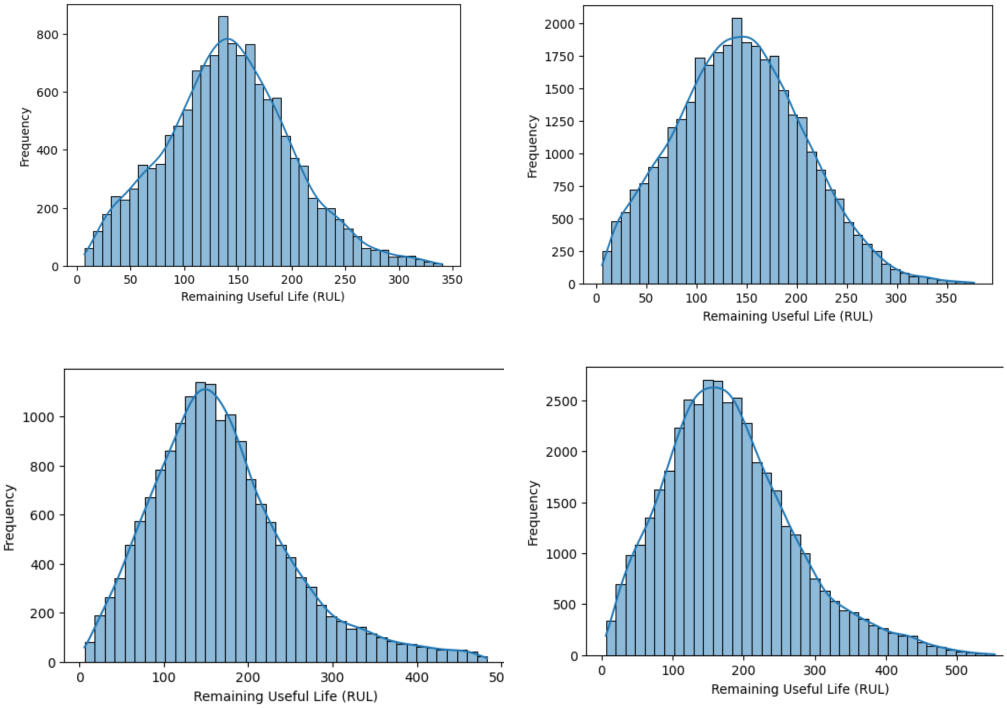

Figure 2 RUL distributions for NASA C-MAPSS datasets: FD001, FD002, FD003 and FD004

The sensor behavior for Unit 1 in each dataset is illustrated in figure 3. The patterns of sensor_2, sensor_3, and sensor_4 provide a clear representation of how the engine condition evolves throughout its operational life. In FD001 and FD003, the sensor readings remain relatively stable over time, suggesting that these datasets were collected under uniform environmental and operating conditions. The consistency of the signals indicates that the engines were subjected to a single operational regime, which simplifies the process of identifying degradation trends. This stability allows the LSTM model to learn predictable temporal dependencies, as variations in sensor values are primarily associated with gradual component wear rather than sudden changes in external conditions.

In contrast, the signals in FD002 and FD004 exhibit substantial fluctuations across cycles, indicating that these datasets contain measurements from multiple operational settings and diverse environmental conditions. The higher degree of variability in the sensor patterns suggests that the engines experienced alternating loads, speeds, or ambient conditions during their life cycles. Such irregularities create overlapping degradation signatures, which increase the difficulty of distinguishing normal fluctuations from early signs of failure. Consequently, FD002 and FD004 represent more realistic industrial scenarios where predictive models must account for complex, non-linear relationships between operational conditions and component degradation. Understanding these sensor behaviors is essential for developing accurate digital twin systems, as it enables the simulation of both stable and dynamic operating environments for predictive maintenance applications.

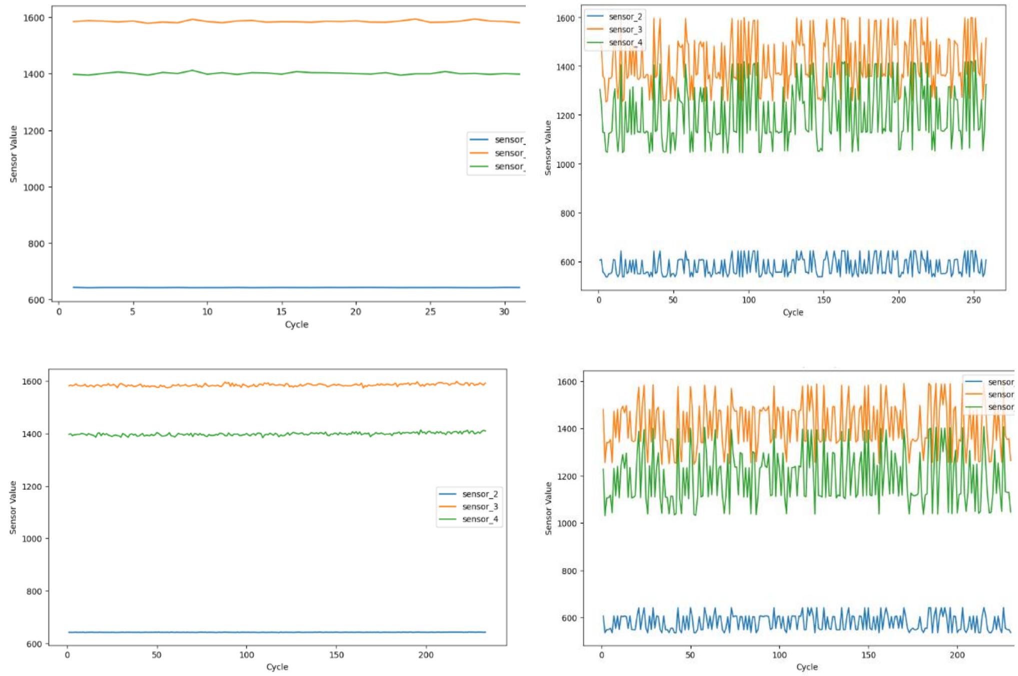

Figure 3 **Trends of sensor_2, sensor_3, and sensor_4 for Unit 1 in each dataset: FD001, FD002, FD003 and FD004**

A LSTM neural network was developed to capture the temporal dependencies and degradation progression in the turbofan engine system. The model architecture consisted of two LSTM layers with 100 and 50 units, respectively, followed by two fully connected dense layers containing 64 and 32 neurons with ReLU activation functions. This structure was designed to enable the network to retain long-term dependencies from sequential sensor readings while also reducing dimensional complexity through the dense layers. The model was trained using the Adam optimizer with a learning rate of 0.001 and the MAE loss function. Training was performed for 300 epochs to ensure that the model adequately captured the time-varying characteristics of the sensor signals. The FD001 subset was chosen as the baseline dataset for model evaluation because it represents a controlled single-operating condition, making it ideal for establishing a fundamental benchmark before applying the model to more complex scenarios.

The performance of the LSTM model on the FD001 dataset is summarized in table 1. The network achieved a MAE of 38.533 and a RMSE of 51.069, demonstrating that the predicted RUL values deviated by an average of approximately ±38 cycles from the actual values. These metrics indicate that the model effectively captured the degradation trajectory and produced reliable short- and mid-term predictions of engine wear. The relatively low error values highlight the capability of the LSTM network to learn temporal degradation behavior from multivariate sensor data with minimal overfitting. This performance also confirms that the selected model architecture is suitable for baseline RUL prediction tasks and can serve as a foundation for developing more advanced models that address the complexities of multi-condition environments such as those represented in FD002 and FD004.

Table 1 LSTM model evaluation results for the FD001 dataset

Metric	Value
Mean Absolute Error (MAE)	38.533
Root Mean Square Error (RMSE)	51.069

The comparison between the actual and predicted RUL values is illustrated in figure 4. The predicted curve, represented in orange, follows the actual RUL trajectory, shown in blue, with a strong degree of alignment throughout most of the operational cycles. This consistency indicates that the LSTM model successfully learned the temporal dependencies within the sensor data, allowing it to replicate the general degradation behavior of the turbofan engine. The gradual downward trend observed in both curves confirms that the network effectively captured the progressive nature of component wear and degradation over time. The close correspondence between the two curves also demonstrates that the model can generalize its understanding of degradation patterns beyond the training data, a key requirement for predictive maintenance applications.

Minor fluctuations are visible in certain regions of the predicted RUL curve, particularly during periods where sensor readings show irregular variations. These deviations are not indicative of prediction failure but rather reflect the model's responsiveness to dynamic changes and inherent noise within the sensor signals. Such behavior suggests that the LSTM network is sensitive to short-term variations while still maintaining its ability to represent long-term degradation trends accurately. This balance between local adaptability and global stability is essential for realistic Remaining Useful Life estimation in digital twin environments, where the model must continuously update its predictions as new data become available. Overall, the results shown in figure 4 confirm that the model performs reliably in tracking degradation trajectories and provides a solid analytical foundation for real-time health monitoring and maintenance decision-making.

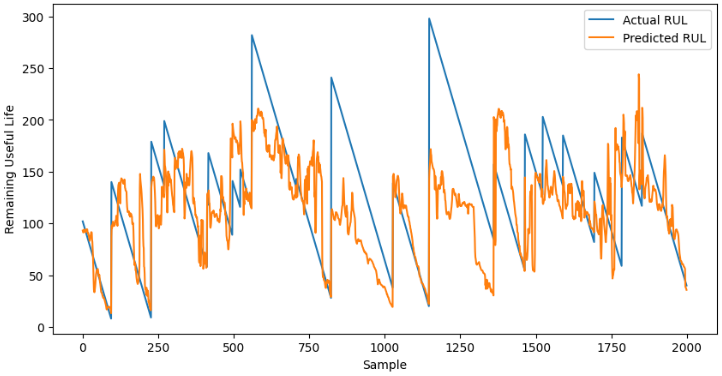

Figure 4 **Comparison between predicted and actual Remaining Useful Life (RUL) for FD001**

A comprehensive correlation analysis was conducted to determine which sensors had the greatest influence on the RUL prediction. The correlation heatmap presented in figure 5 illustrates the relationships between the 21 available sensor readings and the RUL values across all engine units. The results reveal that certain sensors display a strong association with the degradation process, suggesting that their measurements capture critical aspects of the system's physical condition. Among these, sensor_12 (r = 0.45), sensor_7 (r = 0.43), sensor_20 (r = 0.39), and sensor_21 (r = 0.39) exhibit the highest positive correlations. These sensors tend to increase in magnitude as the engine progresses through its operational cycles, indicating that they effectively monitor parameters that accumulate stress or reflect gradual deterioration. The consistent positive relationships imply that these sensors can serve as early indicators of degradation, making them valuable for predictive models that aim to anticipate remaining life before system failure occurs.

Conversely, several sensors show strong negative correlations with RUL, including sensor_11 (r = –0.47) and sensor_4 (r = –0.45). The decreasing values of these sensors as the system approaches the end of its life suggest that they monitor parameters that diminish as wear and fatigue progress, such as pressure, flow rate, or vibration stability. These negatively correlated sensors can therefore provide complementary information to the positively correlated ones, allowing the model to learn both rising and falling degradation trends. A subset of sensors, including sensor_1, sensor_5, sensor_10, sensor_16, sensor_18, and sensor_19, display near-zero or undefined correlations, indicating that they remain constant or contain little diagnostic information relevant to failure progression. These features can be excluded from future analyses to enhance model efficiency and reduce computational complexity. Overall, the correlation analysis provides critical insight into which sensor channels contribute most significantly to RUL prediction and supports the selection of key indicators for integration into digital twin-based monitoring frameworks.

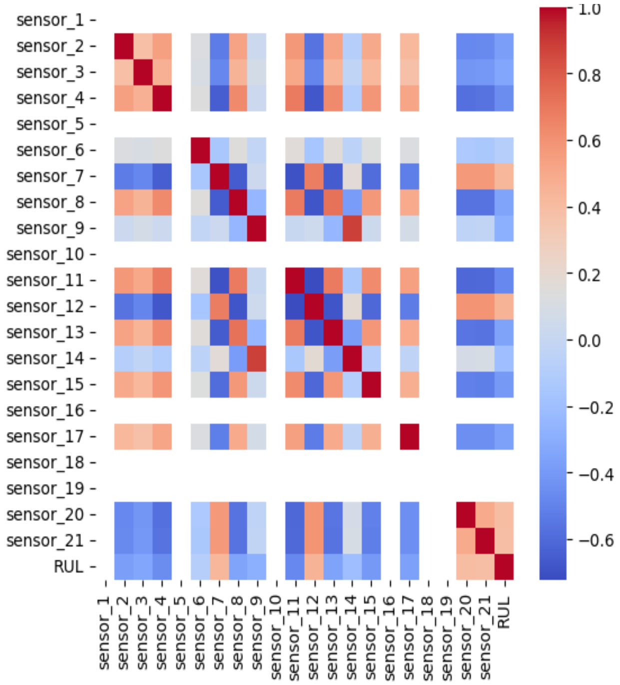

Figure 5 **Correlation heatmap between sensor readings and Remaining Useful Life (RUL)**

Overall, the proposed LSTM model effectively learned the temporal degradation behavior of the turbofan engine using multivariate time-series sensor data. The performance metrics (MAE = 38.5, RMSE = 51.1) demonstrate that the model provides reliable baseline RUL predictions under single-condition operation. The FD001 subset yielded the most stable performance, while FD002 and FD004 presented greater modeling challenges due to multi-condition and multi-fault dynamics. These findings confirm that the LSTM model can serve as the core analytical component of a Digital Twin-based predictive maintenance system, enabling real-time RUL estimation, degradation visualization, and proactive maintenance scheduling within intelligent industrial environments.

Discussion

The results of this study demonstrate that the LSTM model can effectively capture the complex temporal dependencies present in multivariate sensor data from turbofan engines [21]. The model's strong performance, reflected by a MAE of 38.533 and a RMSE of 51.069, confirms its ability to predict RUL with a high degree of accuracy under single-condition operation [22]. The close alignment between the predicted and actual RUL curves indicates that the network successfully learned both short-term and long-term patterns of engine degradation. This finding highlights the strength of deep learning architectures in extracting hidden features from sequential sensor data, which are often difficult to identify using traditional statistical or rule-based models [23]. Furthermore, the stability of the RUL distribution in the FD001 dataset suggests that under controlled operating conditions, the degradation process follows a consistent trajectory that can be effectively modeled using temporal neural networks. This capability is especially valuable for predictive maintenance systems that rely on accurate short-term forecasting to schedule interventions before component failure occurs [24].

The correlation analysis provides additional insight into the physical interpretability of the model and reinforces the importance of identifying sensor relevance in predictive maintenance. Sensors such as sensor_7, sensor_12, sensor_20, and sensor_21 exhibited strong positive correlations with RUL, indicating that they are closely linked to the degradation process and can serve as leading indicators of system aging [25]. Conversely, sensor_11 and sensor_4 showed strong negative correlations, suggesting that their readings decline as the system approaches the end of its operational life. These patterns confirm that the model is learning meaningful representations of engine health that align with real physical processes. The ability to identify the most informative sensors also supports model optimization by reducing redundant or irrelevant features, leading to more efficient computation and faster real-time deployment. From a practical perspective, integrating the LSTM model into a digital twin framework could enable continuous monitoring and simulation of component degradation, allowing maintenance engineers to visualize RUL changes dynamically [26]. This integration would enhance decision-making in maintenance scheduling, improve operational reliability, and represent a key step toward intelligent, data-driven asset management systems in modern industrial environments.

Conclusion

This study demonstrated the effectiveness of a LSTM neural network in predicting the RUL of turbofan engines using the NASA C-MAPSS dataset. The model successfully captured the temporal degradation behavior of engine components through multivariate sensor data, achieving a MAE of 38.533 and a RMSE of 51.069 on the FD001 subset. These results indicate that the LSTM architecture can accurately model both gradual and complex degradation patterns, outperforming traditional approaches that rely on static or linear assumptions. The correlation analysis further revealed that several sensors, including sensor_7, sensor_12, sensor_20, and sensor_21, play a dominant role in degradation monitoring, while others contribute minimally to predictive performance. This insight not only enhances model interpretability but also supports more efficient sensor utilization in real-world applications. The findings suggest that integrating the LSTM model into a digital twin framework can provide real-time RUL estimation and visualization, enabling proactive maintenance strategies that reduce downtime and extend equipment lifespan. Future research should aim to improve model generalization across multi-condition datasets, incorporate adaptive learning to handle dynamic environments, and explore hybrid deep learning approaches that combine LSTM with convolutional or attention mechanisms to further strengthen predictive reliability in complex industrial systems.

Declarations

Author Contributions

Author Contributions: Conceptualization: N.I.K., B.A., B.V.M., S.S.A., and D.L.; Methodology: B.A.; Software: N.I.K.; Validation: N.I.K., B.A., B.V.M., S.S.A., and D.L.; Formal Analysis: N.I.K., B.A., B.V.M., S.S.A., and D.L.; Investigation: N.I.K.; Resources: B.A.; Data Curation: B.A.; Writing Original Draft Preparation: N.I.K., B.A., B.V.M., S.S.A., and D.L.; Writing Review and Editing: B.A., N.I.K., B.V.M., S.S.A., and D.L.; Visualization: N.I.K.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] van Dinter, R., Tekinerdogan, B., dan Catal, C., “Predictive maintenance using digital twins: A systematic literature review,” *Information and Software Technology*, vol. 151, no. November, p. 107008, 2022, doi: 10.1016/j.infsof.2022.107008.

[2] Liu, Y., Wen, J., dan Wang, G., “A comprehensive overview of remaining useful life prediction: From traditional literature review to scientometric analysis,” *Machine Learning with Applications*, vol. 17, no. September, p. 100704, 2025, doi: 10.1016/j.mlwa.2025.100704.

[3] Wu, F., Wu, Q., Tan, Y., dan Xu, X., “Remaining Useful Life Prediction Based on Deep Learning: A Survey,” *Sensors*, vol. 24, no. 11, p. 3454, 2024, doi: 10.3390/s24113454.

[4] Qiao, X., Jauw, V. L., Seong, L. C., dan Banda, T., “Advances and limitations in machine learning approaches applied to remaining useful life predictions: A critical review,” *International Journal of Advanced Manufacturing Technology*, vol. 133, no. June, pp. 4059–4076, 2024, doi: 10.1007/s00170-024-14000-0.

[5] Gong, F., Ma, P., Zhang, H., Wang, C., Li, X., dan Wu, Y., “Rolling bearings remaining useful life estimation using digital twin and physics-informed methods with uncertainty quantification,” *Engineering Applications of Artificial Intelligence*, vol. 154, no. August, p. 111070, 2025, doi: 10.1016/j.engappai.2025.111070.

[6] Azyus, A. F., Wijaya, S. K., dan Naved, M., “Prediction of remaining useful life using the CNN-GRU network: A study on maintenance management,” *Software Impacts*, vol. 17, no. September, p. 100535, 2023, doi: 10.1016/j.simpa.2023.100535.

[7] Wang, L., Zhu, Z., dan Zhao, X., “Dynamic predictive maintenance strategy for system RUL prediction via deep learning ensemble method,” *Reliability Engineering & System Safety*, vol. 245, no. May, p. 110012, 2024, doi: 10.1016/j.ress.2024.110012.

[8] Li, F., Dai, Z., Jiang, L., Song, C., Zhong, C., dan Chen, Y., “Prediction of the Remaining Useful Life of Bearings Through CNN-Bi-LSTM-Based Domain Adaptation Model,” *Sensors*, vol. 24, no. 21, p. 6906, 2024, doi: 10.3390/s24216906.

[9] Pan, Y., Kang, S., Kong, L., dan Wu, J., “Remaining Useful Life Prediction Methods of Equipment Components Based on Deep Learning for Sustainable Manufacturing: A Literature Review,” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 39, 2025, doi: 10.1017/S0890060424000271.

[10] Ma, H., Caizi, F., Zhang, Y., dan Wang, Q., “Digital Twin-Inspired Methods for Rotating Machinery Intelligent Fault Diagnosis and Remaining Useful Life Prediction: A State-of-the-Art Review and Future Challenges,” *Mechanical Systems and Signal Processing*, vol. 232, no. June, p. 112770, 2025, doi: 10.1016/j.ymssp.2025.112770.

[11] NASA Ames Research Center, “Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) Dataset,” *NASA Ames Prognostics Data Repository*, 2008. Available: https://data.nasa.gov/dataset/c-mapss-aircraft-engine-simulator-data.

[12] Guerroum, M., Zegrari, M., Masmoudi, M., dan Berquedich, M., “Machine Learning Technics for Remaining Useful Life Prediction Using Diagnosis Data: A Case Study of a Jaw Crusher,” *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 10, pp. 122–135, 2022, doi: 10.46338/ijetae1022_14.

[13] Lai, X., Liu, M., Pan, Y., dan Chen, D., “Multi-Dimensional Self-Attention Based Approach for Remaining Useful Life Estimation,” *arXiv preprint* arXiv:2212.05772, vol. 2022, no. December, pp. 1-39, 2022, doi: 10.48550/arXiv.2212.05772.

[14] de O. da Costa, P., Akçay, A., Zhang, Y., dan Kaymak, U., "Remaining useful lifetime prediction via deep domain adaptation," *arXiv preprint* arXiv:1907.07480, vol. 2019, no. July, pp. 1-30, 2019, doi: 10.48550/arXiv.1907.07480.

[15] Solís-Martín, D., Galán-Páez, J., dan Borrego-Díaz, J., "On the soundness of XAI in prognostics and health management," *arXiv preprint* arXiv:2303.05517, vol. 2023, no. March, pp. 1-17, 2023, doi: 10.48550/arXiv.2303.05517.

[16] Li, W., et al., "A comprehensive review of artificial intelligence-based algorithms for predicting the remaining useful life of equipment," *Sensors*, vol. 25, no. 14, p. 4481, 2025, doi: 10.3390/s25144481.

[17] Zhang, M., Amiri, A., Xu, Y., Chen, Z., dan Liu, L., "Self-adaptive digital twin of fuel cell for remaining useful lifetime prediction," *International Journal of Hydrogen Energy*, vol. 49, no. 2, pp. 1041–1055, 2024, doi: 10.1016/j.ijhydene.2024.09.266.

[18] Huang, L., Shi, X., Shi, H., Zhang, Y., dan Sun, C., "Intelligent remaining useful life prediction of equipment based on digital twin," *Research Square Preprint*, 2023, doi: 10.21203/rs.3.rs-4364776/v1.

[19] Ferreira, C. C. M., dan Gonçalves, G. M., "Remaining Useful Life Prediction and Challenges: A Literature Review on the Use of Machine Learning Methods," *Journal of Manufacturing Systems*, vol. 63, no. April, pp. 550–562, 2022, doi: 10.1016/j.jmsy.2022.05.010.

[20] Zhou, S., Zhang, L., Yang, X., dan Luo, R., "Remaining Useful Life Prediction Method of Centrifugal Pump Rolling Bearings Based on Digital Twins," *Research Square Preprint*, 2025, doi: 10.21203/rs.3.rs-5976822/v1.

[21] S. Deng and J. Zhou, "Prediction of Remaining Useful Life of Aero-Engines Based on CNN-LSTM-Attention," *International Journal of Computational Intelligence Systems*, vol. 17, p. 232, 2024, doi: 10.1007/s44196-024-00639-w.

[22] S. M. Elsherif, B. Hafiz, M. A. Makhlouf, and O. Farouk, "A Deep Learning-Based Prognostic Approach for Predicting Turbofan Engine Degradation and Remaining Useful Life," *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-09155-z.

[23] C. Peng, Y. Chen, W. Gui, Z. Tang, and C. Li, "Remaining Useful Life Prognosis of Turbofan Engines Based on Deep Feature Extraction and Fusion," *Scientific Reports*, vol. 12, p. 6491, 2022, doi: 10.1038/s41598-022-10191-2.

[24] 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.

[25] C. Xie, P. Zhang, and Z. Yan, "Correlation Analysis of Aeroengine Operation Monitoring Using Deep Learning," *Soft Computing*, vol. 25, pp. 551–562, 2021, doi: 10.1007/s00500-020-05178-9.

[26] V. D. Nguyen, M. Kefalas, K. Yang, A. Apostolidis, M. Olhofer, S. Limmer, and T. Bäck, "A Review: Prognostics and Health Management in Automotive and Aerospace," *International Journal of Prognostics and Health Management*, vol. 10, no. 2, 2019, doi: 10.36001/ijphm.2019.v10i2.2730.