



**HAL**  
open science

# Data-Centric Helicopter Failure Anticipation: The MGB Oil Pressure Virtual Sensor Case

Nassia Daouayry, Ammar Mechouche, Pierre-Loic Maisonneuve, Vasile-Marian Scuturici, Jean-Marc Petit

► **To cite this version:**

Nassia Daouayry, Ammar Mechouche, Pierre-Loic Maisonneuve, Vasile-Marian Scuturici, Jean-Marc Petit. Data-Centric Helicopter Failure Anticipation: The MGB Oil Pressure Virtual Sensor Case. 2019 IEEE International Conference on Big Data (Big Data), Dec 2019, Los Angeles, United States. pp.1784-1793, 10.1109/BigData47090.2019.9006532 . hal-04118362

**HAL Id: hal-04118362**

**<https://hal.science/hal-04118362>**

Submitted on 6 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339478862>

# Data-Centric Helicopter Failure Anticipation: The MGB Oil Pressure Virtual Sensor Case

Conference Paper · December 2019

DOI: 10.1109/BigData47090.2019.9006532

CITATIONS

3

READS

368

5 authors, including:



**Daouayry Nassia**

Institut National des Sciences Appliquées de Lyon

2 PUBLICATIONS 3 CITATIONS

SEE PROFILE



**Vasile-Marian Scuturici**

Institut National des Sciences Appliquées de Lyon

72 PUBLICATIONS 665 CITATIONS

SEE PROFILE



**Jean-Marc Petit**

Institut National des Sciences Appliquées de Lyon

161 PUBLICATIONS 1,198 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



MAESTRO [View project](#)



Knowledge discovery and privacy-preserving data mining [View project](#)

# Data-Centric Helicopter Failure Anticipation: The MGB Oil Pressure Virtual Sensor Case

Nassia Daouayry  
Airbus Helicopters & INSA Lyon

*Laboratoire d'InfoRmatique en Image et  
Systèmes (UMR 5205CNRS)*

Nassia.Daouayry@airbus.com

Ammar Mechouche,  
Pierre-Loic Maisonneuve  
Airbus Helicopters

*International Airport of Marseille Provence  
Marignane, France*

{Ammar.Mechouche, Pierre-  
Loic.Maisonneuve}@airbus.com

Vasile-Marian Scuturici,  
Jean-Marc Petit

*Laboratoire d'InfoRmatique en Image et  
Systèmes*

*UMR 5205 CNRS, INSA Lyon, France*  
{Marian.Scuturici, Jean-Marc.Petit}@insa-  
lyon.fr

**Abstract** - This paper presents a virtual sensor for helicopter Main Gear Box (MGB) oil pressure. It is developed for anticipating failures of systems related to MGB lubrication.

The virtual sensor is built using Machine Learning and massive in-service flight data collected from Airbus helicopters flying world-wide. The correlation between oil pressure values and other flight parameters is learnt during stable phases of flights in which the system is in its nominal state.

At each flight, the values continuously estimated by the virtual sensor are compared to the measured ones, and an alert is raised when the difference becomes higher than a statistically predefined threshold.

The virtual sensor was tested using normal and abnormal flights, and the results obtained so far in terms of anomaly detection performance are promising.

**Keywords** - *Big Data, Virtual Sensor, Helicopter, HUMS, Predictive maintenance, Machine Learning*

## I. INTRODUCTION

Failure anticipation, and more generally predictive maintenance, is an important and growing topic in the industry [1][2][3][4][5], especially in the aeronautic domain, due to strong requirements in terms of safety and high costs of unscheduled events [6][7][8][9][10]. Failure anticipation is a condition-based maintenance concept and consists in anticipating failures - which are unscheduled events - before they happen. In addition to obvious safety enhancement benefits, anticipating failures is also motivated by economic aspects, which may represent a valuable service for customers. It allows them to: 1) prepare and plan for the required maintenance operation instead of suffering it; 2) order the required part in order to have it in stock on time, right before the maintenance operation; and 3) avoid operational interruptions that imply unexpected financial losses for operators, due to mission's cancellation. In summary, increasing helicopters availability at a lower cost is a key element for operators.

Moreover, early failure detection of a component may prevent from causing damages to its neighbouring parts, and it can also make this component repairable before its complete

failure. For example, failures of some components of the helicopter Main Gear Box (MGB) that are not detected at a very early stage may contaminate other non-replaceable components. This leads to the removal of that MGB which will be definitely lost, at the cost of around 250 K€.

During decades Airbus Helicopters (AH), like its competitors, focused its efforts on anticipating failures mainly for ensuring safety of helicopters [9][11]. This is based on 'classical' Health and Usage Monitoring System (HUMS) [12]. HUMS mostly consists of counters monitoring quantities related to the usage of the machine, and high-frequency vibration measurements acquired from multiple sensors, typically located in the proximity of critical mechanical transmission components and dynamic systems [13]. However, from one hand, embedded algorithms for on-board alarming and usage monitoring are rather simple since they are constrained by embedded hardware processing capabilities. They also only analyse "fresh" data generated during the on-going flight, without any trend monitoring exploiting historical flight data. Consequently, alarms and alerts are raised when monitored sensors, quantities and systems are already faulty (I.e. no longer operational, exceeding a critical threshold limit or reaching an advanced level of wear). They either allow an emergency landing or cancel future flights until the required maintenance is completed. On the other hand, on-ground vibration monitoring is restricted to critical mechanical and dynamic components, whereas plenty of other systems, which are less critical from a safety point of view, are still not monitored. However, it is worth noting that whenever they are gathered to consolidate a global view of a helicopter in operation, they could reveal weak signals very useful to failure anticipation. Thus, there is a need for covering more helicopter systems and improving early detection of failures.

Considering harsh market competition and given the operational costs of helicopters (2/3 of the total cost), helicopters manufacturers are more and more investing beyond safety to generate new revenues. To do so, they develop new types of services to support their customers in reducing their costs. The massive continuous data generated by helicopters during every flight represent an important lever for developing

such services. Analysis of these data might help for improving products, offering new services (e.g. to enlarge condition based maintenance to helicopter usage data) or gaining additional business insights.

In this setting, Airbus Helicopters gathers more than 400000 flight hours (FH) data (called in the sequel in-service flight data) for more than 10 years from hundreds of connected helicopters, operated by several customers worldwide and performing a large variety of missions (Figure 1). These time series data are recorded from the beginning until the end of each flight thanks to the Flight Data Continuous Recorder (FDCR) function in HUMS system. Specifically, hundreds of parameters are recorded from the FDCR, coming from each of the vehicle systems connected to the avionics [13].

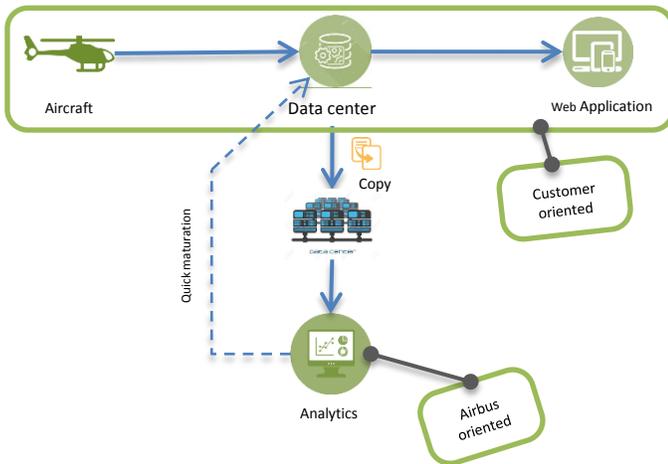


Figure 1: Data collection in the Big Data platform at AH

The overall objective of this paper is to take benefits from a broad variety of collected in-service flight data in order to anticipate and early detect MGB failures, using machine learning techniques. The idea is to build virtual sensors predicting different parameters from other measures as an alternative to physical sensors.

**Problem statement:** Given huge amount of in-service flight data, the objective is to build a virtual sensor for MGB oil pressure at each timestamp of the flight, in order to raise an alert as early as possible when a significant deviation between values measured by the physical sensor and those provided by the virtual sensor occurs.

The choice of MGB is motivated both for its criticality with respect to safety requirements and for economic reasons, i.e. especially in case of unscheduled and non-anticipated removals. Moreover, MGB turns out to be a highly monitored system with many physical sensors producing continuous data at high frequency, helpful for building a virtual sensor.

A supervised machine learning based anomaly detection approach is devised to build this virtual sensor. It is built thanks to learning on multi-varied time series data representing flight parameters. The parameter to be predicted is the MGB oil pressure, for which a physical sensor does exist in every helicopter MGB. To strengthen this data-centric approach, “domain knowledge” of experts is taken into account to

understand the main parameters that should influence the oil pressure of an MGB. Starting from these parameters we have been able to come up with a new relationship, not known from a physical point of view, between the MGB oil pressure and a subset of other indicated flight parameters during a stabilized flight phase<sup>1</sup>. The idea of monitoring MGB oil pressure allows several systems to be covered that are not covered by vibration monitoring, including lubrication system, oil cooling system, oil pump, oil injectors and radiator filter. The point is that any anomaly in the operation of that MGB system should normally be detected by monitoring MGB oil pressure.

The proposed approach presented several challenges. First, correlation and analysis of big amounts of time series data require a very good knowledge of latest big data technologies and an adequate big data platform. At AIRBUS such platforms and tools are available, which facilitates helicopter big data processing and analytics [13]. The second challenge consists in the constitution of a curated training dataset. In the aeronautic domain, failures remain very few compared to the number of flights. Hence, anomaly detection proposed here consists in modelling the normal state of the MGB. Fortunately, most data used for learning the model is composed of “healthy flights”. Moreover maintenance actions exist since they are tagged by customers in their Maintenance Information Systems (MIS). The third challenge lies in the constitution of a complete training dataset covering all flight conditions. As an example, one of the impacting parameters of the MGB oil pressure is the Outside Air Temperature (OAT). Consequently, if a new flight is operated in a different region / continent with an OAT which has never been seen in the past, then the virtual sensor model is likely to diverge. The retained solution consists in learning on the whole fleet data already collected, and monitoring the drift of the model during the execution phase, in order to re-learn it when new and different data is collected.

To sum up, we have made the following contributions:

- A principled approach to address the problem of failure anticipation of an MGB in time series data using state of the art supervised machine learning techniques (Gradient Boosting).
- Integrating domain knowledge as early as possible into the virtual sensor’s construction, to get domain expert as involved and confident as possible in its industrialization process.
- Ongoing deployment of the virtual sensor in the “Flyscan predictive maintenance program” of Airbus Helicopters<sup>2</sup>.

Regarding related work, virtual sensing has been developed in other domains but they do not rely on Machine Learning and big data [14][15][16]. However, similar anomaly detection methods exist, like in [6][7][29], even if they are not presented in the form of virtual sensors. We exploit massive time series data collected from Airbus Helicopters customers, without any reduction as done for example in [9].

<sup>1</sup> In this paper, a stabilized flight phase is one during which the helicopter altitude, the speed and the direction are maintained around specific value.

<sup>2</sup> [https://www.airbus.com/content/dam/corporate-topics/publications/rotor/ROTOR\\_116\\_UK.pdf](https://www.airbus.com/content/dam/corporate-topics/publications/rotor/ROTOR_116_UK.pdf)

The remainder of the paper is organized as follows: first an overview of helicopter data is provided as well as a general view of the problem and the proposed solution. Then, the proposed method is detailed, including validation results and tests on real ‘known’ failures. After that, the deployment in production within FlyScan service, which is the Airbus Helicopters service for predictive maintenance proposed to customers, is described. Finally, the obtained results are discussed and some perspectives are presented.

## II. A PRINCIPLED APPROACH FOR MGB MAINTENANCE

### A. Available data

The data used in this study comes from measures produced by the sensors installed on Airbus’s helicopters. They are recorded continuously during each flight and then transferred, processed and analysed according to the different AH services. The FDCR data collected are recorded at the frequencies of 2 Hz. These time series are gathered on a Big Data platform (Hadoop / Spark).

We focused on a recent aircraft type, for which more than 60 000 FH are already collected. For this helicopter fleet, which encompasses more than 30 different aircraft serial numbers, a dataset has been created with flights data related to the MGB system. The size of the data extracted for the experimentation reported here is around ~180 gigabytes. Table 1 sketches how such data are structured and represented. Each parameter ( $f_0, \dots, f_p$ ) is measured at the same frequency (2 Hz) and the elapsed time is represented as the number of 500 millisecond steps from the beginning of the flight ("time\_step") identified by the “flight id”. Additional information is available concerning for example the operating region, the specific helicopter associated to a flightID etc.

TABLE I. TIME SERIES COMPOSING THE USED DATASET. THE NAME OF ATTRIBUTES IS ANONYMISED

| Flight id | time step | $f_0$      | $f_1$      | $f_2$      | $f_3$      | ...        | $f_p$      |
|-----------|-----------|------------|------------|------------|------------|------------|------------|
| 1         | 1         | $f_0(1,1)$ | $f_1(1,1)$ | $f_2(1,1)$ | $f_3(1,1)$ | $f_4(1,1)$ | $f_p(1,1)$ |
| 1         | 2         | $f_0(1,2)$ | $f_1(1,2)$ | $f_2(1,2)$ | $f_3(1,2)$ | $f_4(1,2)$ | $f_p(1,2)$ |
| ...       | ...       | ...        | ...        | ...        | ...        | ...        | ...        |
| 2         | 1         | $f_0(2,1)$ | $f_1(2,1)$ | $f_2(2,1)$ | $f_3(2,1)$ | $f_4(2,1)$ | $f_p(2,1)$ |
| 2         | 2         | $f_0(2,2)$ | $f_1(2,2)$ | $f_2(2,2)$ | $f_3(2,2)$ | $f_4(2,2)$ | $f_p(2,2)$ |
| 2         | 3         | $f_0(2,3)$ | $f_1(2,3)$ | $f_2(2,3)$ | $f_3(2,3)$ | $f_4(2,3)$ | $f_p(2,3)$ |
| ...       | ...       | ...        | ...        | ...        | ...        | ...        | ...        |
| 3         | 1         | $f_0(3,1)$ | $f_1(3,1)$ | $f_2(3,1)$ | $f_3(3,1)$ | $f_4(3,1)$ | $f_p(3,1)$ |
| ...       | ...       | ...        | ...        | ...        | ...        | ...        | ...        |

### B. Data pre-processing

Since the objective is to build a dataset corresponding to data issued from helicopters operating with a known “normal” state of the MGB.

Significant work was required to address the corresponding issues. Two main approaches are used to reduce the variance in the virtual sensor model. From one hand, abnormal flights are identified in the *data cleaning phase* and, from the other hand, *variations of contextual parameters during a flight* in the contextualization step.

### Data cleaning

Identifying abnormal flight (regarding the MGB system) on the complete historical data is not an easy task. This has been done in collaboration with the MGB system designer. Several workshops have been organized to identify the types of flights that influence the normal operation of the MGB. As a consequence, this process cannot be automated. We started by removing from the dataset flights associated with system failures that were reported in the MIS database. Then, a sub-set of flights preceding each failure was also excluded from the dataset. The number of flights to exclude before each failure tagged in MIS depends on the type of the failure itself.

### Data contextualization

As a matter of fact, the MGB system responds in different ways depending on usage and environmental conditions of the flight. Indeed, the systems such as the MGB are followed differently according to the different phases of the flight such as take-offs, hovers, turns, climbs, descents or landings and therefore react differently. In order to reduce the variance of the virtual sensor model and to maximize the chances of finding a representative model of the normal operation of the MGB system, an intelligent filtering of each time series based on expert knowledge has been applied. Thus, only sub-sequences of each flight have been selected. These sub-sequences, less influenced by the climatic conditions and the usage of a helicopter, correspond to the level flight phase, where the helicopter flies almost at the same speed, same altitude and same direction (*stabilized flight*). Other contextualisation strategies (expert-based), more restrictive on the flight behaviour, have been tested in order to obtain the more accurate model possible. However, they were leading to very small stable sequences of the flight which were not often repeatable over the flights history. For these reasons, it was decided to contextualize the training data relying on the flight sequences corresponding to level flight phase (long and repeatable sequences). This work on contextualization allowed analysing each flight according to the similar context. Consequently, we reduced some parameter variability.

### Data sampling

Furthermore, to perform the splitting of flight history data, a sampling method has been developed to generate the training and a test dataset. The creation of these datasets is one of the most important steps, because it directly impacts the model's performance and robustness. To provide a representative dataset for all flights performed by AH customers (regardless of the geographical operation area, the weather conditions, the season of the year and the type of mission); a methodology of sampling flight data has been set up also for the creation of the test dataset. In this phase, we sampled the complete flight history performed by all AH customers. Consequently, we have carried out this step with a random sampling based on periodicity, geographical and mission type criteria.

### C. Features selection

The features selection of the virtual sensor model has been conducted with the collaboration of experts. Initially, this made it possible to considerably reduce the scope of the analysis and allowed the obvious relationships from the expert's point of

view to be verified through the operational data. The initial motivation for this study was to check the physical laws regarding this system and understand relationships between different contextual parameters.

#### D. Regression with Gradient Boosting

As stated above, the goal is to define a virtual sensor model (regression) from flight data set in normal operation. Then, the algorithm presented here uses this model to detect the moment that the operating system deviates from the normality. The regression was modelled with Gradient Boosting. The Boosting gradient introduced by Friedman [23] is a generalization of Boosting in which a loss function is used in a manner that can be assimilated to a gradient descent. The trees are thus added by reinforcement to the model to minimize the loss function.

The Gradient Boosting regression model has performed well on flight data compared to other usual regression models. An exploratory analysis has been conducted to test the performance of Gradient Boosting in comparison with other regression models. Table 2 represents the results obtained for Random Forest [24], linear regression [26][27][28] and Gradient Boosting on our dataset, using training datasets of different sizes. The two comparison indicators are the coefficient of determination ( $R^2$ ) and the mean squared error (MSE). The ideal values are  $R^2 = 1$  and  $MSE = 0$ . For these two indicators, the Gradient Boosting obtains the highest scores. This is reflected both in the accuracy of the estimate with lowest MSE values regardless of the size of the data sample, and in the prediction quality of the model with the  $R^2$  always the highest.

TABLE II. PERFORMANCES OF DIFFERENT REGRESSION MODELS. FOR EACH TRAINING DATASET SIZE THE EXPERIMENT HAS BEEN REPEATED 10 TIMES WITH DIFFERENT RANDOM DATA RESPECTING OUR CONSTRAINTS AND THE AVERAGE VALUE IS REPORTED IN THE TABLE.

|                   | MSE   | $R^2$ |
|-------------------|-------|-------|
| Gradient Boosting | 0.030 | 0.757 |
| Linear Regression | 0.050 | 0.619 |
| Random Forest     | 0.057 | 0.544 |

#### E. Context influence on model performance

To complete the exploratory analysis, different data samples have been tested in similar ways according to a particular context. The graph shown in the figure below (Figure 2), illustrates the results obtained for different helicopters. Each bar represents the distribution of the  $R^2$  observed on 5 samples analysed from different training and test datasets. We observe in all the different helicopters that the  $R^2$  corresponding to the Gradient Boosting (in green) is higher than the other models. Only two exceptions represented by helicopters 6 and 9 do not validate this observation. These two helicopters obtain low  $R^2$  values for all models. One of the explanations for this anomaly is the change in the geographical area of intervention of these two helicopters. For example, we

have observed changes in operating areas from Asia to Northern Europe.

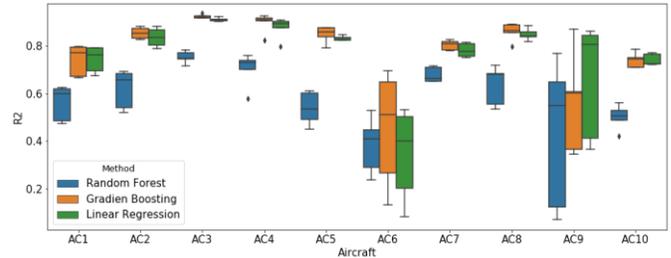


Figure 2: Boxplot represents  $R^2$  calculated for 10 different aircrafts. (Random Forest: blue; Gradient Boosting: orange; Linear Regression: green)

In the same type of analysis, we have compared the Gradient Boosting with the Random Forest as well as the linear regression on different data samples according to the season. We reproduced 5 different samples for winter and summer in the boxplot Figure 3. With respect to other models, Gradients Boosting obtains good results on the  $R^2$  regardless of the season.

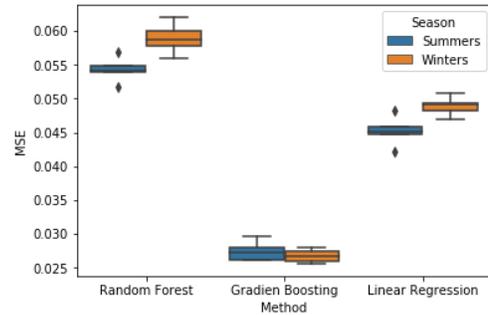


Figure 3: Boxplot represents MSE calculated for summer and winter. (Summer: blue; winter: orange)

#### F. General methodology

The methodology implemented is fairly classical but each step has been developed in depth according to business application.

As illustrated in the Figure 4, the first step is dedicated to the creation of a database for the learning stage of the model and a second for the test step. Gradient Boosting is used on a subset of the data to form the learning stage database. The performance indicators are then evaluated on the test dataset. We use the coefficient of determination and mean squared error to evaluate the best regression model. Once the model performance evaluation step is completed, an anomaly score is calculated from the results obtained with the test dataset.

In order to distinguish flights showing signs of failure from healthy flights, a measurement way has been set up to evaluate the evolution of the results returned by the prediction of the model (virtual sensor) and the values of the real sensor. This measurement, called here "anomaly score", allows us to obtain the distribution of the prediction errors observed within the learning base. For this purpose, we compute the absolute difference between the value predicted by the model and the measured value. The analysis of this distribution is necessary to introduce a threshold that will indicate a divergence between the real sensor and the virtual sensor.

The following figures (Figure 5 and Figure 6) illustrate results obtained with a Gradient Boosting regression for a normal flight and an abnormal flight. In both examples the oil pressure values are represented in blue, the values predicted by the Gradient Boosting in orange and the anomaly score in green. We can easily observe a difference in the distance between the curves representing the virtual sensor and the real sensor depending on the case of a normal or abnormal flight. We have therefore decided to use the anomaly score as a parameter that will indicate whether a flight is normal or abnormal. Thus, when we are in a normal flight situation, the anomaly score is close to zero and when we move away from this value, the flight has a tendency to deviate from normality.

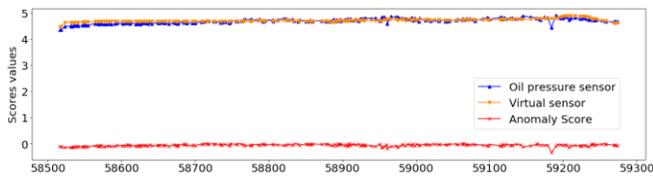


Figure 5: Oil pressure sensor (blue line), virtual sensor (orange line) and anomaly score (red line) for a normal flight

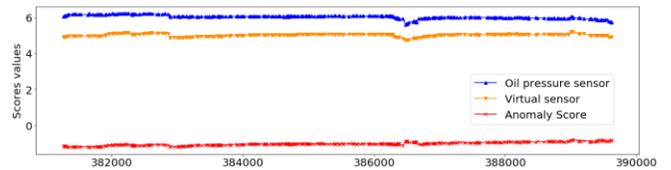


Figure 6: Oil pressure sensor (blue line), virtual sensor (orange line) and anomaly score (red line) for an abnormal

Then, a threshold on this score allows us to detect the flights coming out of the normal operating criteria. In this thresholding step, the "false alarms" question is the main constraint for monitoring tools in the aerospace industry due to the costs they induce.

### III. ANOMALY DETECTION

#### A. Principle

The anomaly detection based on a predictive model consists in identifying atypical observations compared to a reference model. According to [19], an anomaly is an observation that does not conform to the expected standard pattern. The threshold set on anomaly scores calculated from the historical data allows for the model to be validated using samples of real cases regarded as normal.

#### B. Threshold

To set an anomaly detection threshold on the base of the distribution of the anomaly score, the principle of Extreme Values Theory (EVT) was applied. Particularly in recent years, a growing interest has been observed for the application of Extreme Values Theory in the modelling of such events. For a fairly complete presentation of the topic, let us refer to the reference work of Embrechts, Klüppelberg and Mikosch [20] giving the main theoretical results on the EVT and to Reiss and Thomas [21] who proposed a number of practical examples in finance, insurance and environment domains.

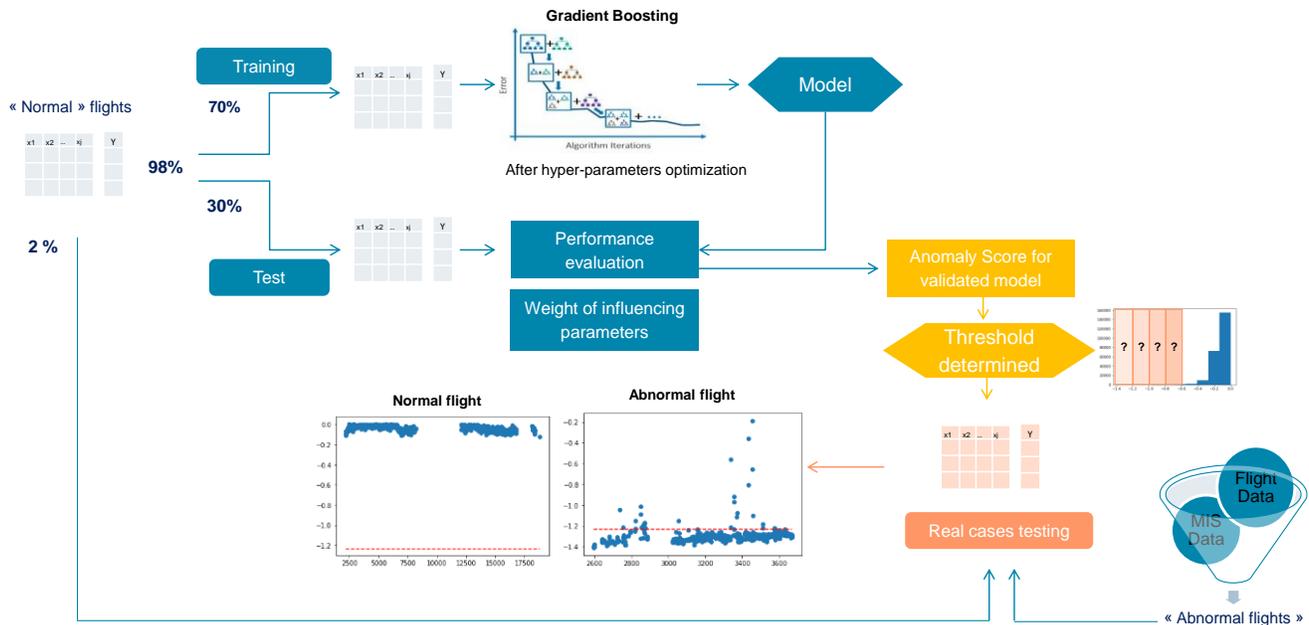


Figure 4: General methodology for the virtual sensor

In this section we are interested in the anomaly score distribution and more precisely the distribution tail of extreme values. The detection method implemented in this article is based on the notion of extreme quantiles. In order to approximate the lowest frequencies representing rare events and therefore abnormal occurrences, we use the classical techniques of EVT. We have used the Generalized Pareto Distribution (GPD) [22][25] to fit the distribution of the frequencies furthest from the graph below (Figure 7). The generalized Pareto distribution is a useful tool in this context because, in addition to accurately reproducing the shape of the empirical distribution above a threshold, the GPD makes it possible to estimate the high quantiles of the distribution. This means that the data can be extrapolated beyond the available sample.

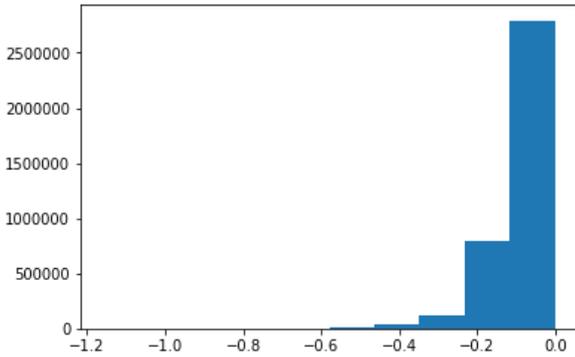


Figure 7: Distribution of the anomaly score calculated on the test dataset. The x-axis represents the values of the absolute difference between the virtual and real sensors and the y-axis corresponds to their frequency.

### C. Experimentations

The use of the virtual sensor to detect abnormal events in the operation of the MGB is essentially based on the hypotheses that the training data represent all normal operating conditions of the helicopter.

The graphical representation below on Figure 8 represents the anomaly score calculated from a set of flight points and the threshold defined to detect the atypical points. The time stamp is represented on the x-axis of the graph and the anomaly score values on the y-axis. The anomaly score is displayed as negative values to be in conformity with the usage of Scikit-Learn python library.

Figure 8 displays the results associated with the anomaly score calculated for each time stamp of the filtered flight sequence on level flight phase. The detection threshold for an abnormal event is set for the complete fleet and represented in red dotted lines. To initialize the anomaly detection step for all helicopters in this study, the threshold was set at an absolute difference of 1.2 bar pressure between the values of the real sensor and the virtual sensor. This threshold has been also validated by the domain expert. Testing with this threshold level is quite "conservative", i.e. it voluntarily detects events more rarely.

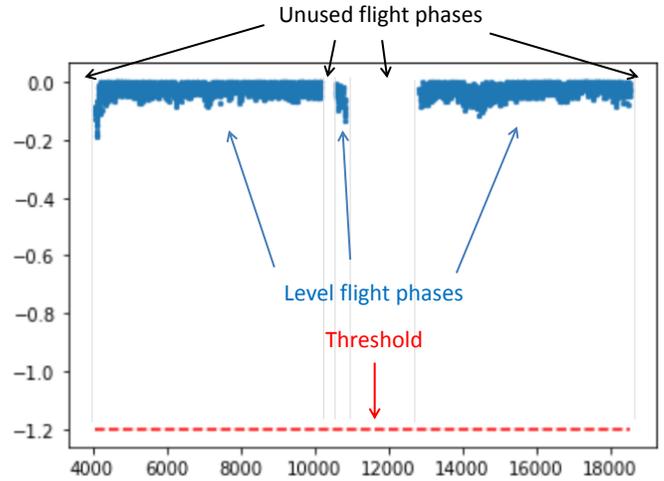


Figure 8: Graphic with the anomaly score and the threshold for a normal flight filtered in level flight phases. The x-axis represents the time-step and the y-axis represents the anomaly score measured in bars (negative value).

On the other hand, the detection model has been evaluated on cases identified as abnormal during the data preparation stage and more precisely, the data cleaning.

Following this data cleaning work, different forms of detections have appeared. For these abnormal cases, we have identified, for example, anomalies with a relatively slow evolution, observed on several consecutive flights before a maintenance operation. We have also identified anomalies impacting only a few flights before the maintenance intervention with a special pattern for the anomaly score.

## IV. MODEL VALIDATION

Data sampling work for the creation of test and training datasets via multiple temporal train-test splits and the application of cross-validation made it possible to check the learning capabilities of the Gradient Boosting model.

In addition, to validate the model's detection performance, different cases of anomalies have been selected by domain experts. These are cases of failures that cannot be detected with vibration analysis. We are particularly interested in anomalies produced by: lubrication system, oil pump and radiator filter. Three flights containing anomalies of this kind are represented in the Figure 9<sup>3</sup>. All these cases are correctly identified by our algorithm.

Several flights considered as normal flights are also proposed by domain experts. All cases are identified as *normal* by our system (Figure 10), except one. After discussion with domain experts, this case seems to have a real anomaly, probably detected by the maintenance team later after the flight.

<sup>3</sup> Please note that for confidentiality reasons, the real values of MGB oil pressure are anonymized.

## V. TESTS WITH MIS DATA

The illustration of an anomaly detection case is shown in Figure 11. The three scatterplots correspond respectively to the anomaly scores calculated on flights before the occurrence of a failure, with the failure and after a corrective maintenance operation. The first figure shows the flight performed by a helicopter shortly before the change of the MGB system. This operation was completed as part of AH's scheduled maintenance plan, which involves changing the system after a number of flight hours performed by the helicopter (TBO: Time Between Overhauls). The second figure shows the first flight operated after the same maintenance operation by the same helicopter. The detection of an abnormal event is obvious, most of the anomaly score points from this flight are below the threshold. The third graphical representation displays the flight performed after the new change of the MGB after an on-board alarm has been triggered.

This helicopter had been flying for one month after the first maintenance operation and had not triggered an on-board alarm for exceeding the critical damage threshold of the system. Although the severity of this failure allowed the helicopter's missions have been performed without incident, the failure propagation was damaging the MGB system and required the removal of the system only one month after its installation. In this example, we were able to detect the anomaly from the first flight operated by the helicopter.

The proposed detection method can thus prevent a malfunction of the monitored system and can prevent its advanced degradation as an expensive system such as the MGB. In this example, threshold detection is easily interpretable and effective. However, the abnormal flights that are detected by this method involving the virtual sensor may have different patterns. Indeed, some cases of detection are not as obvious as the latter example and require analysis to identify the phenomena explaining this variety of patterns.

The two graphs on Figure 12 show other cases of abnormal flight detection which are less obvious than the previous example. Indeed, in both cases, the points representing the anomaly scores calculated on the flights do not all exceed the limit threshold. However, the point distributions in both graphs are very distinctive. For each of these patterns, the following

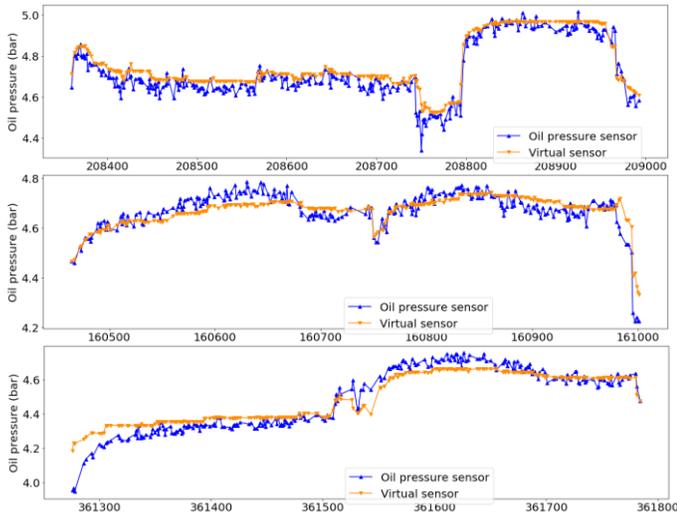


Figure 9: Oil pressure sensor (blue line) and virtual sensor (orange line) for three normal flights

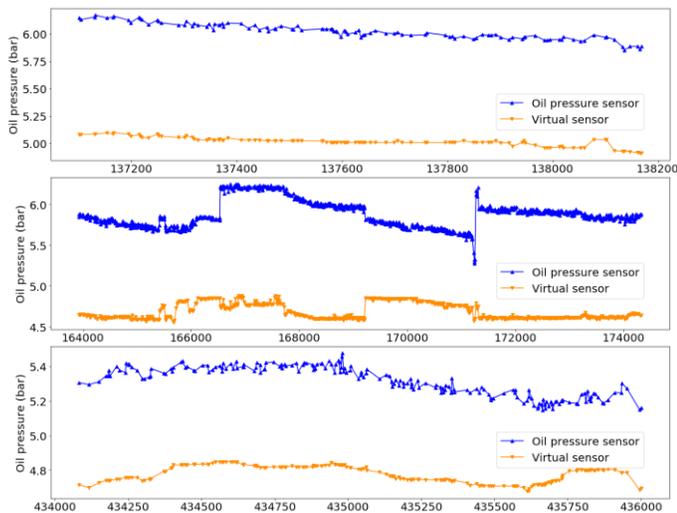


Figure 10: Oil pressure sensor (blue line) and virtual sensor (orange line) for three abnormal flights, corresponding to failures related to lubrication system, oil pump and radiator filter

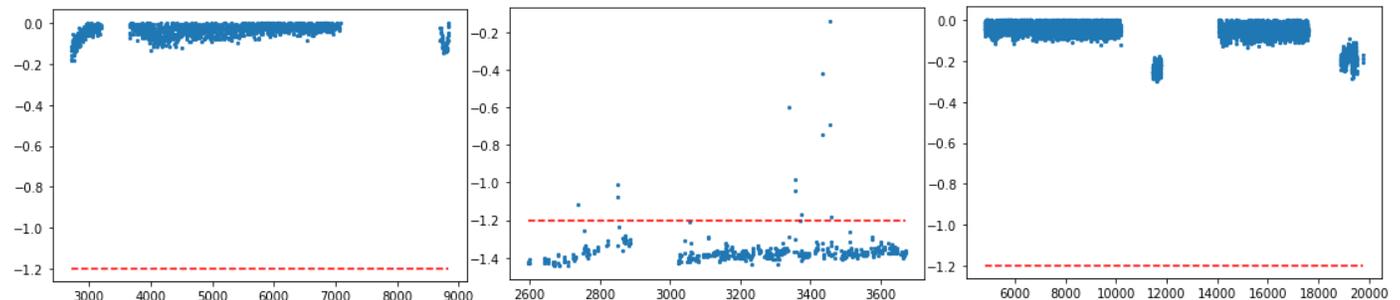


Figure 11: Anomaly detection starting from the first flight operated after the removal of the MGB system. Anomaly Score is represented by the blue points and the threshold by the horizontal red dashed line.

flights have a similar distribution. In some cases, we were able to identify a maintenance operation that allowed us to observe a return to normal system operation. This is the case for the first pattern described by a very noisy distribution of the scatterplot. The second pattern which is observed to follow a trend does not necessarily allow associating this pattern to a type of failure in the MGB system.

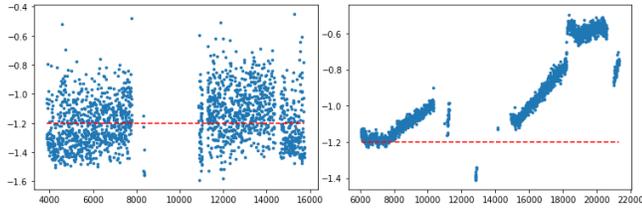


Figure 12: Different abnormal flight pattern. The anomaly score is represented on the y-axis and the time step for each flight on the x-axis.

## VI. DEPLOYMENT WITHIN AIRBUS HELICOPTERS FLYSCAN SERVICE

Since 2007, HUMS / flight data has been collected from various aircrafts during different missions. Since 2012, all data has been automatically treated by HUMS information systems in order to return validated analysis results back to customers via a web portal. For now, valorised results are related to statistics, vibration indicators, KPI, flight spectrum, etc. In addition, HUMS information systems now aims to add predictive maintenance results to their capability. Airbus Helicopters aims to be the first to propose predictive services to customers in order to anticipate issues impacting helicopter availability.

Predictive services consist in monitoring the trend of the results computed by HUMS / Flight data algorithms deployed in the customer oriented data chain. An alerting based on that monitoring is added in order to allow the support team to validate the alert and eventually advise the concerned customer on the maintenance action to perform (Figure 13).

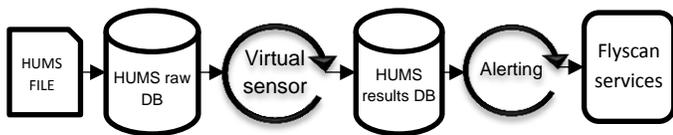


Figure 13: The integration of the virtual sensor in an AH service

### A. Virtual Sensor Integration in HUMS / FlyScan Industrial Process

The virtual sensor presented here is a way to achieve the failure anticipation industrial objective in an aeronautical context. The virtual sensor has been deployed recently in FlyScan. With the predictive model presented, it is possible to detect abnormal usage of the helicopter and predict some failures. However, when an alert is raised no decision will be made automatically. For the moment this step is under the responsibility of a dedicated support team, able to understand the alert, troubleshoot the issue and provide advice to customers in terms of maintenance actions. Nevertheless, on-

going research projects are exploring methodologies to make decisions more rapid and optimal in the future (Figure 14).

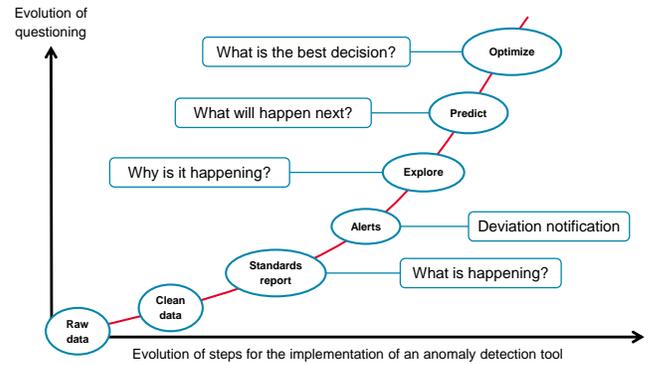


Figure 14: From raw data to autonomous decision making

### B. Operational HUMS Support Room

FlyScan supports operators who have subscribed to the service. Thus, Airbus Helicopters brings its expertise to customers, by monitoring their helicopter fleets based on collected data. This service aims to maximize the customer's operational efficiency through data analytics, enhancing flight safety and increasing fleet availability (eg. decreasing the number of Aircraft On Ground events per aircraft per year).

FlyScan is part of HCare Technical Support<sup>4</sup>, offering 24/7 technical support service, continuous airworthiness management and global technical assistance. An operational control room is able to display all events detected by predictive algorithms (Figure 15). The alerted support team analyses the event and validates the relevance of the detection. Algorithms results combined with experts' knowledge about the monitored systems enable the right decision to be made on whether or not to trigger a maintenance action at the customer end.

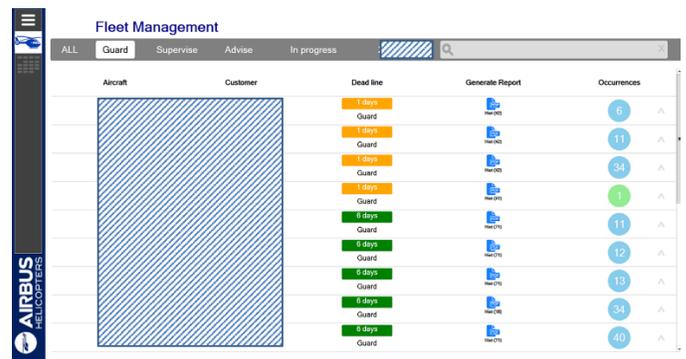


Figure 15: FlyScan Interface

This human step allows a validation of the algorithm output but it is also a means to improve and adjust deployed predictive models. In fact, this allows a labelling of the case as true or false positive. Then, each detected case, rejected or validated by the support team, reinforces the model. Design office

<sup>4</sup> <https://www.airbus.com/helicopters/services/hcare-services.html>

engineers can then re-launch the learning phase and deploy a new version of the model in production.

## VII. DISCUSSION

The aim of the proposed method is to automatically detect malfunctions in the helicopter MGB, by exploiting operational data and business knowledge provided by experts.

The principle of the virtual sensor, in the context of anomaly detection, is based on the use of parameters recorded during flights to model an already existing sensor. The presented MGB oil sensor model, which is based on a Machine Learning approach, allows simulating the oil pressure values from other parameters monitored during flight. Subsequently, an anomaly score is developed, which consists in measuring the difference between the virtual sensor values and the measured ones during flight. Finally, a threshold has been set to allow the detection of abnormal flights.

The method could be improved with respect to several aspects. First, the learning data base cleaning is based only on MIS data, which could be incomplete since it is not mandatory for customers to record/complete it. Another option could be the use of unsupervised machine learning in order to automatically isolate faulty cases. Moreover, data cleaning was a tedious and long task. In fact, MIS data is textual inputs from customers' technicians for describing their maintenance actions on helicopters. It is often difficult for a non expert to understand all technical terms used, especially when some terms are interchangeable. To use the data we needed the help from an expert in order to identify maintenance actions as well as key terms referring to MGB sub-systems' failures. Thus, use of this data was not optimal and not reusable. Moreover, the process was long due to unavailability of the expert. Actually, Airbus Helicopters team responsible of MIS data is developing a natural language and text processing based tool for better use and more automated analysis of the data. This will allow for a better and easier cleaning of our learning database, and consequently results presented here to be improved.

The learning phase is based on the Gradient Boosting algorithm which is recognized and confirmed as a efficient machine learning algorithm compared to classical supervised algorithms. This step, however, could be improved by exploring the latest progress in deep learning. In fact, the presented learning problem could also be modeled using deep learning, based for example on the Long Short Term Memory (LSTM) architecture.

The flight phase actually considered is the one specified by the expert, namely the level flight phase. It could be interesting to use a data analysis approach to identify the best and acceptable flight phase to consider which may improve the results, and where different contextual parameters have less variation.

The influence of the geographical area of operations on the MGB oil pressure requires considering the changes of missions in the learning process. Actually, the solutions considered to deal with these issues would be to integrate helicopters location parameters such as geographical positions among the input parameters or to trigger the relearning of the model when a change in geographical area is identified.

During tests performed on real failures tagged in MIS, we have observed different patterns when a flight is detected as abnormal. Sometimes dispersion is observed (Figure 12), some other times ascending trends are also observed. Actually these different shapes are not analysed in depth, but would be useful in order to study the possibility of enlarging anticipation intervals based on these shapes.

We have chosen to develop the same detection model for the complete fleet of a helicopter type by integrating the identifier of the aircraft into the input variables. This choice is justifiable because of the hypothesis that the systems are designed identically and therefore are supposed to respond in the same way depending on the use of the helicopter. However, operations may differ from one customer to another and the model may be less effective for less standard helicopter use. This problem could be solved by duplicating the detection model for each helicopter.

Finally, features (contextual parameters) considered in the implemented model were selected according to the expert knowledge. This step could be automated using recent works about features selection that maximize the performance of a Gradient Boosting regression [23]. Further experimentation of automatic features selection would be likely to identify additional parameters in order to model the oil pressure sensor more accurately. Moreover, this would also allow for the generalization of this method to other sensors and other helicopter systems.

## VIII. CONCLUSION

This article presents the end-to-end solution for detecting anomalies in industrial preventive maintenance. It addresses the different steps and difficulties in the processing of huge amounts of data for the development of a failure anticipation method. The initial objective of defining the normal operation of a helicopter system was achieved through the cooperation of aeronautical experts. This expert knowledge has been used to exploit HUMS flight data with machine learning methods to model a form of normal operation of the MGB system through the virtual oil pressure sensor. This is a pragmatic solution that allows AH customer support to improve the maintenance service offered to customers.

We are currently working on the drift monitoring of the deployed virtual sensor, in order to trigger its update when new and different data is collected.

**Acknowledgments:** We would like to thank our colleagues Jayant SEN-GUPTA, Olivier REGNIER-COUDERT and Nolwen HUET for their support, and the interesting technical exchanges about this work; done in the framework of the AIRBUS ADVISED project about anomaly detection in time series data.

## REFERENCES

- [1] W. O. L. Vianna, T. Yoneyama: Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. *IEEE Systems Journal* 12(2): 1170-1181, 2018.
- [2] M. Fernandes, A. Canito, V. Bolón-Canedo, L. Conceição, I. Praça, G. Marreiros: Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry. *Int J. Information Management* 46: 252-262, 2019.
- [3] V. Rousopoulou, A. Nizamis, L. Giugliano, P. Haigh, L. Martins, D. Ioannidis, D. Tzovaras: Data Analytics Towards Predictive Maintenance for Industrial Ovens - A Case Study Based on Data Analysis of Various Sensors Data. *CAiSE Workshops*: 83-94, 2019.
- [4] A. N. Jalali, A. Schindler, B. Haslhofer: Understandable Deep Neural Networks for Predictive Maintenance in the Manufacturing Industry. *ERCIM News* 2019(116), 2019.
- [5] F. Giobergia, E. Baralis, M. Camuglia, T. Cerquitelli, M. Mellia, A. Neri, D. Tricarico, A. Tuninetti: Mining Sensor Data for Predictive Maintenance in the Automotive Industry. *DSAA*: 351-360, 2018.
- [6] S. Alestra, C. Bordry, C. Brand, E. Burnaev, P. Erofeev, A. Papanov and C. Silveira-Freixo, "Rare Event Anticipation and Degradation Trending for Aircraft Predictive Maintenance", 11th World Congress on Computational Mechanics, Barcelona, 2014.
- [7] R. Sundareswara, F. D. Betz, and T.-C. Lu, "Interpretable Unsupervised Feature Extraction and Learning of Abnormal System State Transitions in Aircraft Sensor Data", *PHM\_CONF*, 10(1), 2018.
- [8] D. Wade, A. Wilson, A. Reddy, and R. Bharadwaj, "Validating Machine-learned Diagnostic Classifiers in Safety Critical Applications with Imbalanced Populations", *PHM\_CONF*, vol. 10(1), 2018.
- [9] V. Camerini, G. Coppotelli, S Bendisch, "Fault detection in operating helicopter drivetrain components based on support vector data description", *Aerospace Science and Technology* 73, 48-60, 2018.
- [10] R. Sundareswara, F. D. Betz, and T.-C. Lu, "Spectrum Reconstruction for Oil Cooler On Condition Monitoring", *Vertical Flight Society 75th Forum*, USA, 2019.
- [11] CAA (U.K. Civil Aviation Authority), CAA Paper: The application of advanced anomaly detection to tail rotor HUMS data: based on a report prepared for the CAA by GE Aviation Systems Limited () Spiral-bound, 2013.
- [12] U.K. Civil Aviation Authority, CAA Paper 85012: Report of the CAA Working Group on Health Monitoring, 1985.
- [13] A. Mechouche, N. Daouayry and V. Camerini, "Helicopter Big Data Processing and Predictive Analytics: Feedback & Perspectives", *European Rotorcraft Forum (ERF)*, paper n° 35, Poland, 2019.
- [14] P. Kadlec, B. Gabrys: Soft sensors: Where are we and what are the current and future challenges? *International Conference on Intelligent Control and Automation Science*, 572-579, 2009.
- [15] J. Kullaa, Development of virtual sensors to increase the sensitivity to damage. *Procedia Engineering*, 199, 1937-1942, 2017.
- [16] S. Kabadayi, A. Pridgen and C. Julien, "Virtual sensors: abstracting data from physical sensors", *International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06)*, Buffalo-Niagara Falls, NY, pp. 6 pp.-592, 2006.
- [17] N. Daouayry, P-L. Maisonneuve, A. Mechouche, J-M. Petit, and V-M. Scuturici, "Predictive maintenance for helicopter from usage data : application to main gear box", *ERF*, Delft, The Netherlands, , p. 6, 2018.
- [18] I. Arous, M. Khayati, Ph. Cudré-Mauroux, Y. Zhang, M-L. Kersten, S. Stalinov: *RecovDB: Accurate and Efficient Missing Blocks Recovery for Large Time Series*. *ICDE 2019*: 1976-1979.
- [19] V. Chandola, A. Banerjee, and V. Kumar. *Anomaly detection : a survey*. *ACM Computing Surveys*, 41(3), 2009.
- [20] P. Embrechts, C. Kluppelberg, T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Berlin, Springer, 1997.
- [21] R. Reiss, M. Thomas, *Statistical Analysis of Extreme Value with Applications to Assurance, Finance, Hydrology and Other Fields*, Basel, Birkhauser, Verlag, 2001.
- [22] G.R. Dargahi-Noubary, "On tail estimation: An improved method", *Mathematical Geology*. 21 (8): 829-842, 1989.
- [23] H.Friedman, *Greedy Function Approximation A Gradient Boosting Machine*, Feb 24, 1999.
- [24] L. Breiman, *Random forests*, *Machine Learning* 45, 5-32, 2001.
- [25] S. Coles, "An Introduction to Statistical Modeling of Extreme Values", Springer. p. 75, 2001.
- [26] DG. Kleinbaum, LL. Kupper, KE. Muller, A. Nizam, "Applied Regression Analysis and Multivariable Methods", 3rd ed. Boston, Mass: Duxbury Press, 1997.
- [27] SA. Glantz, BK. Slinker, *Primer of Applied Regression and Analysis of Variance*. 2nd ed. New York, NY: McGraw-Hill, 2000.
- [28] MH. Kutner, CJ. Nachtsheim, J. Neter, W. Li, "Applied Linear Statistical Models", 5th ed. New York, NY: McGraw-Hill, 2004.
- [29] M. De Benedetti, F. Leonardi, F. Messina, C. Santoro, A. Vasilakos, "Anomaly detection and predictive maintenance for photovoltaic systems", *Neurocomputing* 310 (59-68), 2018.