19th Australian International Aerospace Congress

Please select category below:
Normal Paper ☐
Student Paper ☒
Young Engineer Paper ☐

Helicopter vibration-based operating regimes identification through the use of mixture models on health indicator

Maxime Meuterlos[1,2], Valerio Camerini[2], Jérôme Antoni[1], Lucas Macchi[2] and Gianni Naccarato[2]

[1] Univ Lyon, INSA Lyon, LVA, EA677, 69621 Villeurbanne, France
[2] Airbus Helicopters, Aéroport International Marseille Provence, 13700 Marignane

Abstract

Understanding the behavior of vibration-based condition indicators for health monitoring of helicopters can be difficult. The high complexity of mechanical assemblies, combined with a high range of possible operating parameters constitute a challenging environment for vibration analysis. Therefore, a data-driven approach based on statistical modeling is proposed in this work to define vibration-based operating regimes. Specifically, parametric clustering based on a Gamma mixture model is considered to associate vibration-based health indicators to helicopter operating regimes through Naïve Bayes classification. An entropic criterion is considered for clustering optimization: namely, the Validity-measure (V-measure).

Keywords: condition indicators, operating regimes, Gamma mixture model, clustering, classification, Validity-measure criterion, vibration monitoring, HUMS.

Introduction

Context and problem formulation

Health monitoring of complex mechanical systems such as helicopter gearboxes is a challenging topic. Indeed, operating parameters have an influence on vibration signature and thus health indicators. Understanding vibration behavior relation to operating regimes is essential to designing an efficient damage detection system, minimizing false alarms due to the effect of specific flight conditions [1]. A probabilistic method to associate health indicator statistics with operating regimes is proposed. Other studies dealing with the influence of operating regimes on vibration patterns can be found in [2] [3] [4] [5]. Vibration signatures of helicopter dynamic components can be qualified as cyclostationary signals [6]. In this framework, the amplitude estimate of harmonic components of first order cyclostationary signals x follows a chi-square distribution, which can be generalized as a Gamma distribution function tuned with two parameters α, β for one data value x as defined in Eqn 1. Detailed explanations can be found in [7].

$$p(x|\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} ; x \in \mathbb{R}^+ ; \alpha \in \mathbb{R}^{+*} ; \beta \in \mathbb{R}^{+*} , \quad (1)$$

where $\Gamma(.)$ is the Gamma function.

In HUMS terminology, the amplitude of these harmonics component belong to a set of scalars called health indicators which are used to monitor specific phenomena: gear wear, shaft unbalance, shaft misalignment, etc. To ensure fault capability detection, frequent signal acquisitions are needed. Health indicators (HI) are then sampled with a periodicity which results from a trade-off between HUMS processing and storing capabilities and the need to keep sufficiently frequent acquisitions for timely fault detection. HI can be seen as a stochastic process with hidden statistical states related to specific operating regimes. The overall data in this paper is described by two vectors containing HI values acquired at each time step t_n such as $X = [x_{1:N}]$ associated with corresponding operating regimes vector defined as $C_X = [c_{1:N}^{1:I}]$ (Fig. 1).

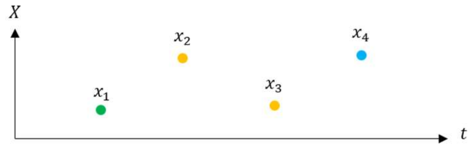

Fig. 1: Example of a Health Indicator $X = [x_1, x_2, x_3, x_4]$ as a function of time t. Colors – operating regimes $C_X = [c_1^1, c_2^2, c_3^2, c_4^3]$

Operating regimes definition

Operating regimes used in this work are defined from loads calculation on helicopter according to operating parameters $\zeta = \{\zeta_{1:p}\}$. Specific ranges (black dots) for each operating parameter are defined. Paths through these ranges determine operating regimes. They will be referred to as classes $C = \{C_{1:I}\}$ in this paper. See an example for three operating regimes definition in Fig. 2.

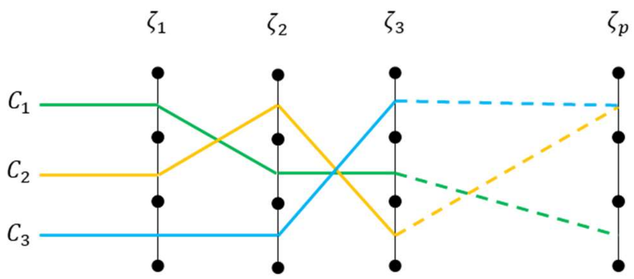

Fig. 2: Definition of three operating regimes C_1, C_2, C_3 as a function of range (black dots) defined on operating parameters $\zeta_{1:p}$

Methodology

The aim of this study is to associate health indicators statistics to helicopter operating regimes. This translates into identifying the association between health indicator clusters S based on a Gamma Mixture model and one or more classes C (i.e. operating regimes) as depicted in Fig. 3. HI statistics are characterized with a mixture model composed of k clusters defined as $S = \{S_{1:K}\}$, as detailed in Sub-section: Health indicator clustering. A maximum a posteriori probability is computed with Naïve Bayes Classifier to associate the operating regimes defined as classes $C = \{C_{1:I}\}$ to clusters S as detailed in Sub-section: Operating regimes classification. The number of clusters K is inferred using a specific criterion as detailed in Sub-section: Model selection.

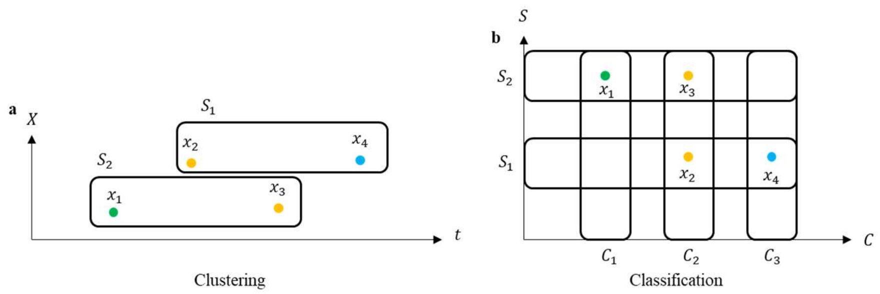

Fig. 3: (a) HI clustering assigning data x_1 to x_4 to clusters S_1, S_2 , (b) association of HI clusters S_1, S_2 (rows) to the operating regimes C_1, C_2, C_3

Health indicator clustering: Gamma mixture model

The hidden HI cluster structure can be obtained through the use of an unsupervised machine learning method called mixture model (see Fig. 4). The hidden structure can be characterized with K clusters $\{S_K\}$. The following variables are introduced: a binary latent variable z_k with k possible states such that $z_k \in \{0,1\}$ with $\sum_{k=1}^{K} z_k = 1$, a mixing coefficient π_k with $0 \leq \pi_k \leq 1$ and $\sum_{k=1}^{K} \pi_k = 1$. The marginal distribution over z can be written as $p(z_k = 1) = \pi_k$ thus for k clusters $p(z) = \prod_k \pi_k^{z_k}$. Consequently, $p(X|z) = \prod_k^K p(X|\alpha_k, \beta_k)^{z_k}$ with α_k, β_k the parameters of the Gamma distribution for the cluster k. By marginalization over latent variable z, the probability density function of the mixture reads (Eqn 2)

$$p(X) = \sum_{k=1}^{K} p(X|z_k)p(z_k) = \sum_{k=1}^{K} p(X|\alpha_k, \beta_k)\pi_k. \quad (2)$$

Determination of the parameters $\{\pi_k, \alpha_k, \beta_k\}$ attached to cluster S_k is done with the Expectation-Maximization (EM) algorithm once the number of clusters K is obtained from the model selection criterion.

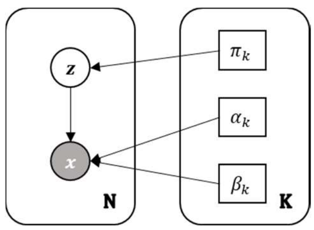

Fig. 4: Graphical representation of a Gamma mixture model (observed data in the shaded circle and small squares indicate fixed parameters) with π_k mixing coefficient, z latent variable attached to data x, α_k and β_k the parameters of the Gamma distribution in cluster k. N is the number of data points and K the number of clusters.

Operating regimes classification: Naïve Bayes Classifier

The most probable cluster associated with an operating regime is estimated using a Naïve Bayes Classifier (Eqn 3). Operational regimes are linked to a cluster by computing the maximum a posteriori probability of the associated HI values belonging to the cluster S_k defined with $\{\pi_k, \alpha_k, \beta_k\}$ parameters (Eqn 4, Eqn 5), i.e.

$$C_i \in S_{k^*};\ \forall c_n^i \in S_{k^*}\ ;\ k^* = \max_k\big(p(S_k|x_n)\big), \tag{3}$$

where

$$p(S_k|x_n) = \frac{\pi_k p(x_n|S_k)}{\sum_{k=1}^{K} \pi_k p(x_n|S_k)}, \tag{4}$$

and

$$p(x_n|S_k) = p(x_n|\alpha_k, \beta_k). \tag{5}$$

In the above equations, S_k is the cluster, c_n^i corresponds to the operating regime associated to x_n and C_i correspond to the operating regimes. Consequently, the operating regimes C_i are attributed to clusters S_k through the corresponding HI values (see Fig. 5).

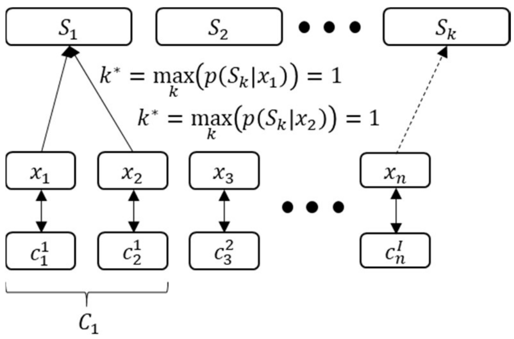

Fig. 5: Maximum a posteriori clustering with naïve Bayes classifier. The N values x_n and the associated I operating regimes c_n^i are attributed to the K clusters S_k through their a-posteriori probabilities.

Model selection: V-measure (Validity-measure)

Recalling the hypothesis for selecting the number of clusters, the idea is to maximize the separation of the clusters based on the operating regimes. Such criterion can be seen as equivalent as to minimizing the conditional entropy on clusters given a set of classes and it is known as the V-measure [8]. The V-measure (Eqn 6) is based on NMI (Normalized Mutual Information) and allows a balance between homogeneity (Eqn 7) or completeness (Eqn 8) within each cluster thanks to a penalty coefficient γ.

$$V_\gamma = \frac{(1+\gamma)hc}{\gamma h + c}, \tag{6}$$

where

$$h = 1 - \frac{H(C|S)}{H(C)} \; ; 0 \leq h \leq 1\,, \tag{7}$$

and

$$c = 1 - \frac{H(S|C)}{H(S)} \; ; 0 \leq c \leq 1. \tag{8}$$

If $\gamma \ll 1$ then $V_\gamma \rightarrow h$, the optimal value of clusters is to favor homogeneity in other words all datapoints in cluster S_k have the same class. At the opposite, when $\gamma \gg 1$ then $V_\gamma \rightarrow c$, the optimal value of clusters is to favor completeness; in other words, all datapoints of a given class have been assigned to the same cluster S_k.
In this work, the concern is to favor completeness, thus $\gamma \gg 1$ in Eqn 6.

Results

In this section, the methodology outlined in the section: Methodology is applied on data from a helicopter fleet composed of a HI representing the amplitude of a harmonic component which is synchronous with the rotor blade frequency and measured in the lateral direction of flight. Classes (i.e. operating regimes) are defined by maneuvers through flight parameters, which are measured contextually to the HI values, aircraft weight, and center of gravity position. Around 300 operating regimes C_i are identified in this application. Fig. 6 shows the method applied to the HI described above. To maximize the V-criterion, three clusters are selected with arbitrary $\gamma = 1e2 \gg 1$. Overlap rate is the ratio between the number of classes associated with more than one cluster and the total number of classes. The probability density function of the HI is given in Fig 6.a. Fig 6.b shows the association of operating regimes (abscissa) with the identified data clusters (ordinates). Red dots represent the regimes belonging to multiple clusters, and blue dots those operating regimes linked with one and only one cluster. The resulting overlap rate is around 26% in this case. Consequently, 74% of operating regimes is associated with a specific vibratory state while 26% of operating regimes do not define a clear vibratory state or do not impact the statistical states of the HI. However, the high density of acquisitions within specific operating regimes limits the information in the rest of the envelope (Fig. 7) for the available data.

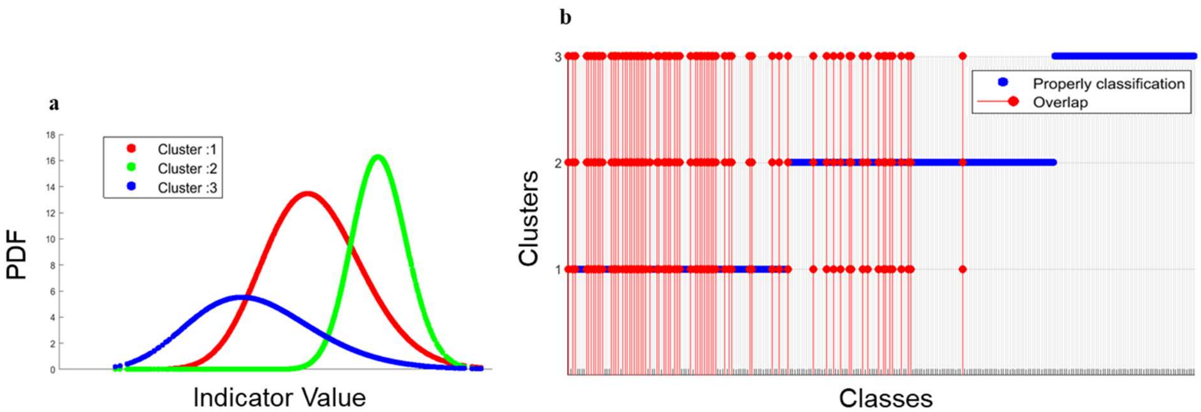

Fig. 6: (a) probability density function of the identified clusters on HI values (clusters: red, green, blue), (b) Correspondence between the classification of operating regimes and clustering of HI (blue dots) and overlap (red dots)

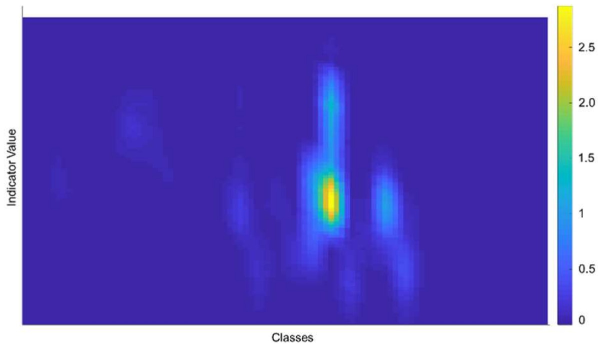

Fig. 7: HI probability density function estimated with a box kernel as a function of classes (i.e. operating regimes)

Conclusion

In this paper, a clustering-classification method is introduced with the objective of capturing the statistical behavior of a health indicator associated with operating regimes. A posteriori characterization of health indicator statistics with a data-driven model is still a complex task because of high dimensionality problem and operating regimes definition. It was found that some operating regimes result in a distinct vibration behavior and are thus linked to one data cluster, whereas another part of them does not affect the health indicators statistics and it is hence associated to multiple clusters.

References

[1] V. Camerini, G. Coppotelli, and S. Bendisch, "Fault detection in operating helicopter drivetrain components based on support vector data description," *Aerosp. Sci. Technol.*, vol. 73, pp. 48–60, Feb. 2018, doi: 10.1016/j.ast.2017.11.043.

[2] E. Huff, I. Tumer, E. Barszcz, M. Dzwonczyk, and J. Mcnames, "Analysis of maneuvering effects on transmission vibration patterns in an AH1 cobra helicopter," *J. Am. Helicopter Soc.*, vol. 47, Jan. 2002.

[3] M. Mosher, E. M. Huff, and E. Barszcz, "Analysis of in-flight measurements from helicopter transmissions," p. 15, 2004.

[4] E. Huff, I. Tumer, E. Barszcz, D. Lewicki, H. Decker, and J. Zakrajsek, "Experimental Analysis of Mast Lifting and Bending Forces on Vibration Patterns Before and After Pinion Reinstallation in an OH-58 Transmission Test Rig," Feb. 2001.

[5] E. Huff, E. Barszcz, I. Tumer, M. Dzwonczyk, and J. McNames, "Experimental Analysis of Steady-State Maneuvering Effects on Transmission Vibration Patterns Recorded in an AH-1 Cobra Helicopter," *undefined*, 2000, Accessed: Jun. 21, 2021. [Online]. Available: https://www.semanticscholar.org/paper/Experimental-Analysis-of-Steady-State-Maneuvering-Huff-Barszcz/03f5486bbfacb307f874dc8399ef86e667c56a04

[6] J. Antoni, F. Bonnardot, A. Raad, and M. El Badaoui, "Cyclostationary modelling of rotating machine vibration signals," *Mech. Syst. Signal Process.*, vol. 18, no. 6, pp. 1285–1314, Nov. 2004, doi: 10.1016/S0888-3270(03)00088-8.

[7] V. Girondin, K. M. Pekpe, J. P. Cassar, and H. Morel, "Bearings fault detection in helicopters using frequency readjustment and cyclostationary analysis," *Mech. Syst. Signal Process.*, vol. NA, no. NA, p. NA, Mar. 2013, doi: 10.1016/j.ymssp.2013.03.015.

[8] A. Rosenberg and J. Hirschberg, "V-Measure: A Conditional Entropy-Based External Cluster Evaluation Measure.," Jan. 2007, p. 420.