Sensor fault detection using machine learning applied on acoustic test bench

Thibaut Le Magueresse(1), Jérémie Derré(2), Florent Mercat(2)

(1) Amiral Technologies, 12 rue Ampère, 38000 Grenoble, France
(2) Airbus Operations S.A.S., Acoustics Testing Team, 316 route de Bayonne, 31060 Toulouse Cedex 09, France

1. Introduction

This paper presents a study of sensor fault detection from an acoustic test bench, performed by machine learning. The concerned rig is based on the modal generation and detection principle, aiming at characterizing the acoustic properties of engine liners. Such a test mean is instrumented with a large number of sensors (i.e. more than a hundred and seventy sensors), and used in the frame of heavy experimental campaign with significant test matrices. In that context, it is of prime interest to ensure that all the sensors are behaving as expected, and to validate the current test point, before moving forward and change the configuration. Data validation becomes therefore fundamental, as the causes of failure are numerous in such a complex environment, and automatic processing avoids time losses, especially regarding the huge data quantity. The proposed approach lies in machine learning software, whose inputs are the temporal raw data, which therefore do not need to be pre-proceeded. First comparisons between the AI and human validations are presented in this paper.

2. ModSquare acoustic test bench presentation

The ModSquare test bench is dedicated to the characterization of aircraft engine liners thanks to the acoustic modal detection method [1,2]. A plane source constituted by 56 loudspeakers generates higher order acoustic transverse modes in a straight duct with rectangular cross-section (see Figure 1). This noise source is then characterized thanks to a first microphones antenna constituted by 114 microphones allowing modal detection and waves sorting between acoustic waves propagating forward and backward. Next section of the test bench includes the acoustic liner to be tested, placed on a wall of the duct. Finally, a second modal detection antenna of 57 microphones characterizes the acoustic field after the test section. This second modal detection antenna does not perform waves sorting as an anechoic ending allows to free from any reflexion.

In the frame of this study, the test section consists in a hardwall section, and the test matrix consists in 118 measurements of 5 seconds each, each measurement corresponding to one generated mode at source. Acquisition is performed with the 171 microphones at 16384 points per second, generating a total of about 20 Go of time-data for a full test matrix.

Three test matrices of 118 measurements each were used in this study, in order to assess the performance of failure detection model, each matrix being used to generate a database respectively for model training, model validation and model test.

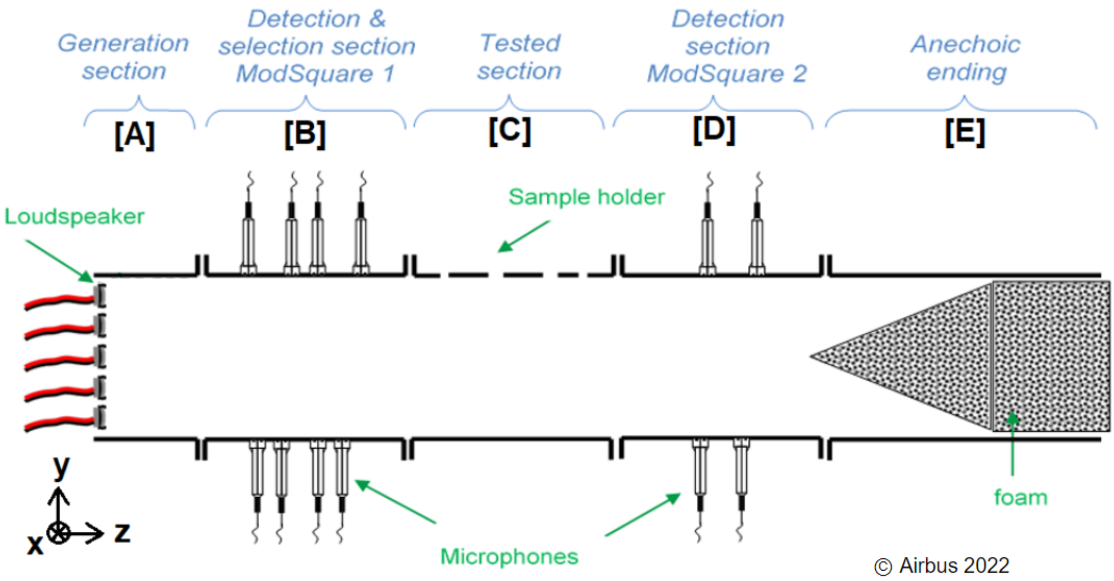

Figure 1. Sketch of the ModSquare test bench (rectangular section duct), with 171 flush-mounted microphones. Acoustic liner to be tested is excited thanks to loudspeakers antenna located on the left.

3. Machine learning model

The machine learning model has been built by DiagFit®, a blind failure prediction software for industrial equipment edited by Amiral Technologies [3]. It consists in a sequence of modules containing a feature extractor, a blind failure detection model and a failure identification model. The blind failure detection model is learnt using the healthy data free from any faults. Coupled with the feature extractor, it detects any deviation from the normality without need of failure history. The failure identification model has the responsibility to classify the anomaly. If the fault occurs for the first time, the model should declare it as novelty.

Four types of failure have been tagged in the training data: an electrical noise characterized by a Dirac comb, a wideband noise, a surrounding noise generated by impulsive events around the bench and instabilities caused by loud-speaker defect.

4. Modal training and validation

The blind failure detection model has been trained on healthy data. At the same time, the tagged data have been exploited to build the failure identification model. In both cases, only one percent of the total amount of the training matrix have been used. This has been possible thanks to the data redundancy and the relevance of the DiagFit® feature extractor. Figure 2 depicts the concatenated time series corresponding to the training and validation sets over time. The validation of the model has been carried out by training the model on the first half of this subset and by validating it on its second half.

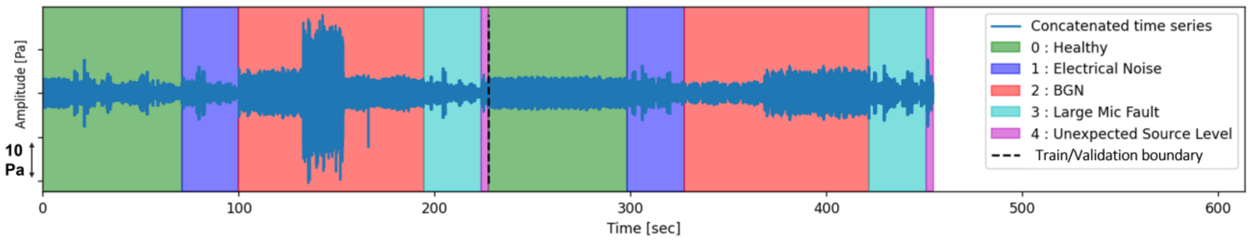

Figure 2: Time series corresponding to training data (around 7 million samples in total). In green, the representative samples of healthy data and successively in mauve, red, blue and purple the defects linked to electrical noise, impulsive noise, wideband noise and loudspeaker instabilities. The dashed black line represents the train/validation separation line.

The model has been applied on the second half part of the dataset presented in Figure 2. Figure 3 presents the results produced by DiagFit® including two types of output. First, the health indicator indicates the presence of failure in the data when it is superior to a certain threshold. The red zone represents the true presence of the failure. Except two punctual events located between 205 and 300 seconds, the detection is nearly perfect. This first output is crucial in order to detect new failures never seen in the past. Second, the model succeeded to identify the failure with a metrics of F1-Score [4] equal to 0.96. The learning time took 20 seconds, enough to learn the behavior of 20 Go of data.

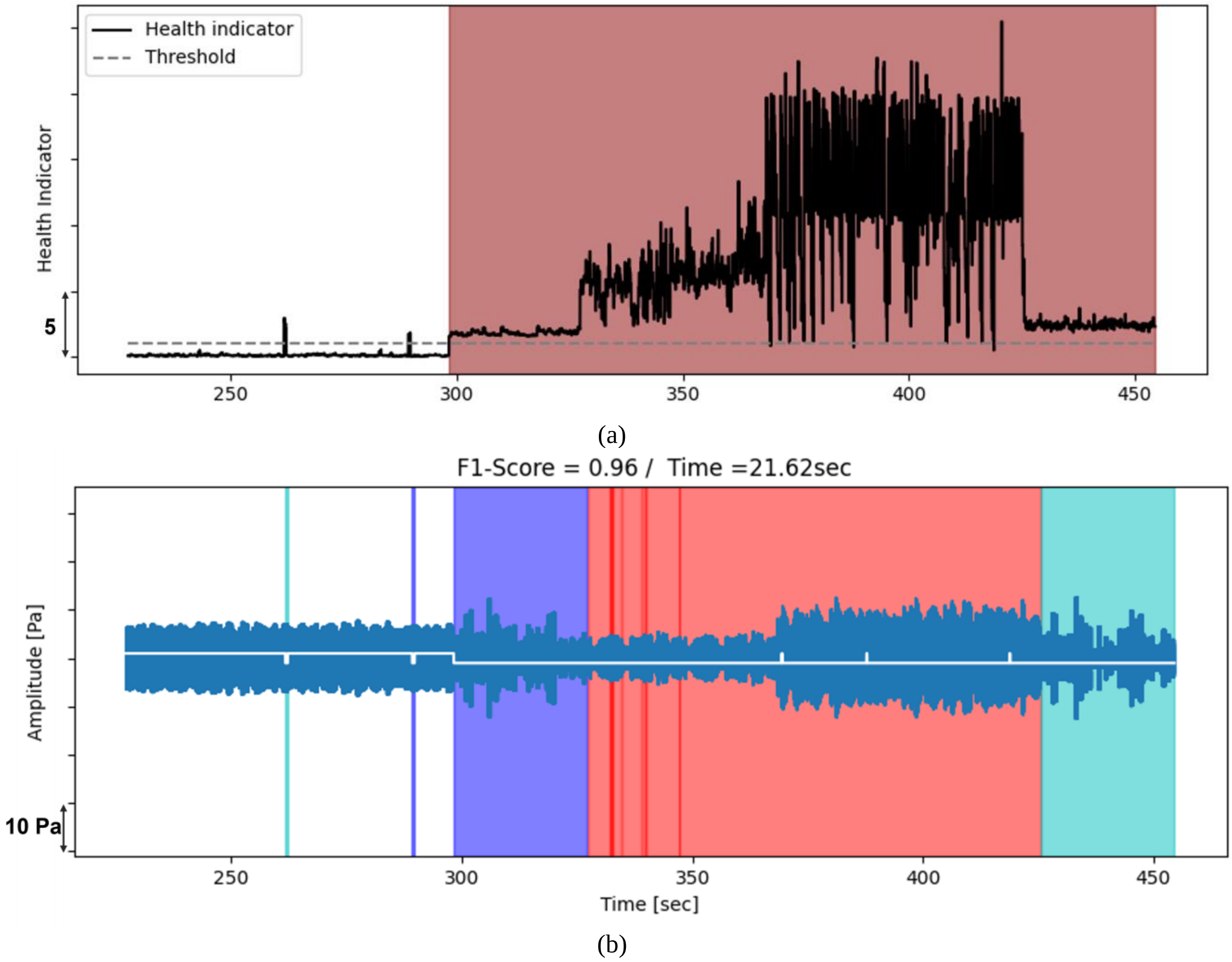

Figure 3: Blind failure prediction as a health indicator (a) and classifier prediction for the failure identification (b). The red zone of the first graph indicates the presence of a real fault. The colors of the second graph follow the same colormap as the Figure 2.

5. Results and discussion

The learnt model has been applied to the test matrix. An example of results corresponding to a measurement of 5 seconds extracted from the test dataset is presented in Figure 4. The model prediction is globally in accordance with the human defect identification. In details, one can observe that the large microphone default (represented by a yellow horizontal line) is not tagged and predicted on the same sensor. After checking, it turns out that the model did the right prediction. Moreover, the model can predict with precision the intermittence presence of the default (see the blue top line), when the human tag is more global. Finally, the model found a new defect at the very beginning of the record, which has been validated thereafter.

DiagFit® model detected automatically with reliability the presence of failures in the test dataset. In addition, it raised new defects which were not tagged, but nevertheless very real ones. It is indeed difficult for a human to be exhaustive in the manual failure detection in such a large dataset.

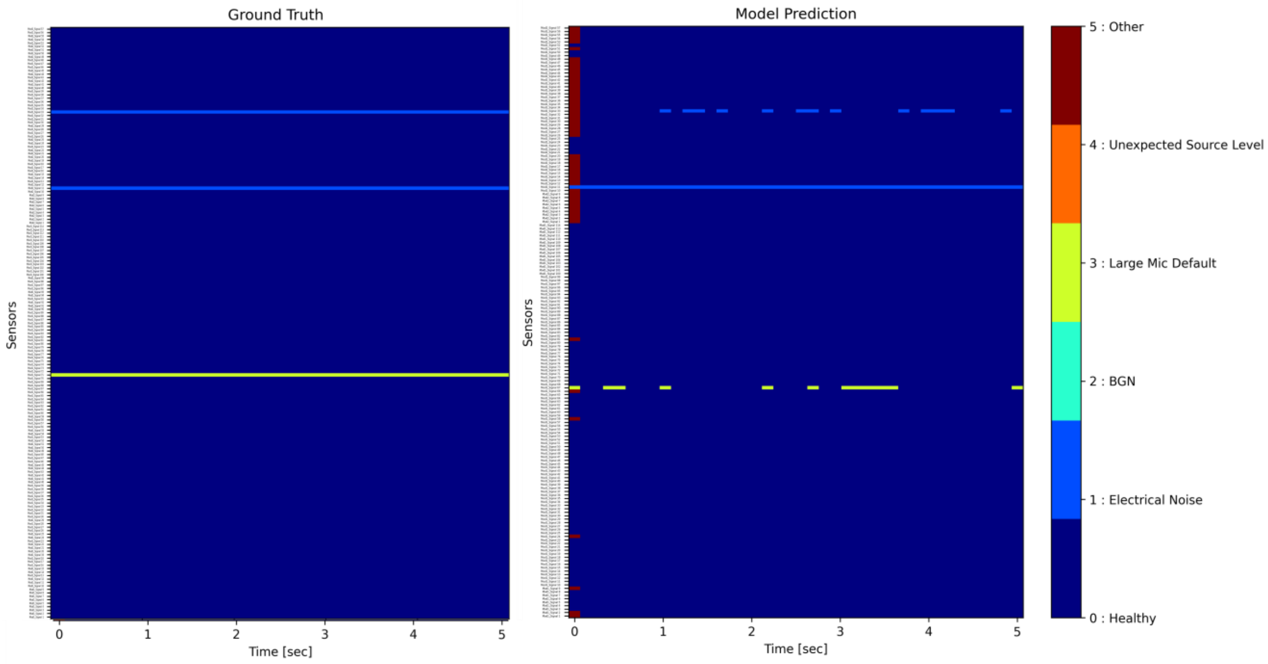

Figure 4: Result comparison between the human identification (left) and model prediction (right). Colors indicate the type of failure to detect over time (x-axis) for the whole 171 microphones (y-axis) for 5 seconds measurement.

6. Conclusion

This study presents the results of a sensor defect detection performed by machine learning and applied to an experimental acoustic test bench. The approach is based on a first phase of learning over manually labelled data, and then the learned model is applied to raw data, which are directly the acquired temporal data in the rig. A comparison between human defect identification and machine learning prediction is proposed. It exhibits interesting information, as for instance the fact that a sensor defect is not necessarily present during all the record time but rather by intermittence. It also permits to limit human error, especially in such an industrial context with large quantity of data.

7. Acknowledgments

This study has been partially realized in the frame of the MAMBO project (sous-flux 1.2), financed by the DGAC (Direction Générale de l'Aviation Civile, the French civil aviation authority) under the Convention 2021-50. The second author wants to thank Marcelle Gomes for her contribution on this topic.

8. References

1. J. DERRÉ, F. MERCAT, O. DEILLE, E. JULLIARD, « Validation of the instrumentation and measurement of an acoustic liner test rig by a combined Frequency and Modal approach», 28th AIAA/CEAS Aeroacoustics 2022 Conference (2022)
2. J. DERRÉ, F. MERCAT, O. DEILLE, E. JULLIARD, « Multi-criteria approach for instrumentation and measurement validation: application to a fully rigid duct with modal detection», 16eme Congrès Français d'Acoustique, Marseille (2022)
3. S. LE GALL, T. LE MAGUERESSE, « Benchmark : Comparison of failure prediction models », White paper, Amiral Technologies (2021)
4. A. GÉRON, Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow, Dunod, 2019.