A Hybrid Machine Learning Framework for AE-Based Anomaly Detection in Pressure Vessels

Yassine Traidi
CERI Digital Systems
IMT Nord Europe
59500 Douai, France

Lala Rajaoarisoa
CERI Digital Systems
IMT Nord Europe
59500 Douai, France
lala.rajaoarisoa@imt-nord-europe.fr

Salim Chaki
CERI Materials and Processes
IMT Nord Europe
59500 Douai, France
salim.chaki@imt-nord-europe.fr

Daniel Njeugna
MAGYAR
R&D Department
21610 Fontaine-Française, France
daniel.njeugna@smff-magyar.com

Fabrice Laurent
MAGYAR
R&D Department
21610 Fontaine-Française, France
fabrice.laurent@magyar.fr

***Abstract*—In the transportation of hazardous materials, ensuring the structural integrity of pressure vessels is essential for safety and regulatory compliance. This paper proposes a hybrid machine learning framework for acoustic emission (AE)–based anomaly analysis in pressure vessels, designed to enable early detection and characterisation of internal damage under static and fatigue loading conditions. The proposed methodology follows a multi-phase workflow comprising data preparation and validation, anomaly detection using an attention-enhanced LSTM autoencoder, adaptive threshold determination, damage type classification using a Light Gradient Boosting Machine (LightGBM), and spatial localisation of anomalies through a distributed sensor network. This hybrid architecture combines unsupervised and supervised learning to address the complexity, non-linearity, and noise inherent in AE signals. Experimental results obtained from both simulated and real-world datasets demonstrate that the proposed framework can reliably identify and differentiate multiple degradation mechanisms, including matrix cracking, fibre/matrix debonding, delamination, and fibre rupture. Furthermore, the use of attention mechanisms and interpretable outputs supports explainability, making the framework suitable for predictive maintenance and informed decision-making in safety-critical applications.**

***Index Terms*—Acoustic emission, anomaly detection, anomaly classification, LSTM autoencoder, composite materials, hydrogen storage tanks**

I. INTRODUCTION

HAZARDOUS material such as hydrogen is increasingly regarded as a clean and sustainable energy carrier within the context of the global energy transition. Among the available storage solutions, Type IV high-pressure composite hydrogen tanks are particularly attractive due to their favourable strength-to-weight ratio and high mechanical performance. However, the reliability and safety of these structures remain critical concerns, as they are subject to progressive degradation mechanisms such as matrix cracking, fibre/matrix debonding, delamination, and fibre rupture during repeated pressurisation and depressurisation cycles [10], [22].

Ensuring early detection of degradation is crucial to prevent catastrophic failures and to support condition-based maintenance strategies [17]. In this context, acoustic emission (AE) monitoring has emerged as a promising non-destructive technique for capturing transient signals related to damage [13], [18], [24]. However, interpreting AE data is challenging because of its high dimensionality, non-stationarity, and sensitivity to noise, which can limit the effectiveness of standard signal processing methods [9], [21]. Netherveless, AE sensor monitoring has shown to be a useful non-destructive technique for detecting transient signals related to damage [13], [18], [24]. Still, the interpretation of AE data can be complicated due to its high dimensionality, non-stationarity, and sensitivity to noise. This complexity can sometimes reduce the effectiveness of standard signal processing methods [9], [21].

This paper adopts an applied research perspective and aims to develop an intelligent diagnostic framework for the early detection and characterisation of anomalies in pressure vessels. The proposed approach combines advanced signal processing-based on artificial intelligence techniques, including an attention-based LSTM Autoencoder for anomaly detection, supervised classification for degradation mechanism identification, adaptive thresholding, and multi-sensor localisation. The objectives of this work are then twofold. From a scientific standpoint, the goal is to propose a robust and explainable methodology capable of detecting and describing internal damage from complex and noisy AE data. From an industrial perspective, the aim is to provide a practical structural health monitoring solution that supports predictive maintenance, extends service life, and enhances the overall safety of hydrogen storage systems.

The remainder of this paper is structured as follows. Section II presents the related works, including acoustic emission monitoring, anomaly detection techniques, LSTM Autoencoders, attention mechanisms, and unsupervised clustering methods. Section III details the methodology and the overall diagnostic pipeline validation. Section IV presents experimental results and performance evaluation. Also, it provides discussion of the findings and limitations. Finally, Section V concludes the paper and outlines perspectives for future work.

II. Related Works

Anomaly detection techniques in composite structures are generally classified into two categories: techniques based solely on the use of acoustic emission (AE) sensors and those using machine learning (ML) methods.

A. Anomaly Detection with Acoustic Emission

AE is a widely adopted non-destructive evaluation method in structural health monitoring (SHM), particularly for composite structures. Piezoelectric sensors, such as PZT ceramics and PVDF polymers, detect transient waves generated by micro-damage events, including matrix cracking, fibre/matrix debonding, delamination, and fibre rupture [4].

AE offers several advantages: it allows continuous, passive monitoring without disassembly, is sensitive to low-energy events, and provides temporal and frequency-based features that correlate with specific damage mechanisms [12], [23]. These include amplitude, duration, energy, mean frequency, RA, and average frequency (AF). Despite its advantages, AE-based monitoring faces several challenges. The signals are highly complex and non-linear, varying significantly with material heterogeneity, sensor placement, and environmental conditions. They are also sensitive to noise from external sources, operational vibrations, and sensor cross-talk, which complicates direct interpretation. Additionally, multi-sensor arrays produce high-dimensional time-series data, making manual inspection impractical. Finally, certain damage mechanisms, such as delamination and matrix cracking, may generate overlapping signal features, creating ambiguity and limiting the reliability of classification based solely on AE signals.

Among other things, these limitations may justify the integration of advanced machine learning techniques to automate anomaly detection and improve interpretability techniques in composite structures.

B. Anomaly Detection with Machine Learning

Data-driven machine learning methods have been widely tested for detecting anomalies and classifying damage in complex systems. These methods fall into two main categories: supervised and unsupervised approaches.

Supervised learning methods, such as classical autoencoders [2], [15] and LSTM networks [1], [5], depend on labeled datasets to identify anomalies and classify damage types. While these methods work well, they have several drawbacks. First, they rely heavily on representative training data, meaning any normal operating condition not included during training may be incorrectly labeled as abnormal. Second, classical autoencoders handle data independently, which reduces their effectiveness on sequential AE signals. LSTM models need large amounts of labeled data and significant computing resources [11]. Finally, complex setups like LSTM autoencoders with attention mechanisms improve performance but add extra hyperparameters and longer training times. Their interpretability is still limited [16], as these models provide only a partial view of the underlying damage mechanisms.

Unsupervised approaches, such as K-means clustering [14] and Gaussian mixture models (GMMs) [3], are most commonly used when sufficient labeled data is unavailable. However, they present several challenges. In particular, optimizing classification quality depends on the chosen hyperparameters, such as the number of clusters and covariance assumptions; these may not accurately reflect the characteristics of actual failures. Furthermore, the clusters created by unsupervised methods provide only pseudo-labels and do not guarantee a match with the actual damage mechanisms. Additionally, when the data is large, this can lead to prohibitive computational costs, and noise in the signals can degrade clustering accuracy. This somewhat limits the reliability of unsupervised classification in real-time structural health monitoring (SHM) situations.

Due to the limitations of previously developed methods based on acoustic emission (AE) and machine learning (ML), this paper presents an integrated and interpretable framework for damage detection in composite pressure vessels. The proposed framework consists of two main components: an AE monitoring system and a hybrid ML pipeline that combines an LSTM autoencoder with attention to capture the temporal dynamics of AE signals and a Gaussian mixture model–based adaptive thresholding strategy to robustly distinguish normal from abnormal behaviour under varying signal characteristics and operating conditions. Following anomaly detection, unsupervised clustering and physically informed feature analysis are employed to associate distinct behavioural patterns with specific damage mechanisms. To support damage identification in scenarios with limited labelled data, a LightGBM classifier is used to provide efficient and robust classification of damage categories. Finally, a graph-based sensor aggregation approach enables spatial localisation of damage, allowing the identification of critical regions within the structure. Then, the novelty of this work lies in the integration of all these modules into a unified multi-stage diagnostic pipeline, adapted to structural integrity monitoring based on acoustic emission sensors. Indeed in this study, collectively, these modules enhance the robustness, interpretability, and practical applicability of AE-based structural health monitoring for the safety assessment of composite storage systems.

III. Methodology

With the aim of implementing an intelligent diagnostic system combining anomaly monitoring with acoustic emission (AE) sensors and machine learning techniques, we propose a multi-phase approach. The methodology follows the flow illustrated in Figure 1 to progressively validate the steps and reduce deployment risks. Each phase is described below.

A. Data Preparation and Validation

Let $\mathbf{X} = \{\mathbf{x}_t \in \mathbb{R}^d\}_{t=1}^T$ denote a multivariate AE time series, where d is the number of extracted features [8]. Prior to

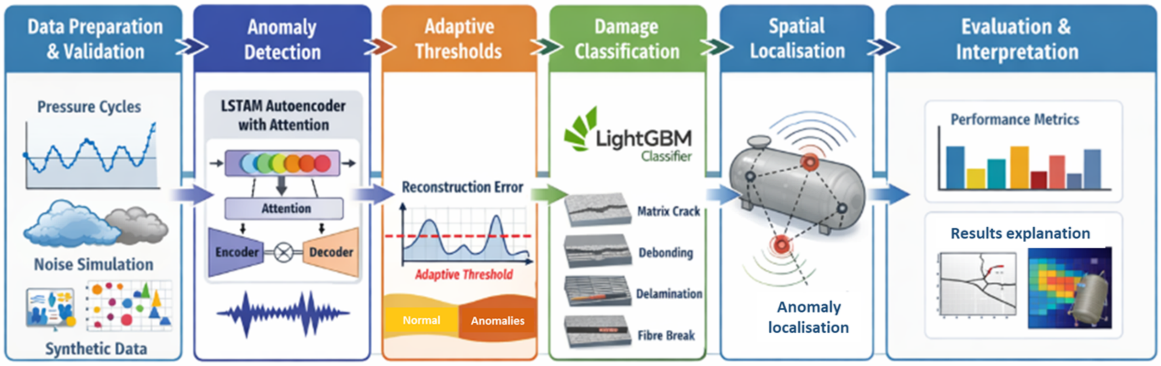

Fig. 1. Overview of the intelligent diagnostic approach pipeline.

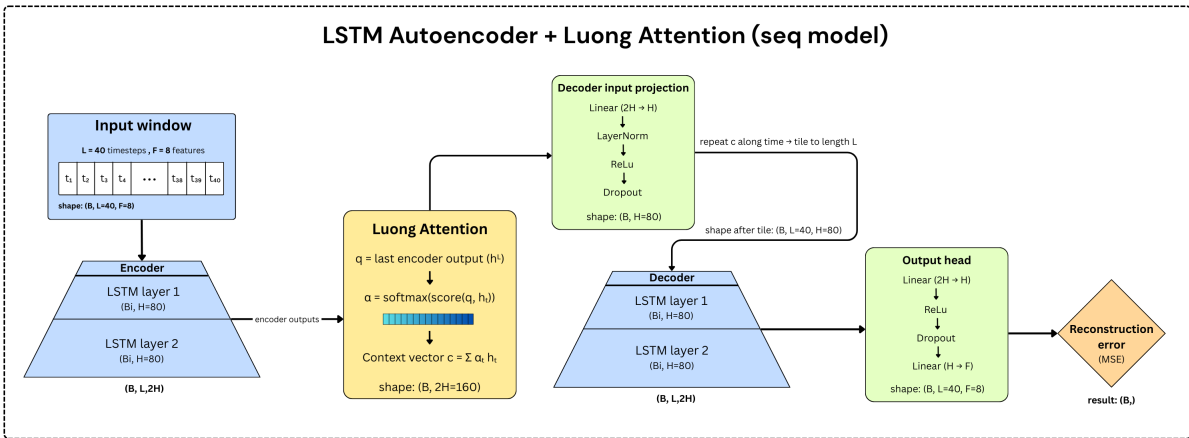

Fig. 2. LSTM autoencoder with Luong attention computation.

experiments, simulated datasets reproduce nominal operational behaviour under pressurisation/depressurisation cycles:

$$\mathbf{x}_t = f_{\text{phys}}(p_t, \varepsilon_t, v_t) + \boldsymbol{\eta}_t, \tag{1}$$

where p_t is pressure, ε_t strain-related effects, v_t vibrations, and $\boldsymbol{\eta}_t \sim \mathcal{N}(0, \sigma^2 I)$ models sensor noise. $f_{\text{phys}}(\cdot)$ represents a simplified surrogate function describing the nominal evolution of AE features under normal operating conditions.

Features are standardised according to:

$$\tilde{x}_{t,i} = \frac{x_{t,i} - \mu_i}{\sigma_i}, \tag{2}$$

and segmented into overlapping temporal windows $\mathbf{X}^{(k)} \in \mathbb{R}^{L\times d}$ of length L. Statistical and temporal features—including sliding-window means, standard deviations, local slopes, and inter-variable correlations—are extracted to enhance representation. The model is trained solely on standard cycles, and the evaluation is performed on a test set containing both normal and anomalous cycles.

B. LSTM Autoencoder with Attention

An LSTM autoencoder with multiplicative attention is employed to detect anomalies in sequential AE data. The LSTM encoder compresses temporal sequences into a latent representation, while the decoder reconstructs the input (Fig. 2). The internal LSTM operations are:

$$\begin{aligned}
g_t &= \sigma(W_f[h_{t-1}, x_t] + b_f),\\
i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i),\\
\tilde{C}_t &= \tanh(W_C[h_{t-1}, x_t] + b_C), \quad C_t = g_t \odot C_{t-1} + i_t \odot \tilde{C}_t,\\
o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \qquad h_t = o_t \odot \tanh(C_t),
\end{aligned} \tag{3}$$

where x_t is the input, h_t the hidden state, C_t the memory cell, W_* weights, b_* biases, σ the sigmoid function, and $\odot$ the element-wise product.

Attention weights α_t highlight informative time steps:

$$e_t = h_t^\top W_\alpha q, \qquad \alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{t} \exp(e_k)}, \tag{4}$$

$$c = \sum_{t=1}^{T} \alpha_t h_t, \tag{5}$$

where q is a query vector, and c is the context vector fed to the decoder. Reconstruction error serves as the anomaly score, while attention provides interpretability.

C. Adaptive Thresholding Computation

To discriminate between normal and abnormal behaviours, reconstruction errors e_i are normalized and modeled using Gaussian Mixture Models (GMMs):

$$p(x) = \sum_{t=1}^{T} \pi_t \mathcal{N}(x \mid \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), \tag{6}$$

where π_t are mixing coefficients, and $\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t$ are the mean and covariance of each component. Components with low mean and variance correspond to nominal operation, while components with high mean or variance indicate anomalous behaviour.

Formally, the adaptive threshold τ_i for observation i is defined as the likelihood-based boundary that accounts for the local distribution of normal operation:

$$\tau_i = \min \left\{ x \;\middle|\; \bar{p}(e_i) < \epsilon \right\}, \tag{7}$$

where $\bar{p}(e_i)$ is the probability density under the Gaussian component representing normal operation, and $\epsilon \in (0, 1)$ is a chosen significance level (e.g., $\epsilon = 0.01$).

An observation is flagged as anomalous if its reconstruction error exceeds the adaptive threshold:

$$\mathcal{F}_i = \begin{cases} 1, & e_i > \tau_i, \quad \text{Anomaly detected} \\ 0, & e_i \leq \tau_i, \quad \text{Otherwise.} \end{cases} \tag{8}$$

This formulation allows us to vary the detection threshold according to the local statistical properties of the reconstruction error, taking into account sensor heterogeneity, operational variability, and measurement noise. Thus, its determination is not based solely on a fixed, global threshold. Consequently, it improves the robustness and sensitivity of anomaly detection for different acoustic emission sensors and under various operating conditions.

D. Damage Classification and Cluster Interpretation

Detected anomalies undergo unsupervised clustering and supervised classification.

1) Unsupervised Clustering and Model Selection: Abnormal events are therefore first grouped using Gaussian mixture models (GMMs), which simultaneously model elliptic clusters with heterogeneous covariances. Furthermore, the optimal number of clusters k is chosen by minimizing the following Bayesian information criterion (BIC):

$$\mathrm{BIC} = -2\ln(\hat{L}) + p\ln(n), \tag{9}$$

with $\hat{L}$ the maximum likelihood, p the number of parameters, and n the number of events.

2) Cluster Analysis in Original Feature Space: Clusters are projected back to the original AE feature space. Key descriptors (linear amplitude $\mathrm{AMP}_{\mathrm{LIN}}$, AF, RA, duration, energy) are analysed. Linear amplitude is preferred:

$$\mathrm{AMP}_{\mathrm{LIN}} = 10^{\mathrm{AMP}_{\mathrm{dB}}/20}, \tag{10}$$

providing direct physical interpretability.

3) Attribution of Physical Labels and Supervised Classification: Clusters are assigned physical labels based on different parameters: (i) median parameters of AEs, (ii) the temporal distribution over fatigue cycles, (iii) the methodology proposed by the literature [6], [19], (iv) and expert validation. LightGBM is then trained on feature vectors $\mathbf{f}^{(k)}$ to classify anomalies:

$$\hat{y} = \arg\max_{c \in \mathcal{C}} P(y = c \mid \mathbf{f}^{(k)}), \tag{11}$$

where $\mathcal{C}$ represents degradation modes. In this case, LightGBM method can ensure robustness and more efficiency.

E. Spatial Localisation of Anomalies

In order to localise the anomaly, let us aggregate the anomaly scores $e_i^{(s)}$ from each sensor s, such that :

$$E^{(s)} = \frac{1}{N_s} \sum_{i=1}^{N_s} e_i^{(s)}. \tag{12}$$

While considering the impact of the anomaly over the sensor network, we use a graph-based method such as :

$$S^{(s)} = (1-\gamma)E^{(s)} + \gamma(\mathbf{D}^{-1}\mathbf{A}E)^{(s)}, \tag{13}$$

with $\mathbf{A}$ the adjacency matrix, $\mathbf{D}$ the degree matrix, and γ a smoothing parameter. The dominant sensor channel is then identified as:

$$s^* = \arg\max_{s} S^{(s)}, \tag{14}$$

providing an interpretable localisation of the damage within the vessel.

IV. Experimental Validation

A. Experimental Setting

The proposed methodology was evaluated on both simulated and real datasets during various static and fatigue tests performed with EA sensors. The simulated data demonstrated that the proposed process is capable not only of detecting anomalies but also of classifying the simulated degradation mechanisms with high reliability. However, it is important to emphasize that these performance levels are achieved under ideal simulation conditions; their application to real data remains the main challenge in ensuring the robustness of the model.

The general configuration of the dataset is described as follows. The data is organised into temporal windows of length $L = 40$ timesteps, each timestep described by $F = 8$ features extracted from the acoustic signal. Thus, each training batch is represented by a tensor: $X \in \mathbb{R}^{B \times L \times F}$, where B is the batch size, L the length of the sequence, and F the number of input variables per timestep. The hyperparamters used in the proposed framework is summarised in Table I. Moreover, each sequence window is transformed into independent samples, one per timestep. To obtain a global classification per sequence, per-timestep predictions are aggregated. At the end, the results are reshaped back to sequential form, preserving the original temporal structure. The quality of the reconstruction is evaluated with the Mean Squared Error (MSE) between the

Parameter	Sequence length L	Features F	Batch size B	Hidden size H	Bi-LSTM layers	Dropout	Activation	Loss	GMM criterion	ϵ	γ
Value	40	8	32	80	2	0.2	ReLU	MSE	BIC	0.01	0.3–0.5

TABLE I

MAIN HYPERPARAMETERS OF THE PROPOSED FRAMEWORK.

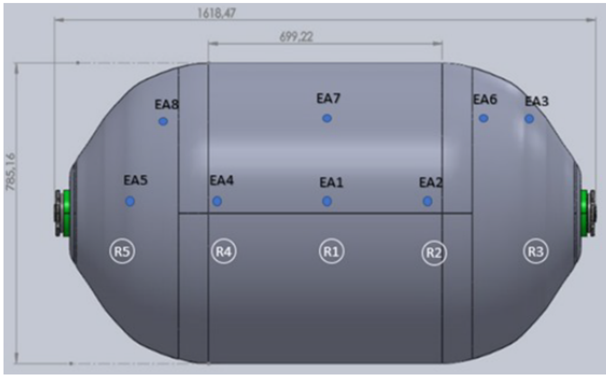

Fig. 3. Pressure vessel under study with mounted AE sensors and their locations

input sequence X and the reconstructed sequence $\hat{X}$, such that:

$$\mathcal{L}_{rec} = \frac{1}{B \times L \times F} \sum_{b=1}^{B} \sum_{l=1}^{L} \sum_{f=1}^{F} (X_{[b,l,f]} - \hat{X}_{[b,l,f]})^2.$$

We used two experimental AE-campaigns conducted on the composite pressure vessel equipped with a network of eight AE sensors, as positioned in Fig. 3:

- a static loading test, consisting of a continuous and progressive increase in pressure up to approximately 570 bar, without interruption;
- a fatigue test, involving intensive pressure cycles from approximately 20 to 440 bar, repeated until the damage of the vessel.

*1) **[Static Test]** Baseline:* It should be noted that the composite vessel remained virtually intact during the static test, despite vibrations caused by its prior movement. During the test, the pressure was gradually increased to nearly 570 bar without causing the tank damage. This test provides an ideal benchmark for evaluating the tank's condition in terms of its acoustic signature.

Furthermore, during several static tests, we identified an inflection point beyond which acoustic activity increases rapidly and almost exponentially. More precisely, below this point, acoustic emissions remain low and primarily reflect microscopic adjustments without significant damage. Beyond this point, acoustic activity increases sharply, indicating the initiation and propagation of early internal damage, such as microcracks, fiber/matrix delamination, and similar defects. This inflection point thus marks the transition between normal elastic behavior and the onset of initial damage.

*2) **[Fatigue Test]** Damage Onset:* In the fatigue test, the tank underwent repeated cycles between approximately 20 and 440 bar until failure. AE analysis revealed a progressive increase in activity over successive cycles, the gradual emergence of representative signatures of internal damage, including microcracks, fibre/matrix debonding, delamination, and fibre breaks; and significant spatial heterogeneity across sensors depending on their positions on the tank. Automated cycle detection identified a total of 224 cycles, with a detailed examination of 10 randomly selected cycles conducted to evaluate the periodicity and consistency of peak values (Cycle 22 illustrated in Fig. 4).

B. *Cluster signatures*

We clustered acoustic emission hits using a Gaussian Mixture Model (GMM) on standardized per-hit descriptors Amplitude (AMP) in dB, Average Frequency (A-FRQ) in MHz, Duration (DURATION) and Rise Time (RISE) in μs, Signal energy (ENERGY) in aJ, and RA use an instrument unit. The objective was to group AE signals into physically meaningful damage mechanisms that could later be used as labels for supervised classification.

Following the identification of $k = 5$ clusters using a GMM (after choosing the principal components via PCA), and using the representative median feature values reported in Table II, each cluster was assigned a physical damage mechanism. This attribution relied on a combined analysis of the characteristic acoustic emission parameters, the clusters' temporal evolution over the fatigue cycles, comparisons with established acoustic emission literature [7], [20], and the validation by mechanical experts. Under this framework, Cluster 0 is interpreted as matrix cracking, as it exhibits moderate amplitudes, low average frequency, intermediate RA values, and moderate signal durations ($\approx 0.9\ ms$), which are typical signatures of progressive microcrack development in the matrix during cyclic loading. Cluster 1 is associated with fibre–matrix de-cohesion and is characterised by higher amplitudes (around 626), high average frequency ($\approx 0.108\ MHz$), very low RA values (about 0.039), and short durations (around 0.15 ms), indicating fast, high-frequency events commonly observed during the early stages of damage initiation. Cluster 2 corresponds to delamination processes, as reflected by longer signal durations (around 1.1 ms), intermediate RA values ($\approx$ 0.27), and relatively low energy, suggesting gradual interlaminar separation or distributed delamination. Cluster 3 represents fibre breakage and is distinguished by extremely high amplitudes (up to 57 000), exceptionally long durations exceeding $10^6\ \mu$s (i.e. more than 1 s), and very high energy levels; although rare (only four events), these signals are indicative of critical fibre failures occurring immediately prior to the catastrophic rupture of the pressure vessel. Finally, Cluster 4 is attributed to advanced crack propagation or delamination, combining high amplitudes (around 2 617), very long durations ($\approx 3.4\ ms$), and intermediate average frequency values, which are consistent with more

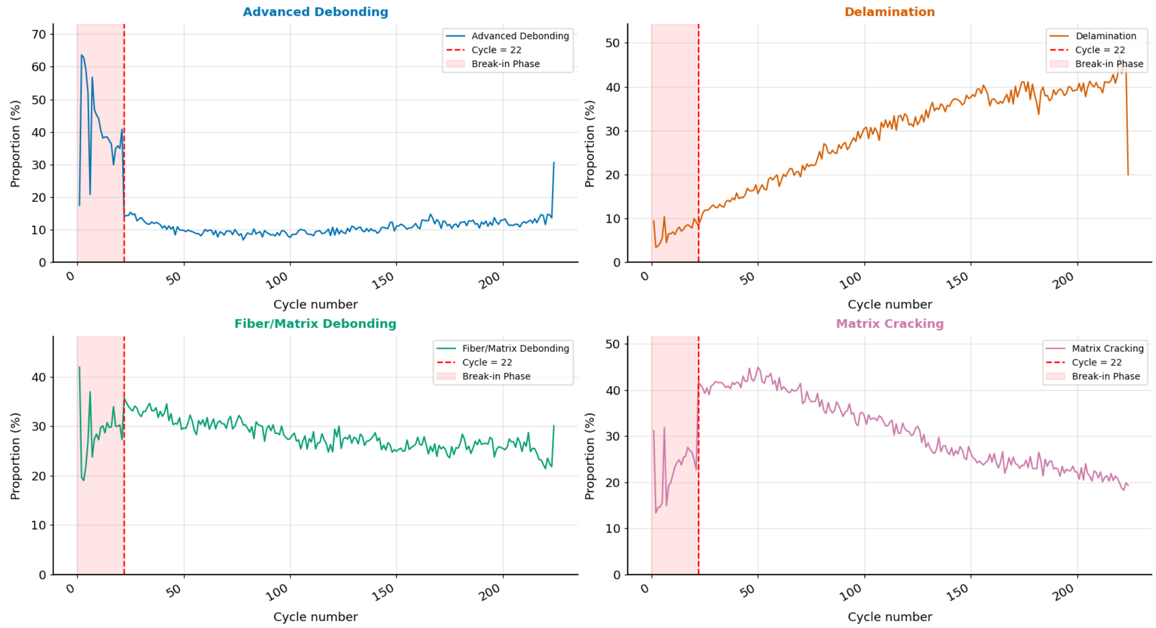

Fig. 4. Temporal evolution of the four damage-related clusters across fatigue cycle, highlighting behaviour consistent with established fatigue damage mechanisms.

Cluster ID	**AMP**$_{LIN}$	**A-FRQ**[MHz]	**Duration**[µs]	**Energy**[aJ]	**RA**	Event number
0	434.9	0.045	907	156.1	0.558	124 014
1	625.9	0.108	150	179.5	0.039	163 880
2	526.6	0.056	1 137	4.1	0.270	65 400
3	56 988.6	0.103	1 393 653	149 837	0.183	4
4	2 617	39	3 459	132.8	0.081	32 234

TABLE II
MEDIAN FEATURE VALUES OF THE FIVE GMM CLUSTERS (IDs 0–4).

	Delamination	Fibre/Matrix debonding	Matrix cracking	Propagation/Advanced delamination
Delamination	**9 109**	0	0	47
Fibre/Matrix debonding	0	**22 237**	437	269
Matrix cracking	0	238	**16 953**	171
Propagation/Advanced delamination	33	59	64	**4 357**

TABLE III
CONFUSION MATRIX FOR ANOMALIES DETECTION.

energetic and mature damage mechanisms than those observed in Cluster 2.

The analysis of the discovered clusters based on when they appeared in fatigue cycles (Fig. 4) shows the expected progression of the various types of damage mechanisms in the composite pressure vessel. Fibre–matrix decohesion reaches its maximum shortly after load application, and then decreases; accordingly, the fibre–matrix decohesion occurs first as a consequence of loading during the previous static load cycle, followed by changes in the matrix due to loading cycles. For example, delaminations tend to occur early, followed by stabilization of the delaminations, indicating that they could have occurred due to static load-induced micro-delaminations prior to the occurrence of the early fatigue cycles. The amount of matrix cracking continues to increase throughout the fatigue process, indicating the important role that matrix cracking plays in terms of cumulative damage mechanisms. Advanced delamination and crack propagation are typically more sporadic during the last few cycles of fatigue, which coincide with the forming of matrix fractures; thus, fiber rupture occurs very infrequently until the final cycle, where they are representative of the ultimate failure of the composite material. The differentiation between early stage and advanced stage of debonding is also evident, as fibers continue to experience advanced debonding as they complete the fatigue process. These findings indicate a clear relationship between the discovered clusters and the physical damage process and support established predictive models of composite fatigue based on the clusters detected through AE data.

C. Damage detection and localisation

The confusion matrix (c.f Table III) demonstrates high classification performance across all four damage mechanisms. The majority of samples are correctly identified, with only minor misclassifications occurring primarily between physically related anomalies. Delamination and Fibre/Matrix debonding exhibit near-perfect detection, while matrix cracking shows limited confusion with Fibre/Matrix debonding. Propagation/Advanced delamination is also well recognized, with minimal misclassification into other classes. Overall, the results indicate that the model reliably distinguishes both early-stage and advanced damage modes.

Class	**Precision**	**Recall**	**F1-score**
Delamination	0.996	0.995	0.996
Fibre/Matrix debonding	0.987	0.969	0.978
Matrix cracking	0.971	0.976	0.974
Propagation/Advanced delamination	0.900	0.967	0.932

TABLE IV
PRECISION, RECALL, AND F1-SCORE PER ANOMALY CLASS.

The metrics presented in the table IV show that the model performs very well in classifying different types of damage. Delamination achieves the highest accuracy and F1 score (≈ 0.996). This reflects minimal false positives and false negatives. Fiber/matrix separation and matrix cracking also perform well, with respective F1 scores of 0.978 and 0.974. This suggests that the model reliably distinguishes between these two physically similar types of damage. Propagated/advanced delamination, while slightly less accurate (0.900), maintains a high recall (0.967) and a high F1 score (0.932); this indicates that the model still effectively detects advanced damage, despite a slight over-classification in this category. Overall, these metrics demonstrate that the model is robust and reliable for automatic anomaly detection. It is capable of detecting both early and advanced damage with high accuracy.

By considering the sensors' placement illustrated in Fig. 3, the spatial distribution of damage was evaluated by aggregating channel anomaly activities per sensor as proxy anomaly scores and smoothing them over the acoustic emission (AE) network using the metrics in equations (12), (13) and (14). With $\gamma = 0.5$, smoothed scores revealed a pronounced spatial heterogeneity, with the AE3 sensor (CH3) showing the highest value ($S^{(3)} = 49\,615$), clearly identifying it as the dominant damage location. Secondary contributions were observed at AE6 and AE7, while the remaining sensors displayed substantially lower anomaly levels, consistent with less active damage regions. To verify robustness, the smoothing parameter was reduced to $\gamma = 0.3$, placing greater emphasis on local activity; AE3 remained the dominant sensor ($S^{(3)} = 67\,778$), and the relative spatial ranking of sensors was preserved. These results indicate that the observed localisation is not sensitive to the choice of smoothing and is primarily driven by intrinsic damage activity, rather than network diffusion effects, providing an interpretable and reliable identification of critical structural zones within the vessel.

D. Comparative Study

To quantify the contribution of the main components of the proposed method, a validation process involving the progressive integration of each module was performed. The following configurations were evaluated:

- **Baseline LSTM-AE**: standard LSTM autoencoder using a fixed global threshold for anomaly detection.
- **LSTM-AE + Attention**: addition of the Luong attention mechanism to enhance temporal feature weighting.
- **LSTM-AE + Attention + Fixed Threshold**: attention-based reconstruction using a static anomaly threshold.
- **LSTM-AE + Attention + Adaptive Threshold**: incorporation of the GMM-based adaptive thresholding strategy.
- **Full Proposed Model**: complete framework including attention, adaptive thresholding, and graph-based spatial smoothing.

Note that each configuration was evaluated on the entire fatigue test dataset using the same training and testing protocol. Table V presents the results obtained in terms of accuracy, precision, recall, F1 score, and false positive rate (FPR).

The comparative results demonstrate the importance of each module within the proposed architecture. The basic LSTM autoencoder already offers satisfactory anomaly reconstruction capabilities; however, the introduction of the attention mechanism significantly improves detection performance by allowing the model to focus on temporally informative acoustic events.

Furthermore, adapting the threshold using a Gaussian Mixture Model (GMM) further improves recall and the false positive rate, confirming that reconstruction errors are better separated when local statistical variations are taken into account rather than with a global threshold. This result is particularly important for monitoring with acoustic emission (AE) sensors, where responses can vary depending on load conditions and local structural heterogeneity.

Finally, integrating graph-based spatial smoothing provides improved overall performance. Indeed, the reduction in the false positive rate from 0.027 to 0.014 indicates that leveraging spatial coherence between neighboring sensors improves robustness against isolated noisy signals while preserving sensitivity to events actually related to damage. Overall, the complete framework achieves the highest accuracy and F1-Score, demonstrating that the various architectural components contribute in a complementary manner to the detection and localization of anomalies.

V. CONCLUSION

This article describes a hybrid machine learning framework for detecting and classifying anomalies in pressure vessels using acoustic emission signals. The proposed method combines an LSTM autoencoder with an improved attention mechanism for strong anomaly detection. It also includes adaptive thresholding to handle changes in operating conditions, a supervised LightGBM model for classifying damage types, and spatial localization of anomalies through a sensor network.

Preliminary evaluations performed on simulated datasets demonstrated the validity of the proposed approach. Its performance was confirmed by thorough validation on experimental datasets from static and fatigue tests, thus evaluating the generalizability, robustness, and practical reliability of the

Model Configuration	Accuracy	Precision	Recall	F1-score	FPR
Baseline LSTM-AE	0.912	0.903	0.894	0.898	0.083
LSTM-AE + Attention	0.941	0.936	0.928	0.932	0.059
LSTM-AE + Attention + Fixed Threshold	0.952	0.948	0.941	0.944	0.046
LSTM-AE + Attention + Adaptive Threshold	0.971	0.969	0.963	0.966	0.027
Full Proposed Model	**0.984**	**0.981**	**0.977**	**0.979**	**0.014**

TABLE V
COMPARATIVE STUDY OF THE PROPOSED FRAMEWORK.

model for real-world structural integrity monitoring applications. The LSTM autoencoder effectively detected deviations from normal behavior, while the LightGBM classifier accurately distinguished several degradation mechanisms, including matrix cracking, fiber/matrix separation, delamination, and fiber breakage. These results highlight the suitability of the selected architectures and confirm the feasibility of the proposed multiphase diagnostic process.

Future work will focus on extending the methodology to real-time implementation and improving the robustness of anomaly detection under complex operating conditions. The ultimate goal is to provide a reliable, interpretable, and operationally deployable solution to support predictive maintenance strategies, improve the safety and lifespan of hazardous materials storage systems, and facilitate informed decision making in industrial applications.

ACKNOWLEDGMENT

The authors gratefully acknowledge MAGYAR for initiating and funding this work, and for providing the materials and prototypes required for its completion. Also, this work was funded by the French State as part of France 2030 operated by ADEME, and funded by the European Union – NextGenerationEU.

REFERENCES

[1] M.-H. Benzaama, F. Alassaad, L. Rajaoarisoa, and K. Touati. Artificial intelligence approaches to predict thermal behavior of light earth cell incorporating pcms: Experimental cnn and lstm validation. *Journal of Energy Storage*, 68:107780, 2023.

[2] A. Borghesi, A. Bartolini, M. Lombardi, M. Milano, and L. Benini. Anomaly detection using autoencoders in high performance computing systems. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):9428–9433, Jul. 2019.

[3] A. Farhidzadeh, E. Dehghan-Niri, and S. Salamone. Gaussian mixture modeling of acoustic emissions for structural health monitoring of reinforced concrete structures. In *Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring*, 2013.

[4] N. Ghadarah and D. Ayre. A review on acoustic emission testing for structural health monitoring of polymer-based composites. *Sensors*, 23(15), 2023.

[5] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 11 1997.

[6] P.-Y. Joubert, E. Vourc'h, A. Tassin, and Y. Le Diraison. Source separation techniques applied to the detection of subsurface defects in the eddy current ndt of aeronautical lap-joints. *NDT and E International*, 43(7):606–614, 2010.

[7] P.-Y. Joubert, E. Vourc'h, A. Tassin, and Y. Le Diraison. Source separation techniques applied to the detection of subsurface defects in the eddy current ndt of aeronautical lap-joints. *NDT & E International*, 43(7):606–614, 2010.

[8] M. Kuk, S. Bobek, B. Veloso, L. Rajaoarisoa, and G. J. Nalepa. Feature importances as a tool for root cause analysis in time-series events. In J. Mikyška, C. de Mulatier, M. Paszynski, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. Sloot, editors, *Computational Science – ICCS 2023*, pages 408–416, Cham, 2023. Springer Nature Switzerland.

[9] C. Leone, G. Caprino, and I. de Iorio. Interpreting acoustic emission signals by artificial neural networks to predict the residual strength of pre-fatigued gfrp laminates. *Composites Science and Technology*, 66(2):233–239, 2006. Experimental Techniques and Design in Composite Materials.

[10] P. Liu, G. Li, B. Li, K. Wang, and X. Ma. Multi-property mechanical degradation mechanisms and reliability assessment of cf/ep composite for hydrogen storage tank winding layers after fire exposure. *International Journal of Hydrogen Energy*, 154:150335, 2025.

[11] M. Luong, H. Pham, and C. D. Manning. Effective approaches to attention-based neural machine translation. *CoRR*, abs/1508.04025, 2015.

[12] E. Njuhovic, M. Bräu, F. Wolff-Fabris, K. Starzynski, and V. Altstädt. Identification of interface failure mechanisms of metallized glass fibre reinforced composites using acoustic emission analysis. *Composites Part B: Engineering*, 66:443–452, 2014.

[13] T. B. Quy and J.-M. Kim. Crack detection and localization in a fluid pipeline based on acoustic emission signals. *Mechanical Systems and Signal Processing*, 150:107254, 2021.

[14] N. A. A. Rahman, Z. May, and M. S. Mahmud. Unsupervised classification of acoustic emission signal to discriminate composite failure at low frequency. In R. Ibrahim, K. Porkumaran, R. Kannan, N. Mohd Nor, and S. Prabakar, editors, *International Conference on Artificial Intelligence for Smart Community*, pages 797–806, Singapore, 2022. Springer Nature Singapore.

[15] L. Rajaoarisoa, M. Kuk, S. Bobek, and M. Sayed-Mouchaweh. Hybrid and co-learning approach for anomalies prediction and explanation of wind turbine systems. *Engineering Applications of Artificial Intelligence*, 133:108046, 2024.

[16] R. Randrianandraina, J. Chapuy, L. Rajaoarisoa, and M. Sayed-Mouchaweh. Wind turbine generator bearing anomaly detection and explanation using rrcf approach. In *2023 International Conference on Machine Learning and Applications (ICMLA)*, pages 1574–1581, 2023.

[17] J. J. Randriarison, L. Rajaoarisoa, and M. Sayed-Mouchaweh. Faults explanation based on a machine learning model for predictive maintenance purposes. In *2023 International Conference on Control, Automation and Diagnosis (ICCAD)*, pages 01–06, 2023.

[18] N. Seemuang, J. Lim, and T. Kaewkongka. Shaft crack monitoring by using acoustic emission technique. In *2018 2nd International Conference on Engineering Innovation (ICEI)*, pages 30–34, 2018.

[19] S. Suresh. *Fatigue of Materials*. Cambridge University Press, 2 edition, 1998.

[20] S. Surish. *Fatigue of Materials*. 2nd ed. Cambridge University Press, 1998.

[21] L. Tomcic, A. Ederer, S. Grabmann, M. Kick, J. Kriegler, and M. F. Zaeh. Interpreting acoustic emissions to determine the weld depth during laser beam welding. *Journal of Laser Applications*, 34(4):042052, 11 2022.

[22] S. Valilis and G. Lampeas. Thermomechanical analysis and progressive damage modelling of composite cryogenic hydrogen tanks. *The Aeronautical Journal*, page 1–25, 2026.

[23] J. H. WilliamsJr and S. S. Lee. Acoustic emission monitoring of fiber composite materials and structures. *Journal of Composite Materials*, 12(4):348–370, 1978.

[24] D. Xiang. Acoustic emission detection of early stages of cracks in rotating gearbox components. *AIP Conference Proceedings*, 1806(1):070010, 02 2017.