Development of a combined Artificial Neural Network and Principal Component Analysis technique for Engine Health Monitoring

M G De Giorgi¹,*, L Strafella¹, N Menga¹, A Ficarella¹

¹ University of Salento. Department of Engineering for Innovation, Lecce, 73100 Italy

*maria-grazia.degiorgi@unisalento.it

Abstract. In aerospace sector, reliability is a crucial point. Modern technologies widely use Artificial Intelligence (AI) algorithms together with detections by sensors in order to design a health-based maintenance plan which stops an aircraft only when needed. In this work, an Engine Health Monitoring (EHM) system was developed by exploiting AI algorithms as Artificial Neural Networks (ANNs) trained to estimate a series of Performance Parameters (PPs) used as index of the health status of the main components constituting an engine. A neural network called Feed-Forward Neural Network (FFNN) in combination with a Principal Component Analysis (PCA) for feature reduction was used in this paper. The software Gas turbine Simulation Program (GSP) was used to generate a series of data containing information about engine performance under different flight conditions and compressor degradation levels. The datasets were subsequently used to train the neural networks to estimate the PPs of the degraded component. The final purpose of the present work is to develop an efficient diagnostic system useful to increase flight safety and decrease maintenance costs and fuel consumption.

1. Introduction
Nowadays, diagnostic is widely used in order to increase reliability of systems. Especially for aircraft engines, a high reliability is required in order to make the flight safe. Diagnostic is a very useful tool for reliability purposes and in the past years different techniques were developed [1]. In the past maintenance plans were based on the number of flight hours and an aircraft was stopped after a predetermined threshold was reached [2]. However, such a strategy is not obviously based on health status of the engine, leading to a less efficient maintenance plan. Degradation is a stochastic event which can be characterized by many factors, such as mission performed, external environment or casual events as ingestion of debris or birds. In [3] and [4] are illustrated respectively the Ingested Debris Monitoring System (IDMS) and the Engine Distress Monitoring System (EDMS) used to observe the amount of debris ingested in a jet engine while, regarding to environmental impact, two studies are reported respectively in [5] on the effect of dust and sand and in [6] on the effect of volcanic ash on engines. Regarding the degradation phenomena, the events mentioned above (ingestion of external bodies and flight in low purity air) have a non-negligible impact and a maintenance plan based on flight hours that does not consider them could lead to a premature maintenance or to an engine failure in flight. EHM discipline is nowadays extremely exploited in order to optimize the maintenance plan. To do this, EHM techniques use a health-based approach, using AI algorithms which in turn use information providing by a series of sensors installed through the powertrain to monitor values as temperatures, pressures, air
mass flow and fuel mass flow and rotational speed to detect any fault in one or more components constituting the engine. Do to their performance, computer algorithms are widely used in diagnostic and prognostic tools. Genetic Algorithms, Decision Trees and Fuzzy Logic are used in [7] for the prediction of the Remaining Useful Life (RUL). In [8] Support Vector Machine and ANNs are used for performance prediction and diagnostic purposes. The relationship between Exhaust Gas Temperature (EGT) and other variables used as input of EHM system is obtained in [9] by using MultiGene Genetic Programming. The EGT thus obtained was subsequently used to create a dataset for the one-step-ahead prediction by means of Nonlinear AutoRegressive with eXogenous inputs neural network. In [10] Long Short-Term Memory and Deep Belief Network are used for RUL estimation, fault classification, sensor prediction and condition assessment. Specific fuel consumption is estimated, in different flight maneuvers, in [11] exploiting two Nonlinear AutoRegressive Neural Networks. ANNs were trained in [12] and [13] in order to distinguish between two different degradation scenarios, i.e. compressor fouling and turbine erosion. The work presented here shows an EHM system developed by training a FFNN to estimate the PPs of the compressor of a turboshaft engine. The network was trained using dataset generated by GSP simulations performed by implementing the degradation in the analyzed component. The degradation was implemented by changing the value of the PPs in the engine model developed with GSP. The degraded values of the PPs used to implement the degradation during simulations were randomly obtained. There are many degrading phenomena that can occur on any engine component in a different way. The principal degradation phenomenon in a compressor is the fouling, i.e. a deposition of air contaminants and sand, dust, dirt and other on compressor blades and walls. This leads to an increase in roughness and change in aerodynamic shape [14] resulting in a decrease in flow parameter and efficiency, obtaining a reduction in performance [15]. Compressor fouling is the compressor degradation phenomena considered in this work. Some examples on the amount of variation of efficiency and air mass flow rate are reported in [16]. Due to their exposure to very high temperature capable of weakening the materials, turbines instead are typically exposed to erosion, i.e. a gradual loss in material from blades and case due to interaction with contaminants in combustion gases. The process leads to a change in aerodynamic shape of the blades [14] [16] and the effect is an increase in flow parameter and a decrease in efficiency [17]. For example, a 4% increased air mass flow rate accompanied with a 2% decreased efficiency is reported in [18]. Finally, a data reduction technique as PCA was used in order to create a reduced dataset subsequently used to train another FFNN with the same purpose of the previous trained FFNN. The results obtained from the two different FFNNs are compared with the aim to observe the impact on the EHM system efficiency of a data reduction technique.

2. Engine modelling

GSP is a popular and useful software used to simulate an engine operating under different conditions. More information about GSP software are available in GSP11 user manual [19].

2.1. Engine description: features and design point

The engine presented here is a lightweight two spool turboshaft engine used for helicopter applications part of the PW200 Pratt & Whitney Canada family. A single stage centrifugal compressor is driven by a single stage High Pressure Turbine (HPT) thanks to a high-pressure spool, while the low pressure spool connects a Low Pressure Turbine (LPT) to the rotor, which have a nominal speed of 6000 rpm. Figure 1 illustrates the block model developed with GSP drag and drop interface.

![Figure 1. Block model of the reported engine developed with GSP.](image-url)
The model was validated assuming the design point corresponding to standard sea level static condition (temperature = 288.25 K, pressure = 1.013 bar, altitude = 0 m, Mach = 0) by iteratively changing fuel flow and air flow rates until obtain the power indicated on engine reference database. Engine parameters in design condition are summarized in table 1.

### Table 1. Engine parameters in design point.

| Parameter                              | Value   |
|----------------------------------------|---------|
| Rotor shaft power ($P_r$)              | 295 KW  |
| Intake air mass flow rate ($w_1$)      | 2.05 Kg/s |
| Compressor pressure ratio ($\beta_c$)  | 7.17    |
| Compressor polytropic efficiency ($\eta_c$) | 0.825 |
| High Pressure Turbine polytropic efficiency ($\eta_{HPT}$) | 0.88  |
| Low Pressure Turbine polytropic efficiency ($\eta_{LPT}$) | 0.88   |
| Burner pneumatic efficiency ($\eta_{b,p}$) | 0.96 |
| Burner combustion efficiency ($\eta_{b,c}$) | 0.985 |
| Fuel mass flow rate ($w_f$)            | 0.0314 Kg/s |
| High Pressure Shaft speed ($\omega_{HPS}$) | 40891 rpm |
| Low Pressure Shaft speed ($\omega_{LPS}$) | 6000 rpm |

### 3. Simulations: build of a dataset containing data on degraded engines

In this section the component chosen to be degraded is illustrated together with the related PPs used as its health index.

#### 3.1. Considered scenario for component degradation implementation

The degradation condition considered in this work is the compressor degradation. Table 2 reports the PPs considered for the examined component together with the maximum and minimum limits between which degraded values of PPs were generated and their values in healthy condition at design point.

### Table 2. Performance Parameters associated with the degraded component.

| Component       | Degraded performance parameters | Healthy condition | Range of percentage variation |
|-----------------|---------------------------------|-------------------|--------------------------------|
| Compressor      | Compressor polytropic efficiency ($\eta_c$) | 0.82              | -20% ÷ 0                      |
|                 | Compressor flow capacity ($f_c$)      | 1                 | -20% ÷ 0                      |

#### 3.2. Obtaining the dataset: simulation details

The simulations were performed considering six different flights working conditions, reported in table 3 and taken from the mission profiles shown in [20].

A set of 4500 steady-state simulations were performed, 750 for each working point. The related dataset contains the results of the just mentioned simulations, 3000 of the 4500 simulations are related to training dataset (to train the neural networks) while the remaining 1500 are related to test dataset (to test their performance). Some of these simulations have been discarded as they relate to extreme working points where the software is unable to converge. In each simulation, degraded state was simulated by changing the value of the PPs of the component degraded. The degraded values of the PPs were obtained in randomly manner by using equation (1).

\[
pp_d = \left[ 1 - \left( 1 - \frac{\text{MIN}}{\text{MAX}} \right) RAN^3 \right] \text{MAX}
\] (1)

\[
pp_d = \left[ 1 - \left( 1 - \frac{\text{MIN}}{\text{MAX}} \right) RAN^3 \right] \text{MAX}
\] (1)
where \( p_{\text{deg}} \) is the degraded value of the PP, \( \text{MIN} \) and \( \text{MAX} \) are the lower and the upper limits for the PP as obtainable from table 2 and \( \text{RAN} \) is a random value in the interval \( 0 \div 1 \). In figure 2 the distribution of the generated PPs with equation (1) is reported. Table 5 shows an example of the effect on Power Turbine Exit Temperature (PTET) and \( w_f \) of the degradation condition considered. The data in table 4 is related to the working condition MAXPOW-B. The effect of the degradation process is an increase in PTET and \( w_f \), resulting in a more severe working environment for turbine (which degrades quickly) and in increased consumptions.

### Table 3. Flight operating conditions.

| Working point name | Speed (m/s) | Altitude (m) | Power (kW) | Low pressure turbine speed (rpm) |
|--------------------|-------------|--------------|------------|----------------------------------|
| START-A            | 30.6        | 0            | 48         | 6000                             |
| START-B            | 0           | 1150         | 173        | 6000                             |
| MAXPOW-A           | 30.6        | 0            | 291        | 6000                             |
| MAXPOW-B           | 3.96        | 1154         | 273        | 6000                             |
| MAXDUR-A           | 30.6        | 492          | 169        | 6000                             |
| MAXDUR-B           | 0           | 2550         | 176        | 6000                             |

### Table 4. Impact on PTET and \( w_f \) of the degradation cases considered.

| Degradation case     | PP’s percentage variation | PTET percentage variation | \( w_f \) percentage variation |
|----------------------|---------------------------|---------------------------|-------------------------------|
| Compressor degradation | \( \eta_c: -4.88\%; f_c: -0.26\% \) | +5.91 %                  | +4.68 %                       |

3.3. Selected variables

Tables 5 summarizes the variables used as input for the EHM system, in the case of the original dataset (the one obtained from GSP simulations). The input variables are related to engine operating condition and flight conditions, while the output variables are the PPs to be estimated by means of neural networks, presented in table 2.
Table 5. Input variables for FFNN trained with original dataset (before using PCA)

| Inputs                                      | Total Temperature at HPT inlet (TT<sub>4</sub>) | Total Pressure at HPT inlet (TP<sub>4</sub>) | Total Temperature at HPT outlet (TT<sub>45</sub>) | Total Pressure at HPT outlet (TP<sub>45</sub>) | Total Temperature at LPT outlet (TT<sub>46</sub>) | Total Pressure at LPT outlet (TP<sub>46</sub>) | LPT torque (T<sub>LPT</sub>) | Overall Pressure Ratio (OPR) | Fuel mass flow rate/Power Lever Angle (ρ) |
|---------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Flight mach (M)                             |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Ambient Total Pressure (TP<sub>a</sub>)     |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Ambient Total Temperature (TT<sub>a</sub>)  |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Power Lever Angle (PLA)                     |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Total Temperature at compressor inlet (TT<sub>2</sub>) |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Total Pressure at compressor inlet (TP<sub>2</sub>) |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Total Temperature at compressor outlet (TT<sub>3</sub>) |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| Total Pressure at compressor outlet (TP<sub>3</sub>) |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |
| High pressure shaft speed (ω<sub>HPS</sub>) |                                               |                                               |                                               |                                               |                                               |                                               |                                 |                                 |                                 |

4. Neural Network setup
In diagnostic, ANNs are used to estimate some values of a determined variable (or more variables) starting from other known variables. Neural networks are characterized by two steps, training and testing. In training phase, the neural network is trained to learn the correlation between the variables to be estimated and the information (variables) given in input, while during testing phase the previously trained network is used to predict some known variables to test its performance. Neural networks are made by layers (one input layer, one or more hidden layers and one output layer) and neurons, which number influences the performance. The input layer has as many neurons as input variables and the same is for output layer and output variables. The number of hidden layers and related neurons (called hidden neurons) influences the performance of the network. The neural networks used in this work are FFNNs, one trained with the original dataset and one trained with the reduced dataset. Table 6 shows the setup used for the neural networks.

Table 6. Setup used for the neural networks.

| Input neurons | Hidden neurons | Output neurons | Hidden layers |
|---------------|----------------|----------------|---------------|
| FFNN without PCA | 18             | 36             | 2             | 1             |
| FFNN with PCA   | 3              | 10             | 2             | 1             |

5. PCA
PCA is a technique which transforms the original set of variables into a set of uncorrelated variables by determining the Principal Components (PCs) of the signal [21]. The features extracted by PCA have the same amount of information as the original set of variables [22] even if the number of PCs is less than the number of original set of data. Given a dataset \(X \in \mathbb{R}^{N \times M}\) where \(N\) is the number of variables and \(M\) is the number of observations with zero mean, PCA gives a mapping from the original space to a feature space using equation (2):

\[
Y = X^T Q
\]

where \(Y \in \mathbb{R}^{M \times a}\) is the score matrix, and \(Q \in \mathbb{R}^{N \times a}\) is the loading matrix, its columns being the \(a\) eigenvectors associated with the \(a\) largest eigenvalues of the covariance matrix of \(X\) [23]. The number of features \(a\) is less than \(N\) and is commonly so the ratio of sum of eigenvalues to the overall sum of eigenvalues is commonly equal to 98%.
6. Results and discussions
In figures 3a and 3b the comparison between the predicted and the target data from GSP simulations has been shown for the FFNN without PCA. A good agreement between the FFNN outputs and the real ones has been obtained, as confirmed by the coefficient of determination $R^2$ that assumes values higher than 0.9. The good results are confirmed also by figures 4a and 4b, which show the related percentage errors. Figures 3a and 3b and figures 4a and 4b underline how the working condition MAXPOW-B, which corresponds to a high-power condition in altitude, leads to worst performance of the FFNN. Finally, the PCA was applied for the reduction of the input dataset. The reduction of dataset size was performed discarding those PCs that contribute less than 2% to the total variation in the dataset, the dimensions of the dataset was reduced from 18 to 3 input variables (PCs), while the number of observations remained equal to the original one. Table 6 reports the features of the FFNNs used for the hybrid PCA-FFNN approach, in terms of number of neurons and hidden layers. Figures 3c and 3d underline that the PCA feature reduction is an important issue in the implementation of prediction algorithms for compressor degradation because it has no impact on the accuracy of the predictions even if it reduces the computational effort. The hybrid PCA-FFNN method leads to reduce the prediction error, as shown in figures 4c and 4d.

![Figure 3. Comparison between network predictions and targets (obtained from GSP simulations) for case without PCA (a,b) and with PCA (c,d). Grey: START-A; Brown: Start-B; Green: MAXPOW-A; Blue: MAXPOW-B; Black: MAXDUR-A; Magenta: MAXDUR-B.](image-url)
7. Conclusions

The aim of the present work is to develop a reliable EHM system with particular focus on compressor degradation. To reach this goal, GSP software has been used to perform a series of simulations useful to obtain data on degraded engines at six different flight working points. Furthermore, different degradation levels, i.e., with different percentage variation of the PPs respect to healthy condition have been considered. These datasets have been used for the implementation of FFNN neural networks combined to PCA for feature reductions. The neural network permits to estimate the compressor performance parameters based on inputs related to engine operating conditions (temperatures and pressures at different stations in the engine, fuel flow rate, high pressure spool rotational speed, LPT torque, overall pressure ratio) and flight conditions (power lever angle, Mach, ambient temperature, and pressure). The comparisons between predicted and target data lead to a high coefficient of determinations $R^2$ and low percentage errors for both FFNN without and with PCA. The analysis underlined the effect of the different flight’s conditions on the errors of the predictions. Results confirmed the suitability of the implemented FFNN for the prediction of the engine components degradation with low computation efforts due to the low dimensionality of the reduced input datasets. A reliable EHM system is a fundamental tool to obtain an optimized maintenance plan, resulting in higher flight safety and reduced consumptions and maintenance costs.

Acknowledgement

This research was funded by Italian Ministry of University and Research, project PON “SMEA”, code PON03PE_00067_5.
References

[1] Tahan M, Tsoutsanis E, Muhammad M and Abdul Karim Z A 2017 Appl. Energy 198 122-44
[2] Vujic D 2005 Diagnostic systems in aircraft engines maintenance
[3] Powrie H and Fisher C E 1999 Monitoring of Foreign Objects Ingested into the Intake of a Jet Engine
[4] Powrie H and Novis A 2006 IEEE Aerospace Conference 8
[5] DeRemer J 1949 J. Am. Soc. Nav. Eng. 62 502-11
[6] De Giorgi M G, Campilongo S and Ficarella A 2015 J. Eng. Gas. Turbines Power 137
[7] Ishibashi R and Nascimento C 2013 1-7
[8] De Giorgi M G, Campilongo S and Ficarella A 2018 Energy Procedia 148 860-67
[9] De Giorgi M G and Quarta M 2020 Aerosp. Sci. Technol. 103 105902
[10] De Giorgi M G and Quarta M 2020 Aerosp. Sci. Technol. 103 105423
[11] De Giorgi M G, Strafella L and Ficarella A 2021 Aerospace 8
[12] De Giorgi M G, Ficarella A and De Carlo L 2019 Aerosp. Technol. 92 296-303
[13] De Giorgi M G, Campilongo S and Ficarella A 2018 MATEC Web Conf. 233
[14] Sogut M Z, Yalcin E and Karakoc T H 2017 Energy 140 1417-26
[15] Asme T et al. 2011 J. Eng. Gas. Turbines Power 134
[16] Fentaye A D, Baheta A T, Gilani S I and Kyprianidis K G 2019 Aerospace 6
[17] Salar A, Hosseini S M, Zangmolk B R and Sedigh A K 2010 IEEE Xplore 240-47
[18] Qingcai Y, Li S, Yunpeng C and Zhao N 2016 V006T05A016
[19] GSP 11 User Manual https://www.gspteam.com/GSPsupport/OnlineHelp/index.html?documentation.htm
[20] Donateo T, de Pascalis C L, Strafella L and Ficarella A 2021 IOP Conf. Ser. Mater. Sci. Eng. 1024 012074
[21] Abdi H and Williams Wiley L J 2010 Interdiscip. Rev. Comput. Stat. 2
[22] Jolliffe I T 1986 Principal Component Analysis (New York: Springer-Verlag )
[23] Khormali A O and Aliyari Sh M 2014 Gas turbine fault detection and identification by using fuzzy clustering methods 70-5