Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-3 (2020) 113–118

A clustering approach to detect faults with multi-component degradations in aircraft fuel systems

Anna Zaporowska [a], Haochen Liu [a], Zakwan Skaf [b], Yifan Zhao [a*]

[a] *Through-life Engineering Services Centre, Cranfield University, Cranfield, MK43 0AL, UK*
[b] *Integrated Vehicle Health Management, Cranfield University, Cranfield, MK43 0AL, UK*
**Corresponding author, email: yifan.zhao@cranfield.ac.uk*

Abstract: Accurate fault diagnosis and prognosis can significantly increase the safety and reliability of engineering systems and also reduce the maintenance costs. There is very limited relative research reported on the fault diagnosis of a complex system with multi-component degradation. The Complex Systems (CS) problem, which features multiple components simultaneously and nonlinearly interacting with each other and corresponding environment on multiple levels, has become an essential challenge in system engineering. In CS, even a single component degradation could cause misidentification of the fault severity level and lead to serious consequences. This paper introduces a new test rig to simulate multi-component degradations of the aircraft fuel system. A data analysis approach based on machine learning classification of both the time and frequency domain features is then proposed to detect and identify the fault severity level of CS with multi-component degradation. Results show that a) the fault can be sensitively detected with an accuracy > 99%; b) the severity of fault can be identified with an accuracy of 100%.

Keywords: Fast Fourier Transform; Clustering analysis; K-means clustering; Fault Diagnosis

1. INTRODUCTION

In the past few decades, engineering systems have become increasingly complex. At the same time, the demands on the reduction in life cycle costs of the engineering systems are increasing. These make MRO (Maintenance, Repair and Overhaul) a more challenging task in the industry. For example, in the aviation market, the MRO cost is expected to rise to $116 billion by 2029, up from $81.9 billion in 2019 (Cooper et al. 2019). How to reduce this magnitude of cost and increase the safety and availability of the aircraft systems at the same time is a crucial issue faced by the aviation industry.

Current maintenance strategy can be categorised into two types, Preventive (or Scheduled) Maintenance (PM) and Condition-Based Maintenance (CBM). As a traditional manner, PM is regularly conducted on the whole system and components to avoid failure. While CBM is widely carried out when certain indicators show fault signs or degradation of the system. For highly critical or very important assets with high costs, compared with PM, CBM includes both fault diagnosis and prognosis, which significantly increase the safety and reliability of the system and also reduce the maintenance costs. Fault diagnosis is one of the main parts of CBM aiming to shut down the system and schedule a maintenance task when an abnormality is detected.

Fault diagnosis is a relatively mature subject when applied at the component level. However, in terms of the effective maintenance on system level, it is still challenging to diagnose faults in a complex engineering system. For instance, quickly identifying a specific LRU (Line-Replaceable Unit) in an aircraft is still a tough task. This paper focuses on diagnosing faults on the system level. In an engineering system, degradation occurs on every component and varies in diverse pace through the whole service life. It is a practical challenge attracting attentions, while previous researches focus on the scenario that only part of the faulty component is considered.

Current fault identification methods can be divided into two categories: qualitative and quantitative (Venkatasubramanian, Rengaswamy, & Kavuri, 2003; Rogers, Guo & Rasmussen, 2019; Li, de Oliveira & Cerrada, 2019). Qualitative methods include two subcategories: graph theory such as fault tree methods (Johnson & Gormley, 2011) and signed digraph (SDG) methods, and expert systems such as conventional expert system, fuzzy expert system and belief rule-based method (Yang, Wang, Dong, & Liu, 2012). Quantitative methods include two subcategories: model-based and data-driven methods (Rogers, Guo & Rasmussen, 2019). Model-based methods such as observer/filter-based methods (Zhang & Pisu, 2014), parameter estimation methods and parity relation methods (Ríos, Davila & Fridman, 2014). Data-driven methods such as signal processing methods, machine learning methods (Saimurugan & Nithesh, 2016), statistical and hybrid methods (Muralidharan & Sugumaran, 2012). Considering identifying a single fault in a system with multi-component degradation, the single method mentioned above is incompetent due to respective limits. A combination of both qualitative and quantitative methods will be promising to fit this scenario. Lin et al. (Lin, Zakwan, & Jennions, 2017) proposed a probabilistic framework to combine multi-component degradation information when diagnosing a fault

10.1016/j.ifacol.2020.11.018

in a system, however, the identification of degradation has not been addressed.

The data-driven methods (Lin et al., 2017) is a powerful alternative to diagnose a fault on system level, but a large amount of historical data captured is inevitable. In practical cases, it is always difficult to capture a long time historic data of all degraded state information of the target system, especially for those advanced and large systems that contain several hundred of degraded components. In addition, this research only considers the faults that are already known by previous experience or knowledge (i.e., the historic data). The method cannot deal with an unknown fault in the system. In addition, the framework in Lin et al., 2017 is incapable to take the severity of multi-component degradation identification into account.

In the case of the systems with multi-component degradations, it is a huge challenge to adequately and quickly identify the location and severity of each degradation because it is a typical Complex System (CS), which features multiple components simultaneously and nonlinearly interacting with each other and corresponding environment on multiple levels. Even a single component degradation could cause misidentification of the fault severity level and lead to serious consequences (Lin et al., 2017). In this case, distinguish different faults in the system is a problematic but crucial task. Current fault diagnostic methods usually employ information from nearby sensors as reinforcement. However, the assumption of reinforcement information captured from healthy nearby components is commonly made in the previous report. Without approach improvements, it would be a huge inaccuracy in diagnosed results when all sensory information considered are from degraded components. And serious mistakes and failure would occur in the identification of the fault severity level. Therefore, an approach which is suitable for a system with multi-component degradation is meaningful and urgent.

Addressing the above challenges, this paper initially introduces a laboratory scale of the aircraft fuel system to replicate failure and degradation modes of real flight-worthy components. A novel data-driven approach based on clustering and classification for two features from time and frequency domains respectively is then proposed to detect and identify the severity level of fault.

2. METHODS

2.1 Experiment setup

The aircraft fuel system has a great impact on flight safety and most of the accidents associated with fuel system have led to hazardous and even catastrophic events. The fault diagnosis for the fuel system can improve aircraft safety and reliability and finally increase the aircraft's availability and efficiency.

This section provides a comprehensive description of the experimental fuel rig which includes the hydraulic system, the control and measuring system, and the fault injection mechanism.

A. Hydraulic System

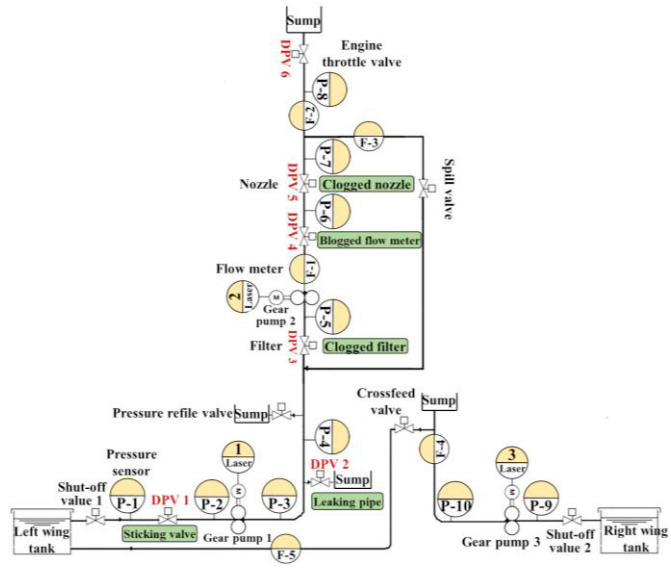

Fig. 1. The layout of the fuel rig system

Fig. 1 illustrates the layout of the fuel rig, which consists of three fuel tanks, three gear pumps, five shut-off valves, ten pressure sensors (marked P1-P10), and six direct proportional valves (DPVs). All the components are connected with pipes and mounted on an aluminium optical breadboard (1.8m x 1.1m x 5cm) which is above a drip tray to catch any unintended leak in the system. Two main tanks act as the left-wing tank and right-wing tank of an aircraft respectively, and a sump tank representing the engine that receives the fuel from the aircraft fuel system. Each gear pumps is driven by an external motor drive and has a pressure-relief valve inside to prevent the gears overstressing. Specifically, the fuel test rig can be divided into three lines:

- As illustrated in Fig. 1, the engine fuel feed line consists of the Shut-off valve 1, a non-return valve (emulated by the DPV1), two gear pumps (Gear pump 1 and Gear pump 2), a pressure-relief valve (a shut-off valve), a filter (emulated by the DPV3), a flow meter (emulated by the DPV4) and a nozzle (emulated by the DPV5). The two gear pumps serve as the low-pressure pump and high-pressure pump, respectively. The pressure-relief valve opens when the overpressure condition happens.
- The cross-feed line includes the shut-off valve 2, a cross-feed valve (a shut-off valve), and Gear pump 3 which transfers the fuel from the right-wing tank to the left-wing tank to help to maintain the central gravity of the aircraft during flight.
- The spill line includes a spill valve (a shut-off valve) which returns the fuel when the engine fueling is adequate. An engine throttle valve (emulated by the DPV6) generates the backpressure when the spill valve is opened.

B. Fault Types

For an aircraft fuel system, the common faults can be classified into three different types, namely process faults, actuator faults and sensor faults. Process faults mainly affect the operational ability of the system such as a leaking pipe or cracked joint. Actuator faults affect the actuated parts of the system such as pump malfunction or a sticking valve, and sensor faults mainly impact the sensor operation. In the aircraft fuel system, fouling, erosion, or corrosion are the main reasons for components degradation. Table 1 defines the faults simulated corresponding to the designed test rig.

Table 1. The faults injected into the test rig

Fault type	Fault
Process fault	Leaking pipe
Actuator fault	Sticking valve
	Clogged filter
	Clogged nozzle
Sensor fault	Blocked flow meter

To inject various faults, with different degrees of severity into the fuel test rig, five DPVs were used. The failure of the shut-off valve (being stuck in amid range position) is implemented by using DPV1 and set to be fully open at the beginning. The fully opened DPV1 represents a healthy valve status while the partially closed DPV1 represents a sticking valve with a certain degree of fault severity. Different degrees of severity can be emulated by varying the opening percentage of the DPV which is controlled through the developed software and can be varied from 0% to 100%. DPV2 emulates a leaking pipe fault into the system and set to be initially fully closed. The fully closed DPV2 represents a healthy pipe while the partially opened DPV2 represents a leaking pipe with a certain degree of severity. DPV3 injects a clogged filter fault into the system. The fully opened DPV3 represents a healthy filter while the partially closed DPV3 represents a clogged filter with a certain degree of severity. DPV4 injects a blocked flow meter fault into the system. The fully opened DPV4 represents a healthy flow meter while the partially closed DPV 4 represents a blocked flow meter with a certain degree of severity. DPV5 injects a clogged nozzle fault into the system. The fully opened DPV5 represents a healthy nozzle while the partially closed DPV5 represents a clogged nozzle with a certain degree of severity.

2.2 Experiment design

In this study, DPV2 simulated the leaking pipe, as a system fault, through varying the opening percentage of this valve with values of 0% (healthy), 30%, 40%, and 50%. For each opening percentage of DPV2, the leaking percentage of four valves (Sticking valve, Clogged filter, Blocked flow meter, Clogged nozzle shown in Fig. 1) was changed from 0%, 10% to 20%, respectively, to simulate a multi-component degradation. In other words, for each of these four valves, there were 3 possible states. There were 3^4=81 possible combinations for each fault severity level. The combination index, expressed as $\sum a_i 3^i$ (a_i indicates the state of the ^{i}th valve), presents an exclusive location and severity of this multi-component degradation. Therefore, the identification of this complex degradation problem was simplified to the estimation of the combination index. Firstly, this study aims to detect and identify the fault when the severity level was just above the threshold (30%) of fault and degradation, which allowed preventing further damage propagation at the earliest possible stage. Additionally, faults with a severity level exceeding 50% were significantly easier to be detected and identified than the lower ones.

2.3 Data analysis

Due to the complexity of this system, applying a model-based approach to monitor the system condition is difficult because an analytical or numerical model cannot represent the complex interaction of the whole system, although each individual component can be modelled successfully. Additionally, any change of system design will require the model to be re-estimated, which limits the universality of such an approach. This paper proposes a model-free clustering method based on two features extracted from the time and frequency domains respectively for fault detection, classification and identification.

The proposed method has three objectives including a) fault detection: determining if the system is faulty or healthy, b) fault classification: dividing the sampled data into a number of groups using an unsupervised machine learning approach, which mainly evaluates the performance of selected features identifying different severity levels of fault, and c) fault identification: determining the severity level of fault. Initially, all pressure sensors are considered, and then unnecessary sensors are gradually screened to identify an optimal channel eventually. The details of each process are described below.

A. Feature extraction

Assuming that the system is stationary under a specific severity level of fault and degradation combination, the mean amplitude of each channel or each pressure sensor (P1, P2, …, P8 in Fig. 1) is regarded as the first feature. The channels P9 and P10 are not considered because they are in a different branch with the components that have fault and degradation. It has been observed that the amplitude of pressure in the time domain is a good feature due to its sensitivity to different severity levels of fault. On the other hand, the frequency response or Power Spectrum Density (PSD) of each channel is regarded as the second feature.

B. Fault detection and classification

In the clustering analysis, grouping a set of objects into multiple categories is the main task. The most common method is the k-means clustering. It is an unsupervised analysis based on the theory of the maximising similarity inside a class while the minimising similarity outside the class. Given a set of observations ($x_1, x_2, \cdots, x_n$), where each

observation is a d-dimensional real vector, the k-means clustering aims to partition the n observations into $k(\leqslant n)$ sets $S=\{S_1,S_2,\cdots,S_k\}$ so as to minimise the within-cluster sum of squares. In this study, the observation is a 2-dimensional vector $\{\bar{P}_l, f_l\}$ and sampled number n is 4×81=324 which include four fault severity levels (0%, 30%, 40%, and 50%) and each level includes 81 combinations of degradation.

For the task of fault detection, assuming a Gaussian distribution of features, an 95% confidence ellipse in the 2-dimensional feature space is calculated based on the healthy data. An ellipse can be written as Eq. (1)

$$\frac{(x-c_x)^2}{a^2}+\frac{(y-c_y)^2}{b^2}=1 \tag{1}$$

where (c_x, c_y) is the centre, and a and b are two radiuses of horizontal and vertical directions. To extend the equation to a more general case where the ellipses could be rotated, the Principle Component Analysis is applied to calculate the 2 by 2 coefficients PC. If a feature vector (v_1, v_2) is located inside the ellipse or determined as healthy, the following condition shown in Eq. (2) must be satisfied or this vector is located outside the ellipse or determined as faulty.

$$\frac{((v_1-c_x)\cdot PC')^2}{a^2}+\frac{((v_2-c_y)\cdot PC')^2}{b^2}=1 \tag{2}$$

For the task of fault classification, the number of group k is chosen as 4 and the unsupervised k-means method is applied. Given the knowledge of the ground truth of classes, the performance of classification can be evaluated by Adjusted Rand Index, Mutual Information, Silhouette Coefficient.

For the task of fault identification, one approach is to establish the 95% confidence ellipse for each level of fault (healthy, 30%, 40% and 50%). Another approach is to use the supervised learning approaches, such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Complex Tree and Boosted Tree. The confusion matrix is used to measure the identification results and helps select the optimal channels for monitoring regarding the experiments conducted in this paper. Typical derivations from a confusion matrix include true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The accuracy of classification is given by Eq. (10) to describe the percentage of the correctly classified samples.

$$ACC=\frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

3. RESULTS

In order to demonstrate the feasibility of the proposed framework, the leaking pipe case study has been tested. The threshold for fault decision is set as 30% of degradation level (i.e. the opening percentage of DPV2).

3.1 Fault detection

The clustering results using the mean amplitude and peak frequency are shown in Fig. 2, where different colours indicate different severity level of fault, and each level of fault includes 81 cases. It is suggested from visual inspection that P1, P3, P7 and P8 have good performance to separate these four groups (healthy, 30%, 40% and 50%). Considering P3 for example, the mean amplitude can separate them into 4 groups but the groups of 30%, 40% and 50% are not very well separated. The peak frequency can separate 30%, 40% and 50%, but healthy and 50% are overlapped. The clustering performance is significantly improved if both features are considered.

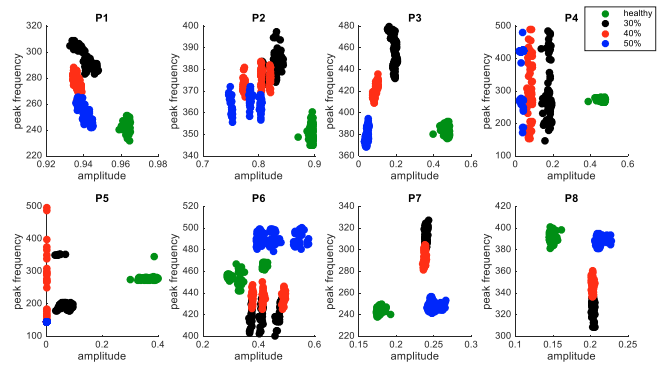

Fig. 2. Clustering visualisation for all 8 considered channels, where each faulty case has 81 combinations of degradation.

Based on the above observation, the channel P3 was selected for fault detection. The 95% confidence ellipse for the healthy group was calculated and illustrated by Fig. 3. If a tested vector locates inside this ellipse, the system is determined as healthy, otherwise, the system is faulty. The similar approach can be applied to other 7 channels and the fault detection results are shown in Table 2. All 8 channels produce a sound performance with the false positive number of less than 5 of 324 (1.54%).

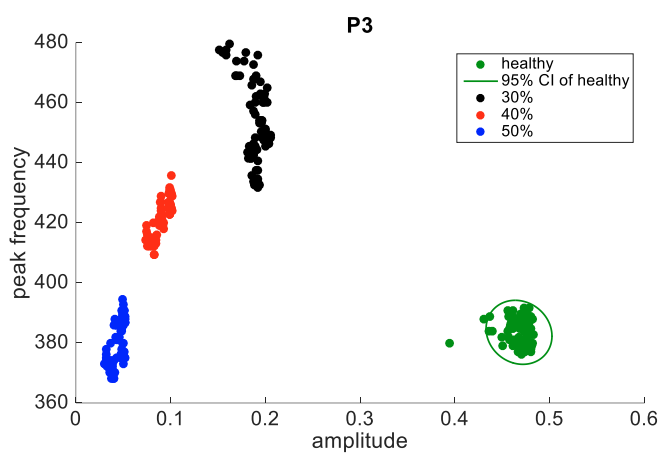

Fig. 3. An example of fault detection based on the channels P3, where the eclipse represents the 95% confidence level of healthy behaviour.

Table 2. Performance quantification of fault detection

Fault type	Fault	Accuracy
P1	4	97.86%
P2	2	99.38%
P3	2	99.38%
P4	2	99.38%
P5	2	99.38%
P6	3	99.07%
P7	3	99.07%
P8	4	97.86%

3.2 Fault classification

Base on the above results, the channels P1, P3, P7 and P8 were selected for the fault classification and identification. The k-mean approach was applied to cluster the data shown in Fig. 2, where k was chosen as 4. The classification results are shown in Fig. 4. Since it is an unsupervised approach, the corresponding between the colour and severity level is still unknown. Comparison of Fig. 2 (ground truth) and Fig. 4 (classification results) suggests that P3 produce a 100% classification result while the other three channels produce sound results. To quantify the performance, Table 3 shows the Adjusted Rank Index (ARI), Mutual Information (MI) and Silhouette Coefficient (SC) between the true tags and the produced tags from the k-means methods for these four channels. If two features are used, all three criteria suggest that P3 has the best performance. P8 has the second-best performance and P1 and P7 have slightly worse performance. Table 3 also compares the clustered performance between the single feature and two featured. For P3, the amplitude itself produces satisfactory results. While for the other three channels, the performance is significantly improved when two features are used rather than a single feature.

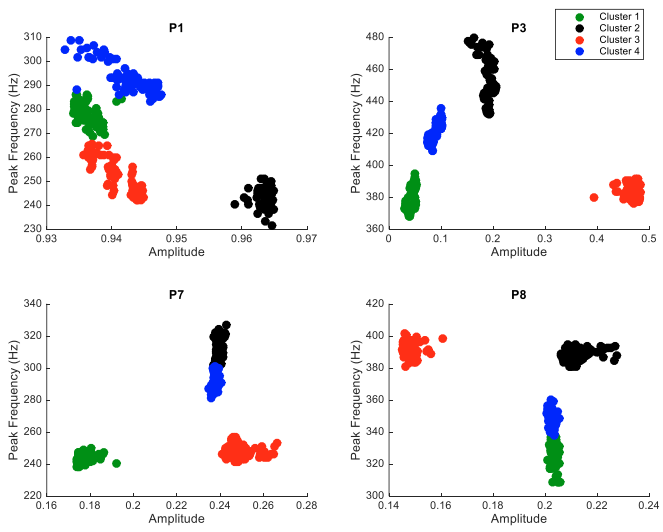

Fig. 4. Classification results of fault severity level using the k-means approach (k=4).

Table 3. Performance quantification of fault classification

Channel	Amplitude + PSD			Amplitude			PSD		
	ARI	MI	SC	ARI	MI	SC	ARI	MI	SC
P2	0.960	1.892	0.765	0.482	1.086	0.823	0.680	1.471	0.770
P3	1.000	2.000	0.955	1.000	2.000	0.959	0.603	1.407	0.868
P4	0.960	1.915	0.870	0.573	1.318	0.892	0.686	1.485	0.701
P5	0.984	1.958	0.893	0.511	1.271	0.804	0.594	1.411	0.822

3.3 Fault identification

To identify the severity level of fault, the 95% confidence ellipses were estimated for each group in P1, P3, P7 and P8, and the results are shown in Fig. 5. It can be observed that P3 produces the best result where there is no overlap between ellipses. For P1, there is a small region of overlap between 40% and 50%. For P7 and P8, there is a small region of overlap between 30% and 40%. The detailed performance of this approach can be described by the confusion matrixes shown in Table 4. The total accuracy for each considered channel is shown in Table 5, which again proves that P3 produces the best performance of fault identification.

Other learning-based classification approaches, namely KNN, SVM, Complex Tree and Boosted Tree, have also been applied to identify the fault level and results are shown in Table 6. The channels P3 still has the best performance. It should be noted that although the performance of some of these methods is better than the 95% confidence ellipse, the models are lack of transparency and cannot be really written down.

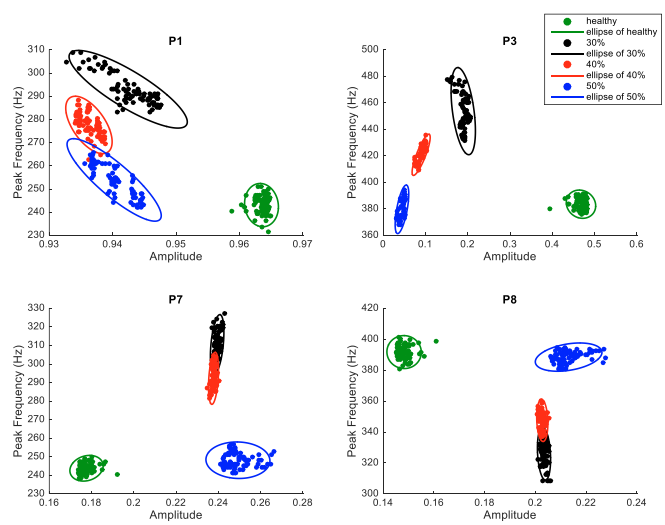

Fig. 5. The result of fault identification using the 95% confidence ellipse.

Table 4. The confusion matrix of the fault severity level identification using the 95% confidence ellipse

		Predicted					
		Healthy	30%	40%	50%	Unknown	Total
P1 Actual	Healthy	77	0	0	0	4	81
	30%	0	78	0	0	3	81
	40%	0	0	79	2	0	81
	50%	0	0	0	80	1	81
	Total	77	78	79	82	8	
P3 Actual	Healthy	79	0	0	0	2	81
	30%	0	76	0	0	5	81
	40%	0	0	79	0	2	81
	50%	0	0	0	81	0	81
	Total	79	76	79	81	9	
P7 Actual	Healthy	78	0	0	0	3	81
	30%	0	78	18	0	3	99
	40%	0	25	79	0	2	106
	50%	0	0	0	78	3	81
	Total	78	103	97	78	11	
P8 Actual	Healthy	77	0	0	0	4	81
	30%	0	76	5	0	5	86
	40%	0	16	78	0	3	97
	50%	0	0	0	78	3	81
	Total	77	92	83	78	15	

Table 5. The accuracy of the fault severity level identification using 95% CI ellipse

Channel	Accuracy
P1	96.9%
P3	97.2%
P7	85.3%
P8	89.6%

Table 6. The accuracy of the fault severity level identification using selected supervised learning-based approaches.

Channel	KNN	Linear SVM	Complex Tree	Boosted Tree
P1	99.1%	99.1%	96.0%	87.3%
P3	100%	100%	100%	100%
P7	97.5%	98.1%	97.8%	69.4%
P8	98.8%	99.1%	98.5%	74.7%

4. CONCLUSIONS

To address the challenge of faulty and degradation diagnosis of a complex engineering system where the multi-component degradations are presented, this paper has presented a test rig and corresponding data analytical approach for three tasks: fault detection, fault classification and fault identification. The results allow the following conclusions:

- The discrepancy of the pressure amplitude between healthy and faulty scenario depends on the fault location. The P3-P5 have more than 40% amplitude discrepancy while other sensors have less than 6% amplitude discrepancy when the fault is emulated between P3 and P4. The change of amplitude of certain channels is capable to detect, classify and identify the fault.
- For the cases where the severity level of fault is the same while the degradation level increases, there is no regular pattern of amplitude change, which suggests the amplitude cannot be used for degradation identification. The shift of the frequency peak, showing a linear trend of decrease following the increment degradation level, is an effective feature to identify the degradation level.
- By clustering analysis using two features of time-domain amplitude and frequency-domain peak, the fault can be detected with an accuracy > 97%; the severity of fault can be identified with an accuracy > 99%.

A limitation of this approach is that it only works on stationary systems where the fault or degradation severity level does not change when shifted in time. To widen its applications, future studies will focus on its extension on dynamical systems and real aircraft components.

REFERENCES

Cooper, T., Smiley, J., Porter, C. and Precourt, C. (2019), Global Fleet & MRO Market Forecast Commentary, Olover Wyman, New York.

Johnson, S., & Gormley, T. (2011). System Health Management: with Aerospace Applications . Wiley.

Li, C., de Oliveira, J., Cerrada, M., Cabrera, D., Sanchez, R. and Zurita, G., (2019). A Systematic Review of Fuzzy Formalisms for Bearing Fault Diagnosis. IEEE Transactions on Fuzzy Systems, 27(7), 1362-1382.

Lin, Y., Zakwan, S., & Jennions, I. (2017). A Bayesian Approach to Fault Identification in the Presence of Multi-component Degradation. International Journal of Prognostics and Health Management, 8, 4–12.

Muralidharan, V., & Sugumaran, V. (2012). A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Applied Soft Computing, 12(8), 2023–2029.

Ríos, H., Davila, J., Fridman, L. & Edwards, C., (2014). Fault detection and isolation for nonlinear systems via high-order-sliding-mode multiple-observer. International Journal of Robust and Nonlinear Control, 25(16), 2871-2893..

Rogers, A., Guo, F. & Rasmussen, B., (2019). A review of fault detection and diagnosis methods for residential air conditioning systems. Building and Environment, 161, 106-236.

Saimurugan, M., & Nithesh, R. (2016). Intelligent Fault Diagnosis Model for Rotating Machinery Based on Fusion of Sound Signals. International Journal of Prognostics and Health Management, 7, 18–27.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. Computers & Chemical Engineering, 27(3), 313–326.

Zhang, X., & Pisu, P. (2014). An Unscented Kalman Filter based on-line Diagnostic approach for PEM fuel cell Flooding. International Journal of Prognostics and Health. International Journal of Prognostics and Health Management, 5(1).