MONITORING OF MGB LUBRICATION AND COOLING SYSTEM BASED ON BIG DATA NORMALITY MODELS AND FUZZY EXPERT RULES

Ammar MECHOUCHE, Data Science Expert, Airbus, Marignane, France (Ammar.Mechouche@Airbus.com)
Matthis HOULES, Data Engineering Apprentice, Airbus, Marignane, France
Jérôme BELMONTE, Drive Systems Lubrication & Cooling Expert, Airbus, Marignane, France
Pierre-Loic MAISONNEUVE, HUMS / CBM Expert, Airbus, Marignane, France

Abstract

This paper presents an original method for the monitoring of the helicopter main gearbox lubrication and cooling system. The method relies on sensor data of oil pressures and temperature, as well as on domain expert knowledge. It combines oil pressures and temperature normality models built from a significant amount of training data collected from customers' helicopters, with fuzzy expert rules which allow precise root cause identification in case of lubrication or cooling sub-systems anomalies. The results have been validated based on known maintenance findings, showing that the proposed method allows to accurately detect anomalies related to the lubrication and cooling sub-systems as soon as they appear.

1. ABBREVIATIONS

AI – Artificial Intelligence
FDCR – Flight Data Continuous Recorder
FH – Flight Hours
FTI – Flight Test Instrumentation
H/C – Helicopter
HUMS – Health and Usage Monitoring System
ISIR – In-Service Incident Report
MGB – Main GearBox
MIS – Maintenance Information System
ML – Machine Learning
MRO – Maintenance Repair & Overhaul
RUL – Remaining Useful Life

2. INTRODUCTION

Predictive maintenance is gaining more and more interest within multiple industries (Ref. 1), including those in the aeronautical domain (Ref. 2-5). It uses (advanced) data analytics to estimate when an equipment is going to fail so that corrective maintenance can be scheduled before the failure. Within the helicopter industry, a huge volume of (time series) data is collected during flight tests through Flight Test Instrumentation (FTI), and also from customer helicopters through the Health and Usage Monitoring Systems (HUMS) and the Flight Data Continuous Recorder (FDCR). The (joint) analysis of all these data brings considerable benefits in terms of safety enhancement, maintenance optimization, system design improvement and in-service incidents support (Ref. 6-7). As an example, Airbus Helicopters FlyScan (Ref. 8) has been recently developed as a service with the objective of proactively anticipating mechanical components degradation based on available data. It consists of data-based health indicators, monitoring different helicopter systems and raising alerts as soon as degradation symptoms are detected. This results in increased fleet availability and reduced maintenance burden, along with enhanced safety (Ref. 8).

The work presented here describes a new FlyScan health indicator which monitors the Main GearBox (MGB) lubrication and cooling system. This monitoring exploits expert knowledge, machine learning normality models built from big amounts of customer flight data and fuzzy logic for uncertainty management. To the best of our knowledge, it is the first time that such a combination is used in health monitoring of aeronautical systems.

The remainder of the paper is organized as follows. In section 2, the MGB lubrication and cooling system is presented and the problem defined. In section 3 the whole method followed to build our monitoring system is detailed. In section 4 the obtained results are presented regarding both ML models performance and the whole system accuracy obtained with real anomalies. Finally, in section 5, the results, the methodology and some of its perspectives are discussed.

3. PROBLEM DEFINITION

This section describes the monitored MGB lubrication and cooling system (Figure 1). It also provides an example of the expert anomaly root cause identification rules, and defines the anomaly detection and fault diagnosis problem.

The lubrication system of a MGB has the following main functions:

- Lubrication of parts in contact for power transmissions,
- Collection of the heat generated at each contact and its evacuation at the oil cooler level,
- Participation to "health monitoring" of the mechanical elements by collecting the potential metallic particles from degradation and to bring them to the chip detectors (element 15 of the Figure 1), and by pressure (elements 1, 2 and 3) and temperature sensors (elements 6 and 7),
- Collection of other pollutants and bringing them to the oil filter.

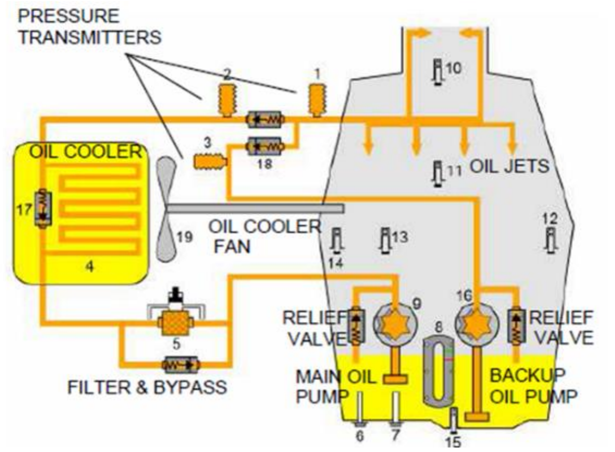

Figure 1: Monitored MGB lubrication / cooling system architecture.

For safety reasons, the MGB lubrication system is composed of 2 pressurized circuits. The first one is equipped with an oil filter and an oil cooler, and the second one is integrated to the housing to minimize the sources of failure leading to external leakages.

As it can be seen on Figure 1, the lubrication and cooling system is composed of several sub-systems, and the objective here is to detect any anomaly and anticipate any failure that can occur on each of them. For this purpose, the lubrication expert has provided a set of troubleshooting rules (anomaly root causes identification) which exploit normality states of four parameters measured on the system, and which result from the knowledge of the system, including the knowledge of its design and its testing results. These parameters are the MGB oil temperature (*MGBT*) and three oil pressures measured in different areas on the system: *MGBP* Principal (*MGBP*), *MGBP* rail (*MGBPr*) and *MGBP* backup (*MGBPb*).

- ***MGBT*** (elements 6 and 7 of the scheme) is the MGB oil sump temperature. This temperature represents the equilibrium between thermal power produced by friction inside the MGB and the cooling via heat exchanger (element 4) and the thermal exchanges by the housings.
- ***MGBPr*** is the rail pressure measured by sensor (1). It represents the pressure at the oil nozzles which deliver oil into gears and bearings.
- ***MGBP*** is the main (or principal) pressure measured by sensor (2). It represents the pressure inside the main circuit. In nominal operation, this pressure is slightly higher than *MGBPr* with the same profile as the flow from the main circuit crosses through the rail.
- ***MGBPb*** is the back-up (or emergency) circuit pressure measured by sensor (3). It represents the pressure inside the emergency circuit. In nominal condition, this pressure must be at a constant value due to lubrication system architecture. The pump (16) is directly delivering oil through its overpressure valve as the main pump is feeding the rail.

As an example of such expert rules the following:

IF *MGBP_is_high* ***AND*** *MGBPb_is_normal* ***AND*** *MGBPr_is_high* ***AND*** *MGBT_is_normal* ***THEN*** *ANOMALY_IS_MGB_JETS_CLOGGING*.

This rule states that if *MGBP* and *MGBPr* are abnormally high while the *MGBPb* and *MGBT* are normal, then there should be potentially a clogging of the MGB jets.

The provided expert rules allow identifying more than ten different root causes of anomalies related to the MGB lubrication and cooling sub-systems.

Nevertheless, the challenge is to know when a parameter state is normal, abnormally high or abnormally low. Specifically, oil temperature and pressure values, since they are depending on flight parameters and environmental conditions, such as the outside air temperature.

The solution proposed in this paper consists of building machine learning reference (normality) models using data collected from customers' helicopters. Moreover, in order to manage uncertainties related to those models, to sensors' precision and other uncertainties such as those linked to the production and manufacturing of the system components, a fuzzy logic formalism has been adopted.

4. METHODOLOGY

4.1. Overall Methodology

The overall general methodology for building data based health indicators is depicted on Figure 2.

First, there is data, and for this several data sources are considered:

- *Customer data:* it consists of helicopter data, mainly vibration and flight / usage data. It is the main data used either to build or feed health indicators.
- *Engineering data:* it can be (flight) test data, simulation data, bench test data, design data etc. It includes digital twins like physics-based health indicators and virtual sensors. It is very useful, especially when customer data does not contain the right information needed to monitor a given system.
- *Maintenance data:* it consists of Maintenance Repair & Overhaul (MRO) data; customer Maintenance Information System (MIS) data, in addition to incident data. It is useful to validate health indicators on past data. It provides installation and removal dates of different systems, the reasons for removals and the observation on removed parts (such as degradations or wears levels). It allows for example explaining and validating possible alerts when running health indicators on historical data.
- *Manufacturing data:* consists of data related to the parts themselves such as manufacturing or installation deviations of pinion tooth pitch, which could affect health indicators signatures (Ref 9). These data allow avoiding false alerts of health indicators deviations which are not due to system or mechanical failures.

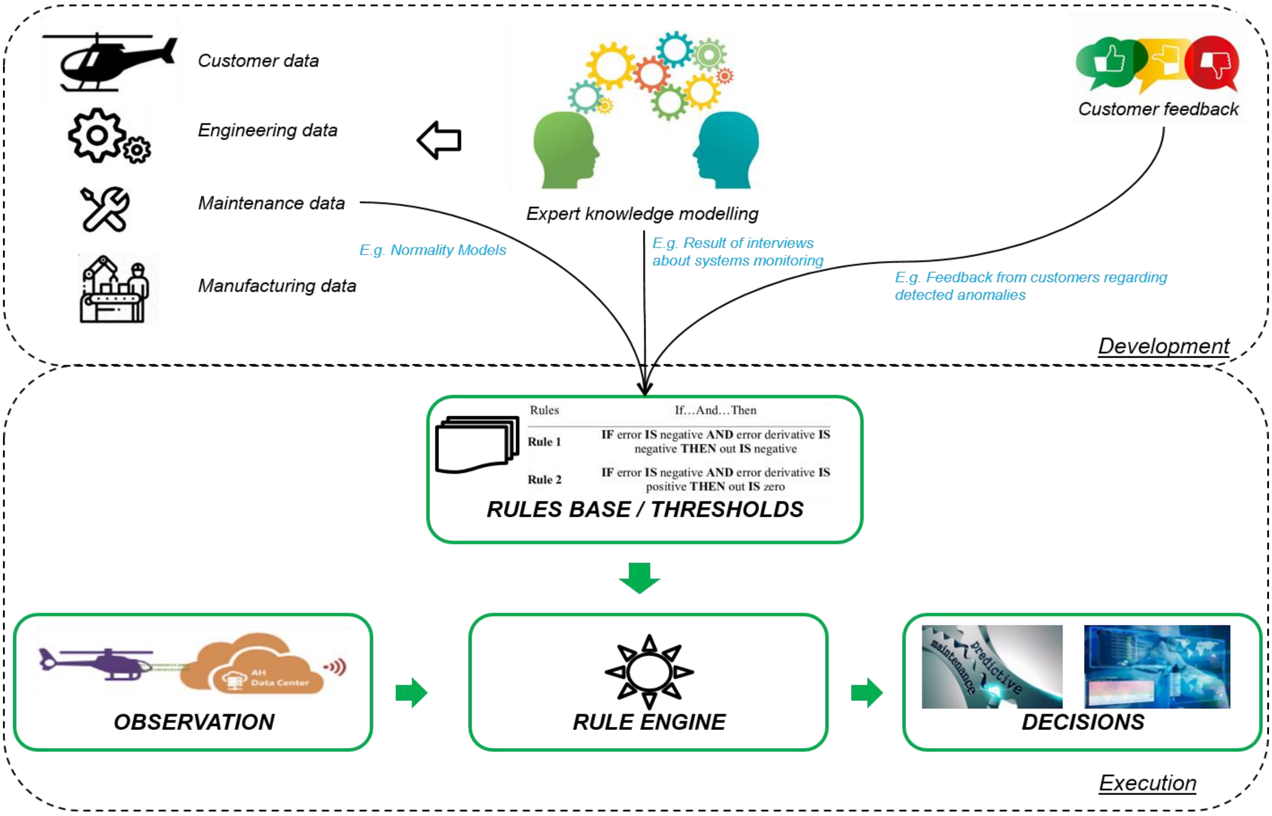

Figure 2: Methodology and Information Sources for Health Monitoring Indicators Construction.

The second element to be considered when building health indicators is the expert knowledge about the systems, the data they generate, and the way they have to be monitored. For example, looking for a particular combination of data in specific conditions could be an obvious health indicator for an expert of a system, which could be, conversely, not obvious to automatically extract from data - even using machine learning - especially when there are few anomalies to characterize abnormal behaviors of those systems.

The last element to be considered is the customer feedback, especially regarding health indicators alerts. This allows for example maturating and adjusting alerting thresholds.

Once health indicators maturated then they are deployed in production for continuous monitoring of H/C. In general, alerts raised by health indicators are validated by experts before maintenance recommendations are made to customers, in order to minimize false alerts risk. So, health indicators act often as a decision support for experts and managers.

4.2. MGB Lubrication / Cooling System Case

As said previously, the heart of our MGB lubrication and cooling system monitoring consists of the normality models for the MGB oil temperature and pressures as well as their combination with expert troubleshooting rules. So, the following sub-sections detail first how those models are built and second how they are integrated with the expert rules.

4.2.1. Normality Models Learning

The whole process for building the normality models of *MGBP*, *MGBPr* and *MGBT* is shown on Figure 3. The dependency of *MGBPb* reference values on environmental and flight parameters being negligible, normality bounds can be given based on design expectations rather than defined from observed data.

First, potential anomalies of the MGB lubrication and cooling system are filtered from the big amounts of in-service data collected and used in this study. This filtering is based on MIS data. The objective is to avoid including anomalies within reference training data and preventing their detection in the future.

Then, a number of pre-processing is applied to the 'normal' flights: 1) only parameters recommended by the expert are retained and filtered on a stable phase; 2) possible remaining anomalies (from the previous step) are partially removed using thresholds which should not be exceeded. At this stage the learning database is composed of more than 500 million data points; and 3) finally, before its normalization, data are reduced thanks to a parameter rounding method that allows lower processing / learning time and which also prevents the learning from overfitting. This method consists in rounding each considered parameter using a dedicated function taking into account the parameter range. For example, *MGBP* is rounded to one decimal place (3.67 bar rounds to 3.7 bar), outside air temperature and *MGBT* are rounded to the nearest whole number (15.4°C rounds to 15°C), and altitude rounds up or down with a maximum step of 500 ft. This way, after duplicates removal (by grouping by similar rows and averaging the target parameter values), data size is divided by ten. Figure 4 shows an example of such data reduction.

Original Data

P1	*P2*	*P3*	*T*
15.5	65.5	2600	3.12
15.6	65.2	2750	3.17
15.8	58.5	3000	4.20
17.1	58.0	3120	4.27
17.1	57.9	3200	4.29
17.4	57.7	3300	4.30
17.3	57.2	3345	4.31

Rounded Data

P1	*P2*	*P3*	*T*
15	65	2500	3.12
15	65	2500	3.17
15	58	3000	4.20
17	58	3000	4.27
17	57	3000	4.29
17	57	3000	4.30
17	57	3000	4.31

Grouped Data

P1	*P2*	*P3*	*T*
15	65	2500	3.14
15	58	3000	4.20
17	58	3000	4.27
17	57	3000	4.30

Figure 4: Rounding Based Data Reduction. *P1, P2* and *P3* are the rounded parameters. *T* is the target parameter whom values are averaged.

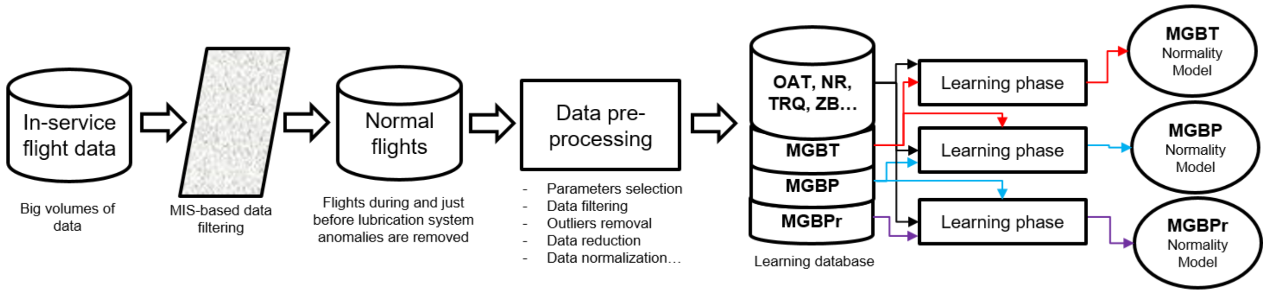

Figure 3: Learning Process of Normality Models (Data Filtering and Pre-Processing, then Machine Learning).

The ultimate step is the normality models learning from big and various data. Big in terms of volumes, and various in terms of mission diversity representation. It is worth mentioning that the *MGBT* normality model is built independently of MGB pressures. The *MGBP* model takes as input the 'normal' *MGBT* as well as a set of other flight parameters. And *MGBPr* takes 'normal' *MGBP* and other parameters as input.

4.2.2. Integration of Normality Models with Expert Rules

The normality models of the previous section produce inputs that trigger expert rules. The overall integration method is depicted on Figure 8. In fact, for each flight received from a customer helicopter the following steps are carried out:

1. The flight is filtered on stable phases specified by the expert, then the normality models predict what should be the values of each parameter of the expert rules if no anomaly exists in the system.

2. For each parameter the predicted values are compared to the measured ones (average value of the difference, Eq. (1), Figure 7), in order to determine if it is normal, abnormally low or abnormally high.

$$(1) \qquad d = \sum_{i=1}^{N} (y_i - \widehat{y_i})/N$$

d: Average value of the difference between predicted and measured values during a stable phase

y_i: Value of the measure n° i

$\widehat{y}_i$: Corresponding predicted value

N: number of data points in the stable phase

Ideally the measured values should be equal to predicted ones if no anomaly exists in the system. However, different uncertainty sources (mentioned previously) make it impossible (see statistical dispersion in the result section). In order to manage those uncertainties, fuzzy logic formalism is used (Ref. 10). This latter is well adapted for rule based systems, and it is widely adopted in the industry thanks to its efficiency and ease of implementation. So, for each parameter, values between zero and one are computed for its three possible states: normal, high or low. The higher the value is the more likely this parameter state is to be true. As an example of such fuzzy membership functions the one shown on Figures 5-6 for the *MGBP* parameter. As we can see, depending on the difference between the prediction and the measurement, each state of the parameter (green for normal, blue for low and red for high) takes a value more or less high (*y-axis*), so that relevant information is captured even in presence of uncertainties (*x-axis*). The thresholds used in these membership functions (-0.5, -0.3, 0, 0.3 and 0.5) are fixed following an analysis of few known past anomalies. As an example, the pressure is considered as fully (100%) normal if the difference between the prediction and the measurement is comprised between [-0.3 bar, 0.3 bar]. It is considered partially normal when this difference belongs to [-0.5 bar, -0.3 bar] or [0.3 bar, 0.5 bar]. And it is considered as completely abnormal if the difference is higher than 0.5 bar of lesser than -0.5 bar.

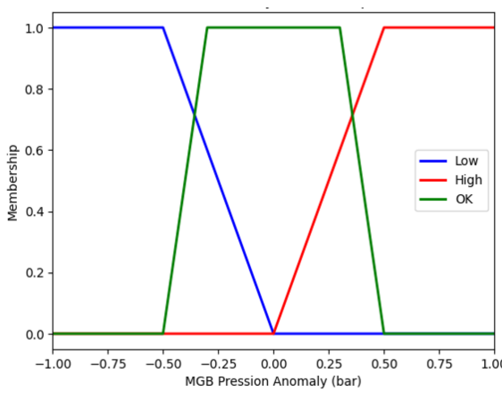

Figure 5: *MGBP* Anomaly Membership Functions (Schema).

$$\mu_{low} = \begin{cases} 1 & \text{if } x < -0.5 \\ -2x & \text{if } -0.5 \le x < 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{high} = \begin{cases} 0 & \text{if } x < 0 \\ 2x & \text{if } 0 \le x < 0.5 \\ 1 & \text{otherwise} \end{cases}$$

$$\mu_{normal} = \begin{cases} 5x + 2.5 & \text{if } -0.5 \le x < -0.3 \\ 1 & \text{if } -0.3 \le x < 0.3 \\ -5x + 2.5 & \text{if } 0.3 \le x < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Figure 6: *MGBP* Anomaly Membership Functions (Formulas).

3. All fuzzy values computed in the previous step are given as input to the fuzzy rule engine, in which the expert rules are also transformed from crisp and binary rules into fuzzy rules. The output of the engine is an anomaly risk level (normal, low or high) associated with each of the expert rules. This result is stored in the FlyScan database for continuous monitoring of in-service helicopters. In fact, after three consecutive high anomaly risks regarding the same expert rule a FlyScan alert is raised.

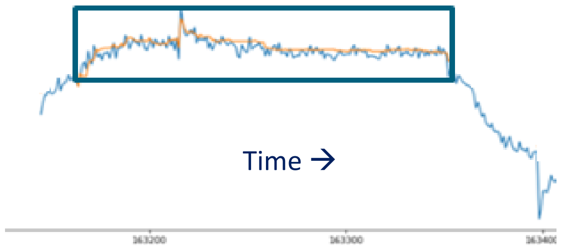

Figure 7: Example of measured (blue) vs predicted (orange) values of *MGBP* during a stable flight phase.

5. EXPERIMENTS AND RESULTS

The proposed system was implemented and tested using real data and anomalies.

First, the normality models were built using the Gradient Boosting algorithm. The choice of this algorithm is a result of a benchmarking done in (Ref. 3) as well as other benchmarking done by other industries and academics (Ref. 2,13)[1] regarding (structured) tabular and time series data. The learning phase was done using the parallelized Gradient Boosting implementation in MLLib of Spark engine[2] run on a big data cluster composed of tens of nodes on which Hadoop eco-system technologies were installed (Ref. 6). Despite the high computing performance of the cluster, the learning phase of each normality model took several hours. A restricted hyper-parameters optimization phase was performed due to the big volume of learning data considered and the resulting long processing times. A classical learning process was applied, with 80% of data used for learning, 20% for 5-folds cross-validation. 2 years separate and independent data was kept for testing. The obtained results on test data in terms of precision

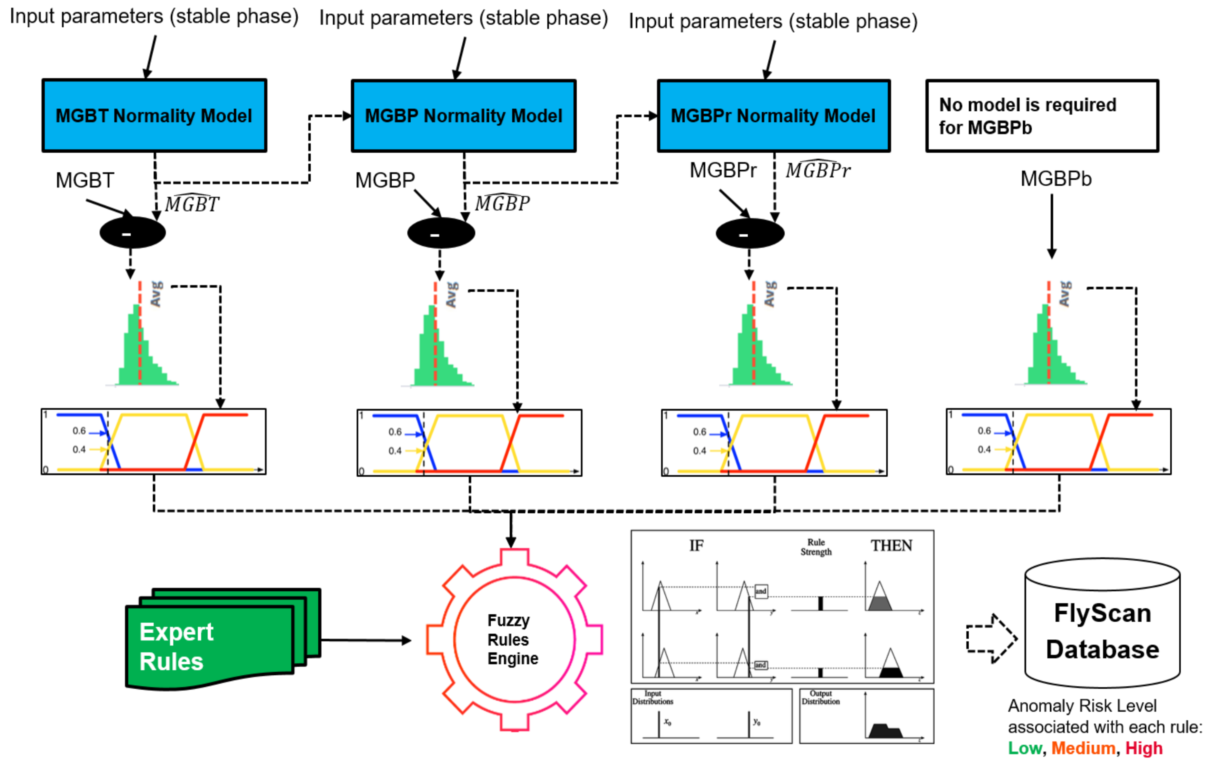

Figure 8: Normality Models Integration with Expert Rules Using Fuzzy Logic.

[1] https://mljar.com/machine-learning/neural-network-vs-xgboost/

[2] https://spark.apache.org/

and dispersion are shown on Table 1 for the three normality models. The precision is measured using the coefficient of determination, denoted *R2* (Eq. (2)). It is the proportion of the variation in the dependent variable that is predictable from the independent variable(s)[3].

$$(2) \qquad R2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

y_i*: value of the measure n° i*
$\hat{y}_i$ *: corresponding predicted value*
$\bar{y}$*: mean value of the measures*

N: number of measures

Dispersion is measured using Root Mean Square Error (*RMSE*). It shows how far predictions fall from measured true values using Euclidean distance (Eq. (3)).

$$(3) \qquad RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}$$

As it can be seen the obtained results show quite good performance of the models.

Table 1: Normality Models Performance Results: Precision vs Dispersion.

Model / Metrics	R2 (%)	RMSE
MGBP	74.4	0.27 (bar)
MGBT	85.1	3.89 (°C)
MGBPr	96.0	0.10 (bar)

The expert rules have been implemented using the Mamdani fuzzy inference system (Ref. 11). The whole system combining normality models and expert rules was tested on the flying period corresponding to the test data (2-years period). For each flight of this period the system was applied, and an anomaly level was calculated for each expert rule. In the whole period fourteen potential anomalies of the MGB lubrication and cooling system were highlighted (high anomaly level). Five were related to helicopters for which either MRO or MIS data was not collected. For the remaining ones MRO, MIS and in-service incident databases were exploited to check if there is a maintenance or an event that can explain or validate the anomaly raised by the system (Figure 9). For example, if an anomaly regarding a sub-system disappears just after a maintenance action (tagged in MIS) performed on this sub-system, then we consider that the raised anomaly as valid.

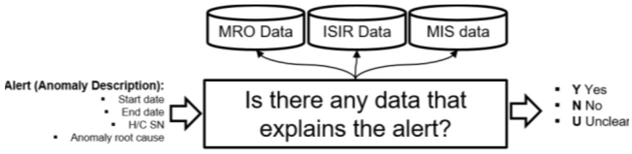

Figure 9: Data Sources Considered to Validate Anomalies Raised by the System.

The results obtained with the nine anomalies are reported on Table 2. Only one anomaly was not validated, which can be a false alert or which can correspond to a real anomaly not tagged in the considered databases. One unclear situation was encountered. In fact, the system raised an alert regarding a clogging of one of the MBG jets, and during the overhaul of this MGB one oil jet was scrapped, but the same oil jet part number was also scrapped on other helicopters without being alerted by the system. So, either it is a false alert or the clogging of the oil jet was more important for the helicopter on which the alert was raised.

Table 2: Overall System Performance: Recall versus Precision like Assessment.

#H/C	#FH	#Y	#N	#U
14	62000	7	1	1

6. DISCUSSION

This paper described an original method for the monitoring of the helicopter MGB lubrication and cooling system. It showed how ML and AI in general could be of high interest for health monitoring when combined with expert knowledge. It clearly opens new perspectives in this domain.

The proposed system is designed for one of the Airbus Helicopters MGB lubrication and cooling system architectures. However, it is extendible to any other architecture, provided that the expert specifies the corresponding troubleshooting rules.

This system relies on normality models learned from big amounts of data, thanks to latest big data

[3] https://en.wikipedia.org/wiki/Coefficient_of_determination

technologies that offer high performance parallelized computing capabilities. It is often better to learn from more and bigger data in order to better explore the operation domain of a system. But, even with big data technologies, there is still some limitations for doing that. First, hyper-parameters optimization on big data takes a very long time, even when they are restricted to few values. It is very difficult or even impossible to explore a large number of possibilities that lead to a real optimization of hyper-parameters and then to a better optimized normality models. Another way of learning normality models is to do it on sampled data. This consists in selecting 'small' and representative subset of data from the original big dataset, and then using classical ML libraries such as scikit-learn[4] to learn a model. The challenge here is to get the best representative data sample which allows the learned model to generalize to all data. The advantage of sampling, in turn, is the possibility to explore more possibilities for optimizing hyper-parameters and also to test more ML algorithms - including deep learning approaches - which are not yet parallelized in big data processing engines such as Spark. Different sampling strategies have been developed in the literature such as random sampling, stratified sampling, cluster sampling etc. This has been slightly tested in the work presented here. A subset of data (~1 million data points) was sampled thanks to Spark sampling function, then this data was used to learn our normality models after a tuned hyper-parameters optimization. The obtained models got higher precision and lesser dispersion than those built using Spark MLLib, but they did not allow obtaining better results when integrated to the whole system and tested against real anomalies. Definitely, learning on big data versus sampled data still is an open question which is worth exploring.

The data considered for normality models construction was restricted to a stable phase specified by the expert, in order for example to limit the effect of parameters that are not recorded. Actually, this phase covers the majority of the flights, but it is worth considering all the data points of a flight in order to be sure that every flight will be monitored. Further research should be done for that, especially for establishing the good features engineering and testing more complex algorithms that will better capture evolution of each parameter during transition phases.

[4] https://scikit-learn.org/stable/

The proposed system allows accurately detecting symptoms of failures at an early stage, but it does not yet allow estimating the Remaining Useful Life (RUL) of a component which is going to fail. RUL is probably the most difficult aspect in health monitoring (Ref. 12), but it allows a better optimization of customer operations. Consequently, it is important to work on this aspect in a future version of the system.

Last but not least, in order for AI based systems to be accepted by the end users and authorities, it is important to develop means that ensure their generalization, stability and robustness. This will also contribute to reducing the number of false positives and non-detections of health indicators, and open new perspectives for using AI in critical applications requiring certification such as maintenance credit.

7. ACKNOWLEDGMENTS

We are grateful to Shanta SHASTRI for her contribution to the evaluation of the normality models performance and Valerio CAMERINI for his review of the article.

8. REFERENCES

1. Hermansa, M., Kozielski, M., Michalak, M., Szczyrba, K., Wróbel, L., Sikora, M., Sensor-Based Predictive Maintenance with Reduction of False Alarms - A Case Study in Heavy Industry, Sensors 22(1): 226 (2022).

2. Pavao, A., Guyon, I., Stéphane, N., Lebeau, F., Ghienne, M., et al. Aircraft Numerical "Twin": A Time Series Regression Competition, ICMLA, 20th IEEE International Conference on Machine Learning and Applications, Pasadena / Virtual, United States, Dec 2021.

3. Daouayry, N., Mechouche, A., Maisonneuve, PL., Scuturici, VM., Petit, JM., Data-centric helicopter failure anticipation: The mgb oil pressure virtual sensor case, IEEE International Conference on Big Data (Big Data), 2019, 1784-1793.

4. Bartolone, A., Trujillo, A., Glass-Cockpit Pilot Subjective Ratings of Predictive Information, Collocation, and Mission Status Graphics: An

Analysis and Summary of the Future Focus of Flight Deck Research Survey, NASA, 2002.

5. Ranasinghe, K., Sabatini, R., Gardi, A., Bijjahalli, S., Kapoor, R., Fahey, T., Thangavel, K., Advances in Integrated System Health Management for mission-essential and safety-critical aerospace applications, Progress in Aerospace Sciences, Volume 128, 2022.

6. Mechouche, A., Daouayry, N., Camerini, V., Helicopter big data processing and predictive analytics: Feedback and perspectives, Proceedings of the 45th European Rotorcraft Forum, Warsaw, Poland, 2019.

7. Mechouche, A., Helicopter Data Science: From Concepts to Applications, Proceedings of the European Rotorcraft Forum 47th Forum, Virtual, 2021.

8. Boutaleb, A., Diaz, A., HUMS Proactive Analysis for Predictive Maintenance, 48th European Rotorcraft Forum, Winterthur, Switzerland, 2022.

9. Camerini, V., Aubin, V., Mechouche, A., Helicopter Gearbox Mechanical Classification based on Vibration Pattern Recognition, Proceedings of the Vertical Flight Society 78th Forum, Fort Worth, U.S.A.

10. Zadeh, L. A., “Fuzzy Sets”, Information and Control, 8(3): 338–353, 1965.

11. Mamdani, E.H., Assilian, S., An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. International Journal of Man-Machine Studies, 7, 1-13, 1975.

12. Jianhai, Y., Zhen H., Shuguang H., A deep learning framework for sensor-equipped machine health indicator construction and remaining useful life prediction, Computers & Industrial Engineering, Volume 172, Part A, 2022.

13. Grinsztajn, L., Oyallon, E., Varoquaux, G., Why do tree-based models still outperform deep learning on tabular data? 36th Conference on Neural Information Processing Systems – NeurIPS - Track on Datasets and Benchmarks, 2022.

CLASSIFICATION & COPYRIGHT