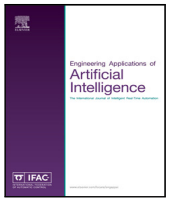
Engineering Applications of Artificial Intelligence 114 (2022) 105016

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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



Flight regimes recognition in actual operating conditions: A functional data analysis approach

Jessica Leoni [a,*], Francesco Zinnari [a], Eugenia Villa [a], Mara Tanelli [a,b], Andrea Baldi [c]

[a] Politecnico di Milano, Dipartimento di Elettronica Informazione e Bioingegneria (DEIB), Milan, Italy
[b] Istituto di Elettronica e Ingegneria dell'Informazione e delle Telecomunicazioni, Torino, Italy
[c] Leonardo S.p.A., Helicopters Division, Cascina Costa di Samarate (VA), Italy

ARTICLE INFO

Keywords:
Regime recognition
Unsupervised learning
Functional data analysis
Helicopters

ABSTRACT

Helicopters need adequate monitoring to prevent dynamic failures from excessively affecting components' health status, increase the level of safety, and reduce operative costs. Health and Usage Monitoring Systems have been developed to monitor helicopters during their lifetime in the last few decades. Recent works demonstrated that despite analyzing physical components' behavior over time, tracking the regimes performed during each flight contributes to estimating the aircraft's health and usage status, paving the way for designing accurate prognostics algorithms. However, today, most regime recognition systems rely on data recorded during certification flights. It follows that the training regimes differ from the ones proposed in the prediction phase, which are acquired during helicopter actual operating conditions. This affects these recognition system performances. Aiming at overcoming this limitation, in this work, we proposed an unsupervised regimes recognition system capable of better handling the actual helicopter usage spectrum. In detail, we proposed a system based on an unsupervised learning paradigm, which leverages a soft-membership classification technique to account even for mixed regimes and transitions. In addition, the system represents data according to functional data analysis theory, which allows for considering the temporal relationship between samples in the classification process, often neglected in state-of-the-art approaches. The proposed system was tested on experimental data, collected by Leonardo Helicopter Division, assessing outstanding capabilities in recognizing correctly standard and mixed regimes and transients. Also, the presented results demonstrate the approach capabilities in paving the way for the definition of new regimes, more consistent with the actual helicopter usage spectrum.

1. Introduction

Helicopters need adequate monitoring to prevent dynamic failures from excessively affecting components' health status, increase the level of safety, and reduce operative costs. Due to dynamic loads and vibrations, mechanical parts may fail if not adequately maintained. In order to prevent mechanical failures, the manufacturers provide an estimate of the component's design life, and, based on this value, maintenance operations are scheduled to ensure flight safety. This estimate is, however, formulated using an assumed helicopter usage spectrum, which may not reflect the actual operating conditions (Lombardo, 1998). Thus, different scenarios can be outlined depending on the actual rotorcraft usage, as shown in Fig. 1, retrieved from Romero (1996).

When the helicopter is used more severely than the manufacturers anticipated at design time, a potential safety risk arises due to fatigue. Also, helicopters are often used in contexts different from those assumed at design time (Şenipek and Kalkan, 2019). It follows that aging and fatigue may compromise the helicopter.

To help this situation, during the North Sea operations conducted in the second half of the '80s, the UK Government started promoting the development of Health and Usage Monitoring Systems (HUMS). The first generation of these systems was designed to detect mechanical faults, preventing early accidents. Through sophisticated signal processing techniques, these systems primarily focus on the analysis of the components' vibration signature acquired through a set of accelerometers. Although the introduction of these systems more than halved the number of accidents caused by structural failures, disasters such as the crashes in Scotland in 2009 and Norway in 2016 demonstrate that there is still room for HUMS improvement (Branch, 2009).

In this sense, one of the main directions undertaken foresees to integrate the dynamic components usage with information related to

* Corresponding author.
E-mail addresses: jessica.leoni@polimi.it (J. Leoni), francesco.zinnari@polimi.it (F. Zinnari), eugenia.villa@polimi.it (E. Villa), mara.tanelli@polimi.it (M. Tanelli), andrea.baldi01@leonardocompany.com (A. Baldi).

https://doi.org/10.1016/j.engappai.2022.105016
Received 17 January 2022; Received in revised form 27 May 2022; Accepted 1 June 2022
Available online 16 June 2022

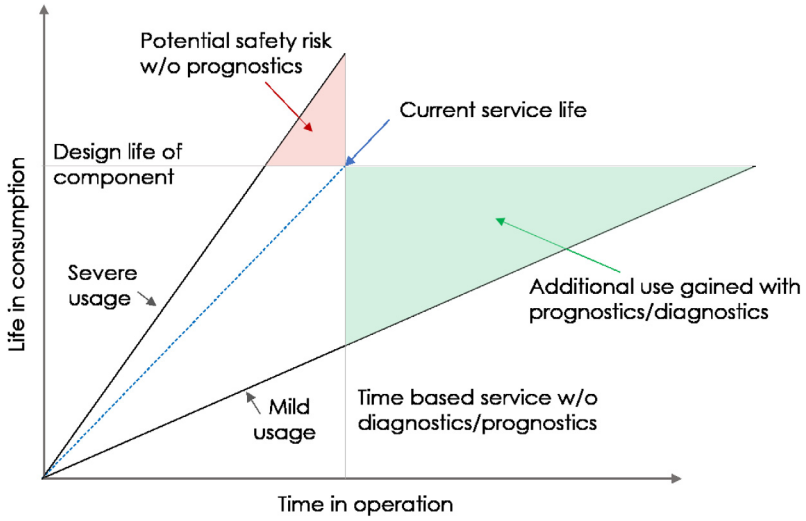


Fig. 1. Component Design and Actual Life. This figure illustrates the current service life trend for a component respect to the one estimated at design time, based on the actual helicopter usage.

the actual helicopter usage. Aware of this necessity, the Federal Aviation Administration, in a 2007 document, pointed out that a high priority is to be attributed to the development of a robust method for regime recognition (Le and Cuevas, 2007). An accurate method would allow attributing a damage factor to each component relative to the identified severity of the flight regime, resulting in a more precise estimation of the actual state of fatigue usage. Besides improving the diagnostic capabilities of HUMS, this method paves the way for the development of the second generation of systems, also capable of precise prognostics (He et al., 2010). In this sense, monitoring and tracking fatigue accumulation would allow managing another interesting case depicted in Fig. 1, *i.e.,* when the helicopter's actual usage is milder than what was assumed at design time, delaying the need for components replacement. Consequently, it would be possible to achieve the long-standing goal of switching from a time-based to a condition-based maintenance scheduling, allowing both a considerable operating costs reduction and a flight safety increase (Berry et al., 2006).

Since the late 1990s, several regime recognition methods have been presented in the literature. To help the reader outline a general framework, two main approaches can be distinguished: threshold-based and machine-learning-based. For both categories, the most relevant works are reported in Table 1.

The first category reflects an approach commonly shared in the avionics field, *i.e.,* combining the analysis of the collected data with a priori domain knowledge to define a set of thresholds suitable for the intended purpose (Berry et al., 2006; Teal et al., 1997). The main limitation of these methods is that criteria overlap in the distinction between two or more regimes can lead to improper recognition. Additionally, this approach is not easily scalable with the number of analyzed signals and detectable regimes due to the increasing number of rules required for more advanced recognition.

On the other hand, machine-learning-based approaches aim to improve flight regime recognition by training a classifier to automatically identify underlying patterns in the provided signals that characterize each regime. Several supervised methods have been proposed in the literature. In He et al. (2010), Hidden Markov Models are used to recognize 50 of the Army UH-60L flight card dataset regimes, obtaining an accuracy of 99%. In Şenipek and Kalkan (2019), the authors evaluated the effectiveness of different techniques in recognizing 57 different regimes. The supervised approach obtained a 69.2% accuracy using Neural Networks.

In addition, recently, deep-learning approaches overcome the performance of machine-learning algorithms, relying on neural networks with increasingly complex architectures. As an example, in Wu et al. (2022), an LSTM is proposed to recognize 6 different maneuvers, assessing 94.50% accuracy with a signal-to-noise ratio of 20. Moreover, in HanYang et al. (2020) the authors present a deep-variational autoencoder architecture, capable of recognizing 16 maneuvers with an accuracy of 99.94% and 91.03% on the two examined case studies, respectively. However, although there is no standard definition for the regimes, some of the 16 classes predicted are easily recognized by analyzing a single signal, *e.g.*, the *low-speed state* class. Also, deep-learning-based approaches rely on decision-making processes that lack human interpretability, affecting their compliance with in-flight safety requirements.

Regardless their undoubted performances, these supervised learning methods can only be trained using labeled flight conditions, which are usually recorded during load survey flights. Indeed, in order for a helicopter to be marketed, a certification phase is mandatory. Standard regimes are performed and recorded on the flight card, while on-board sensors record a set of parameters. On-ground, the measured data are analyzed to verify, for each regime, their compliance with the a priori requirements defined by the certification institution. However, the regimes performed in this context are executed according to precise instruction (regarding the duration of the regime, angles, speeds, altitudes), unlike during the actual rotorcraft usage, characterized by mixed regimes and frequent transitions. Despite the attempts to mathematically define each flight regime, as in Thomson and Bradley (1990), a standard definition was never provided. This compromises these methods' effectiveness, affecting their generalization capabilities (Warner and Rogers, 2019). Therefore, even today, the proposed regime recognition methods' main problem is that of suffering from an overreliance on the training data. Indeed, the difference between load survey conditions and actual operating conditions renders them unreliable when handling actual regimes that do not exactly match the learned ones.

Aware of these limitations, the most recent methods, including those mentioned above, abandoned *hard classification* approaches, which attribute to each sample a single label, in favor of *soft classification* ones, which return the probability of a sample to belong to each possible category. However, as reported in Table 1, these approaches are affected by poor accuracy. We believe that even better performances may be achieved by separating the learning process from the labels, *i.e.,* moving to unsupervised approaches. This would allow obtaining a more robust method for handling the actual regime spectrum, even when this does not precisely match the load survey data.

Table 1
State of the Art. This table depicts the major achievements in flight regimes recognition considering both threshold-based and machine-learning-based approaches.

Ref	Supervised	Algorithm	# Regimes	Accuracy	Data	Mixed regimes
Teal et al. (1997)	✓	Logical Threshold	Not Specified	90.0%	Real	✗
Lombardo (1998)	✓	K-Nearest Neighbors	5	94.4%	Real	✗
Gene et al. (2007)	✓	Logical Threshold	Not Specified	Not Specified	Real	✓
Berry et al. (2006)	✓	Artificial Neural Networks	11	76.2%	Real	✓
He et al. (2010)	✓	Hidden Markov Models	50	99.0%	Real	✗
Şenipek and Kalkan (2019)	✗	Gaussian Mixture Models	21 57	98.2% 91.5%	Simulated	✗
Musso and Rogers (2020)	✓	Interacting Multiple Models	10	84.1%	Simulated	✓
Wu et al. (2022)	✓	Convolutional Neural Networks	6	99.94%	Simulated	✗
Wu et al. (2022)	✓	Variational Autoencoders	16	94.50%	Simulated	✗

Therefore, in this paper, we propose a method capable of performing regime recognition based on hidden patterns in the collected signals. The labels will not be used in the training phase but will be employed only afterward in the clustering process to interpret and evaluate the obtained results. To the best of the author's knowledge, this is one of the first attempts to use an unsupervised approach for regime recognition purposes. Indeed, the only work proposing an unsupervised approach leverages Gaussian Mixture Models to identify 57 regimes, claiming a 91.5% accuracy (Sheridan et al., 2020). However, it seems that every sample of the collected time-series is treated independently, thus losing the temporal relationship between consecutive samples.

In our approach, instead for first time the temporal relationship between consecutive samples is considered in the classification process, thanks to the multivariate functional data representation of the signals of interest. This allows to better represent complex maneuvers, and to manage their spread both in terms of patterns and of duration.

In detail, we propose the combination of Functional Data Analysis (FDA) and Fuzzy C-Means Clustering, an approach capable of performing regime recognition accurately and robustly. FDA representation allows considering each signal as a continuum, emphasizing temporal dependencies among samples. This is a relatively recent technique, which has proved effective in several applications (Ullah and Finch, 2013). This work is the first application of FDA to regime recognition, despite it being particularly suitable. Indeed, using the functional representation allows us to entirely retain the signals' dynamic information, which would be (at least partially) lost by collapsing them into statistical feature vectors (*e.g.*, mean, variance).

Regarding the clustering algorithm, Fuzzy C-Means' was chosen instead of the classic K-Means due to the advantage of soft classification techniques compared to hard ones. By investigating the obtained clusters', we can often match them to existing flight regimes. Perhaps even more valuable is the possibility of discovering new regimes, more compatible with the actual helicopter usage. Moreover, we define a suitable similarity metric between clusters to investigate the positions assumed in the multidimensional space by standard regimes and their respective transitions, to investigate the similarity between regimes.

The proposed method was tested on a real dataset, collected on a selected helicopter model from Leonardo Helicopters Division (LHD). The dataset structure was fundamental to demonstrate the effectiveness of our method in handling mixed regimes and transitions since, in addition to standard regimes, it contains data acquired during transitions between several regime pairs.

The rest of this article is structured as follows: Section 2 depicts the structure of the dataset employed as a case study, along with the designed pre-processing phase. Then, the proposed method is illustrated in Section 3, dedicated to the FDA representation description, and in Section 4, which details the actual regime recognition process. Section 5 defines the considered evaluation metrics, respect to which reports and discusses the obtained results. Finally, Section 6 exposes the final remarks of the produced work and the future perspectives that it paves the way for.

2. Data pre-processing and selection

2.1. Dataset description

The HUMS installed on the selected helicopter acquires 30 signals with a sampling frequency in the order of ten Hertz. Such signals provide an extensive description of the helicopter's dynamics during flight time, including information on the aircraft's position, motion, attitude, and environmental conditions.

We analyzed more than a hundred load survey flights, for a total amount of 30 h flight. The available data referred to two fully instrumented prototypes of a specific LHD helicopter model. It is important to remember that only the regimes that the pilot was tasked to carry out are reported (everything in between is unknown a priori), while other flight spans are unlabeled. For our purposes, we only analyzed the labeled sections of each flight, considering each contiguous section of the flight with the same label as a single regime occurrence. The selected portion corresponds to more than 30% of the initial data, representing therefore a total duration of 10 h flight. The flights contain 19 different types of labeled flight regimes. The regimes were executed several times, based on the intrinsic characteristic of the flights: for example, takeoff, landings, and ground operations were executed fewer times compared to forward level flight — which can be executed more than once within the same flight. Details on the occurrence numerosity is available in Table 2.

For a better comprehension of the regimes labels, the meaning of the employed acronyms is reported in Table 3. The available data were divided in training data, employed in the training phase of the regime recognition algorithms, and test data exploited to evaluate performances of the proposed algorithm. The dataset division was balanced devoting, for each regime, 75% of the instances for training and 25% for testing.

Table 2
Regimes Representation. This table reports for each provided regime, the number of times that it was performed in all the flights considered in the original, in the cleaned and in the reduced dataset.

Regime	Original dataset	Cleaned dataset	Reduced dataset
Acceleration From OGE to 50K/VMAX Fwd	54	52	39
Autorotation 1.0 VNE Steady	82	82	39
Bank Turn Right 45 DEG 0.6 VNE	86	81	39
Climb VY Top	87	83	39
Deceleration From 0.9 VNE to VY	81	74	39
Deceleration From 50K/VMAX Fwd to OGE	54	53	39
Forward Flight 0.4 VNE	99	90	39
Forward Flight 0.6 VNE	105	102	39
Forward Flight 1.0 VNE	113	108	39
Forward Flight 1.1 VNE	106	95	39
Normal Landing (from VY to Ground)	41	39	39
Normal Mpog	47	42	39
Sideways Flt Port 50K/VMAX Rec IGE/OGE	61	55	39
Sideways Flt Port 50K/VMAX Stdy IGE/OGE	58	55	39
Spot Turn 30 Deg/s RH (360) Ent-St-Rec	63	59	39
Transition From Hige to VY Top	49	47	39
Transition From VY Level to VY Climb MCP	73	69	39
Transition From VY Level to VY Descent 1500FPM	63	63	39
Vertical Take Off to H. IGE	47	44	39

Table 3
Acronyms Explanation. This table shows the meaning of the acronyms employed in the regime labels.

Acronym	Explanation
VNE	Velocity Never-exceed
VY	Best Rate Of Climb Speed
MPOG	Minimum Pitch On Ground
FLT	Flight
FWD	Forward
MCP	Maximum Continuous Power
(H)IGE	(Hover) In Ground Effect
(H)OGE	(Hover) Out Ground Effect
Ent-St-Rec	Entry, Steady and Recovery
RH	Right Hand
LH	Left Hand
FPM	Feet Per Minutes

2.2. Explorative analysis

The first mandatory phase that precedes any data manipulation is the Exploratory Data Analysis (EDA). It allows the data analyst to explore the dataset provided, investigating individual signals' properties and their associations, leveraging univariate and multivariate techniques, respectively (Jebb et al., 2017).

The analyzed dataset contains instances related to different flight regimes. For each regime, more than one instance was recorded. The start and end instants for each instance are defined by the flight test engineer, who on-board starts and stops signal acquisition so that each corresponds to a regime performed by the pilot.

Therefore, a first analysis may focus on the similarity that characterizes different instances of the same flight regime. Indeed, time-series referred to the same regime may differ consistently from one repetition to another. The causes are various: undoubtedly, the environmental conditions and the piloting style have a crucial impact. Fig. 2 shows some of the signals referred to pilot commands, engine behavior, and flight dynamics angles, acquired during four different instances of the same regime, *i.e.*, transition VY Level to VY Climb. Very different trends characterize the instances. Similar differences can be found by monitoring multiple instances of all flight regimes, demonstrating that the same regime may not be univocally defined.

This variability represents one of the leading causes affecting the performances of the supervised classification methods (Bradley, 2015; Musso and Rogers, 2020). An unsupervised classification method, on the contrary, is less affected since it is not bound to the regime label in grouping the signals.

2.3. Correlation analysis

We also investigated the overall correlation between the signals provided in the EDA phase, considering all regimes. This information can be further employed during the features selection phase since highly correlated signals provide redundant information, which allows keeping one of them, discarding the others. In statistics, several methods are presented to calculate the correlation between two variables. Among the parametric methods, the best known is the Pearson's coefficient, while among the non-parametric ones, the most commonly used are Spearman's and Kendall's coefficients. Since we cannot make any a priori assumptions about the distribution of the given signals, we decided to employ a non-parametric method. In this work, we chose to use the correlation based on the Kendall τ coefficient, which, compared to Spearman, is more robust to the presence of outliers (Chok, 2010). This correlation coefficient is computed as:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \tag{1}$$

where n is the total number of samples pair, and n_c and n_d are the number of concordant and discordant sample pairs, respectively.

The result obtained is shown in Fig. 3. The heatmap was used to help the reader understand the correlation degree. The intensity of the color is proportional to the degree of correlation; an intense red corresponds to a solid direct correlation, blue inverse. It is possible to notice the strong correlation between the signals measured on the two engines (>0.90), which means that they mostly work synchronously. There is also a high correlation between the engine variables and the main rotor speed (>0.85), which are also physically correlated. The other correlations highlighted are less than 0.75; therefore were not considered significant enough to justify a feature discharge.

2.4. Data cleaning and pre-processing

As already mentioned, the process of labeling the flight regimes is entirely manual and, as such, prone to human error. It is not uncommon for a right-hand bankturn to be labeled as a left-hand one or for a takeoff to be mixed up for a landing. For this reason, a thorough visual inspection of the data was carried out to identify the mislabeled occurrences, which might result in lower performance metrics.

Furthermore, Fig. 4 reports the boxplots of the each instance duration for each regime. Observing the reported distributions, several outliers can be noticed (which might be due, for example, to an inaccurate regimes termination flagging), that need to be removed. As a result of this process, 76 regime occurrences were removed. The

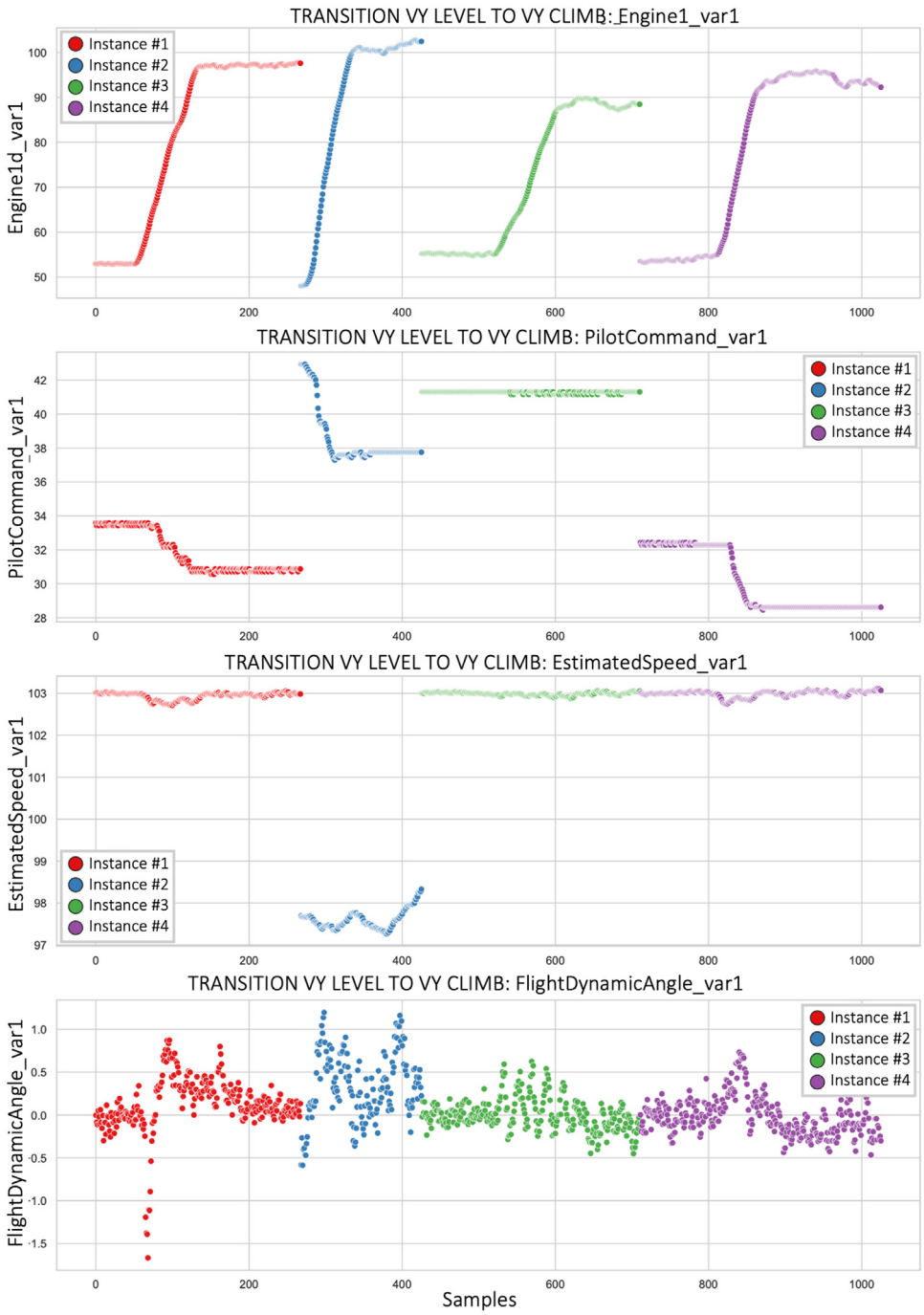


Fig. 2. Transition VY Level to VT Climb: Signals Trend. This figure illustrate the time trends for some of the provided signals during four different instances of transition between VY Level to VY climb.

final result of this data cleaning is visible in Table 2. However, as mentioned in Section 2.1, the dataset is intrinsically unbalanced. We performed a random undersampling by class to prevent the classifier performances from being affected by the dataset composition, keeping only 39 occurrences for each one (see Table 2), i.e., those available for the least represented regime in the dataset. This step is necessary to prevent the learned model from being affected by the regime's representativeness difference. Indeed, it is well known in the literature that class unbalance may lead to biased classifiers (Abd Elrahman and Abraham, 2013). Finally, all the signals were filtered with a 1st order low-pass filter to avoid sudden signal variations and spikes caused by acquisition errors.

2.5. Features engineering and selection

As the set of acquired signals was already informative, it was unnecessary to carry out extensive feature extraction or engineering. The only necessary adjustment was transforming the measured ground speeds (which refer to the magnetic north) into more meaningful speed and acceleration estimates. Then, an ad hoc features selection was carried out. Especially when dealing with unsupervised learning approaches, this phase is crucial. Indeed, the algorithm cannot automatically evaluate each feature's importance autonomously, contrary to the case of supervised learning. Supervised approaches can weigh features according to their explanatory power by using the target labels. However, in unsupervised learning, we do not leverage the label information

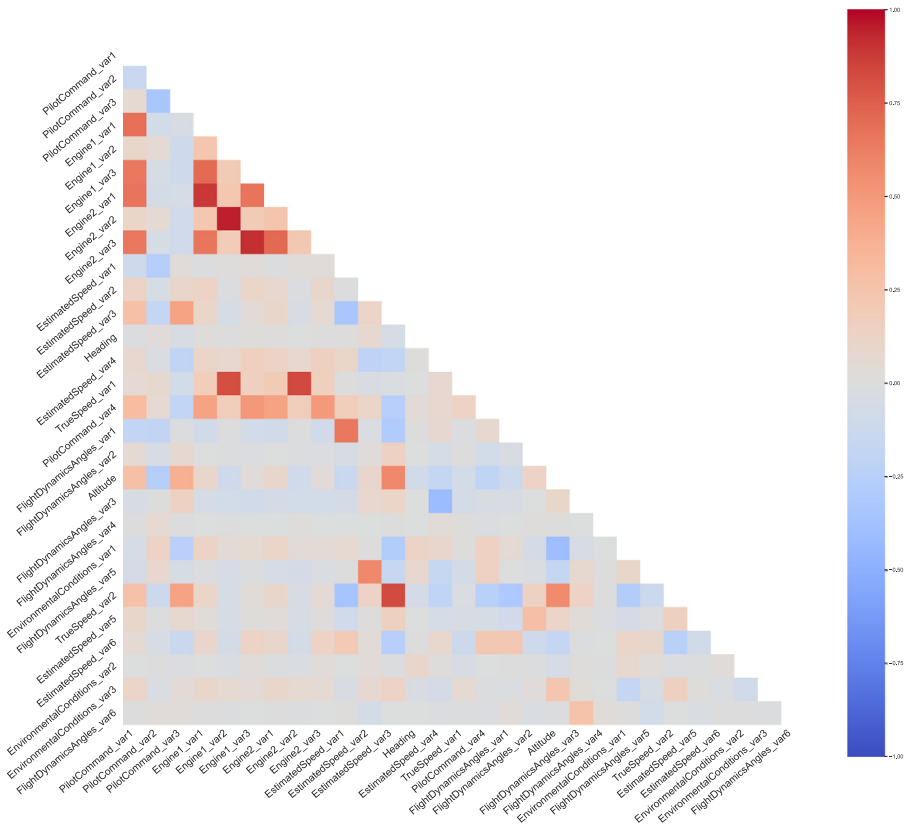


Fig. 3. Provided Signals Correlation Matrix. This figure provides the correlation estimated according to Kendall method between each provided signal pair.

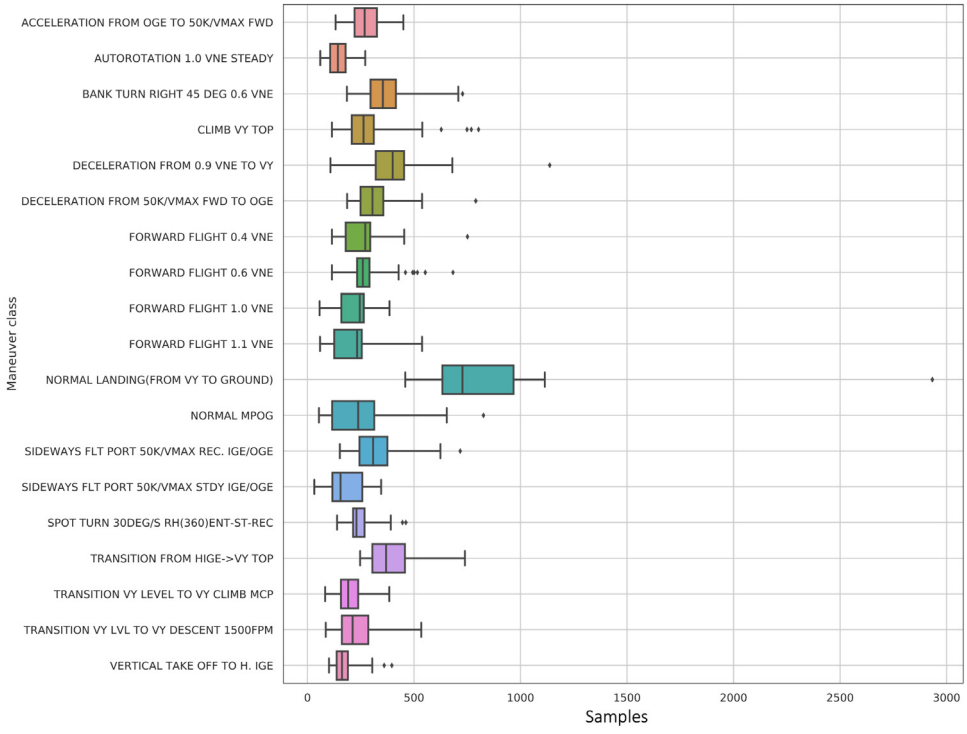


Fig. 4. Boxplot of maneuver durations, highlighting the outliers.

during the learning process. The inclusion of irrelevant features in the unsupervised learning procedure is problematic since the insignificant variables are as equally weighted as the important ones, leading to misleading distance computations and erroneous results. First, redundant information is discarded. Indeed, as shown in Fig. 3 many of the signals included in the original dataset are strongly correlated. For example, the set of features measured on the two engines is redundant, allowing for discharging engine 2 variables without any information loss. Then

Table 4
Dataset Features. This table shows the features categories provided and the ones considered after features engineering and selection phases.

Feature category	# Variables	Original dataset	Cleaned dataset
Pilot Commands	4	✓	✗
Engine 1	3	✓	✓
Engine 2	3	✓	✗
Measured Ground Speed	2	✓	✗
Flight Dynamics Angles	6	✓	✓
Heading	1	✓	✗
True speed	2	✓	✓
Estimated Acceleration and Speed	6	✗	✓
Environmental Conditions	3	✓	✗
Altitude	1	✓	✓
Problem dimensionality		25	18

a further selection was carried out, with the supervision of domain experts. This way, we discarded features such as the environmental variables, which may be misleading for flight recognition purposes. At the end of this phase, the problem dimensionality was reduced from 25 to 18. The final result of the feature engineering and selection is reported in Table 4

3. Functional data representation

The problem under study can be included in the mathematical framework of Functional Data Analysis (FDA), firstly defined by Ramsay in 1982 (Ramsay, 1982). Thanks to the data collection technology, the available data are densely sampled over time, so it is reasonable to assume that the continuous underlying processes can be adequately estimated from the discrete observations. This kind of data, defined in a continuous domain such as a time interval, can be efficiently analyzed with FDA methods.

In our scenario, we assume that each instance, namely each regime in the dataset, can be represented by multiple continuous signals evolving in a time interval $T = [t_{min}, t_{max}]$. Data are considered regular enough to be embedded in the Hilbert space of L^2 functions. Therefore we assume that each instance x is a realization of a functional random variable from a probability space to L^2 and so can be described as:

$$x(t) : T \to \mathbb{R}^d \tag{2}$$

where d is the total number of signals considered, namely the multiple attributes of each instance in the dataset.

3.1. Smoothing

Since signals are recorded at regular time samples in the interval T, the resulting available observations are actually discrete. Each instance $x_i(t)\ \forall i = 1, \dots, N$ consists indeed of a set of n_i pairs $(t_{i,j}, \boldsymbol{y}_{i,j})$ with $j = 1, \dots, n_i,\ t_{i,j} \in T,\ \boldsymbol{y}_{i,j} \in \mathbb{R}^d$.

The continuous process and the discrete observations can be related through an error term $\epsilon_{i,j}$ by the following expression:

$$\boldsymbol{y}_{i,j} = x_i(t_j) + \epsilon_{i,j} \tag{3}$$

A smoothing procedure is thus required to estimate the underlying continuous functional process $x_i(t)$ from the discrete available observations. For this purpose, different smoothing methods could be applied. Such methods can be divided into two macro-categories: kernel and basis smoothers. The firsts compute the approximation $x_i(t)$ at each point t by considering the influence of surrounding points, while basis smoothers calculate the continuous data representation by expanding bases with truncation to a defined order.

In our application, the K-Nearest Neighbors estimator was adopted. It is a non-parametric linear kernel smoothing method that estimates the function value at each point in the time interval T as the average of the m nearest observations.

Given the set of n_i observations for the instance i, the value of the function $\hat{x}_i(t_0)$ estimating the process x in the point $t_0 \in T$ is given by:

$$\hat{x}_i(t_0) = \frac{\sum_{j=1}^{n_i} K_m(t_0, t_j) x_i(t_j)}{\sum_{j=1}^{n_i} K_m(t_0, t_j)} \tag{4}$$

where $K_m(t_0, t_j)$ is the kernel function assigning a weight to t_j based on its distance from t_0. The parameter m defines the width of the neighborhood. The kernel function for the nearest neighbors smoothing method is given by:

$$K_m(t_0, t_j) = \begin{cases} \frac{1}{m} & \text{if } \frac{\|t_j - t_0\|}{\|t_{[m]} - t_0\|} \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $t_{[m]}$ is the *m-th* time sample closest to t_0.

Besides the relative simplicity of this smoothing method, it showed to work well in our application and thus preferred over other kernel or basis smoothers.

For a more immediate and intuitive understanding, the smoothing procedure is presented in Fig. 5 for a fictitious set of data in one dimension.

3.2. Registration

The procedure adopted to collect the available data is strongly human dependent; therefore, shifts between the first observed sample and the regime's start are frequent. Such shifts, variable between different instances, do not provide any information about the regime itself, but instead they could negatively affect subsequent data analysis.

Therefore, the registration procedure is necessary to align the smoothed data and thus make the analysis robust to shift in phase.

In our application, a shift in the domain, *i.e.,* a shift registration, was performed to align peaks in the recorded signals. The procedure was applied, after smoothing, for the 19 labeled regimes separately. Since registration requires data to have equal length, all the available signals were padded to the length of the regime's instance with longer duration and then registration was performed.

The registered function for the instance i is indeed given by:

$$x_i^*(t) = x_i(t + \delta_i) \tag{6}$$

where the time shift δ_i is the one minimizing a least squared expression, *i.e.,* the Registered Sum of Squared Errors:

$$REGSSE = \sum_{i=1}^{N} \int_T [x_i(t + \delta_i) - \hat{\mu}(t)]^2 ds \tag{7}$$

Here $\hat{\mu}(t)$ is the mean of the registered functions.

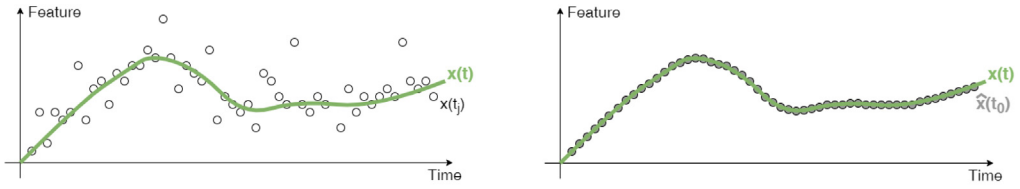
The minimization problem is solved iteratively using the Newton–Raphson algorithm (Ramsay and Silverman, 2008).

For a better understanding, the effect of registration on a fictitious set of data in one dimension is presented in Fig. 6.

4. Flight regimes recognition

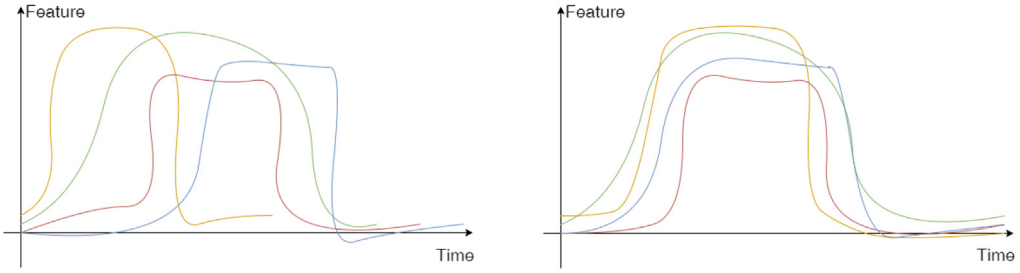
As already mentioned, flight regime recognition is fundamental to monitoring a helicopter's actual usage and thus enabling a condition-based maintenance scheduling considered more efficient than a time-based one in terms of costs and safety. Here, for this purpose, an unsupervised approach is proposed. Unsupervised clustering, indeed, released from regime's labels defined during the certification phase, allows to identify, in addition to standard regimes, mixed regimes and transitions which are frequent in actual flight conditions although not included in certification flights.

A soft classification technique was adopted to cluster the available functional data in this perspective. Specifically, the fuzzy C-Means clustering algorithm was chosen for our application, ad-hoc reformulated to handle functional data (Tokushige et al., 2007). Fuzzy C-Means



(a) Raw data and estimated continuous function. (b) Data after smoothing

Fig. 5. Smoothing. The figures report an example of Smoothing procedure on one dimensional fictitious data.



(a) Data before registration (b) Registered data after padding

Fig. 6. Registration. The figure report an example of Registration procedure on one dimensional fictitious data.

allows each instance to belong to more than one cluster by assigning a membership grade for each of the available classes. This membership grade denotes the confidence level with which the algorithm associates the instance to the specific cluster. In other words, a membership grade close to 1 denotes a high similarity between that instance and the given cluster, while a membership grade close to 0 implies little similarity. A suitable distance metric was employed to measure such similarity in a functional space.

Therefore, the soft clustering algorithm adopted for the current application, preventing a hard data partition, accounts for possible similarities between different regimes, leading to higher overall performance compared with other clustering procedures. Indeed, as reported in Section 1, one important aspect related to certification and real flights is that often the latter re-characterized by new regimes, not represented in certification flights, but typically obtained by the mixture of known regimes performed in certification flights. The adoption a soft classification technique has the main advantage to offer the possibility to identify in real flights such new regimes determined by the mixture of two (or potentially more) known regimes, which will be those associated with higher predicted probabilities.

4.1. Fuzzy C-means algorithm

Given the set of functional observations $X = x_1, x_2, \dots, x_N$ soft cluster analysis consists in identifying a fuzzy partition of the data, namely in assigning to each datum $i = 1, \dots, N$ the membership grades $u_{i,j}$ to class $j = 1, \dots, C$. The grades $u_{i,j}$ must satisfy the following conditions:

$$\sum_{j=1}^{C} u_{ij} = 1 \ \forall i = 1, \dots, N, \quad \sum_{i=1}^{N} u_{ij} > 0 \ \forall j = 1, \dots, C \tag{8}$$

The algorithm finds the optimal fuzzy partition of the data by minimizing the following cost function:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - v_j \right\|_{*}^{2} \tag{9}$$

where v_j is the center of cluster j, m is a weighting exponent determining the degree of fuzziness and $\|\cdot\|_*$ denotes any norm well defined in the embedding functional space.

With regard to the fuzziness parameter m it can be proven that with $m = 1$ the algorithm leads to hard clustering where v_j are the geometric centroids of each class, while as m tends to infinity the membership grades will tend to be equally distributed over the C clusters providing no information about any possible data grouping.

In our application, after fine-tuning of this parameter, the degree of fuzziness was set equal to 2. This choice, indeed, appears to offer a good compromise between hard clustering and the non informative equal probability distribution over the regimes.

The norm adopted in this case study was the L^1 norm given by

$$\left\| x_i - v_j \right\|_{L^1}^{2} = \int_T |x_i(t) - v_j(t)| dt \tag{10}$$

which showed a better performance with respect to the L^2 norm, both well defined in the embedding L^2 space.

The algorithm solves the minimization problem iteratively and, at each iteration, the cluster centroids and the membership degrees are update as follows:

$$u_{i,j} = \left[\sum_{k=1}^{C} \left(\frac{\left\| x_i - v_j \right\|_{L^1}^{2}}{\left\| x_i - v_k \right\|_{L^1}^{2}} \right)^{\frac{2}{m-1}} \right]^{-1} \tag{11}$$

$$v_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \tag{12}$$

The iterative process stops when the fuzzy partition do not change between consecutive iterations, namely when, given a threshold ϵ, at iteration step $k+1$ we have:

$$\max_j \left\| v_j^{(k+1)} - v_j^{(k)} \right\|_{L^1} < \epsilon \tag{13}$$

The described algorithm was applied on the available data, after smoothing and registration procedure.

Since the clustering algorithm works on data with fixed length, all the regimes' instances were sliced in windows of equal size set to 350, after fine tuning. Fig. 8 shows the results of functional clustering for a fictitious set of smoothed and aligned data, highlighting that the prediction is assigned to the entire maneuver instance, accounting for temporal relationship between samples.

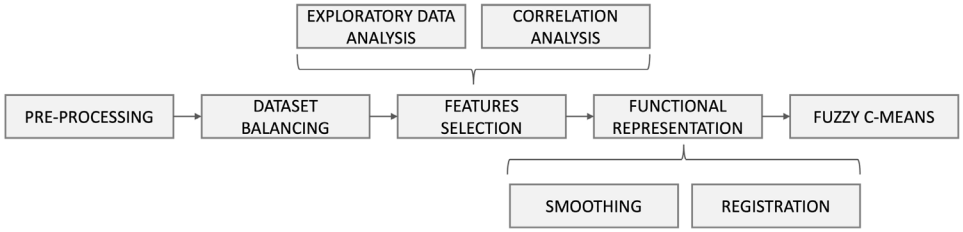
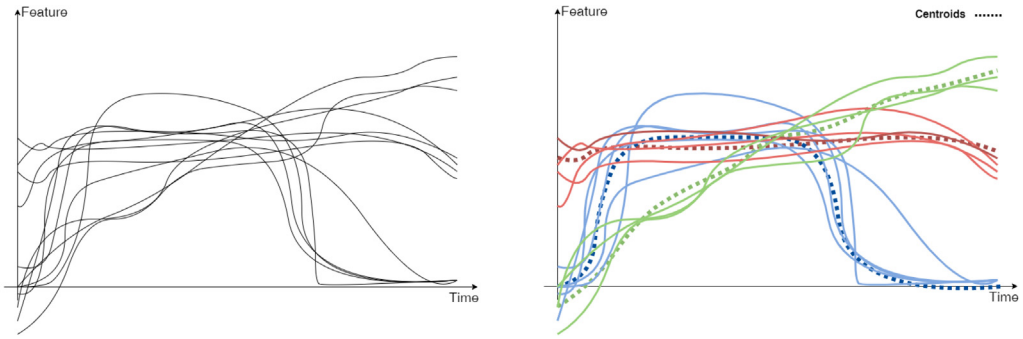


Fig. 7. Methodology Pipeline Architecture. The reported figure shows the phases composing the proposed regimes recognition approach.



(a) Functional data from different classes

(b) Grouped data with number of classes C=3

Fig. 8. Functional Clustering. The figures report an example of Functional Clustering on one dimensional fictitious data.

To help the reader to better understand the proposed methodology, its architecture is depicted in Fig. 7.

The results obtained applying it to the provided data are presented in Section 5.

4.2. Cluster-regime mapping

Fuzzy C-Means returns the probability of belonging to each cluster for each instance. However, to carry out the subsequent lifetime monitoring tasks, *i.e.*, estimate the usage status, it is necessary to associate each regime with a cluster. To this end, downstream of the training process, we mapped each cluster identified by Fuzzy C-Means to the corresponding regime. Accordingly, each instance was considered as belonging to the cluster predicted with maximum likelihood. At this point, we analyzed the frequency of true regime labels in each cluster. This was possible as, despite never being leveraged during the training process, the actual labels were available in the provided dataset. Also, we noticed that one true regime label was predominant in each cluster, and no regime label was the most represented in two or more clusters. This allowed us to define a 1:1 correspondence between identified clusters and true regimes, which was crucial for handling the usage monitoring task associated with new instances.

As reported in Section 1, leveraging an unsupervised method allows our flight regime recognition approach to be reliable even when considering a real usage spectrum, including transitions and mixed maneuvers, which combine different categories. It follows that a unsupervised analysis could also assist the revision process of the regimes composing the true helicopter usage spectrum. In addition, leveraging an unsupervised approach is also valuable in the military sector, where it can provide useful insights about combat maneuvers, regardless the low regimes repeatability and the unclear segmentation criteria respect to the civil scenario.

5. Experimental results and discussion

This Section first defines the metrics used to evaluate clustering performance on test data, pre-processed and selected according to the methods presented in Section 2. According to them, obtained results are then reported and discussed.

5.1. Evaluation metrics

The total Sum of Squared Error within-cluster (SSE) was employed to infer the optimal clusters' number. This metric measures the dispersion for each of the N_k points belonging to a cluster x_{ik} with respect to its centroid v_k, and is defined as:

$$SSE = \sum_{k=1}^{N_C} \sum_{i=1}^{N_k} (x_{ik} - v_k)^2 \tag{14}$$

where N_C is the number of classes, *i.e.,* of the considered regimes.

The lower the SSE, the better the clustering produced, characterized by cohesive groups. However, as the number of clusters increases, fewer instances will fall within each group, resulting in a total SSE decrease, not resulting in a better clustering. The *Elbow method* provides precise instructions to identify the point that optimizes the trade-off between SSE and the number of clusters, preventing overfitting (Thorndike, 1953). The clustering was repeated according to its specifications, varying the number of clusters and calculating SSE each time. The results obtained were then plotted in a SSE vs. Number of Clusters plane. The optimal clusters number corresponds to the bend's location, i.e., the knee of the obtained curve. Adding clusters would not lead to a significant decrease in total SSE from this value.

Considering the clustering output and each instance regime label, neglected in the clustering process, additional metrics can be computed. Please notice that Fuzzy C-Means assigns a probability of belonging to each cluster to each instance. Therefore, each instance was attributed to the high probability cluster to evaluate the produced results. The most commonly used is the Rand Index (RI), which corresponds to an accuracy measured between pairs of clustering instances (Rand, 1971). Given the clustering outcome, RI is calculated as the ratio of the number of instances pairs having the same label and cluster, denoted by the letter a, and of those having different labels and clusters, denoted by

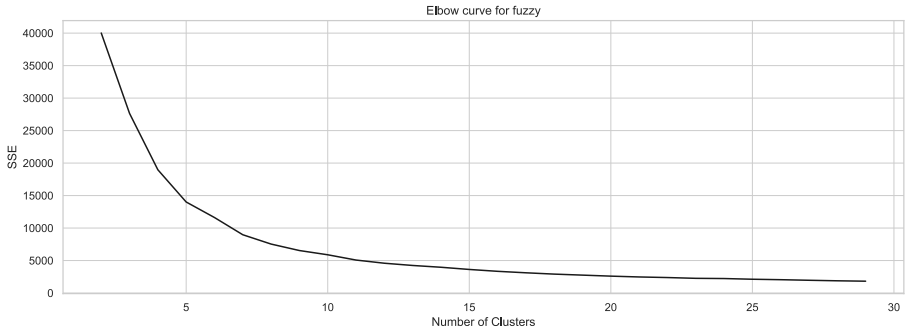


Fig. 9. Elbow Plot. The figure illustrates the total SSE within clusters' trend varying the clusters' number.

the letter b, over the total number of possible pairs, i.e., the binomial coefficient $\binom{N}{2}$. N indicates the total instances number.

$$RI(C) = \frac{a+b}{\binom{N}{2}} \tag{15}$$

RI value ranges from 0, *i.e.*, no correspondence between Fuzzy C-Means clustering and the one obtained by grouping the instances according to their label, to 1, *i.e.*, the two clusterings coincide.

The Adjusted Rand Index (ARI) is a for chance adjustment of RI, and is defined as:

$$ARI(C) = \frac{RI(C) - \overline{RI(C)}}{max(RI(C)) - \overline{RI(C)}} \tag{16}$$

where $\overline{RI(C)}$ is the expected RI for each cluster.

Another useful metric is the Mutual Information (MI) (Shannon, 1948).

This metric considers the consistency between the clustering output C and the true labels L. To this extent, the label attributed by Fuzzy C-Means to each cluster, indicated as c_k, is compared to true value l_j for each sample of belonging. In detail, considering a dataset composed of N instances, the MI is defined as:

$$MI(C,L) = \sum_{k=1}^{N_C}\sum_{j=1}^{N_k} \frac{|c_k \cap l_j|}{N} log\left(\frac{N|c_k \cap l_j|}{|c_k||l_j|}\right) \tag{17}$$

where N_C is the number of clusters, *i.e.*, of the regimes, and N_k is the number of samples attributed to each one.

Also in this case, it is possible to adjust the metric by chance, producing the Adjusted Mutual Index (AMI):

$$AMI(C,Y) = \frac{MI(C,Y) - \overline{MI(C,Y)}}{max(MI(C,Y)) - \overline{MI(C,Y)}} \tag{18}$$

where $\overline{MI(C)}$ is the expected MI for each cluster.

Both RI and MI, as their adjusted values ARI and AMI, are independent of the mapping between labels and clusters. However, performing this mapping is essential to produce a confusion matrix, and thus to compute all the metrics commonly used to evaluate a supervised algorithm's performance (Hay, 1988). Unlike ARI and AMI, these metrics vary as the mapping established between labels and clusters varies. In this work, we associated each cluster to the label most represented by its belonging instances. Once the mapping is performed, the confusion matrix quantities can be estimated, namely True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Among the metrics that can be computed considering the produced confusion matrix, the most commonly used is undoubtedly the accuracy, *i.e.*, the ratio of correct predictions over the total number of produced ones.

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{19}$$

Besides, precision and recall are also necessary. In detail, the precision quantifies the number of instances of a regime correctly attributed to the corresponding cluster over the total number of instances attributed to that cluster.

$$precision = \frac{TP}{TP+FP} \tag{20}$$

On the other hand, the recall measures the number of instances of a regime correctly attributed to the corresponding cluster over of the total number of instances belonging to that regime.

$$recall = \frac{TP}{TP+FN} \tag{21}$$

When the classes considered are not equally balanced, as in multi-class classification problems, the F1 Score is a more robust metric than accuracy. Indeed, F1 Score allows also for considering the number of instances not correctly classified, since it is defined as the harmonic mean between accuracy and recall.

$$F1\ Score = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{22}$$

Also, macro and micro-averages of precision, recall and F1 Score were considered. Macro-average is estimated computing the metric independently for each regime, and taking the overall average; Micro-average is obtained aggregating the contributions of all regimes to compute the average metric.

Last, to measure the similarity between clusters, a novel metric is proposed. Second cluster label, y_2, is defined as the vector containing for each instance the ID of the second most probable cluster, according to Fuzzy C-Means predictions. The similarity between clusters' pair, A and B, is then defined as the ratio of A's instances whose second cluster label is B, indicated as $|x_A(y_2 == B)|$ over the total number of A's instances, indicated as $|A|$.

$$\mathrm{S(A,B)} = \frac{|x_A(y_2 == B)|}{|A|} \tag{23}$$

This metric ranges from 0, *i.e.*, clusters are very dissimilar, to 1, *i.e.*, clusters are very similar. Leveraging the mapping between regimes and clusters, it is possible to interpret the similarity between clusters as a similarity between regimes.

5.2. Experimental results

As previously mentioned, the total SSE within clusters was employed to estimate the optimal clusters' number. In detail, SSE was evaluated by varying the number of clusters between 1 and 30. The results obtained are shown in Fig. 9.

The knee of the curve corresponds to the point that optimize the trade-off between the two considered variables. Accordingly, it is possible to notice that the maximum clusters' number that provides a significant SSE reduction ranges from 16 to 20. Passing from 2 to 3 clusters decreases the SSE by 30.92%, and this gain progressively reduces as the clusters' number increases. In the range between 16 and 20 clusters, adding one cluster reduce the SSE by 6.42%, on average. Passing from 20 to 21 clusters, instead, reduces the SSE by 3.65%, and the gain is dramatically reduced for further increases. Since the actual number of regimes collected in the considered dataset is 19, which also

Table 5
Clustering Evaluation. This table aims at evaluating the outcomes of Fuzzy C-Means clustering when mapping each cluster to a regime. In detail, precision, recall, and F1 Score are reported for each regime.

Regime	Precision	Recall	F1 Score
Acceleration From OGE to 50K/VMAX Fwd	0.93	0.95	0.94
Autorotation 1.0 VNE Steady	1.00	1.00	1.00
Bank Turn Right 45 DEG 0.6 VNE	1.00	1.00	1.00
Climb VY Top	0.95	1.00	0.97
Deceleration From 0.9 VNE to VY	1.00	1.00	1.00
Deceleration From 50K/VMAX Fwd to OGE	0.78	0.97	0.86
Forward Flight 0.4 VNE	0.90	0.97	0.94
Forward Flight 0.6 VNE	0.97	1.00	0.99
Forward Flight 1.0 VNE	0.79	0.95	0.86
Forward Flight 1.1 VNE	0.93	0.72	0.81
Normal Landing (from VY to Ground)	1.00	0.97	0.99
Normal Mpog	0.91	0.82	0.86
Sideways Flt Port 50K/VMAX Rec IGE/OGE	0.95	0.97	0.96
Sideways Flt Port 50K/VMAX Stdy IGE/OGE	1.00	1.00	1.00
Spot Turn 30 Deg/s RH (360) Ent-St-Rec	1.00	1.00	1.00
Transition From Hige to VY Top	0.91	1.00	0.95
Transition From VY Level to VY Climb MCP	1.00	0.90	0.95
Transition From VY Lebel to VY Descent 1500FPM	0.97	0.90	0.93
Vertical Take Off to H. IGE	1.00	0.77	0.87
Weighted Average	0.95	0.94	0.94

falls in the optimal elbow range, we set the clusters' number to this value. Considering the clustering produced by Fuzzy C-Means when the clusters' number is set to 19, the ARI and AMI metrics assess to 0.843 and 0.903, respectively.

Each cluster is then mapped to the regime most represented by its belonging instances. This allows for computing a confusion matrix for each regime and estimating precision, recall, and F1-Score.

The results obtained are reported in Table 5. For each performance metric, the weighted average is computed weighting each score by the number of samples actually belonging to the respective regime.

Besides, the heatmap obtained from the mapping produced is shown in Fig. 10. Each cell is colored according to the percentage of samples belonging to the corresponding row cluster whose label is the one of the corresponding column.

It can be noticed that most of the regimes correspond perfectly to one of the clusters produced by Fuzzy C-Means (each of the classes produces a single and distinct cluster). Considering the F1 Score, 14 of the 19 regimes map to a cluster with more than 90.0%, and 18 to more than 85.0%. A lower F1 Score is obtained for Forward Flight 1.1 VNE, with an F1 Score of 81.0%. Analyzing Fig. 10 helps understand the reasons behind this lower performance: a portion of the instances collected in this cluster belong to the Forward Flight 1.0 VNE class. The two classes have similar characteristics, and our algorithm has helped identify the regimes labeled as Forward Flight 1.0 VNE that have been flown at higher speeds. Considering precision and recall, 14 out of 19 regimes map to their respective clusters more than 90.0%. These results are consistent with the global estimate of the macro and micro averages, which for the precision are 0.95, while for recall and F1 Score are 0.94. It is important to note that the F1 Score, the precision, and the recall of the three transitions' regimes are greater than 90.0%. This shows that our method can distinguish and recognize them correctly.

Another interesting evaluation considers actual and predicted labels for the same flight regimes, obtaining the so called mission profile.

Indeed, as reported in Section 4 the predicted cluster can be associated to the corresponding regime label considering the mapping produced at the end of the training process. Fig. 11 shows the labeled part of a whole mission profile for one of the collected flights.

The result obtained shows that the proposed method is perfectly able to recognize the helicopter flight regime, estimated considering the estimated cluster-regime mapping.

Finally, we define a new metric to estimate the similarity between clusters and, consequently, between regimes. The produced results are reported in Fig. 12.

Fig. 10. Clusters-Regimes Heatmap. This Figure highlights for each cluster the percentage of samples belonging to each possible regime. Red squares correspond to a perfect cluster-regime correspondence.

In detail, it shows the similarity degree for each clusters pair according to the reported colorbar; indeed, the darker the square, the closer the clusters, i.e., the regimes.

The results obtained are consistent with the expert a priori knowledge. For example, it is shown that the regime most similar to Forward Flight 1.1 VNE is Forward Flight 1.0 VNE. Also, Sideways Flt Port 50K/VMAX Stdy IGE/OGE, and Spot Turn 30DEG/S RH(360)Ent-St-Rec are most similar to Sideways Flt Port 50K/VMAX Rec IGE/OGE.

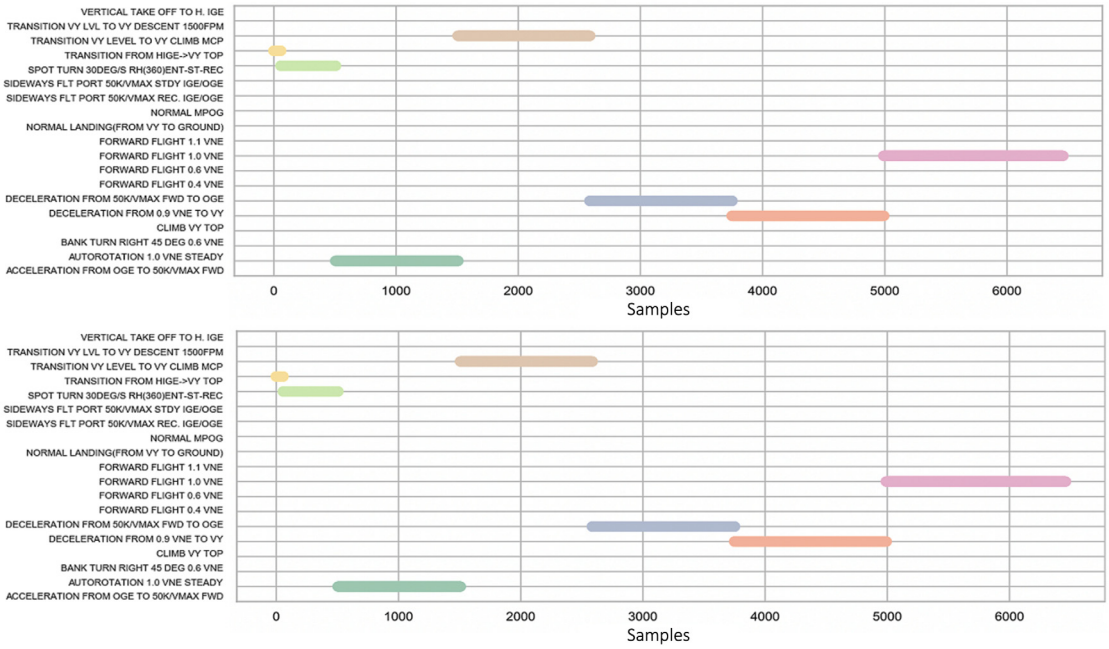


Fig. 11. Mission Profile. This figure reports for a given flight the actual recorded regimes (top), and the ones predicted by our method (bottom).

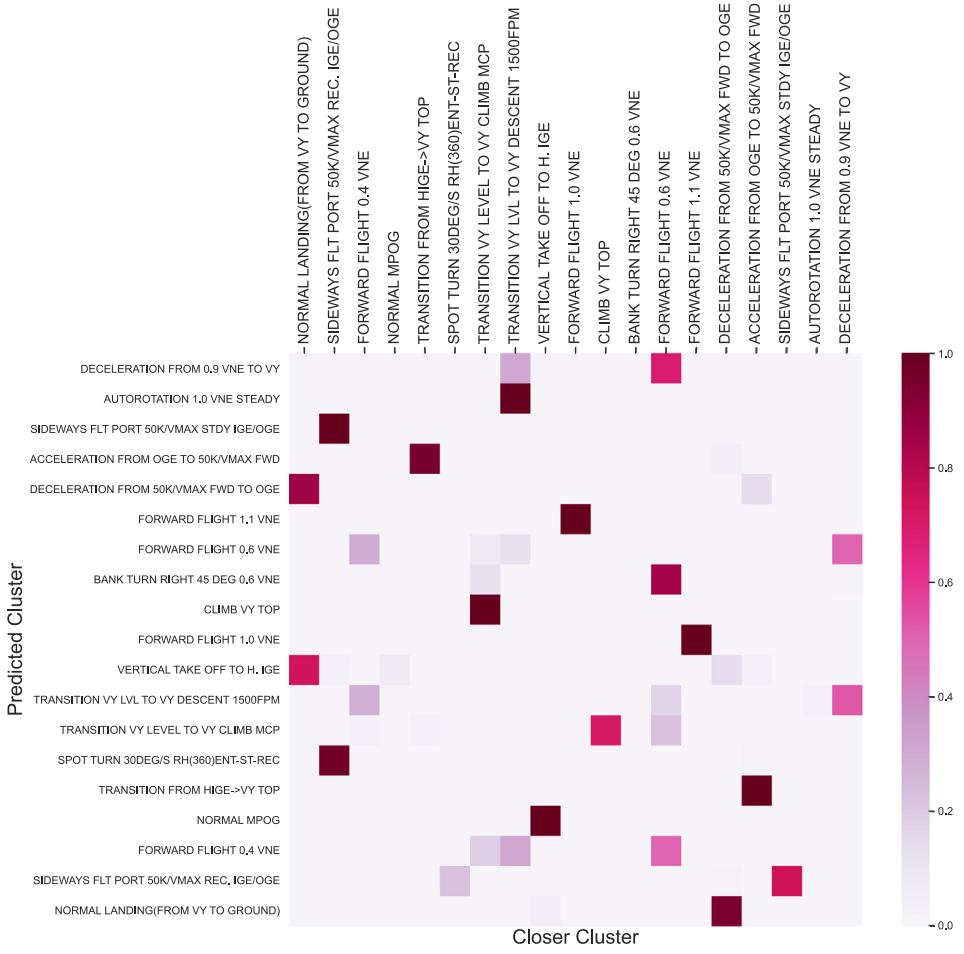


Fig. 12. Clusters Similarity. Considering the results of clusters-regimes matching, we map each cluster to a regime. In this figure, leveraging this information, the similarity between clusters is estimated.

Another significant similarity is between Climb VY Top and Transition From VY Level to VY Climb MCP. Besides, Acceleration from OGE to 50K/VMAX Fwd is most similar to Transition from Hige− >VY Top. The similarity between Normal Landing (from VY to Ground) and Deceleration From 50K/VMAX Fwd to OGE is also relevant since the two regimes are often causally related. The same occurs considering the causality that links Vertical Take Off to H.Ige to Normal Mpog. In addition to demonstrating that the clustering produced is consistent with actual helicopter operating conditions, the results obtained pave the way to the possibility of identifying new flight regimes, which can, however, be related to the known ones thanks to the introduced similarity measure.

6. Concluding remarks

This paper proposes an approach for flight regime recognition that leverages the FDA for instances' representation and Fuzzy C-Means as the clustering algorithm. To the best of the authors' knowledge, this is the first time these two techniques have been applied for regime recognition purposes. The FDA representation allows for considering the temporal dependence between the samples in the recognition process, improving the clustering algorithm's performance. Besides, employing an unsupervised algorithm prevents the identification process from affecting the provided flight regime labels. Indeed, the datasets available for the training of regime recognition algorithms are often collected during certification test flights, where the regimes are performed according to standard procedure, unlike during helicopter's actual usage. Supervised learning approaches force the algorithm only to identify those regimes collected during test flights, producing inconsistent results when the behavior is different from such classes. Instead, our approach is unconstrained by such limitations and can potentially deal with any usage. The efficacy of our approach was validated on actual flights, obtaining an average F1 score of 92.0% over 19 flight regimes.

Also, we proposed an ad hoc similarity metric between regimes in this paper. This metric proves to be effective in considering similar regimes comparable from a physical point of view or related by a causal relationship. This result is significant in defining new regimes, more consistent with the actual helicopter usage. According to this perspective, future developments foresee applying this method on data collected during actual civil and military operations to analyze their positioning concerning the clusters identified here. The proposed method would recognize new clusters referred to existing regimes and characterize them based on their similarity with the standard ones already defined.

Therefore, the results obtained pave the way for more effective usage monitoring systems. Collecting precise flight reports would allow HUMS systems to be more accurate in their diagnostics and increase their prognostic capabilities. This last achievement is fundamental to move from time-based to condition-based maintenance, which would lead to a significant operating cost reduction and safety gain.

CRediT authorship contribution statement

Jessica Leoni: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Francesco Zinnari:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Eugenia Villa:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Mara Tanelli:** Conceptualization, Methodology, Resources, Writing – review & editing, Visualization, Supervision, Project administration. **Andrea Baldi:** Conceptualization, Resources, Writing– review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mara Tanelli reports financial support was provided by Leonardo Helicopters Division Cascina Costa. Jessica Leoni, Francesco Zinnari, Eugenia Villa, Mara Tanelli, Andrea Baldi has patent METODO E SISTEMA PER RILEVARE REGIMI DI VOLO DI UN AEROMOBILE, SULLA BASE DI MISURE ACQUISITE DURANTE UN VOLO DELL'AEROMOBILE pending to Leonardo Helicopter Division Cascina Costa.

References

Abd Elrahman, S.M., Abraham, A., 2013. A review of class imbalance problem. J. Netw. Innov. Comput. 1 (2013), 332–340.

Berry, J., et al., 2006. Automatic regime recognition using neural networks. In: Proceedings of American Helicopter Society 60th Annual Forum.

Bradley, N.C., 2015. Pilot Variability Study for Federal Aviation Administration Health and Usage Monitoring Mock Certification. Tech. Rep., Army Research Lab Aberdeen Proving Ground MD Vehicle Technology Directorate.

Branch, A.A.I., 2009. Report on the accident to Aerospatiale (Eurocopter) AS332 L2 super puma, registration G-REDL 11 nm NE of peterhead, Scotland on 1 April 2009. Air Accidents Investigation Branch.

Chok, N.S., 2010. Pearson's Versus Spearman's and Kendall's Correlation Coefficients for Continuous Data (Ph.D. thesis). University of Pittsburgh.

Gene, B., Sarkar, S., Miller, C., 2007. Maneuver regime recognition development and verification for H-60 structural monitoring. In: Proceedings of Annual Forum Proceedings-American Helicopter Society.

HanYang, F., Hongming, F., RuiYuan, G., 2020. Research on air target maneuver recognition based on LSTM network. In: 2020 International Workshop on Electronic Communication and Artificial Intelligence. IWECAI, pp. 6–10. http://dx.doi.org/10.1109/IWECAI50956.2020.00009.

Hay, A., 1988. The derivation of global estimates from a confusion matrix. Int. J. Remote Sens. 9 (8), 1395–1398.

He, D., Wu, S., Bechhoefer, E., 2010. A regime recognition algorithm for helicopter usage monitoring. In: Aerospace Technologies Advancements. InTechOpen Croatia, pp. 391–404.

Jebb, A.T., Parrigon, S., Woo, S.E., 2017. Exploratory data analysis as a foundation of inductive research. Hum. Resour. Manag. Rev. 27 (2), 265–276.

Le, D., Cuevas, E., 2007. United States federal aviation administration health and usage monitoring system R&D strategic plan and initiatives. In: Proceedings of the 5th DSTO International Conference on Health and Usage Monitoring.

Lombardo, D., 1998. Helicopter flight condition recognition: A minimalist approach. In: Australian Joint Conference on Artificial Intelligence. Springer, pp. 203–214.

Musso, D., Rogers, J., 2020. Interacting multiple model estimation for helicopter regime recognition. J. Aircr. 57 (6), 1134–1147.

Ramsay, J., 1982. When the data are functions. Psychometrika 47 (4), 379–396.

Ramsay, J.O., Silverman, B., 2008. Funtional data analysis. Internet adresi: http.

Rand, W.M., 1971. Objective criteria for the evaluation of clustering methods. J. Amer. Statist. Assoc. 66 (336), 846–850.

Romero, R., 1996. Feasibility Study of a Rotorcraft Health and Usage Monitoring System (HUMS): Results of Operator's Evaluation. US Department of Transportation, Federal Aviation Administration.

Şenipek, M., Kalkan, U., 2019. Learning-based clustering for flight condition recognition. In: Proceedings of the 45th European Rotorcraft Forum.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (3), 379–423.

Sheridan, K., Puranik, T.G., Mangortey, E., Pinon-Fischer, O.J., Kirby, M., Mavris, D.N., 2020. An application of dbscan clustering for flight anomaly detection during the approach phase. In: AIAA Scitech 2020 Forum. p. 1851.

Teal, R., Evernham, J., Larchuk, T., Miller, D., Marquith, D., White, F., Deibler, D., 1997. Regime recognition for MH-47e structural usage monitoring. In: Annual Forum Proceedings-American Helicopter Society, Vol. 53. pp. 1267–1284.

Thomson, D., Bradley, R., 1990. Modelling and classification of helicopter combat manoeuvres. In: 17th International Council of the Aeronautical Sciences, Stockholm, Sweden.

Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18 (4), 267–276.

Tokushige, S., Yadohisa, H., Inada, K., 2007. Crisp and fuzzy k-means clustering algorithms for multivariate functional data. Comput. Statist. 22 (1), 1–16.

Ullah, S., Finch, C.F., 2013. Applications of functional data analysis: A systematic review. BMC Med. Res. Methodol. 13 (1), 1–12.

Warner, J., Rogers, J., 2019. Novel diagnostic metrics for regime recognition verification and validation. In: 75th Annual Forum of the Vertical Flight Society.

Wu, J., Sun, C., Zhang, C., Chen, X., Yan, R., 2022. Deep clustering variational network for helicopter regime recognition in HUMS. Aerosp. Sci. Technol. 124, 107553. http://dx.doi.org/10.1016/j.ast.2022.107553.