Estimation of Health Indicators using Advanced Analytics for Prediction of Aircraft Systems Remaining Useful Lifetime

Daniel Azevedo[1], Alberto Cardoso[1], and Bernardete Ribeiro[1]

[1] *University of Coimbra, CISUC, Department of Informatics Engineering, Portugal*
dazevedo@student.dei.uc.pt
alberto@dei.uc.pt
bribeiro@dei.uc.pt

Abstract

A valuable asset for the improvement of aviation maintenance is the correct assessment of the aircraft systems health condition, for a more accurate planning and execution of maintenance routines. As such, the creation of a Prognostic and Health Management (PHM) system, supported by Condition Based Maintenance (CBM) can bring important benefits to the aeronautical field. The ultimate goal of a PHM system is the correct prediction of the Remaining Useful Lifetime (RUL) of a certain aircraft system, using the sensors data extracted during flights. Nevertheless, a relevant stage in the PHM pipeline concerns the estimation of the system condition, expressed by the Health Indicator (HI). The HI value reflects the health condition of a specific aircraft system, which can possibly be affected by degradation, failures or external conditions occurred during flight time. Henceforth, due to the relevancy of the HI assessment for the accuracy of the PHM model, this paper aims to propose a new formulation for the HI computation, derived from raw anonymized data retrieved from different sensors within the aircraft system. The proposed formulation combines information from the different variables (like sensors) that have an impact on the overall system condition, by assigning a positive or negative weight to each variable depending on the influence on the system behaviour. The weights are determined based on the typical and standard data patterns. Thus, the estimated HI aims to reflect the number of hours of flight expected to be flown, based only on raw data extracted from the system. Furthermore, considering that the available sensors data is anonymized, a study of the relevancy of the different sensors features for the degradation assessment is performed, using specific metrics. Considering a scenario where the HI ground truth is unknown, the failure data of each aircraft system is used to evaluate and discuss the formulation suitability. The HI formulation is applied in real datasets, on the environmental systems of two wide body aircraft types.

1. Introduction

Aircraft maintenance has an important role in the aircraft activity. Due to the increasing importance of air transportation in the modern world, aircraft are expected to operate under a higher utilization, which results in a significant number of operation hours over varied conditions. This leads to an increasing degradation over time of the aircraft systems or components which need to be monitored in order to prevent them to reach a state where they are likely to fail. It is the goal of aircraft maintenance to assure that the conditions for the aircraft to operate normally are gathered, in order to guarantee the safety of the airplane and its passengers. To accomplish this, different types of tasks may be needed to be performed by the maintenance engineers, like pre-flight and post-flight inspections, routines check and parts replacement due to malfunctions (Papakostas, Papachatzakis, Xanthakis, Mourtzis, & Chryssolouris, 2010). One of the main problems faced in the maintenance routines concerns the unplanned maintenance tasks. These can take up to 50% of the maintenance work (Samaranayake & Kiridena, 2012), depending on the aircraft and system, which also incurs in additional costs. For instance, if a large aircraft flight, like Boeing 747, is canceled, it can cost the airline US $140,000, also, a delay at the gate can have a cost of US $17,000 per hour (Dalkilic, 2017). Regarding the different types of maintenance routines, two of the most used methodologies are Preventive Maintenance (PM) and Corrective Maintenance (CM) (Wu, Liu, Ding, & Liu, 2004). The PM performs the inspections and interventions based on previously established fixed time intervals, the interventions are scheduled regardless of the condition of the component and with the goal of anticipating failure. The CM, as the name suggests, takes place after the failure has occurred with the goal of repairing or replacing the component or part with defect or malfunction. This type of maintenance is more related to failures that do not compromise the aircraft's normal operation. Although these two approaches are valid and widely used, they are not optimized in terms of costs, scheduling or plan fulfillment. In recent years, a different approach, designated as CBM, has been introduced to improve aircraft maintenance. CBM focuses on the monitoring of the system's degradation condition for predicting the time of failure and, therefore, planning and performing the interventions at the right moment before failure. This condition-based approach may help reduce the maintenance costs of unscheduled work, as well as, the intervention's duration as the malfunction is addressed in time. As such, a way to improve aircraft maintenance is to provide more relevant information of a system or component degradation condition in order to better detect and predict failures. This would lead to more adequate planning and execution of the maintenance routines, following the CBM standards.

A possible approach for assessing the system condition is through the HI computation. The HI value reflects the health condition of a particular system or component of an aircraft, which can be influenced by multiple factors like wear, physical damage and failures (Lei et al., 2018). In some cases, it can be challenging to determine the HI of a specific system, due to the lack of relevancy or absence of the available data (normally sensors data) or the inadequate system behavior when faced with a failure (Guo, Lei, Li, Yan, & Li, 2018). According to the literature, there are two main categories of HI, the Physical Health Indicator (PHI) and the Virtual Health Indicator (VHI) (P. Wang, Youn, & Hu, 2012). The PHI uses physical features directly extracted from the system as a direct metric for the HI. This requires a physical understanding of the system behavior and the presence of a signal or feature representative of system degradation condition. Traditional statistical methods and signal processing techniques are common approaches to compute the PHI, where features like Root Mean Square (RMS), peak and kurtosis are used as relevant features (Lei et al., 2018), (Li, Lei, Lin, & Ding, 2015). The VHI is used when there is not a singular feature representative of the system degradation. Therefore, VHI combines and fuses different features with the goal of representing the global condition of the system (Lei et al., 2018). Generally, the values obtained for the VHI do not have a physical meaning associated as they are a result of a combination of various features. Some common techniques to compute the VHI are Principal Component Analysis (PCA) (Mina & Verde, 2005), Self-Organizing Map (SOM) (Hai Qiu & Yu, 2003) and Linear Regression (T. Wang, Yu, Siegel, & Lee, 2008).

The HI can be used in different systems and over different representations, generally, the choice of the technique used for the HI computation should depend on the system being analyzed and the characteristics of the available data. When working with time domain features, several techniques for computing the HI were presented using features like kurtosis, RMS and peak. In (Zhang, Si, & Hu, 2015), a band-pass filter was applied to the kurtosis feature with the goal of estimating the bearing degradation path. The authors in (Li et al., 2015) extracted the RMS and kurtosis features from vibration signals, where kurtosis was used for health monitoring and RMS for the RUL prediction. The frequency domain of the data has also been a subject of study for determining the HI. One of the most used methods for analyzing the frequency spectrum of the data signals is the Fast Fourier Transform (FFT). (Liao, 2014) decomposed a signal into the different frequency bands, and used the energy of each as input features to compute the HI. The HI was obtained by running a genetic algorithm, which identified the best combination of features that better described the fault evolution. In (Zhang, Wang, & Wang, 2013), after decomposing the signal with the FFT method, the relevant peaks were selected from the obtained series and were used as input to a Neural Network (NN) trained to predict the RUL in manufacturing systems. Some also used Envelope analysis for detecting faults using the frequency domain. The authors in (Boškoski, Gašperin, & Petelin, 2012), combined Gaussian process models with Envelope analysis for computing the RUL in bearings. With regards to time-frequency analysis, significant research in this domain was also performed. (He, Miao, Azarian, & Pecht, 2015) proposed an approach where the Wavelet Transform is used for computing the HI in vibration signals. The HI consisted of the sum of the amplitudes of the relevant signal characteristics extracted from bearing data.

All these case scenarios share a common characteristic, which is the knowledge of the type of data being used for computing the HI, nevertheless, there are some cases where the data is anonymized and thus there is no information concerning the data characteristics.

In this work, a new formulation is proposed for computing the HI based on raw and anonymized data. An anonymization scenario is used in order to represent a generic and agnostic approach. In the formulation, different data features are combined by assigning a weight to each feature that reflects its positive or negative impact in the HI computation. Two datasets regarding two different aircraft systems are used in this work to test and validate the proposed approach. Furthermore, these datasets contain real sensors data extracted directly from the aircraft and were not subject to any preprocess. This way, the health condition (expressed by the HI) is extracted entirely based on the sensors data of the respective system, using a complete black-box approach with regards to the data characteristics.

This paper is organized as follows. Section 2 presents the HI formulation proposed in this work, as well as some important aspects with regards to the formulation. Section 3 describes the experimental scenario and in Section 4 the results are presented and discussed. Finally, the conclusions drawn and future work are specified in Section 5.

2. HEALTH INDICATOR APPROACH

As stated before, the HI aims to reflect the aircraft system health condition, which can be influenced positively or negatively by several factors. The most basic and direct factor concerns the flight hours (duration), as the number of flight hours operated by an aircraft increases, the accumulated degradation in that aircraft also increases. Other important factors that affect the system condition are the flight conditions, the weather conditions and the flight trajectory. As the sensors data are expected to express the system degradation, the aircraft systems condition is extracted using only the respective sensors data, which comprise the available datasets. Hence, the HI is computed by assigning a weight to each sensor feature considered relevant for the degradation assessment.

In a general form, the HI formulation proposed in this paper is the following:

$$HI = \sum_{k=1}^{n} \sum_{j=1}^{p} duration_{j,k} \cdot (1 + (\alpha_{1_{j,k}} + \alpha_{2_{j,k}} + ... + \alpha_{m_{j,k}})) \tag{1}$$

where n is the number of flights, p is the number of aggregated phases in a single flight, $duration_{j,k}$ is the duration of the aggregated phase j of flight k and m is the number of features considered relevant. Moreover, the α's vector corresponds to the weights assigned to each considered feature.

This way, the HI (expressed in flight hours) corresponds to the flights duration with possible additions due to degradation reflected in the chosen features, whose impact is indicated by the respective α value. Therefore, the HI value aims to reflect the number of flight hours likely to be flown based only on the sensors values, where the α's determine the increase that should be added to the flights duration (baseline).

It should be noted that the proposed HI formulation is made under the assumption that the aircraft system being analyzed is operating during the entire flight (example: turbofan engine and air conditioning system) and not only during a specific time period, as the brake and wheels system.

2.1. Weights Assignment

An important step for the correctness of the formulation corresponds to the determination of the weight associated to each selected feature, represented by α's vector.

The value of $\alpha_{i_{j,k}}$ represents the weight assigned to feature i regarding the aggregated phase j of flight k. The values of $\alpha_{i_{j,k}}$ should be interpreted as follows:

- If $\alpha_{i_{j,k}} > 0$: Means that in flight k and aggregated phase j, the value of feature i is too deviated from the typical values for this feature; therefore, it corresponds to a situation of **Extra Degradation**;
- If $\alpha_{i_{j,k}} = 0$: Means that in flight k and aggregated phase j, the value of feature i is within the typical range of values for this feature; therefore, it corresponds to a situation of **Normal Degradation**.

It is worth noting that the situation of "Lesser Degradation" could also be added, which would represent the scenario where the degradation is lesser than the expected. Nevertheless, in the absence of information regarding the features expected behavior, this scenario was not considered in this work.

By analyzing the HI formulation (Eq. (1)), one can conclude that, if for a specific flight and aggregated phase, the sum of the α values is greater than 0 it means that during this time interval there was some anomaly reflected in the sensors data that leads to an extra increase in the HI value. Otherwise, if the sum of the α's is equal to 0, it means that this time interval corresponds to a situation of normal degradation, and so the HI is equal to the time interval $duration$.

For the determination of the α's values, the flowchart in Figure 1 is proposed:

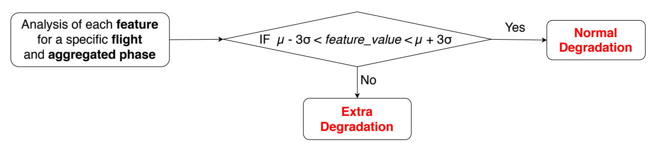


Figure 1. Flowchart for α determination

For assessing what is a typical or abnormal value for each feature, flights reflecting a healthy state are taken as a base for comparison. Therefore, the mean (μ) and standard deviation (σ) are computed with respect to each feature and these values represent the typical and usual values for the respective feature. These variables (μ and σ) are computed from a set of flights considered healthy, which means, they occurred after an inspection or a part replacement was performed and their sensors assume regular and standard data patterns. This approach of using healthy data as a base for comparison has already been applied in literature, in similar contexts (Nguyen et al., 2018), (Jardine, Lin, & Banjevic, 2006) , (Sun, Li, Liu, Gong, & Wang, 2018).

The Piecewise Function 2 illustrates the way the α are determined for a specific feature ft.

$$a_{ft_{j,k}} = \begin{cases} 0 & \text{if} \quad \mu - 3\sigma < F < \mu + 3\sigma \\ \gamma[\frac{F-(\mu+3\sigma)}{MAX-(\mu+3\sigma)}] & \text{if} \quad F > \mu + 3\sigma \\ \gamma[\frac{(\mu-3\sigma)-F}{(\mu-3\sigma)-MIN}] & \text{if} \quad F < \mu - 3\sigma \end{cases} \tag{2}$$

In this function, F is the value of feature ft for flight k and aggregated phase j, MAX and MIN represent the maximum and minimum value, respectively, known for feature ft and γ corresponds to a weight that influences the α magnitude and should be chosen according to the considered system.

The idea is that α is 0 when facing a situation of Normal Degradation. Otherwise, the α should be greater than 0 and can reach the value of γ depending on the distance to the extreme boundaries (Max boundary or Min boundary). Figure 2 illustrates an example of the values of a specific feature for 30 consecutive flights.

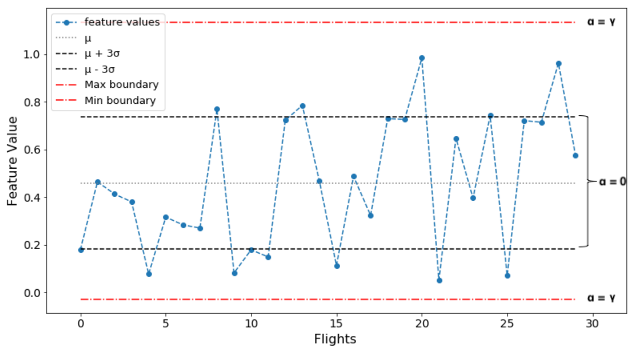


Figure 2. Example of α determination for a specific feature

In this work, the choice of the γ value (0.6) was made without any awareness of the system behavior, as there is no information regarding the HI ground truth, no special tuning was performed with regards to the choice of γ value.

2.2. Features Relevancy

Another important aspect is the choice of the most relevant features, extracted from the sensors data, for the degradation assessment. Although all the sensors available for an aircraft system are located in the respective system, not all of them are relevant for extracting the health condition. Therefore, the choice of the most relevant sensors and respective features is important for the HI computation.

In the literature, some evaluation metrics have been proposed and can be adapted to use for the computation of the feature relevance for the degradation assessment (Coble & Hines, 2009), (Lei et al., 2018), namely:

- **Monotonicity**: The monotonicity metric expresses the positive or negative trend reflected in feature values over time. Ideally, in order for the feature to be relevant for the degradation assessment, its values should have a clear monotonic increase/decrease over time. It can be computed by analyzing the difference between the positive and negative derivatives along the feature values vector, as follows (Coble & Hines, 2009):

$$Monotonicity(F) = \left| \frac{\# \frac{dF}{df} > 0}{n-1} - \frac{\# \frac{dF}{df} < 0}{n-1} \right| \tag{3}$$

  where n is the number of observations for a given period of time, F represents the feature values obtained for a set of consecutive flights, f represents the flights considered and $\frac{dF}{df}$ corresponds to the derivative of F with respect to f. An important step for this formula's correctness is applying the correct smoothing of the feature values. Another possible method for computing the monotonicity of a feature is to apply the Spearman's rank correlation coefficient to the feature vector.

- **Prognosability**: The prognosability metric aims to evaluate the variance in the failure threshold of different trajectories or systems. With a lower value it would be easier to extrapolate or predict the degradation evolution more accurately. It can be computed using the following formula (Coble & Hines, 2009):

$$Prognosability(F) = exp\left(\frac{-std(F_{fail})}{mean(|F_{start} - F_{fail}|)} \right) \tag{4}$$

  where F represents the feature values.

- **Trendability**: The trendability metric indicates the degree of similarity of the shape or form of the values of a feature, regarding different trajectories or systems. The computation of the similarity between different forms can be challenging, nevertheless, a proposed formulation is (Coble & Hines, 2009):

$$t_F = \left| \frac{\# \frac{dF}{df} > 0}{n-1} + \frac{\# \frac{dF^2}{df^2} > 0}{n-2} \right| \tag{5}$$

$$Trendability(F) = 1 - std(t_F) \tag{6}$$

  where n is the number of observations for a given period of time, F represents the feature values, f represents the flights considered and t_F represents the trend of feature F in the different trajectories.

Besides these metrics focusing on degradation assessment, another potential criteria for selecting the most relevant features is their capability in exposing the faults occurred during flights. Using a statistical test, the relevant features are the ones that can better distinguish faulty flights from normal flights, by detecting the faulty data patterns. Nevertheless, the high capability to distinguish faulty behavior may not indicate that a feature is adequate for capturing the degradation over time.

This way, two different criteria are proposed for selecting the most relevant features. The first is based on the degradation assessment (monotonicity, prognosability and trendability) and the second is based on the fault detection. In both situations, a previous study is required in order to understand which features are more relevant according to the aircraft system and criterion used, this step is described in Section 3.1.

3. EXPERIMENTAL SCENARIO

The HI formulation proposed in Section 2 is tested and validated using real sensors data retrieved from 2 distinct wide-body aircraft systems, henceforth referred as System 1 and System 2. System 1 and System 2 have similar functions within the aircraft although each system operates in a particular type and model of airplane. They are responsible for an important procedure in the airplane operation which is the air conditioning task. The air coming from the outside is processed and altered in the adequate conditions of pressure and temperature so that it can be sent to the flight deck and to the passenger cabin. Although the systems have similar functions, the way they perform the air processing and conditioning can be distinctive, as such, the failure modes and the systems reaction to failures are singular, thus, the degradation behavior is assumed to be different. Also, the sensors positioned in each system are different, therefore the data retrieved from each system is different in quantity and meaning.

In System 1, the data was retrieved from 20 sensors with a sampling rate of 4Hz. In System 2, the data was retrieved from 90 sensors with a sampling rate of 1Hz. As it corresponded to raw sensors data some preprocessing and data transformation steps were performed, namely flights filtering and flight phases aggregation.

By analyzing the data some variation of the flights duration was found, sometimes associated with inconsistencies regarding the flight phases information. Therefore, if the flight duration was too short (maybe a maintenance routine) or there were some missing data regarding an important flight phase (like the cruise phase) the flight was filtered and not considered for analysis.

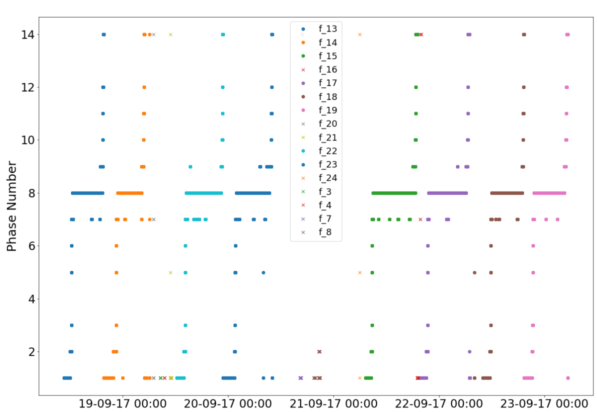


Figure 3. Flights labeling before filtering (System 1)

Figure 3 illustrates an example of the original flights labeling, obtained directly from the raw data. Figure 4 illustrates the flights labeling after filtering, where a removal of 8 "noisy flights" can be seen.

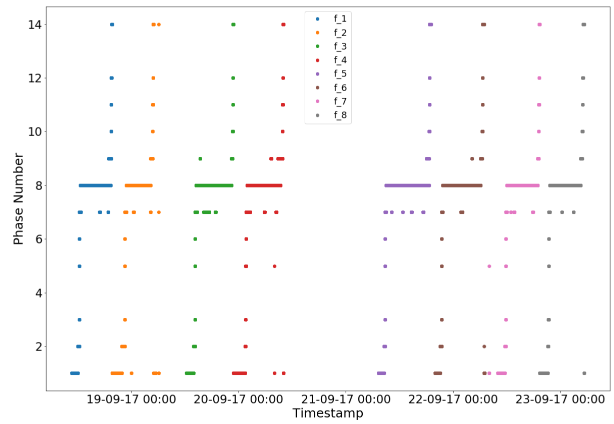


Figure 4. Flights labeling after filtering (System 1)

Another preprocessing step was regarding the flight phases aggregation. In order to perform a more adequate and accurate analysis of the sensors data, the 14 original phases of each flight in the dataset were combined in a set of 5 aggregated phases, as demonstrated in Figure 5.

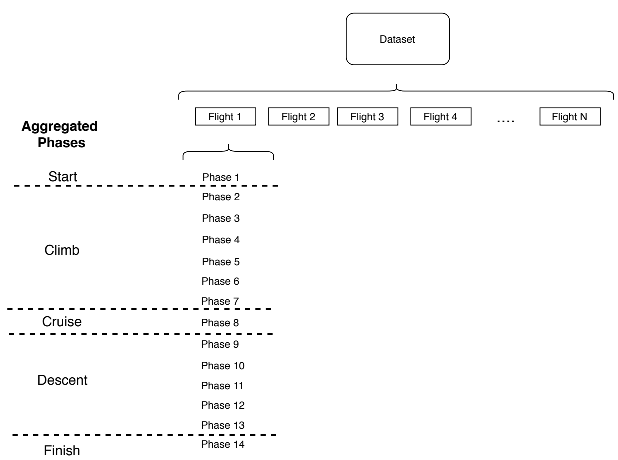


Figure 5. Flight phases aggregation

This way, the different data patterns in the sensors of singular flights are isolated, as exemplified in Figure 6.

As the presence of outliers can be a consequence of failure or degradation, no special technique was applied to remove the data outliers.

3.1. Identification of the relevant features

In order to compute an accurate HI, the most relevant sensor features for the degradation assessment should be selected

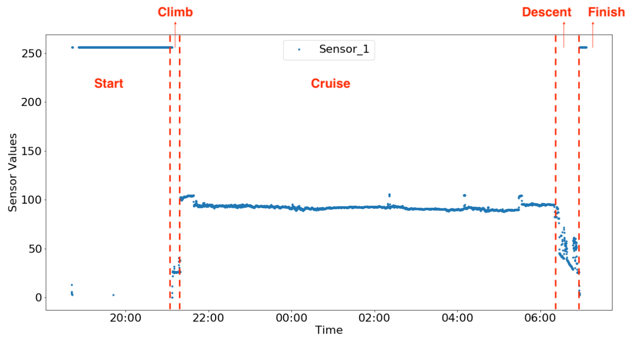


Figure 6. Example of flights phases aggregation

and used in the formulation (Eq. (1)).

Considering the data sampling rate, the extraction of features in the frequency domain or time-frequency domain should be meaningless as the relevant information would already be filtered. As such, the features that are extracted from the sensors data and whose relevance is analysed are related to the time domain. These correspond to common time domain features like median, variance, kurtosis, maximum, minimum, sample entropy, quartiles, amongst others (Yan, Qiu, & Iyer, 2008). The features extraction and analysis is performed for all the sensors in the respective system.

Some of the enunciated metrics were used to evaluate the features relevancy in detecting the degradation evolution over time. By applying the monotonicity formula (Eq. (3)) and also the Spearmans rank correlation coefficient method, no conclusive results were obtained. This was due to the fact that none of the extracted sensor features in System 1 or System 2 showed a clear monotonic increasing/decreasing trend in its values over time.

The prognosability metric was not considered adequate to be used in this context, as the features values range is the same over time. In the same way, trendability metric was also not used due to the nonexistence of distinguishable forms or shapes in the features values evolution over time, for different trajectories of the aircraft.

The next step was to assess the features relevancy for the HI based on the capacity of discriminating flights where a fault had occurred (faulty flights) from normal flights, where no fault was registered. To achieve this, first, a visual analysis was performed by applying non linear feature reduction techniques (Isomap and Locally Linear Embedding) to the entire set of features extracted from all the sensors. Figure 7 and Figure 8 illustrate the results for a given period of time.

As can be concluded in Figure 7, there is a low capability of discrimination of the different types of failures in System 1. On the contrary, in System 2, there is a high separability between normal flights and faulty flights (Figure 8). It should be

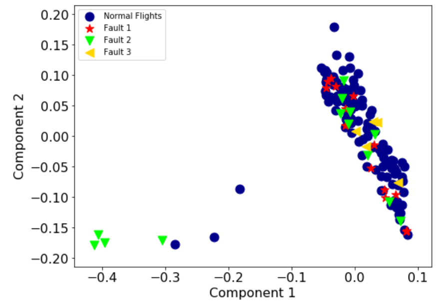


Figure 7. Flights representation - System 1

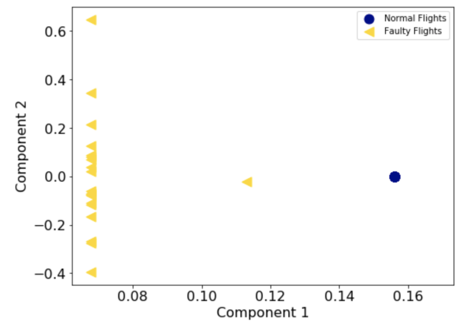


Figure 8. Flights representation - System 2

noted that there may exist redundancy or correlation between the different faults, as, although the fault names are different, they may be encompassed in the same failure mode.

For the selection of the most relevant features of each sensor aiming to detect faults, a statistical test was used. By applying a two-sample T-test, the goal was to identify the features in which the difference between the mean of normal flights and faulty flights was statistically more significant. The most relevant features identified were:

- **System 1:**
  - **Sensor_1:** standard deviation, variance
  - **Sensor_2:** standard deviation, variance
  - **Sensor_3:** maximum, median, 3rd and 4th quartiles
  - **Sensor_4:** sample entropy, minimum, mean, 3rd and 4th quartiles
- **System 2:**
  - **Sensor_3:** minimum, median, 1st and 2nd quartiles
  - **Sensor_4:** minimum, median, 1st and 2nd quartiles
  - **Sensor_6:** median, 1st and 2nd quartiles
  - **Sensor_7:** minimum, median, 1st and 2nd quartile

Figure 9 and Figure 10 illustrate the reduced data points using only the identified relevant features.

Through Figure 9, one can conclude that the different failure classes in System 1 are more easily separable using only the

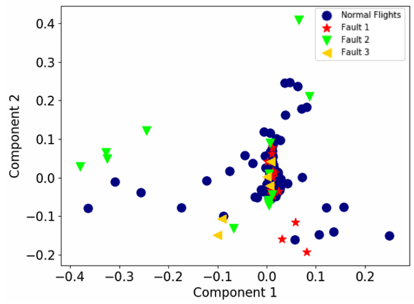


Figure 9. Flights representation using only relevant features-System 1

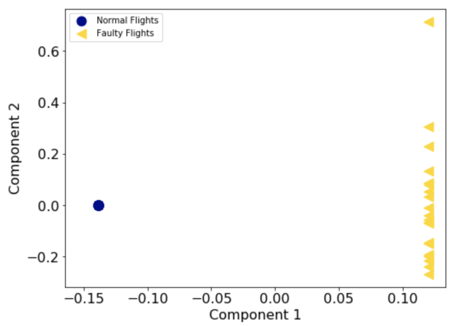


Figure 10. Flights representation using only relevant features-System 2

relevant features, although there is still embedded noise, that is, normal flights near the faulty flights and vice-versa. Regarding System 2, the data separability remains higher, with a slight increase in the classes margins, as can be depicted from Figure 10. As it turns out, when a fault occurs in System 2, the entire system is shut down which leads to a change in the sensors values. Therefore, it is reasonable and expected that the sensors features can accurately separate the normal flights from faulty flights.

Regarding the α determination, the methodology described in Section 2.1 is applied in both Systems. The overall idea is illustrated in Figure 1 and is complemented by the Piecewise Function 2. As one can conclude, the α are determined based on the deviation of the sensors features values, with respect to the typical and reference values of each feature.

4. Results of the HI application

The proposed HI formulation has been applied to different time intervals of System 1 and System 2 with the goal of presenting the HI evolution in real case scenarios and discussing the correctness of the obtained results. In the absence of the HI ground truth, the presence of the faults occurred over time is used for validation purposes. As the faults are suppose to represent some anomaly in the system, a larger increase of the HI is expected in the flights where the faults were recorded, as well as, in the preceding flights.

For an easier interpretation of the HI graphs, the faults are indicated with the letter "F" in the respective flight where they took place. The vertical blue lines represent the consecutive flights, where each vertical line defines the end of a flight and the beginning of the next one. Also, it is worth noting that each time interval considered does not correspond to the start of the degradation cycle. Ideally, the HI should be 0 when the system is new, which corresponds to the starting point of the degradation evolution. For comparison purposes, the "Normal Behavior" curve is illustrated in each graph, this line corresponds to the function $y = x$ and reflects a typical and ideal scenario, where due to absence of anomalies in the data, the HI is equal to the flights $duration$.

4.1. System 1

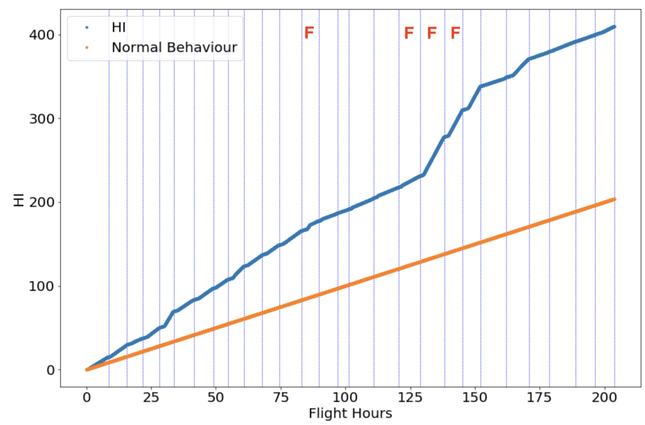


Figure 11. Time Interval 1 - System 1

Figure 11 represents the HI evolution for a specific time interval of System 1. As observed, the majority of the generated faults are related to an extra increase of the HI value, which is a good evidence of the formulation suitability with regards to the degradation detection. Nevertheless, this relation is not always true, as there are some flights where a fault was reported but there was no evidence of it by looking at the HI curve. The contrary may also happen, an extra increase in the HI value without the presence of faults may be an indicator that a failure will occur in a near future, hence, even without the occurrence of a fault, it can be an important indication of degradation.

4.2. System 2

The analysis of the HI evolution in System 2 can be more challenging due to the fact that there is a minor presence of faults over time, with a greater accumulation of faults in the final stage of degradation where, due to the impact of failure, the system is shut off and the sensors assume a default value.

In Figure 12, a higher increase in the HI during the flight where a fault was generated, is observed. This means that, in

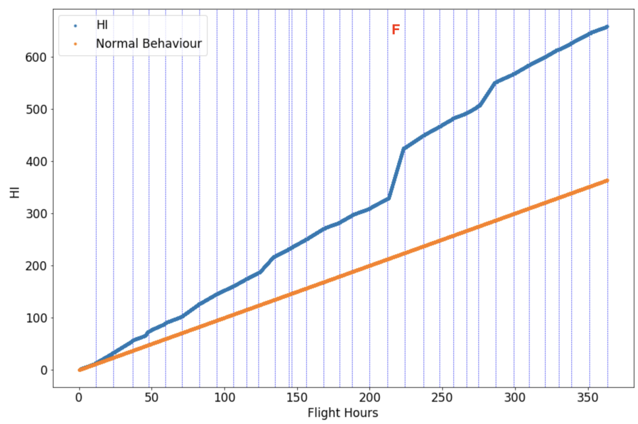


Figure 12. Time Interval 1 - System 2

that specific flight, there were some anomalies in the sensors data, expressed in its features, that are a consequence of the occurrence of a fault. Although this corresponds to the ideal scenario, it may not happen in all the case scenarios.

Figure 13 and Figure 14 represent time intervals in different stages of degradation, Figure 13 illustrates a final stage in the degradation cycle (before failure) and Figure 14 illustrates an initial stage of degradation, after a successful parts replacement, required due to the presence of failure.

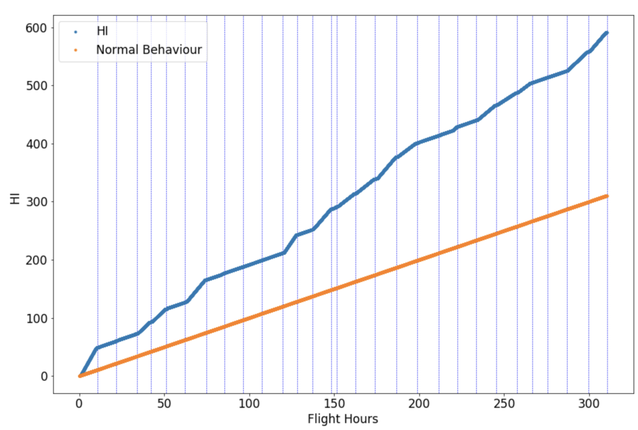


Figure 13. Time Interval 2 - System 2

One can verify that in the initial stage of degradation (Figure 14) the HI evolution is smoother than in the final stage of degradation (Figure 13). This behavior is expected as the newer the system is, the lower the degradation rate should be. Therefore, even without the presence of faults, HI increase should be higher in the final stages of a system lifetime.

Overall, by analyzing the HI evolution in the considered time intervals, interesting results were obtained with regards to the relation between the presence of a fault and the increase of HI. However, there are some possible reasons for this association not being always true. The fact that an extra increase of HI

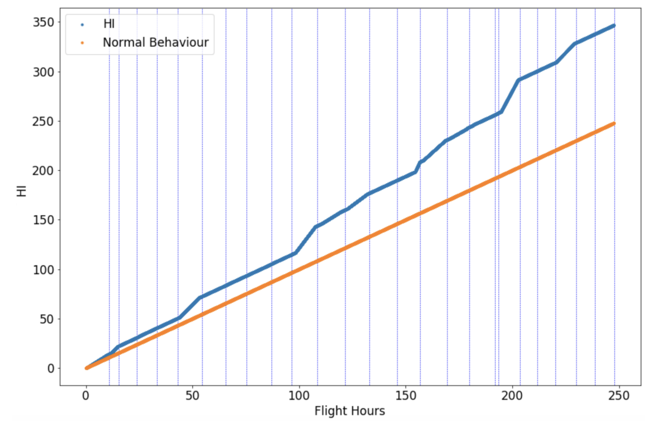


Figure 14. Time Interval 3 - System 2

is not always related to the occurrence of a fault can be due to an error in the data logging or to the fact that it may be contributing to an accumulation of degradation that will cause a fault in a near future. Opposingly, when the presence of a fault is not associated with an extra increase in the HI, it can be mainly because of two reasons. Firstly, the sensors data do not reflect the fault, and, consequently, the HI will not reflect it either. Secondly, it could be that the fault corresponds to a false positive, which means that, although a fault was raised there is not a real reason for it.

5. CONCLUSION AND FUTURE WORK

In this work, a novel approach was developed for computing the Health Indicator (HI) of a given aircraft system. The novelty in the HI formulation (Eq. (1)) concerns the combination of different sensor features, each with an assigned weight that reflects its importance and impact in the health condition of the system. The weights, expressed by the α, reflect the degradation found on the sensors features and were determined based on anomalies and irregular patterns found in the data, using reference values. As such, predefined thresholds were used to assess the different levels of degradation. This way, the overall HI aims to reflect the evolution of the degradation embedded in the system over time and is computed based totally on a data driven approach using anonymized sensors data.

In this unsupervised context, the validation of the proposed approach may be challenging, however, the results of its application to real case scenarios (System 1 and System 2) are interesting. In general terms, they show a relationship between the presence of a fault and the increase in the HI value, which confirms the capability of the proposed formulation to detect anomalies in the system being analyzed. This approach was developed with the goal of more accurately estimating the system degradation, expressed by the HI, which can be used for improving the effectiveness and accuracy of aircraft maintenance interventions. It should be used as a complementary tool for performing the diagnosis of a particular system health condition as it is not intended to substitute the aircraft engineers expertise. Besides the diagnostic aspect, this proposed approach can also be useful for the prognostic task. In particular, using a proper HI, the prediction of the RUL should also be more accurate. This may lead to a more adequate and correct planning of interventions of the maintenance team, therefore, reducing the number of unscheduled intervention tasks and the costs thereto associated.

Thus, future work may encompass the prediction of the RUL, based on the obtained HI. Also, the proposed formulation should be applied to other type of systems in order to evaluate its appropriateness to other aircraft systems and potentially improve the formulation. In this work, the relevancy of the features was assessed through their capacity of detecting faults, nevertheless, other adequate metrics may also be tested and explored in later work. Finally, another relevant aspect, which was not addressed in detail in this work, is the choice of γ value in the α computation. The value of γ is important as it determines the magnitude of the HI added in each flight. Therefore it should be also investigated in future work.

ACKNOWLEDGMENT

This work is partially funded by national funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020. Also, this Paper is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°769288.

REFERENCES

Boškoski, P., Gašperin, M., & Petelin, D. (2012). Bearing fault prognostics based on signal complexity and Gaussian process models. *PHM 2012 - 2012 IEEE Int. Conf.on Prognostics and Health Management*.

Coble, J., & Hines, J. (2009). Identifying optimal prognostic parameters from data: A genetic algorithms approach. *Annual Conference of the Prognostics and Health Management Society, San Diego, CA, September*.

Dalkilic, S. (2017). Improving aircraft safety and reliability by aircraft maintenance technician training. *Engineering Failure Analysis*, *82*, 687 - 694.

Guo, L., Lei, Y., Li, N., Yan, T., & Li, N. (2018). Machinery health indicator construction based on convolutional neural networks considering trend burr. *Neurocomputing*, *292*, 142 - 150.

Hai Qiu, J. L., Jay Lee, & Yu, G. (2003). Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, *17*(3-4), 127–140.

He, W., Miao, Q., Azarian, M., & Pecht, M. (2015). Health monitoring of cooling fan bearings based on wavelet filter. *Mechanical Systems and Signal Processing*, *64-65*, 149–161.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483 - 1510.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, *104*, 799 - 834.

Li, N., Lei, Y., Lin, J., & Ding, S. X. (2015). An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings. *IEEE Transactions on Industrial Electronics*, *62*(12), 7762–7773.

Liao, L. (2014). Discovering prognostic features using genetic programming in remaining useful life prediction. *IEEE Transactions on Industrial Electronics*, *61*(5), 2464–2472.

Mina, J., & Verde, C. (2005, Sep.). Fault detection using dynamic principal component analysis by average estimation. In *2005 2nd international conference on electrical and electronics engineering* (p. 374-377).

Nguyen, T. P. K., Khlaief, A., Medjaher, K., Picot, A., Maussion, P., Tobón, D. A. R., … Cheron, R. (2018). Analysis and comparison of multiple features for fault detection and prognostic in ball bearings. *PHM Society European Conference*, *4*(1), 1–9.

Papakostas, N., Papachatzakis, P., Xanthakis, V., Mourtzis, D., & Chryssolouris, G. (2010). An approach to operational aircraft maintenance planning. *Decision Support Systems*, *48*(4), 604 - 612.

Samaranayake, P., & Kiridena, S. (2012, 10). Aircraft maintenance planning and scheduling: An integrated framework. *Journal of Quality in Maintenance Engineering*, *18*.

Sun, J., Li, C., Liu, C., Gong, Z., & Wang, R. (2018). A data-driven health indicator extraction method for aircraft air conditioning system health monitoring. *Chinese Journal of Aeronautics*, *32*(2), 409–416.

Wang, P., Youn, B. D., & Hu, C. (2012). A generic probabilistic framework for structural health prognostics and uncertainty management. *Mechanical Systems and Signal Processing*, *28*, 622 - 637. (Interdisciplinary and Integration Aspects in Structural Health Monitoring)

Wang, T., Yu, J., Siegel, D., & Lee, J. (2008, Oct). A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In *2008 international conference on prognostics and health management* (p. 1-6).

Wu, H., Liu, Y., Ding, Y., & Liu, J. (2004, 02). Methods to reduce direct maintenance costs for commercial aircraft. *Aircraft Engineering and Aerospace Technology*, *76*, 15-18. doi: 10.1108/00022660410514964

Yan, W., Qiu, H., & Iyer, N. (2008, 01). Feature extraction for bearing prognostics and health management (phm) - a survey (preprint).

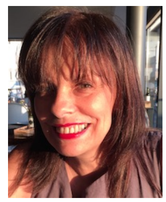
Zhang, Z., Si, X., & Hu, C. (2015). An Age- and State-Dependent Nonlinear Prognostic Model for Degrading Systems. *IEEE Transactions on Reliability*, *64*(4), 1214–1228.

Zhang, Z., Wang, Y., & Wang, K. (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing*, *24*(6), 1213–1227.

BIOGRAPHIES

Daniel Azevedo holds a Master Degree in Informatics Engineering obtained in 2019 at the University of Coimbra, Portugal, specializing in Intelligent Systems. He is a member of Center for Informatics and Systems of the University of Coimbra (CISUC) and is currently involved in the ReMAP H2020 project, as a researcher. His research interests include sensors data analysis, data-based diagnosis of fault condition and remaining useful lifetime prognosis, as well as, other topics in the fields of Machine Learning, Data Science and Intelligent Systems.

Alberto Cardoso is Assistant Professor at the Department of Informatics Engineering of the University of Coimbra, where he teaches Programming, Data Analytics and Intelligent Systems, among other subjects. He is a senior researcher of the Center for Informatics and Systems of the University of Coimbra (CISUC) and his research interests include intelligent systems, fault-tolerant control, wireless sensor and actuator networks, data analysis, online experimentation, and remote and virtual laboratories. He has coordinated and participated in several national and international projects.



Bernardete Ribeiro is Full Professor at the Department of Informatics Engineering at the University of Coimbra, where she teaches Programming, Pattern Recognition, Business Intelligence among other subjects. She is Director of the Center for Informatics and Systems of the University of Coimbra (CISUC). Her research interests are in the areas of Machine Learning, Pattern Recognition, and their applications to a broad range of fields. She has coordinated and participated in several national and international projects.