Ref. Ares(2019)7068845 - 15/11/2019

ReMAP

Real-time Condition-based Maintenance for Adaptive Aircraft Maintenance Planning

Deliverable D5.1

Report on exploratory data analytics

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769288

Document History

Revision Nr	Description	Author	Review	Date
1	Version 1	UTRC-I	NA	15 Sept 2019
2	Version 2	UTRC-I, UC	NA	30 Sept 2019
3	Version 3	UTRC-I, Embraer, UC, Embraer, ONERA, KLM	NA	3 Oct 2019
4	Version 4	UTRC-I, Embraer, UC, Embraer, ONERA, KLM	NA	4 Oct 2019
5	Version 5	UTRC-I	NA	23 Oct 2019
6	Version 6	UTRC-I, Embraer, UC, Embraer, ONERA, KLM	TUD, KLM	27 Oct 2019
7	Version 7	UTRC-I, Embraer, UC, Embraer, ONERA., KLM		04 Nov 2019
8	Version 8	TU Delft (Coordinator Input)		08 Nov 2019
9	Version 9 (Final)	UTRC-I		11 Nov 2019

Index

1. Introduction

1.1. Project Summary

ReMAP "Real-time Condition-based Maintenance for adaptive Aircraft Maintenance Planning" (hereinafter also referred as "ReMAP" or "the project"), is a European project started on the 1st of June 2018 and has a duration of four years. The project addresses the specific challenge to take a step forward into the adoption of Condition-Based Maintenance in the aviation sector. In order to achieve this, a data-driven approach will be implemented, based on hybrid machine learning & physics-based algorithms for systems, and data-driven probabilistic algorithms for systems and structures. A similar approach will be followed to develop a maintenance management optimisation solution, capable of adapting to real-time health conditions of the aircraft fleet. These algorithms will run on an open-source IT platform, for adaptive fleet maintenance management. The proposed Condition-Based Maintenance solution will be evaluated according to a safety risk assessment, ensuring its reliable implementation and promoting an informed discussion on regulatory challenges and concrete actions towards the certification of Condition-Based Maintenance.

1.2. Purpose of this Document

This document is the Deliverable D5.1 report of the ReMAP Project. The deliverable is part of Work Package 5 and related to the Task 5.1 on "exploratory data analytics and specific scope & requirement definition". The main objectives of this task, which are described in detail in this document are to: (a) perform a detailed literature survey to identify the scope and requirements of the prognostics and health management (PHM) algorithms, which are to be developed in the ReMAP project, (b) perform exploratory analysis of the available data to understand the characteristics and (pre-) processing requirements, (c) perform exploratory experiments with the different PHM algorithms and identify the scope and requirements for value addition to the state of the art, (d) generate initial PHM algorithm concepts, (e) understand and implement pre-processing requirements in the context privacy and security for sharing data by the airlines (e.g., KLM) with the ReMAP consortium.

1.3. Context

As shown in Figure 1, presently most aircraft systems either have preventive or reactive maintenance policies. In preventive maintenance, maintenance is pre-scheduled as per usage specifications, which potentially lead to over-maintenance and overuse of time & resources. In reactive paradigm, a maintenance action is taken when the aircraft or its crew detect an internal fault in an aircraft system/component, indicating a potential (partial) loss in functionality. The reactive maintenance potentially may result in high cost, unexpected downtime and inefficient use of labor in case of a failure.

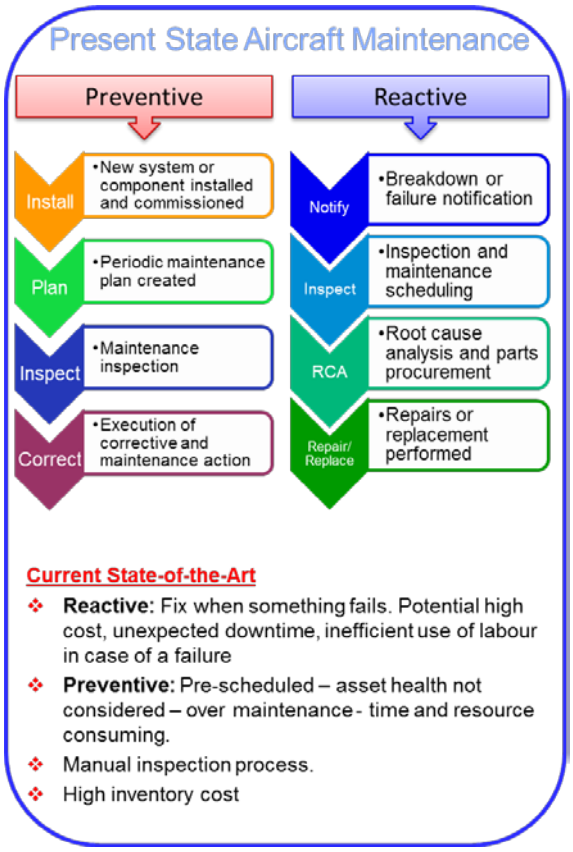

Figure 1 State of the Art in Aircraft Maintenance

In the ReMAP project, a novel predictive condition-based aircraft maintenance system is being developed which is envisaged to overcome the shortcomings of the aforementioned reactive and predictive paradigms. The complete envisioned tool chain is shown in Figure 2, where we see that in Work Package 5 the core analytics methodologies for system and component level diagnostics, prognostics and health management (PHM) technologies are planned to be developed. The diagnostics & PHM analytics will use data from different sources (sensors in aircraft, external data like weather, pollution etc.) to predict and detect degradation in the aircraft systems and components while estimating the remaining useful life and quantifying health. The results of these methodologies are fed into the next processing block as shown in Figure 2 for using in a predictive condition-based maintenance system.

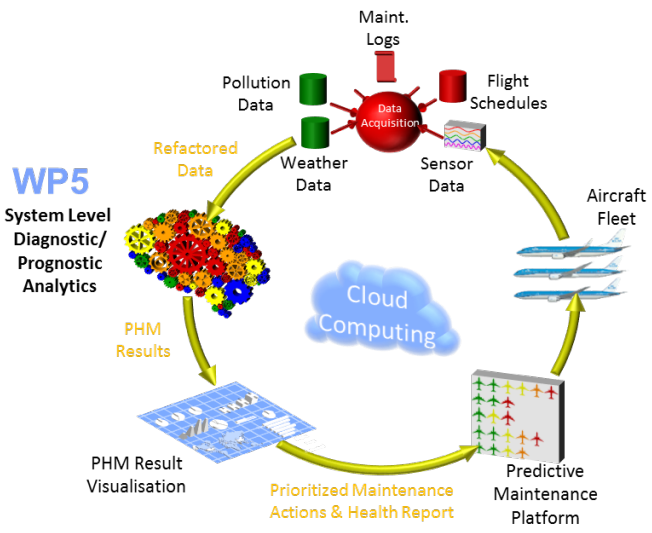

Figure 2 ReMAP Condition Based Maintenance Tool

For development of the diagnostics and PHM algorithms in WP5, in Task 5.1, detailed exploratory analysis is performed with the following objectives: (a) understanding and characterization of the different data sets, (b) identify pre-processing requirement, (c) define requirements for diagnostics and PHM algorithms through literature survey, (d) feature analysis and data requirement for diagnostics and PHM algorithms.

As shown in Figure 2, data from many different sources are expected to be used in the process flow: (1) sensor data from aircraft, (2) external weather and pollution data, from public domain sources like NASA (https://earthdata.nasa.gov/), (3) maintenance logs, (4) flight schedules. Given the diverse nature of the problem, the work was split amongst among partners which comprised of the following main tasks:

a. Literature Survey on PHM methodologies for failure prediction, health quantification and RUL estimation using time series data from aircraft and specific scope & requirements identification (Partners: UTRC-I, UC, ONERA).
b. A framework for efficiently acquiring of external data for hundreds of thousands (and possibly scaling up to millions) flights for external failure driver analysis (Partner: UTRC-I).
c. Development of framework to share KLM aircraft data to be used in WP5 (Partners: KLM, TU Delft).
d. Exploratory analysis for Embraer E-jet system selection (Partners: Embraer, KLM).
e. Automated analytics methodologies for maintenance logs and messages for 747 Bleed Air systems (Partners: Embraer).
f. Exploratory analysis on edge computing for health estimation (Partners: UC).
g. Exploratory analytics for Health Indicator computation for 747 bleed air and 787 brake system (Partners: UC, ONERA),

Before describing the details and the outcomes of the aforementioned tasks, we would like to emphasize that, as a Research and Innovation Action (RIA) project, a significant amount of focus has been on novelty and innovation. The results of the research carried out in WP5 during the past 12 months period resulted in: *(1) one master's **thesis**, (b) a **journal paper** accepted for publication in International Journal of Prognostics and Health Management (Shahid & Ghosh, 2019) and invited for presentation at "annual*

conference of the prognostics and health management society 2019" at Scottsdale, Arizona, 23rd Sept to 27th Sept 2019. Due to the quality of the work, the lead author of the paper has been invited to deliver a session closing keynote talk in the PHM2019 conference. (c) a ***journal paper*** *being submitted to "Aerospace" journal published by MDPI (Multidisciplinary Digital Publishing Institute)* (Basora, Olive & Dubot, 2019).

2. Literature Survey on PHM Methodologies for Failure Prediction, Health Quantification and RUL estimation using Time Series Data from Aircraft and specific scope & requirements identification

2.1. Literature Survey

Prognostics for complex systems and components, such as those in an aircraft, have attracted significant interest of industrial and academic research community in the last few years. In the industrial field, a widely accepted definition of 'prognostics' is the ability to predict the Remaining Useful Life (RUL) of a component after a fault has occurred (Aizpurua & Catterson, 2015). Here fault refers to degradation and failure refers to total loss of functionality. While significant importance has been given in the academic research to improve the RUL estimates with complex algorithms, little effort is being done to explain these algorithms in terms of how or why they achieve a certain level of performance. This is in contrast to the future industrial requirement of explainable safety critical PHM system (Elattar, Elminir, & Riad, 2016). Satisfying the aforementioned definition of prognostics, most PHM literature focuses on estimating RUL via physics-based or data-driven approaches. In *Figure 3*, a high-level description of the state of the art of the in PHM research is presented.

Data-driven approaches perform RUL prediction from the operational run-to-failure raw time series data, collected from the sensors mounted on (or near) the component or system under consideration. There are predominantly two types of data-driven approaches in the literature; direct and indirect. The direct approach relies on training of an analytics model to learn the RUL directly from the run-to-failure time series data. Many methods which use Bidirectional Long-Short Term Memory Networks (Bi-LSTM) (A. Zhang et al., 2018), (J. Wang, Wen, Yang, & Liu, 2018), a combination of Convolutional Neural Network (CNN) and LSTM (Li, Li, & He, 2019) or a Deep Belief Network (DBN) (C. Zhang, Pin, K. Qin, & Chen Tan, 2016) have been proposed in this context. These methods have achieved promising performance but are uninterpretable due to the direct mapping of the time series to a life estimate. Furthermore, many of the aforementioned approaches use a piecewise linear function of RUL for each time series to train the model. It is not possible to define this function for a broad range of systems and application scenarios if the underlying characteristics are highly non-linear.

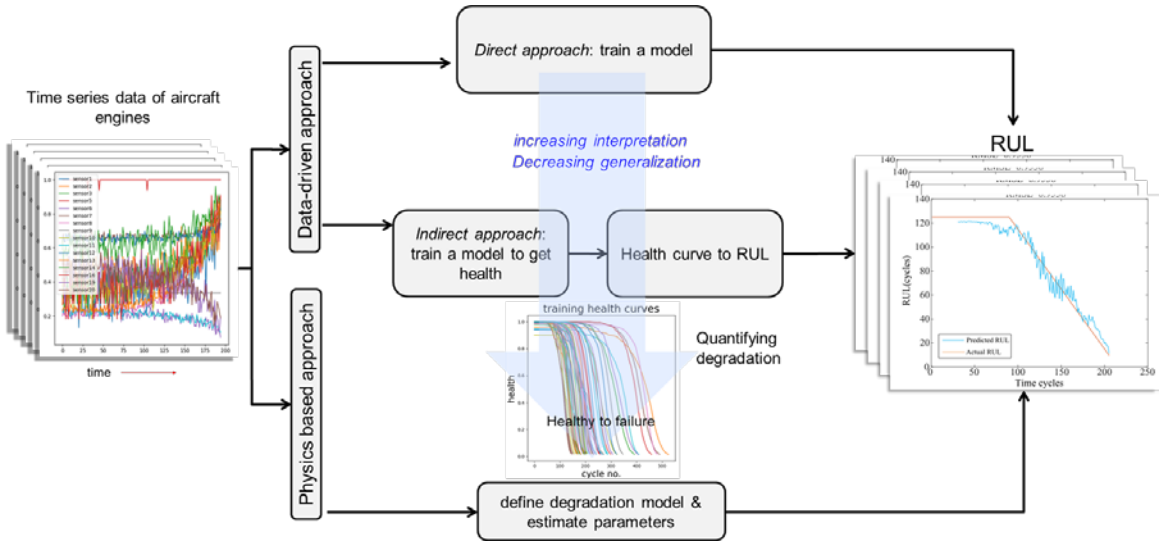

Figure 3 Overview of the state of the art in prognostics and health management research based on the literature survey performed in Task 5.1 of Work Package 5 of the ReMAP project

The indirect approach first maps the time series data into a one-dimensional health index HI or health curve (ranging from 1 to 0), which decreases monotonically and proportionally to the time series degradation (Mosallam, Medjaher, & Zerhouni, 2016), (Ramasso, 2014), (Mosallam, Medjaher, & Zerhouni, 2015). This task is also known as health monitoring. Similar to the direct approaches, deep learning methods such as Restricted Boltzmann Machines, DBNs and CNNs have been used very frequently (Zhao et al., 2019), (Akintayo, Lore, Sarkar, & Sarkar, 2016), (Reddy, Venugopalan, & Giering, 2016) for health monitoring purpose. More recently RNN based approaches have become even more popular (Zhao, Wang, Yan, & Mao, 2016), (Zhao, Yan,Wang, & Mao, 2017), (Wu, Yuan, Dong, Lin, & Liu, 2018), (Ellefsen, Bjørlykhaug, Æsøy, Ushakov, & Zhang, 2019) for monitoring health degradation. In particular, (Malhotra et al., 2016) and (Gugulothu et al., 2017) propose to learn health index from the time series reconstruction error computed with an autoencoder or distance from healthy embedding in low dimensional space. Once the health index is computed, RUL is estimated as a weighted average of RULs of matching HI curves (T. Wang, 2010) of all the time series in the training dataset.

We are interested in the indirect approaches, which estimate RUL via health monitoring. Such approaches are easy to interpret and do not require the specification of piecewise linear function of RUL or other thresholds which cannot be estimated without a-priori knowledge on the data domain.

In addition to prognostics and RUL estimation, PHM encompasses fault diagnostics, which includes fault detection, isolation (i.e. which component has failed), failure mode identification (i.e. what is the cause of the fault) and quantification of the failure severity. Fault detection is typically based on the quantification of the inconsistencies between the actual and the expected behaviour of the system in nominal conditions. In this context, effective anomaly detection techniques to predict incipient failures from historical data are important to estimate time-to-failure and help schedule maintenance activities. Although anomaly detection is an active field of research, a recent survey on recent techniques in this domain (Basora, Olive & Dubot, 2019) points out a limited number of applications of anomaly detection methods in aviation with the goal to improve PHM and predictive maintenance processes.

2.2. Gaps in the state of the art PHM approaches

Although the one-dimensional health curve is very commonly used for health degradation monitoring and RUL estimation, it provides very limited information about the complex dynamics of the degradation, such as the connection between different failure modes. Therefore, the health curve can only be used to quantify the component degradation without providing any further insight into the complex failure dynamics.

2.3. Scopes and Requirement for Innovation

Given this detailed literature survey, some the specifics *scope and requirements* for PHM methodologies to be developed in WP5 were identified.

Due to the mere quantitative nature of the health curve and as some models, like deep neural networks, are capable of extracting information richer than just a one-dimensional health curve, we believe that there is a need for an intermediate block in the health analysis and RUL computation pipeline.

This block should be capable of summarizing the neural network model information in a low-dimensional space, which can be leveraged for maintenance decision making. Motivated by the fact that the notion of explainability in prognostics is strongly connected to visualization (Phillips, 2012), we propose this intermediate block to be a visualization of degradation of raw time series from healthy to failure classes, in the form of a smooth trajectory in a 2D space. The schematics of the envisaged innovation is shown in Figure 4. Following are the main technical contributions of the methodology conceived through the detailed literature survey, exploratory analysis and brainstorming:

1. A set of novel RNN autoencoder architectures, called *TrajecNets*, which embed the operational run-to-failure time series data in a 2D metric space in the form of a smooth trajectory, capturing the evolution of system from healthy to failure state.
2. A novel and interpretable health index computation method from 2D trajectories, which is a fusion of “local-health”, i.e., the health of the system (e.g., engines) with respect to its nascence and “global-health” - health with respect to other similar systems.
3. An unsupervised methodology to compute failure probability using the aforementioned health index.
4. A fusion of local and global health curve matching strategies for RUL estimation by using a more informative distance metric.

The framework shown in the Figure 4 is developed and validated with NASA PHM turbofan (https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/) engine dataset as part of the Task 5.2. The results beat the state of the art especially for remailing useful life estimation. The next steps are to further improve and validate the methodology with aircraft operational sensor and actuation data. The details of the methodology are beyond the scope of this deliverable report and will be included in D5.2.

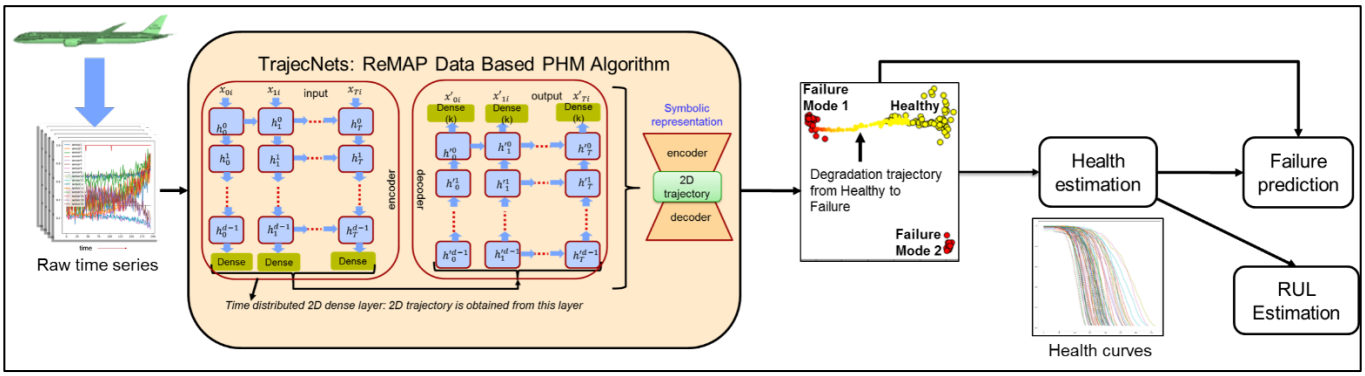

Figure 4 Proposed Generic Data Based PHM Methodology Tool Chain

2.4. Physics based Fault Isolation and Root Cause Analysis

Further exploratory analysis on the capabilities of the framework shown in Figure 4 was performed. The methodology produces results in visualization of degradation, health estimation, failure prediction and remaining useful life. In addition, the requirement for fault isolation and root cause analysis (RCA) were investigated and analysed. Through a detailed brainstorming an initial framework was developed based on physics-based analysis. The schematic of the framework is shown in Figure 5. Following this initial concept, in Task 5.2 the methodology will be improved, implemented and validated.

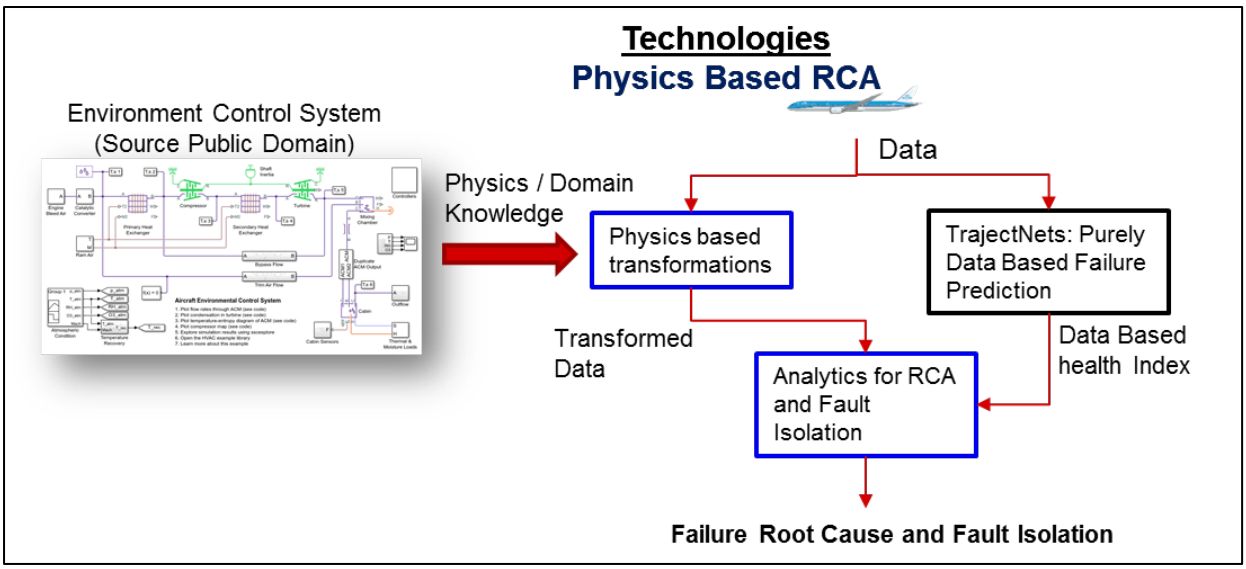

Figure 5 Physics Based Fault Isolation and Root Cause Analysis

It should be noted that the combination of methodologies described in Figure 4 and Figure 5 leads to the hybrid approach envisaged in ReMAP proposal.

2.5. Feature Analysis:

After a detailed study of the possible feature extraction methodologies it was decided, that at the initial stage the use of deep learning methodology in “TrajecNets” would not require any explicit feature extraction. However, features will be extracted as motif in the time series data for different failure modes from the outputs of “TrajecNets”, which will be fed into Physics Based RCA block. The types of data to be used for the PHM algorithm development are shown in Figure 2.

2.6. Data Pre-processing Requirement

At this point normalization with max is performed as data pre-processing and promising results are achieved. It should be noted that, in the present PHM algorithm development it is assumed that some run to failure data are available to train the models.

3. A framework for efficient acquiring of external data for hundreds of thousands (and possibly scaling up to millions) flights for external failure driver analysis

Aircraft operate in various types of routes that can present varying weather conditions, levels of pollution and hours of flight per flight-cycle. When exposed to environmental conditions with high concentration of pollutants or extreme weather, aircraft systems may develop accelerated degradation compared to aircraft that don't get exposed to such conditions.

There is a need for an automated methodology identifying the specific conditions, in terms of weather and pollutants, which lead to an accelerated deterioration of aircraft systems. A schematic description of the envisaged methodology is shown in Figure 6. A necessary requirement for this analysis is to acquire and process weather and pollution data corresponding to the airports (take-off and landing) for to large number of flights.

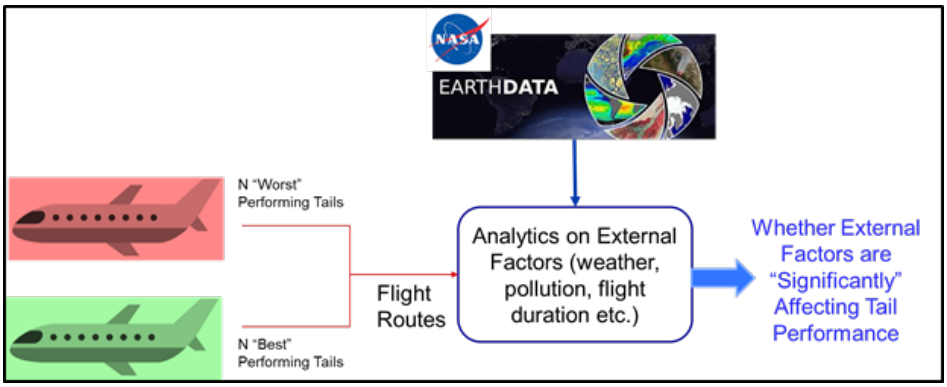

Figure 6 External Failure Driver Analysis

An efficient and robust computational framework was developed to achieve this data acquisition and processing requirement. The framework was developed in Python programming language using the APIs provided by the public domain websites which contains weather and pollution data.

The weather and pollution data are publicly available from the NASA Earth data website (https://earthdata.nasa.gov/). NASA promotes full and open sharing of all its data to research and applications communities, private industry, academia, and the general public. The Earth Observing System Data and Information System (EOSDIS) provides various ways to discover, access and use the data.

Furthermore, METAR data (https://www.aviationweather.gov/metar) can also be used, which gives us temperature and dew point at the airport. With that, humidity can be calculated.

Temporal and geographical information are extracted from the flight routes: for each tail, the location (latitude and longitude) of each airport and time where the tail landed, are collected and stored in a database. With this temporal and geographical information for hundreds of thousands of flights, a query is performed to the cloud-based NASA database on weather conditions (e.g. temperature,

air density, sea salt etc.) and pollutants (e.g. ozone, carbon monoxide, dust, sulphate etc.) to retrieve their value for the given coordinates and time. A parallelization approach is preferred due to the significant number of requests to be made to the database.

The proposed framework aims at identifying the environmental factors leading to an increased maintenance operation by comparing a calculated health metric for each aircraft tail, knowing the conditions to which each tail was exposed to.

The flight data required include the time and location where each tail landed for a certain time period. Additional information can be derived such as the number of flight hours and flight cycles. The flight data also lists the number of parts that failed or were removed for servicing or replacement during the same time period. A score, which is a representative health metric of the tail, is computed based on the number of removals normalized by the flight hours or flight cycles, for instance.

Initial analysis using the framework revealed intriguing insight into impact of external conditions affecting degradation and failure patterns. The details of the results are beyond the scope of this initial deliverable.

4. Development of framework to share aircraft data to be used in WP5

This task mainly focused on identification of various input and labelling of data to be used in WP5, including an assessment on data quality and usability for the work package.

Sharing data generated by the aircraft during flights is non-trivial mainly because of the following: (a) commercial sensitivity, (b) potential privacy violations, (c) regulatory & company compliance procedures. A significant amount pre-processing is required before KLM data can be shared with the consortium. Furthermore, approvals were required from the airline management for the data to be shared with the ReMAP consortium.

The KLM ReMAP team have been granted approval to share CPL data of selected systems of the Boeing 787 and 747 fleet, under the following conditions:

1. Parameter names (headers) are anonymized
2. City pairs are anonymized
3. Tail and Flight codes are anonymized
4. Date and time stamps are offset by a random number of minutes

Similarly, to sensor-data, an anonymization and approval process is required to share maintenance data as well. KLM is currently working with KLC to get similar approval to share sensor data of the Embraer fleet.

4.1. IT Infrastructure

KLM has established a secure connection to the sFTP server of TU Delft. With this connection, it is possible to share anonymized datasets of aircraft sensor data, which can be used for developing and validating prognostic and diagnostic models. Currently, data will be transferred as a one-off action, initiated by KLM. However, the infrastructure is also suitable for a continuous (e.g. daily) stream of data. The schematics of the data sharing infrastructure and methodology is shown in Figure 7. Please note that this infrastructure is of temporary nature. Once Work Package 2 has implemented a stable version of the IT Platform, data will be shared through that vehicle.

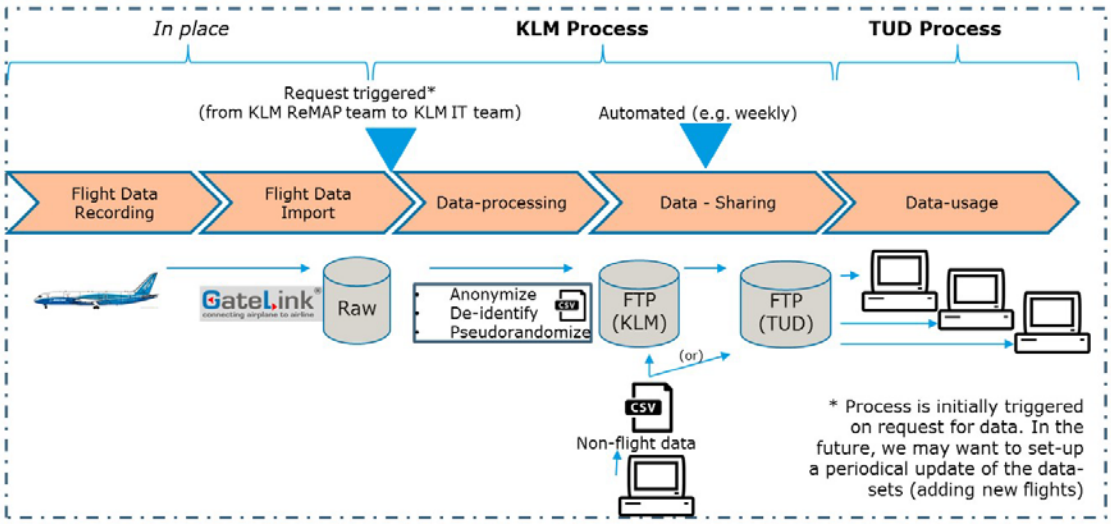

Figure 7 Data Sharing Framework for Aircraft Operational Data

4.2. Data sets

The following datasets are shared with the consortium.

The first set concerns the Boeing 787 brake system, and consisted of the following subset of data:

Input data: Sensor data
Failure data: Removal data

In addition to removal data, also maintenance logs were studied, but it was concluded that removal data was currently better suited for the task of prognostic modelling, due to its better structure.

The second dataset concerns the Boeing 747 Bleed Air System, and consisted of the following subset of data:

Input data: Sensor data
Failure data: Flight Deck Effect (FDE) data, removal data, and maintenance logs.

This data-set is substantially larger than the first set. Also, the sensors are not as easily linked to the degradation of the system, hence it asks for better physical understanding of the system, and/or for more advanced data analysis.

Finally, data from the 787 CACTCS system was also shared with ReMAP partners.

Preliminary data-analysis showed that there are various data-sources that can be used for labelling system-conditions. For example, FDE data is very rich in volume and is often the earliest identification of a fault. However, the corresponding fault may be resolved after a reset or a simple maintenance action. Hence, this fault had not led to operational disturbance or significant unscheduled maintenance that ReMAP intends to reduce, and is therefore less relevant than other faults.

Removal data is generally well structured, and one may assume that an unscheduled removal is a significant maintenance activity that could lead to a potentially large operational disturbance. However, the removal date may not be the same as the fault date, which could lead to mislabelling of the data.

Maintenance logs include faults with high operational impact and log the maintenance actions and findings from the first moment a fault was detected until the fault has been resolved. These logs contain a more reliable assessment of the aircraft condition than FDEs and removal data, but the data is unstructured and therefore more difficult to process in large volume. In addition, the root cause of the problem is not found until it is reported in the Shop Report. In some situations, the removed part may be assessed as 'No Fault Found'.

From the preliminary data-analysis, it is concluded that a good labelling of the data requires integrating multiple data-sources, which is the topic of Section 5.

5. Exploratory analysis for Embraer E-jet system selection

5.1. E-jets System Selection

As a starting point for Embraer activities in the WP5 working group, several discussions were held with *KLM Cityhopper (KLC)* regarding the selection criteria for E-jets components and systems. Considering the scope of the project, the best candidates were ranked considering:

- Operational Impact and Unscheduled Maintenance:
  - Unscheduled Ground Time (UGT)
  - Technical Dispatch Reliability (TDR)
  - MEL Dispatch Disposition / Repair Category
- Reliability:
  - Mean Time Between Removals (MTBR)
  - Mean Time Between Failures (MTBF)
- Scheduled Maintenance:
  - Maintenance Plan Tasks
  - Other Recommended / Required Tasks

5.2. Exploratory analysis with data from KLM B747

Since we had access to the data, Embraer team has been working on pre-processing raw data and summarization. In a first moment, *KLM* provided a sample of the data to start this activity. *At present the analysis is done using data from B747 bleed air system*[1].

The data contained in this sample are relative to the sensors and it allowed a preliminary interpretation. An exploratory data analysis was performed (including pre-processing and manipulation) extracting the first visibilities and understanding of the data content. The exploration was carried out throught the following approaches:

- Segregation and analysis per flight;
- Detection and correction of missing data;
- Sensor correlation;
- Analysis per flight phase;
- Identification of suspicious variations.

[1] *The analysis will be extended to be used with data from KLC E-jet aircraft, when the data will be made available to ReMAP consortium in the coming period.*

Index | Flight Phase | Sensors Output | Flight Number

	I	TimeStamp	FltPhase	S1	S2	S3	S4	S5	S6	S7	...	S15	S16	S17	S18	S19	S20	Tail	Origin	Destination	FltNbr
0	0	01-04-2018 13:57	1	0.0	0.0	26.8	25.6	0.0	0.0	64.7	...	NaN	NaN	40.0	NaN	NaN	NaN	73646573	.	.	.
1	1	01-04-2018 13:57	1	0.0	0.0	27.1	25.4	2.8	2.6	64.7	...	NaN	NaN	NaN	47.0	NaN	NaN	73646573	.	.	.
2	2	01-04-2018 13:57	1	0.0	0.0	27.3	25.3	4.9	4.8	64.8	...	35.0	NaN	NaN	NaN	111.0	NaN	73646573	.	.	.
3	3	01-04-2018 13:57	1	0.0	0.0	27.4	25.1	6.8	6.9	64.8	...	NaN	34.0	NaN	NaN	NaN	102.0	73646573	.	.	.
4	4	01-04-2018 13:57	1	0.0	0.0	27.3	25.1	8.5	8.8	64.7	...	NaN	NaN	60.0	NaN	NaN	NaN	73646573	.	.	.

Figure 8 Raw sample data layout

5.2.1. Raw Data

The sample data is divided into several columns, each containing information about: time, flight phase, tail number and the sensors measurements. The data cames from 4 identical systems with 5 sensors each (total 20 sensors). The **flight phase** refers to a period within a flight (ex: standing, taxi, takeoff, initial climb, cruise, descent, approach, landing, etc).

5.2.2. Missing Data

Different sensors storage rates results in some missing data (cells without any value) (Figure 9). These missing values were filled using the *forward fill method*, which fill the cell with the immediate previous value for that respective sensor.

5.2.1. Correlation Analysis

Before proceeding with the analysis, data were segregated by flight. A statistical and correlation analysis between the sensors was performed. Some of the results are presented in the following figures.

The Figure 10 map represents the correlation between the sensors and it is possible to identify a relation between several sensors. For instance, sensors number 9, 10, 11 and 12 present a very close correlation, as evidenced by their plot, where they present identical behaviour. Probably, each of these sensors gives information about the same magnitude. In Figure 10 we can detect that the sensors correlation is organized into groups of four. We suspect that this fact may be related to the number of engines of the aircraft. In Figure 11 raw values of a subset of sensors are shown to facilitate improved understanding of the data characteristics.

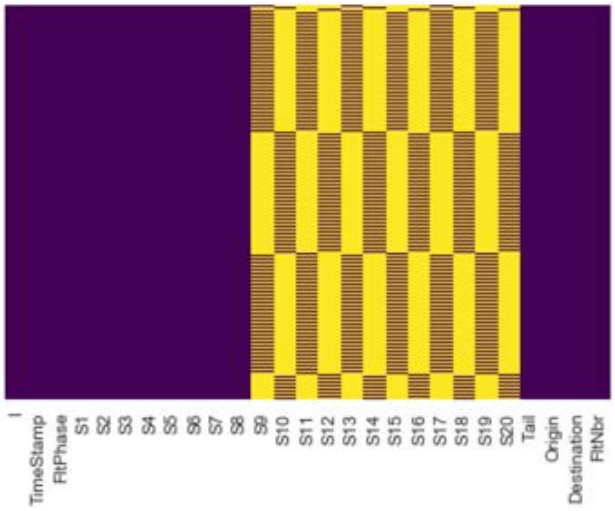

Figure 9 Missing values map (presented by yellow color).

	I	FltPhase	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
I	1	0.57	0.3	0.3	0.3	0.3	0.22	0.21	0.22	0.21	-0.11	0.0086	-0.099	-0.093	-0.23	0.35	0.27	-0.3	-0.17	0.42	0.17	0.16
FltPhase	0.57	1	0.89	0.89	0.89	0.88	0.84	0.83	0.86	0.85	0.71	0.78	0.73	0.73	0.28	0.51	0.33	0.13	0.59	0.89	0.72	0.74
S1	0.3	0.89	1	1	1	1	0.98	0.98	0.99	0.99	0.82	0.85	0.82	0.83	0.65	0.66	0.56	0.53	0.81	0.88	0.75	0.83
S2	0.3	0.89	1	1	1	1	0.98	0.98	0.99	0.99	0.82	0.85	0.82	0.83	0.65	0.65	0.56	0.53	0.81	0.88	0.75	0.83
S3	0.3	0.89	1	1	1	1	0.96	0.97	0.99	0.99	0.81	0.84	0.81	0.83	0.65	0.66	0.56	0.53	0.8	0.87	0.73	0.81
S4	0.3	0.88	1	1	1	1	0.97	0.97	0.99	0.99	0.81	0.84	0.81	0.83	0.66	0.66	0.56	0.53	0.81	0.87	0.73	0.82
S5	0.22	0.84	0.98	0.98	0.96	0.97	1	1	0.97	0.97	0.84	0.87	0.84	0.85	0.66	0.61	0.53	0.55	0.85	0.87	0.78	0.86
S6	0.21	0.83	0.98	0.98	0.97	0.97	1	1	0.97	0.98	0.84	0.87	0.84	0.85	0.68	0.62	0.54	0.56	0.85	0.86	0.77	0.85
S7	0.22	0.86	0.99	0.99	0.99	0.99	0.97	0.97	1	1	0.84	0.86	0.84	0.85	0.69	0.63	0.54	0.57	0.84	0.86	0.74	0.83
S8	0.21	0.85	0.99	0.99	0.99	0.99	0.97	0.98	1	1	0.84	0.86	0.85	0.86	0.68	0.62	0.53	0.56	0.85	0.86	0.75	0.84
S9	-0.11	0.71	0.82	0.82	0.81	0.81	0.84	0.84	0.84	0.84	1	0.99	1	1	0.55	0.31	0.23	0.42	0.89	0.78	0.85	0.87
S10	0.0086	0.78	0.85	0.85	0.84	0.84	0.87	0.87	0.86	0.86	0.99	1	0.99	0.99	0.49	0.33	0.22	0.36	0.87	0.83	0.88	0.89
S11	-0.099	0.73	0.82	0.82	0.81	0.81	0.84	0.84	0.84	0.85	1	0.99	1	1	0.53	0.3	0.21	0.4	0.88	0.77	0.83	0.85
S12	-0.093	0.73	0.83	0.83	0.83	0.83	0.85	0.85	0.85	0.86	1	0.99	1	1	0.56	0.33	0.25	0.43	0.89	0.78	0.83	0.86
S13	-0.23	0.28	0.65	0.65	0.65	0.66	0.66	0.68	0.69	0.68	0.55	0.49	0.53	0.56	1	0.66	0.74	0.97	0.74	0.41	0.38	0.55
S14	0.35	0.51	0.66	0.65	0.66	0.66	0.61	0.62	0.63	0.62	0.31	0.33	0.3	0.33	0.66	1	0.65	0.63	0.46	0.51	0.34	0.48
S15	0.27	0.33	0.56	0.56	0.56	0.56	0.53	0.54	0.54	0.53	0.23	0.22	0.21	0.25	0.74	0.65	1	0.74	0.41	0.38	0.23	0.39
S16	-0.3	0.13	0.53	0.53	0.53	0.53	0.55	0.56	0.57	0.56	0.42	0.36	0.4	0.43	0.97	0.63	0.74	1	0.66	0.27	0.24	0.43
S17	-0.17	0.59	0.81	0.81	0.8	0.81	0.85	0.85	0.84	0.85	0.89	0.87	0.88	0.89	0.74	0.46	0.41	0.66	1	0.76	0.77	0.85
S18	0.42	0.89	0.88	0.88	0.87	0.87	0.87	0.86	0.86	0.86	0.78	0.83	0.77	0.78	0.41	0.51	0.38	0.27	0.76	1	0.85	0.89
S19	0.17	0.72	0.75	0.75	0.73	0.73	0.78	0.77	0.74	0.75	0.85	0.88	0.83	0.83	0.38	0.34	0.23	0.24	0.77	0.85	1	0.95
S20	0.16	0.74	0.83	0.83	0.81	0.82	0.86	0.85	0.83	0.84	0.87	0.89	0.85	0.86	0.55	0.48	0.39	0.43	0.85	0.89	0.95	1

Figure 10 Representation of sensors (S) correlation.

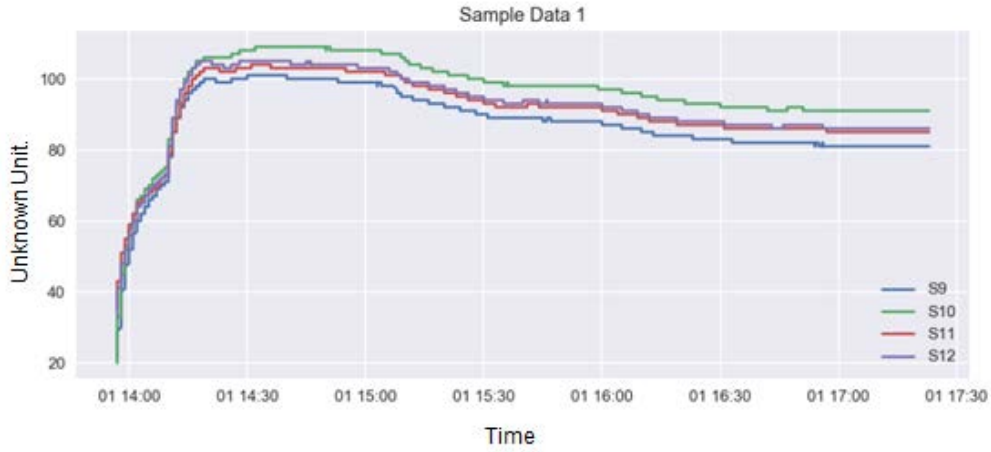

Figure 9 Sensors plot (S9, S10, S11 and S12).

5.2.2. Analysis per Flight Phase

Figure 12 shows the flight phase over time, probably the take-off until get the cruise flight (flight phase = 8). The sensors output varies according to the flight phase. The flight takes 30 minutes to stabilize at flight phase 8, as with sensors output. In cruise flight, sensors output remains stable, except in sensor S15 where we can find a **spike**.

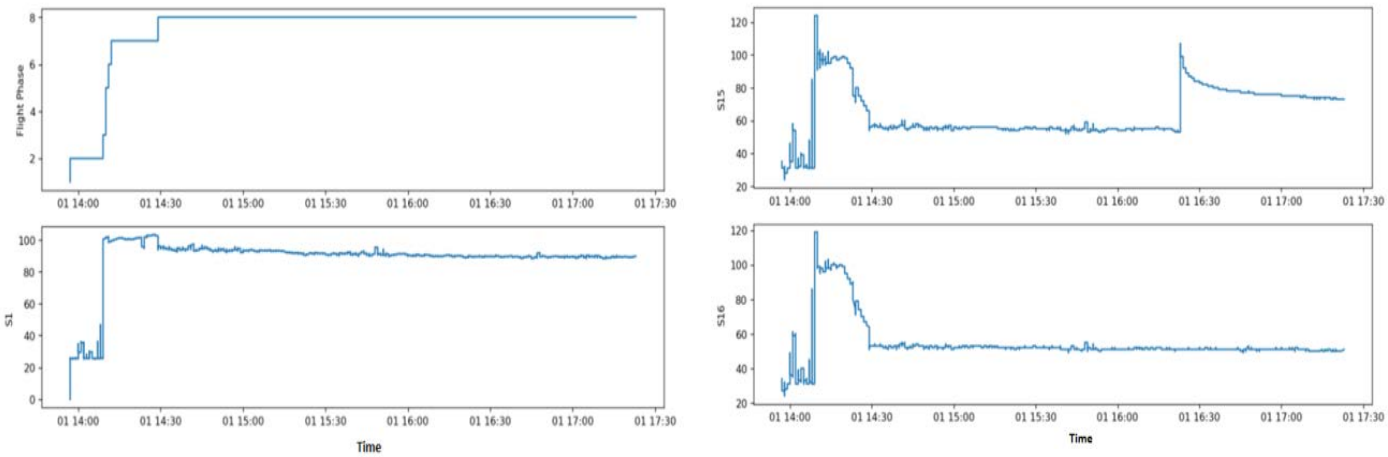

Figure 12 Flight Phase and S1, S15 and S16 output plots.

5.2.3. Identifying suspicious variations

In order to automatically identify the spikes, a numeric method was developed. As we can see in Figure 12 and Figure 13, sensor S15 has a positive spike during flight phase 8. This behaviour is automatically identified creating a plot with a dirac at the moment it occurs. At the moment it is not possible to draw any conclusion about the causes that lead to spike emergence.

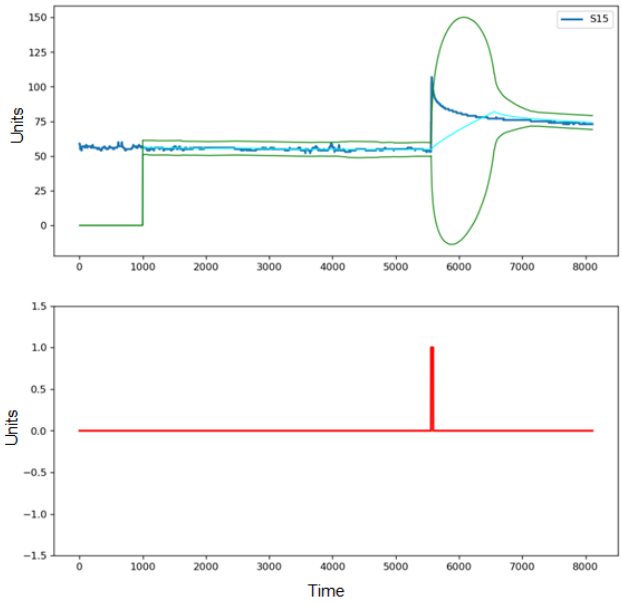

Figure 13 Detected spike out of the green line boundary (top figure). The red plot identifies the instant when spike arises (figure below).

5.3. Conclusion

The work described in Section 5 results from a first sensors data assessment carried out in order to allow a familiarization and understanding of data content. A preliminary data pre-processing is essential to capture the relevant information to conduct the system evaluation. A similar inspection was performed for data related to Flight Deck Events (FDE) and Maintenance Messages (MMSG), as it is described in Section 6.

The general conclusion from this exploratory analysis is that the B747 Bleed Air System data set received can be used for developing the proposed studies of PHM, since some proper pre-processing methods are applied. Most relevant manipulation for preparing the data are related to the segregation of data per flight and flight phase and, if necessary, treatments for missing values, as required by specific algorithms. Extraction of some features based on statistical analysis for sensors data is also possible and can be useful for generating data-driven models.

6. Automated analytics methodologies for maintenance logs and messages for 747 Bleed Air systems

In early April, a first data set related to the bleed air system from the Boeing 747 fleet was provided to the partners. Besides the sensor readings, data to be used to develop the PHM methodologies also includes failure information: **Flight Deck Events (FDE)** and **Maintenance Messages (MMSG)**. These messages can give indirect information about system failures. Embraer team has been conducting an exploratory method in order to better identify the failure events. The method consists of detailed analysis of fault/failure messages in conjunction with **components removal** data for each tail number. **The objective of this method is to determine actual system failures in order to be considered as input to the algorithms for PHM**.

6.1. Flight Deck Events (FDE)

A sample of the Flight Deck Events data received, as shown in Figure 14, has two columns with information of interest: “Message Text” and “Fault Text”. Both of them give information about system number and type of occurrence/component. Because it has a shorter pattern, “fault text” column was the focus of our study, but the same work can be done for “Message Text” column.

	A	B	C	D	E	F	G	H	I	J
1	Message Co	Message Text	Tail #	Leg Date	Flight Ph	ATA	Fault Code	Fault Text	Fault Report Code - Monitor Coc	Priority
2	NO COR		88022223	2/26/2015 16:19		36-00-00	36100700	BLEED ISLN APU		LOW
3	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/15/2015 15:17	ER	36-11	36101300	BLEED 3 OVHT		LOW
4	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/15/2015 15:17	ER	36-11	36101200	BLD 3 OVHT/PRV		MEDIUM
5	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/16/2015 4:48	ER	36-11	36101300	BLEED 3 OVHT		LOW
6	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/16/2015 4:48	ER	36-11	36101200	BLD 3 OVHT/PRV		MEDIUM
7	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/21/2015 18:54	DC	36-11	36101300	BLEED 3 OVHT		LOW
8	36268	BLEED-3 PRSOV TEMPERATURE TOPPING FAIL	88022223	4/21/2015 18:54	DC	36-11	36101200	BLD 3 OVHT/PRV		MEDIUM
9	36269	BLEED-4 PRSOV TEMPERATURE TOPPING FAIL	88022223	5/13/2015 13:11	DC	36-11	36101400	BLD 4 OVHT/PRV		MEDIUM
10	36289	BLEED-4 FAN AIR MODULATING VALVE/FATS FAIL CLOSED	88022223	5/13/2015 13:11	DC	36-12	36101400	BLD 4 OVHT/PRV		MEDIUM
11	36269	BLEED-4 PRSOV TEMPERATURE TOPPING FAIL	88022223	5/13/2015 13:11	DC	36-11	36101500	BLEED 4 OVHT		MEDIUM
12	36289	BLEED-4 FAN AIR MODULATING VALVE/FATS FAIL CLOSED	88022223	5/13/2015 13:11	DC	36-12	36101500	BLEED 4 OVHT		MEDIUM

Figure 14 FDE original data (provided by KLM)

The FDEs messages were organized on a timeline, as presented in Figure 15, where each colour represents a different event. It is important to mention that each FDE can still result in Non-Fault Found (NFF), furthermore each FDE can re-occur in several flights, all linked to the same problem (while there is a fault the FDE still appearing).

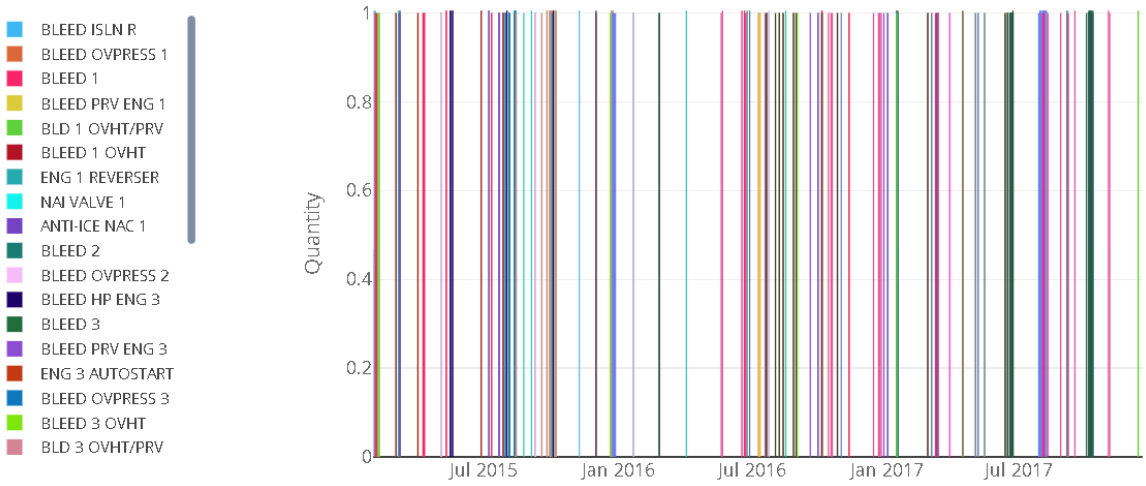

Figure 10: Flight Deck Events (FDE) over time.

Figure 16 shows the typical concentration of Fault Messages (FDE) of each system and from it we can distinguish two different failure modes, one related to pressure and another one related to temperature. This perception of different failure modes can be noticed in every system.

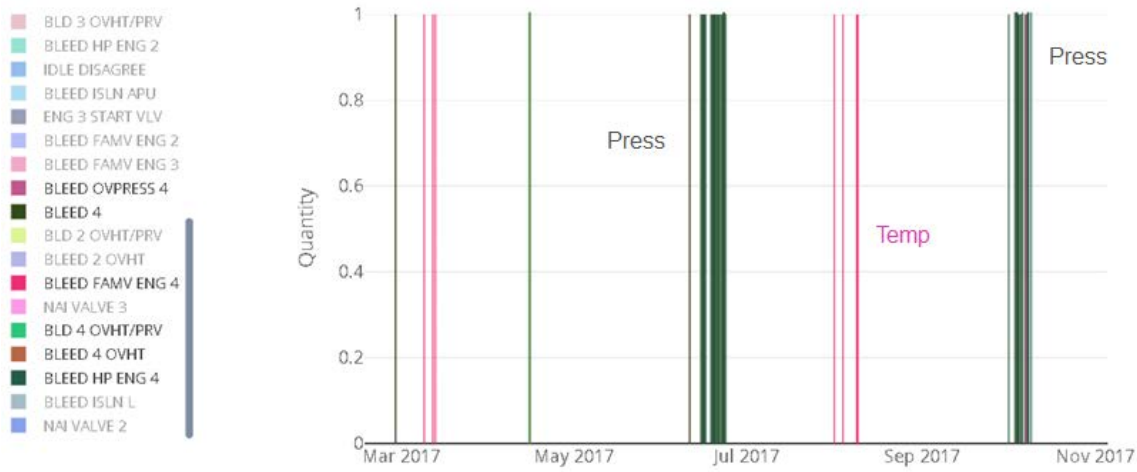

Figure 11: Typical Fault Messages (FDE) concentration.

6.2. Components Removals

A sample of the component removals data is shown in Figure 17, where we can extract information about the removed component. The “Position” column indicates the system number which the respective component belongs. In the “Rem.Code” column the number may varies between 1 to 4, sorted by removal type as:

- Removal Code 1 - Removed because component is on a time-limit (scheduled maintenance);
- Removal Code 2 - Removed for a cross-exchange;
- Removal Code 3 - Unscheduled removal (i.e. due to a fault);
- Removal Code 4 - Shop check (may be due to fault but could also be an administrative input).

Bearing in mind that the purpose of this work is to study fault related removals, only code 3 was considered in the method. Also, a component removal could not solve the existing problem and removal can still result in Non-Fault Found (NFF).

1	Name	Rem/Inst Date	IN/OUT	AC Reg	Position	Rem.Code	TSI Hours	TSI Cycles	ATA	Airline	FF?
2	VLV HPSOV	13-03-2015	OUT	38017358	100	3	118123	15961	361201	KLM	
3	VLV HPSOV	13-03-2015	OUT	38017358	100	3	118123	15961	361201	KLM	
4	CNTRL PRV	13-03-2015	OUT	38017358	100	3	1092	139	361120	KLM	
5	CNTRL PRV	08-04-2015	OUT	38017358	300	3	312	37	361120	KLM	

Figure 12: Removals original data (provided by KLM).

6.3. Failure Assessment (Date and Component)

In most of cases, combining the information from FDEs with the removals data we are able to estimate the data failure and its affected component. The graphics in Figure 18 represents a real case study that we are using to demonstrate the developed method. From the picture it is possible to observe that the same FDE is occurring every day since 28th July. Immediately after the sensor temperature removal, the occurrence of FDEs ceases, allowing us to conclude that this should be the component responsible for the malfunction and the date of first FDE should be the beginning of this failure (28th July).

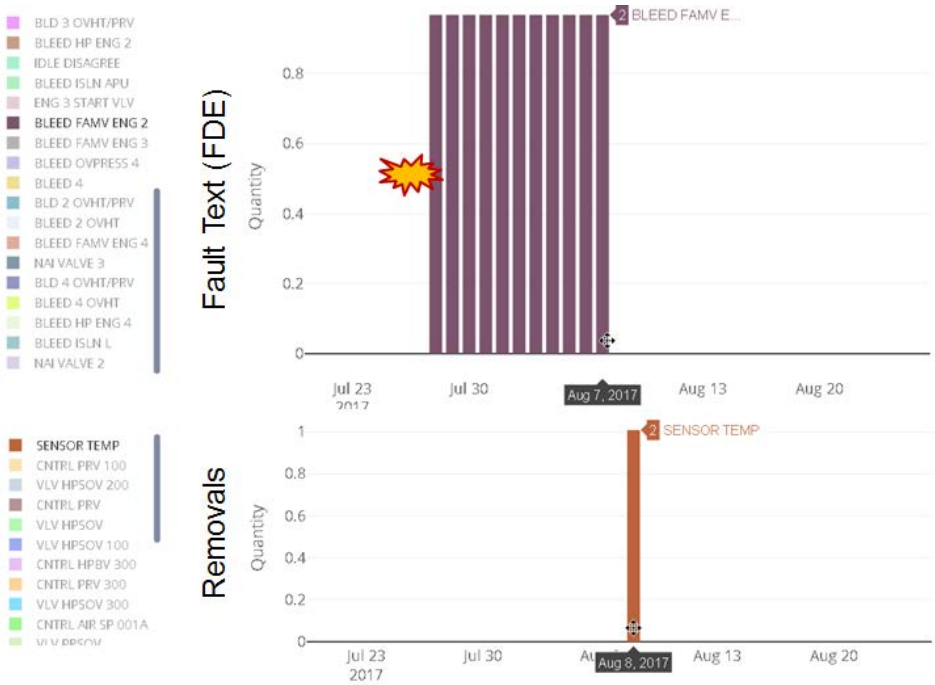

Figure 13: Use case #1 in determining failure date.

In some cases, the failure assessment becomes more complex, as the case presented in Figure 19. In this case, failure causes several FDEs which are followed by three different component removals. The FDEs stop appearing after removing component "CNTRL PRV 300", suggesting that it should be the broken component and its fault occurred on 21st December. Nevertheless, the same component had already been replaced once and the FDEs start showing up again some days after. Probably this first component failed early in life.

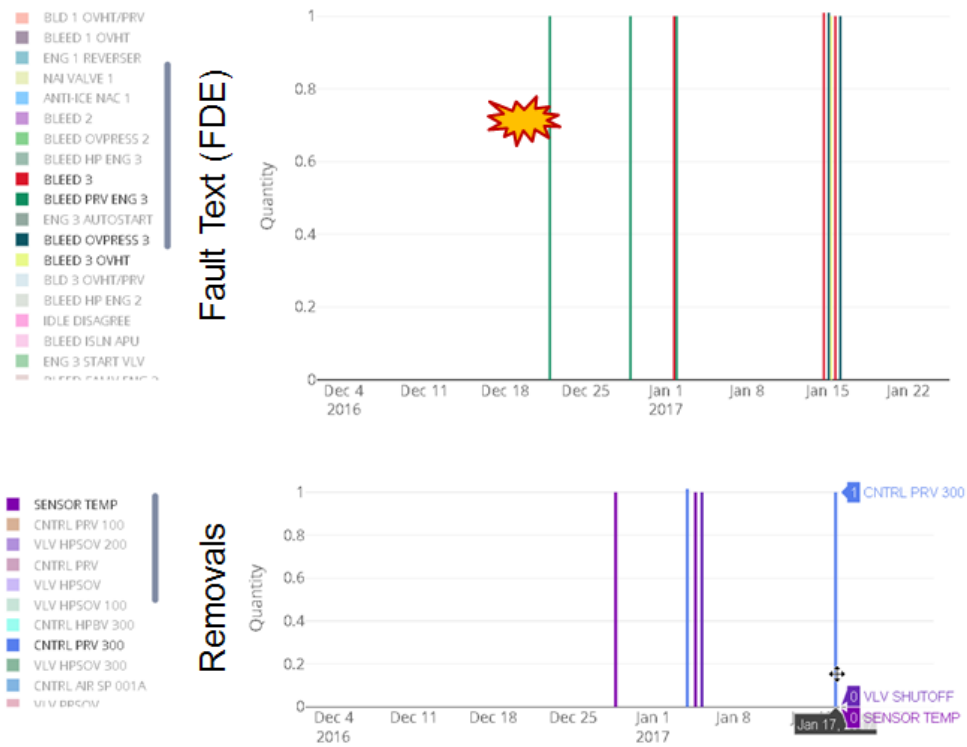

Figure 14: Use case #2 in determining failure date.

6.4. Fault Identification Tool

This work culminated on the development of a **tool capable of automatically identify the failure date** of a given component. This script has as input all FDE's data messages (for all aircraft tails) and returns the start and end date of each FDEs concentration, for each type of FDE. Figure 20 represents a timeline distribution of a specific FDE type where two different messages concentration can be detected. The tool has two main tuning parameters:

- Constant of Time (*ct*) – maximum time (in days) to search for a new FDE, considering it belongs to the same concentration group.
- Constant of Frequency (*cf*) – minimum amount of FDE that defines a concentration.

For most cases the best performance was obtained for: $ct = 25\ days$ and $cf = 3$. It means that when the software detects a FDE, it will search for another one in the following 25 days. In this time window it repeats this process every time a FDE is found. As *cf* is

set to 3, each FDE concentration consists of 3 or more occurrences. In Figure 25, it can be observed that a single message in July 2016, which is not defined as fault concentration.

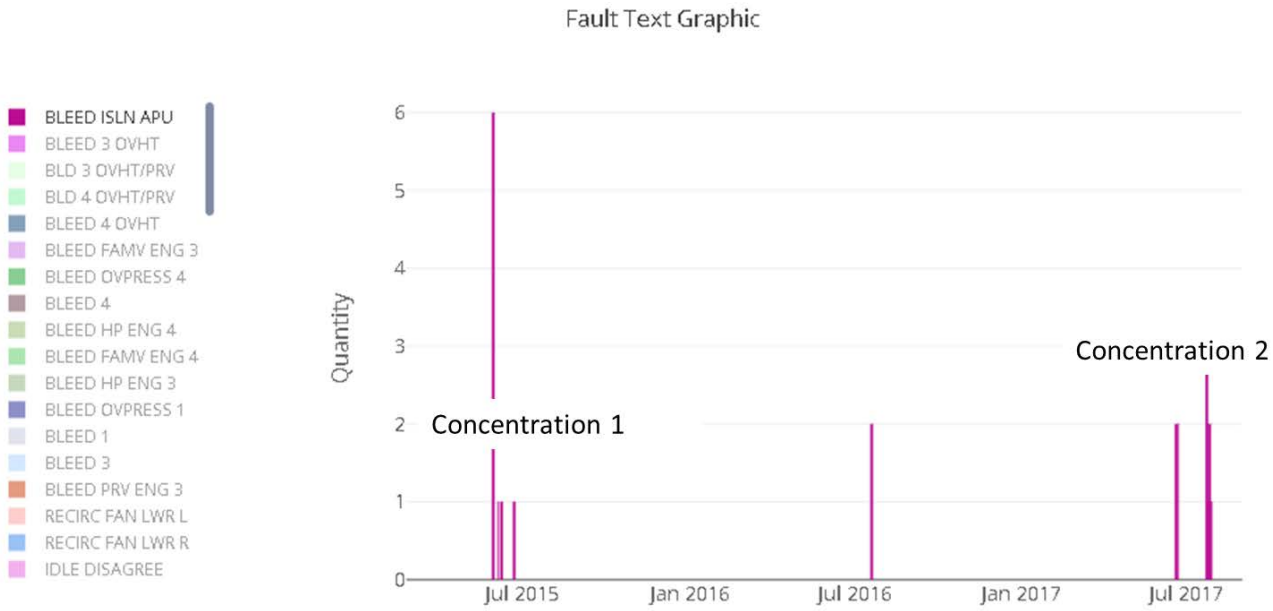

Figure 15 Distribution of a specific FDE type overtime.

Tail #	FDE	date init	date end
4388298	BLEED ISLN APU	2015-05-30	2015-06-22
4388298	BLEED ISLN APU	2017-07-24	2017-07-29
4388298	BLD 4 OVHT/PRV	2016-02-17	2016-03-04
4388298	BLEED 4 OVHT	2016-02-17	2016-03-04
4388298	BLEED FAMV ENG 2	2017-12-12	2017-12-16
4388298	BLEED FAMV ENG 1	2015-07-12	2015-07-20
4388298	BLEED PRV ENG 2	2015-03-03	2015-03-07
4388298	BLEED PRV ENG 2	2016-06-25	2016-08-01
4388298	BLEED PRV ENG 2	2016-09-25	2016-10-04

Figure 16 Output from Fault Identification Tool

The developed tool informs about the failure date(s) of each component (Figure 21), serving as input to the algorithms for PHM. An identical tool to perform the Maintenance Messages data (MMSG) was developed.

6.5. Matching Failures with Sensors Output

After analysing system failures from the information present in the FDEs and removals, it would be helpful to understand their relationship with the behaviour of the signal coming from the sensors. Aggregating the work developed by UC, focused on sensor data analysis, with Embraer's work on fault data, that assessment can be done. Figure 22 shows the match between the use case #2 with the two sensor outputs.

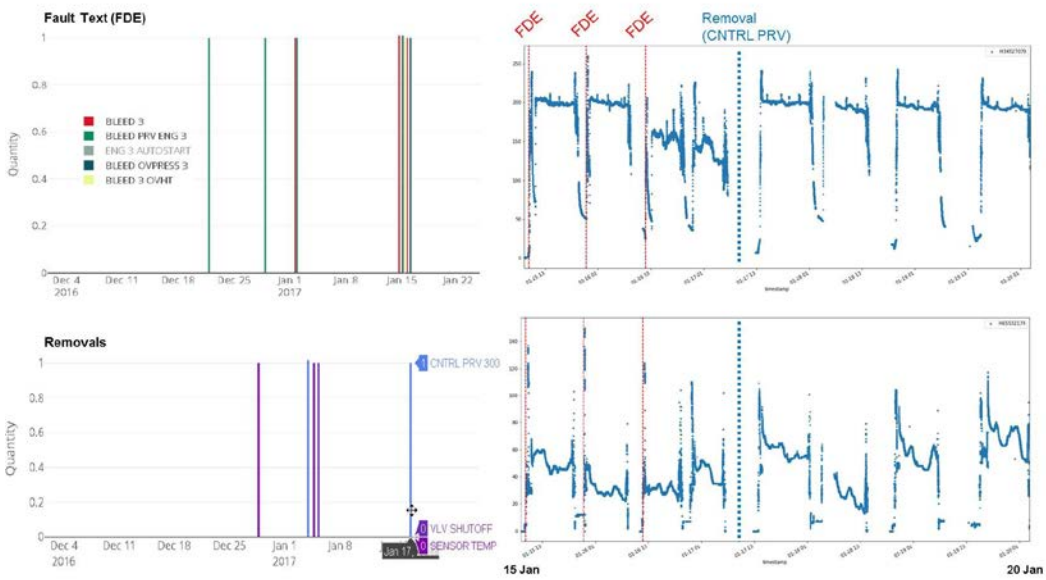

Figure 17 Visualization of FDEs on sensors outputs.

From Figure 22 graphics we can visually unveil an anomaly on the sensors' outputs after FDEs occurrence. The sensors outputs get a new behaviour as soon as the affected component is removed. As a comparison, Figure 23 presents some sensors outputs during a period of time without any FDE or removal, showing their patterns for a system without any anomaly. As expected, the healthy system presents a regular sensor output patterns, as it is also normal its operation.

With the work presented in this section, we can conclude that the data from FDE and MMSG allow us to assess components historical operating status, helping to estimate the failure start date and how long this failure remained. The failure timeframe can be correlated with the data coming from the sensors and thus identify the output sensor signals that are related to fault. This information is crucial as input for the diagnostic and prognostic algorithms.

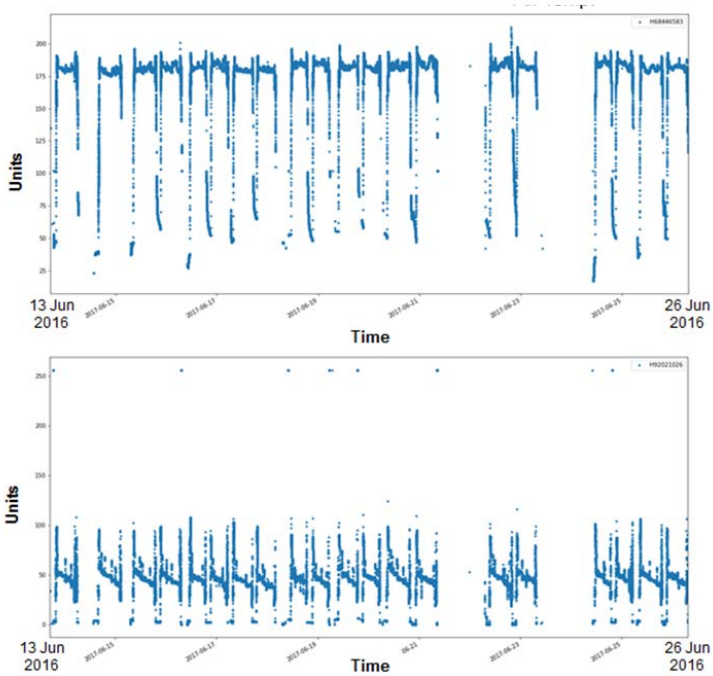

Figure 18: Sample sensor outputs during a period without any FDE or Removal.

7. Exploratory analysis on edge computing for health estimation

ReMAP proposes an exploratory analysis on the use of edge computing principles for the estimation of health. The principle to be explored consists on the use of on-board, lean algorithms able to diagnose the existence of system faults that can be immediately communicated to the maintenance control centre. Towards this goal we implemented a simulation environment for development, testing and validation of Edge computing methodologies. The platform is being used evaluated for some algorithms, with the goal of improving them in terms of speed, data usage, and computation ability. The simulation environment, that can be representative of a real workflow system, is used to evaluate the potential of implementing edge computing concepts in an aircraft system. In order to better understand how the output of the simulation interact with this environment, we implemented an interface that displays the main parameters under study.

This simulator study aims to answer these three questions about edge-computing:

1. Can it contribute to improving overall prediction accuracy?
2. Can it be used to reduce the response time to faults?
3. Can it contribute to the reduction of data being transferred?

To answer these questions, we implemented the tracking of specific information in the simulator, which can show us, for example, the time from a fault being detected until the faulty component is replaced. Figure 24 presents the main interface of this simulator: on the left side, the individual state of each individual airplane; at the top right the RUL estimated by the central model from the component being simulated; and, at the bottom right, the graphical log of the amount of data being transferred per flight, and the MSE on the RUL estimation where each vertical blue line represents a discrete event of training (or retraining) the classifier with the available data so far.

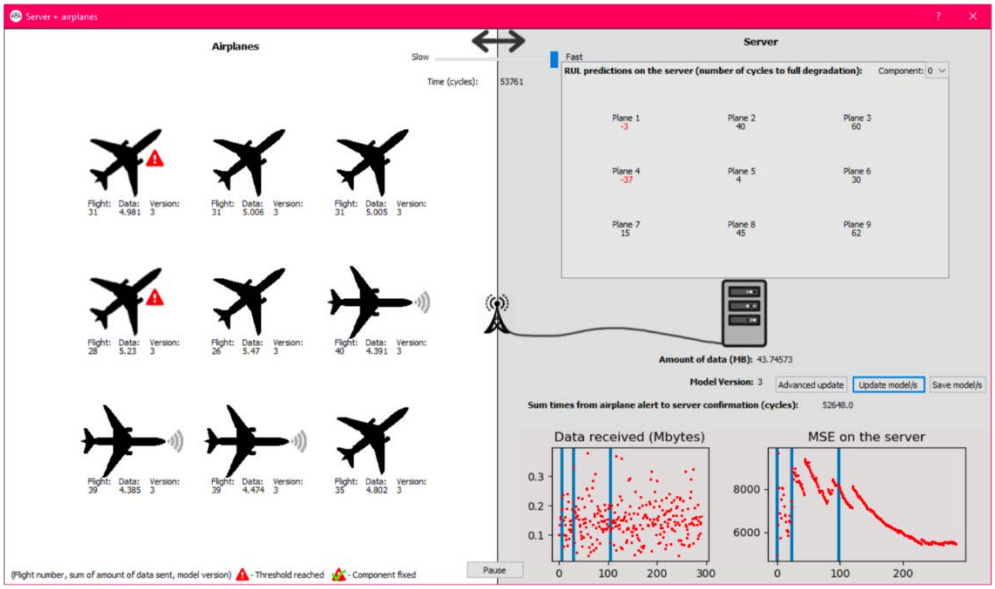

Figure 19 General interface view of the implemented simulator.

7.1. Improving accuracy while the system is online

This test serves as a demonstration of the ability of the system to improve itself using new data that is not yet available to any possible pre-existent regression model. Even though these models usually require a substantial verification and deployment processes, this test serves to show a possible response time reduction in deployment time while also having the ability to be verified before deployment and with the added benefit of being easily updated.

From the Mean Squared Error (MSE) graph on Figure 24:

- Up to the third training event, the predictions have a consistently high average MSE (when compared with posterior predictions).
- After the third, the predictions are consistently better, eventually stabilizing around the optimal value for the dataset.

7.2. Improvement of response time to faults

One problem being addressed on ReMAP is how to improve the response time to faults and how to translate that into savings in the maintenance and planning activities.

On Figure 29 the parameter “Sum time from airplane alert to server confirmation” represents the number of cycles between the RUL estimated at the airplane and the RUL estimated at the central server. This can be understood as a time estimate that could have been used to manage new maintenance schedules and prepare for the eventual maintenance procedure for the targeted component. This is most beneficial when the component has an unexpected problem mid-flight.

Even though this alert could not be sent directly to the company, in normal flight conditions; where the plane has no connection to the server; it shows the benefit of making use of the very low bandwidth connection to the server, where the aircraft can communicate with the server at any point during the flight.

7.3. Reduction on the bandwidth needed to data transfer

One of the goals of this task is to show the advantages of some edge computing principles, such as performing extra processing on the aircraft, this could lead to better predictions mid-flight, but it could also be used to reduce the bandwidth used to transfer data to the server, using methods such as compression or any type feature selection. This test serves to show the benefits of using feature selection in this system in terms of amount of data transmitted.

In Figure 30, a model was trained, at the start of the simulation, then, later, updated when more data was available. This update included an F-Score feature selection.

When compared with the results of Figure 25, in the data graph we see:

- The first few samples are in order of size, due to the fact that the faster flights finish first.
- The samples before the second update are consistently larger than the ones afterwards.
- This reduction in transfer size is obtained with no perceptible loss on the MSE

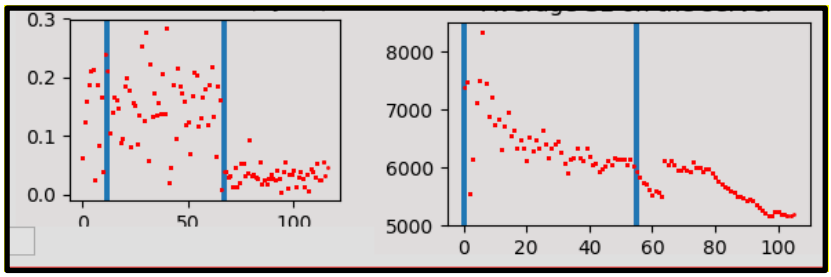

Figure 20 Testing results when updating the model with feature selection

7.4. Future developments

Considering the preliminary results obtained, a few possible improvements to the current

system come up:

- The usage of more models for even more components, to test the viability of full-scale models on the aircraft and to assess how much we have to reduce the size of the models.
- And, with more data and more information on the systems, better and smaller models could be generated in order to get better accuracy and make the system more appealing

8. Exploratory analytics for Health Indicator computation for 747 bleed air and 787 brake system

The exploratory data analytics, performed by the UC, was focused on the brake system of the Boeing 787 and on the Bleed Air System of Boeing 747, whose data was provided by *KLM*.

8.1. Brake System (Boeing 787)

The provided data, which concerned the information of the brake wear of 8 different brake positions over time, was used to develop a methodology for the diagnosis and prognosis of the system health condition. Due to the data characteristics no special data processing was performed.

8.1.1. Diagnostic – Health Indicator (HI) computation

By analysing the brake data, it was found that the data regarding the system condition was already provided. As the data features correspond to the wastage percentage (health condition) of the different brake positions, these values correspond directly to the Health Indicator of each position.

8.1.2. Prognostic – RUL computation

Considering the HI evolution of the brake position over time, a **Linear Regression** was used for modelling and predicting the HI evolution, and consequently, estimating the RUL value.

The **Linear Regression** is formulated as follows:

$$y = \alpha + \beta x$$

Where, the intercept, α, represents the initial percentage of brake health of the specific brake position being analyzed and the slope, β, represents the degradation ratio.

Figure 26 illustrates the application of a Linear Regression for predicting the future degradation behaviour of a single degradation lifecycle regarding a specific brake position.

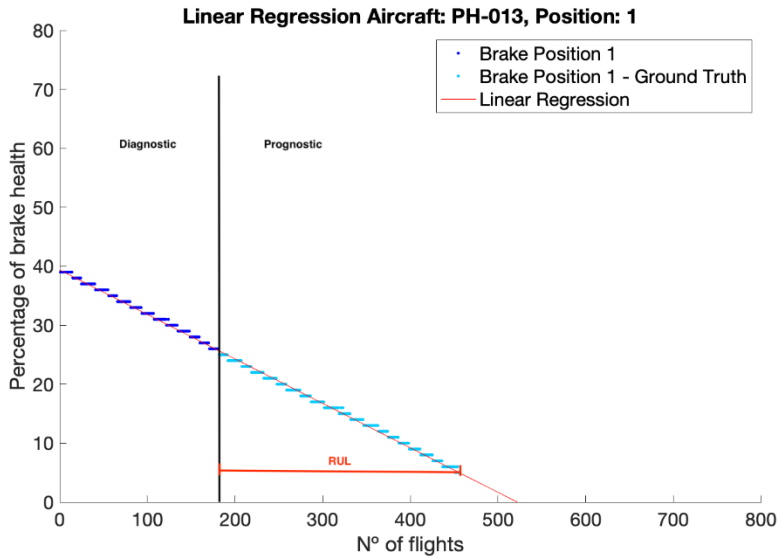

Figure 21 RUL estimation for Brake system

The considered health threshold for determining the RUL was 4%. This 4% corresponds to the average of the health percentage (HI) at the removal instant of all the brake deterioration trajectories present in the dataset.

As observed, the predicted degradation curve correctly represents the ground truth, which results in an accurate estimation of the RUL value, represented in the Figure 31.

8.2. Bleed Air System (Boeing 747)

For this system a significant amount of data was provided, which is represented in three groups:

- **Sensors data:** Raw data extracted directly from the Bleed Air sensors.
- **Flight Deck Events (FDE):** Information regarding the alerts (FDEs) triggered during the flights.
- **Removals data:** Information and details regarding the removals (parts replacement) performed in the Bleed Air system.

8.2.1. Preprocessing on the Bleed Air data

Taking in consideration the **sensors data** characteristics and importance, some pre-processing steps were applied to this data, namely **flight phase aggregation**, **detection and accommodation of outliers** and **flight labelling**.

The **phase aggregation** was performed with the intention of isolating the different patterns presented in the sensor values during a single flight. This way, a more correct and appropriate analysis could be performed.

The 14 initial phases were aggregated in 5 phases, in the following way:

1. **Start**: Flight Phase 1
2. **Climb**: Flight Phase 2-7
3. **Cruise**: Flight Phase 8
4. **Descent**: Flight Phase 9 - 13
5. **Finish**: Flight Phase 14

Next, an algorithm for **detecting and adjusting possible outliers** was delineated. The developed algorithm uses a sliding window approach for detecting outliers based in the [μ-3*σ, μ+3*σ] interval, where the μ and σ represent the mean and standard deviation, respectively, of the windowed values. If the values are outside this interval they are considered an outlier and thus are adjusted to the interval [μ-2*σ, μ+2*σ].

Also, a new **flight labelling** was performed due to the inconsistencies found in the labelling provided in the initial dataset. The new labelling was based on the combination of two criteria: **Data rows timestamps** and **Phase transition 14 → 1**. In a first iteration over the data, the time criterion is used for the flight labelling, where if the difference between the timestamp of two consecutive data rows was higher than a specific threshold (ex: 5 hours) a break in the current flight is performed. Then in a second iteration, the phase transition 14 → 1 is used. The goal of the second iteration is identifying 'sub-flights' not detected in the first iteration. This way, if there are 'sub-flights' within the same flight label, these can be detected and isolated through the identification of the transition 14 → 1.

8.2.2. Diagnostic – Health Indicator (HI) computation

For the health condition diagnosis in the Bleed Air system, a new approach was used.

The HI was formulated as a combination of different variables that affect the system condition. Each variable has a weight assigned which reflects its impact in the HI value calculated. Examples of these variables are the flight hours, flight conditions, flight destination and deviation and variance in the sensors data.

Based on the available data, the HI was formulated using three variables:

- **Duration:** Flight hours (baseline)
- **α_1:** Sensors deviation from typical values. It is analysed using the **mean** feature
- **α_2:** Sensors deviation from typical values. It is analysed using the **std** feature

This way, the HI formula is the following:

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

Where, the *n* corresponds to the number of flights and *p* to the number of aggregated phases. Regarding the **α_1** values these can be:

- **α_1 = 0.6:** If the sensors values are too deviated from the typical range.

- **α_1 = 0:** If the sensors values are within the expected range.

With respect to **α_2** values these can be:

- **α_2 = 0.6:** If the sensors value variation is too deviated from the usual range.
- **α_2 = 0:** If the variation of the sensors value is within the typical range.
- **α_2 = -0.6:** If the sensors value variation is below the expected values.

The values for α1 and α2 were chosen using a trial and error approach. Due to the absence of knowledge regarding the system operation or its age evolution, no special parameter tuning could be performed. Thus, the choice of the α1 and α2 values focus more on their interpretability, rather than explainability.

Using this formulation, Figure 27 illustrates the HI computation for a specific time interval.

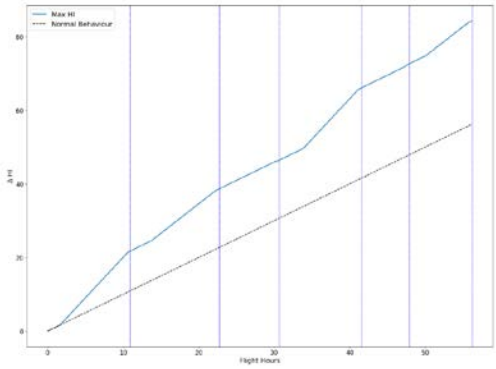

Figure 22 HI estimation based on the proposed formulation

As can be observed in the Figure 27, there are fluctuations in each flight (time interval between two consecutive vertical lines). These are due to the different α values determined. The black curve represents a situation of normal degradation where the α values are 0, thus the HI values is equal to the number of flight hours.

8.2.3. Prognostic – RUL computation

Regarding the RUL computation, a Similarity-based approach is used. In this approach, the Euclidean distance is used for assessing which degradation presented in the training trajectories better describes the degradation presented in the test trajectories. Then, based on the most similar trajectories the RUL of the test trajectory is computed. Each trajectory aims to represent a total degradation cycle of a component of the Bleed Air system.

Generally, the RUL computation algorithm is defined in three steps:

1. Computation of Similarities between the Test trajectory and each of the Train trajectories: For the similarity assessment, the Euclidean distance is used.

2. Selection of the most similar trajectories to the Test trajectory: The trajectory selection is performed by combining two criteria:
    a. **Euclidean Distance:** A fixed distance threshold is used to separate the similar and non-similar trajectories.
    b. **Minimum Distance:** The trajectories with lower distance errors are used as the most similar to the Test trajectory.
3. **Computation of the RUL interval:** After selecting the most similar trajectories the RUL interval is calculated using the minimum and maximum RUL of the selected trajectories.

Figure 28 illustrates the estimation of the RUL interval for a certain time interval containing 106 consecutive flights.

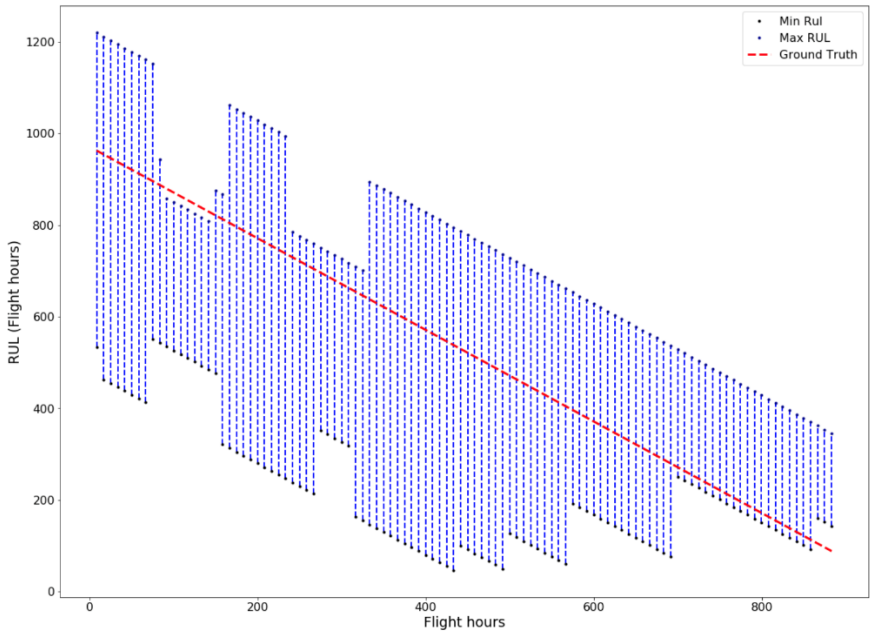

Figure 23 RUL estimation in the Bleed Air system

In Figure 28, the blue zone corresponds to the predicted interval and the red line to the ground truth. As can be observed, the ground truth is within the predicted interval. Also, there are oscillations in the RUL interval boundaries, these are due to the fact that the set of similar trajectories is updated over time, which changes the minimum and maximum RUL obtained. The significant variance associated with the interval predicted is mainly due to the small set of trajectories (7 trajectories) used as Train trajectories. In order to validate the models created, the inclusion (or not) of the ground truth in the predicted RUL interval and its variance should be analysed.

9. Summary and Conclusions

As part of task 5.1, exploratory analytics were performed to: (a) study the state of the art in PHM for aerospace systems, (b) identify and defined specific scope and requirements for the ReMAP PHM methodologies which will be part of the condition-based tool chain shown in Figure 2. The main focus of the work has been on: (1) degradation training and visualization, health index computation and remaining useful life estimation using time series data, (2) analysis framework to study the effect of external factors (e.g., weather, pollution etc.) on failure patterns, (3) use of edge computing for health estimation, (4) automated analysis of maintenance and removal logs for analysing failure patterns. Furthermore, a detailed framework was developed to share KLM aircraft operational data with the partners, who would develop PHM algorithms for aircraft components and systems. At this initial phase of exploratory analysis, the following data sets were used: (a) NASA turbofan engine data, (b) 747 bleed air system data, (c) 787 brake system data.

These datasets were analysed and characterized for a thorough understanding. In the context of pre-processing requirements, removal of missing data and normalization of the data were the main methodologies explored. At this stage of the work, the initial exploratory analysis show that the quality of data is adequate for the proposed objectives of WP5 when some pre-processing methods are applied. With regards to the volume of data, it is not yet possible to determine at this stage, if the quantity of tail numbers, flights and faults contained in the datasets (more specifically in the B747 bleed air system data) are sufficient for achieving the desired levels of accuracy when developing PHM models. In the next phases of the project, as access to more and more data is established for the tasks of WP5, further requirement-based data pre-processing methodologies will be explored.

A detailed literature survey of the state of the art of PHM research performed to define requirements for the algorithms to be developed in ReMAP. After a detailed study of the possible feature extraction methodologies it was decided, that at the initial stage the use of deep learning methodologies in would not require any explicit feature extraction. However, during the technology readiness level progression of the PHM algorithms in Task 5.2, as per requirements, further feature analysis methodologies might be explored.

The results were found to be very promising. The details and results of the framework shown in Figure 4 were invited to be presented in PHM 2019 conference at Scottsdale Arizona, 21-26 September 2019. A paper based on the framework shown in Figure 4 has been accepted for publication in peer reviewed International Journal of Prognostics and Health Management with some revision. Furthermore, a master thesis based on the work described in section 7 was submitted. In addition, a paper comprising of a detailed literature survey of state of the art PHM algorithms and approaches is prepared and is being submitted to “Aerospace” journal published by MDPI (Multidisciplinary Digital Publishing Institute) (Basora, Olive & Dubot, 2019).

Also, some preliminary work related to edge computing explored its potential to improve overall prediction accuracy, reduce the response time to faults and reduction of volume of data required to be transferred.

The next steps are: (1) based on the defined requirements and specifications in Task 5.1, further development, testing and validation of the databased PHM algorithms with operational data from some the 12 aircraft systems mentioned in the ReMAP proposal scope, (2) further development of Physics based fault isolation and root cause analysis to complement the developed data based PHM algorithms – this combination of data based and physics based methodologies will constitute the envisaged hybrid approach to PHM outlined in the scope of the ReMAP program.

References:

Aizpurua, J. I., & Catterson, V. M. (2015). Towards a methodology for design of prognostic systems. In Annual conference of the prognostics and health management society 2015 (pp. 504–517).

Akintayo, A., Lore, K. G., Sarkar, S., & Sarkar, S. (2016). Early detection of combustion instabilities using deep convolutional selective autoencoders on hi-speed flame video. arXiv preprint arXiv:1603.07839.

Basora, L., Olive, X., & Dubot T., Recent Advances in Anomaly Detection Methods applied to Aviation. Sept 2019. Submitted to Aerospace Journal (ww.mdpi.com/journal/aerospace).

Elattar, H. M., Elminir, H. K., & Riad, A. (2016). Prognostics: a literature review. Complex & Intelligent Systems,2(2), 125–154.

Ellefsen, A. L., Bjørlykhaug, E., Æsøy, V., Ushakov, S., &Zhang, H. (2019). Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. Reliability Engineering & System Safety, 183, 240–251.

Gugulothu, N., TV, V., Malhotra, P., Vig, L., Agarwal, P., & Shroff, G. (2017). Predicting remaining useful life using time series embeddings based on recurrent neural networks. arXiv preprint arXiv:1709.01073.

Li, J., Li, X., & He, D. (2019). A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction. IEEE Access, 7, 75464-75475. doi: 10.1109/ACCESS.2019.2919566

Malhotra, P., TV, V., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). Multi-sensor prognostics using an unsupervised health index based on LSTM encoder-decoder. arXiv preprint arXiv:1608.06154.

Mosallam, A., Medjaher, K., & Zerhouni, N. (2015). Component based data-driven prognostics for complex systems: Methodology and applications. In 2015 first international conference on reliability systems engineering (icrse) (pp. 1–7).

Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. Journal of Intelligent Manufacturing, 27(5), 1037– 1048.

Ramasso, E. (2014). Investigating computational geometry for failure prognostics. International Journal of Prognostics and Health Management, 5(1), 005.

Reddy, K. K., Venugopalan, V., & Giering, M. J. (2016). Applying deep learning for prognostic health monitoring of aerospace and building systems. In 1st ACM SIGKDD workshop on ml for phm.

Shahid, N., Ghosh, A., TrajecNets: Online Failure Evolution Analysis in 2D Space, Accepted for Publication in International Journal of Prognostics and Health Management, ISSN2153-2648. 2019.

Wang, J., Wen, G., Yang, S., & Liu, Y. (2018). Remaining useful life estimation in prognostics using deep bidirectional LSTM neural network. In 2018 prognostics and system health management conference (phmchongqing) (pp. 1037–1042).

Wang, T. (2010). Trajectory similarity-based prediction for remaining useful life estimation (Unpublished doctoral dissertation). University of Cincinnati.

Wu, Y., Yuan, M., Dong, S., Lin, L., & Liu, Y. (2018). Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. Neurocomputing, 275, 167–179.

Zhang, A.,Wang, H., Li, S., Cui, Y., Liu, Z., Yang, G., & Hu, J. (2018). Transfer learning with deep recurrent neural networks for remaining useful life estimation. Applied Sciences, 8(12), 2416.

Zhang, C., Pin, L., K. Qin, A., & Chen Tan, K. (2016, 07). Multi-objective deep belief networks ensemble for remaining useful life estimation in prognostics. IEEE Transactions on Neural Networks and Learning Systems, PP, 1-13. doi: 10.1109/TNNLS.2016.2582798.

Zhao, R., Wang, J., Yan, R., & Mao, K. (2016). Machine health monitoring with LSTM networks. In 2016 10th international conference on sensing technology (icst) (pp. 1–6).

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213–237.

Zhao, R., Yan, R., Wang, J., & Mao, K. (2017). Learning to monitor machine health with convolutional bidirectional LSTM networks. Sensors, 17(2), 273.