International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. V, No. 05, pp. 1-10, May

Artificial Rabbits Optimized Mamba-Transformer Network for Aircraft Turbofan Engine Remaining Useful Life Prediction

D. Lakshmi[3], D. Karthikeyan[4]

[3]Professor, Department of Electrical and Electronics Engineering,
AMET University Chennai, Tamil Nadu, India-603112.
[4]Assistant professor, Department of Electrical and Electronics Engineering,
SRM Institute of Science and Technology, Kattankulathur, Chennai, India-603203.

Corresponding Author E-mail

ABSTRACT: Air hazards caused by component deterioration is evaded by forecasting the Remaining Useful Life (RUL) of turbofan engines, which is a vital process in health management and prognostics. Thus, an Artificial Rabbits Optimised Mamba-Transformer Network (AROA-MTN) for accurate turbofan engine RUL estimate is developed in this paper. To find important patterns, variations, and deterioration features, the framework first pre-processes the data gathered from CMAPSS Jet Engine Simulated Data using data visualisation and Exploratory data analysis (EDA). The MTN model efficiently captures both complicated temporal linkages and long-term sequential dependencies in engine operating data by combining the Mamba architecture with a Transformer network. The Artificial Rabbits Optimisation (ARO) is employed to optimise important network parameters and enhance the learning process to further raise model performance. To assess the accuracy, robustness, and generalisation potential of the developed method, the outcomes are evaluated in Python tool, reveals the RMSE of 0.0155. Therefore, the AROA-MTN framework offers an efficient data-driven solution for aircraft engine prognostics and permits reduced operating risk and enhanced engine reliability.

Keywords: RUL, Data visualization, Feature engineering, EDA, AROA-MTN.

1. Introduction

The increasing focus on the dependability and preservation of intricate systems, necessitates intelligent and self-sufficient approaches [1]. One method is to monitor on the condition of turbofan engines via autonomous anomaly detection [2]. Turbofan engine operators avert catastrophic failures by taking corrective action when anomalies are detected early [3]. Precisely determining an aviation engine's RUL is vital to lowering maintenance costs while maintaining reliability [4]. The degradation tendencies among the different circumstance monitoring sensors are the base for the development of the RUL prediction model [5]. For the system to be employed as efficiently as possible, early anomaly identification and prompt failure warning are vital [6].

2. Recent Works

To predict the system's RUL, data-driven approaches employ a variety of Machine learning (ML) algorithms. When paired with efficient pre-processing techniques, an LSTM [7] network estimate the RUL with a high degree of accuracy. Numerous factors, like hyperparameters for the LSTM model and parameters for the initial RUL computations, affect the prediction accuracy. High data hunger, sensitivity to sensor noise, black-box opacity and high computing costs during extended flight cycles are some of the difficulties it poses. Also, the RF algorithm [8] is employed to efficiently excerpt features. This reduced the effect of data redundancy on the accuracy.

For regression problems, the Convolutional Auto encoder and Attention-based LSTM (CAELSTM) [9] is appropriate. Further developments involve expanding the model using domain adaption techniques to increase performance to increase CAELSTM's generalization. The large dimensions from a turbofan engine degradation process are addressed by the stacked sparse Auto encoder-Temporal Convolutional Network (SAE-TCN) [10]. It takes a lot of processing resources and fine-tuning overhead to train both a deep TCN sequence forecaster and an unsupervised SAE for feature extraction. In high-dimensional issues, Bayesian Optimization [8] become computationally expensive and less effective, but it provides effective global search with fewer function evaluations. Strong exploration-exploitation balance and good global optimization capability are provided by improved GWO [11], nonetheless, its performance is reliant on parameter. Adam [12] Optimization converges fast, and automatically adjusts learning rates, yet, it necessitate careful hyperparameter adjustment and converge to less-than-ideal solutions. Thus, the AROA-MT classifier is employed in this research for RUL prediction. Transformer self-attention learns intricate relationships amid multivariate engine sensors, while Mamba effectively captures long-range temporal deterioration patterns with linear-complexity sequence modelling. Furthermore, sequence length, learning rate, Mamba state dimension, network depth, attention heads, and dropout rate are among the crucial hyperparameters of the Mamba–Transformer network that are optimised by ARO. The main contributions are,

- By using data visualization and EDA to find trends, correlations, missing values and aberrant observations, the preprocessing stage enhances the quality and consistency of aircraft turbofan-engine sensor data.
- To capture inter-sensor dependencies and long-term temporal degradation trends in turbofan-engine data, the classifier combines Mamba and Transformer network. An ideal set of MT hyperparameters is tuned by the ARO algorithm.

3. Proposed Methodology

The first stage in the aircraft turbofan engine RUL prediction process is the collection of engine sensor input data, which is then pre-processed to enhance the consistency and quality of the data. To find substantial patterns, correlations, distributions and anomalies in the engine running data, preprocessing comprises data visualization and EDA. Conversely, pre-processed data into an appropriate representation, feature engineering is employed to extract and produce informative features linked to degradation. The Mamba–Transformer hybrid network receives the training data with Mamba capturing long-term temporal dependencies and the Transformer using attention processes to learn substantial sequential relationships. To increase convergence and prediction accuracy, the AROA optimizes the model's hyperparameters. The optimized network yields the anticipated RUL after learning the turbofan engine's degrading behaviour.

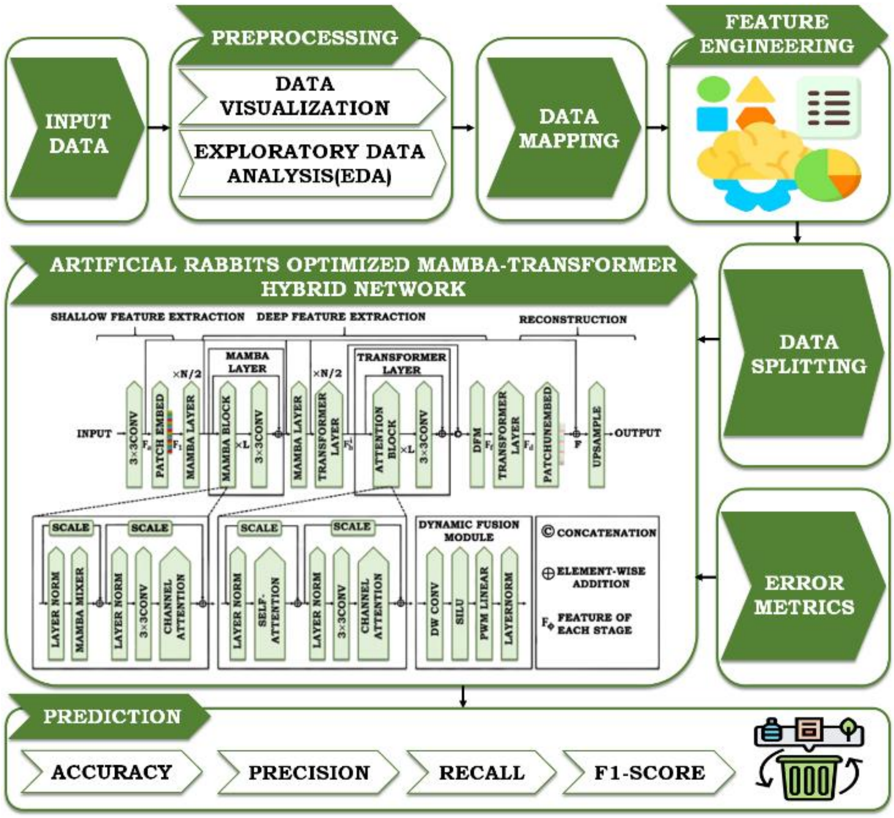

Figure 1: *Proposed framework*

The ARO-optimized Mamba–Transformer hybrid network enhances prediction accuracy, convergence, and predictive maintenance decision dependability by providing a reliable and efficient framework for turbofan engine RUL forecasting.

3.1 Data preprocessing

Data preprocessing manages missing or noisy observations and eradicates unrelated or inconsistent measurements. It comprises 2 stages:

3.1.1 Data visualization:

Data visualisation permits analysts to examine and understand the data visually. Histograms, box plots, scatter plots, bar charts and heat maps are instances of visualisations that provide clear depictions of the data while highlighting trends, outliers, patterns and correlations between variables.

3.1.2 EDA

Finding patterns, connections and insights in unprocessed data is the aim of EDA. Before using more sophisticated analytical approaches, it entails using a variability of statistical and visualisation techniques to analyse and comprehend the underlying structure of the data.

3.2 AROA-MT Classifier

The Mamba–Transformer Network learns inter-sensor connections and temporal correlations by processing the mapped multivariate sensor sequences. Using state-space sequence modelling, the Mamba module effectively captures long-range degradation trends during engine operation cycles. To predict substantial relationships amid sensor features and temporal representations, the Transformer module employs self-attention. To compute mistakes in the pixel and frequency domains during training, a unique dual-domain loss is employed as,

$$\mathcal{L}\ominus = \sum_{i=1}^{S} \left\| i_{HR}^{(i)} - i_{SR}^{(i)} \right\|_2 + \lambda \left\| F\left(I_{HR}^{(i)}\right) - F\left(I_{SR}^{(i)}\right) \right\|_2 \quad (1)$$

Here $F(\cdot)$is for the fast Fourier transform, λ is adjusted to 0.1 to establish balance during training,

S is the training samples count, i_{HR} is the ground truth images and Ψ signifies the learnable parameters. Long-range dependencies are captured by Self-Attention or Mamba Mixer blocks, as presented in Figure 2.

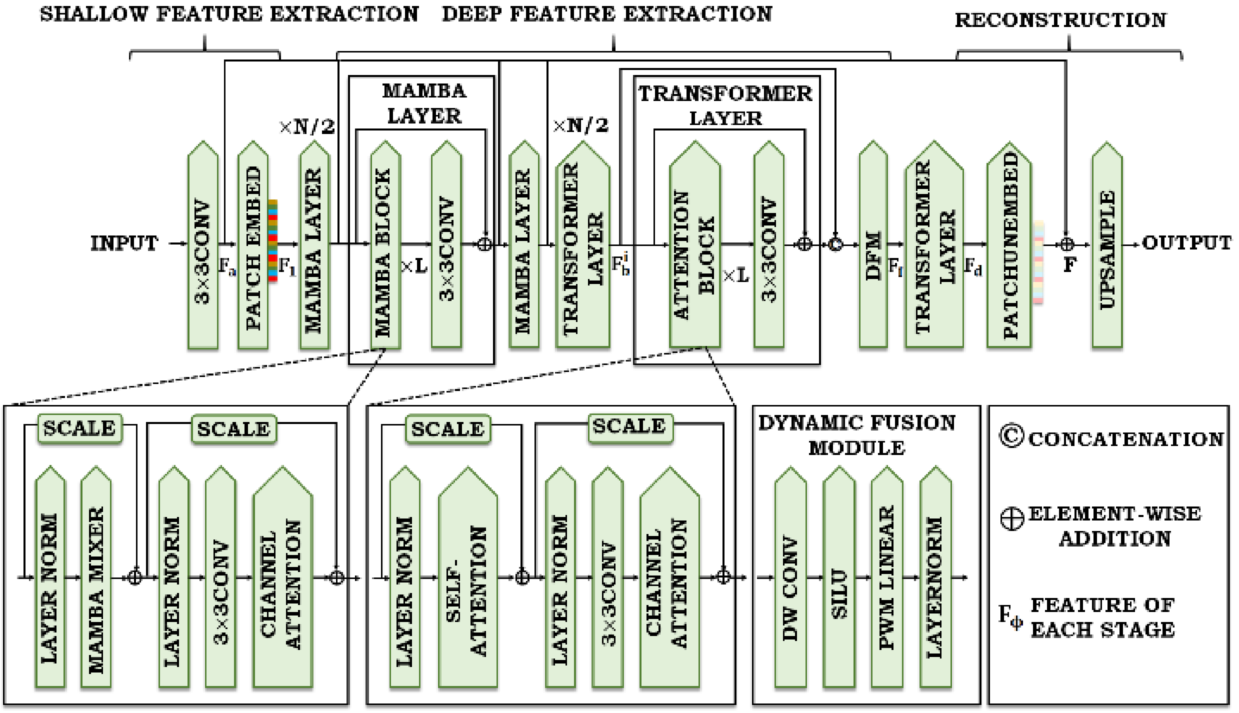

Figure 2: *AROA-MT architecture*

The characteristics $\hat{X}_{i,l}$ in the intermediate is calculated by,

$$\begin{cases} \hat{X}_{i,l} = Mamba\left(Norm(X_{i,l})\right) + S_1 \times X_{i,l}, i \leq \frac{N}{2} \\ \hat{X}_{i,l} = Attention\left(Norm(X_{i,l})\right) + S_1 \times X_{i,l}, else \end{cases} \tag{2}$$

A Channel Attention Block (CAB) is employed to rise the expressive capability of numerous channels and avert channel redundancy.

$$X_{i,l+1} = CAB\left(Conv\left(Norm(\hat{X}_{i,l})\right)\right) + S_2 \times \hat{X}_{i,l} \tag{3}$$

The DFM enhances connection within the U-Net architecture while preserving efficiency by aggregating hierarchical information in terms of channel and geographic space. The process is normalized as,

$$X_1 = \sigma\left(Conv\left(Linear\left(C, \frac{C}{2}\right)(X_{in})\right)\right) \tag{4}$$

$$X_2 = DMSSM\sigma\left(Conv\left(Linear\left(C, \frac{C}{2}\right)(X_{in})\right)\right) \tag{5}$$

$$X_{out} = Linear\left(C, \frac{C}{2}\right)\left(Concat(X_1, X_2)\right) \tag{6}$$

Here $Linear$ $(\cdot)$is the linear operation, $Conv(\cdot)$ for the convolution operation, $\sigma(\cdot)$ for the activation function, $DMSSM$ $(\cdot)$is the multidirectional selective scanning operation and $Concat$ $(\cdot)$is employed the convolution operation.

$$Attention\ (Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_h}}\right)V \tag{7}$$

Where K is the key, V is the value and Q specifies query.The number of self-attention heads is denoted as d_h. To initialise the ARO population, n rabbits are generated, each of which signifies a potential set of Mamba-Transformer hyperparameters for turbofan-engine RUL prediction. Learning rate, sequence length, convolution kernel size, batch size, expansion factor, Mamba layers, Transformer layers, Mamba state dimension, attention heads, embedding dimension and dropout rate are among the parameters comprised in the position matrix as,

$$X = [X_1, X_2, \cdots X_i, \cdots X_n]^T = \begin{bmatrix} X_{11} & \cdots & X_{1d} \\ \vdots & X_{ij} & \vdots \\ X_{n1} & \cdots & X_{nd} \end{bmatrix} \quad (8)$$

Here, d is the total number of hyperparameters being optimised.

$$F_X = \begin{bmatrix} f(X_{11} \cdots X_{1d}) \\ \vdots X_{1j} \vdots \\ f(X_{n1} \cdots X_{nd}) \end{bmatrix}_{n*d} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}_{n*1} \quad (9)$$

The RUL prediction performance is signified by the fitness functionF_X,The flowchart of AROA-MT model is depicted in Figure 3.

$$\vec{V}_i(t+1) = \vec{X}_j(t) + R.\left(\vec{X}_i(t) - \vec{X}_j(t)\right) + round\big(0.5 + (0.05 + r_1)\big).n_1 \quad (10)$$

$$R = L.C \quad (11)$$

$$L = \left(e - e\left(\frac{t-1}{T}\right)^2\right).sin(2\pi r_2) \quad (12)$$

To avert local optima, ARO simultaneously investigate numerous combinations of Transformer parameters, Mamba architecture parameters, sequence parameters and training parameters.

$$C(k) = \begin{cases} 1, k = g(l) \\ 0, Else \end{cases}; k = 1, \cdots d, l = 1, \cdots [r_3, d], g = randperm(d) \quad (13)$$

Local exploitation around prospective Mamba–Transformer configurations obtained during exploration is carried out using the random concealment method.

$$\vec{b}_{i,u}(t) = \vec{X}_i(t) + H.g_u.\vec{X}_l(t); i = 1, \cdots n, j = 1, \cdots d \quad (14)$$

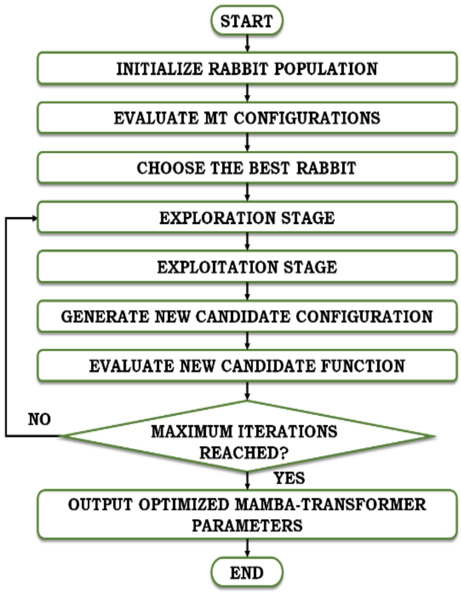

Figure 3: *Flowchart of AROA-MT model*

Here, a candidate hiding position is represented by $\vec{b}_{i,u}(t)$, the hyperparameter dimension chosen for local modification is determined by g_u and the hiding/search magnitude is managed by H.

$$H = \frac{T-t+1}{T}.n_2; n_2 \sim N(0,1) \quad (15)$$

A certain hyperparameter vector dimension is chosen for exploitation by the parameter g_u.

$$g_u(k) = \begin{cases} 1, k = [r_5.d] \\ 0, Otherwise \end{cases}; k = 1, \cdots d \quad (16)$$

Equation (16) is employed to update the candidate rabbit position following either random hiding or detour foraging.

$$\vec{V}_i(t+1) = \vec{X}_i(t) + R.\left(r_4.\vec{b}_{i,u}(t) - \vec{X}_l(t)\right); i = 1, \cdots n \quad (17)$$

The RUL prediction performance of the resultant network is evaluated using the turbofan-engine validation data.

$$\vec{X}_i(t+1) = \begin{cases} \vec{X}_i(t), f\left(\vec{X}_i(t) \le \left(\vec{V}_i(t+1)\right)\right) \\ \vec{V}_i(t+1) \; Otherwise \end{cases} \quad (18)$$

Whether the newly produced Mamba–Transformer configuration is superior to the present configuration is determined by the fitness-based selection. Thus, precise RUL forecasts for aircraft turbofan engines are attained by the final optimised model.

4. Results and Discussion

This section analyses an outcomes of developed model in Python software. The original TXT files are transformed and combined into a single CSV dataset, yielding a dataset of 265,256 samples and 29 features. The target variable is RUL and the input features are engine ID, cycle. An 80:20 split is employed to distinct the dataset into 212,204 training samples and 53,052 testing samples.

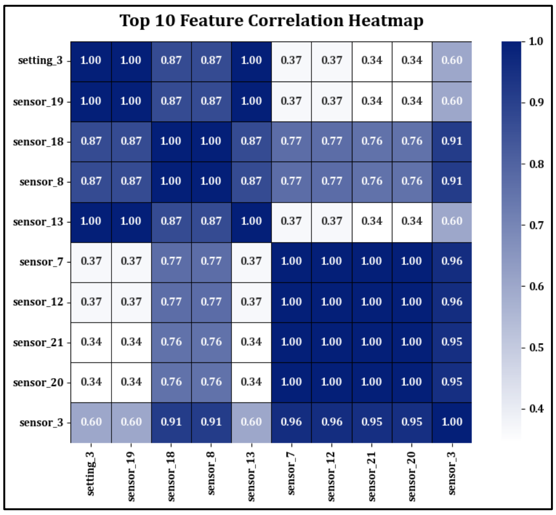

Figure 4: *Correlation matrix*

The correlation matrix of the ten most significant sensor characteristics is depicted in Figure 4. By recognizing pertinent inter-sensor dependencies for precise RUL prediction, the analysis enables effective feature representation in the AROA-MTN Network.

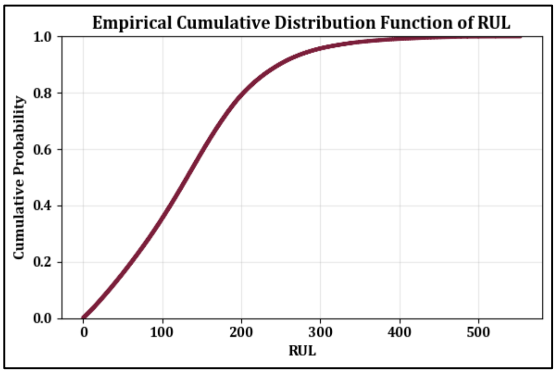

Figure 5: *Empirical cumulative distribution*

The cumulative distribution of RUL, which describes the cumulative probability of engine observations across numerous RUL values is presented in Figure 5.

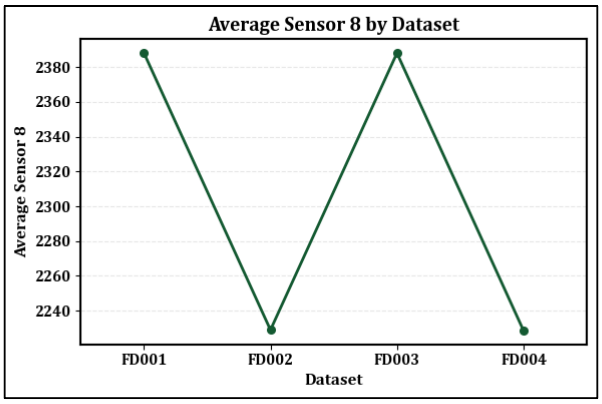

Figure 6: *Average Sensor 8 values*

The average Sensor 8 values for FD004, FD002 FD003, FD001 and are displayed in Figure 6. While FD002 and FD004 display lower values close to 2230, FD001 and FD003 show higher average values of 2388.

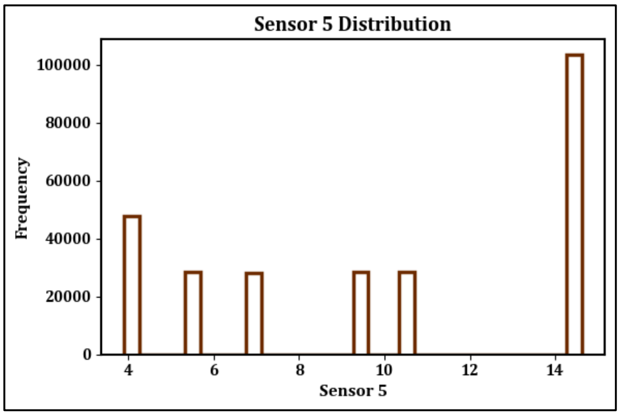

Figure 7: *Frequency distribution of Sensor 5*

The frequency distribution of Sensor 5 is represented in Figure 7. This distribution displays numerous operating states and emphasises in what way vital it is to acquire pertinent sensor-state demonstrations for precise turbofan RUL estimate.

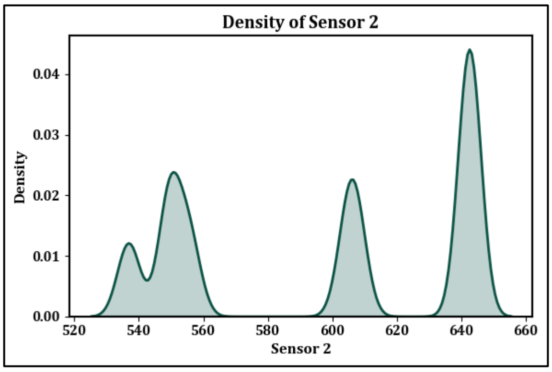

Figure 8: *Density distribution of Sensor 2*

The probability density distribution of Sensor 2, which has numerous different peaks is shown in Figure 8. Over the AROA-MTN architecture, these nonlinear and multimodal features highlight the importance of sophisticated representation learning.

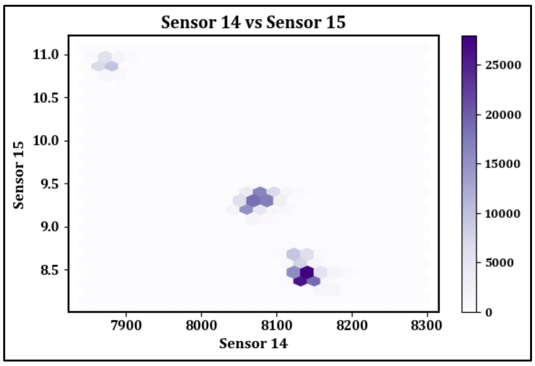

Figure 9: *Joint relationship between Sensors 14 and 15*

The joint relationship between Sensors 14 and 15 is depicted in Figure 9. The Mamba-Transformer framework for RUL prediction effectively take use of these patterns, which display substantial cross-sensor interdependence.

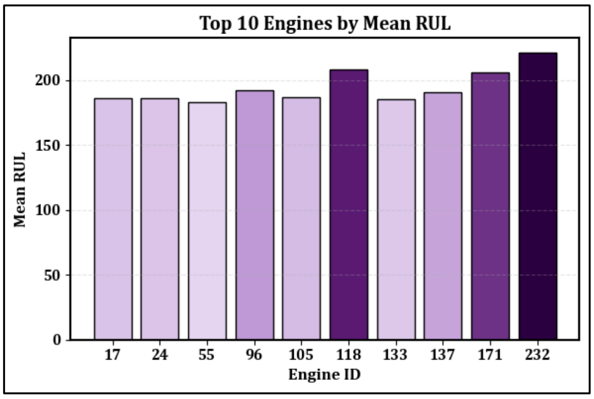

Figure 10: *Mean RUL*

The ten engines with the greatest mean RUL are shown in Figure 10. This alteration highlights the consequence of reliable temporal modelling for precise RUL prediction and signifies the various degradation behaviour among turbofan engines.

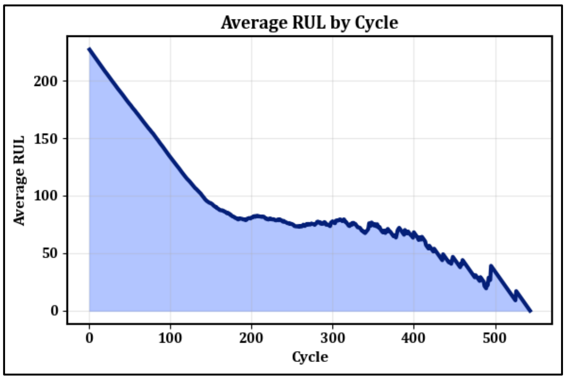

Figure 11: *Average RUL*

The average RUL variation across operating cycles is presented in Figure 11. This pattern provides a vital foundation for temporal RUL learning by the AROA-MTN and sufficiently demonstrates the increasing deterioration behaviour of turbofan engines.

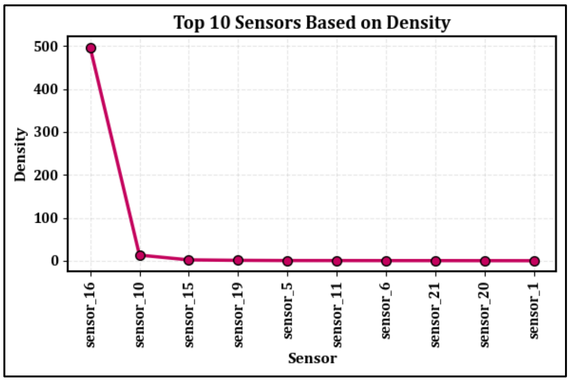

Figure 12: *Density for top 10 sensors*

The ten sensors are depicted in Figure 12 based on their estimated data density. Sensor 16 has a significantly higher density than the other sensor variables. The other chosen sensors offer slightly low-density observations with sensors 10 and 15 following with lower density values.

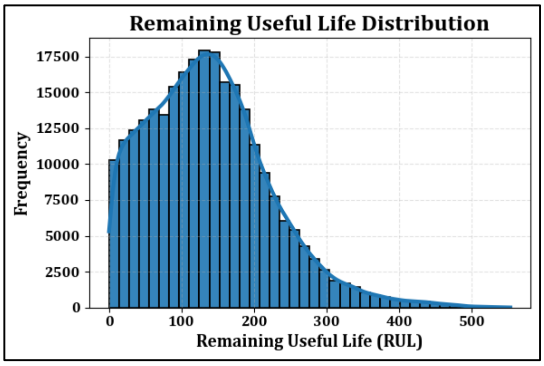

Figure 13: *Distribution of RUL*

The total distribution of RUL values is represented in Figure 13. With fewer observations stretching in the direction of higher RUL values beyond 300 cycles, the distribution displays a positively skewed pattern.

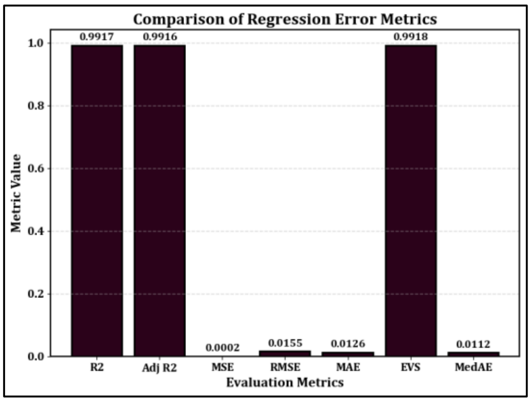

Figure 14: *Regression performance metrics*

The primary regression performance metrics obtained by the developed model are depicted in Figure 15. The model attaining 0.0155 RMSE, 0.9918 EVS, 0.9916 adjusted R^2 0.0002 MSE, 0.9917 R^2, 0.0126 MAE and 0.0112 MedAE, reveals an exceptional predictive accuracy and low estimate error.

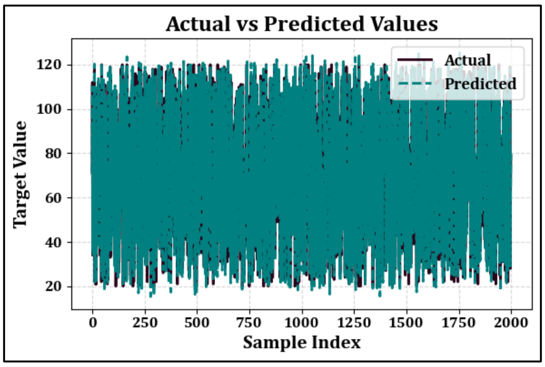

Figure 15: *RUL values*

The actual and predicted RUL values are represented in Figure 15, which validates substantial overlap amid the two sequences across the evaluation range.

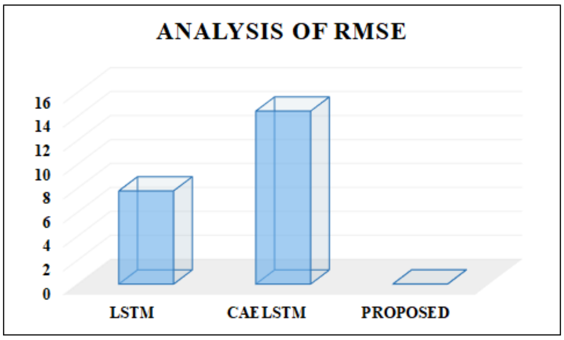

Figure 16: *Analysis of RMSE*

An analysis of RMSE is presented in Figure 16, which reveals the AROA-MT attains the lowest RMSE of 0.0155 than LSTM [7] and CAELSTM [9] approaches.

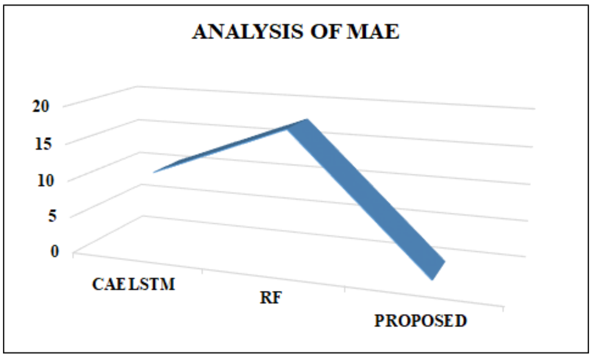

Figure 17: *Analysis of MAE*

Figure 17 depicts the MAE analysis of AROA-MT, RF [8] and CAELSTM [9] approaches. Here, AROA-MT attains the lowest MAE of 0.0126.

5. Conclusion

To precisely anticipate the RUL of aircraft turbofan engines, this paper develops an AROA-MTN Network. From the gathered engine operating data, the preprocessing stage makes it easier to recognise substantial sensor properties and degradation trends. By detecting pertinent data related to progressive engine degradation, feature engineering enhances the quality of the input representation. The developed model learns both global temporal linkages and long-term dependencies within multimodal sensor sequences owing to the combination of Mamba and Transformer architectures. The developed model's prediction accuracy and dependability are evaluated through performance evaluation utilising suitable error measures. Thus the AROA-MTN method shows the promise of fusing sophisticated sequence-learning methods with optimisation inspired by nature for aircraft engine prognostics. Repetitive model training is essential for ARO-based hyperparameter optimisation, which might raise training time and computing costs. To assess the AR-MTN model's capability for generalisation, future studies concentrate on testing it across numerous datasets, operating conditions and cross-domain situations. Explainable AI (XAI) and uncertainty quantification are employed to further expand the model to produce interpretable and confidence-aware RUL forecasts.

References

1. Unnati Thakkar; Hicham Chaoui, Year: 2022, MDPI "Remaining useful life prediction of an aircraft turbofan engine using deep layer recurrent neural networks", In Actuators, Vol: 11, No. 3, pp. 67.
2. Zhengyang Fan; Wanru Li; Kuo-Chu Chang, Year: 2024, Sensors, "A two-stage attention-based hierarchical transformer for turbofan engine remaining useful life prediction", Sensors, Vol: 24, No: 3, pp. 824.
3. Sizhe Deng; Jian Zhou, Year: 2024, "Prediction of remaining useful life of aero-engines based on CNN-LSTM-attention", International Journal of Computational Intelligence Systems, Vol: 17, No: 1, pp. 232.
4. Wenting Zha; Yunhong Ye, Year: 2024, "An aero-engine remaining useful life prediction model based on feature selection and the improved TCN", Franklin Open, Vol: 6, pp. 100083.
5. Jiusi Zhang; Jilun Tian; Pengfei Yan; Shimeng Wu; Hao Luo; Shen Yin, Year: 2024, "Multi-hop graph pooling adversarial network for cross-domain remaining useful life prediction: A distributed federated learning perspective," Reliability Engineering & System Safety, Vol: 244, pp. 109950.
6. Yıldırım Özüpak, Year: 2026, "A hybrid CNN–GRU approach for robust and explainable remaining useful life prediction of turbofan engines", International Journal of Energy Research, Vol: 2026, No: 1, pp. 7022062.
7. Owais Asif; Sajjad Ali Haider; Syed Rameez Naqvi; John FW Zaki; Kyung-Sup Kwak; SM Riazul Islam, Year: 2022, "A Deep Learning Model for Remaining Useful Life Prediction of Aircraft Turbofan Engine on C-MAPSS Dataset", IEEE Access, Vol: 10, pp. 95425-95440.
8. Hairui Wang; Dongwen Li; Dongjun Li; Cuiqin Liu; Xiuqi Yang; Guifu Zhu, Year: 2023, "Remaining useful life prediction of aircraft turbofan engine based on random forest feature selection and multi-layer perceptron", Applied Sciences, Vol: 13, No: 12, pp. 7186.
9. Samiha M. Elsherif; Bassel Hafiz; M. A. Makhlouf; Osama Farouk, Year: 2025, "A deep learning-based prognostic approach for predicting turbofan engine degradation and remaining useful life", Scientific Reports, Vol: 15, No: 1, pp. 26251.
10. Xiaofeng Liu; Liuqi Xiong; Yiming Zhang; Chenshuang Luo, Year: 2023, "Remaining useful life prediction for turbofan engine using SAE-TCN model", Aerospace, Vol: 10, No: 8, pp. 715.
11. Lihua Shen; Yucheng Wang; Baorui Du; Hailong Yang; He Fan, Year: 2025, "Remaining useful life prediction of aero-engine based on improved

GWO and 1DCNN", Machines, Vol: 13, No: 7, pp. 583.

12. Hairui Wang; Dongjun Li; Ya Li; Guifu Zhu; Rongxiang Lin, Year: 2024, "Method for remaining useful life prediction of turbofan engines combining adam optimization-based self-attention mechanism with temporal convolutional networks", Applied Sciences, Vol: 14, No: 17, pp. 7723.