Engine health monitoring with fuzzy data: lessons learned from aircraft industry

Luciano Sánchez
Universidad de Oviedo
Departamento de Informática
Gijón, Asturias, Spain
luciano@uniovi.es

Inés Couso
Universidad de Oviedo
Departamento de Estadística
Gijón, Asturias, Spain
couso@uniovi.es

***Abstract*—The use of data-driven techniques for health monitoring and prognosis of engines is often plagued by the lack of information about failure scenarios. This problem is aggravated when there are missing or partially missing records in the log data. Unluckily, the chance that a sequence of events leading to a failure is not properly stored is higher when the system exhibits an abnormal behaviour. In absence of patterns that can be related to deterioration, failure detection is often guided to find novelties in the records, or uncommon combinations of variables that may indicate an anomalous performance. Finding novelties in sequences of low quality data is a complex problem where numerous ambiguities must be resolved. Aircraft industry is a paradigmatic case in that it involves monitoring and prognosis with less than perfect data, partly because many different control and security subsystems prevent that failure conditions are actually reached. The use of fuzzy technologies for modelling and diagnosing aero-engines is discussed in this paper, along with the most challenging issues and future research lines, according to our own experience in this field.**

***Index Terms*—Fuzzy data, Engine health monitoring, Fault diagnosis, Fuzzy model**

I. Introduction

The main purpose of engine data is to monitor engine parameters in order to avoid running the engine under undesired conditions. Engine instrumentation is configured to trigger alerts related to operator action, maintainance action or, if a significant condition is found, shutting down the engine.

The development over time of the engine data is also monitored in a process that is called Engine Health Monitoring (EHM). EHM management systems have diagnostic and prognostic purposes, as not only individual working conditions but also the trend over time are examined in order to identify rapid levels of deterioration. EHM management systems estimate the Remaining Useful Life (RUL) of an engine, anticipating certain events or findings and therefore reducing the number and degree of engine refurbishments [1]. In the simplest case, EHM management systems compare engine data against those parameters identified to be characteristic of known engine conditions or against design limits [2]. However, predicting the engine parameter deterioration levels over time is complex. There are multiple methods of EHM data assessment developed, whose range of application depends on the type of engine and the amount and quality of the engine data.

This study concerns the application of fuzzy technologies for EHM in aircrafts, where multi-sensor information is used to report failures and predict the RUL of commercial turbofan engines. Aircraft industry has specific requirements for EHM systems, derived from the fact that gas turbines are inherently fail-safe and the reliability of the turbine may be higher than that of its sensors [3]. Sensor degradation cannot be tolerated because deviations in sensor readings can be mistaken for engine degradations and this can cause secondary failures because of wrong control decisions [4]. Hence, aircraft engines tend to be under-sensorized and a substantial uncertainty in the knowledge of the operating point of the engine is accepted. Uncertainty in EHM data is accounted for via mathematical models, statistical or intelligent techniques, and in many cases with the help of fuzzy technologies, which will be reviewed in the forthcoming sections.

The structure of this paper is as follows: Section II reviews the use of fuzzy technologies in aircraft EHM, discusses the strengths are weakness of these techniques, and states some open problems. A new family of fuzzy diagnostic tools is proposed in Section III. An explanatory application of the new tools is developed in Section IV. The conclusions of the paper are presented in Section V.

II. Review of the use of fuzzy technologies in aeroengine health monitoring

Most of the aircraft EHM assessment methods are based around Gas Path Analysis (GPA) [2]. The gas path components are all air-washed parts within the engine gas path: the compressors, the combustor and the turbines. The gas path components are susceptible to distinct different issues, such as worn seals, excessive tip clearances, burning, cracking or missing parts or sections of parts, etc. Changes in the internal working conditions of the engine are detected either by direct observation of EHM parameters or indirectly, through a suitable transform of the EHM data.

This work was supported by the Science and Innovation Spanish Ministry under Projects TIN2014-56967-R and TIN2017-84804-R

A. Direct methods: detection of novelties with pattern recognition

The most common GPA methodology consists in detecting changes or *novelties* [5] in the internal working conditions of the engine as early as possible, by finding specific combination of values or *signatures* of the different defects. Fuzzy expert systems have been used to identify engine trends and step changes [6]. These algorithms rely on an experience database comprising GPA variables sampled at engines with known faults. Recent approaches learn fuzzy models from clean engine data and deteriorations are detected as novelties in the EHM data or in the residuals between the measurements and the model predictions [7], [8].

The diagnosis of an engine is formulated as a supervised pattern recognition problem, often via a case-based reasoning algorithm in which the measurements are compared to the known signatures of defects that have been observed in the past [9]. Thus, pattern recognition techniques are only efficient if the experience database is complete; otherwise, unseen defects will remain undetected. This is a requirement hard to fulfil, as most of times the Full Authority Digital Engine Control (FADEC) and other safety mechanisms will prevent that the engine reaches these abnormal conditions. Also, there are faults that do not have a well defined signature: for instance, a higher than normal turbine temperature may have different causes, such as an incipient deterioration or a higher load of the engine. Pattern recognition cannot tell apart these root causes if temperature, fuel flow, pressures and thurst are treated in isolation.

B. Subspace methods and manifold learning

Since most of the deteriorations alter more than one GPA variable, some authors consider that each of these variables must be regarded as a mix of different factors [10]. In this respect, subspace methods project EHM variables into a low-dimensional factor space and the signatures of the defects are sought in this projection [11], [12]. In other cases, a non-linear transformation is used; for instance, neural-network based autoencoders condense the EHM information into a reduced number of variables [13]. Recently, other manifold learning techniques have been extended to uncertain data and applied to this problem. In [14], Blind Source Separation (BSS) for interval-valued data data has been applied to the EHM problem thus the independence among the factor is maximized. This last reference addresses the presence of epistemic uncertainty in the data. In any case, note that although the dimensionality reduction simplifies the aforementioned pattern recognition problem, the need for an experience database is still present.

C. Multivariate trend analysis of delta signals

One of the major sources of uncertainty in EHM is related to the scarcity of sensors. Measurements are taken under different flying conditions. Hence, certain changes in the monitored variables may be caused by either a deterioration in the engine or a change in the flight conditions. Furthermore, there is a wide variability between GPA related magnitudes for different engines, as the monitored variables of two engines flying in the same conditions might not be concordant from flight to flight. Following [15], this second variability is solved if defects are not sought in GPA variables but in the gradient of the so-called "delta" functions, which are the differences between GPA variables and the theoretical values of a single reference engine. The gradient of these delta functions are used, because useful information about the deterioration speed is conveyed by the slopes of these curves. However, since GPA variables combine measurements taken in different flying conditions, delta signals are very noisy. This poses a complex computational problem, as determining the gradient of a noisy signal is strongly dependent on the properties of the noise filter. A fuzzy-bandwidth filter was used to compute these slopes in [16], where a fuzzy rule-based classifier was subsequently used to monitor the engine on the basis of the multivariate gradient signal, and also in [17] fuzzy estimations of the RULs of compressor and turbine were produced for prognosis purposes. These methods still make use of a sample of engines with a known condition, that is used for learning classifier and RUL models. In contrast, the experience database with the signatures of the known defects are not needed anymore, as the monitoring and prognosis decisions are taken by data-driven fuzzy rule-based systems.

D. Sequence mining

Multivariate trend analysis transforms variable-length series of gradients into fixed-length feature vectors that are the inputs to a fuzzy classifier or model. This feature selection process discards the sequence order of the deteriorations. For example, suppose that two different motors #1 and #2 have bad compressors and turbines: the compressor of #1 failed first and its turbine was subsequently damaged because of the compressor bad state. In contrast, the turbine in #2 failed first and its compressor was damaged later. It may happen that their feature vectors in trend analysis are identical. In reference [18] the fixed-length transform is avoided, and specific sequential pattern mining-based classifiers (that can operate with variable-length fuzzy inputs) were leveraged. Sequence mining discovers conditions that are not to be found with multivariate trend analysis. In any case, other drawbacks exist: a sample of engines with known conditions is needed (although in this case the decision system does not require an experience database). Furthermore, frequent pattern matching techniques are used for identifying the signatures of the deffects. Hence, isolated deteriorations will not be found.

Summarizing, the conclusions of this review in fuzzy EHM for aircrafts are:

- Direct methods operate with experience databases comprising GPA values of engines with known deteriorations. These methods are sensible to uncertainties originated in the variability from engine to engine, the changing flight conditions and the lack of sensors.
- Fuzzy subspace methods alleviate the uncertainty caused by the lack of sensors, as different imprecise signals are

combined to extract independent factors. The experience database is still needed.
- Multivariate trend analysis relax the requisite of an experience database to a list of engines with known RULs, thus not-yet-seen defects can be discovered. These methods cannot detect the precise time when a defect appeared.
- Sequence mining cannot find defects with a low occurrence, because the signatures are obtained via fuzzy frequent pattern matching.

III. Isotonic fuzzy health models

Diagnosis and prognosis methodologies either depend on an experience database, a sample of engines with known RULs or a model of the engine whose residual can be monitored. The main advantage of model-dependent assessments lies in the possibility of finding defects not seen before, but being dependent on an engine model is a stringent requisite. In this section a new isotonic fuzzy health model of an aeroengine is proposed, whose residual can be related to the location of the deteriorations. This method is novel because it makes use of *the lowest possible amount of domain knowledge*, which is (i) certain variables in the engine are comonotonic, and (ii) the state of health of the engine decreases with time.

The input variables of the proposed model are the same "delta" variables introduced in Section II-C. In the particular case of the aeroengines concerned by this study (see Figure 1), these variables are called $\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}$ and ΔN2. The list of points of interest in the engine are:

- Station 3: This is the High Pressure Compressor (HPC) exit and the entry into the combustion system. The conditions at this point are key for the correct functioning of the engine. The main variables measured at this station are P30 (pressure) and T30 (temperature).
- Station 4: This is the combustion chamber exit and High Pressure Turbine (HPT) entry. The temperature at this point is one of the main engine parameters. T4, may also be known as Turbine Gas Temperature (TGT)
- Station 5: This is the Low Pressure Turbine (LPT) exit. The main variable at this station is P50. This pressure is used to define EPR, which is subsequently used to determine the overall engine thrust. EPR is the relation of P50 to P20.

The Low Pressure (LP) system is the combination of the fan and the LPT. The speed at which the LP system turns is defined as N1. The High Pressure (HP) system is the combination of the HPC and the HPT. The speed at which the HP system turns is known as N2. In addition, the amount of fuel consumed is also monitored through Fuel Flow (FF).

EPR is the relation of P50 (pressure at the Low Pressure Turbine exit) to P20 (pressure at the fan inlet). In two shaft high bypass ratio turbo fans, the thrust is performed by the air compressed by the fan blades and pushed through the engine bypass. The air pushed through the core of the engine is solely used to turn the fan. This is, the air is compressed by the high pressure compressor (HPC) so that the optimum conditions are reached within the combustion chamber to subsequently turn the high pressure turbine (HPT) to maintain the high pressure (HP) system and subsequently turn the low pressure turbine (LPT) which moves the fan and produces the engine thrust.

A. Proposed isotonic fuzzy model

Let the thrust of the considered engine be described by the following function:

$$\text{epr}(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2}, \mathbf{fc}) \tag{1}$$

where $\mathbf{fc} \in \text{FC}$ is a vector defining the flight conditions, ΔP30 is the difference between the gas pressure at the entry of the HPC and that of the reference engine, etc. Note that the thrust of the reference motor is $\text{epr}_0(0,0,0,0,0,\mathbf{fc})$. For engines that are in a good condition,

- $\text{epr}(\cdot)$ is comonotonical with ΔP30, ΔFF, ΔTGT, ΔN2 (the higher the compressor pressure, the consumed fuel, the turbine temperature and the turbine speed, then the higher the thrust is)
- $\text{epr}(\cdot)$ is antimonotonical with ΔT30 (the higher the compressor temperature, the less dense the air is, hence the thrust is lower)

Thus, the following fuzzy-valued model is proposed:

$$\begin{aligned}&\mu_{\widetilde{\text{EPR}}}(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2})(e) = \\ &\quad \sup_{\alpha}\{e \in [\overline{\text{EPR}}]_\alpha(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2})\}\end{aligned} \tag{2}$$

where the set $[\overline{\text{EPR}}]_\alpha$ is a confidence interval of the thrusts of the engine,

$$\begin{aligned}[\overline{\text{EPR}}]_\alpha = \{&\text{epr}_i(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2}, \mathbf{fc}) \\ &: \mathbf{fc} \in \text{FC}_\alpha\}\end{aligned} \tag{3}$$

and FC_α is the smallest subset of FC such that $\text{P}(\mathbf{fc} \in \text{FC}) \geq 1-\alpha$ and $\text{FC}_\alpha \subseteq \text{FC}_\beta$ for $\alpha > \beta$. Lastly, let the parametric definition of $\text{epr}(\cdot)$ be

$$\begin{aligned}&\text{epr}(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2}, \mathbf{fc}) = \\ &\quad f(\kappa_1 \cdot \Delta\text{P30} - \kappa_2 \cdot \Delta\text{T30} + \kappa_3 \cdot \Delta\text{TGT} + \kappa_4 \cdot \Delta\text{FF} + \\ &\quad\quad \kappa_5 \cdot \Delta\text{N2}) + g(\mathbf{fc})\end{aligned} \tag{4}$$

and

$$\begin{aligned}&[\overline{\text{EPR}}]_\alpha(\Delta\text{P30}, \Delta\text{T30}, \Delta\text{TGT}, \Delta\text{FF}, \Delta\text{N2}) = \\ &f(\kappa_1 \cdot \Delta\text{P30} - \kappa_2 \cdot \Delta\text{T30} + \kappa_3 \cdot \Delta\text{TGT} + \kappa_4 \cdot \Delta\text{FF} + \kappa_5 \cdot \Delta\text{N2}) \\ &\quad + [\gamma - \mu(1-\alpha), \gamma + \mu(1-\alpha)]\end{aligned} \tag{5}$$

with $\kappa_i \geq 0$, $g(\mathbf{fc}) >= 0$ for all $\mathbf{fc} \in \text{FC}$ and $f(x) \geq f(y)$ for $x > y$.

It can be safely assumed that engines are without deffects in the first cycles after a shop visit, and $\kappa_1, \ldots, \kappa_5$ and γ, μ can be found by maximizing the interval-valued extension of Kendall's tau [19] between $[\overline{\text{EPR}}]_\alpha$ and the actual thrust values

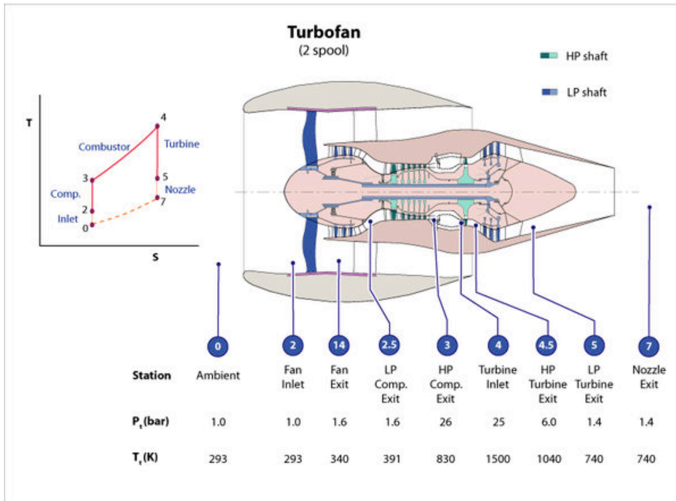

Fig. 1. Outline of a two shaft high bypass ratio turbo fan

EPR^t of the engine at the first $t = 1 \ldots N$ cycles, constrained by $\kappa_i \geq 0$, $\gamma > 0$, $\mu > 0$ and

$$\#\{\text{EPR}^t \in [\widetilde{\text{EPR}}]_\alpha(\Delta\text{P30}^t, \Delta\text{T30}^t, \Delta\text{TGT}^t, \Delta\text{FF}^t, \Delta\text{N2}^t)\} \geq (1-\alpha) \cdot N \quad (6)$$

The expression of the function f is obtained by interpolating the pairs $(\kappa_1 \cdot \Delta\text{P30}^t - \kappa_2 \cdot \Delta\text{T30}^t + \kappa_3 \cdot \Delta\text{TGT}^t + \kappa_4 \cdot \Delta\text{FF}^t + \kappa_5 \cdot \Delta\text{N2}^t + \gamma, \text{EPR}^t)$ with a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP).

Lastly, once the membership function of the fuzzy model is known,

$$\mu_{\widetilde{\text{EPR}^t}}(e) = \mu_{\widetilde{\text{EPR}}}(\Delta\text{P30}^t, \Delta\text{T30}^t, \Delta\text{TGT}^t, \Delta\text{FF}^t, \Delta\text{N2}^t)(e) \quad (7)$$

the residuals of this model can be regarded as a health indicator, as will be shown in an illustrative example in the next Section:

$$\mu_{\widetilde{\text{HEALTH}^t}}(e) = \mu_{\widetilde{\text{EPR}^t}}(e - \text{EPR}^t). \quad (8)$$

In words, the residuals of the proposed model will be centered at zero if the engine is not deteriored, but as soon as a degradation happens (decrease of ΔP30, alone or in combination of an increase of ΔT30, ΔTGT, etc.) the predicted thrust will be lower than that of the non-degraded engine, because of the signs of the coefficients and the monotonicity of the function f. Given that the health of an engine cannot be increased unless it undergones maintainance, isotonic regression techniques are also used to estimate the centerpoints of the aforementioned residuals.

IV. Illustrative example

The new method has been validated on a sample of 330 turbofan engines. A representative case has been chosen that illustrates the properties of this method.

Figure 2 depicts the GPA delta variables of an engine with a light deterioration in the HPT that cannot be found with direct methods, manifold learning, trend analysis neither sequence mining. A sixth variable DN2VIB has been added to the graph; this variable is not being used by the model but serves for validation purposes.

The isotonic fuzzy model has been fitted to the first $N = 1000$ cycles. The whole data spans more than 4000 cycles. Each GPA value is colored according to the EPR (between 1.20 and 1.50). The synthetic HEALTH signal is obtained by fitting a line with isotonic regression to the centerpoints of the residual of the fuzzy model proposed in the preceding section. Observe that the health tends to increase in the first half of the monitored period, but the isotonic regression forces that the signal is monotonically decreasing. Hence, the HEALTH signal is constant in the initial cycles. In the second part of the diagnostic, the residual is decreased and a deterioration in the second half of the figure becomes evident. This signal has been built without supplying the model any kind of information about the signature of a degradation in a similar engine: the only domain knowledge is the monotonicity or antimonotonicity of the GPA delta signals and the thrust. As an additional validation of the method, note that the signal DN2VIB measures the vibrations in the high speed shaft. This information has not been used by the isotonic model, that was nonetheless able to detect a deterioration immediately before the vibrations were measurable.

V. Concluding remarks and future work

Most of times, fault detection are regarded as pattern recognition methods, that can be solved with data-driven techniques with a different amount of domain knowledge. If a large database of engines with different degrees of deterioration is available, direct and subspace methods are the most common fuzzy techniques, however these cannot detect defects that are not in the experience database.

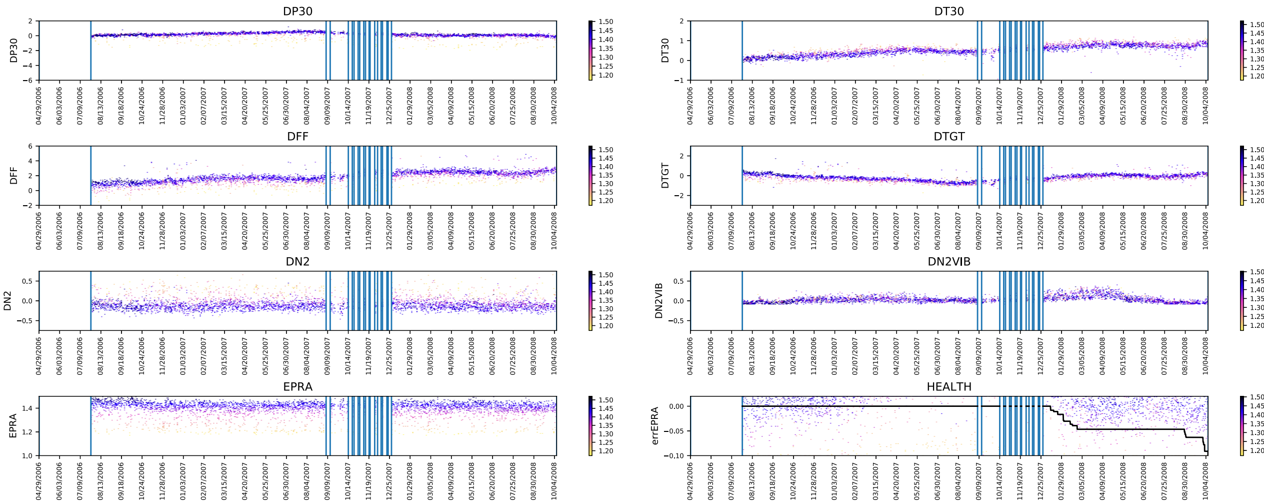

Fig. 2. Example of the application of the isotonic fuzzy model to an actual aeroengine. From left to right, top to bottom: Deltas of P30, T30, FF, TGT, N2, N2VIB (vibrations of the engine, not contemplated in this model) and synthetic HEALTH signal formed by the centerpoints of the isotonic fuzzy model. Observe that the model detects a deterioration of the engine in the last cycles, and this prediction is consistent with the independently measured vibration N2VIB.

The use of delta variables with respect to a reference engine allows applying machine learning algorithms to learn an EHM management system from data, where fuzzy technologies are used for designing the noise filters. The main drawbacks of these methods are the impossibility of detecting the order of the events and the precise cycle when the deterioration initiates. The use of sequence mining solves the first issue, and the isotonic algorithm introduced in this study solves the second problem. Besides its simplicity, the use of a model with a minimal amount of domain knowledge reduces the systematic error of specific models, and at the same time its generalization capabilities improve that of black boxes.

This is a work in progress; there are some aspects of the application of isotonic learning to EHM monitoring that are still unsolved. In particular, it is not yet clear how to attribute the descent in the thrust to the turbine or the compressor. Once this issue is solved, the next step in this research will be to assess the dependence between the RUL of the engine and the jumps in the health signal.

References

[1] Y. Li, "Gas turbine performance and health status estimation using adaptive gas path analysis," *Journal of Engineering for Gas Turbines and Power*, vol. 132, no. 4, p. 041701, 2010.

[2] L. A. Urban, "Gas Path Analysis Applied to Turbine Engine Condition Monitoring," *Journal of Aircraft*, vol. 10, no. 7, 2012.

[3] A. R. Behbahani, "Need for robust sensors for inherently fail-safe gas turbine," in *52nd International Instrumentation Symposium: Engine Controls, Monitoring, and Prognostics. Cleveland, OH*, 2006.

[4] ATSB, "Australian transport safety bureau: Power plant failures in turboprop-powered aircraft 2012 to 2016," https://www.atsb.gov.au/publications/2017/ar-2017-017/, accessed: 2018-06-06.

[5] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Processing*, vol. 99, pp. 215–249, 2014.

[6] S. Demirci, C. Hajiyev, and A. Schwenke, "Fuzzy logic-based automated engine health monitoring for commercial aircraft," *Aircraft Engineering and Aerospace Technology*, vol. 80, no. 5, pp. 516–525, 2008.

[7] A. Benyounes, A. Hafaifa, A. Kouzou, and M. Guemana, "Gas turbine modeling using adaptive fuzzy neural network approach based on measured data classification," *Mathematics-in-Industry Case Studies*, vol. 7, no. 4, 2017.

[8] Z. Li, S.-S. Zhong, and L. Lin, "Novel gas turbine fault diagnosis method based on performance deviation model," *Journal of Propulsion and Power*, pp. 1–10, 2016.

[9] C. H. Fontes and O. Pereira, "Pattern recognition in multivariate time series–a case study applied to fault detection in a gas turbine," *Engineering Applications of Artificial Intelligence*, vol. 49, pp. 10–18, 2016.

[10] Z. Chen, H. Fang, and Y. Chang, "Weighted data-driven fault detection and isolation: A subspace-based approach and algorithms," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3290–3298, 2016.

[11] R. Baklouti, M. Mansouri, M. Nounou, H. Nounou, and A. B. Hamida, "Iterated robust kernel fuzzy principal component analysis and application to fault detection," *Journal of Computational Science*, vol. 15, pp. 34–49, 2016.

[12] I. Gueddi, O. Nasri, K. Benothman, and P. Dague, “Fault detection and isolation of spacecraft thrusters using an extended principal component analysis to interval data,” *International Journal of Control, Automation and Systems*, vol. 15, no. 2, pp. 776–789, 2017.

[13] W. Yan and L. Yu, “On accurate and reliable anomaly detection for gas turbine combustors: A deep learning approach,” in *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, 2015.

[14] A. Martinez, L. Sanchez, and I. Couso, “Interval-valued blind source separation applied to ai-based prognostic fault detection of aircraft engines.” *Journal of Multiple-Valued Logic & Soft Computing*, vol. 22, 2014.

[15] A. Volponi and B. Wood, “Engine health management for aircraft propulsion systems,” in *Proceedings of the Forum on Integrated System Health Engineering and Management (ISHEM) in Aerospace, Napa, CA, November*, 2005, pp. 7–10.

[16] A. Martínez, L. Sánchez, and I. Couso, “Engine health monitoring for engine fleets using fuzzy radviz,” in *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–8.

[17] A. Martinez, L. Sánchez, and I. Couso, “Aeroengine prognosis through genetic distal learning applied to uncertain engine health monitoring data,” in *Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1945–1952.

[18] A. Palacios, A. Martínez, L. Sánchez, and I. Couso, “Sequential pattern mining applied to aeroengine condition monitoring with uncertain health data,” *Engineering Applications of Artificial Intelligence*, vol. 44, pp. 10–24, 2015.

[19] O. Hryniewicz and K. Opara, “Efficient calculation of kendall’s tau for interval data,” in *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*. Springer, 2013, pp. 203–210.