Measurement 227 (2024) 114212

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Coupling principal component analysis-based sensor data reduction techniques and multi-net systems for simultaneous prediction of multi-component degradation levels in hybrid electric rotorcraft engines

Maria Grazia De Giorgi *, Teresa Donateo, Antonio Ficarella, Nicola Menga, Ludovica Spada Chiodo, Luciano Strafella

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

A B S T R A C T

Hybrid Electric Power Systems (HEPS) have gained popularity as a more efficient and eco-friendly alternative. However, with increasing system complexity, fault potential rises.

The work aims on implementing a diagnostic system for rotorcraft engine health within a hybrid-electric system. Health monitoring tools are still understudied for HEPS, so this work can represent a valid contribution in the literature. The main goal is assessing degradation and monitoring multi-component simultaneous degradation. Various machine learning techniques for Engine Health Monitoring (EHM) have been compared, varying in network architecture and data reduction.

A dynamic model of the entire HEPS generated a dataset including fault information. This dataset trained FFNNs to predict performance parameters (PPs) of degraded components from sensor data. The proposed EHM system's efficacy was evaluated by comparing neural network predictions to dynamic model data. Results show that the Multi-net architecture, with distinct networks for each PP, works more effectively reducing training time.

1. Introduction

Environmental pollution is a pressing issue, and researchers have explored eco-friendly and sustainable alternatives to conventional fuels [1,2]. In aeronautics, efforts to reduce environmental impact focus on aerodynamics, structures, materials, fabrication, maintenance, modern engines, and operational practices [3,4]. The drive to cut fuel consumption and pollution has spurred interest in alternative aircraft propulsion technologies and energy sources, moving away from relying solely on fossil fuels. Recent studies have seen a significant rise in exploring this topic [5–9]. Traditional aircraft propulsion uses combustion processes to extract energy from fossil fuels with high energy density. However, they exhibit low conversion efficiency and emit significant pollutants and greenhouse gases during combustion [10,11]. Electric propulsion eliminates local emissions using efficient electric machines as converters, powered by stored electricity in batteries, and managed with an appropriate strategy [12]. Transitioning from conventional to electric systems requires significant effort [13]. Current battery limitations in power and energy density make all-electric solutions unviable for large-scale aviation. Lithium-ion batteries are preferred for electric energy storage due to their characteristics: high output voltage, long lifespan, high energy density, low self-discharge rate, and wide operating temperature range [14]. Lithium-ion batteries' energy density increased by 2.5 times, with costs reduced by 81% since 2008 [15].

1.1. Hybrid electric power systems (HEPS) for greener propulsion and battery storage

To leverage strengths and address shortcomings, Hybrid Electric Power Systems (HEPSs) were initially developed for automobiles and more recently adopted in aviation [16] with piston or turbine engines. In aviation, HEPSs combine fuels and electricity, supplementing traditional turbine engine components with electric drives (electric motors and battery packs). This results in a more efficient propulsion system, reducing fuel consumption and environmental impact [17]. However, the process of hybridizing actual aircrafts is still a challenging task [18,19]. Compared to conventional systems, HEPS provides better fuel efficiency and enhanced safety during engine failure. However, the increased complexity raises reliability concerns, necessitating monitoring of component health and fault detection within the propulsion chain for safe and efficient operation. For informed decisions during condition-monitoring system design, assessing the impact of individual subsystems on reliability is crucial. In HEPS, batteries are susceptible to

* Corresponding author.
E-mail address: mariagrazia.degiorgi@unisalento.it (M.G. De Giorgi).

https://doi.org/10.1016/j.measurement.2024.114212
Received 31 October 2023; Received in revised form 23 December 2023; Accepted 20 January 2024
Available online 26 January 2024
0263-2241/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

health issues and aging effects, leading to safety concerns and permanent damages [20]. Issues arise from irreversible changes in the electrolyte, anode, cathode, and battery structure. These lead to reduced capacity, power, and increased Peukert effect. Key factors in battery degradation include environment temperature, discharge/charge rate, depth of discharge (DoD), and time intervals between full charge cycles [21]. Improper management of lithium-ion batteries can result in fires and explosions [22]. The Battery Management System (BMS) in large lithium batteries for propulsion monitors battery health, assessing State of Health (SOH) using measurable data like voltage and current [23]. Diagnostic methodologies are also being developed for batteries [24]. Remaining Useful Life (RUL) prediction is also used for lithium-ion batteries and is strongly based on the battery's capacity [25–28]. The operations of the BMS are strongly related to the information deriving from battery sensors [29,30]. A large variety of methods for the assessment of SOH of lithium batteries has been proposed in the scientific literature (see for example [31]). However, due to relative modernity of the system, health monitoring for the entire HEPS is still a relatively underexplored area, in contrast to the well-established health monitoring systems for thermal engines.

1.2. Performance deterioration and health monitoring

High-performance aircraft engines are exposed to extreme pressure, temperature, and speed conditions, which greatly increases the chance of failures [32]. The performance over time of a given engine depends on how efficiently its components work, as analysed in [33]. Degradation affects engine components and impacts mechanical, thermodynamic, and aerodynamic properties [34]. Fouling in compressors is a common problem caused by the buildup of air contaminants on blades and walls. This reduces flow passage area, changes airfoil shape, and lowers efficiency [35]. Turbines are prone to erosion due to blade and wall interaction with air contaminants and combustion gases, causing gradual material loss and changes in aerodynamic blade shape. This leads to decreased efficiency and increased flow capacity [36]. As an example, a fouled compressor can result in a 5% reduction in flow capacity and a 1.8% decrease in isentropic efficiency [37]. Turbine erosion may cause an overall performance loss of approximately 5% [38]. Sensors and actuators assume significant importance from this point of view [39]. In modern high-performance aircraft, thermal loads play an increasingly relevant role in health monitoring [40]. Fatigue and rubbing are crucial considerations for rotating components like compressor and turbine blades [41,42]. Engine Health Monitoring (EHM) systems are extensively used in aircraft engines to continuously monitor engine health and enable proactive maintenance before critical conditions arise. Furthermore, EHM is used not only to predict the current health status of the engine, but also for RUL prediction, which is one of the most commonly used methodologies [43–45]. Health monitoring methods are becoming more widespread, and increasingly efficient systems are being developed [46]. Early issue identification through health monitoring ensures safe and efficient aircraft operation, preventing costly downtime [47–49]. Using health monitoring for maintenance leads to more efficient plans, performing actions when needed, not just based on flight hours. Degradation varies probabilistically among engines due to flight missions and air quality, resulting in different levels of degradation for similar engines with the same flight hours. If degradation is slow, the engine might be stopped for maintenance before reaching a critical condition. In severe degradation cases instead, a breakdown could happen in flight before the flight hours limit, jeopardizing safety. Studies on sand and dust ingestion [50] and volcanic ash ingestion [51] reported greater erosion rates at the tips in both cases. Ref. [52] describes typical issues in aircraft electric motor batteries.

1.3. Machine learning in health monitoring

Machine learning techniques are commonly used in diagnostic and prognostic tools [53]. Combining multiple computer algorithms can also yield good results, resulting in a complex and useful model [54]. In [55] a study similar to the one proposed in this paper is reported but focused on a pure thermal engine. In the just cited study, Artificial Neural Networks (ANNs) are exploited to perform performance prediction by predicting Exhaust Gas Temperature (EGT) and fuel flow rate (w_f) and to perform an engine fault prediction, considering three different degradation conditions, i.e. compressor fouling, turbine erosion and compressor fouling and turbine erosion together. Ref. [56] implemented ANNs and Support Vector Machines (SVMs) for performance prediction and classification in the same three degradation scenarios. ANNs excel in prediction, while SVMs perform better in classification, showcasing the strengths and limitations of each method. Choosing the right technique is crucial for specific tasks. Ref. [57] presents a two-step tool using ANNs. The first step predicts measurable parameters using flight data. If discrepancies arise between predicted and measured data, the second step detects degraded components in the propulsion system. Ref. [58] employs Levenberg-Marquardt Feed-Forward Neural Network (FFNN) and Radial Basis Function Network for exhaust gas temperature prediction in EHM applications. Results shown better prediction performance obtained from the first machine-learning technique. Ref. [59] proposes a method using transfer learning coupled with classical gas path analysis for health monitoring. This approach allows the gas path analysis model to adjust its training dataset, ensuring consistent diagnostic accuracy in life-cycle monitoring. Ref. [60] introduces a novel offline extension of a simple Wiener model. The approach, called hybrid Wiener model, is used to estimate some relevant parameters such as thrust, surge margin and turbine entry temperature and is furthermore periodically updated to match the new conditions of a degrade engine. The fault detection model presented in [61] is based on the isolation forest algorithm, which, as reported in the cited work, has the advantage of being able to process large-scale and high-dimensional data effectively. It has also shown good results when anomalies are not available in the training sample. The results showed high accuracy and a short running time. Fault detection and isolation is performed by the LSTM-DAE (Long Short-Term Memory-Denoising AutoEncoders) model presented in [62]. In particular, detection is carried out using the LSTM component, while the DAE is used for fault isolation. Furthermore, fault detection is crucial not only for engines. For example, fault diagnosis for an electric rudder is reported in [63–64], for a Q-pulse diode electrifier in [65] and for actuators in [66]. HEPS component degradation and failure lack the comprehensive understanding, reliability data, and health monitoring models that conventional thermal aircraft propulsion systems possess. The limited availability of data and models for EHM in HEPS creates a knowledge gap in predicting and preventing failures. To enhance HEPS reliability and safety, investing in research and development for accurate data and models in EHM is crucial.

The EHM technique developed in this paper uses virtual sensors throughout the powertrain to gather engine health data. Pre-trained ANNs process these sensor measurements to predict engine degradation. In more details, this study explores using FFNNs to estimate Performance Parameters (PPs) of degraded components in HEPS for helicopters. The FFNN models were trained and tested using datasets from simulations with a Matlab & Simulink-based model of the entire HEPS. The study examines a degradation scenario where several components, such as compressor, burner, mechanical transmission, electric motor, and battery, degrade simultaneously. The data obtained from simulations aim to simulate the data acquired in a real-case scenario using sensors. In this study, the effects of measurement noises are not taken into consideration in each scenario, but a related analysis with results is presented in Section 5.2. Data availability for degraded engines is vital for EHM system development. Nonetheless, some efforts have been made to detect faults in gas turbines even in the absence of historical fault data [67]. ANNs were trained to estimate PPs (or their Correction Factors - CFs) for degraded components, such as isentropic efficiency and flow parameter for the compressor, pneumatic and

combustion efficiency for the burner, mechanical efficiencies for spools, efficiency for the electric motor, and internal resistance for the battery. These parameters serve as health indices, indicating component degradation levels. Different EHM systems with various architectures were developed. The simplest approach involved selecting a set of simulation parameters as input variables for FFNNs to predict the PPs. Advanced approaches included data reduction techniques like Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA).

The choice of using a FFNN for developing the goal systems is attributed to its simplicity. A FFNN, in fact, is one of the simplest types of ANNs and is particularly suitable for regression problems. Unlike the more complex Recurrent Neural Network, in an FFNN information flows in one direction without returning to the input, necessitating a simpler network architecture. Further details about FFNN are provided in Section 4.1. PCA was used to perform Feature Selection (FS), while KPCA performs Feature Extraction (FE), which are two different data reduction approaches. Data reduction is a useful tool for eliminating redundant and unnecessary data, which can lead to a decrease in processing time and an improvement in the prediction performance of FFNNs when utilizing the reduced dataset. More details are available in Section 4.2. Two architectures, "Single-net" and "Multi-net," were considered. The choice between them depends on specific requirements. Multi-net suits cases with distinct and unrelated outputs, allowing independent optimization for better overall performance. A multiple output neural network (Single-net architecture) is preferable when outputs are related and share an underlying structure. It allows joint learning and optimization of shared features, leading to better performance and more efficient training. Both approaches are effective in predicting hybrid electrical engine performance, with the choice depending on analysis requirements. When choosing between Single-net and Multi-net architectures, consider dataset size, complexity, output variables, and desired accuracy and efficiency. This study compares the two architectures coupled with dimensionality reduction methods. Single-net uses one FFNN for all PPs, while Multi-net has separate FFNNs for each PP. Performance was evaluated by comparing predicted values to target values from simulations. Results from each EHM system were compared to assess their effectiveness. Fig. 1 illustrates the comprehensive technical roadmap of the current study.

1.4. Work goals

The goal of the present work is to create a reliable EHM system for HEPS and understand how different approaches and architectures impact prediction performance. Health monitoring techniques for HEPS, like those presented here, are less explored in the literature. The main contributions of this study are summarized as follows:

- The study of a HEPS has significantly expanded the existing literature on EHM applied to hybrid propulsion, providing valuable insights for future research in this field.
- The development of a new EHM method capable of assessing the real-time performance of a HEPS in the presence of simultaneous degradation of all components under different operating flight conditions is a significant contribution to the field, and its ability to handle increasing complexity and nonlinearity in performance diagnosis is noteworthy.
- The study's assessment of the impact of engine degradation on the HEPS's performance, particularly on fuel consumption and economic considerations, is a valuable contribution to the literature on hybrid propulsion.
- The application of PCA and KPCA, typically used in literature with pure thermal engines, to a complex HEPS for data reduction purposes demonstrates the versatility and applicability of these methods to a wider range of propulsion systems, including very complex types such as HEPS, which, in addition to the classical components of a thermal engine, incorporate the components of an electrical one, serving as a potential source of faults, providing useful insights for future research in this field.
- Focus revolves around a comparative analysis of two distinct cutting-edge techniques: feature extraction and feature selection, both

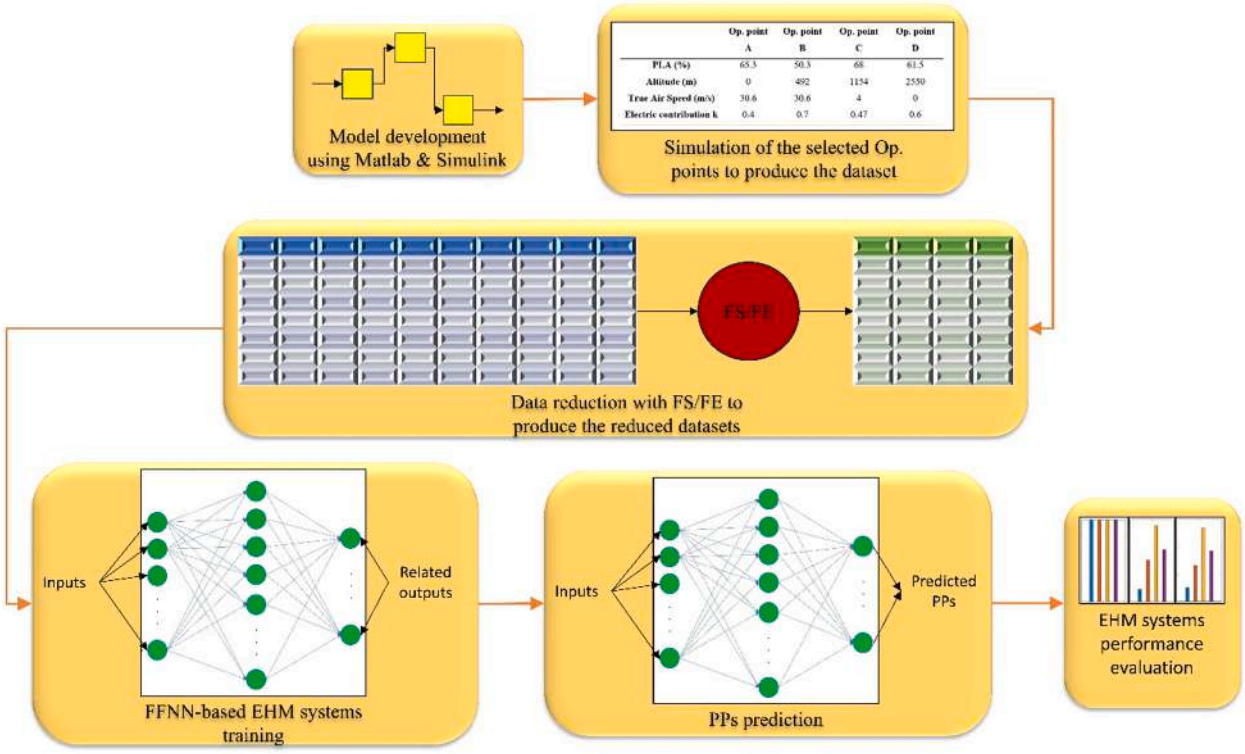

Fig. 1. Overall technical roadmap of the current study.

integrated with Artificial Neural Networks (ANNs). Specifically, the study delves into enhancing accuracy and optimizing computational time within this domain.
- The comparison centers on evaluating the efficiency and effectiveness of employing Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA) in feature extraction and feature selection processes.
- The ultimate objective is to optimize the operation, enhance efficiency, and ensure the safety and reliability of rotorcraft hybrid-electric propulsion through a fusion of cutting-edge technology and aerospace engineering expertise.

In Section 2 of the present work, the Simulink model of the HEPS is illustrated and its functioning explained. Section 3 is dedicated to the description of the simulations performed and of the degradation implemented. The systems are described in Section 4, together with the techniques used to develop them (FFNN, PCA and KPCA) and with the error metric adopted to estimate their performance. The results are shown and discussed in Section 5 and finally the conclusions of the present work are reported in Section 6.

2. The hybrid electric power system

The present investigation considers a HEPS in a parallel configuration, designed for a VTOL rotorcraft for Urban Air Mobility (Fig. 2). The system is composed of a turboshaft engine mechanically coupled to two identical electric motors to move the rotor shaft. A parallel configuration is preferred over the series-hybrid because of the lighter arrangement resulting from the elimination of the generator, the smaller size of the electric motor that is not required to satisfy the entire power demand, and the higher efficiency that in series configuration is penalized by the chain of conversion processes [68].

The size of the engine has been selected to match the peak power demand experienced during typical operating conditions (295 kW); the electric machines and the battery are sized for electric back-up operation in case of engine failure (for more details please read [69]). A rated power of 120 kW and a nominal revolutionary speed of 6000 rpm are assumed for all machines. Therefore, the proposed HEPS system has an overall degree of hybridization of 0.45.

A Li-ion battery is used as a storage device on the electric path. The suitability of such kind of batteries for electric vehicles and hybrid electric vehicles is commonly assumed, given their higher energy density, reduced self-discharge, long calendar life, and increased safety [70] than other commercial types of batteries. However, the details of the machines and of the battery are not reported for the sake of confidentiality.

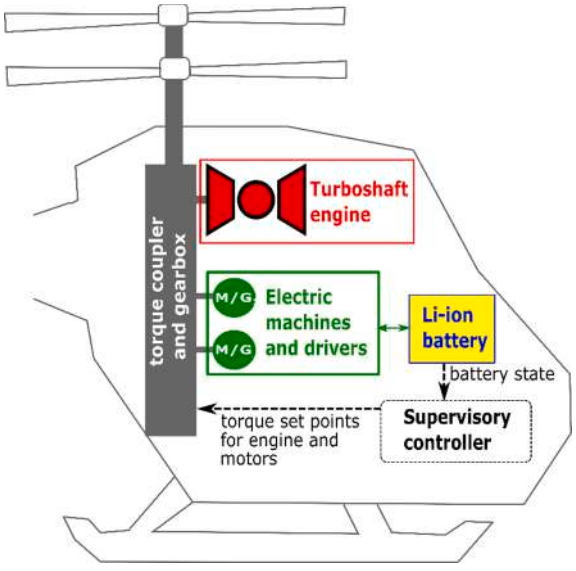

Fig. 2. Parallel HEPS for coaxial rotor vehicle.

The modelling approach used in the development of the Simulink model is explained in the next subsection.

2.1. Modelling approach and energy management strategy

The modelling approach implemented in the Matlab & Simulink environment is shown in Fig. 3. A pilot interpreter block translates the Power Lever Angle (PLA) input into a power request, based on a 3D map which includes altitude and Mach effects. The resulting power demand in turn is given in input to the supervisory controller module, which sets the power split between thermal and electric power sources at any time instant. This is regulated by means of the control parameter k, which defines the electric contribution as the ratio of electric to total torque request and is called power split:

$$k = \frac{Q_{EM}}{Q_{req}} \tag{1}$$

In the present investigation, the selection of k depends on the kind of simulation performed for the degradation scenario as reported in Table 1.

2.2. Turboshaft engine

The engine model takes as input the required torque Q_{ICE} and calculates air and fuel flow rates, gas-generator speed and thermodynamical properties at each station and finally actual power output on the Low-Pressure Spool (LPS). The turboshaft (Fig. 4) is a dual spool engine modelled with a 0D approach, consisting of a High-Pressure Spool (HPS) connected to the compressor and the High-Pressure Turbine (HPT), and a LPS with the power turbine (or Low-Pressure Turbine (LPT)). The Simulink dynamic model is simulated by solving mass flow rate balances between each component and work balance on the HPS. At the current stage, LPS dynamics is neglected, since its speed is assumed constant and the dynamic rotor model is not considered yet. For interested readers, a full dynamic model of the hybrid electric power system has been implemented in recent papers of some of the authors [72,73].

The inlet determines a pressure rise as expressed in the following equation with a constant efficiency $\eta_r = 0.9$, although the limited flight speed makes it negligible:

$$\Delta p = \eta_r \frac{1}{2} \rho V_0^2 \tag{2}$$

where Δp represents the pressure rise, ρ is the air density and V_0 is the air velocity at inlet.

The combustion process is simulated in the burner block to calculate the combustor exit temperature from the fuel flow rate according to the following equation:

$$T_4 = T_3 + \frac{\eta_{b,c} H_i}{c_p \alpha} \tag{3}$$

where $\eta_{b,c}$ is the combustion efficiency, H_i is the lower heating value, and α the air–fuel ratio. c_p is averaged among air and burned gas values of constant pressure specific heat, which, in turn, are polynomial functions of the temperature.

The burner exit pressure is given by the following relation, given the compressor outlet pressure p_3 and the burner pneumatic efficiency:

$$p_4 = \eta_{b,c} \cdot p_3 \tag{4}$$

The healthy values of combustion and pneumatic efficiency are given later, together with the corresponding considered range of degradation.

The compressor and the two turbine modules are implemented in the

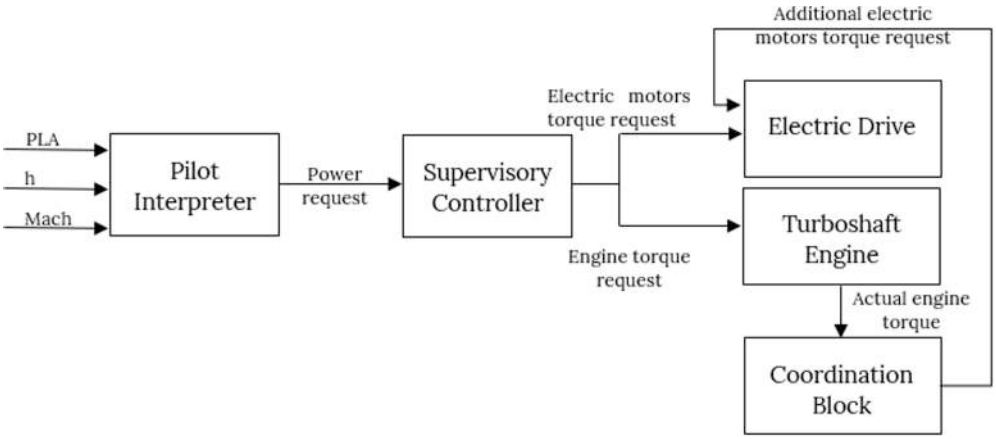

Fig. 3. Flowchart of the overall model.

Table 1
Design of experiments for the degradation analysis.

Operating conditions	Power Split k
Steady-state tests	Constant
Small perturbations	Constant
Mission analysis	Defined by a fuzzy logic supervisory controller previously developed in [71]

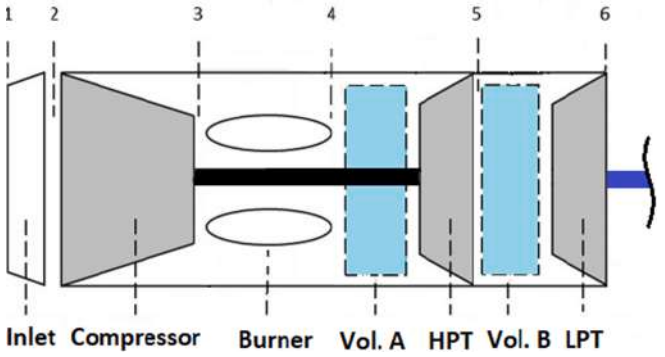

Fig. 4. Scheme of the turboshaft.

Simulink model by means of lookup tables which map the component pressure ratios and efficiencies as a function of the corrected mass flow rates and corrected speeds. Each rotating component has been mapped by previously scaling the available maps of a reference turboshaft model included in GSP libraries, to match the power requirements of the current application. Transient phases are modelled through HPS acceleration and mass accumulations in inter-component volumes (marked as vol. A and vol. B in). Such features make the model suitable to perform transient simulations [74]:

$$\dot{N}_{gg} = \left(\frac{30}{\pi}\right)^2 \frac{1}{IN_{gg}} (P_{HPT} - P_C) \tag{5}$$

where the HPS speed instantaneous variations $\dot{N}_{gg}$ are expressed as a function of current speed N_{gg} and power mismatch between compressor (P_C) and HPT (P_{HPT}). The symbol I in Eq. (5) is the moment of inertia around the axis of rotation. As previously mentioned, such balance was not applied to the LPS whose speed is currently considered constant, since the application of a dynamic load represents a further development of the model. The following mass flow rate accumulation equation models pressure rise in the plenum V_p on the basis of perfect gas equation and mass flow unbalance between inlet and exit sections of the plenum:

$$\frac{dp}{dt} = \frac{RT}{V}\frac{dm}{dt} = \frac{RT}{V}(\dot{m}_{in} - \dot{m}_{out}) \tag{6}$$

where R is the elastic constant of the gas, T is the fluid temperature, V is the volume of the plenum and $\dot{m}_{in}$ and $\dot{m}_{out}$ represent the mass flow rate of fluid at inlet and exit sections, respectively. The value of HPS inertia I, as well as the values of inter-component volumes V_p, have been derived from the previously cited reference GSP turboshaft. LPT exit pressure is mapped as a function of fuel flow rate with a correction to account for altitude effects. The fuel flow rate is governed by a PID controller which works on LPT power error, stated in other words it acts to keep LPT power output equal to shaft power demand. The PID gains have been tuned through Simulink automated tuning tool at the current stage, being an accurate characterization of fuel control and dynamics a further development.

2.3. Electric path

On the electric path, two electric motors equally contribute to Q_{Ems}, taking in input their own fraction of torque command and giving as output actual torque output and battery current as explained below.

The Simulink Mapped Motor block is used for the electric machine where the output torque tracks the torque reference command with a given time constant [75]. Being the motors connected to the same shaft as the LPT, their speed is constant too. Knowing motors speed and delivered power (P_{mech}) and power loss (P_{loss}), the battery current I(t) is calculated as:

$$I(t) = \frac{P_{mech} + P_{loss}}{V(t)} \tag{7}$$

This requires an iterative procedure, since the battery is modeled as an equivalent circuit characterized by voltage V(t):

$$V(t) = OCV - R_i \cdot I(t) \tag{8}$$

where OCV is the open circuit voltage mapped as a function of battery (SoC) and R_i is the battery internal resistance. Moreover, the battery behavior is modeled by including Peukert effect too, by means of a Peukert coefficient n which represents a further performance parameter of the battery. The Peukert effect is accounted for in the calculation of the SoC which is performed with the well-known coulomb-counting method but an effective current (I_{eff}) is used instead of the actual battery current [76]:

$$I_{eff} = I \cdot \left[\frac{I}{I_{nom}}\right]^{n-1} \tag{9}$$

where I_{nom} is the nominal current and I is the actual battery current. The effective current is used to include the Peukert effect on the battery discharge process. For a new and healthy battery, this effect is less relevant but as the battery ages or is degraded, the Peukert coefficient increases, and a significant reduction of electric range is observed. For more details, please read [69]. The effect of T is also taken into account for the evaluation of the SoC. A temperature dependence is modelled through the ratio of reference to current operating temperature and the exponent β, as suggested in [76]. T_{ref} is assumed to be 300 K and β is set equal to 2.9. Therefore, the state of charge is evaluated as:

$$SoC(t) = SoC(t_0) - 100 \cdot \int_{t_0}^{t} \frac{I_{eff}(t)}{C} \left(\frac{T_{ref}}{T}\right)^{\beta} dt \tag{10}$$

where t_0 is the initial time of the simulation, t is a generic time instant at which the SoC is being calculated, T is the operating temperature of the battery, T_{ref} is a reference temperature. The employed battery has a nominal (healthy) capacity of 130 Ah and a voltage of 270 V, and consists of 73 cells in series. To take into account its degradation, the main specifications of the battery (capacity, internal resistance and Peukert coefficient) are expressed as a function of the "cycle number", which is defined as the number of complete discharge-recharge cycles. A battery is conventionally said to have reached its end of life when the capacity reaches 80% of the nominal value. This usually happens, for a Li-ion battery, after 300–500 discharge-recharge cycles [77]. The model employed here is based on the results of [78], which experimentally characterized parameter variation over a LiFePO4 battery life. On the basis of the above-mentioned research, some of the authors have developed the following equations for the dependence of battery parameters on cycle number *N* [69]. For each parameter P of the battery (namely nominal capacity C, Peukert coefficient *n* and internal resistance R_i), a correction factor (CF) related to battery age can be defined as follows:

$$CF = \frac{P(N)}{P^0} \tag{11}$$

where the superscript 0 denotes the initial condition. The dependence of the correction factor CF on battery cycle number N is expressed as a double exponential:

$$CF = a \cdot exp(b \cdot N) + c \cdot exp(d \cdot N) \tag{12}$$

where the coefficients a, b, c, d have been found interpolating experimental results from [69] with least square error method. The fitting curves are depicted in Fig. 5 together with experimental data. Since battery end of life is conventionally set to 20 % capacity reduction with respect to the nominal value, the employed battery has a useful life of 436 cycles. In the developed model, the battery SOH is calculated as the ratio of undegraded internal resistance Ri^0 to current internal resistance $R_i(N)$ multiplied by a polynomial function of operating temperature. In a previous work of some of the authors [79], the impact of each parameter degradation on the main outputs of the model, namely the fuel flow rates, the battery residual energy, the battery voltage, and the speed of the HP shaft was analyzed by means of the Pearson product-moment correlation coefficient. This analysis showed that the performance of the engine is not affected by the degradation of the electric machines and vice versa. However, this is not entirely true when a mission-based analysis is performed. In fact, in this case, the degradation of the battery has an indirect effect on the behavior of the engine because the rules of energy management strategy reduce the electric contribution or even exclude it when the battery becomes degraded.

2.4. Validation issues

The lack of experimental data in the literature about hybrid electric power systems for rotorcraft makes it impossible to perform a full validation. However, each of the sub-models used for the powertrain was separately validated in the referred studies available in the scientific literature. Moreover, the approach used in this investigation is generally accepted in the analysis of hybrid electric propulsion systems for road vehicles [80]. In particular, the results of the dynamic model of the stand-alone turboshaft engine were assessed by comparison with the results of the same engine modelled in GSP [81]. The two models were run over the same dynamic inputs (Power Level Angle, Altitude, and True Air Speed) and the signals of brake power, flow rate, and compressor speed obtained with the two models were compared. The results confirmed the validity of the proposed model. As for the battery, the authors in [69] compared the results of the proposed electric circuit model with literature models validated against experimental data. The validation was performed using as input the power request (that was increased from 0 to the nominal power of the battery) and as output the discharge time. Therefore, the results of the numerical analysis can be considered, according to the authors of this investigation, reasonable and suitable for the goal of this investigation.

3. Datasets development: simulations and degradation scenario

The systems developed in this work require training on the physical correlation between the variables related to the engine operation and condition, which are used to inform the FFNNs, and the variables to be predicted. To achieve this, suitable datasets containing relevant variables are essential. However, obtaining a dataset containing real flight data can be quite challenging. One solution to this problem is to design a virtual model of the target engine using appropriate software and conduct a series of simulations from which useful data can be obtained. This section will describe the simulations performed and the datasets used in this work, highlighting the variables selected from the simulation results.

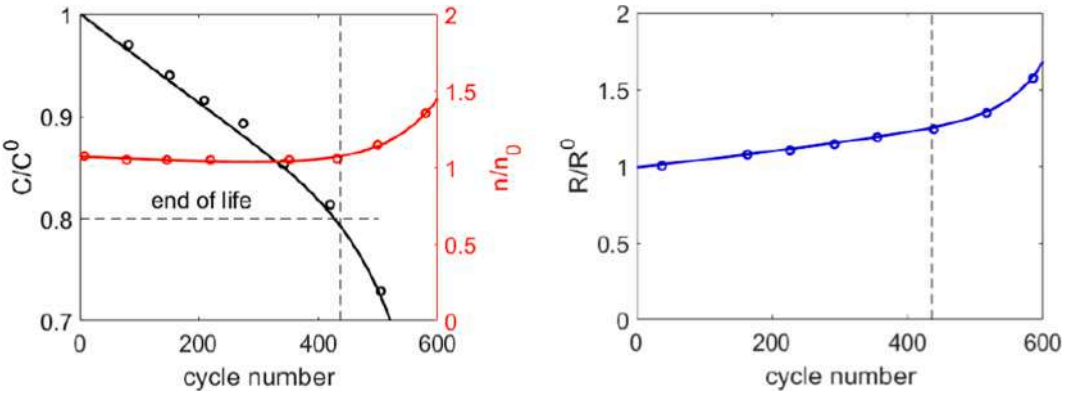

Fig. 5. Capacity, Peukert coefficient and internal resistance variation with battery age.

3.1. Degraded components

The degradation scenario presents the degradation of components related to both the thermal and electrical portions of the HEPS, including the compressor, burner, mechanical transmissions of LPS and HPS, electric machine, and battery. For each component, the degradation was implemented as a percentage variation of certain chosen PPs (or their CFs) that represent the health state of the components compared to their healthy condition. Table 2 lists the degraded components, along with the related PPs to be predicted, the values of the PPs in healthy condition, and the maximum and minimum limits of the considered percentage variations. It is important to note that the term "CF" refers to a change made to the PPs by applying a correction factor rather than by changing the parameter itself.

3.2. Simulation details

As previously stated, the objective of this paper is to develop an EHM system capable of predicting the values of PPs (or their CFs) under both healthy and degraded conditions. To achieve this, FFNNs were trained to recognize the specified parameters for each component based on sensor data from the engine and flight conditions. The datasets required for training the FFNNs were generated using the engine model described in earlier sections. The simulations were conducted using four different operating points. For each operating point, a series of steady-state simulations were run, each with a different level of percentage variation for the degraded parameters. Table 3 provides an overview of the four simulated operating points, referred to as Op. Point A, Op. Point B, Op. Point C, and Op. Point D. The just cited operating points were chosen following a logic of optimization of emissions obtained from the burn of the chemical fuel [82].

These operating points were taken from real missions, shown in Fig. 6 and taken from [83]. The data calculated from the simulations for each of the four operating points listed in Table 3 have been used to create the dataset. The mixing process helps to ensure a dataset composed of various scenarios on which the ANNs can be trained. The complete dataset consists of the results of 773 steady-state simulations, conducted at four different operating points and with different levels of component degradation. To configure the neural networks, 526 of the total 773 observations were selected for the training dataset, while the remaining 247 observations were used for the test dataset. This subdivision results from a sensitivity analysis conducted on the training dataset size, where its dimensions were varied until a specific stop criterion was achieved. During training, the neural network learns solely from the training set, while the testing set is employed to calculate the testing error, providing an approximation of the generalization error. The stopping criterion is established such that, when the testing error begins to rise, or equivalently, when the training error and testing error start to diverge, the optimal number of observations in the training dataset has been reached. Table 4 provides additional information about the training and test datasets.

As previously mentioned, the degraded values of the PPs (or their CFs) were randomly generated within a chosen range, determined by the maximum and minimum values of the percentage variations listed in Table 2. The equation below was used to calculate the degraded values of the PPs:

$$pp_d = \left[1 - \left(1 - \frac{MIN}{MAX}\right) RAN^3\right] MAX \tag{13}$$

where pp_d is the degraded value of the PP, *MIN* and *MAX* are respectively the lower and the upper limits considered for the PP for which the equation is being used and *RAN* is a randomly generated number in the interval 0–1. Eq. (11) was used to calculate the degraded value of each degraded PP for each of the 526 training points and 247 testing points. Fig. 7 shows the distribution of the PPs (or their CFs) generated with Eq. (11). In particular, each graph in Fig. 7 shows the number of values that fall within the ranges indicated in Table 2.

Table 5 summarizes the effects of degradation on the Total Temperature at LPT outlet (TT_6) and w_f at Op. Point B. The ones reported in Table 5 are typical values of a real case. It can be observed that degradation leads to an increase in TT_6 and w_f. To give an idea of the increase in emitted pollutants, Table 6 reports the value of the emissions both in clean and in the degraded conditions described by Table 5. These values were obtained by referring to the emission indexes (defined as the amount of the contaminant normalized to the mass of fuel burned) obtained from equations available in [84], and in detail we refer to the sections related to the single turboshaft engine. The fuel flow was also calculated using the appropriate equation from the aforementioned work, resulting in a value of 0.0223 kg/s (in clean condition), which is comparable to the value obtained from our simulations (0.0251 kg/s). Op. point B, which is referred to in Table 5, has a PLA of 50.3% and a k value of 0.7 (70% of electric contribution to the total required power). As a result, the thermal portion of the engine is considered to be operating at 35.2% (0.7 × 0.503) of the total deliverable shaft power, which is equal to 295 kW. Therefore, the thermal part produces 0.352 × 295 kW = 103.87 kW at Op. point B. The values reported in Table 6 were obtained by multiplying the emission indexes taken from [84] by the fuel flow data from our simulations. The non-negligible increase in contaminants visible in Table 6 underscores the urgency of transitioning to greener energy sources in order to effectively address the issue of pollution. Other suggested researches about aircraft engine pollutions are available in [85] and [86]. Moreover, to gain an understanding of cost increases resulting from higher fuel consumption, one could estimate by referring to the available data at [87], which refers to kerosene-type jet fuel, such as Jet-A, which is typically used for turboshaft engines [88,89]. Taking, for example, a 10 km cruise at a flight speed of 30.6 m/s (refer to Table 3, Operational Point B), the required time equals (10 ×

Table 3
Operating conditions for the tests.

	Op. point A	Op. point B	Op. point C	Op. point D
PLA (%)	65.3	50.3	68	61.5
Altitude (m)	0	492	1154	2550
True Air Speed (m/s)	30.6	30.6	4	0
Electric contribution k	0.4	0.7	0.47	0.6

Table 2
Degraded components with related PPs.

Component	Degraded performance parameters	Healthy condition	Range of percentage variation
Compressor	Isentropic efficiency (η_c)	From compressor map	0 to −20%
	Corrected mass flow rate (f_c)	From compressor map	0 to −20%
Burner	Pneumatic efficiency ($\eta_{b,p}$)	0.96	0 to −20%
	Combustion efficiency ($\eta_{b,c}$)	0.985	0 to −20%
Shaft transmissions	High-pressure shaft mechanical efficiency (η_{HPS})	0.99	0 to −20%
	Low-pressure shaft mechanical efficiency (η_{LPS})	0.99	0 to −20%
Electric machine	Efficiency ($\eta_{e,m}$)	From the motor map	0 to −20%
Battery	Internal resistance (R_i)	79 $m\Omega$	0 to +20%

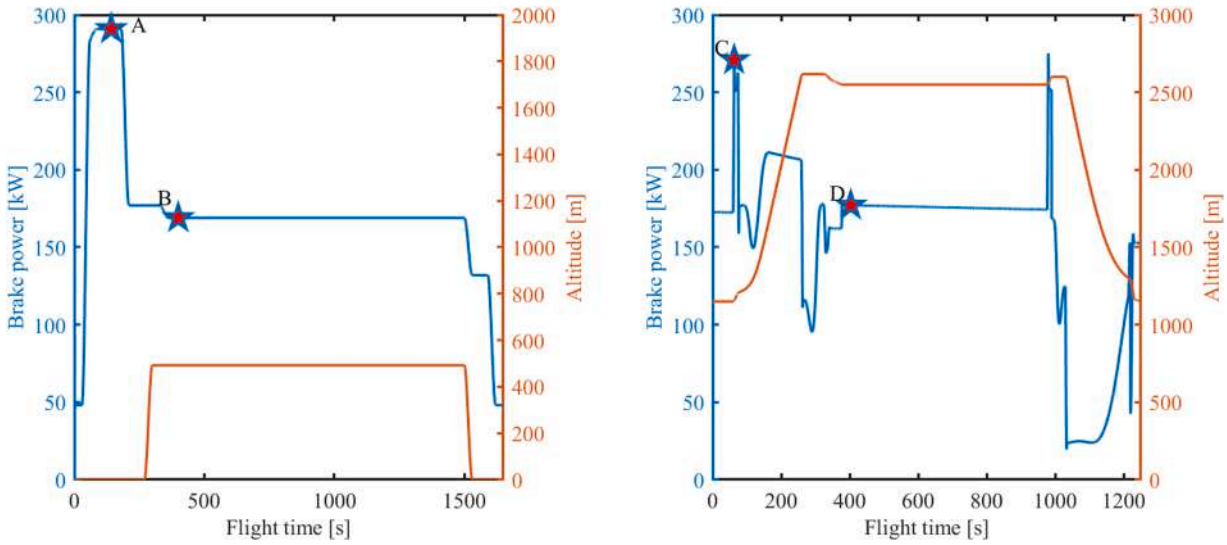

Fig. 6. Mission profiles from which the four working points reported in Table 3 were obtained.

Table 4
Details of train and test datasets.

	Training	Testing
Op. point A	102 steady-state simulations	48 steady-state simulations
Op. point B	170 steady-state simulations	80 steady-state simulations
Op. point C	87 steady-state simulations	40 steady-state simulations
Op. point D	167 steady-state simulations	79 steady-state simulations

1000)/30.6 = 327 s, where 1000 is the conversion factor from km to m. Considering a clean state fuel consumption of 0.0251 kg/s (derived from our simulation, distinct from the value of 0.0223 kg/s estimated from the equation available in [84]), the fuel burned is approximately 0.0251 × 327 = 8.21 kg. Similarly, in the degraded state reported in Table 5, characterized by a fuel consumption of about 0.0273 kg/s, the fuel burned is approximately 0.0273 × 327 = 8.93 kg. Finally, considering a mean fuel density of 811.47 kg/m^3 [90] (Table 1 of the referenced study, at temperature of 283.15 K, which is close to the temperature at the flight altitude of Op. Point B (492 m), according to the International Standard Atmosphere.) and using the most recent data available in [87] (dic 04, 2023), i.e. a price of 2.401 dollars per gallons, the costs for a 10 km cruise in clean and degraded states are approximately (8.21/811.47) × 264.2 × 2.401 = 6.42 dollars and (8.93/811.47) × 264.2 × 2.401 = 6.98 dollars, respectively, where 264.2 is the conversion factor from m^3 to gallons reflecting an increase of 0.56 dollars.

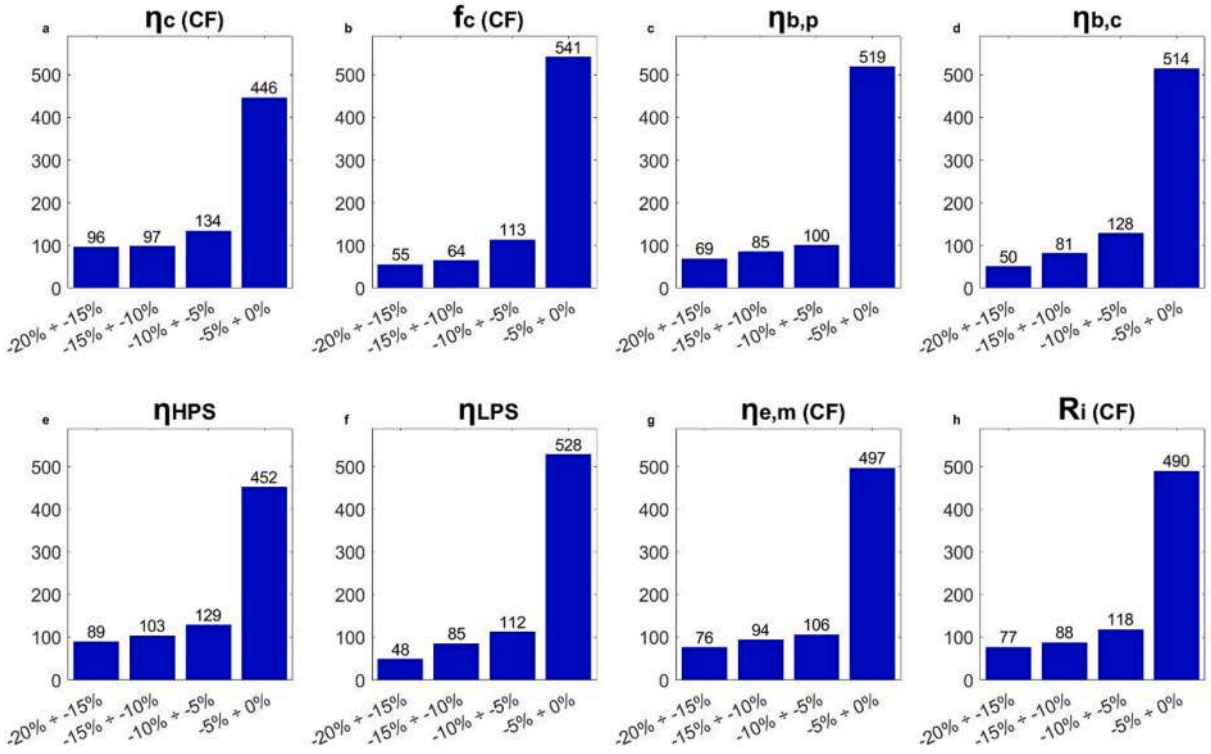

Fig. 7. Distribution of generated PPs (or related correction factors) in degraded conditions.

Table 5
Effects on TT_6 and w_f of the degradation condition analyzed in this work. The percentage variations of TT_6 and w_f are typical of a real case. Working point: Op. point B. % var.: percentage variation.

PP	% var.	TT_6 % var.	w_f % var.
η_c (CF)	−4.543 %	11.041 %	9.084 %
f_c (CF)	−0.037 %		
$\eta_{b,p}$	−1.153 %		
$\eta_{b,c}$	−1.468 %		
η_{HPS}	−8.691 %		
η_{LPS}	−0.0005 %		
$\eta_{e,m}$ (CF)	−17.923 %		
R_i (CF)	−19.592 %		

Table 6
Values of contaminants [g/s] in clean state and in the two degraded conditions considered for Table 5. Obtained with emission indexes available in [84].

	CO	NO_x	HC	PM
Clean	5.92e-1	8.73e-2	4.63e-1	3.45e-3
Condition of Table 5	6.45e-1	9.52e-2	5.05e-1	3.76e-3

3.3. Selected variables

As previously stated, the ANNs were trained to predict PPs (or their CFs) that act as an indicator of the health status of the components. In an online EHM system designed for diagnostic purposes, a pre-trained ANN is continuously updated with data from various stations of the power-train (such as pressures, temperatures, speeds, etc.) to estimate other parameters that indicate the health status of the components and engine. To function properly, the data used to inform the neural network about the conditions present in the engine must be measurable through appropriate sensors or easily derivable from measurable variables. In this paper, a total of twenty variables were selected as input for the ANNs from the ones obtained from simulations. The chosen variables are all measurable with sensors or obtainable from measurable ones. Table 7 reports the variables chosen to create the input datasets for the ANNs while Table 8 summarizes the PPs to be predicted, i.e. the output variables of the ANNs which constitute the output datasets. The variables presented in Table 7 represent the ones constituting the entire input datasets, which will be subject to data reduction techniques to obtain other subsets.

4. Developed health monitoring systems

Various approaches and architectures have been used to develop EHM systems. To address the large amount of data that needs to be processed by the algorithms, data reduction techniques have been implemented and compared in order to decrease computational requirements. The techniques of PCA and KPCA have been utilized for this purpose, to perform data reduction, either through FS by exploiting PCA or FE by exploiting KPCA. In FS, the input variables that are less correlated with the output ones are removed, resulting in a reduction of the dimensionality of the dataset. In FE, the information contained within the dataset is collected and rearranged into a new dataset, formed by new variables that are numerically different from those in the original dataset, but dependent on them. The new variables, called principal components, are ordered by the amount of information they contain in the original dataset, with the first principal component containing the most information and the last one containing the least. Data reduction in this case is achieved by discarding the least important principal components, which contain a minimal amount of information in the original dataset [91,92]. Further details are available in Section 4.2. In this paper, four different approaches have been implemented. Additionally, these approaches have been applied using two different architectures, which are shown in Fig. 8. Fig. 8a illustrates the Single-net architecture, in which only one network is used to predict all eight PPs (or their CFs). Conversely, Fig. 8b illustrates the Multi-net architecture, in which each PP (or its CF) is predicted by a separate network. Table 9 details the four different approaches used. By combining the approaches and architectures mentioned, six different EHM systems have been designed and developed. These EHM systems are listed in Table 10. Notice that the Systems FFNN + FS approach with Single-net architecture and the FFNN + FS + FE approach with Single-net architecture lack. This is because applying FS (which selects only the input variables physically related to the variables to be predicted) with a Single-net architecture does not result in any dataset reduction, because there are no input variables that are uncorrelated with all the output variables.

Furthermore, Fig. 9 reports a flow-chart which explains how diagnostic procedure based on the developed EHM works.

The following paragraphs are dedicated to a description of the methods used to develop the four approaches described in Table 9.

Table 7
Variables selected to create the entire input dataset.

Input	
Power Lever Angle (PLA)	Total press. at burner outlet (TP_4)
Altitude (h)	Total temp. at HPT outlet (TT_5)
Flight speed (v_f)	Total press. at HPT outlet (TP_5)
Electric contribution (k)	Total temp. at LPT outlet (TT_6)
Battery delta temp. respect to clean status (ΔT)	Total press. at LPT outlet (TP_6)
Total temp. at compressor inlet (TT_2)	Compressor shaft speed (ω_{HPS})
Total press. at compressor inlet (TP_2)	Fuel mass flow rate (w_f)
Total temp. at compressor outlet (TT_3)	Overall pressure ratio (OPR)
Total press. at compressor outlet (TP_3)	Battery current (I)
Total temp. at burner outlet (TT_4)	Battery voltage (V)

Table 8
Variables selected as output for ANNs. These represent the PPs (or their CFs) to be predicted, used as health indicator of the components.

Output
η_c (CF)
f_c (CF)
$\eta_{b,p}$
$\eta_{b,c}$
η_{HPS}
η_{LPS}
$\eta_{e,m}$ (CF)
R_i (CF)

4.1. Feed-forward neural network

ANNs have been developed to mimic the capacity of the human nervous system to learn from examples. ANNs are able to understand the correlation between two sets of variables, an input set and an output set, that are physically related to each other. The fundamental components of an ANN are neurons. These neurons are organized in layers. ANNs are composed of an input layer, an output layer, and one or more hidden layers. The input layer has a number of neurons equal to the number of input variables, and the output layer has the same number of neurons as output variables. The number of hidden layers and neurons is adjustable by the user and affects the prediction performance of the ANNs. Neurons send information to the neurons of the next layer through links, each with its own "weight". A FFNN, the type employed in this study,

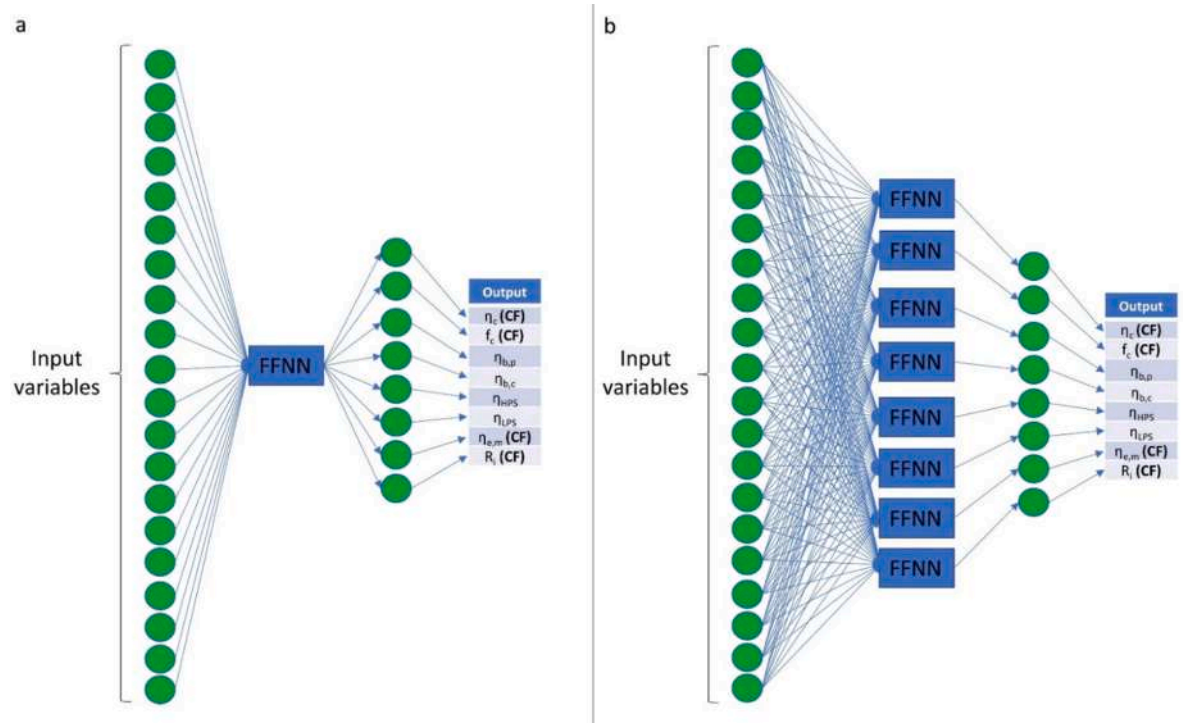

Fig. 8. The two architectures used. a: Single-net; b: Multi-net. The input variables change from case to case.

Table 9
The four approaches used in this work.

#	Approach	Description
1	Simple FFNN	The nets are informed with the entire input dataset, formed by all the twenty variables reported in Table 7.
2	FFNN + FS	A FS process is applied to the entire input dataset formed by all the twenty variables reported in Table 7 and the resulting subset is used to inform the nets.
3	FFNN + FE	A FE process is applied to the entire input dataset formed by all the twenty variables reported in Table 7 and the resulting subset is used to inform the nets.
4	FFNN + FS + FE	Both FS and FE processes are applied to the complete dataset and the resulting reduced dataset is used to inform the nets.

represents one of the two main categories of ANNs, distinguished by the direction of information flow between its layers. In FFNNs, information flows in unidirectional manner, meaning that it moves only in one direction, from the input nodes to the output ones, passing from the hidden nodes, if any, without any cycles or loops. In recurrent neural networks instead, the flow is bidirectional. Fig. 10 illustrates the typical structure of a FFNN, the type used in this work.

Information is input into the network via the input layer and propagates through the network until it reaches the output layer. Each neuron located in the hidden (or hiddens) and output layers processes the information received from the previous layers, each with its own weight, and adding a bias, as shown in Eq. (14):

$$z = \sum_{i=1}^{s} W_i J_i + \gamma \tag{14}$$

where z represents the neuron output, W_i is the weight of the link between the neuron in question and the *i-th* neuron that sends it information, J_i is the information sent by the *i-th* neuron, s is the number of neurons that send information to the neuron in question and γ is the bias. Finally, an activation function is applied to normalize the results of Eq. (14). FFNNs are characterized by two phases, a training phase and a testing phase. In the training phase, the FFNN is provided with an input dataset and a corresponding output dataset. The algorithm iteratively

Table 10
The six systems developed, created combining the four approaches and the two architectures presented in Table 9 and Fig. 8 respectively.

#	System	Description
1	Simple FFNN approach with Single-net architecture	The input dataset for each FFNN used is constituted by all the twenty variables reported in Table 7. Only one FFNN is used to predict all the eight variables reported in Table 8.
2	Simple FFNN approach with Multi-net architecture	The input dataset for each FFNN used is constituted by all the twenty variables reported in Table 7. A different FFNN is used for each of the eight variables reported in Table 8.
3	FFNN + FS approach with Multi-net architecture	The input dataset for each FFNN used is constituted by a subset resulting from a FS process applied to the entire dataset constituted by all the twenty variables reported in Table 7. A different FFNN is used for each of the eight variables reported in Table 8.
4	FFNN + FE approach with Single-net architecture	The input dataset for each FFNN used is constituted by a subset resulting from a FE process applied to the entire dataset constituted by all the twenty variables reported in Table 7. Only one FFNN is used to predict all the eight variables reported in Table 8.
5	FFNN + FE approach with Multi-net architecture	The input dataset for each FFNN used is constituted by a subset resulting from a FE process applied to the entire dataset constituted by all the twenty variables reported in Table 7. A different FFNN is used for each of the eight variables reported in Table 8.
6	FFNN + FS + FE approach with Multi-net architecture	The input dataset for each FFNN used is constituted by a subset resulting from a FS and subsequently a FE process applied to the entire dataset constituted by all the twenty variables reported in Table 7. A different FFNN is used for each of the eight variables reported in Table 8.

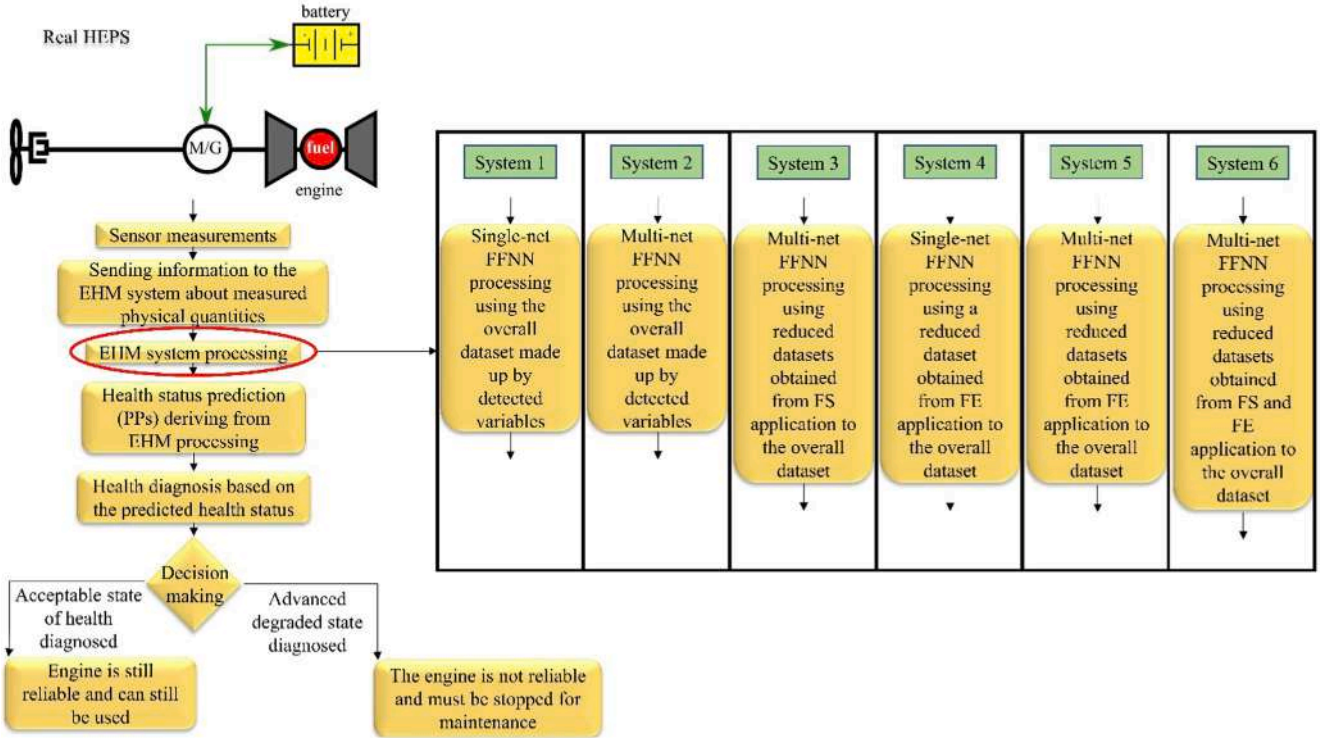

Fig. 9. Flow-chart of the diagnostic procedure that exploits the developed EHM systems.

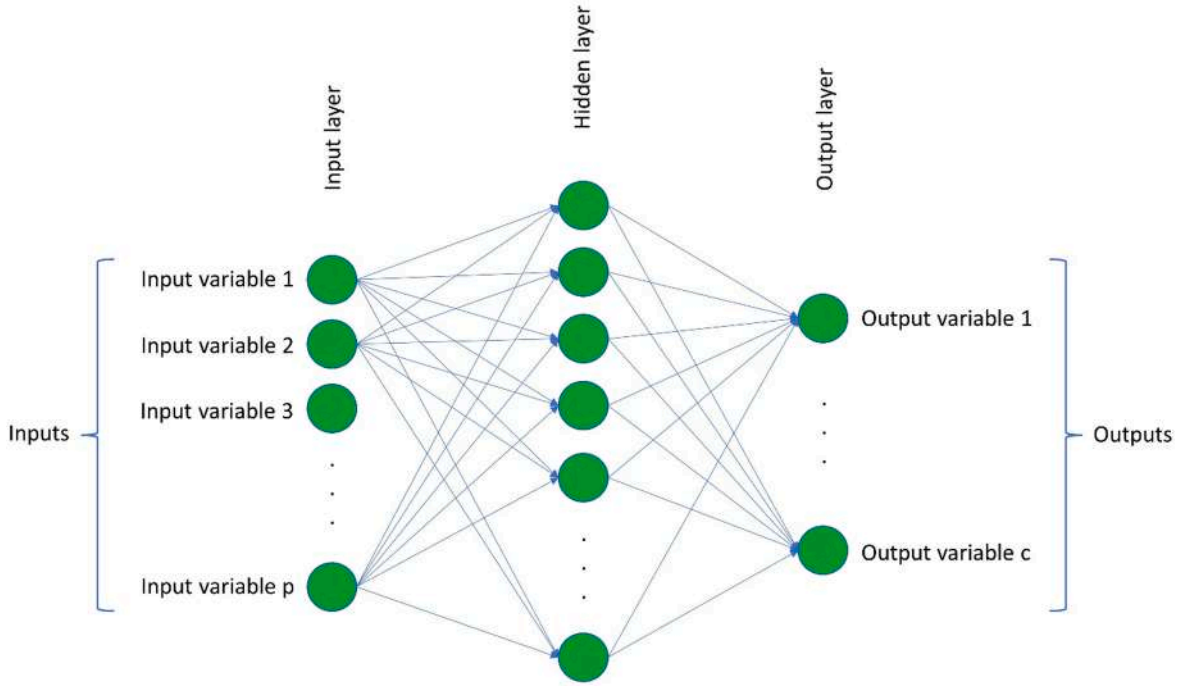

Fig. 10. A typical structure of a FFNN, the type used in this paper.

calculates the connection weights and biases in order to minimize the error between the network's output and the output variables (also called target variables) during this phase. In the testing phase, the pre-trained network (with the weights and biases calculated during the training phase) is provided with new input data to predict the corresponding output variables. The obtained network outputs are then compared with the actual values to evaluate the prediction performance. The networks used in this work were implemented using the Matlab neural network tool and the Bayesian regularization algorithm for training. The networks were trained for a maximum of 5000 epochs and the training process stops when the gradient reaches $1e^{-7}$ or the 5000 epochs are reached.

4.2. PCA and KPCA for FS and FE

PCA is a data analysis technique that is commonly used to analyse large datasets by identifying the most significant trends and associated variables. PCA is also used to determine how the total variation of the data is distributed among the principal components [93]. PCA reduces dimensionality in machine learning, transforming data into a lower-dimensional representation while retaining essential information. Both FE and FS can be achieved using PCA, but there are key differences between the two approaches. Specifically, in this work PCA was utilized to perform FS on the input variables. The PCA is widely used for FE as well, but in more complex systems, such as non-linear systems, the KPCA is more appropriate than the traditional PCA. Therefore, in this work, KPCA was used for FE. FE involves transforming the original features of the input data into a new set of features called principal components. These principal components are linear combinations of the original features that capture the most important information in the data. The number of components determines the reduced feature space dimensionality. On the other hand, FS involves keeping the most relevant original features, evaluating their importance, and retaining the most informative ones for analysis. FE is useful for reducing dimensionality in highly correlated and redundant data. FS is suitable when only a relevant subset of features is needed. Both techniques could improve ANN performance, with the choice based on analysis requirements. This study used these techniques to identify and extract relevant features from input data, enhancing prediction tool performance. The KPCA proceeds by mapping the input data X_k into a feature space F_k by means of a non-linear mapping φ and next performing a linear PCA in F_k [94] (Fig. 11 [95]).

Assuming centered mapped data, i.e. $\sum_{i=1}^{s}\varphi_{(x_i)} = 0$ where s is the number of observations and x_i is the single observation, KPCA diagonalizes the covariance matrix of the mapped data $\varphi_{(x_i)}$ defined as:

$$C = \frac{1}{s}\sum_{i=1}^{s}\varphi_{(x_i)}\varphi_{(x_i)}{}^{T} \tag{15}$$

To do this, the following equation must be solved for eigenvalues $\lambda > 0$ and eigenvectors $\nu \in F_k\backslash\{0\}$:

$$\lambda v = Cv \tag{16}$$

the right-hand member of which, using Eq. (15) becomes:

$$Cv = \frac{1}{s}\sum_{i=1}^{s}\left(\varphi_{(x_i)}\cdot v\right)\varphi_{(x_i)} \tag{17}$$

All the solutions v with $\lambda \neq 0$ must lie within the span of $\varphi_{(x_1)}, \cdots, \varphi_{(x_s)}$ and the coefficients $\alpha_i (i = 1, \cdots, s)$ exist such that:

$$v = \sum_{i=1}^{s}\alpha_i\varphi_{(x_i)} \tag{18}$$

The set of equation defined as:

$$\lambda\left(\varphi_{(x_i)}\cdot v\right) = \varphi_{(x_i)}\cdot Cv \qquad \text{for i} = 1, \cdots, \text{s} \tag{19}$$

can be considered. After substituting Eqs. (15) and (18) in 19 and defining an s × s matrix K by $k_{ij} \equiv k(x_i, x_j) = \varphi_{(x_i)} \cdot \varphi_{(x_j)}$ an eigenvalue problem it is obtained:

$$s\lambda\alpha = K\alpha \tag{20}$$

(with α denotes the vector with entries $\alpha_1, \cdots, \alpha_s$) to be solved for nonzero eigenvalues λ_l and eigenvectors $\alpha^l = (\alpha_l{}^1, \cdots, \alpha_l{}^s)^T$ subject to normalization criterion $\lambda_l(\alpha^l\cdot\alpha^l) = 1$. To obtain the principal components, the projections of input data X_k are computed onto the eigenvectors v^l. More details are available in [94–97].

4.3. Systems setup

In this section, we provide an overview of the configuration used for the developed systems. This includes the number of neurons selected for the ANNs, the number of principal components retained in the FE process, and the criteria used to perform the FS process. Additionally, the evaluation criteria used to assess the prediction performance of the developed health monitoring systems are presented.

4.3.1. Error metrics

Prediction performance has been measured by using different error metrics, typically used to obtain the goodness of the models. More in detail, the just cited metrics are:

- Mean Absolute Error (MAE)
- Coefficient of Determination (CoD)
- Maximum Absolute Error (MaxAE)
- Mean Squared Error (MSE)
- Root Mean Square Error (RMSE)
- Normalized Error (n.e.)

defined as:

$$MAE = \frac{\sum_{i=1}^{s}|p_i - t_i|}{s} \tag{21}$$

$$CoD = \left(\frac{s\left(\sum_{i=1}^{s}t_ip_i\right) - \left(\sum_{i=1}^{s}t_i\right)\left(\sum_{i=1}^{s}p_i\right)}{\sqrt{\left[s\left(\sum_{i=1}^{s}p_i{}^2\right) - \left(\sum_{i=1}^{s}p_i\right)^2\right]\left[s\left(\sum_{i=1}^{s}t_i{}^2\right) - \left(\sum_{i=1}^{s}t_i\right)^2\right]}}\right)^2 \tag{22}$$

$$MaxAE = MAX_{i=1}^{s}\left(|p_i - t_i|\right) \tag{23}$$

$$MSE = \frac{\sum_{i=1}^{s}(p_i - t_i)^2}{s} \tag{24}$$

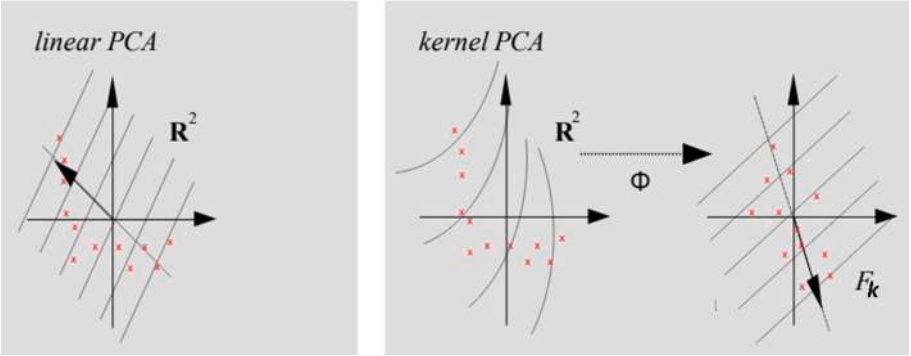

Fig. 11. Linear PCA vs Kernel PCA.

$$RMSE = \sqrt{\frac{1}{s}\sum_{i=1}^{s}(p_i - t_i)^2} \tag{25}$$

$$n.e._i = \frac{(p_i - t_i)}{std(t)} \tag{26}$$

where in our work p_i is the prediction of the i-th observation, t_i is the target value of the i-th observation, $std(t)$ is the standard deviation of the targets and s is the number of observations.

4.3.2. System 1: Simple FFNN approach with Single-net architecture

In this system, all twenty variables presented in Table 7 were used as the input dataset for the FFNN, without performing any data reduction techniques. A single FFNN was used to predict all eight output variables. Each FFNN in this work consisted of only one hidden layer. Therefore, the only parameter that required tuning was the number of neurons in the hidden layer, which was determined through a sensitivity analysis to identify the optimal number of hidden neurons for each scenario. For the single FFNN used in System 1, 20 hidden neurons were used. In all training processes, 5% of the input dataset was allocated for validation purposes.

4.3.3. System 2: Simple FFNN approach with Multi-net architecture

This system is similar to the one previously described, with the exception that in this system eight different FFNNs were used, one for each output variable to be predicted. The complete dataset of twenty variables presented in Table 7 was used as the input dataset for each of the FFNNs used in this system. The parameters that needed to be set in this system were the number of hidden neurons for each of the FFNNs used, which are summarized in Table 11. Note that due to the fact that each FFNN in the Multi-net architecture is independent, a separate sensitivity analysis must be conducted for each different FFNN, resulting in a more time-consuming sensitivity analysis.

4.3.4. System 3: FFNN + FS approach with multi-net architecture

In system 3 FS has been applied. PCA was used to understand the correlation between the input variables and the output variable to be predicted. Fig. 12 illustrates the observations of the first two principal components of all relevant variables, both input and output. Each variable is represented by a line and each observation is represented by a dot. In this graph, the more perpendicular the lines corresponding to two variables are, the less correlated they are.

In Fig. 12, the green and red lines represent the input variables, while the magenta lines represent the output variables, i.e. the PPs (or their CFs) to be predicted. As previously noted, the correlation between variables is indicated by the angle between them. The closer the angle is to 90 degrees, the more uncorrelated the variables are. Taking this into account, the criterion used for the FS process is to remove all input variables whose direction falls within a specified angle around 90 degrees relative to the direction of the output variable in question. FS was only applied using the Multi-net architecture, in which only one variable needs to be predicted by each FFNN, as it is challenging to find an input variable that is lowly correlated with all eight output variables when using the Single-net architecture, where all eight output variables must be predicted by the same FFNN. In more detail, the correlation between the twenty variables in the entire input dataset was studied for each of the eight output variables (one at a time), and a subset was obtained for each case. Table 12 lists the input variables that were discarded for each of the eight FFNNs used in System 3. The parameters that must be set in System 3 are the angular tolerance between the perpendicular direction of the output variable and the input variables for the FS process and the number of hidden neurons in the FFNN used. The angle chosen is 10 degrees, thus any input variable whose direction falls within the range of −10 to 10 degrees relative to the perpendicular direction of the output variables were removed. Table 13, on the other hand, indicates the number of hidden neurons employed.

Table 11
Number of hidden neurons in the FFNNs used in System 2.

Net number	Predicted PP	Hidden neurons
1	η_c (CF)	6
2	f_c (CF)	4
3	$\eta_{b,p}$	10
4	$\eta_{b,c}$	6
5	η_{HPS}	6
6	η_{LPS}	6
7	$\eta_{e,m}$ (CF)	8
8	R_i (CF)	8

4.3.5. System 4: FFNN + FE approach with Single-net architecture

In system 4, KPCA was employed to perform FE on the entire input dataset, consisting of all 20 variables listed in Table 7. KPCA accomplishes this by constructing a new dataset composed of new variables referred to as principal components, which are numerically distinct from the original variables, but linearly dependent on them. The new dataset retains the same amount of information present in the original one. The principal components are sorted in order of the amount of information they contain, with the first component having the most and the last component having the least. Data reduction is achieved by discarding some of the last principal components, i.e., those containing the least amount of information, depending on the desired degree of reduction. This results in a reduction of the dataset size and in the computational effort without a lost in prediction performance. In System 4, the Single-net architecture was employed, meaning that only one FFNN was utilized to predict all eight PPs (or their CFs). In this scenario, the parameters that must be set are the number of hidden neurons for the single FFNN used, and the number of principal components to retain during the FE process. A sensitivity analysis was conducted to determine the optimal number of principal components to retain. In System 4, 20 hidden neurons were utilized, and 10 principal components were retained.

4.3.6. System 5: FFNN + FE approach with Multi-net architecture

In system 5 a multi-net architecture was employed. This entails the use of a distinct FFNN for each of the eight variables to be predicted, and the implementation of a FE process for each of these FFNNs. Table 14 provides an overview of the number of hidden neurons and the number of principal components retained for each of the FFNNs used in System 5. These parameters have been chosen based on a sensitivity analysis on the influence of their values on the predictive performance.

4.3.7. System 6: FFNN + FE + FS approach with Multi-net architecture

This is the most complex system developed in this work. It involves the use of the multi-net architecture, which entails the use of eight distinct FFNNs. Therefore, eight different subsets have been derived from the initial dataset, which consists of all twenty variables listed in Table 7, by implementing both a FE and FS process. In more detail, a FS process was initially performed to eliminate the input variables that are less correlated with the output variables, considered one at a time, similar to the approach used in System 3. Subsequently, the FE process was conducted on the eight subsets obtained from the FS process. The input variables that were discharged during the FS process are the same ones that were removed during the FS process in System 3 (Table 13). In this complex case, the parameters that must be set are the angular tolerance between the perpendicular direction of the output variable and the input variables for the FS process, the number of hidden neurons, and the number of principal components retained for each of the FFNNs used. As in previous case, the angle considered is 10 degrees, thus any input variable whose direction falls within the range of −10 to 10 degrees relative to the perpendicular direction of the output variables were removed, while Table 15 provides information on the number of hidden neurons and principal components retained for each FFNN.

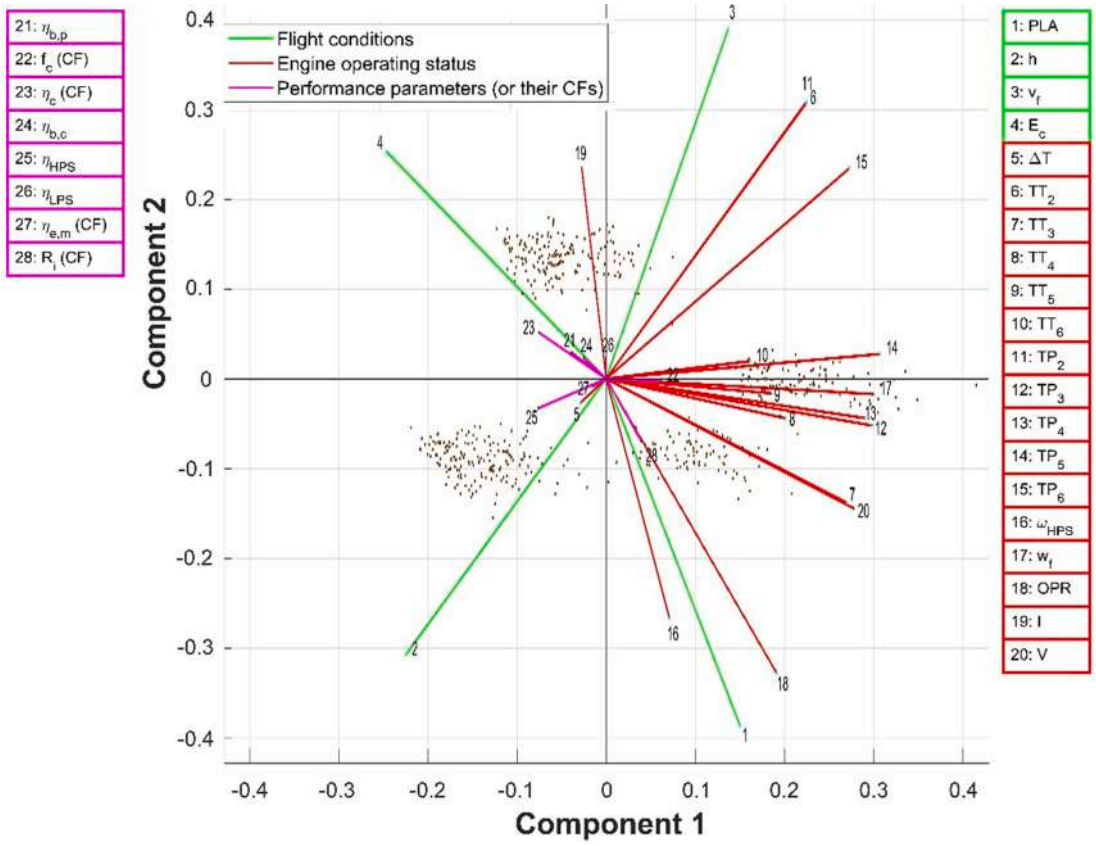

Fig. 12. Correlation between input and output variables.

Table 12
Discharged input variables after application of a FS process.

Net number	Variable to predict	Discharged inputs
1	η_c (CF)	h, TT_2, TP_2
2	f_c (CF)	I
3	$\eta_{b,p}$	h, TT_2, TP_2
4	$\eta_{b,c}$	h, ΔT, TT_2, TP_2, TP_6
5	η_{HPS}	PLA, OPR
6	η_{LPS}	TT_5, TP_5, w_f
7	$\eta_{e,m}$ (CF)	PLA, ω_{HPS}, OPR
8	R_i (CF)	TP_6

Table 13
Number of hidden neurons in the FFNNs used in System 3.

Net number	Predicted PP	Hidden neurons
1	η_c (CF)	6
2	f_c (CF)	10
3	$\eta_{b,p}$	6
4	$\eta_{b,c}$	6
5	η_{HPS}	6
6	η_{LPS}	6
7	$\eta_{e,m}$ (CF)	4
8	R_i (CF)	4

Table 14
Number of hidden neurons and principal components used in System 5.

Net number	Predicted PP	Hidden neurons	Principal Components
1	η_c (CF)	6	10
2	f_c (CF)	4	15
3	$\eta_{b,p}$	10	10
4	$\eta_{b,c}$	6	11
5	η_{HPS}	6	10
6	η_{LPS}	6	12
7	$\eta_{e,m}$ (CF)	8	7
8	R_i (CF)	8	10

Table 15
Number of hidden neurons and principal components used in System 6.

Net number	Predicted PP	Hidden neurons	Principal Components
1	η_c (CF)	6	10
2	f_c (CF)	10	9
3	$\eta_{b,p}$	6	10
4	$\eta_{b,c}$	6	10
5	η_{HPS}	6	10
6	η_{LPS}	6	13
7	$\eta_{e,m}$ (CF)	4	7
8	R_i (CF)	4	10

5. Results and discussions

In this section, the results obtained from the systems previously described are presented and discussed. A total of six different systems were utilized and their results are compared. For further information, please refer to Table 10.

Figs. 13–16 display the results obtained in the prediction of the eight PPs (or their CFs) using the six developed systems. The results will now be analyzed. In the prediction of η_c (CF), the results are very good, as indicated by the CoD metrics, which are all equal to 1. The best results are provided by System 2 and System 3, with a MAE of 0.00097 and 0.00088 respectively. The worst results are yielded by System 4, with a MAE of 0.00207. The results for the prediction of f_c (CF), the second parameter selected for the compressor, are still good. The CoD values are slightly lower than the previous case, ranging from 0.99 to 1. In this scenario, System 3 continues to exhibit one of the best results with a

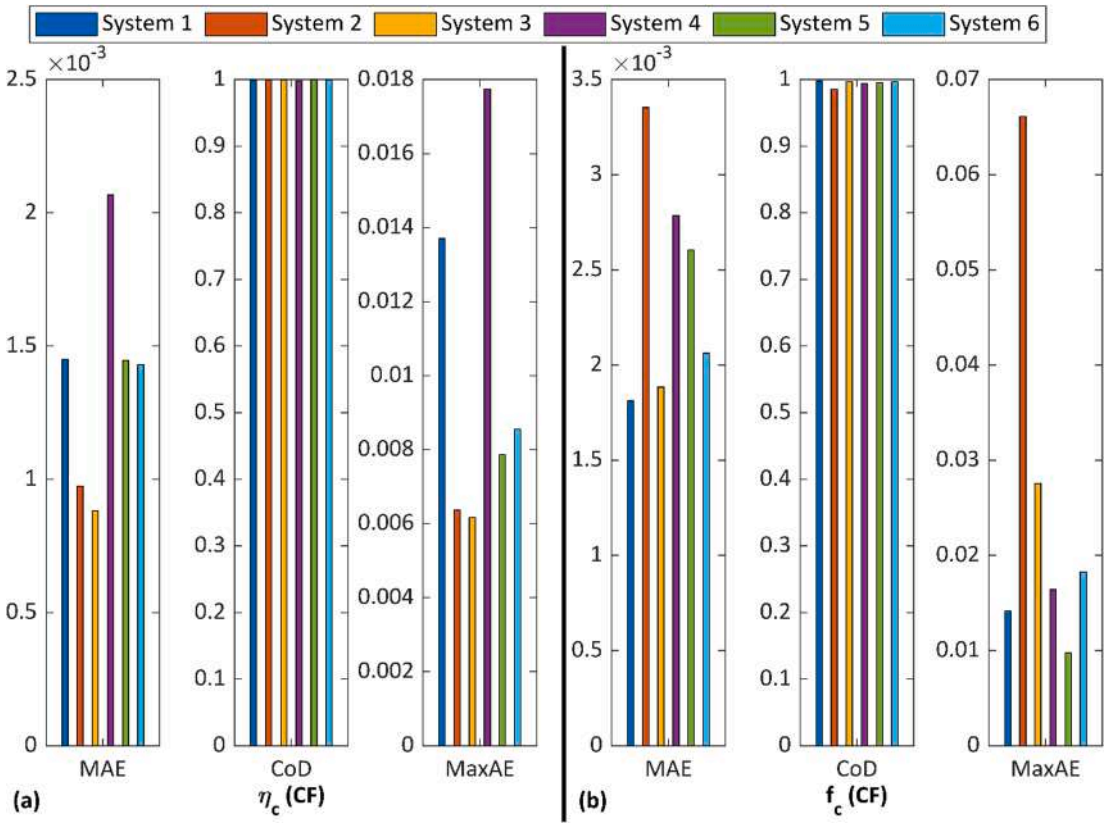

Fig. 13. Prediction results for η_c (CF) and f_c (CF). All the six developed systems are compared. The used error metrics were obtained by using Eqs. (21)-(23).

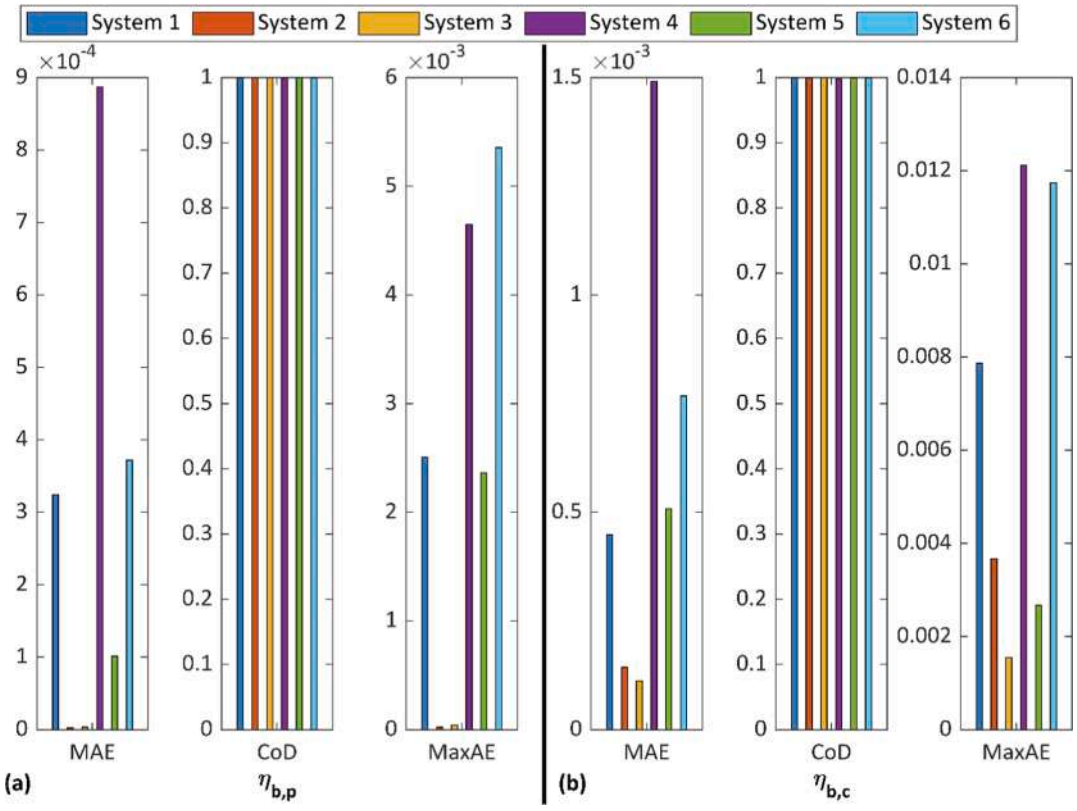

Fig. 14. Prediction results for $\eta_{b,p}$ and $\eta_{b,c}$. All the six developed systems are compared. The used error metrics were obtained by using Eqs. (21)-(23).

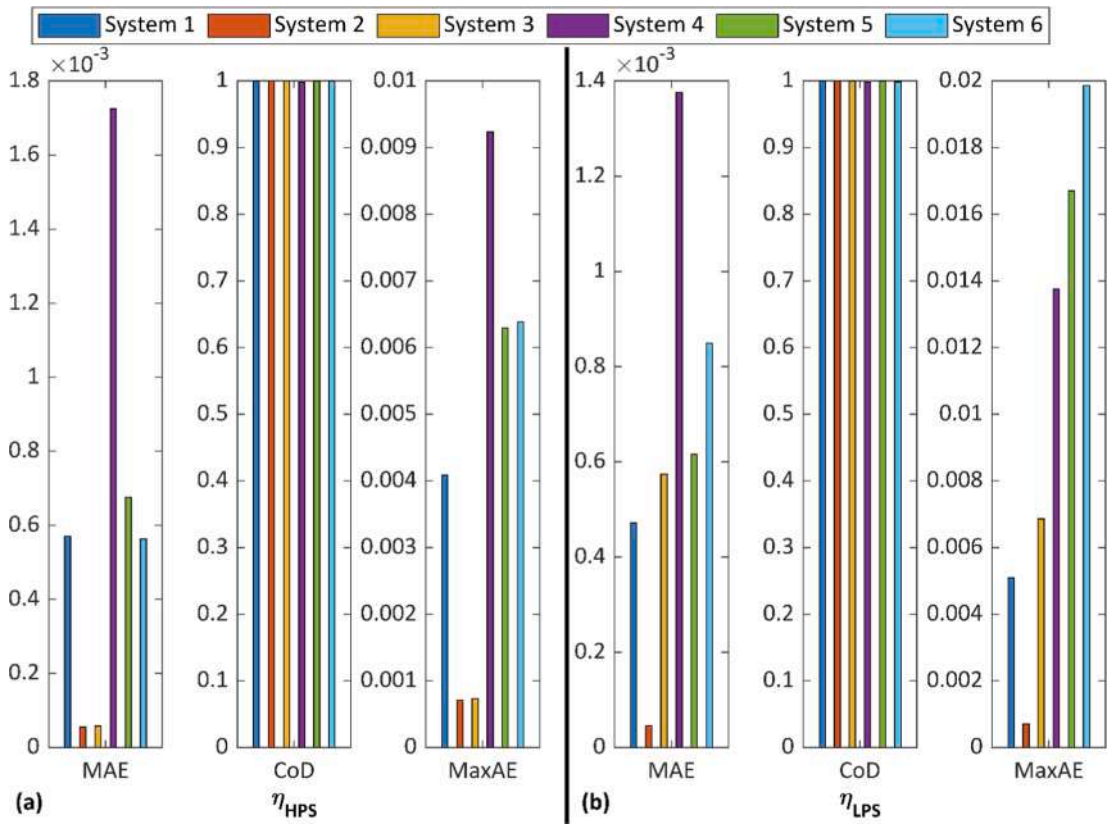

Fig. 15. Prediction results for η_{HPS} and η_{LPS}. All the six developed systems are compared. The used error metrics were obtained by using Eqs. (21)-(23).

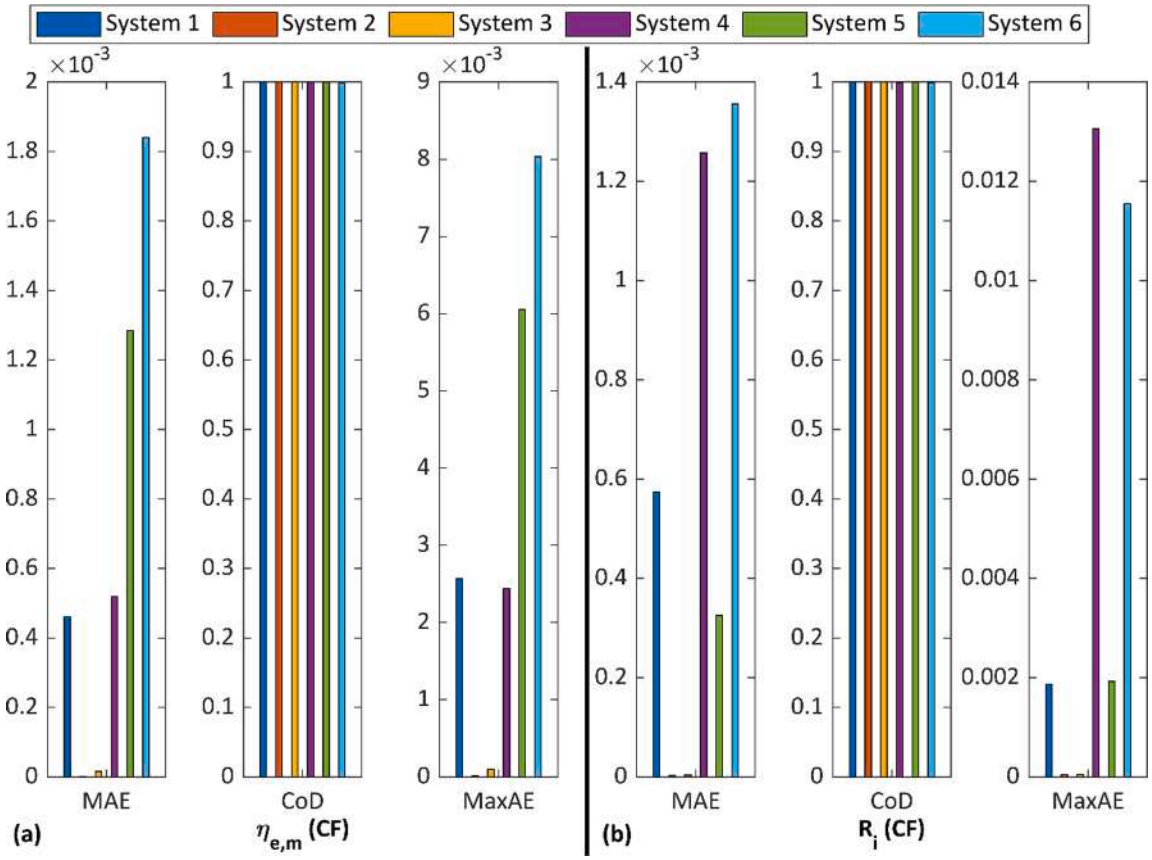

Fig. 16. Prediction results for $\eta_{e,m}$ (CF) and R_i (CF). All the six developed systems are compared. The used error metrics were obtained by using Eqs. (21)-(23).

MAE of 0.00188. The prediction performance of System 2, however, is visibly worse as seen from all error metrics used. Better performance is provided by System 1, with a MAE of 0.00181. Very good results are also obtained for $\eta_{b,p}$, with CoD values all equal to 1. From Fig. 14, a marked difference between Systems 2 and 3 is evident. The performance of these two systems is very good, with a MAE of 0.0000025 and 0.0000035 respectively. System 4 shows the worst results, with a MAE of 0.00089. In $\eta_{b,c}$ prediction, the same pattern is observed. Systems 2 and 3 are the best, with a MAE of 0.00014 and 0.00011 respectively, while System 4 shows the worst results, with a MAE of 0.0015. In η_{HPS} prediction, CoD values remain all equal to 1. In this case as well, Systems 2 and 3 continue to provide the best results, with a MAE of 0.000055 and 0.000058 respectively. Similarly, System 4 continues to perform poorly, with a MAE of 0.0017. A marked difference between the systems is evident. The results for the prediction of η_{LPS} are similar, with CoD values remaining equal to 1. System 4 continues to perform poorly, with a MAE of 0.0014. However, in this scenario, there is a significant difference in the performance of Systems 2 and 3. System 2 emerges as the best, with a MAE of 0.000045. In the prediction of $\eta_{e,m}$ (CF), CoD values remain equal to 1. In this case, however, the worst system is the 6th, which has a MAE of 0.0018. The best performance is still provided by Systems 2 and 3. For these two systems, the MAEs are 0.0000012 and 0.000016 respectively. Although both systems provide excellent results, the superior performance of System 2 is evident. Finally, the results for the prediction of R_i (CF) are also very good, with CoD values remaining equal to 1. In this scenario, the worst performing systems are the 4th and 6th, which exhibit similar performance, unlike previous cases where a specific system performed poorly across all error metrics. The best results are once again provided by Systems 2 and 3, which show very similar performance. However, System 2 is the best, with a MAE of 0.0000034. Systems 2 and 3 are the top performers in almost all cases, with a clear distinction from the other developed systems. An exception occurs in the prediction of f_c (CF), where System 2 is the worst performer and System 1 is the best. In the prediction of η_{LPS} instead, System 2 is the best performer, while System 3 shows a significant drop in performance. Fig. 17 compares the target values to the predicted values by all six developed systems, along with the normalized errors. It is evident that the normalized errors remain below 0.3, except for f_c (CF) where a peak in the normalized error of System 2 is present and for η_{LPS} where a peak of the errors of System 6 is apparent, reaching the value of −0.38. Lastly, Fig. 18 shows the total training time required for each of the six systems presented, highlighting that Systems 1 and 4 have the longest time, primarily due to their use of the Single-net architecture. Specifically, the use of FS process in System 4 results in a decrease in input variables and subsequently leads to 8.8 s reduction in training time. The use of Multi-net architecture also leads to a decrease in training time, as seen from the required time for Systems 2, 3, 5, and 6. In the Multi-net architecture, the total training time reported in Fig. 18 is the sum of the training times required by each FFNN that makes up the systems in question. Unsurprisingly, the least time-consuming system for the training process is the 6th, which is the most complex system developed in this work, incorporating both data reduction techniques. This system took 12.75 s for the training process, 24.17 s less than System 1, which has the longest training time.

The superior performance of System 2 can be attributed to its simple structure, employing a straightforward feedforward neural network

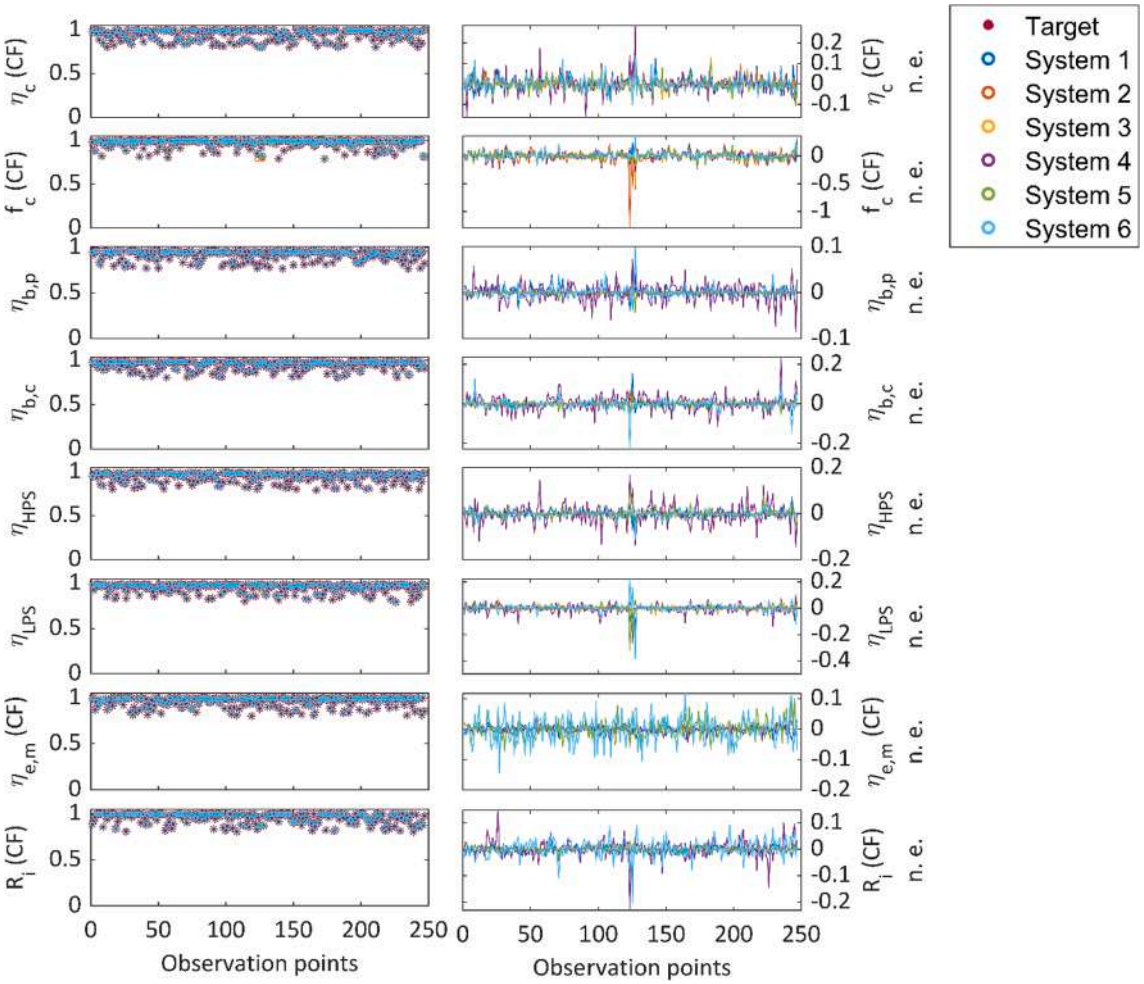

Fig. 17. Comparison between target and predicted values, together with the related normalized errors (n.e.). The normalized errors were obtained by using Eq. (26).

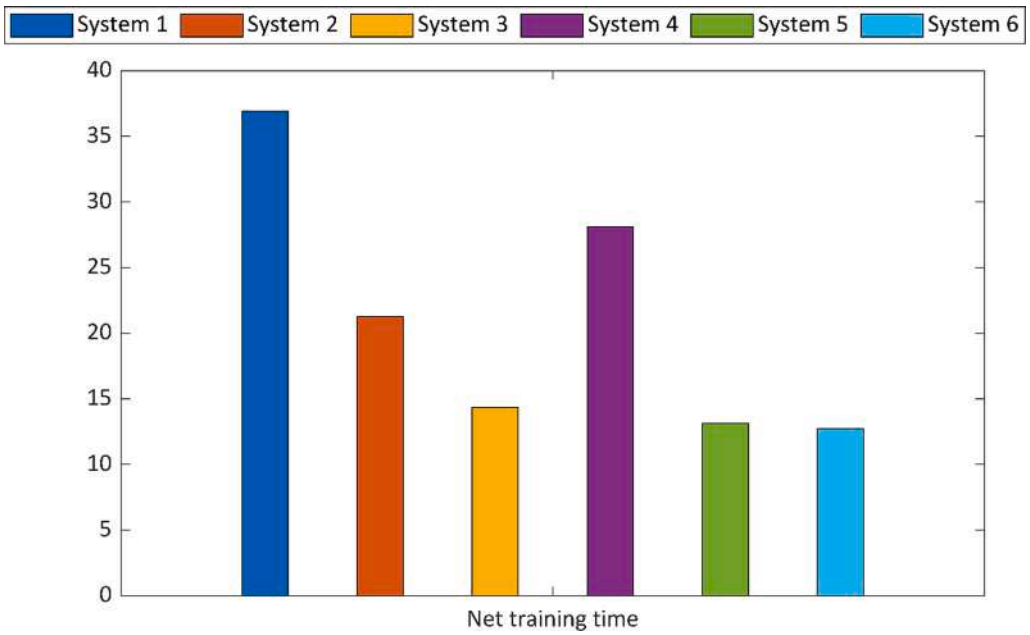

Fig. 18. Total training time of the six developed systems in seconds.

(FFNN) within the Multi-net architecture. This combination enhances training efficiency by requiring the network to learn the physical correlations with only one output variable.

On the other hand, System 3 demonstrates commendable performance, occasionally outperforming System 2. This success can be credited to the Multi-net architecture and the impact of the feature selection (FS) process. FS optimizes training by removing unnecessary input variables, streamlining the learning process and preventing potential confusion for the network.

Conversely, System 4 often exhibits subpar performance, indicating a lower-quality training process. This is likely due to the challenge of training the network to learn the relationship between the inputs and all output variables simultaneously, as seen in the Single-net architecture. Additionally, comparing with System 1, the application of the FE process appears to intensify the learning challenge for the network, as evidenced by the comparison between System 2 and System 5.

5.1. Comparison with other prediction methods

To demonstrate the efficiency of the present prediction method, following a sensitivity analysis, an in-depth comparison with various machine learning methods deriving from the matlab regression learner tool has been performed. This comparison is detailed in Table 16, with the models trained using the original data (without feature selection or feature extraction), well described in the associated references listed in Table 16, referring to the associated Matlab tutorials.

The comparison includes commonly used regression methods such as linear regression (LR) models, decision trees (DTs), Gaussian process regression (GPR) models, support vector machines (SVMs), and ensembles of tree models. The default parameters were utilized for the models in this study. It's worth noting that more favourable outcomes can be attained through the optimization of model parameters.

As evident in Section 5, System 2 consistently demonstrates superior performance across various cases. The comparisons provided in this section focus on System 2 as outlined in this work, utilizing different prediction methods, including our own and those reported in Table 16. The predicted variable is $\eta_{b,p}$, for which the best prediction performance was achieved.

In Table 17, cells highlighted in green signify Machine Learning Techniques that exhibit superior performance, whereas those in red indicate the techniques with comparatively subpar performance. Notably, Table 17 highlights the superiority of the FFNN utilized in this study (indicated in bold at the top). The superiority of the ANNs compared with other Machine Learning Techniques is highlighted in many works available in literature. ANNs shown a high versatility, and can be adequately used in different scenarios, how demonstrated by many researches [56 98–104]. ANNs instead are one of the most used Machine Learning Techniques, as can be observed through adequate bibliographic research, and this underscores the satisfactory results achievable with this tool.

The comparison reveals a remarkable improvement in performance, with a MAE lower by two orders of magnitude, MSE lower by four orders of magnitude, and RMSE lower by two orders of magnitude compared to the other machine learning methods.

In the test dataset, Gaussian Kernel regression models (KE) were the worst models for predicting the parameters.

From Table 17, one can discern the most effective prediction methods for each macro class of Machine Learning Techniques (LR, RT, SVM, EN, GPR, and KE). Fig. 19 supplements this information by offering additional insights into the prediction performance of these top methods. It highlights their prediction capacity across various ranges of percentage differences (p.d.) concerning the healthy condition of the predicted variable. Upon detailed analysis of the figure, a discernible pattern emerges regarding the influence of percentage difference (p.d.) on the prediction performance, particularly notable for RT, EN, and KE-based models.

Most models demonstrate satisfactory predictive accuracy for degradation levels lower than 5%, and the predicted values closely align with the measured data. Nonetheless, as degradation increases, the predictive error rises.

Specifically, when comparing predictions for $\eta_{b,p}$ values below 5% deviation from the healthy condition to those falling within the range of 15% to 20% deviation, there is a notable increase in Mean Absolute Error (MAE) for RT, EN, and KE-based methods. For example, the MAE for RT increases from 0.0114 to 0.0218, for EN from 0.0087 to 0.0231, and for KE from 0.0310 to 0.0973.

Significantly, the worst performance of the RT and KE-based models was readily discernible in Table 17, emphasizing the paramount importance of this observation, notably for the KE-based model. To provide a more comprehensive comparison, the fifth plot in Fig. 19

Table 16
Some details about the method provided in the RL tool of MATLAB 2023a used for comparison with FFNN. The "Description" column provides a brief overview of the regression model, while the "Options" column offers additional information about the details enclosed in parentheses in column "Method".

Method	Description	Options	
FFNN (used in this work)	A simple FFNN model. More detailed descriptions in Sections 4.1, 4.3.2-4.3.7		[105]
LR (linear) LR (Interactions linear) LR (Robust linear)	A simple linear regression model.	modelspec (Linear, Interactions linear) and RobustOpts (Robust linear): different specifications used by the algorithm.	[106]
SLR (Linear)	A stepwise linear regression model. At each step, predictors are added or removed from a constant model, based on a certain criterion.	modelspec: different specifications used by the algorithm.	[107]
RT (Fine) RT (Medium) RT (Coarse)	A decision tree model for regression. A regression model with a hierarchical structure made by edges and nodes, these latter which make decisions.	MinLeafSize: minimum number of observations of each leaf (final nodes with no children nodes.) Fine: 4; Medium: 12; Coarse: 36.	[108]
SVM (Linear) SVM (Quadratic) SVM (Cubic) SVM (Fine Gaussian) SVM (Medium Gaussian) SVM (Coarse Gaussian)	A SVM regression model, which is a nonparametric one based on Kernel Functions (KFs).	KernelFunction (Linear, Quadratic, Cubic, Gaussian): KF used by the algorithm. Linear: Linear KF; Quadratic: polynomial KF of order 2; Cubic: polynomial KF of order 3; Fine/Medium/Coarse Gaussian: Gaussian KF. Kernelscale (Fine, Medium, Coarse) : value by which the values of the predictors matrix are divided. Fine: 1.1; Medium: 4.5; Coarse: 18.	[109]
EN (Boosted trees) EN (Bagged trees)	An EN regression model, which is a predictive model composed of a weighted combination of multiple regression trees.	Method: represents the ensemble aggregation method, Boosted trees or Bagged trees.	[110]
GPR (Squared Exponential) GPR (Matern 52) GPR (Exponential) GPR (Rational quadratic)	A GPR model is a probabilistic model, nonparametric and based on KFs.	KernelFunction: KF used by the algorithm.	[111]
KE (SVM) KE (Least Squares Regression)	A Gaussian Kernel regression model.	Learner: represents the linear regression model type.	[112]

(labeled "Overall") displays the MAE values obtained by considering all $\eta_{b,p}$ values. These values align with the MAE values presented in Table 17 for the best models under consideration.

LR models feature predictors that encompass linear model parameters, which are not only straightforward to interpret but also boast rapid predictive capabilities. These qualities have significantly boosted the popularity of LR models. However, the constrained model structures associated with LR models often result in lower predictive accuracy.

GPR model is very appropriate for developing the prediction model, especially in the case of scarce data sets.

The poorest performance is observed in the case of kernel approximation models, generally used for nonlinear regression of data with many observations. For large in-memory datasets, kernel approximation models tend to train and predict more quickly than SVM models with Gaussian kernels. Gaussian kernel regression models map predictors from a low-dimensional space to a high-dimensional space and then fit a linear model to the transformed predictors in the high-dimensional space. Kernel-based models, particularly with high-dimensional kernels, can lack interpretability, making it challenging to understand and explain the relationships they capture. These models are sensitive to the choice of hyperparameters, such as the kernel type and its parameters, which can require careful tuning. Furthermore Kernel-based models may struggle with generalization when the dataset has a high degree of noise or when the underlying relationships are highly complex and in the case of small datasets, may be prone to overfitting and not generalize well to unseen data.

Low prediction performance has been obtained also using Tree-based. For small to moderately sized datasets with clear feature importance requirements, tree-based models might be preferred, while for very large datasets with complex patterns, as in this case, neural networks outperform. For tree-based models, it's crucial to have an extensive training dataset that encompasses the full spectrum of changing trends. Without this comprehensive dataset, the tree-based models may struggle to deliver accurate predictions during testing phases.

The decrease in performance, however, could also be attributed to the limited availability of data for training the system in cases of higher degradation. The decision to supply networks with reduced data as the level of degradation intensifies is motivated by the aim to mimic real-world conditions, where degradation typically remains below 5% before any maintenance intervention and only escalates to higher levels in extreme scenarios. As the level of degradation increases, the occurrence of such cases becomes rarer. Despite the decline in performance for predicting high degradation cases in the mentioned models, the results remain within an acceptable range.

Additionally, a noteworthy detail is observed: all the MAE values obtained by considering only the $\eta_{b,p}$ values below 5% deviation from the healthy condition are higher than those obtained when considering all the $\eta_{b,p}$ values. This information suggests a greater ability of these models to accurately predict the $\eta_{b,p}$ variable when dealing with lower levels of degradation.

5.2. *Effect of measurement noise on performance prediction*

In a real-world scenario, measurements from sensors are often accompanied by a small source of disturbances that can slightly reduce accuracy in detection. The presence of measurement noises may influence the prediction performance of diagnostic systems. This section provides an analysis conducted to observe the impact of measurement noise on prediction performance. The analysis is conducted by introducing a noise source to the input data obtained from simulations and predicting the values of only one PP (as in a Multi-net scenario), i.e., $\eta_{b,p}$, for which the best prediction performance was achieved, as evidenced in Section 5. The introduced noise was simulated as a maximum percentage variation of the base value (absence of noise). The maximum percentage variations are derived from those available in [113,114] and are presented in Table 18.

The value of 0.1375% was considered as the average of the other values. In the present analysis, four different noise scenarios were considered: 25%, 50%, 75%, and 100% of the maximum added noise. In the case of 25% of the maximum added noise, a random percentage variation falling within the range {–25% of the maximum added noise: +25% of the maximum added noise} was applied to each base data point. For example, considering the TT2 variable, which is associated with a maximum added noise of ±0.4%, in the scenario with 25% of the maximum added noise, a random percentage variation within the range {–25% of 0.4%: +25% of 0.4%} is applied to each base value of TT2. The same process is followed for the other three cases. Using this approach, we investigate the impact of noise levels on prediction performance. The results are presented in Section 5.2.

Table 17
Comparison between the model used in this work (FFNN) and the ones available in the RL tool from Matlab 2023a. LR: Linear Regression; RT: Regression Tree; SVM: Support Vector Machine; EN: Ensemble; GPR: Gaussian Process Regression; NN: Neural Network; KE: Kernel. The used metrics are defined in Section 4.3.1.

Method	Matlab function	MAE	MSE	RMSE	COD
FFNN (used in this work)	fitnet()	2.47E-06	1.48E-11	3.85E-06	1.00E+00
LR (Linear)	fitlm()	4.98E-03	4.94E-05	7.03E-03	9.83E-01
LR (Interactions linear)	fitlm()	6.05E-04	7.82E-07	8.84E-04	1.00E+00
LR (Robust linear)	fitlm()	5.06E-03	1.30E-04	1.14E-02	9.74E-01
LR (Stepwise)	stepwiselm()	6.27E-04	8.24E-07	9.08E-04	1.00E+00
RT (Fine)	fitrtree()	1.51E-02	4.76E-04	2.18E-02	8.41E-01
RT (Medium)	fitrtree()	1.72E-02	6.12E-04	2.47E-02	7.96E-01
RT (Coarse)	fitrtree()	2.26E-02	1.04E-03	3.23E-02	6.52E-01
SVM (Linear)	fitrsvm()	5.54E-03	5.72E-05	7.56E-03	9.81E-01
SVM (Quadratic)	fitrsvm()	3.20E-03	1.51E-05	3.88E-03	9.95E-01
SVM (Cubic)	fitrsvm()	4.26E-03	3.14E-05	5.60E-03	9.89E-01
SVM (Fine Gaussian)	fitrsvm()	2.20E-02	9.93E-04	3.15E-02	7.04E-01
SVM (Medium Gaussian)	fitrsvm()	7.79E-03	1.42E-04	1.19E-02	9.57E-01
SVM (Coarse Gaussian)	fitrsvm()	2.25E-02	9.48E-04	3.08E-02	8.33E-01
EN (Boosted trees)	fitrensemble()	3.76E-02	1.65E-03	4.06E-02	9.27E-01
EN (Bagged trees)	fitrensemble()	1.20E-02	2.97E-04	1.72E-02	9.11E-01
GPR (Squared Exponential)	fitrgp()	3.51E-04	2.52E-07	5.02E-04	1.00E+00
GPR (Matern 52)	fitrgp()	3.42E-04	2.44E-07	4.94E-04	1.00E+00
GPR (Exponential)	fitrgp()	4.79E-03	4.40E-05	6.64E-03	9.86E-01
GPR (Rational quadratic)	fitrgp()	3.30E-04	2.21E-07	4.70E-04	1.00E+00
KE (SVM)	fitrkernel()	4.70E-02	4.16E-03	6.45E-02	6.72E-02
KE (Least Squares Regression)	fitrkernel()	3.99E-02	2.62E-03	5.12E-02	1.37E-01

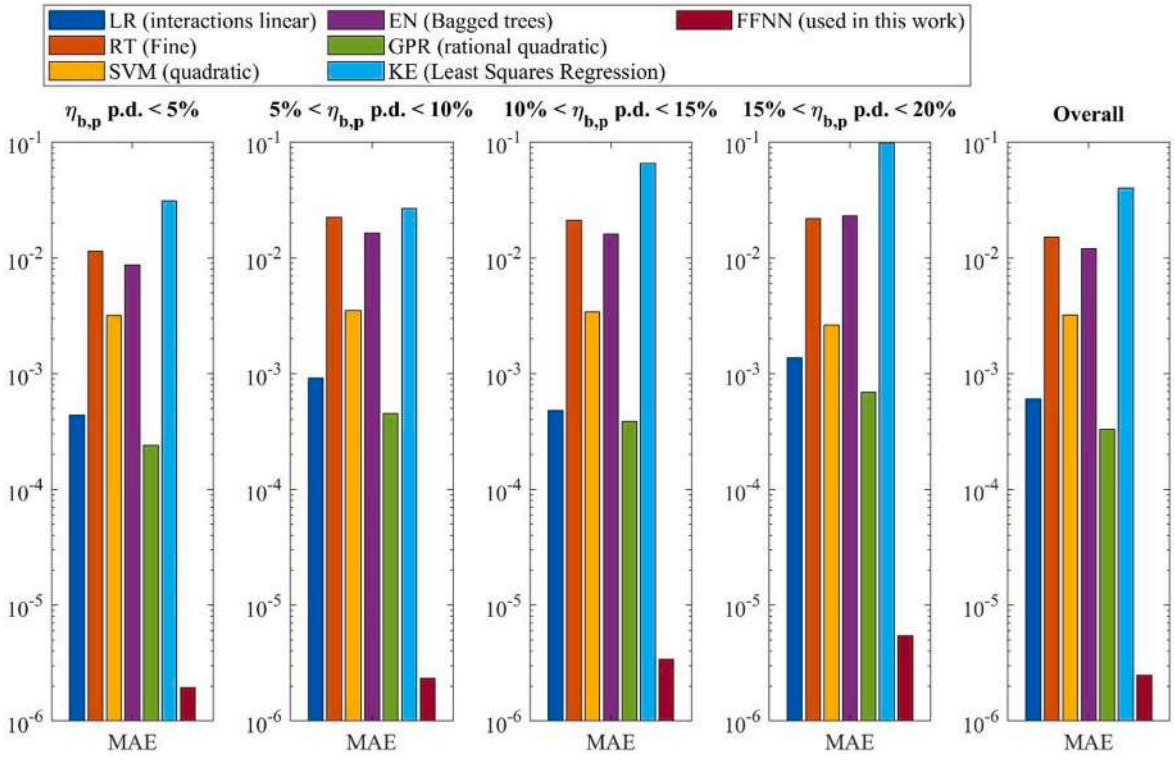

Fig. 19. MAE of the six best prediction methods for different p.d. of the predicted variable.

Table 18
Maximum noise levels considered for this analysis for each variable.

Input variable	Maximum added noise (as percentage variation)
PLA, h, v_f, k, I, V	±0.1375 %
ΔT, TT_2, TT_3, TT_4, TT_5, TT_6, w_f	±0.4 %
TP_2, TP_3, TP_4, TP_5, TP_6, OPR	±0.1 %
ω_{HPS}	±0.02%

In this section, the results of the analysis conducted to assess the impact of the presence of measurement noise on prediction performance are presented. The details of the analysis are available in Section 3.4. Fig. 20 illustrates the variations in MAE, CoD, and MaxAE metrics across the four different degradation scenarios considered, compared with the case of the absence of measurement noise (No noise). From Fig. 20, it is evident that there is a performance deterioration with the increase in noise level, as suggested by the corresponding increase in MAE. However, prediction performance remains acceptable, with MAE staying at the order of 10^{-4}. Additionally, CoD, consistently remaining close to 1, indicates high performance at each noise level.

6. Conclusions

The goal of this study is to develop a dependable EHM system for HEPSs, which are currently under-researched, by investigating the impact of various data reduction techniques. The proposed method has demonstrated the capability to efficiently and accurately predict component degradation in hybrid aircraft engines operating under different operating conditions. This method can assist in condition-based maintenance by enhancing the assessment of short-term performance status in such engine architectures.

The results suggest that a similar health monitoring approach could be applicable in a real-case scenario and serves as a starting point for the scientific community to develop more advanced diagnostic tools for hybrid propulsion systems, which are an emerging technology. The present work describes six EHM systems utilized to predict the health status of a HEPS employed in lightweight helicopters. Currently, health monitoring is crucial to ensure flight safety and cost savings, particularly with more complex systems like a HEPS. The cost increase estimate provided in Section 3.2 indeed highlights an increase of approximately 0.56 dollars per 10 km of cruise, passing from a healthy condition (6.42 dollars) to a degraded one (6.98 dollars). The developed systems were evaluated by comparing their predictions with target values obtained from simulations. Furthermore, an analysis of the effect of measurement noise was conducted, and the results suggest a negligible decrease in prediction performance with the considered noise levels. All the systems showed very good results with high accuracy. However, the main findings highlighted from the study are as follows:

- The Multi-net architecture requires less training time compared with the Single-net one.
- The Multi-net architecture provides improved prediction performance for nearly all of the predicted parameters, as evidenced by comparing the performance of System 1 and System 4 with their counterparts utilizing the Multi-net architecture, i.e., System 2 and System 5, respectively.
- The FE process seems to result in a worsening of the prediction performance, as evidenced by comparing System 1 and System 3 with System 4 and System 6, respectively.
- The FS process does not consistently yield positive effects in all cases, as evidenced by comparing System 2 and System 5 with System 3 and System 6, respectively.
- Due to the exceptionally high prediction performance provided by all the tested systems, the primary advantage of using a Multi-net architecture and incorporating FE and FS techniques becomes evident when comparing the total training times of each system.

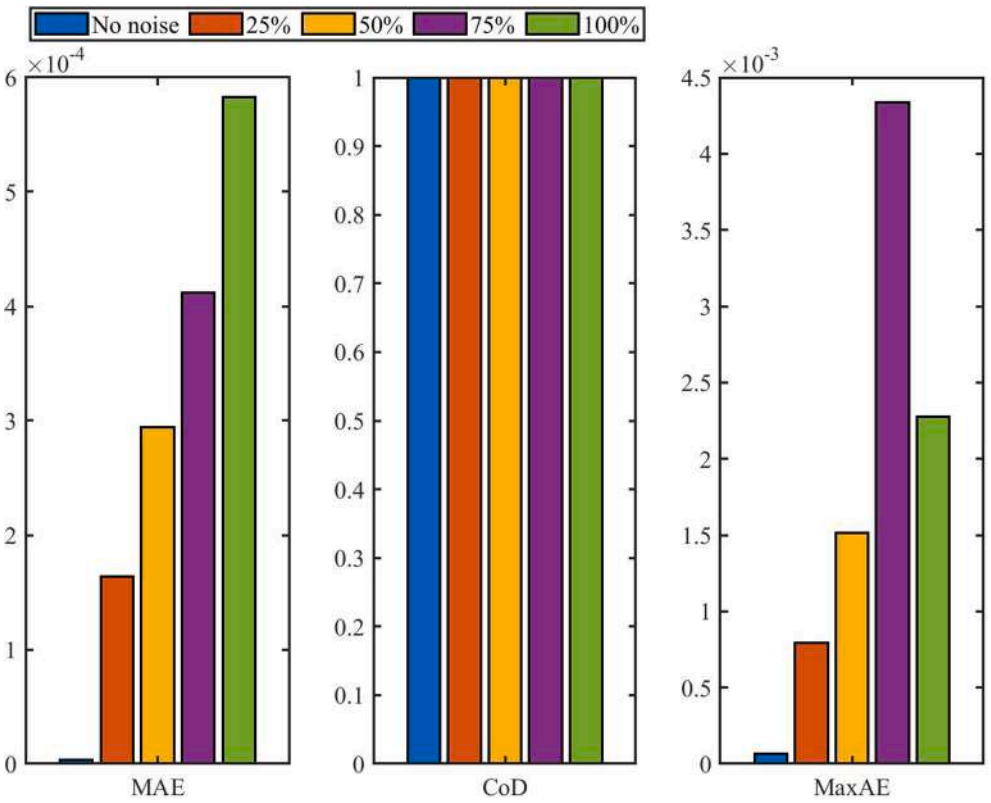

Fig. 20. MAE, CoD and MaxAE in $\eta_{b,p}$ prediction in different noise scenarios: No noise (base data), 25%, 50%, 75% and 100% of the maximum noise levels indicated in Table 18.

- The performance decrease resulting from the application of FE and, frequently, from the application of FS, might be compensated by the reduction in training time. This is especially relevant given that the overall performance remains quite high. The choice to apply a data reduction technique is ultimately up to the user.
- The time savings from incorporating a second data reduction technique in a system that already utilizes one may be minimal.

CRediT authorship contribution statement

Maria Grazia De Giorgi: Conceptualization, Methodology, Formal analysis. **Teresa Donateo:** Methodology, Software. **Antonio Ficarella:** Funding acquisition, Project administration. **Nicola Menga:** Data curation, Investigation, Software, Writing – original draft. **Ludovica Spada Chiodo:** Software, Writing – original draft. **Luciano Strafella:** Software.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

References

[1] S. Singh, S. Jain, V. Ps, A.K. Tiwari, M.R. Nouni, J.K. Pandey, S. Goel, Hydrogen: a sustainable fuel for future of the transport sector, Renew. Sustain. Energy Rev. 51 (2015) 623–633.

[2] N. Yilmaz, A. Atmanli, Sustainable alternative fuels in aviation, Energy 140 (2017) 1378–1386, https://doi.org/10.1016/j.energy.2017.07.077.

[3] J. Ribeiro, F. Afonso, I. Ribeiro, B. Ferreira, H. Policarpo, P. Peças, F. Lau, Environmental assessment of hybrid-electric propulsion in conceptual aircraft design, J. Clean. Prod. 247 (2020) 119477.

[4] A. Bravo, D. Vieira, G. Ferrer, Emissions of future conventional aircrafts adopting evolutionary technologies, J. Clean. Prod. 347 (2022) 131246, https://doi.org/10.1016/j.jclepro.2022.131246.

[5] J. Zhang, I. Roumeliotis, A. Zolotas, Nonlinear model predictive control-based optimal energy management for hybrid electric aircraft considering aerodynamics-propulsion coupling effects, IEEE Trans. Transport. Electrif. 8 (2022) 2640–2653, https://doi.org/10.1109/TTE.2021.3137260.

[6] A.A. Recalde, M. Lukic, A. Hebala, P. Giangrande, C. Klumpner, S. Nuzzo, P. H. Connor, J.A. Atkin, S.V. Bozhko, M. Galea, Energy storage system selection for optimal fuel consumption of aircraft hybrid electric taxiing systems, IEEE Trans. Transport. Electrif. 7 (3) (2021) 1870–1887.

[7] H. Zhao, C. Liu, Z. Song, W. Wang, Analytical modeling of a double-rotor multiwinding machine for hybrid aircraft propulsion, IEEE Trans. Transport. Electrif. 6 (4) (2020) 1537–1550.

[8] Y. Chen, T. Yang, M.R. Khowja, A.L. Rocca, U. Nasir, S. Chowdhury, D. Evans, D. Kember, T. Klonowski, Y. Arnaud, L. Apostin, T. Liger, G. Cossart, G. Vakil, C. Gerada, S. Bozhko, S. Detry, C. Diette, P. Wheeler, Mild hybridization of turboprop engine with high-power-density integrated electric drives, IEEE Trans. Transport. Electrif. 8 (4) (2022) 4148–4162.

[9] E. Sayed, M. Abdalmagid, G. Pietrini, N.-M. Sa'adeh, A.D. Callegaro, C. Goldstein, A. Emadi, Review of electric machines in more-/hybrid-/turbo-electric aircraft, IEEE Trans. Transport. Electrif. 7 (4) (2021) 2976–3005.

[10] M. Masiol, R.M. Harrison, Aircraft engine exhaust emissions and other airport-related contributions to ambient air pollution: a review, Atmos. Environ. 95 (2014) 409–455, https://doi.org/10.1016/j.atmosenv.2014.05.070.

[11] D. Adolfo, D. Bertini, A. Gamannossi, C. Carcasci, Thermodynamic analysis of an aircraft engine to estimate performance and emissions at LTO cycle, Energy Proc. 126 (2017) 915–922.

[12] Y.u. Wang, F. Xu, S. Mao, S. Yang, Y. Shen, Adaptive online power management for more electric aircraft with hybrid energy storage systems, IEEE Trans. Transport. Electrif. 6 (4) (2020) 1780–1790.

[13] L. Dorn-Gomba, J. Ramoul, J. Reimers, A. Emadi, Power electronic converters in electric aircraft: current status, challenges, and emerging technologies, IEEE Trans. Transport. Electrif. 6 (4) (2020) 1648–1664.

[14] H. Tian, P. Qin, K. Li, Z. Zhao, A review of the state of health for lithium-ion batteries: research status and suggestions, J. Clean. Prod. 261 (2020) 120813.

[15] Y. Kang, X. Yang, Z. Zhou, B. Duan, Q. Liu, Y. Shang, C. Zhang, A comparative study of fault diagnostic methods for lithium-ion batteries based on a standardized fault feature comparison method, J. Clean. Prod. 278 (2021) 123424.

[16] C. Friedrich, P.A. Robertson, Hybrid-electric propulsion for automotive and aviation applications, CEAS Aeronaut. J. 6 (2015) 279–290, https://doi.org/10.1007/s13272-014-0144-x.

[17] D.P. Decerio, D.K. Hall, Benefits of parallel hybrid electric propulsion for transport aircraft, IEEE Trans. Transport. Electrif. 8 (2022) 4054–4066, https://doi.org/10.1109/TTE.2022.3193622.

[18] M.A. Rendón, M. Assato, V.A.C. Martins, P.H. Hallak, A.S. Altgott, R. Graça, Z. Landy, N.L. Oliveira, R.G.P. Delmonte, Design method and performance analysis of a hybrid-electric power-train applied in a 30-passenger aircraft, J. Clean. Prod. 339 (2022) 130560.

[19] F.R. Gimenez, C.E.K. Mady, I.B. Henriques, Assessment of different more-electric and hybrid-electric configurations for long-range multi-engine aircraft, J. Clean. Prod. 392 (2023) 136171, https://doi.org/10.1016/j.jclepro.2023.136171.

[20] Y. Tang, K. Yang, H. Zheng, S. Zhang, Z. Zhang, Early prediction of lithium-ion battery lifetime via a hybrid deep learning model, Measurement 199 (2022) 111530.

[21] S.M. Rezvanizaniani, Z. Liu, Y. Chen, J. Lee, Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility, J. Power Sources 256 (2014) 110–124.

[22] M.-F. Ge, Y. Liu, X. Jiang, J. Liu, A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries, Measurement 174 (2021) 109057.

[23] W. Zhou, Q. Lu, Y. Zheng, Review on the selection of health indicator for lithium ion batteries, Machines 10 (2022) 512, https://doi.org/10.3390/machines10070512.

[24] S. Laribi, F. Zohra Arama, K. Mammar, N. Aoun, T. Ghaitaoui, M. Hamouda, State-of-health diagnosis of lithium-ion batteries using the fractional-order electrochemical impedance model, Measurement 211 (2023) 112653.

[25] G. Hong, W. Song, Y. Gao, E. Zio, A. Kudreyko, An iterative model of the generalized Cauchy process for predicting the remaining useful life of lithium-ion batteries, Measurement 187 (2022) 110269.

[26] W. Zhang, X. Li, X. Li, Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation, Measurement 164 (2020) 108052, https://doi.org/10.1016/j.measurement.2020.108052.

[27] M. El-Dalahmeh, M. Al-Greer, Ma'd El-Dalahmeh, I. Bashir, Physics-based model informed smooth particle filter for remaining useful life prediction of lithium-ion battery, Measurement 214 (2023) 112838.

[28] Z. Zhang, W. Zhang, K. Yang, S. Zhang, Remaining useful life prediction of lithium-ion batteries based on attention mechanism and bidirectional long short-term memory network, Measurement 204 (2022) 112093.

[29] Y. Cheng, M. D'Arpino, G. Rizzoni, Fault diagnosis in lithium-ion battery of hybrid electric aircraft based on structural analysis, in: 2022 IEEE Transportation Electrification Conference & Expo (ITEC), 2022, pp. 997–1004, doi:10.1109/ITEC53557.2022.9813976.

[30] H.M. Khalid, Q. Ahmed, J.C.H. Peng, Health monitoring of li-ion battery systems: a median expectation diagnosis approach (MEDA), IEEE Trans. Transport. Electrif. 1 (2015) 94–105, https://doi.org/10.1109/TTE.2015.2426431.

[31] A. Thelen, Y.H. Lui, S. Shen, S. Laflamme, S. Hu, H. Ye, C. Hu, Integrating physics-based modeling and machine learning for degradation diagnostics of lithium-ion batteries, Energy Storage Mater. 50 (2022) 668–695.

[32] L.S. Langston, Fahrenheit 3,600, Mech. Eng. 129 (2007) 34–37.

[33] W. Balicki, P. Głowacki, S. Szczecinski, R. Chachurski, J. Szczeciński, Effect of the atmosphere on the performances of aviation turbine engines, Acta Mechanica et Automatica 8 (2) (2014) 70–73.

[34] J. Li, Y. Ying, Gas turbine gas path diagnosis under transient operating conditions: a steady state performance model based local optimization approach, Appl. Therm. Eng. 170 (2020) 115025, https://doi.org/10.1016/j.applthermaleng.2020.115025.

[35] U. Igie, M. Goiricelaya, D. Nalianda, O. Minervino, Aero engine compressor fouling effects for short- and long-haul missions, Proc. Inst. Mech. Engin., Part G: J. Aerosp. Eng. 230 (7) (2016) 1312–1324.

[36] A.D. Fentaye, A.T. Baheta, S.I. Gilani, et al., A review on gas turbine gas-path diagnostics: state-of-the-art methods, Challeng. Opportun., Aerosp. 6 (2019) 83, https://doi.org/10.3390/aerospace6070083.

[37] I.S. Diakunchak, Performance deterioration in industrial gas turbines, J. Eng. Gas Turbines Power 114 (2) (1992), https://doi.org/10.1115/1.2906565.

[38] M.S. Grewal, Gas Turbine Engine Performance Deterioration Modelling and Analysis, 1988.

[39] K. Schadow, W. Horn, H. Pfoertner, Sensor and actuator needs for more intelligent gas turbine engines, in: Turbo Expo: Power for Land, Sea, and Air, 2010, pp. 155–167, doi:10.1115/GT2010-22685.

[40] L. Zhuang, G. Xu, B. Dong, et al., Study on performance and mechanisms of a novel integrated model with power & thermal management system and turbofan engine, Appl. Therm. Eng. 219 (2023) 119481, https://doi.org/10.1016/j.applthermaleng.2022.119481.

[41] V.K. Shanmuganathan, A.P. Haran, N. Gayathri, Condition monitoring maintenance of aero-engines through LUMS – a method for the implementation of Lean tools, Measurement 73 (2015) 226–230, https://doi.org/10.1016/j.measurement.2015.05.017.

[42] C. Guo, Study on the recognition of aero-engine blade-casing rubbing fault based on the casing vibration acceleration, Measurement 65 (2015) 71–80, https://doi.org/10.1016/j.measurement.2014.12.038.

[43] C. Zhao, X. Huang, Y. Li, S. Li, A novel remaining useful life prediction method based on gated attention mechanism capsule neural network, Measurement 189 (2022) 110637.

[44] Y. Wen, M.D. Fashiar Rahman, H. Xu, T.-L. Tseng, Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective, Measurement 187 (2022), https://doi.org/10.1016/j.measurement.2021.110276.
[45] H. Tian, L. Yang, B. Ju, Spatial correlation and temporal attention-based LSTM for remaining useful life prediction of turbofan engine, Measurement 214 (2023) 112816, https://doi.org/10.1016/j.measurement.2023.112816.
[46] C. Yang, H.u. Wu, J. Du, H. Zhang, J. Yang, Full-engine simulation of micro gas turbine based on time-marching throughflow method, Appl. Therm. Eng. 217 (2022) 119213.
[47] M.T. Yildirim, B. Kurt, Aircraft gas turbine engine health monitoring system by real flight data, Int. J. Aerosp. Eng. 2018 (2018) 1–12.
[48] H.E.G. Powrie, C.E. Fisher, Engine health monitoring: towards total prognostics, in: 1999 IEEE Aerospace Conference. Proceedings (Cat. No.99TH8403), 1999, doi:10.1109/AERO.1999.789759, pp. 11–20.
[49] Y. Ying, J. Li, An improved performance diagnostic method for industrial gas turbines with consideration of intake and exhaust system, Appl. Therm. Eng. 222 (2023) 119907, https://doi.org/10.1016/j.applthermaleng.2022.119907.
[50] J. DeRemer, Sand and dust erosion in aircraft gas turbines, J. Am. Soc. Nav. Eng. 62 (1950) 505–511.
[51] M.G. De Giorgi, S. Campilongo, A. Ficarella, Predictions of operational degradation of the fan stage of an aircraft engine due to particulate ingestion, J. Eng. Gas Turbines Power 137 (2015), https://doi.org/10.1115/1.4028769.
[52] S. Sripad, A. Bills, V. Viswanathan, A review of safety considerations for batteries in aircraft with electric propulsion, MRS Bull. 46 (2021) 435–442, https://doi.org/10.1557/s43577-021-00097-1.
[53] B. Rezaeianjouybari, Y. Shang, Deep learning for prognostics and health management: state of the art, challenges, and opportunities, Measurement 163 (2020) 107929, https://doi.org/10.1016/j.measurement.2020.107929.
[54] W. Jiang, Y. Xu, Z. Chen, N. Zhang, X. Xue, J. Zhou, Measurement of health evolution tendency for aircraft engine using a data-driven method based on multi-scale series reconstruction and adaptive hybrid model, Measurement 199 (2022) 111502.
[55] M.G. De Giorgi, A. Ficarella, L. De Carlo, Jet engine degradation prognostic using artificial neural networks, Aircr. Eng. Aerosp. Technol. 92 (2019) 296–303, https://doi.org/10.1108/AEAT-01-2018-0054.
[56] M.G. De Giorgi, S. Campilongo, A. Ficarella, A diagnostics tool for aero-engines health monitoring using machine learning technique, Energy Procedia 148 (2018) 860–867, https://doi.org/10.1016/j.egypro.2018.08.109.
[57] M.G. De Giorgi, S. Campilongo, A. Ficarella, K. Kontis, S. Pantelakis, Development of a real time intelligent health monitoring platform for aero-engine, MATEC Web Conf 233 (2018), https://doi.org/10.1051/matecconf/201823300007.
[58] M.T. Yildirim, B. Kurt, Engine health monitoring in an aircraft by using Levenberg-Marquardt feedforward neural network and radial basis function network, in: 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), 2016, pp. 1–5, doi:10.1109/INISTA.2016.7571847.
[59] S. Tang, H. Tang, M. Chen, Transfer-learning based gas path analysis method for gas turbines, Appl. Therm. Eng. 155 (2019) 1–13, https://doi.org/10.1016/j.applthermaleng.2019.03.156.
[60] Z. Wei, S. Jafari, S. Zhang, T. Nikolaidis, Hybrid Wiener model: an on-board approach using post-flight data for gas turbine aero-engines modelling, Appl. Therm. Eng. 184 (2021) 116350.
[61] H. Wang, W. Jiang, X. Deng, J. Geng, A new method for fault detection of aero-engine based on isolation forest, Measurement 185 (2021) 110064.
[62] K. Wang, Y. Guo, W. Zhao, Q. Zhou, P. Guo, Gas path fault detection and isolation for aero-engine based on LSTM-DAE approach under multiple-model architecture, Measurement 202 (2022) 111875.
[63] H. Ren, C. Guo, R. Yang, S. Wang, Fault diagnosis of electric rudder based on self-organizing differential hybrid biogeography algorithm optimized neural network, Measurement 208 (2023) 112355.
[64] H. Qin, R. Yang, C. Guo, W. Wang, Fault diagnosis of electric rudder system using PSOFOA-BP neural network, Measurement 186 (2021) 110058.
[65] A.Z. Nejad, A. Dastfan, Signal-based fault diagnosis in Q-pulse diode rectifier with fast and accurate detection of faulty switch, Measurement 207 (2023) 112331, https://doi.org/10.1016/j.measurement.2022.112331.
[66] Z. Jia, Z. Liu, Y. Cai, A novel fault diagnosis method for aircraft actuator based on ensemble model, Measurement 176 (2021) 109235, https://doi.org/10.1016/j.measurement.2021.109235.
[67] M. Bai, J. Liu, J. Chai, X. Zhao, D. Yu, Anomaly detection of gas turbines based on normal pattern extraction, Appl. Therm. Eng. 166 (2020) 114664.
[68] D.F. Finger, R. de Vries, R. Vos, et al, A comparison of hybrid-electric aircraft sizing methods, in: AIAA Scitech 2020 Forum, American Institute of Aeronautics and Astronautics, 2020, doi:10.2514/6.2020-1006.
[69] T. Donateo, A. Ficarella, A modeling approach for the effect of battery aging on the performance of a hybrid electric rotorcraft for urban air-mobility, Aerospace 7 (2020) 56, https://doi.org/10.3390/aerospace7050056.
[70] D.I. Stroe, M. Swierczynski, S.K. Kær, et al., A comprehensive study on the degradation of lithium-ion batteries during calendar ageing: the internal resistance increase, in: 2016 IEEE Energy Conversion Congress and Exposition (ECCE), 2016, pp. 1–7, doi:10.1109/ECCE.2016.7854664.
[71] T. Donateo, A. Terragno, A. Ficarella, An optimized fuzzy logic for the energy management of a hybrid electric air-taxi, in: E3S Web Conf, 2021. doi:10.1051/e3sconf/202131207004.
[72] T. Donateo, L. Spada Chiodo, A. Ficarella, A. Lunaro, Improving the dynamic behavior of a hybrid electric rotorcraft for urban air mobility, Energies 15 (20) (2022) 7598.
[73] T. Donateo, A. Ficarella, L. Spada Chiodo, Dynamic modeling and degradation study of a hybrid electric power system for urban air mobility, American Society of Mechanical Engineers Digital Collection, doi:10.1115/GT2023-101597.
[74] S.R. Gaudet, Development of a dynamic modeling and control system design methodology for gas turbines, 2008.
[75] Mapped motor and drive electronics operating in torque-control mode - Simulink - MathWorks Italia, https://it.mathworks.com/help/autoblks/ref/mappedmotor.html (accessed 4 July 2022).
[76] A. Hausmann, C. Depcik, Expanding the Peukert equation for battery capacity modeling through inclusion of a temperature dependency, J. Power Sources 235 (2013) 148–158, https://doi.org/10.1016/j.jpowsour.2013.01.174.
[77] What's the Best Battery? Battery University, https://batteryuniversity.com/article/whats-the-best-battery (2010, accessed 1 January 2022).
[78] M. Dubarry, B.Y. Liaw, Identify capacity fading mechanism in a commercial LiFePO4 cell, J. Power Sources 194 (2009) 541–549, https://doi.org/10.1016/j.jpowsour.2009.05.036.
[79] T. Donateo, L. Spada Chiodo, A. Ficarella, A dynamic model of a hybrid electric propulsive system for degradation assessment, J. Phys.: Conf. Ser. 2385 (1) (2022), https://doi.org/10.1088/1742-6596/2385/1/012060.
[80] L. Guzzella, A. Sciarretta, Vehicle Propulsion Systems, Introduction to Modeling and Optimization, third ed., Springer-Verlag, Berlin Heidelberg, Germany, 2013.
[81] T. Donateo, L. Cucciniello, L. Strafella, A. Ficarella, Control oriented modelling of a turboshaft engine for hybrid electric urban air-mobility, E3S Web Conf. 197 (2020) 05003.
[82] T. Donateo, C.L. De Pascalis, L. Strafella, A. Ficarella, Off-line and on-line optimization of the energy management strategy in a Hybrid Electric Helicopter for urban air-mobility, Aerosp. Sci. Technol. 113 (2021) 106677.
[83] T. Donateo, C.L.d. Pascalis, L. Strafella, A. Ficarella, Optimal energy management of a hybrid electric helicopter for urban air-mobility, IOP Conf. Ser.: Mater. Sci. Eng. 1024 (1) (2021) 012074.
[84] T. Rindlisbacher, L. Chabbey, Guidance on the determination of helicopter emissions, Edition 2, Federal office of civial aviation FOCA. (2009) Bern, Swiss.
[85] A. Filippone, N. Bojdo, Statistical model for gas turbine engines exhaust emissions, Transp. Res. D Transp. Environ. 59 (2018) 451–463, https://doi.org/10.1016/j.trd.2018.01.019.
[86] J. Cain, M.J. DeWitt, D. Blunck, E. Corporan, R. Striebich, D. Anneken, C. Klingshirn, W.M. Roquemore, R. Vander Wal, Characterization of gaseous and particulate emissions from a turboshaft engine burning conventional, alternative, and surrogate fuels, Energy Fuels 27 (4) (2013) 2290–2302.
[87] U.S. Gulf Coast Kerosene-Type Jet Fuel Spot Price FOB (Dollars per Gallon), https://www.eia.gov/dnav/pet/hist/eer_epjk_pf4_rgc_dpgD.htm (accessed 13 December 2023).
[88] A. Hasan, O.J. Haidn, Combustion of kerosene jet a fuel and superheated steam injection in an aviation turboshaft engine: improving power output and reducing emissions, J. Inst. Eng. India Ser C. 102 (2021) 275–281, https://doi.org/10.1007/s40032-020-00643-x.
[89] A. Hasan, O.J. Haidn, Jet A and Propane gas combustion in a turboshaft engine: performance and emissions reductions, SN Appl. Sci. 3 (2021) 471, https://doi.org/10.1007/s42452-021-04468-w.
[90] S. Outcalt, A. Laesecke, M.B. Freund, Density and speed of sound measurements of jet A and S-8 aviation turbine fuels, Energy Fuels. 23 (2009) 1626–1633, https://doi.org/10.1021/ef800888q.
[91] H. Hanachi, C. Mechefske, J. Liu, A. Banerjee, Y. Chen, Performance-based gas turbine health monitoring, diagnostics, and prognostics: a survey, IEEE Trans. Reliab. 67 (3) (2018) 1340–1363.
[92] S. Khalid, T. Khalil, S. Nasreen, A survey of feature selection and feature extraction techniques in machine learning, in: 2014 Science and Information Conference, 2014, pp. 372–378.
[93] C. Bloomer, G. Rehm, Using principal component analysis to find correlations and patterns at diamond light source, 2014.
[94] K.I. Kim, K. Jung, H.J. Kim, Face recognition using kernel principal component analysis, IEEE Signal Process Lett. 9 (2002) 40–42, https://doi.org/10.1109/97.991133.
[95] V. Dentamaro, D. Impedovo, G. Pirlo, LICIC: less important components for imbalanced multiclass classification, Information 9 (2018) 317, https://doi.org/10.3390/info9120317.
[96] B. Schölkopf, A. Smola, K.R. Müller, Nonlinear component analysis as a kernel eigenvalue problem, Neural Comput. 10 (1998) 1299–1319, https://doi.org/10.1162/089976698300017467.
[97] D. Feng, M. Xiao, Y. Liu, H. Song, Z. Yang, L. Zhang, A kernel principal component analysis–based degradation model and remaining useful life estimation for the turbofan engine, Adv. Mech. Eng. 8 (5) (2016), https://doi.org/10.1177/1687814016650169.
[98] X. Hu, N. Eklund, K. Goebel, A data fusion approach for aircraft engine fault diagnostics, in: Turbo Expo: Power for Land, Sea, and Air, 2007, pp. 767–775, doi:10.1115/GT2007-27941.
[99] A. Khemphila, V. Boonjing, Comparing performances of logistic regression, decision trees, and neural networks for classifying heart disease patients, in: 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM), 2010, pp. 193–198, doi:10.1109/CISIM.2010.5643666.
[100] E. Tosun, K. Aydin, M. Bilgili, Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures,

Alex. Eng. J. 55 (4) (2016) 3081–3089, https://doi.org/10.1016/j.aej.2016.08.011.
[101] T.e. Han, D. Jiang, Q.i. Zhao, L. Wang, K. Yin, Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery, Trans. Inst. Meas. Control 40 (8) (2018) 2681–2693.
[102] Z. Ibrahim, D. Rusli, Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression, in: 21st Annual SAS Malaysia Forum, 2007.
[103] M. Elforjani, S. Shanbr, Prognosis of bearing acoustic emission signals using supervised machine learning, IEEE Trans. Ind. Electron 65 (2018) 5864–5871, https://doi.org/10.1109/TIE.2017.2767551.
[104] S.E. Jozdani, B.A. Johnson, D. Chen, Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification, Rem. Sens. 11 (14) (2019) 1713, https://doi.org/10.3390/rs11141713.
[105] https://it.mathworks.com/help/deeplearning/ref/fitnet.html (accessed 8 August 2023).
[106] https://it.mathworks.com/help/stats/fitlm.html#bt0ck7o-mdl (accessed 8 August 2023).
[107] https://it.mathworks.com/help/stats/stepwiselm.html#bt0cpr2-mdl (accessed 8 August 2023).
[108] https://it.mathworks.com/help/stats/fitrtree.html#bt6cr84-tree (accessed 8 August 2023).
[109] https://it.mathworks.com/help/stats/fitrsvm.html (accessed 8 August 2023).
[110] https://it.mathworks.com/help/stats/fitrensemble.html (accessed 8 August 2023).
[111] https://it.mathworks.com/help/stats/fitrgp.html (accessed 8 August 2023).
[112] https://it.mathworks.com/help/stats/fitrkernel.html (accessed 8 August 2023).
[113] S.O.T. Ogaji, Y. G. Li, S. Sampath, et al, Gas path fault diagnosis of a turbofan engine from transient data using artificial neural networks, in: Turbo Expo: Power for Land, Sea, and Air, 2003, pp. 405–414. doi:10.1115/GT2003-38423.
[114] R.B. Joly, S.O.T. Ogaji, R. Singh, S.D. Probert, Gas-turbine diagnostics using artificial neural-networks for a high bypass ratio military turbofan engine, Appl. Energy 78 (4) (2004) 397–418.