

# TURBOFAN ENGINE HEALTH STATUS PREDICTION WITH ARTIFICIAL NEURAL NETWORK

Slawomir SZRAMA  , Tomasz LODYGOWSKI 

*Aviation Division, Poznan University of Technology, Piotrowo 3, 60-965 Poznan, Poland*

## Article History:

- received 3 May 2024
- accepted 27 August 2024

**Abstract.** The main purpose of this study is to present the concept of the aircraft turbofan engine health status prediction with artificial neural network augmentation process. The main idea of engine health status prediction is based on the engine health status parameter broadly used in the aviation industry as well as propulsion technology being the performance and safety margin. As a result of research engine health status index is calculated in order to determine the engine degradation level. The calculated parameter is then used as a response parameter for the machine learning algorithm. The case study is based on the artificial neural network which was two-layer feedforward network with sigmoid hidden neurons and linear output neurons. Network performance is evaluated using mean squared error and regression analysis. The final results are analyzed using visualization plots such as regression fit plot and histogram of errors. The greatest achievement of this elaboration is the presentation of how the entire process of engine status prediction might be augmented with the use of an artificial neural network. What is the greatest scientific contribution of the article is the fact that there are no scientific studies available, which are based on the engine real-life operating data.

**Keywords:** aircraft turbofan engine, health status prediction, artificial neural network, prognostic health monitoring, engine diagnostics and health monitoring.

 Corresponding author. E-mail: [slawomir.szrama@put.poznan.pl](mailto:slawomir.szrama@put.poznan.pl)

## 1. Introduction

Aircraft turbofan engine is the most popular aircraft propulsion type in the modern air transport. However, aircraft operation, reliability and safety strongly depend on the engine health status and its condition. Modern aircraft turbofan engine is a very complex construction which is controlled with very sophisticated electronic engine computer called FADEC (Full Authority Digital Engine Controller). This electronic device is trying to schedule engine operating parameters on the basis of the surrounding atmospheric conditions and signals provided by engine sensors. These sensors measure many signals like engine spool rotations, position of the engine actuators as well as pressure and temperature in several engine vertical cross-sections. All the engine data could be divided into the categories. Signal categories and types sensed and recorded during turbofan engine operation are as follows:

### Temperature Inputs

- Engine Inlet Total Temperature (Tt2),
- Compressor Exit Temperature (TT3),
- Fan Turbine Inlet Temperature (FTIT).

### Speed Inputs

- Low Rotor Speed (N1),
- High Rotor Speed (N2).

### Pressure Inputs

- Engine Inlet Static Pressure (Ps2),
- Burner Pressure (Pb, Pt4, or Ps3),
- Augmentor Inlet Total Pressure Mixed (Pt6m).

### Position and movement signals

- Compressor Inlet Variable Vanes resolver position transmitter (CIVV),
- Rear Compressor Variable Vanes resolver position transmitter (RCVV),
- Convergent Engine Nozzle Controller position transmitter (CENC).

### Aircraft Flight Condition Inputs

- Mach Number (Mn or Mo),
- Power Lever Angle (PLA).

The main object of the conducted research is the modern turbofan engine. This power plant is a low bypass, high compression ratio, dual spool, turbofan engine incorporating a mixed flow augmentor. The engine is of an advanced design with high thrust-to-weight ratio.

In this article, an effort was taken to define and determine engine health status model which is based on two of the recorded parameters representing engine aging and wearing. Step-by-step procedure was created which presents Engine Health Status Prediction algorithm taking advantage of the Machine Learning. On the example of

the F100-PW-229 engine, elaborated model and algorithm were used to train and test machine learning algorithms.

## 2. Literature review

As for the literature, there are some publications concerning engine health status prediction and there are also some elaborations concerning engine fault detection, but only very few propose the idea of artificial neural network implementation into engine condition prediction, especially that there is limited access to real life aircraft engine operational data, which could be used to build the model of prediction. For instance, Huang et al. (2016) proposed a prediction method of nonlinear time series analysis with the simulation example and validation of the algorithm. Song et al. (2009) proposed an aeroengine performance parameters forecasting method using multiple linear regression model, in which the effect of multiple factors was considered. Ji et al. (2020) presented the method of how the Convolutional Neural Network (CNN) could be used in engine fault detection of hot components. Shi et al. (2012) on the basis of the research of complexity and non-linearity of aero-engine Exhaust Gas Temperature (EGT) system, proposed a regularization chaotic prediction model to build short time forecasting model of EGT. Machine learning for engine remaining useful life estimation and prediction were presented by Lan et al. (2018), Liu et al. (2021), Thakkar and Chaoui (2022) or Zheng et al. (2017). Machine learning applications in engine faults detections and predictions were presented by: Almasi (2016), Brotherton et al. (2000), De Giorgi et al. (2023), Liu et al. (2018), Lu et al. (2014), Saxena et al. (2008), Xiangyang (2019), Song et al. (2009) or Zhang et al. (2022). The other neural network applications could be also found in the literature. Wang et al. (2023) proposed a model for engine Remaining Useful Life (RUL) that is entirely based on the attention mechanism. The attention model is divided into the multi-head self-attention and timing feature enhancement attention models. Zhang et al. (2022) proposed a dual-task network structure to realize engine health status assessment and RUL prediction, which was based on bidirectional gated recurrent unit (BiGRU) and multi-gate mixture-of-experts (MMoE). Some of the elaborations are based on the engine health status index, while for instance Chen et al. (2022) proposed data-driven method which does not require to know the physical nature of degradation mechanism, and used a bidirectional long short term memory (Bi-LSTM) network to construct the health state evaluation model. Ensemble learning has been used by Cheng et al. (2023) for aircraft engine RUL prediction by combining multiple methods with diverse weights to achieve better prediction accuracy and universality. Researchers have introduced ensemble learning-based prognostic approaches to model the degradation process of aircraft engines and predict their remaining useful life. Huang et al. (2022) proposed a dynamic probability (DP) model and a long short-term memory neural network (LSTM) to estimate the remaining useful life (RUL) of the engine. Tirovolas and

Stylios (2022) proposed Fuzzy Cognitive Maps (FCMs) as a Health Indicator (HI) prognostics method for engines RUL prediction. Another type of neural networks proposed for the engine health status prediction was proposed by Liu et al. (2023). Residual life prediction model based on Autoencoder and a Temporal Convolutional Network (TCN) was proposed. Wang et al. (2023) discussed how the multilayered perceptron network with random forest feature selection could be used to assess engine remaining useful life. Peng et al. (2022) proposed a dual-channel long short-term memory (LSTM) neural network model to adaptively select the time feature and then perform first-order processing on the time feature value and use LSTM to extract the time feature and first-order time feature information.

## 3. Problem description

Nowadays, aircraft engines are usually maintained in accordance with the condition-based maintenance strategy. It means that it is possible to continue engine flight operations as long as its condition, health and performance comply within the designed range. What is the background and motivation of the study? It results out of the over 25 years of experience in the air force and aviation industry. One of the most crucial decisions maintenance managerial personnel must make is answering the question: is it still safe to continue engine flight operations, or should the engine be removed from the aircraft, before it fails. Too early engine removal decreases efficiency of engine useful life usage, and it is not financially effective. In addition to this, it results in grounding aircraft, mission capability rate degradation and reduces aircraft fleet operations. On the other hand, continuing engine flight operations while it should be removed from service might result in aircraft accident or even catastrophe. That is why it is extremely crucial to predict the moment when it is absolutely necessary to stop engine operations. All airline operators would like to take advantage of the engine useful life up to the maximum. It also helps airliners in the planning process of the engine replacement, repairs and overhauls, which are extremely expensive and usually require spare engines which are not available at any moment.

There have been many attempts in aircraft engine health status predictions. Still there are not too many models of prediction to be implemented into the engine preventive maintenance strategy and system. One of the reasons for this is the limited access to the real engine operational flight data, which allows us to analyze this data and use this as the input data into the created artificial neural network. This would allow us to train the neural network on the real operational data which comes from different regions of the world, various year seasons, variety of the in-flight atmospheric conditions as well as conditions at the ground level. Engine simulated data is not very reliable data as it does not take under consideration different engine health conditions and in-time degradation. This definitely affects engine data with the engine wearing and deterioration. This is why it is a great opportunity to

take advantage of the real engine operational data which was collected for several years both in flights and on the ground during engine runs and tests.

One of the problems is the engine health status parameter determination. This data would be the crucial parameter in predicting the whole engine condition and degradation. In this article one of the possible ways of engine health status prediction was presented.

### 4. Research methodology

In Figure 1 there has been presented the proposed algorithm of the neural network application in aircraft engine health status prediction.

#### 4.1. Engine Health Status Prediction algorithm taking advantage of the neural network

The main purpose of this algorithm is to design, train designed neural network architecture to distinguish (differentiate) between healthy operation (“safe operation”) and two other engine health status conditions (“middle” and “unsafe”).

Algorithm of the engine health status prediction consist of the following steps:

1. Engine data exploration (identify what are the signals sensed as well as what is the data which was recorded),
2. Analyze the engine data and preprocess it (analyze engine parameters, what they comprise, which data is absolutely necessary),
3. Extract data features, which engine sensed param-

- eters are crucial for engine degradation assessment,
4. Select, define or determine engine parameter which could be used to evaluate engine health status index (in our case Health Index results from Performance Margin (PMAR) and Specific Fuel Consumption Margin (SMAR) engine parameters),
5. Setting Engine Health Status Classes based on the PMAR, SMAR (“safe”, “middle”, “unsafe”),
6. Split the whole data into training set and testing set (usually the 2/3<sup>rd</sup> of the whole set is taken as a training data and 1/3<sup>rd</sup> of the testing data),
7. Design neural network, shallow feedforward network with sigmoid hidden neurons and linear output neurons,
8. Train engine training dataset using all possible algorithms,
9. Compare learning algorithms, check not only what is the learning accuracy but also what is the cost of the misprediction. What is the crucial misprediction is the fact that for the engine which has reached unsafe engine health condition, the prediction value is safe,
10. Test the designed neural network architecture on the separate engine testing dataset,
11. Classify the whole engine fleet into three categories: Safe, Middle and Unsafe,
12. Export the most optimal training algorithm and implement neural network architecture into the maintenance strategy software.

The main idea of the case study as well as the neural network architecture have been presented in Figure 2. This case study methodology was based on the Two-layer

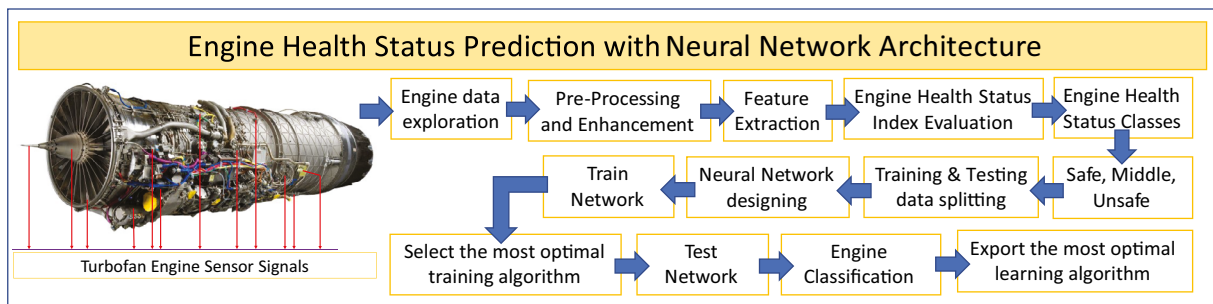


Figure 1. Neural network engine health status prediction algorithm

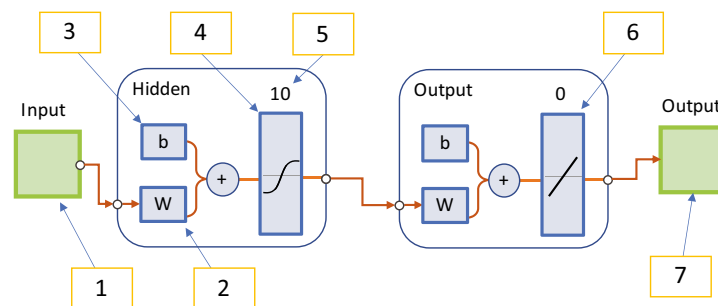


Figure 2. Two-layer feedforward network with sigmoid hidden neurons and linear output neurons: 1 – input data, 2 – network weight, 3 – network bias, 4 – Sigmoid hidden neurons, 5 – number of sigmoid hidden neurons, 6 – linear output neurons, 7 – output data

feedforward network with sigmoid hidden neurons and linear output neurons. The reason for this was the fact that usually there is no exact linear relation between engine parameter input data and the resulting output data. If the activation function is linear, then no matter how many hidden layers in the neural network you generate, the final output is still a linear combination of the original input data. That is why the sigmoid hidden neurons function was selected to create this artificial neural network.

If neuron activation function would be noted as  $z$ , this could be calculated in accordance with Equation (1):

$$z = \sum_{i=1}^m w_i x_i + b, \quad (1)$$

where:  $z$  – neuron activation function,  $m$  – number of input features (engine input parameters),  $w_i$  – the following weight,  $x_i$  – the following input data,  $b$  – bias.

Then the sigmoid function  $\sigma$  could be noted in accordance with Equation (2):

$$\sigma(z) = \frac{1}{1 + e^{-z}}. \quad (2)$$

## 4.2. Engine data exploration, preprocessing and features extraction

Engine performance data exploration relies on identification of the signals sensed by the engine sensors and analysis what kind of data is being recorded. On the basis of this analysis, it is possible to determine which signals represent engine health status and how to calculate engine degradation level and what is more important which engine parameters to predict while assessing engine deterioration and health condition.

Engine performance data was collected for the aircraft turbofan engine type. Engine data collection selected for this research consisted of the real-life engine operation data. This amount of data has been collected for 12 years of engine flight and ground operations. In result there were 29999 records uploaded into the artificial neural network. Each record comprised 46 engine data parameters. What is really important is the fact that all the records were sensed at the same engine thrust level requirement being the maximum augmentation level.

As a result of the artificial neural network structure presented in Figure 2, the whole engine data set was divided into three parts. 70% of the whole set (20999 observations) were randomly assigned as Training Data. 15% of the whole data set (4500 observations) were selected for the validation process and the rest 15% (4500 observations) were dedicated to the test process of the neural network.

Each artificial neural network was trained, validated and tested with three different algorithms being: Levenberg-Marquandt, Bayesian Regularization and Scaled Conjugate Gradient. The first one is working quite fast, the second one is working slower but generalizes well and the third one is very well in memory managing and efficiency. In addition to this, each neural network was trained, vali-

dated and tested with different number of hidden layers. For each algorithm three deep learning processes were performed with 10, 20 and 50 hidden layers.

## 4.3. Engine health status parameter selection

On the basis of the domain knowledge and experience it was determined that there were two engine parameters which could be appropriate to build engine health status index. The reason why PMAR parameter was used was the fact that it is one of the crucial engine operation parameters. Exceeding EGT allowed limits could result in a severe engine damage and in the worst-case scenario aircraft catastrophe. What is even more interesting it is the fact that this parameter is a very good indicator of the turbine efficiency decrement and engine degradation. In time of the engine operation, engine efficiency decreases, which means that to achieve the same level of thrust more fuel is needed in the combustion chamber. More fuel sprayed into the combustion chamber results in higher exhaust gas temperature. When engine parts become worn or suffer damage then parts like the turbine blades can not harvest the energy from the hot compressed air expanding as efficiently as they should. As a result, engine has less power being taken from the turbine to the compressor, which leads to a slower speed or less thrust.

In order to counteract this, the Engine Control Unit (ECU) adds more fuel in, so more energy is created, and the turbine extracts the required torque to maintain thrust. This will cause EGT rise as more fuel is being burned to create the same level of thrust or core operating speed. EGT is an important factor in engine trending and monitoring to indicate engine issues and that is why it was selected to calculate PMAR parameter which is going to be used to predict engine health status. Engine parameter called performance margin PMAR reflects the difference between  $EGT_{redline}$  and  $EGT_{max}$ . During engine operations all engine and aircraft data is being sensed and collected by so called Full Authority Digital Engine Control. All the engine data which reflects engine thermodynamic processes is sent to Engine Diagnostic Unit (EDU), where PMAR is calculated. This calculated PMAR parameter was used to design Engine Health Status Index in accordance with the concept presented in Figure 3b.

In Figure 3a Engine degradation range presented by EGT change was presented. In Figure 3b Engine Health Status Index Model calculated for the PMAR parameter and engine degradation was presented.

As it might be deduced from the description above it is the fact that the lower the PMAR is the more degraded engine condition is. Where it is close to or reaches zero value, it might lead to engine construction damage and result in aircraft severe accident or even catastrophe.

$$PMAR \cong EGT_{redline} - EGT_{max}, \quad (3)$$

where:  $EGT_{redline}$  – marginal temperature of the hot gases in engine which results in engine part damage;

$EGT_{max}$  – maximum operational temperature which is reached in normal flight operations.

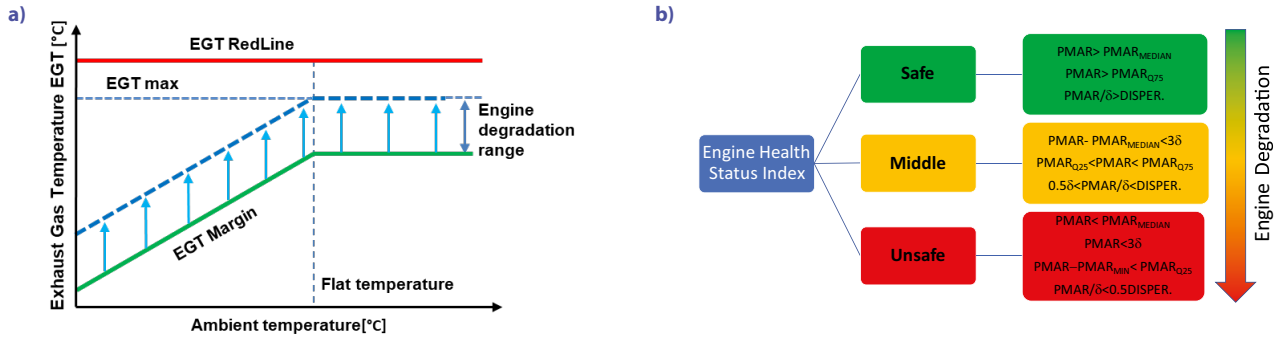


Figure 3. Engine degradation concept: a – presented by EGT change, b – Engine health status index model

This parameter is used as a parameter to be predicted with the artificial neural network. That is why, this is going to be selected in the MATLAB software as a response data.

Another engine data parameter used to design engine health status index was SMAR. Why SMAR data could be used to define engine health status and condition? This is the data which tells us what the fuel consumption rate in relation to the engine thrust (power) is. In reality, this parameter is measured as the relation of the fuel mass flow rate provided to the engine combustion chamber and engine thrust generated while converting this amount of fuel into the heat energy. It is evident that specific fuel consumption rate increases when the engine compressor compression rate decreases. If the engine compressor compression rate decreases in time for the same power level requirement, it means that engine compressor efficiency has degraded.

#### 4.4. Network design

As presented in Figure 2 there has been two-layer feed-forward artificial neural network created. The number of sigmoid hidden neurons was changing and during case study it was set at three different levels: 10, 20 and 50. The input data was the engine performance data which was collected for the aircraft turbofan engine type. Engine data collection selected for this research consisted of the real-life engine operation data. Engine performance data is an extremely significant data, as it is collected during aircraft take-offs. This is the moment when engine is operating at the highest power or thrust and very close or at the parameter margin. At this moment engine suffers the highest thermal and mechanical loads which might result in engine degradation or even failure. Each record of the performance data usually means one flight. In order to be able to train neural network plenty of engine operational data is required. As an output data prediction of the selected and calculated performance parameter is achieved, which identifies engine health status. However, the main goal would be to assess how well the created artificial neural network is working in predicting engine parameter value. How close the predicted values are to the actual target data.

In order to do this, there is a need to train artificial neural network in engine health status prediction. In order to do this one of the applications from MATLAB Works 2023 was selected. This application is called Neural Net Fitting and could help in solving neural network fitting problems using two-layer feedforward networks. With this application it is possible to create, visualize and train generated network to solve data fitting problems. There has been MATLAB live script written which allowed to import engine data from previously prepared file. Then this data set was split into training, validation and test sets. As a following step training neural network was defined using three selected algorithms: Levenberg-Marquandt, Bayesian Regularization and Scaled Conjugate Gradient. The number of hidden layers was set at three different levels; 10, 20, 50. As a result of the research, created network performance was evaluated using mean squared error and regression analysis. The final results were analyzed using visualization plots such as regression fit plot and histogram of errors.

#### 4.5. Artificial neural network performance measures

The mean squared error MSE of the trained neural network is measured with respect to the testing samples. This will give us a sense of how well the network will do when applied to another real turbofan engine operational data. MSE is calculated in accordance with Equation (4).

$$MSE = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2, \tag{4}$$

where:  $N$  – number of observations,  $e_i$  – the following error in the prediction,  $t_i$  – the following target value,  $y_i$  – the following predicted value.

Comparison of the MSE and number of epochs required to achieve the best performance for each algorithm and the size of the artificial neural network is presented in Table 1.

From Table 1 it could be easily deduced how the size of the layers (the number of hidden layers) affected the whole neural network performance. It might be concluded that in all cases except Bayesian Regularization, the increment of the layer size improved the Mean Squared Error for both Training, Validation and Test. The best MSE was



**Table 1.** Mean Squared Error MSE and number of epochs required to achieve the best performance for each algorithm and the size of the artificial neural network comparison (source: Authors' own elaboration, 2024)

Performance metric / Training length	Neural network life cycle phase	Leverent-Marquandt			Bayesian Regularization			Scaled Conjugate Gradient		
		Number of hidden layers								
		10	20	50	10	20	50	10	20	50
MSE	Training	11.9774	8.6299	5.2376	6.1771	4.2623	2.8425	64.0062	58.9715	44.5310
	Validation	14.0952	13.1410	7.1607	NaN	NaN	NaN	67.2213	62.8588	48.3333
	Test	14.6726	9.8665	6.8064	7.3819	7.0726	12.4208	67.2904	66.1934	48.5074
	<i>Additional Test</i>	–	–	25.7897	–	–	95.3641	–	–	99.2125
Number of Epochs	Validation	50	55	30	–	–	–	135	105	235
	Training	–	–	–	765	1000	1000	–	–	–

achieved for the Bayesian Regularization based network. However, in this case the Validation process resulted in NaN (Not a Number) results. About 50% lower accuracy was achieved for the Leverent-Marquandt based neural network. However, in this case the Validation process was fully accomplished and confirmed. In addition to this, this artificial network worked quite well in additional test achieving 25.7897MSE. This is the reason why this model-based network was selected to be exported and tested on another set of data. The highest (worst) MSE was achieved for the Scaled Conjugate Gradient algorithm base neural network. The results were 4–8 times higher than for the other two networks. However, in additional test the final MSE was close to the one achieved in the Bayesian Regularization algorithm neural network.

Another aspect of the artificial neural network training, validation and testing is comparison of the number of the epochs required to achieve the best training, validation and testing performance (in this case MSE). In Table 1 the number of epochs required to achieve the best performance for each algorithm and the size of the network were presented. The maximum number of epochs set for each type of the algorithms and the size of the network was 1000. As it was presented in Table 1 for the Bayesian Regularization type of the algorithm for 20 and 50 hidden layers the process of training reached maximum number of epochs. It is worth noting that for the Leverent-Marquandt type algorithm, the best validation performance, meaning the lowest MSE was achieved after 50, 55 and 30 epochs, adequately for 10, 20 and 50 hidden layers.

As it was presented in Table 1, for Bayesian Regularization algorithm the number of epochs was provided only for the training since for the validation the network output data was NaN. Generally, the error reduces after more epochs of training, but might start to increase on the validation data set as the network starts overfitting the training data. The training stops after six consecutive increases in validation error, and the best performance is taken from the epoch with the lowest validation error.

Another comparison of the results might be performed by comparing R-squared results. R-squared is the calculated by artificial neural network coefficient of determination. It is always smaller than 1 and usually larger than 0.

It compares the trained model with the model where the response is constant and equals the mean of the training response. In this case the closer this value to 1 is the better the neural network performance is. In the context of regression, it is a statistical measure of how well the regression line approximates the actual data. R-squared coefficient is calculated in accordance with Equation (5).

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^N (t_i - y_i)^2}{\sum_{i=1}^N (t_i - \bar{y})^2}, \quad (5)$$

where:  $N$  – number of observations,  $SSR$  – sum squared regression is the sum of the residuals<sup>1</sup> squared,  $SST$  – total sum of squares is the sum of the distance the data is away from the mean all squared,  $t_i$  – the following target value,  $y_i$  – the following predicted value,  $\bar{y}$  – the mean of the predicted value.

Comparison of the R-squared coefficient for each neural network as well as different number of hidden layers is presented in Table 2.

As it might be noticed in the Table 2 both Leverent-Marquandt and Bayesian Regularization based neural network are working very well. Even for the smaller 10-layer size it was possible to achieve 0.99 R-squared value. There was a slight improvement with the increment of the size of the hidden layers, but in general it was still very close to 1. Unfortunately, still the Bayesian Regularization based network was unable to perform validation process and in addition to this this network achieved lower R-squared value for the additional test. At the same time Leverent-Marquandt based neural network worked almost perfectly for both: Training, Validation, Test and Additional Test with the average result close to 0.99 of the R-squared. It may be concluded that selection of this type of neural network on basis of the Mean Squared Error was the best choice and it was also confirmed by the R-squared comparison.

Another measure of how well the neural network has fit the data is the regression plot. Here the regression is plotted across all samples. The regression plot shows the actual network outputs plotted in terms of the associated

<sup>1</sup> residual value = actual  $t$  value – predicted  $y$  value

**Table 2.** R-squared coefficient comparison of the artificial neural network performance (source: Authors’ own elaboration, 2024)

Performance metric	Neural network life cycle phase	Leverent-Marquandt			Bayesian Regularization			Scaled Conjugate Gradient		
		Number of hidden layers								
		10	20	50	10	20	50	10	20	50
R-squared	Training	0.9851	0.9892	0.9935	0.9923	0.9947	0.9965	0.9165	0.9247	0.9431
	Validation	0.9823	0.9838	0.9910	NaN	NaN	NaN	0.9165	0.9182	0.9392
	Test	0.9816	0.9879	0.9916	0.9911	0.9914	0.9849	0.9139	0.9153	0.9384
	<i>Additional Test</i>	–	–	0.9684	–	–	0.8875	–	–	0.8733

target values. If the network has learned to fit the data well, the linear fit to this output-target relationship should closely intersect the bottom-left and top-right corners of the plot. If this is not the case then further training, or training a network with more hidden neurons, would be advisable.

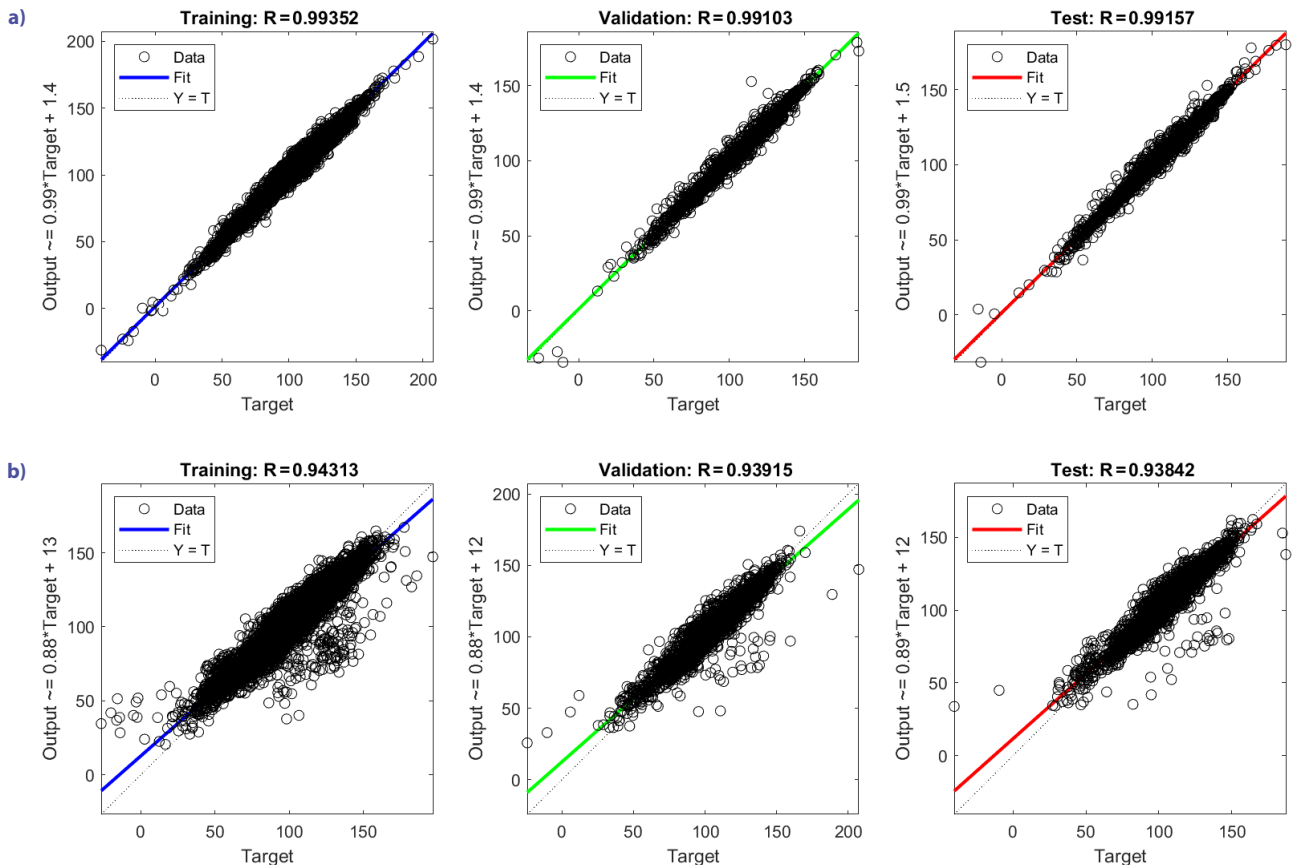
In the Figure 4 below there have been presented regression plot for all the trained validated and tested data.

The comparison of two main algorithms has been presented: 50-layer size neural network Leverent-Marquandt Figure 4a and Scaled Conjugate Gradient Figure 4b.

From the analysis of the plots presented in Figure 4 it might be quite easily deduced which artificial neural network and the type of the algorithm is fitting the prediction

the best. In each of the regression plots there is the line “Y = T”, which represents the situation when all the predicted values are the same as the target (real) values. The more output results are located around this line, the better is our neural network working. For all the processes: training, validation and test, network based on the Leverent-Marquandt type algorithm is working better all most of the predictions are located close to the real values.

Another, the fourth measurement of how well the neural network fits our data is the error histogram. This shows how the error sizes are distributed. Typically, most errors are near zero, with very few errors far from that. On the X-axis there are the errors values, which are the differences between the targets and output values meaning



**Figure 4.** Regression plots for the trained validated and tested data comparing two main algorithms: a – 50-layer size neural network Leverent-Marquandt, b – Scaled Conjugate Gradient



**Figure 5.** Error Histogram with 20 bins for 50-layer size neural network Scaled Conjugate Gradient

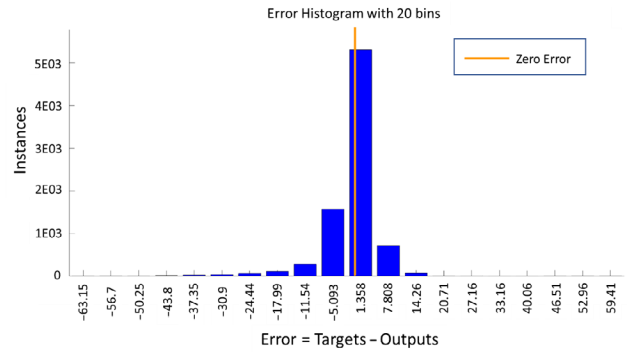
the actual values and predicted values. On the Y-axis are the numbers of instances being the number of samples (records) which fell into each bin. Each bin represents the range of the error. The whole maximum range is divided into 20 smaller ranges which are called "bins". The best result is when the highest number of samples fall into the bin which is close to "zero" axis vertical line (yellow color).

In Figure 5 Error Histogram with 20 Bins for 50-layer size neural network and Scaled Conjugate Gradient type algorithm was presented. As it might be easily deduced from Figure 5 was the fact that neural network based on this type of algorithm and with 50 hidden layers is working quite efficiently. For all the processes: training, validation and testing, the greatest numbers of the results fitted into the bin which was close to "Zero error" value (0.2912).

As a following step all the results were compared, and the comparison was presented in the Tables 1 and 2. Analyzing results from the Table 1 and 2 it was noticed that the best (lowest) Mean Squared Error and the R-squared coefficient was achieved for the Leverent-Marquandt algorithm. This was the reason why it was decided to export the best working artificial neural network as a script into the MATLAB.

#### 4.6. Artificial Neural Network performance confirmation

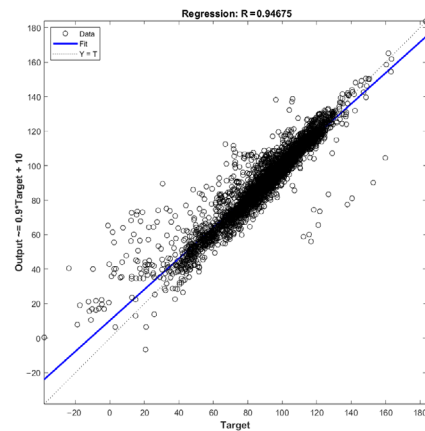
To confirm the results achieved for the first engine data set, there was another (completely) different data test was selected and prepared. The reason for this was the fact that the first training, validation and testing were performed for the same data and range of the parameters. They were just divided into three groups. The next data set was also the collection engine performance data which comprised the same features (engine sensor signals), but it was from a different time period and for the engine with a various engine operating hours. The training validated and tested network was used on the different engine test data which consisted of 12172 records and still 46 features. (For one selected and the best performing neural network. In this case this was the Leverent-Marquandt algorithm-based network). The results of the test were presented in Figure 6.



**Figure 6.** Error Histogram with 20 bins for 50-layer size neural network Leverent-Marquandt

As presented in Figure 6 over 5000 observations were assigned to the bin with 1.358 Error (target-prediction) value. For about 1700 observations error value was around 5.093 and for about 700 observations error value was around 7.808.

In Figure 7 regression plot for 50-layer size neural network based on the Leverent-Marquandt algorithm, for the additional engine data set, was presented.



**Figure 7.** Regression plot for the additional engine data set

As presented in Figure 7 being the regression plot for the additional engine data set, the output value  $\cong 0.9 \times \text{Target value} + 10$ , which gave the total R-squared value  $R = 0.94675$ . The following results were achieved for the previously trained, validated and tested artificial network, which was the 50-layer size neural network based on the Leverent-Marquandt algorithm. Mean square error MSE for this final test was calculated as 42.6760, while the R-squared coefficient was calculated as  $R = 0.94675$ .

## 5. Discussion and conclusions

This article is the introduction to the series of scientific studies which are aimed at Artificial Intelligence (AI) applications for the turbofan engine health status prediction. The whole research is based on the real-life engine operational data and is called data-driven engine health monitoring. The key element of the data-driven engine



health status prediction was the Engine Health Status Index designing. This is usually the most difficult and challenging part of the research. If the design health status index is very reliable it could be implemented into the engine maintenance strategy and planning. Such methodology could help propulsion maintenance management in engine big data analysis and avoid any situations when the engine health status degraded below acceptable level, especially for the large engine fleets. Engine health status prediction based on the Artificial Neural Network ANN is an extremely strong tool in aircraft accident and incident prevention. The next step of the methodology improvement is the application of the various architectures of deep neural networks.

As a result of the research case study, it was confirmed that artificial neural network is a very powerful tool which could be used to augment the process of prediction engine health status and its degradation. With fairly simple artificial neural network it was possible to predict engine health status parameter with a very high accuracy.

Analyzing the predicted values, it was discovered that some of the predicted parameters were marked as NaN. MATLAB represents values that are not real or complex numbers with a special value called NaN, which stands for "Not a Number". NaN (Not a Number) is a numeric data type that means an undefined value or value that cannot be represented, especially results of floating-point calculations. What might be the reason for this? After the in-depth analysis of the training data, it was found out that there was some missing data in the whole training data set, or some data was treated as a text type data. That is why the result of the prediction was not calculated properly. So, this is one of the crucial factors and very important conclusion from this article that trying to take advantage of the artificial neural network in predicting real operational engine health status, training and testing data must be properly configured and data which is corrupted or missing has to be removed from the set. The corrupted or missing data was also the reason why the Mean Squared Error was definitely higher for the test data in comparison with the training and validation data and the R-squared final value was lower.

The question might be raised, what is the advantage of the proposed shallow Neural Network (NN) predicting the engine health status parameter compared to some other parametric methods? Even though, NN performance results, in some cases, in comparison to the machine learning algorithms or other stochastic methods might be to some extent quite similar, there are still some significant factors in favor of NN.

The greatest advantage between artificial neural networks and other parametric methods is the fact that artificial neural networks work great with all types of data, either parametrical or categorical. It also works better for various numeric types, while machine learning algorithms might fail to converge. The reason for this is that neural networks work with diverse types of predictors. They work with both numerical data, vectors and the categorical data,

while ML algorithms (Discriminant and KNN types) in some case might fail or result in very "weak" (with low performance) prediction models.

In summary it is extremely important to mention the fact that, there is nothing more important than safety in aviation. And from this perspective it is crucial to predict the moment when it is still safety to fly the aircraft, to continue engine flight operations. If it was possible to predict the moment when it was needed to stop flight operations and to perform engine maintenance, it would allow to save a huge amount of money for the flight operators. The reason for this is the fact that airliners are able to take advantage of the engine useful life, up to the maximum, and what is also really important for airline operators is the planning process of the engine replacement, repairs and overhauls, which are extremely expensive and usually require spare engines which are not available at any moment.

Due to the fact that the modern turbofan engine is a very complex and sophisticated machinery, and it is very common and popular in air transport it is extremely important to develop engine health status prediction model. As a result of the conducted research and analysis, as a future work, implementation of the developed neural network architecture model into the airliners maintenance system might be proposed. After every flight engine data could be uploaded into the computerized maintenance system to update the engine remaining useful life. This could help in scheduling engine maintenance, repairs and overhauls and avoid any unnecessary maintenance. Such system reinforced by AI would be a great support for the airline operators.

## References

- Almasi, A. (2016). Latest lessons learned, modern condition monitoring and advanced predictive maintenance for gas turbines. *Australian Journal of Mechanical Engineering*, 14(3), 199–211. <https://doi.org/10.1080/14484846.2015.1093252>
- Brotherton, T., Jahns, G., Jacobs, J., & Wroblewski, D. (2000). Prognosis of faults in gas turbine engines. In *2000 IEEE Aerospace Conference Proceedings* (Cat. No.00TH8484, Vol. 6, pp. 163–171). IEEE. <https://doi.org/10.1109/AERO.2000.877892>
- Chen, C., Lu, N., Jiang, B., & Xing, Y. (2022). A data-driven approach for assessing aero-engine Health Status. *IFAC-PapersOnLine*, 55(6), 737–742. <https://doi.org/10.1016/j.ifacol.2022.07.215>
- Cheng, Y., Zeng, J., Wang, Z., & Song, D. (2023). A Health state-related ensemble deep learning method for aircraft engine remaining useful life prediction. *Applied Soft Computing*, 135, Article 110041. <https://doi.org/10.1016/j.asoc.2023.110041>
- De Giorgi, M. G., Menga, N., & Ficarella, A. (2023). Exploring prognostic and diagnostic techniques for jet Engine Health monitoring: A review of degradation mechanisms and advanced prediction strategies. *Energies*, 16(6), Article 2711. <https://doi.org/10.3390/en16062711>
- Huang, Q., Su, H., Wang, J., Huang, W., Zhang, G., & Huang, J. (2016). A prediction method for aero-engine health management based on nonlinear time series analysis. In *2016 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 1–8). IEEE. <https://doi.org/10.1109/ICPHM.2016.7542816>

- Huang, Y., Tao, J., Sun, G., Zhang, H., & Hu, Y. (2022). A prognostic and health management framework for aero-engines based on a dynamic probability model and LSTM network. *Aerospace*, 9(6), Article 316. <https://doi.org/10.3390/aerospace9060316>
- Ji, S., Han, X., Hou, Y., Song, Y., & Du, Q. (2020). Remaining useful life prediction of airplane engine based on PCA-BLSTM. *Sensors*, 20(16), Article 4537. <https://doi.org/10.3390/s20164537>
- Lan, G., Li, Q., & Cheng, N. (2018). Remaining useful life estimation of turbofan engine using LSTM neural networks. In *2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC)*. IEEE. <https://doi.org/10.1109/GNCC42960.2018.9019107>
- Liu, L., Wang, L., & Yu, Z. (2021). Remaining useful life estimation of aircraft engines based on deep convolution neural network and LightGBM combination model. *International Journal of Computational Intelligence Systems*, 14, Article 165. <https://doi.org/10.1007/s44196-021-00020-1>
- Liu, J., Liu, J., Yu, D., Kang, M., Yan, W., Wang, Z., & Pecht, M. (2018). Fault detection for gas turbine hot components based on a convolutional neural network. *Energies*, 11(8), Article 2149. <https://doi.org/10.3390/en11082149>
- Liu, X., Xiong, L., Zhang, Y., & Luo, C. (2023). Remaining useful life prediction for turbofan engine using SAE-TCN model. *Aerospace*, 10(8), Article 715. <https://doi.org/10.3390/aerospace10080715>
- Lu, F., Chen, Y., Huang, J., Zhang, D., & Liu, N. (2014). An integrated nonlinear model-based approach to gas turbine engine sensor fault diagnostics. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 228(11), 2007–2021. <https://doi.org/10.1177/0954410013511596>
- Peng, C., Wu, J., Wang, Q., Gui, W., & Tang, Z. (2022). Remaining useful life prediction using dual-channel LSTM with time feature and its difference. *Entropy*, 24(12), Article 1818. <https://doi.org/10.3390/e24121818>
- Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. In *2008 International Conference on Prognostics and Health Management* (pp. 1–9). IEEE. <https://doi.org/10.1109/PHM.2008.4711414>
- Shi, Y., & Yue, J., & Song, Y. (2012). Application of the regularization chaos prediction model in aero-engine performance parameters. *Advanced Materials Research*, 424–425, 347–351. <https://doi.org/10.4028/www.scientific.net/AMR.424-425.347>
- Song, Y., Zhang, K., & Shi, Y. (2009). Research on aeroengine performance parameters forecast based on multiple linear regression forecasting method. *Journal of Aerospace Power*, 24(2), 427–431.
- Thakkar, U., & Chaoui, H. (2022). Remaining useful life prediction of an aircraft turbofan engine using deep layer recurrent neural networks. *Actuators*, 11(3), Article 67. <https://doi.org/10.3390/act11030067>
- Tirovolas, M., & Stylios, C. (2022). Introducing fuzzy cognitive map for predicting engine's health status. *IFAC-PapersOnLine*, 55(2), 246–251. <https://doi.org/10.1016/j.ifacol.2022.04.201>
- Wang, H., Li, D., Li, D., Liu, C., Yang, X., & Zhu, G. (2023). Remaining useful life prediction of aircraft turbofan engine based on random forest feature selection and multi-layer perceptron. *Applied Sciences*, 13(12), Article 7186. <https://doi.org/10.3390/app13127186>
- Wang, X., Li, Y., Xu, Y., Liu, X., Zheng, T., & Zheng, B. (2023). Remaining useful life prediction for aero-engines using a time-enhanced multi-head self-attention model. *Aerospace*, 10(1), Article 80. <https://doi.org/10.3390/aerospace10010080>
- Xiangyang, Sh. (2019). Research on aero-engine fault diagnosis based on integrated neural network. *Mathematical Models in Engineering*, 5(2), 41–47. <https://doi.org/10.21595/mme.2019.20636>
- Zhang, Y., Xin, Y., Liu, Zh.-w., Chi, M., & Ma, G. (2022). Health status assessment and remaining useful life prediction of aero-engine based on BiGRU and MMoE. *Reliability Engineering & System Safety*, 220, Article 108263. <https://doi.org/10.1016/j.ress.2021.108263>
- Zheng, S., Ristovski, K., Farahat, A., & Gupta, C. (2017). Long short-term memory network for remaining useful life estimation. In *Proceedings of the 2017 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 88–95). IEEE. <https://doi.org/10.1109/ICPHM.2017.7998311>