GAS LEAK DETECTION USING ACOUSTICS AND ARTIFICIAL INTELLIGENCE

Masson, F.[1], Martelet, Y.[2], Preuilh, J.[3], Houssin, D.[4] and Watremez, X.[5]
[1] Acoem, Metravib Engineering, 200 Chemin des Ormeaux, Limonest, 69760, France, florent.masson@acoem.com
[2] Acoem, Metravib Engineering, 200 Chemin des Ormeaux, Limonest, 69760, France, yann.martelet@acoem.com
[3] Acoem, R&D, 200 Chemin des Ormeaux, Limonest, 69760, France, julien.preuilh@acoem.com
[4] Air Liquide Research and Development, Innovation Campus Paris, 1 chemin de la Porte des Loges, 78354 Les Loges-en-Josas, France, deborah.houssin@airliquide.com
[5] TotalEnergies, One Tech – Research & Development, CSTJF – Avenue Larribau, Pau Cedex, 64018, France, xavier.watremez@totalenergies.com

ABSTRACT

Gas leak detection on a production site is a major challenge for the safety and health of workers, for environmental considerations and from an economic point of view. In addition, flammable gas leaks are a safety risk because if ignited, they can cause serious fires or explosions. For these reasons, Acoem Metravib in collaboration with TotalEnergies One Tech R&D Safety has developed for the past four years a system called AGLED for the early detection, localization and classification of such leaks exploiting acoustics and artificial intelligence driven by physics. Numerous tests have been conducted on a theater representative of gas production facilities created by TotalEnergies in Lacq (France) to build a robust learning database of leaks varying in flowrates, exhaust diameters and also types (hole, nozzle, flange...). Moreover, to limit the number of false alarms, a relearning strategy has been implemented to handle unexpected disturbances (wildlife, human activities, meteorological events...). The presented paper describes the global architecture of the system from noise acquisition to the gas leak probability and coordinates. It gives a more in-depth look at the relearning algorithm and its performance in various environments. Finally, thanks to a complementary collaboration with Air Liquide, an example of test campaign in a real industrial environment is presented with an emphasis on the improvement obtained through relearning.

1.0 INTRODUCTION

As part of a project on major risks prevention, TotalEnergies searches technologies and expertise to detect, localize and quantify potential gas leaks on facilities. In plants with pressurized flammable gases, it is relevant to early detect gas leaks to prevent build-up of dangerous concentrations. Accumulation of gas can lead to explosions and fires, which are the main major accidents. TotalEnergies has created a theater representative of these facilities (TADI - TotalEnergies Anomaly Detection Initiatives) serving as technological showcase. On these facts and with its relevant experience in acoustic products and services, Acoem Metravib proposed to develop a system based on acoustics to achieve this challenge, with a particularly focus on gas leak detection and localization.

Acoustical system, because of its presence even in places hostile to man and its objectivity, enables to cover a wider perimeter while limiting the human factor. This technology is different from conventional and commonly used detectors, especially for hydrogen detectors [1]. It could bring a real aid to the exploitation by preventing major accident. Currently, human still detects close to half of gas leaks in industrial environments as the conventional gas detectors, widely used, require gas to enter a fixed-point gas or pass through the path of a linear detector. These existing technologies can be unreliable in open areas especially due to wind direction. The possible non-detection of gas leaks constitutes one of the major limits of these current monitoring strategies. According to the analysis of 7 years of data relating to hydrocarbon major and significant releases on offshore oil and gas installations in the United Kingdom, an effective detection rate of approximately 45% has been recorded [2].

This paper does not intend to make a review of gas leak detection by acoustic and signal processing as Adnan et al. did [3]. Current existing commercial systems based on acoustic for gas leak detection are mostly using ultrasound technology. They can be made of autonomous punctual sensors without the capability of localize a gas leak. As reported by end-users (as TotalEnergies and Air Liquide R&D teams), these systems produce a lot of false alarms and are not always reliable. Some devices use acoustic imaging with ultrasound microphones to localize a gas leak but these systems need human action to scrutinize the soundscape.

In 2018, Acoem Metravib developed a demonstrator taking advantage of some technological bricks of its innovation portfolio with adaptions to gas leak detection and localization. This demonstrator was based on standard microphones (i.e. not ultrasound microphones, 20-20.000 Hz microphones) allowing a wider directivity for sound recording. It proposed a different approach from existing solutions for gas leak detection based on acoustic. Tests conducted on TADI platform in 2018 showed promising results with a satisfying localization and detection of leaks generated through various scenarios and types of gases (CH_4, N_2, CO_2). On the basis of signals recorded during the test campaigns, statistical processing tools (clustering) and neural network were developed to improve the system efficiency and accuracy. In 2019, Metravib proposed to complete developments performed in 2018 to introduce an industrial prototype. To this end, a dedicated prototype of antenna has been developed and constructed, associated to a specific monitoring system (HMI, services, NTP server, etc.) in order to be able to localize the position of a gas leak thanks to a network of several antennas. Localization and identification performances of this dedicated system have been demonstrated in representative gas leak conditions, on TADI platform, on a long-term period (minimum four years). In 2020, Acoem Metravib managed to strengthen the hardware and to improve the performance of the system by reducing the detection time to 12 s, by improving the localization precision, and by dropping the false alarm rate dramatically through the introduction of the relearning process. The elevation was also introduced in the localization algorithm, allowing the 3D localization of gas leaks with an overall error of less than 5 m. Finally, a first version of the gas leak classification was also developed in order to predict the magnitude order of the detected leak with 3 classes according to gas leak flow rates.

This report first describes the full computation chain from the noise recording to the leak prediction. Then the relearning process used for the reduction of the false alarm rate is described. Finally, a test campaign performed in an Air Liquide production site is presented and the results are discussed.

2.0 DESCRIPTION OF THE FULL PROCESS

For confidentiality reasons, the exact process cannot be described. The full deployment of the solution can be summarized in two groups. First, several 4-microphone antennas (in the audible frequency domain) are installed and listen to their surroundings. Then, one central unit (server) gather the data given by the antennas and merge them to communicate the results of fusion to an operator. This overall process is described in Fig. 1 representing the architecture of the AGLED (Acoustic Gas Leak Early Detection) system.

The first part of the analysis happens in the antennas. Each twelve seconds, a recording of a few seconds is made. A processing step is launched and computes more than 50 acoustic features from this file. These are time and frequency descriptors of the noise. These features fed to a neural network trained from a large database that give a gas leak detection probability. A value close to one being a gas leak and a value zero being anything but a gas leak. The database is composed of more than 1400 real gas leaks (so a database of more than 5600 signals recorded by 4 antennas) varying in flowrates, type of gas (mostly CH_4, CO_2, N_2), distance, gas leaks exhaust diameters and also types (hole, nozzle, flange...). Thanks to this large amount of gas leak sounds in the database, the neural network is enough trained in order to detect any type of gas leaks that would create a sound not masked by the soundscape. In parallel, thanks to the four microphones, the direction of arrival of the noise is calculated. This data is then sent to the central unit.

When the computer receives the analysis from all the antennas, it computes a feature called Confidence Index (CI) that will average the probabilities from each antenna with several other time-related treatments to ensure a good reactivity but also to limit the presence of one-time false alarms. The localization information is also gathered to pinpoint the estimated leak location in a specific monitoring area and a feature called Cluster Density (CD) is computed to evaluate the geometrical stability of the estimated position. Several zones of monitoring composed by independent antennas can be configured with the server. At least two antennas are needed to detect a gas leak to compute a fusion. No fusion is authorized out of a monitoring zone.

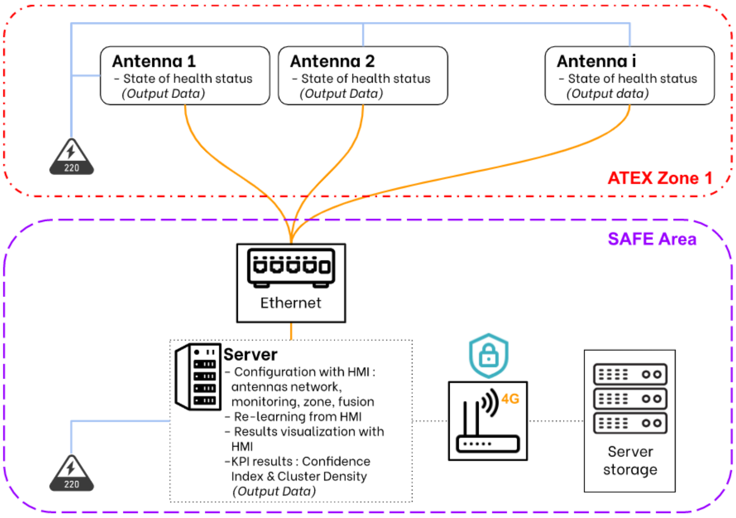

Figure 1. AGLED system architecture

3.0 RELEARNING PROCESS

It is impossible to build a noise database comprehensive enough so that the neural network can learn all the different non-leak noises that can be heard on any industrial site. From one site to another or even from one zone to another in the same site, the local acoustic landscape can be very different. That means that it is nearly impossible to have a neural network that will have a low false alarm rate because there will always be a type of noise that has no close match in the learning database. Therefore, it is very important to tailor the database to a new industrial site where antennas will be installed. Moreover, in machine learning the representativeness of the learning database is crucial so that new samples can be correctly classified. In that regard, the best way to build the database is to include noise recordings directly gathered from the location where the antennas are installed.

A relearning algorithm has been developed to cover that aspect. The global description of this algorithm is represented in Fig. 2. A main neural network is produced from a generic database of gas leaks and various non-leak noises. This neural network will at some point encounters a new noise source that is wrongly classified. The algorithm consists in substituting a portion of the initial database with the new data and reproducing a new neural network. The data is in a way substituted and not added to the initial database to keep the same leak/non-leak ratio.

The loss of some of the initial data is considered negligible with respect to the amount of data present. The new neural network is then tested to ensure both the non-regression on the leak detection capabilities and correct classification of the new noise source. This process can be repeated each time a new noise source disrupts the gas leak monitoring producing false alarms.

An example of scoring for a new site installation (site of Section 4.1) is shown in Fig. 3, with for each plot the old neural network in dashed orange and the new one in blue. The result of this validation process has to be validated by human.

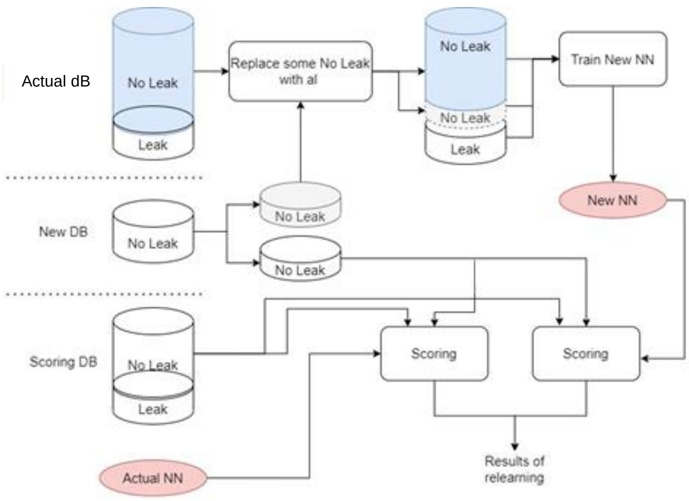

Figure 2. Description of the relearning process

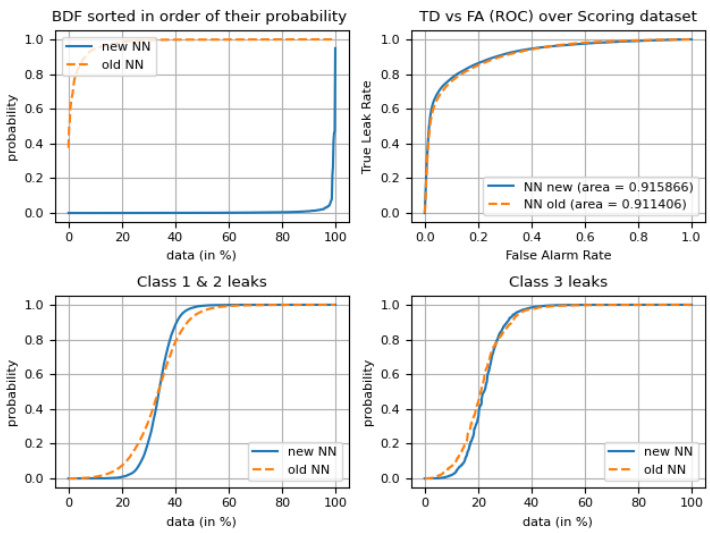

Figure 3. Example of the scoring of a relearning

The exact nature of the data cannot be disclosed but in terms of quantity, less than 15 minutes of audio was used in that relearning, split in two for both the learning and the scoring parts. The top left plot gives the probability associated with the new data sorted in ascending order. In that case, with one being a gas leak and zero being everything else, the old neural network in dashed orange does not understand the new data and incorrectly identifies it as a leak. The new neural network, in blue, however correctly predicts almost all the new data as non-leak with a threshold of 0.8 and will not ever give false alarms for similar events. The three other plots compare the performance of both old and new neural networks against a scoring database. The top right plot shows the ROC (Receiver Operating Characteristic) curves for the full database composed of non-leak noises and leaks with a mass flowrate range from 0 to hundreds of grams per seconds. This database amounts for more than 50 hours of data. The rule of thumb of a ROC curve is that the best performance is obtained when the area under the curve is maximum, i.e. close to one. In that case, both neural network have equivalent performance on the scoring database.

The two bottom row plots show the probability obtained with both neural networks but using only gas leaks. The bottom left plot uses class 1 and 2 leaks (<10 g/s) and the bottom right only class 3 leaks (>10 g/s). Here again, the relearning has little impact on the leak detection capabilities of the three classes but highly improves the behavior of the system in the new environment. In that example, the performances are the same as before the relearning with a much better handling of the false alarms that were encountered. The new neural network can be substituted for the old one for better overall behavior.

4.0 GAS LEAK DETECTION IN AN INDUSTRIAL TEST SITE

4.1 Test site presentation

One of the industrial pilots used in this project is the Air Liquide oxygen production plant in Pierrelatte (France) showed in Fig. 4. In terms of noise level, near equipment one could measure more than 70 dB between 2 kHz and 20 kHz, the typical frequency range of gas leak noises in the audible spectrum. Four antennas were installed on that site.

(a) (b)

Figure 4. (a) Monitored part of the Air Liquide oxygen plant in Pierrelatte – (b) Antenna 10

To ensure the integrity of the devices, no gas leak with explosive gas were created on site the day of a test campaign. Instead, pressurized air cylinders were used (Fig. 5). A regulator controlled the air flowrate coming out from pressurized cylinders. The gas leak program is given in Table 1 and a map of the site with the position of the leaks (blue dots) and antennas (red crosses) are given in Fig. 6. The squares on the map represent a scale of 5 x 5 m. The red cloud represents a small permanent gas leak in the installation that could disturb the system not detected before the test campaign. Mainly class 2 leaks were performed, meaning gas leaks with a flowrate between 1 g/s and 10 g/s with a focus on the localization of the leak. Indeed, the objective was for the AGLED system to study the impact of sound masking by local equipment on both the detection and localization and the limits of detectability with respect to the fairly high background noise. The monitoring zone 1 included the 4 antennas (Fig. 6) were of almost 45 x 25 m.

As pressurized air was used, it is important to keep in mind that the leaks created during the test campaign are generally less noisy than the gas leaks expected to be detected with this type of system based on acoustics. All other things being equal, the type of gas does not directly interfere on gas leaks sound signatures, but rather on sound power.

Figure 5. Air cylinders used during the tests for leak simulation

Table 1. Description of the leaks performed during the tests

Test Number	Release Diameter (mm)	Release Pressure (bars)	Release flowrate (g/s)	Release Location	Release Start	Release Stop
1	1	10	2.48	A	16:53:16	16:53:56
2	1	20	3.51	A	16:53:56	16:55:56
3	1	30	4.29	A	16:55:56	16:56:56
4	1	30	4.29	D	17:01:26	17:02:26
5	1	20	3.51	D	17:02:38	17:03:38
6	1	10	2.48	D	17:03:50	17:04:38
7	1	30	4.29	C	17:06:26	17:07:26
8	1	20	3.51	C	17:07:38	17:08:38
9	1	10	2.48	C	17:08:50	17:09:50
10	1	30	4.29	E	17:25:26	17:26:56
11	1	20	3.51	E	17:27:10	17:28:10
12	1	10	2.48	E	17:28:26	17:29:26
13	1	30	4.29	F	17:31:26	17:32:38
14	1	20	3.51	F	17:32:58	17:33:56
15	1	10	2.48	F	17:34:10	17:35:11

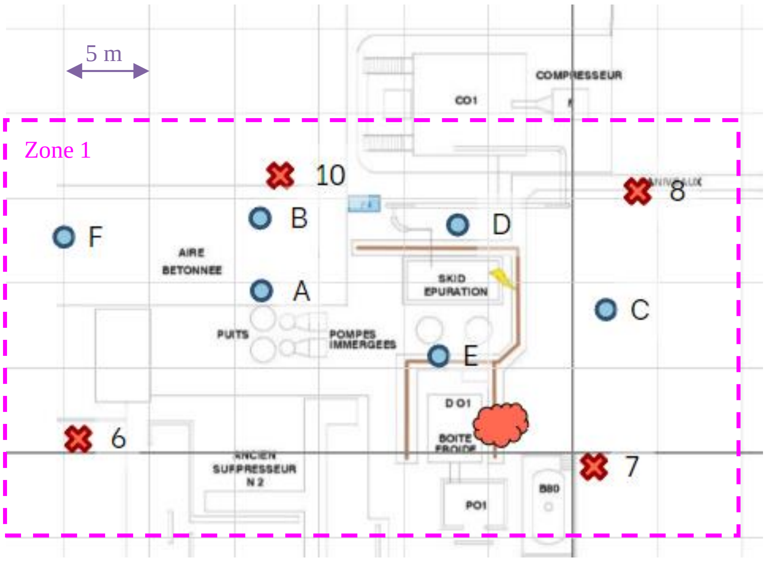

Figure 6: Map of the site with location of antennas and the release points

In terms of background noise and congestion, the central area near points D and E is the most complicated. This area is shown in Fig. 4. Poorer results are expected for leaks performed there. Moreover, the point F - far from the antenna and even further from the production equipment - was chosen to test the limits of detectability.

4.2 Test campaign results

Three neural networks are compared. The first (NN1.1) consists in a neural network based on a neural network NN1 with a relearning performed a few days prior to the tests, when a gas leak could be heard but was not detected before the test (red cloud in Fig. 6). This neural network was the one used the day of the tests. The second one (NN1.2) is also based on the neural network NN1 but with a relearning of data on the day of the tests after the gas leak was patched. Finally a third neural network (NN2.2) is used. It is based on a different neural network (NN2) that has a more complete learning database with industrial sounds, as well as leaks superimposed with such sounds, improved with relearning. This neural network comes from a data augmentation work not explained here but allowing a better behavior of the system in an industrial installation noisiest than TADI. The same data as for NN1.2 was used for the relearning. The results using the base neural networks NN1 and NN2 are not shown here because in rich environments such as the Air Liquide Pierrelatte plant, the probabilities are always close to one before the first relearning, as discussed below.

Fig. 7 shows the CI the day the neural network NN1.1 is introduced. First, the NN1 is in place, it continuously gives a CI close to one, showing that it does not understand the environment. The background noise is closer to the recordings of leaks present in its database than any other non-leak noise. A relearning is done and deployed onsite at 7 am. The CI drops immediately below 0.2 because now NN1.1 can link the new recordings (i.e. the soundscape) to the samples that were added to its database and labelled as non-leak.

As explained in Section 2, the probability obtained by each antenna is translated into an indicator named CI from the moment a fusion occurred in the monitoring zone. The results presented hereafter in Table 2 are the confidence indexes averaged and rounded on the duration of the leak.

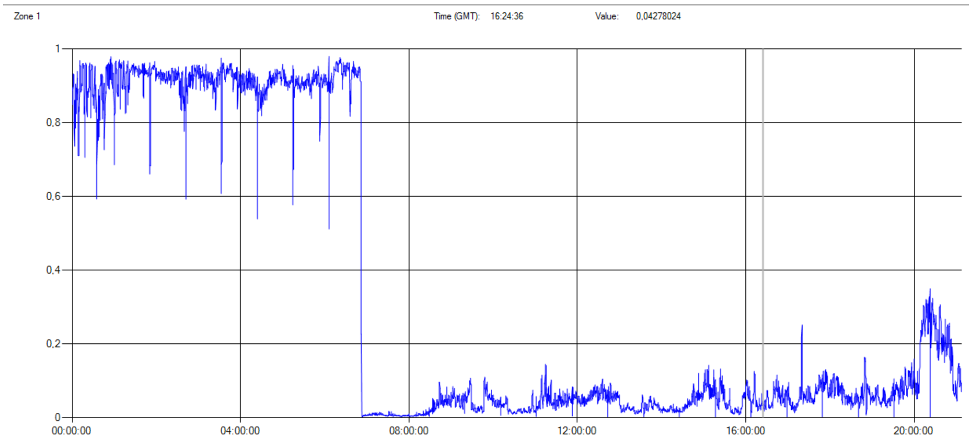

Figure 7. Confidence Index before and after relearning

During the tests, the operators identified a small leak. It was located at the red cloud near the antenna 7 in Fig. 6. This had several impacts on the results. The first one is that before the gas leak was patched, the antenna 7 could hear a real gas leak. This leak was however small enough that the data from the three other antennas were not polluted. The data from antenna 7 was then removed from the study. Moreover, for the neural network NN1.1, the relearning had falsely labelled data due to this gas leak. The impact of this is discussed below. Between gas leak 9 and 10, clean data were recorded for the relearning used in NN1.2 and NN2.2. The gas leak 10 to 15 were not estimated with NN1.1.

The first neural network (NN1.1) does not work. Even if the leaks of the test campaign were clearly audible, the CI remained at zero; the leaks were not detected. This is due to the fact that the relearning that was performed a few days before the test campaign integrated data corresponding to a gas leak noise that was added to the database and labelled as non-leak. This field test confirmed that using not well-labelled data degraded greatly the performance. It showed that it is important to be completely sure that the data used for relearning does not include any type of gas leak sound. Moreover, it highlights the need to perform a gas leak campaign with pressurized air cylinders after an antennas network installation to check that after relearning (adaptation of the system to it soundscape) the AGLED system can still detect a real gas leak.

The second neural network shows better results with a satisfying detection of all leaks for an upstream internal pressure of 30 bar, except the one at location F. This is mainly due to the distance of point F to the nearest antennas and to the geometric spreading of sound. The performance drops for the other leaks except number 8. Even though it is a step forward, the less noisy gas leaks are not properly detected. This is explained by the fact that the emergence of the leaks is not as high as for the 30 bar leaks. In this case, the recording of the leak is very similar to the data used in the relearning. Moreover, in the database used for the creation of NN1, no leak superimposed with industrial noises were present. This configuration is then new for the neural network and could explain the low values of gas leak detection probabilities and of confidence indexes for such leaks.

Finally, the neural network NN2.2 is compared. The difference with the NN1.2 is that the database of the base neural network NN2 is more complete and consider gas leaks superimposed with typical industrial noise. It should be noted that no data from the Pierrelatte plant of Air Liquide was used in the development of NN2. With NN2.2, the results are even better with a near perfect detection of all leaks. Only the quieter leaks in the central area (e.g. gas leaks 6 and 12) and far from the antennas in location F (e.g. gas leaks 13, 14 and 15) are not detected. After listening and analyzing the different raw data of these gas leaks, it could be confirmed that they do not emerge from the background noise for at least two antennas. The fusion could then not be performed. We are close to physical limits for gas leak detection based on acoustic. With appropriate data, the system shows satisfying performances for leaks as low as 2.5 g/s (depending on the leak-sensor distance). For all the gas leaks detected, the accuracy in localization was under 5 m. With NN2.2 only one false alarm was obtained during the test campaign. In the following 2 months after the installation the false alarm rate have been lower than 0.5 per day.

Table 2. Confidence Indexes according to the simulated leaks

Test Number	Release Diameter (mm)	Release Pressure bars)	Release flowrate (g/s)	Release Location	NN1.1	NN1.2	NN2.2
1	1	10	2.48	A	0	0	1
2	1	20	3.51	A	0	0.4	1
3	1	30	4.29	A	0	0.7	1
4	1	30	4.29	D	0	1	1
5	1	20	3.51	D	0	0.2	1
6	1	10	2.48	D	0	0	0.1
7	1	30	4.29	C	0	1	1
8	1	20	3.51	C	0	1	1
9	1	10	2.48	C	0	0.3	1
10	1	30	4.29	E	x	0.3	1
11	1	20	3.51	E	x	0.1	0.8
12	1	10	2.48	E	x	0	0.1
13	1	30	4.29	F	x	0	0
14	1	20	3.51	F	x	0	0
15	1	10	2.48	F	x	0	0

The same test has been performed in other industrial sites showing similar results and a false alarm rate roughly lower than 0.5 false alarm per day was obtained during the following days after relearning and validation test campaigns with pressurized cylinders.

5.0 CONCLUSION

This paper presents one of the first industrial pilot for the AGLED system which was developed to detect gaseous leaks based on acoustics. The relearning process - which is an important aspect in acoustic-based detection - was presented. It showed great capabilities in reducing the false alarms. In fact, without it, the gas leak monitoring is not possible. The results obtained during the test campaign were very different from one neural network to the other. But it can all be summed up by the link between the representativeness of the data used for the learning of the neural networks against the data used during the assessment of the system. With clean data from the site and representative leak recordings on an industrial platform, the results are very encouraging even for relatively low flowrates. This study demonstrates that in addition to the background work required for the development of an innovative detection system, collaborative work with industry brings valuable information and conditions fostering the efficiency, the reliability and the resilience of such equipment. However, the capability of detection of a gas leak with such system will always depends on the sound emergence of the gas leak, of the distance of a leak to the sensors and of the background noise level. Several complementary studies should be performed to ensure the limits of detectability of the system according to those parameters.

The last point that it has to be mentioned is that Acoem Metravib succeeds in developing an ATEX version of the AGLED which is now available for specific sites having this kind of certification requirement. Some of these ATEX prototypes are currently under validation tests in other industrial site to improve the understanding of the system behavior. The same methodology of validation than the one exposed here is being used. The tests performed up to now (not showed in this paper) are giving encouraging results since the sound signatures of all the gas leaks in the actual database are enough different from the industrial site soundscapes (particularly for highest mass flowrates). Preliminary tests have confirmed that for a same type of gas leaks (mass flowrate, distance,…) H_2 gas leaks are louder than CH_4 gas leaks that are louder than N_2 gas leaks. It confirms that it gives a favorable condition for the system to detect such gas leaks that are not easy to create under safe conditions. Controlled H_2 leaks will soon be carried out on the TADI infrastructure to confirm this point.

This work was carried out in close collaboration with TotalEnergies, Air Liquide and Acoem Metravib.

REFERENCES

1. Buttner, W.J., Post, M.B., Burgess, R. and Rivkin, C., An Overview of Hydrogen Safety Sensors and Requirements, ICHS 3th, 16-18 September 2009, Ajaccio.
2. McGillivray A., Hare J., Offshore hydrocarbon releases 2001-2008, RR672 Research Report, HSE, 2008.
3. Adnan N.F. et al., Leak detection in gas pipeline by acoustic and signal processing - A review, IOP Conf. Ser.: Mater. Sci. Eng. 100 012013, 2015.