Structural Control and Health Monitoring

WILEY

RESEARCH ARTICLE OPEN ACCESS

Damage Categorization in Full-Scale, Full-Composite Ship Hull Under High-Energy Impacts by Unsupervised Learning–Enabled Acoustic Emission Monitoring and Laser Shearography Inspection

Pedro A. Ochôa[1] | H. J. den Ouden[1] | Michiel Hagenbeek[1] | Nan Tao[2] | Andrei G. Anisimov[2] | Roger M. Groves[2]

[1]Mobility and Built Environment, TNO, Delft, the Netherlands | [2]Aerospace Structures and Materials Department, Delft University of Technology, Delft, the Netherlands

Correspondence: Nan Tao (n.tao@tudelft.nl) | Andrei G. Anisimov (a.g.anisimov@tudelft.nl)

Received: 30 December 2024 | **Revised:** 16 January 2026 | **Accepted:** 6 February 2026

Academic Editor: Hoon Sohn

Keywords: acoustic emission | composite ship | condition-based maintenance | damage categorization | structural health monitoring | unsupervised learning

ABSTRACT

A large, full-scale (6 m tall, 2.5 m wide) full-composite ship hull section was subjected to three consecutive impacts with an impact energy level (~20 kJ) mimicking a realistic heavy crash. The structure was continuously monitored with acoustic emission (AE) sensors during the impacts, which allowed the possible degradation of the composite hull to be assessed. Unsupervised learning was applied to AE data features to enable the categorization of the damage accumulated during the consecutive impacts. The implemented unsupervised learning routine was a combination of automatic Laplacian data feature selection followed by density-based spatial clustering of applications with noise (DBSCAN), for which the hyperparameters were automatically optimized with a silhouette-driven approach. Three predominant damage mechanisms (sandwich-core crushing/cracking, skin-core debonding, and matrix cracking) were identified through the clustering solution (with 0.9 mean silhouette and 6% outliers) and verified by AE feature bands found in the literature. The AE-based damage categorization was validated with impact slow-motion videos and postimpact digital laser shearography inspections. The study highlights the capability of the developed categorization methodology to be deployed for online monitoring, where SHM algorithms are retrained during ship operation to keep improving diagnosis accuracy. The results show that in the online training scenario, where Impacts 1 and 2 were used for training and Impact 3 was used for testing, the categorization performance was comparable to when data from all three impacts had been used for training, with a mean silhouette of 0.884 and only 2% outliers. Altogether, the damage categorization routine demonstrated reliability and stability for handling realistic AE data variability and different data availability scenarios. This study is an important step toward a complete condition diagnosis (comprising detection, localization, categorization, and quantification) of full-scale composite ship structures, which is crucial for estimating their remaining useful life in real time and thereby for enabling their condition-based maintenance.

1 | Introduction

There is a direct need for increased sustainability and efficiency in the maritime sector [1, 2]. More specifically, it is urgent to drastically reduce fuel consumption and emissions of large ships (shipping accounted for 15% of NO_x, 13% of SO_x, and 3% of CO_2 global annual anthropogenic emissions from 2007 to 2012 [2]) while optimizing the cost-effectiveness of their operation and maintenance. The use of fiber-reinforced polymer (FRP) composite materials for shipbuilding is recognized as an attractive solution for these issues, as it would tackle both aspects simultaneously. On the one hand, FRP composites have the potential to enable a weight reduction of up to 50% [3] with respect to the use of steel and a 15% decrease in fuel consumption and costs relative to current levels [4], thus having a substantial contribution to reducing emissions. On the other hand, composites are corrosion resistant and can, therefore, reduce vessel downtime and eliminate the costs associated with corrosion damage repairs [3].

However, several obstacles still exist to adopting composites as structural materials for large ships above 500 ton (~25 m). First, although there has been rapid progress in the production of composite boats, there are still some technological challenges in cost-effectively manufacturing large full-composite ships. Second, current International Maritime Organization (IMO) Safety of Life at Sea (SOLAS) regulations on fire resistance and structural reliability do not allow FRP composites to be used for structural applications [3]. In particular, it is still difficult to predict damage progression in composites [5, 6], which affects structural design approaches and adds uncertainty to structural evaluations for safety-critical scenarios and to structural assessments for maintenance activity planning.

One of the solutions to address this structural uncertainty is to resort to structural health monitoring (SHM) to acquire real-time data, which can be used to perform the four tasks of preliminary diagnosis: detection, localization, categorization, and quantification [7]. This information can then be periodically fed into degradation models that recompute the remaining load-bearing capacity and remaining useful life of the structure, while accounting for structural and measurement uncertainties. When the load-bearing capacity reaches a certain minimum allowable threshold, a warning can be launched and a detailed nondestructive inspection (NDI) can be issued for the most affected structural areas. Based on the NDI results, a final decision can be made about the appropriate maintenance action. Hence, SHM is a means to optimize the planning and the deployment of detailed NDIs, rather than replacing them [8], and by doing so, it is also a means to optimize the planning and execution of maintenance programs. Thus, it is very important to validate SHM studies with NDI data in order to increase the reliability of SHM diagnosis and, thereby, to ensure that maintenance is performed at the desired locations and moments. On top of that, studies with combined utilization of SHM and NDI improve the understanding of the synergy between the two, which is necessary for condition-based maintenance (CBM) practices [8].

In the last decade, research and industrial efforts at an European level have been set in motion precisely to overcome the aforementioned obstacles to the adoption of composites as structural materials for large ships. The European project "Realization and Demonstration of Advanced Material Solutions for Sustainable and Efficient Ships" (RAMSSES) ran from 2017 to 2021 and consisted of a consortium of 36 partners from 13 European countries. The aims of the RAMSSES project [9] were to tackle the challenges of upscaling the technology for designing and building large FRP-composite vessels that comply with the IMO-SOLAS regulations and to demonstrate fire resistance, impact resistance, and structural robustness with cost-effectiveness across the process [10].

Very few studies have approached the problem of real-time SHM of full-scale composite ship structures. It is at this scale that several inter-related factors—e.g., realistic dimensions and boundary conditions (which affect measured quantities), uneven material quality and associated heterogeneities, existence of joints, a higher number of partially uncontrolled measurement variability parameters (e.g., varying quality of sensor network instrumentation and installation), and a larger number of noise sources (including those from operational-environmental conditions)—make damage diagnosis a more complex task, thereby influencing the reliability and robustness of SHM systems. At the beginning of the 2000 s, the Royal Norwegian Navy conducted tests on a couple of 47-m-long, 13.5-m-wide composite twin-hull ships weighing around 270 tons [11, 12]. The strain in the FRP-composite sandwich hull was measured with fiber-optic Bragg grating (FBG) sensors. Wave impact detection was demonstrated; however, no damage detection capabilities were developed. Some years later, Herszberg et al. [13] employed FBG strain sensing to detect the presence of Teflon inserts in the adhesion interface of long, thick adhesively bonded glass FRP beams, representing typical composite ship joints. In static and dynamic tests, only small differences between pristine and artificially damaged specimens were detected. Recently, Mieloszyk et al. [14] have deployed FBG strain sensing on the FRP-composite sandwich (glass fiber skins and polyvinyl chloride foam core) hull of a 7-m-long fast patrol boat. After acquiring data from the intact structure, a delamination was artificially inserted in the area of the hull instrumented with FBG sensors. The comparison between pristine and damaged states revealed minor differences for only some of the sensors.

All these studies are examples of using dynamic strain response anomalies to determine changes in, e.g., structural natural frequencies and damping factors, which can be correlated to stiffness changes and thereby to damage. However, such SHM approaches are relatively insensitive to early small damage locations in composite structures, because these tend to have a very limited effect on the global structural stiffness [15]. Hence, it is possible to state that a reliable and robust SHM approach for a detailed identification of damage in composite ship structures is still lacking. This constitutes the knowledge gap addressed in this article.

According to Wang et al. [11], "*damage detection will probably have to be based on detailed knowledge of the mechanical properties and damage mechanisms of the actual structure, and a careful characterization of the types of damage that are to be detected.*" This is especially true for composite ship structures, where damage can occur in the form of sandwich-core crushing/cracking, skin-core debonding, fiber–matrix disbonds, matrix cracking, delaminations, fiber pull-out, and fiber breakage. This is precisely what acoustic emission (AE)-based SHM can be the key for. AE corresponds to ultrasonic waves generated by sudden

changes and/or degradation in the material. Besides the possibility of directly using the AE signals to detect the presence of damage and localize it, there is proven evidence that different AE signal features are sensitive to different damage mechanisms in composite structures [16–19]. By resorting to data clustering analysis (i.e., unsupervised machine learning), it is possible to explore AE signal feature patterns to identify different damage locations and types [18]. Taking all this into account, AE sensing has great potential for early-stage detection, localization, and categorization of damage and damage progression in composite ships.

The research goal was to

- Develop an approach for reliable categorization of damage mechanisms in composite structures based on AE-SHM and unsupervised machine learning that can be employed without prior knowledge of the occurring damage mechanisms and that can be applied in a systematic way, irrespective of structure size and complexity.
- Demonstrate the damage categorization approach on a full-scale structure subjected to realistic loads.
- Validate the damage categorization approach with NDI results.
- Demonstrate the robustness of the damage categorization approach in variable data availability scenarios.
- Showcase the importance of combining SHM and NDI for CBM programs.

The research described in this article was performed within the framework of the RAMSSES project and employed a full-scale (1:1 scale, 6 m high, 2.5 m wide, 10 *t* mass) full-composite ship hull section subjected to consecutive realistic high-energy (∼20 kJ) impacts.

To the best of our knowledge, this is the first published validated study addressing the first three SHM functions of preliminary diagnosis (detection, localization, and categorization) for a full-scale full-composite ship structure subjected to realistic loading. The article demonstrates the synergy between SHM and NDI for an efficient and effective damage diagnosis, which is crucial for devising adequate CBM programs for composite ship structures.

2 | Composite Ship Hull Demonstrator

The tested structure was a full-scale demonstrator of an 80-m composite ship hull section, as illustrated in Figure 1. It consisted of a composite sandwich wall and three adhesively bonded composite sandwich decks, as shown on the right side of Figure 1. The design was developed by Damen Gorinchem and InfraCore. The hull wall was manufactured by Airborne UK and the decks were manufactured by InfraCore, all with resin infusion technology.

The hull wall was formed by a foam core and glass fiber-reinforced vinylester laminate skins. The deck consisted of multiple two-flanged web structures immersed in a foam core and enclosed in glass fiber-reinforced laminate skins [10]. The demonstrator had a height of 6 m and a width of 2.5 m and weighed approximately 10 *t*. The impacts on the hull side wall were performed in between the first (lower) and second (middle) decks, approximately 2.5 m above the keel, approximately at the level of the waterline.

3 | Impact Testing

The impact tests in this work were part of a vast and diverse test campaign in the RAMSSES project [9]. The goal of that test campaign was to demonstrate fire resistance, impact resistance, and structural robustness of large FRP-composite vessel structures manufactured with cost-effective materials and processes. In this context, three impact tests were designed to mimic impact conditions as realistically as possible, i.e., harsh and uncontrolled impact events. A basic level of impact control was adopted, by carefully verifying the pendulum mass, drop height, alignment, and contact area before the first impact.

For the three impacts on the hull side wall, the demonstrator was kept upright by a support steel structure, which weighed approximately 5 tons and was not fixed to the floor, as depicted in Figure 2(a). According to design requirements, minor delamination in the outer skin is allowed after repeated limit loads, provided that delamination does not immediately lead to full hull failure. A crash/impact load case should also not result in full hull penetration. To apply the impact load case, a pendulous steel assembly of two weights and a spherical head, with a total mass of 1550 kg (see Figure 2(b)), was dropped from a height of around 2 m with respect to the impact location (3.5 m relative to the ground) and swung against the structure (see Figure 2(b)), resulting in an impact energy of around 20 kJ. The total length of the impactor was ∼2 m, and the diameter was ∼40 cm. The impactor head height and thickness were ∼12 and 2.5 cm, respectively. The first video in the support footage (see the Data Availability section) shows one of the impacts filmed from the same general perspective as in Figure 2(a).

The impact tests were not designed to ensure the induction of a specific, well-controlled progression of composite damage mechanisms. And yet, the damage categorization framework demonstrated strong stability by always producing consistent results, despite the unconstrained damage progression during the impacts, the random data shuffling prior to training, the different training/testing data splits, and the different AE data features used for clustering. In the end, the NDI results validated the AE categorization results.

4 | Measurement Technology

4.1 | SHM: Acoustic Emission

As briefly explained in Section 1, AE corresponds to ultrasonic waves generated by sudden changes and/or degradation in the material. AE signal features are known to be sensitive to different damage mechanisms and/or damage progression stages. Several studies [20–23] have attempted to manually establish unambiguous correlations between AE signal features and damage mechanisms in composites by testing carefully produced coupons under specific and well-controlled conditions. In short, this involves isolating the different damage modes or mechanisms responsible for failure and observing the corresponding AE signal amplitude and frequency patterns.

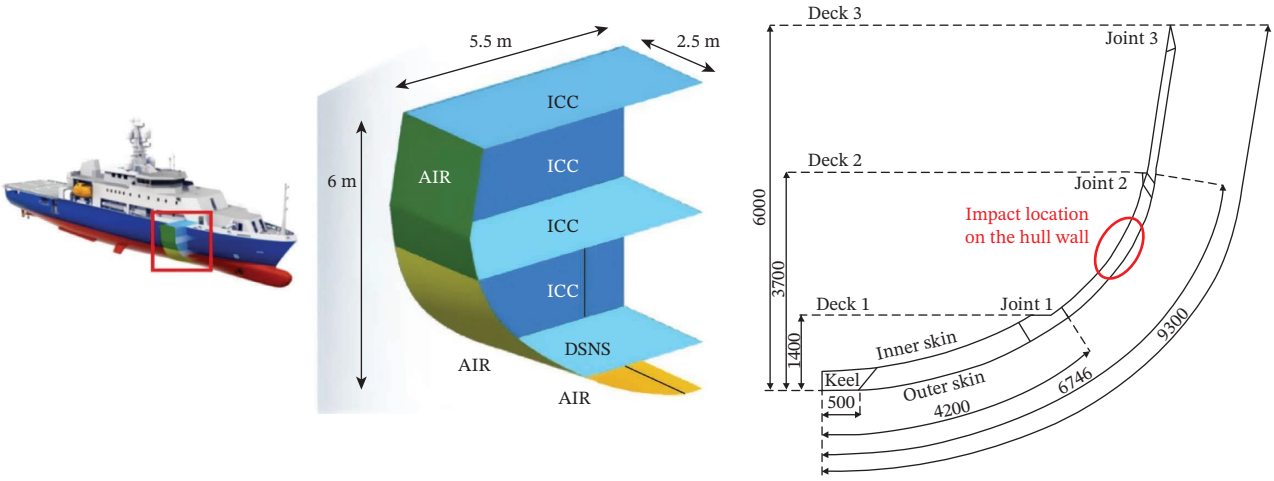

FIGURE 1 | Setting of the full-composite ship hull demonstrator, manufactured in collaboration with InfraCore Company (ICC) and airborne international (AIR) at Damen Schelde Naval Shipbuilding (DSNS) facilities.[10] Cross-sectional diagram of the hull demonstrator with the positions of the three decks and the impact location on the hull side wall. All dimensions are in mm.

However, in realistic scenarios, the loading and the occurrence of damaging events are not well-controlled. Consequently, degradation is not governed by a single, predominant damage mechanism, but rather a combination of them, as explained in Section 1 of the publicly available support material [24]. As a result, AE signals contain the signature of these interactions, and thus, one should analyze multiple AE features simultaneously to obtain a more differentiated identification of the multiple damage mechanisms. This constitutes a case of high-dimensional analysis, which is difficult to accomplish manually.

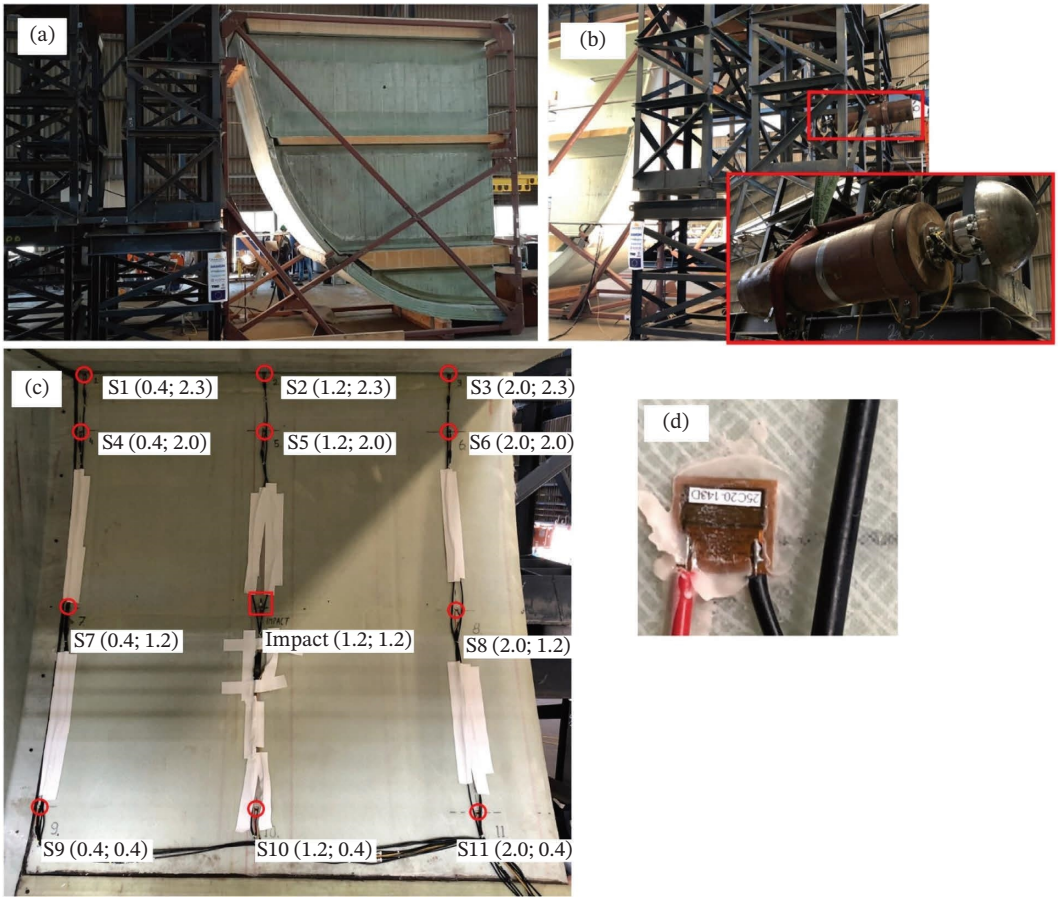

FIGURE 2 | (a) Demonstrator in a steel frame, next to the impactor supporting structure. (b) Impactor at a lifted position. (c) Sensor network on the inner surface of the hull. At the impact location (approximately at the center), an accelerometer was attached to the structure. Sensor location coordinates are in meters and measured from the bottom left corner. (d) Detailed view of the AE sensor at Position S8.

Therefore, the application of clustering algorithms to AE data has been researched by several authors [18, 25–29] and is considered a powerful approach to find complex patterns that can be used to detect and identify different damage mechanisms and/or damage progression stages when their category is not known a priori.

The hull was instrumented with 11 Smart Material M0714-P2 macrofiber composite (MFC) piezoelectric transducers (see Figure 2(d)) for AE sensing during the impacts. The sensors were adhesively bonded with cyanoacrylate to the inner surface of the hull wall, between Decks 1 and 2, at Positions S1–S11 identified in Figure 2(c) (the impact location is also identified in Figure 2(c); the spatial sensor distribution relative to the impact location and triangulated damage locations can be found in Figure 3). The AE monitoring system consisted of 11 AE sensors connected to a Mistras Sensor Highway II acquisition unit. After sensor installation, the AE amplitude detection threshold was calibrated through pencil-lead break measurements for each channel. The pencil-lead breaks were sequentially performed in between each two sensors, and the corresponding AE signals were acquired. The threshold was tuned to 55 dB, which enabled the detection of the pencil-lead breaks in all sensors. Since the high-energy impacts took place in the middle of the hull wall (Figure 2(c)), the rectangular sensor grid represents an optimal configuration, with sensors no more than 80 cm from each other. The upper most line of sensors (S1–S3) was placed close to the joint with the middle deck, especially to enable the detection of AE possible damage events at the joint. The built-in band-pass filter was employed for all channels, in a frequency band of 20–1000 kHz. The hit definition time (HDT), hit lockout time (HLT), and peak definition time (PDT) were 800, 1000, and 200 μs, respectively.

Additionally, both GoPro and high-speed cameras captured the impact test series on video. The impact acceleration generated by the impactor pendulum was measured by means of two Endevco 2262A accelerometers on the impactor mass. These accelerometers have a measuring range of 2000 g and were connected to a DAQP-STG conditioner/amplifier with a 20 KHz sampling rate. An accelerometer was also placed on the structure, on the inside of the wall across the impact location (see the square in Figure 2(c)) in order to capture the acceleration on the hull wall itself.

Multiple different AE signal features can be used for analysis [16–19], some of them being directly extracted and others being the result of mathematical operations between different directly extracted features. Based on the experience of the authors and the reviewed literature about AE monitoring of composites [16–19, 24], the following six AE signal features were chosen for analysis in the current study: (1) maximum amplitude; (2) duration; (3) peak frequency; (4) centroid frequency; (5) risetime-to-amplitude ratio; and (6) waveform information entropy. More information about the definition of these features can be found in Section 2 of the publicly available support material [24].

4.2 | NDI: Shearography

In this work, shearography inspection was conducted on the impacted hull section to validate and complement the SHM results. Ultrasonic testing and thermography were also used to benchmark this on-site inspection. However ultrasonic testing failed to produce sufficient inspection results due to the attenuation, multiple reflections, and scattering of ultrasound signals, as well as the high surface roughness of the object when inspecting thick composite structures. Thermography also yielded poor inspection results, as it is difficult to heat a large structure evenly and to avoid rapid heat dissipation. Therefore, ultrasonics and thermography are not included in this paper. Shearography [30, 31] is a well-known full-field optical NDI method, which has gained considerable industrial acceptance for inspecting composite structures up to 50 mm thick [30, 32, 33]. Shearography is capable of revealing defects including impact damage and delaminations. By applying an appropriate thermal or mechanical load to a test object, the defects can be localized and characterized by the analysis of anomalies in the inspection phase maps. The theory of shearography is well documented in the literature [30–32]. Normally, the in-plane imaging resolution of shearography for detecting internal damage (e.g., fiber breakage or matrix cracking) is up to the shear distance [31], which is around 9 mm in this paper; a phase measurement sensitivity of $2\pi/10$ rad can be achieved in a practical shearography experiment. For this study, one channel from the previously developed 3D shape shearography system [34] was adopted with the main focus on the out-of-plane deformation in which defect-induced deformation is expected to be dominant.

The shearography setup used for inspecting the composite ship hull section with impact damage is shown in Figure 4. The impacted area of the hull section was illuminated with a Torus 532 laser source (optical power of 200 mW and wavelength of 532 nm), forming a speckle pattern. The speckle interferograms were recorded by a shearing camera consisting of a Michelson interferometer with temporal phase shifting and a Basler Pilot piA2400 camera with a Linos MeVis-C 1.6/25 lens. A three-step phase-shifting algorithm was selected for the thermal loading scenario (Figure 4(b)) due to its fast speed, and a five-step phase-shifting algorithm was chosen for the mechanical loading scenario (Figure 4(c)) due to its resulting accuracy [30]. The shear distance was determined experimentally to ensure the quality of phase maps.

During the inspection with thermal and mechanical loadings, the shearography setup was fixed on a height-adjustable motorized rigid platform. The field of view (FOV) of the camera is about $600 \times 600\ \text{mm}^2$. Thermal loading was applied by three 1000-W halogen lamps. The effective surface temperature during heating was monitored with a FLIR A655 infrared (IR) camera. Mechanical loading was applied by a 7-ton metal block on the top deck of the structure, lifted by a gantry crane. This mechanical loading represented the landing of a medium helicopter on the top deck.

5 | Clustering Framework

A set of algorithms was selected based on their prevalence in the reviewed literature [18, 25–29, 35, 36], including both hard- and soft-clustering algorithms. The investigated set of algorithms included K-means, K-medoids, fuzzy C-means (FCM), and "*density-based spatial clustering of applications with noise*" (DBSCAN). The research process that guided us to the final adopted algorithm was incremental. The first trials were with K-means and K-medoids (two hard-clustering algorithms), as a way to grasp the clustering potential of the available data. The results showed that several subgroups of points were considerably

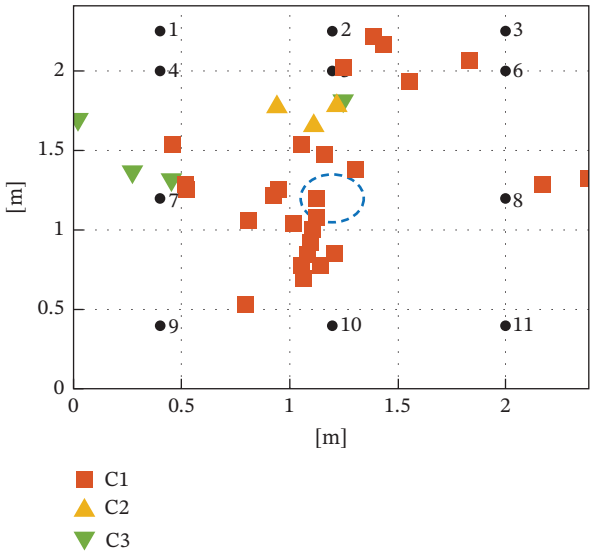

FIGURE 3 | Acoustic emission sensor positions and estimated locations of the damage events for all impacts. The approximate impact location is indicated by the blue dashed circle.

distant from the centroid of the cluster they were assigned to. These distant points themselves could potentially form a cluster. To explore whether these distance points could indeed form a new cluster, it was decided to try FCM (soft-clustering algorithm), which allows data points to be members of different clusters. We implemented a two-step approach to investigate a robust implementation of the FCM algorithm, as follows: firstly, iteratively determining the optimal number of clusters that maximizes the silhouette score while using a minimal fuzziness exponent and, secondly, simultaneously optimizing the hyperparameter fuzziness exponent and the cutoff degree of membership by maximizing the silhouette score.

In the process of trying to solve the issue of outlier detection for FCM, DBSCAN presented itself as a good compromise between the desired unambiguous cluster definition and the statistically correct identification of noisy points (i.e., outliers). Unlike the other studied algorithms, DBSCAN groups points that are closely packed to one another (i.e., in densely populated areas), without requiring a fixed hyperdefinition of the number of clusters. This is a very important characteristic for our research, as it is crucial to be able to classify the damage occurring in a physically meaningful way, in a real scenario where the number and type of damage mechanisms are not known in advance. Additionally, DBSCAN is inherently capable of identifying outliers, which is very useful for handling potentially noisy AE datasets in a statistically correct way. On top of all this, the already demonstrated robustness of the DBSCAN algorithm [37–39] makes the implemented damage categorization framework inherently robust, which was crucial to accomplish our research goal.

The DBSCAN algorithm requires two hyperparameters, which are the search vicinity radius (*eps*) and the minimum number of points (*minPts*) required to form a cluster. Optimizing the two hyperparameters accordingly and automatically is of essence. If the *eps* parameter is too small, all the data will be classified as noise, whereas if the *eps* parameter is too large, all the data will be seen as one big cluster. In a similar sense, if the *minPts* parameter is too small, all the noise points are classified as clusters, and if the *minPts* parameter is too large, then actual clusters will be classified as noise. In our implementation, we applied the grid search optimization technique in two phases in order to find the optimal set of two hyperparameters that maximize the silhouette score.

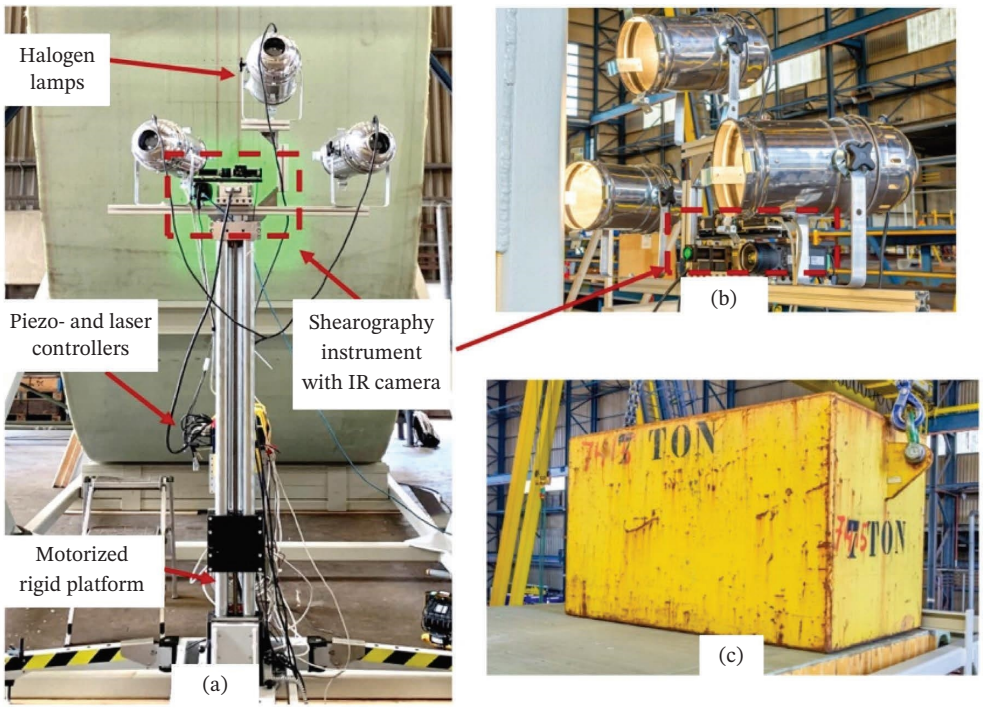

FIGURE 4 | (a) Shearography inspection setup; (b) thermal loading; (c) mechanical loading (on top deck).

The implemented framework establishes a series of operations, namely, for automatic hyperparameter optimization, automatic feature selection, outlier detection, and solution evaluation, aimed at increasing the reliability and stability of the training procedure, even in scenarios of limited data availability. An explanation of the implemented framework (including the hyperparameter optimization), along with a basic explanation of the core clustering algorithm, can be found in Section 3 of the publicly available support material [24].

It is worth noting that many clustering methods are already embedded in the software accompanying AE devices. However, to the best of our knowledge and experience, the clustering functions implemented in commercially available AE equipment often do not employ state-of-the-art algorithms as those investigated in this research (e.g., DBSCAN) and are lacking robustness. This is precisely one of the gaps that this research intends to bridge. More specifically, the approach implemented in this work provides extra adaptiveness by leveraging on automatic statistical relevance–based feature selection and on the inherent capability of DBSCAN for finding outliers and the best number of clusters. All of this is achieved without requiring prior, uninformed hard definitions by the user.

6 | Results and Discussion

6.1 | SHM

The performance evaluation of clustering implementations should focus on data availability and utilization scenarios that mimic the real operation of the algorithm. Therefore, it was decided to organize the AE data in the following two different ways in order to address two different scenarios. Use Cases 1 and 2 are presented in Sections 6.1.1 and 6.1.2, respectively.

- Use Case 1: x% of all data for training, 100 – x% for testing (offline training).
- Use Case 2: Mimic chronological data availability for training and testing (online training).

6.1.1 | Damage Identification: Use Case 1

6.1.1.1 | Automatically Selected Features. An algorithm can be trained before deployment by using historical data (offline training). A recommended practice in machine learning to avoid overfitting issues is to use an x% fraction of the available data for algorithm training and the remaining 100 – x% fraction for validation and/or testing. Typically, the data are randomly shuffled before being split, which makes clustering solutions dependent on that random data shuffling. Therefore, it is relevant to look at the evolution of the obtained number of clusters, mean silhouette, and outlier percentage with increasing number of algorithm executions, for different data splits. For one thousand algorithm executions, the results of clustering with the three most relevant features, for 60/40, 70/30, and 80/20 data splits, are summarized in Table 1. The reader should note that the three most relevant features are amplitude, peak frequency, and centroid frequency, regardless of the method used for their automatic selection.

Firstly, it is relevant to mention that the average mean silhouette is almost 0.9 and the average outlier percentage is below 5%, which shows that the clustering framework is capable of obtaining well-formed solutions. Secondly, the statistics show that, with random shuffling of the present dataset prior to data splitting, the implemented DBSCAN framework can find either two or three clusters.

A common recommendation for relatively small datasets is to use 70% of the available data for training and 30% for validation or testing [40]. An example of a two-cluster solution obtained with this split and the three most relevant features is presented in Figure 5(a). If amplitude (A_{max}) and centroid frequency (f_{peak}) are considered, the separation between Cluster 1 (C1) and Cluster 2 (C2) is clear, apart from some fuzzy points in between the two clusters. Looking carefully at C2, it is possible to distinguish two regions within that cluster, one between 125 and 160 kHz, and another one between 190 and 220 kHz. By looking at a three-cluster solution obtained with the 70/30 data split and the three most relevant features plotted in Figure 5(b), one can immediately recognize the two previously identified subregions as now being Clusters 2 and 3. Hence, it is possible to conclude that because the separation between these two subregions is much smaller than the separation between them and C1, the random data point shuffling prior to clustering has a strong influence in the identification of these two regions as separate clusters.

We consider it insightful to compare the clustering solutions presented above with the frequency content intervals found by Pashmforoush et al. [25] and Saeedifar and Zarouchas [18] to be associated with different damage mechanisms in composite sandwich specimens. In particular, Pashmforoush et al. [25] found that the intervals 35–65, 100–130, 170–250, and 350–450 kHz were associated with core crushing/cracking, skin-core debonding, matrix cracking, and fiber breakage, respectively. The correlation of the categorization clusters with frequency bands from the reviewed literature allowed us to come to a first set of possible damage mechanisms. It is important to emphasize that this correlation was not solely based on the "blind" matching of clusters and frequency bands, but rather a physics-informed process enabled by the authors' knowledge of damage mechanics in composite materials. Given the facts explained in Section 4 of the support material [24] and the high mean silhouette value of the clustering solution in Figure 5(b), it is possible to hypothesize that there is a strong correspondence between C1 (19–58 kHz) and core crushing, between C3 (115–147 kHz) and skin-core debonding, and between C2 (191–217 kHz) and matrix cracking. This correspondence is supported by the chronological analysis made in Section 6.1.1.2. In particular, the progressions plotted in Figures 6(d), 6(e), and 6(f) are consistent with the principles of composite fracture mechanics explained in Section 1 of the support material [24], thereby giving physical meaning to the hypothetical correspondence between the cluster frequency bands and the results by Pashmforoush et al. [25]. It is, therefore, hard to consider such correspondence accidental or coincidental.

These results point to the possibility that damage occurred predominantly in the form of core crushing/cracking during the impacts, which may have subsequently induced some skin-core debonding, along with some matrix cracking events. It is good to

TABLE 1 | Statistics for one thousand DBSCAN executions with the three most relevant features for three different data splits.

Data split	60/40	70/30	80/20
Two-cluster solution percentage	78.2%	66.5%	69.7%
Three-cluster solution percentage	21.1%	33.4%	30.3%
Average mean silhouette	0.888	0.887	0.890
Average outlier percentage	4.46%	3.96%	4.04%

remind that the impact tests were not designed to ensure the induction of a specific, well-controlled progression of composite damage mechanisms "upstream." Therefore, we adopted a "downstream" validation approach: A postimpact validation of the AE categorization results was conducted with NDI techniques and detailed video footage. The validation of the damage mechanism categorization proposed above is presented in Section 6.2 and Section 6.3.

At this point, it is relevant to highlight two aspects. Firstly, for both two-cluster and three-cluster solutions, with both automatically and manually selected features (solution obtained with manually selected features can be found in Section 5 of the publicly available support material [24]), the testing data points are well categorized according to the trained clusters. This gives confidence that the implemented clustering framework is capable of performing damage mechanism identification reliably, even with limited training datasets. Secondly, in our study,

- composite damage progression occurred unconstrainedly, on a 1:1 scale complex structure;
- random data shuffling prior to training is incorporated in the implemented method;
- different data splits for training and testing were evaluated;
- different AE data features were used for clustering.

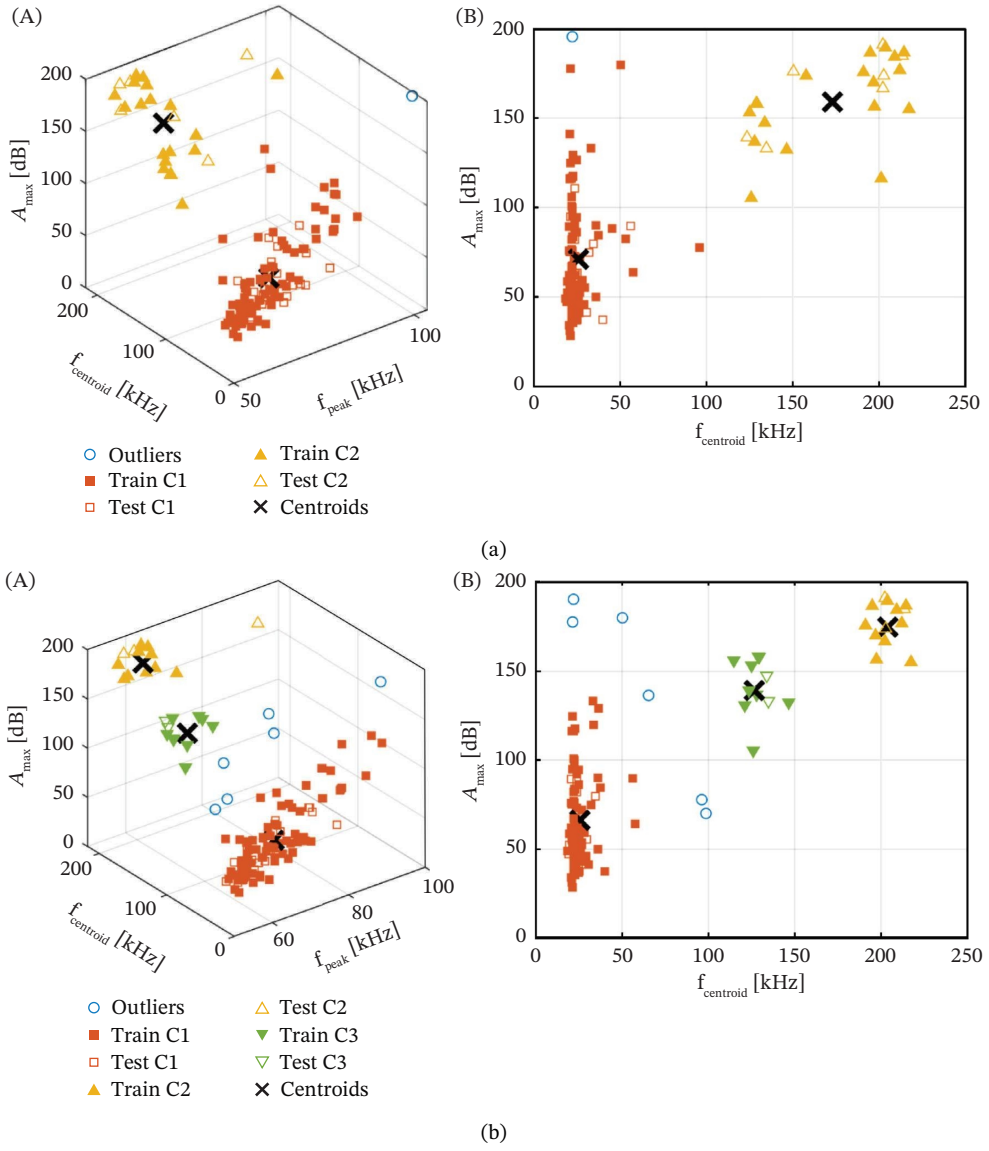

FIGURE 5 | 3D ((a), A) and 2D (A, right) views of the two-cluster solution obtained with 70/30 data split and the three most relevant features. 3D ((b), B) and 2D (B, right) views of the three-cluster solution obtained with 70/30 data split and the three most relevant features. A few apparent data point discrepancies between (a) and (b) relate to how the entire clustering framework has been implemented*.

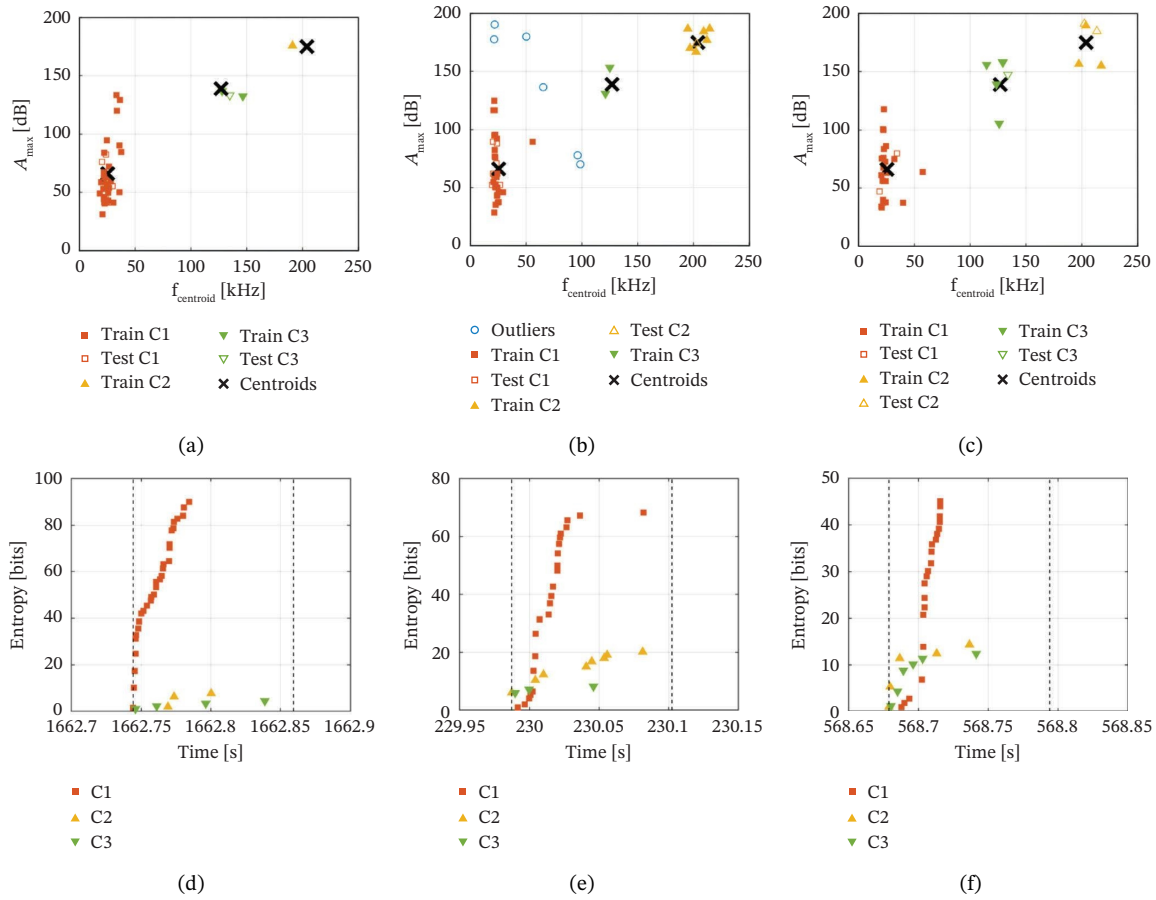

FIGURE 6 | (a–c) 2D views of the three-cluster solution obtained with 70/30 data split and three most relevant features, separately plotted for each impact. (d–f) Cumulative information entropy as a function of time for the training and testing datasets, for each impact. The time interval indicated by the vertical dashed lines corresponds to the duration of the impact.

- and yet, for all the different evaluations, the damage categorization framework always produced consistent results, which is strong evidence of stability.

6.1.1.2 | Chronological Analysis. For the three-cluster solution obtained with the three most relevant features and a 70/30 data split (plotted in Figure 5(b)), data points can be separately plotted according to the impact during which they were acquired, as in Figures 6(a), 6(b), 6(c). One can see that in Impact 1, there is a clear predominance of C1 points (19–58 kHz), with very few C3 (115–147 kHz) and C2 (191–217 kHz) points. Hence, there seems to be a clear predominance of core crushing/cracking over skin-core debonding and matrix cracking. In Impacts 2 and 3, the number of C1 occurrences decreases, while that of C3 and C2 substantially increases, remaining approximately the same for both impacts. This situation seems to indicate the growth of skin-core debonding and matrix cracking after their initiation due to the weakening effect of core crushing/cracking. This last observation highlights the importance of investigating the distinction between damage initiation and growth. For that, it is useful to analyze the clustered AE hits chronologically. Firstly, it is helpful to determine the duration of each impact, i.e., the time interval during which the impactor was in contact with the hull section. The impactor acceleration response during the impacts is plotted in Figure 7. It is possible to see that the duration of all the impacts was the same and it was roughly equal to 0.12 s. Knowing this, it is then possible to plot the cumulative information entropy for the clustered data (training and testing sets together) as a function of time for each impact (see Figures 6(d), 6(e), and 6(f)).

For Impact 1, it is clearly visible that C1 points occur right from the start of the impact and at a high rate, while C2 and C3 occur at a very low rate and C2 points only start 0.025 s later. This sequence of occurrences is in accordance with the damage mechanics explained in Section 1 of the support material [24]. For relatively large core thicknesses and in out-of-plane loading scenarios, core crushing/cracking tends to have the lowest onset energy threshold. The significant reduction of core stiffness weakens the core-skin bond, thereby leading to skin-core debonding when the bond strength is exceeded by interface stresses. At the same time, high skin stresses caused by impactor contact and skin flexion are expected to induce skin matrix cracking. So, skin-core debonding (C3) and skin matrix cracking (C2) only occur after some accumulation of core crushing/cracking (C1). For Impacts 2 and 3, points from all three clusters occur right from the start of the impact and at comparable rates, which seems to indicate the progression of the damage mechanisms initiated in previous impacts.

6.1.1.3 | Damage Localization. The built-in location function of the Mistras AEwin software [41] was used to find the

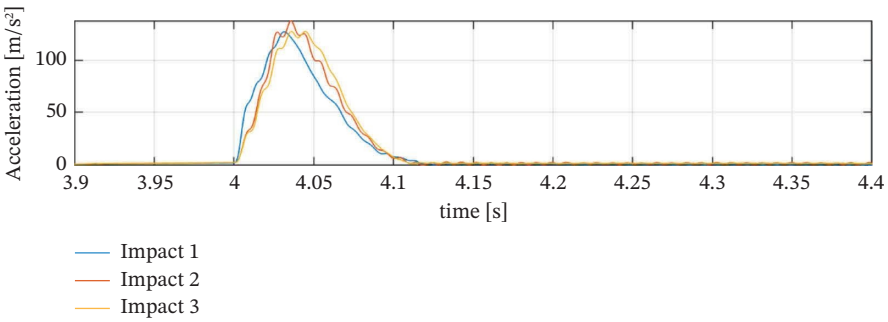

FIGURE 7 | Impactor acceleration response during each impact.

coordinates of the AE events associated with damage locations. Basically, the location of an AE source is estimated by taking the triangulation problem for multiple sets of hits and determining a solution by means of multiple regression optimization analysis. When that optimization converges, the AE event is considered to be localized to some degree of certainty. The damage events whose localization converged in the multiple regression analysis for all three impacts are presented in Figure 3. The events associated with the different damage mechanisms are identified according to the categorization obtained from the clustering solution presented in Figure 5(b). It should be noted that the localization map of the present study should not be taken as ground truth, as there are always errors related to signal time-of-flight picking and group velocity estimates used for distance calculation.

Core crushing/cracking (C1) occurrences formed most of the high event density region around the impact location, although some of them could also be located in the top and right sides of the hull wall. This was the predominant damage mechanism during all the impacts, as previously demonstrated in Figure 5(b). Some matrix cracking (C2) events were located relatively close to the majority of the core crushing events, just slightly above the center of the hull wall. Finally, some skin-core debonding (C3) events were found mostly along the left side of the hull wall.

As explained in Section 1 of the publicly available support material [24], due to the highly anisotropic constitution of composite structures, the sudden release of energy can trigger damage spots at multiple locations of the component. As damage progresses from these multiple spots, the damage accumulation ends up being nonlocal. The damage progression sequence can take place at different rates, at different locations, depending on the local strain rate induced by the loading. Consequently, it is possible to simultaneously have different damage mechanisms at different locations of the structure.

The localization map is an important demonstration of the combination of the localization capabilities of AE systems and the implemented unsupervised categorization framework. It provides valuable insights about the location where the predominant damage mechanism has accumulated, thereby helping with the identification of the region(s) to be inspected in detail with NDI techniques.

6.1.2 | Damage Identification: Use Case 2

Use Case 1 (Section 6.1.1) represents a scenario where the SHM system is trained with historical data before deployment, as AE hits from all three impacts are employed for algorithm training. Use Case 1 is relevant to demonstrate the reliability of the implemented clustering framework. An alternative (or additional) approach is to have an SHM trained online during ship operation to keep improving diagnosis accuracy. So, every time new structural events take place, a new dataset is acquired and interpreted with the trained capabilities at that moment. Then, that new dataset is used for retraining the algorithm, thereby building up experience until the next new dataset. In Use Case 2, it was decided to mimic this chronological data availability and the process of online training. Therefore, the executed sequences were as follows:

1. Impact 1 for training and Impact 2 for testing;
2. Impacts 1 and 2 for training and Impact 3 data for testing.

6.1.2.1 | Impact 1 for Training and Impact 2 for Testing. The clustering solution obtained with Impact 1 as the training dataset and Impact 2 as the testing dataset and with the three most relevant AE features is presented in Figure 8. As seen in Figure 8(A1), only one cluster is found and 37.2% of the Impact 1 data points are considered outliers, meaning that the DBSCAN algorithm could not find more than one properly formed dense region. With only one cluster, it is not possible to compute silhouette values, because there is no intercluster distance to be taken, and as such, there is no quantitative measure of the suitability of the clustering solution. Nevertheless, the obtained cluster corresponds to the same damage mechanism type that was identified in Section 6.1.1 as being core crushing/cracking.

When testing the categorization of the Impact 2 dataset according to trained Cluster C1 (see Figures 8, A2), only a small fraction of the points is successfully classified as core crushing/cracking, while around 92.5% is considered outliers. The existence of only one cluster is very detrimental for the correct categorization of the damage events that occurred during Impact 2, because there is basically no flexibility to account for the inherent variability within each damage mechanism type. As a result, Impact 2 data points that are not within the optimal vicinity radius (hyperparameter *eps* as determined during the DBSCAN training process, see more about the hyperparameter optimization in the publicly available Supporting Material [24]) of the closest cluster centroid are considered testing outliers.

6.1.2.2 | Impacts 1 and 2 for Training and Impact 3 for Testing. When both Impacts 1 and 2 datasets are used for training, two properly formed clusters are found, with a very low outlier percentage (see Figures 8, B1). All the points considered

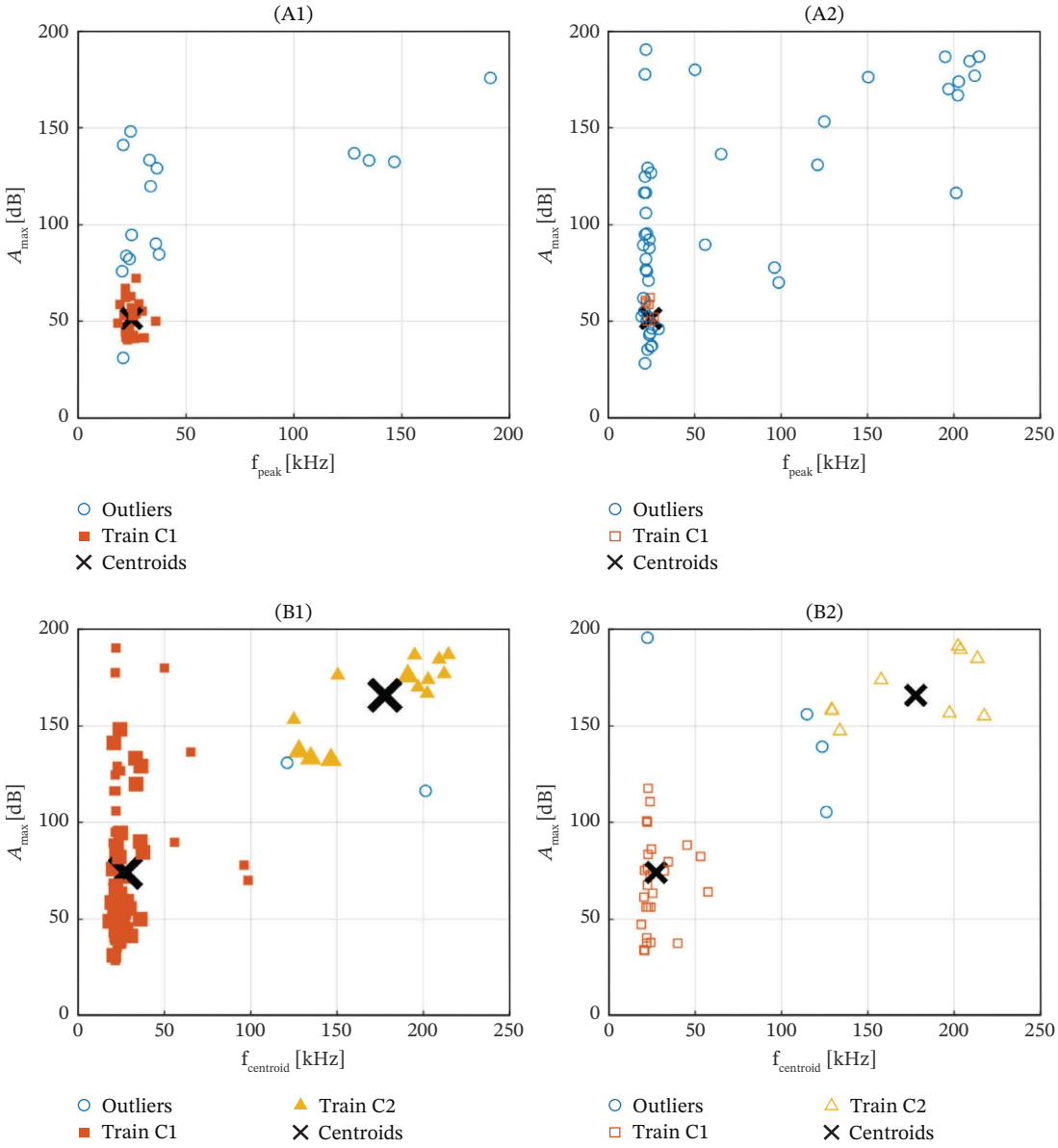

FIGURE 8 | Row A—2D views of cluster solution obtained with Impact 1 as the training dataset and Impact 2 as the testing dataset, and with the three most relevant features: (A1) trained cluster; (A2) tested categorization of Impact 2 dataset. Row B—2D views of cluster solution obtained with Impacts 1 and 2 as the training dataset and Impact 3 as the testing dataset, and with the three most relevant features: (B1) trained clusters; (B2) tested categorization of the Impact 3 dataset.

outliers when the training was performed with only Impact 1 data are now included in Clusters C1 and C2, and only two Impact 2 points are considered outliers. However, while C1 is quite compact (apart from six points on the diagonal transition between C1 and C2) and has a frequency signature that is strongly correlated to core crushing/cracking, C2 is more dispersed and cannot be unambiguously correlated to a single damage mechanism. Instead, it includes frequency signatures typically associated with skin-core debonding and matrix cracking (as explained in Section 6.1.1.1). This situation is due to the considerably lower number of points in C2 than in C1 and shows that core crushing was by far the predominant damage mechanism taking place during the impacts.

The existence of two clusters allows the great majority of the Impact 3 points to be classified, leaving only 10% of outliers in the region between clusters C1 and C2. Because the trained C2 centroid was between 150 and 200 kHz, and no third cluster was found below 150 kHz, the points below 125 kHz that are still far away from C1 remain ambiguous.

While Use Case 1 was especially important to demonstrate the reliability of the implemented clustering framework, Use Case 2 is very useful to highlight its stability again. It shows that in Online Training Scenario 2 (Impacts 1 and 2 were used for training and Impact 3 was used for testing), the categorization performance was comparable to when data from all three impacts had been used for training (i.e., as in Use Case 1, in Section 6.1.1), with an outlier percentage of only 2.08% (Figures 8, B1, B2). Due to the inherent principles of the DBSCAN algorithm and the criteria used for testing categorization with the trained clusters, ambiguous points are not forced to a specific cluster.

Instead, they are considered outliers, which can eventually work as a flag or alarm. Furthermore, the very significant improvement in damage mechanism categorization accuracy obtained by just adding 53 extra points (i.e., going from training with only Impact 1 to training with Impacts 1 and 2) is a clear demonstration of the adaptability of the implemented framework to different data availability scenarios.

6.2 | NDI

6.2.1 | Shearography

Although it could be beneficial to use shearography detection before and after each impact, shearography was only performed after the third impact because of the project complexity and logistic issues. Before the on-site inspection of the composite ship hull section, we did a preliminary test in the lab environment on a representative sandwich panel with a similar structure, material composition, and impact damage. The impact damage in the test panel was reliably detected with 5–10 min of heating with 6–12 mm shearing distance. Those preliminary results provided guidance for the on-site inspection. Damages including fiber breakage along with matrix cracking, skin-to-core debonding, and delaminations were detected in the test panel. Based on the test panel, we expected to get similar detection capabilities of shearography for the on-site inspection.

During the on-site inspection, first, a whole-field inspection of the hull shell was performed using shearography for an overall evaluation of the structure. The area of the whole-field inspection was large enough (around $1500 \times 900\ mm^2$) and most of the area of the outer hull shell looked healthy, excluding the regions near the impacts (Figure 9(e)) and one local region near the edge. It was assumed that the condition of the ship hull surface before the impacts was pristine, i.e., as manufactured.

Then, the detailed shearography inspection of the impacted area on the composite ship hull shell was performed. The impacted area of the ship hull section was thermally as well as mechanically loaded. For thermal loading, 240-s heating was performed over the region of interest with the impacted area in the center. The selection of the heating time was based on preliminary results from a representative sandwich panel with impact damage. During the inspection, the surface temperature increase was in the range of 8°C–15°C. During the cooling of the structure, sets of phase-shifted speckle interferograms were continuously recorded and later processed. Their total build-up results in phase maps that indicate the evolution of the out-of-plane displacement derivative with shear direction. Phase maps in both x and y directions were measured (see Figures 9(b) and 9(c)).

As explained in Section 4.2, the mechanical loading was applied by lifting and placing a 7-ton metal block on the top deck of the hull section. Sets of phase-shifted interferograms with the y-shear were recorded at the end of each loading step as the vertical compression was expected to be dominant. The processing of the recorded phase-shifted interferograms yields the overall mechanical response of the structure (Figure 9(d)).

For both loading types, total phase maps (e.g., Figure 9(a)), which contain defect-induced deformation as well as overall deformation (thermal expansion or mechanical bending), were obtained. To separate the two types of deformation, the defect-induced deformation was extracted from the overall phase map by a curve fitting process [33], which provided compensated phase maps (Figures 9(b), 9(c), 9(d)). The compensated phase maps for thermal loading and x and y shears (Figures 9(b), 9(c), respectively) both reveal the presence of the strain anomalies due to the impact damage (indicated by the oblique red arrows in Figures 9(a), 9(b), 9(c), 9(d)). The results for mechanical loading (Figure 9(d)) reveal the same damage. The strain anomalies found are likely to correspond to skin-to-core debonding.

The fact that the same defect was detected by both thermal and mechanical loading demonstrates the reliability of shearography inspection. It can be noted that the inspection results from mechanical loading (Figure 9(d)) seem to be blurry, which is likely to be due to the rigid body motion of the hull section induced by the loading procedure. The vertical orange arrows in Figure 9(b) and *c* are more difficult to interpret. They could indicate some skin matrix damage that may have been caused by the scraping of the impactor against the hull wall during the bouncing of the impactor after actual impact (red ellipse in Figure 9(e)). In addition to impact damage, in the compensated phase map for thermal loading and x shear, the vertical line–shaped anomalies (indicated by three white horizontal arrows in Figure 9(b)) seem to be associated with composite plie overlaps.

As a result, shearography captured three damage types (delaminations, fiber breakage, and matrix cracking) due to its high sensitivity to stiffness variations caused by impact damage. At macroscopic damage of this scale and severity (high impact energy of around 20 kJ), all these failure modes occur simultaneously and are therefore detected together.

6.2.2 | Coin Tapping

The contact area of the impactor on the hull wall was clearly marked with rust dust left by the impactor head. After each impact, the area on and around the impactor contact area was inspected visually (with the naked eye) and by coin tapping [42–46]. After the second impact, evidence of some skin-core debonding was identified by coin tapping, as a duller/more damped sound in the impact area could be heard compared to the less damped sound outside the untouched area. This area was less than 20×30 cm, as depicted in Figure 9(e).

6.3 | Video Footage

In the slow-motion video of Impact 2 (included as the second video in the support footage, see the Data Availability section), it is possible to witness the generation of local damage locations in two regions along the hull-wall edge facing the high-speed camera (indicated in Figure 10(A)).

Those occurrences are described in this section by resorting to video screenshots from the instant before the impact, the moment of impact, and the instant immediately after the impact. In Figure 10, Row B, one can observe that there is some sandwich-core crushing/cracking visible on the edge of the hull wall. This cracking may be connected to some skin-core debonding. In Figure 10, Row C, it is possible to see that the flexural action during the impact induces a whip effect, which in turn generates some peeling, and thereby some skin-core debonding at the edge of the hull wall.

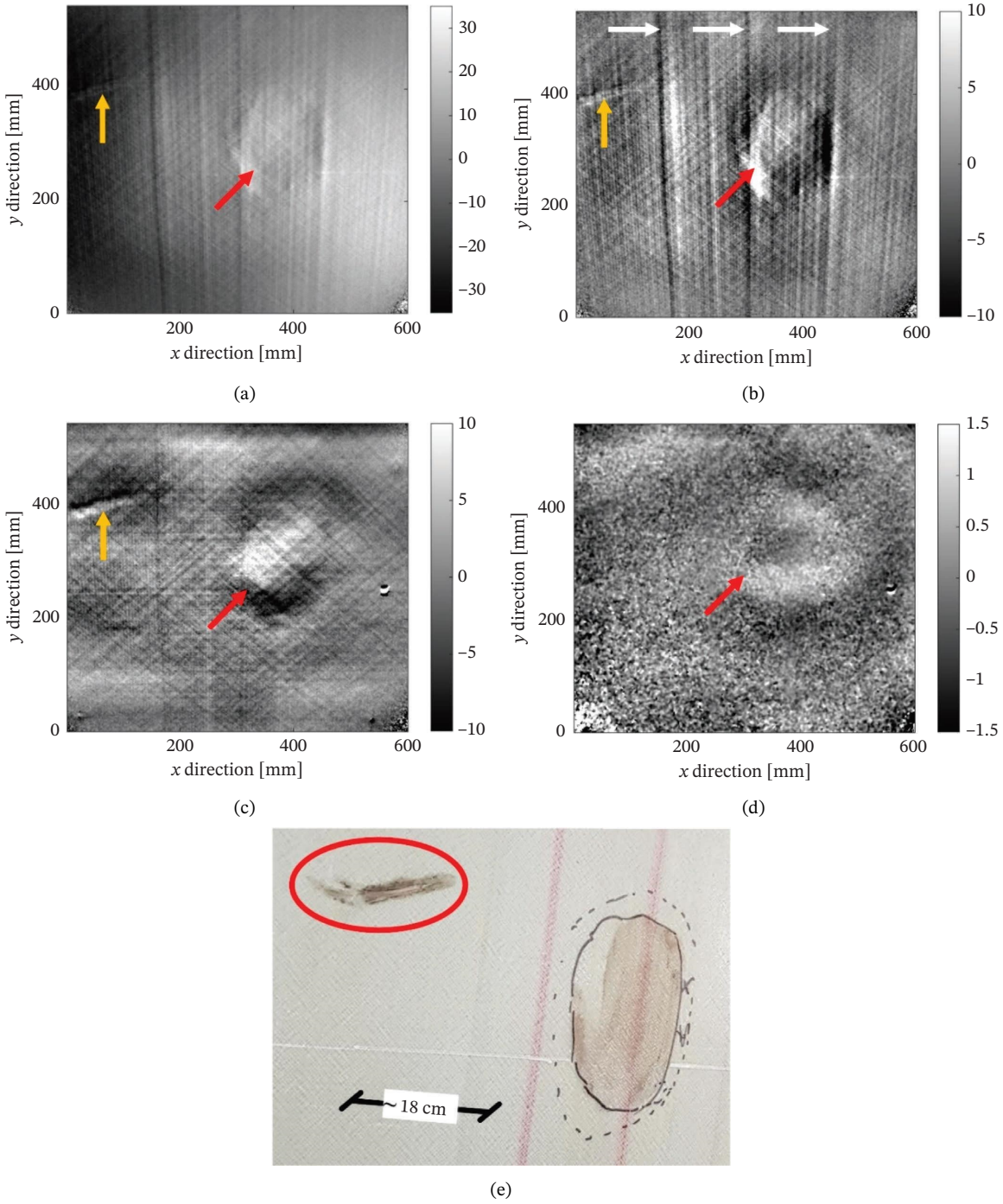

FIGURE 9 | (a–d) Shearography inspection results (units in radian): (a) Original phase map with thermal loading and x shear; (b) compensated phase map with thermal loading and x shear; (c) compensated phase map with thermal loading and y-shear; (d) compensated phase map with mechanical loading and y-shear. The red and yellow arrows indicate possible damage locations discussed in the text. (e) Rust marks left by the impactor on the hull wall. The scratch due to bouncing is highlighted by a red ellipse. The mark on the right is due to the actual impact. The sound difference limits from the coin-tapping inspection (see Section 6.2.2) were marked: continuous line for unambiguous differences, dashed line for ambiguous differences.

At this point, it is important to highlight that, although the reliability of the implemented damage categorization framework was not quantitatively estimated by means of a reliability index, we consider that the slow-motion video footage presented above together with the results of the NDIs presented in Section 6.2 enables a suitable qualitative reliability assessment, and thus validation, of the damage categorization based on the clustering results discussed in Section 6.1.1.

6.4 | Considerations on the Limitations of the Study and Implemented Method

It is important to highlight some limitations of the study and methods. We attempted to generate the most realistic dataset possible with all the diversity that comes with a full-scale structure and representative impacts. However, due to project cost and scope limitations, only one hull section demonstrator

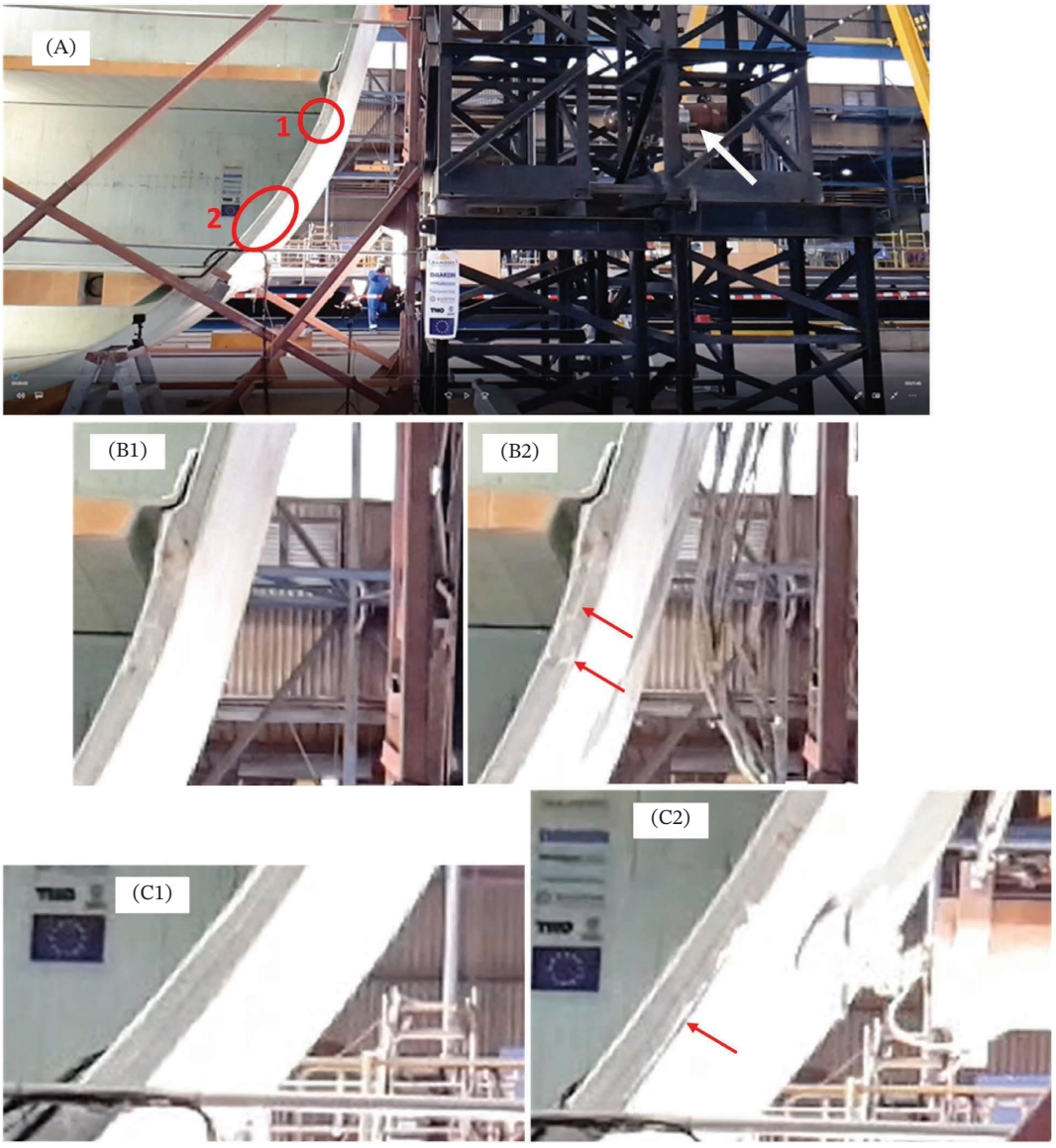

FIGURE 10 | (A) Screenshot from high-speed camera video of Impact 2, before impact (note the impactor in falling position, indicated by the white arrow on the right). Areas of Interests 1 and 2 are highlighted by red ellipses. Row B—zoomed-in views of Area 1: (B1) before the impact; (B2) immediately after the impact. Arrows indicate the damage. Row C—zoomed-in views of Area 2: (C1) before the impact; (C2) at the moment of impact. The arrow indicates the damage.

was manufactured and only impact scenarios were tested with AE instrumentation. This prevented the evaluation of performance of our damage categorization method on multiple specimens, under different loading conditions (e.g., fatigue) and during a representative part of the useful life of the structures. Consequently, it was not possible to study aspects related to integrating our methods into a complete SHM system and making it fit-for-purpose.

In thick composite sandwich structures, which inherently have a high material heterogeneity, damage progression is complex (see more about the complex damage progression in the publicly available Support Material [24]). In realistic loading scenarios, material heterogeneity leads to degradation not being governed by a single, predominant damage mechanism, but rather a combination of them. In such cases, AE signals contain the signature of complex damage progression interactions, and thus, one should analyze multiple AE features simultaneously to obtain a more differentiated identification of the multiple damage mechanisms. Hence, we decided to develop a categorization method that incorporates automatic relevant feature selection and automatic hyperparameter tuning, coupled with a robust clustering algorithm such as DBSCAN, which has the inherent capability of detecting outliers.

It should be noted that the approach developed and implemented in this work does not pertain to mitigate the interference of noise in AE data acquisition for practical SHM applications. However, as explained above, the approach is capable of analyzing noisy data in a stable and robust way, because the DBSCAN method can inherently identify outliers. Possible sources of environmental noise (e.g., sensor bonding loss with the structure during

impact) or nondamage-related events (e.g., friction between the impactor and hull) may influence the results. It was checked that all the transducers were fully bonded to the structure by the end of the three consecutive high-energy impacts. Moreover, if bonding loss with the structure had occurred, the frequencies associated with it would have been under the kilohertz range and would constitute an oscillatory drift of the mean amplitude. This was not observed in any of the acquired waveforms. Regarding friction between the impactor and hull, such types of events cannot be excluded. However, the different AE data analyses conducted in this study prove that this type of event did not influence the categorization results. The chronological analysis in Section 6.1.1.2 shows that the duration and the evolution of the different clusters during each of the impacts match the structure dynamics captured by the accelerometers during the impacts. This proves that the clustered data points are related to what was taking place inside the structure and not to impactor–hull friction.

In cases of slower structural degradation (e.g., fatigue), there might be occurrences in which different damage mechanisms have comparable contributions. The corresponding AE data points might then lie in between cluster "clouds." Because DBSCAN does not allow data points to have shared membership in different clusters (i.e., hard-clustering algorithm), in these cases, the points would likely be classified as outliers. To tackle this limitation, it would be beneficial to adopt the soft variant of the DBSCAN algorithm, which would allow fuzzy cluster membership. Hence, future research could enhance the applicability for real-time SHM applications by incorporating more diverse datasets and exploring more sophisticated soft-clustering techniques to handle ambiguous data points.

7 | Conclusions

This article focused on the AE-based SHM of a full-scale (1:1) full-composite ship hull section subjected to three consecutive realistic high-energy (~20 kJ) impacts. The investigations built up on the recognized potential of AE sensing for damage detection, localization, and categorization of damage in composite structures. The research goal was to develop and validate a reliable and robust approach to classify different damage mechanisms without prior knowledge of their occurrence, and that can be applied in a systematic way, irrespective of structure size and complexity.

To this end, the DBSCAN algorithm was combined with adapted routines for silhouette-driven hyperparameter tuning and for Laplacian data feature selection in order to enable a systematic deployment of the unsupervised damage categorization tool. The tool was able to identify three properly formed clusters, with around 6% outliers and 0.9 mean silhouette. The clustering solution was found to have a very good match with evidence from the literature about the AE frequency content for sandwich-core crushing/cracking, skin-core debonding, and matrix cracking in composite sandwich specimens made of similar materials. The damage categorization framework demonstrated strong stability by always producing consistent results, despite the unconstrained damage progression during the impacts, the random data shuffling prior to training, the different training/testing data splits, and the different AE data features used for clustering.

In the chronological analysis of the clustering results, it was possible to observe damage progression as an interdependent process of the three damage mechanisms. The onset of skin-core debonding and matrix cracking depended on the onset and accumulation of core crushing/cracking. Subsequently, the accumulation rate of the three damage mechanisms remained interdependent. This analysis reinforced the confidence that the implemented tool can identify all the predominant damage mechanisms. The localization capabilities of the AE acquisition system were combined with the clustering solution in post-processing, demonstrating the usefulness of a categorized damage localization map in providing extra insights into the spatial trends in damage progression.

On top of that, when the algorithm was trained consecutive times (with the training dataset incremented with an extra impact for each time), as it moved from training with only Impact 1 to training with Impacts 1 and 2 (i.e., by just adding 53 data points to the training set), two properly formed clusters were found, with a mean silhouette of 0.884 and only 2% outliers. In the subsequent testing with Impact 3, the great majority of the points were categorized, leaving only 10% of outliers in the region between the two clusters. In short, in one retraining iteration, it went from no categorization capability to a categorization capability that was comparable to when data from all three impacts were used for training. Hence, it was concluded that the incorporation of automatic relevant feature selection and automatic hyperparameter tuning, coupled with inherent DBSCAN outlier detection capability, provides a stable approach to handle realistic AE data variability and to adapt to different data availability scenarios that can be found in real operation, thereby showing potential for online monitoring. This increased the confidence in the ability of the framework to perform its mission, even when it is necessary to retrain the algorithm during ship operation to continuously improve damage diagnosis accuracy.

The AE-based damage mechanism categorization results were validated with postimpact NDI and analysis of slow-motion video footage captured during the impacts. Through digital laser shearography and coin-tapping, it was possible to confirm the presence of skin-core debonding at and around the impact location. This indirectly confirmed the occurrence of core crushing/cracking, as an extensive area of this damage mechanism must always precede skin-core debonding. Shearography measurements also contained some indications that could potentially be associated with skin matrix cracking, although this could not be assured. By carefully looking at slow-motion video footage, it was possible to identify events of core crushing/cracking and skin-core debonding occurring along one of the side edges of the hull section. In short, the reliability of the implemented unsupervised damage categorization framework was qualitatively demonstrated.

To the best of our knowledge, this is the first published validated study addressing the first three SHM functions of damage diagnosis (detection, localization, and categorization) for a full-scale full-composite ship structure subjected to realistic loading. The article demonstrates that the unsupervised damage categorization framework can perform its mission in an adaptive and totally automatic way. The study also highlights the capability of the developed categorization methodology to be deployed for

online monitoring, where SHM algorithms are retrained during structure operation to keep improving diagnosis accuracy.

Finally, this research provides an example of how SHM alerts generated during operation can help directing detailed inspections to when and where truly necessary (e.g., with shearography, ultrasonics, thermography). The implementation of SHM should aim at deploying a sparse transducer network capable of providing reliable quantitative indication about the condition of structural areas that are critical for the load-bearing capabilities, especially those with difficult access. When the quantitative damage indicator reaches a certain threshold, advice should be issued to direct NDI in a more efficient way. This makes inspections more cost-effective than when they are conducted without having any information about the operational history of the structure, which is crucial for optimizing a maintenance program.

Acknowledgments

The authors would like to thank Marcel Elenbaas from Damen Schelde Naval Shipbuilding for his valuable collaboration during the RAMSSES project and for providing the opportunity for testing SHM and NDT technologies on the composite ship hull demonstrator. The authors are also grateful to Erik Slis and Peter van der Meer from TNO for their precious help with setting up and operating the measurement equipment during the test campaign.

Funding

The activities described in this article were part of the RAMSSES project that has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 723246. Nan Tao was supported by the European Regional Development Fund and the province of Noord-Brabant, in the context of the Operationeel Programma Zuid-Nederland (OPZuid) as part of the Dutch Composite Maintenance Center (DCMC).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The support video footage of some impacts is available through the following link: https://doi.org/10.5281/zenodo.11486467.

The acoustic emission data acquired during the impacts are available through the following link: https://doi.org/10.5281/zenodo.10925610.

References

1. S. Mansouri, H. Lee, and O. Aluko, "Multi-Objective Decision Support to Enhance Environmental Sustainability in Maritime Shipping: A Review and Future Directions," *Transportation Research Part E: Logistics and Transportation Review* 78 (2015): 3–18, https://doi.org/10.1016/j.tre.2015.01.012.

2. Z. Wan, M. Zhu, S. Chen, and D. Sperling, "Pollution: Three Steps to a Green Shipping Industry," *Nature* 530, no. 7590 (2016): 275–277, https://doi.org/10.1038/530275a.

3. S. Job, "Why Not Composites in Ships?" *Reinforced Plastics* 59, no. 2 (2015): 90–93, https://doi.org/10.1016/j.repl.2014.12.047.

4. G. Garfield, "Material Advantage: Seeking Alternatives to Reliance on Steel," *TradeWinds* (2018): https://www.tradewindsnews.com/safety/material-advantage-seeking-alternatives-to-reliance-on-steel/2-1-395361.

5. C. G. Dávila and C. Bisagni, "Fatigue Life and Damage Tolerance of Postbuckled Composite Stiffened Structures With Initial Delamination," *Composite Structures* 161 (2017): 73–84.

6. C. Socci and C. Kassapoglou, "Prediction of Matrix Cracking Initiation and Evolution and Their Effect on the Stiffness of Laminates With Off-Axis Plies Under in-Plane Loading," *Composites Science and Technology* 200 (2020): 108427, https://doi.org/10.1016/j.compscitech.2020.108427.

7. S. Gopalakrishnan, M. Ruzzene, and S. Hanagud, *Computational Techniques for Structural Health Monitoring* (Springer, 2011).

8. L. Bond and N. Meyendorf, "NDE and SHM in the Age of Industry 4.0," in *Proceedings of the 12th International Workshop on Structural Health Monitoring* (DEStech Publications, 2019).

9. CORDIS European Commission, "Realisation and Demonstration of Advanced Material Solutions for Sustainable and Efficient Ships," (2022), https://cordis.europa.eu/project/id/723246.

10. G. Gardiner, "Removing the Barriers to Lightweighting Ships With Composites," (2019), https://www.compositesworld.com/articles/removing-barriers-to-lightweighting-ships-with-composites.

11. G. Wang, K. Pran, G. Sagvolden, et al., "Ship Hull Structure Monitoring Using Fibre Optic Sensors," *Smart Materials and Structures* 10, no. 3 (2001): 472–478, https://doi.org/10.1088/0964-1726/10/3/308.

12. G. Sagvolden, K. Pran, L. Vines, H. E. Torkildsen, and G. Wang, "Fiber Optic System for Ship Hull Monitoring," in *15th Optical Fiber Sensors Conference Technical Digest* (IEEE, 2002), 435–438.

13. I. Herszberg, H. Li, F. Dharmawan, A. Mouritz, M. Nguyen, and J. Bayandor, "Damage Assessment and Monitoring of Composite Ship Joints," *Composite Structures* 67, no. 2 (2005): 205–216, https://doi.org/10.1016/j.compstruct.2004.09.017.

14. M. Mieloszyk, K. Majewska, B. Puchowski, and T. Wandowski, "Fibre Bragg Grating Sensor Application for Composite Fast Patrol Boat," *Proceedings of SPIE* 11381 (2020): 1138128.

15. J. Moughty and J. Casas, "A State-of-the-Art Review of Modal-Based Damage Detection in Bridges: Development, Challenges and Solutions," *Applied Sciences* 7, no. 5 (2017): 1–24, https://doi.org/10.3390/app7050510.

16. C. Barile, C. Casavola, G. Pappalettera, and P. Vimalathithan, "Damage Characterization in Composite Materials Using Acoustic Emission Signal-Based and Parameter-Based Data," *Composites, Part B: Engineering* 178 (2019): 107469, https://doi.org/10.1016/j.compositesb.2019.107469.

17. F. Ciampa and M. Meo, "A New Algorithm for Acoustic Emission Localization and Flexural Group Velocity Determination in Anisotropic Structures," *Composites Part A* 41, no. 12 (2010): 1777–1786, https://doi.org/10.1016/j.compositesa.2010.08.013.

18. M. Saeedifar and D. Zarouchas, "Damage Characterization of Laminated Composites Using Acoustic Emission: A Review," *Composites, Part B: Engineering* 195 (2020): 108039, https://doi.org/10.1016/j.compositesb.2020.108039.

19. S. Sikdar, P. Mirgal, S. Banerjee, and W. Ostachowicz, "Damage-Induced Acoustic Emission Source Monitoring in a Honeycomb Sandwich Composite Structure," *Composites, Part B: Engineering* 158 (2019): 179–188, https://doi.org/10.1016/j.compositesb.2018.09.071.

20. P. De Groot, P. Wijnen, and R. Janssen, "Real-Time Frequency Determination of Acoustic Emission for Different Fracture Mechanisms in Carbon/Epoxy Composites," *Composites Science and Technology* 55, no. 4 (1995): 405–412, https://doi.org/10.1016/0266-3538(95)00121-2.

21. S. Woo and T. Kim, "High-Strain-Rate Impact in Kevlar-Woven Composites and Fracture Analysis Using Acoustic Emission," *Composites, Part B: Engineering* 60 (2014): 125–136, https://doi.org/10.1016/j.compositesb.2013.12.054.

22. S. Woo and T. Kim, "High Strain-Rate Failure in Carbon/Kevlar Hybrid Woven Composites via a Novel SHPB-AE Coupled Test," *Composites, Part B: Engineering* 97 (2016): 317–328, https://doi.org/10.1016/j.compositesb.2016.04.084.

23. X. Zhuang and X. Yan, "Investigation of Damage Mechanisms in Self-Reinforced Polyethylene Composites by Acoustic Emission," *Composites Science and Technology* 66, no. 3-4 (2006): 444–449, https://doi.org/10.1016/j.compscitech.2005.07.013.

24. P. A. Ochôa, H. J. Den Ouden, M. Hagenbeek, N. Tao, A. G. Anisimov, and R. M. Groves, "Support Material for the Article Damage Categorization in Full-Scale, Full-Composite Ship Hull Under High-Energy Impacts by Unsupervised-Learning-Enabled Acoustic Emission Monitoring and Laser Shearography Inspection," *Zenodo* (2024): https://doi.org/10.5281/zenodo.11544936.

25. F. Pashmforoush, R. Khamedi, M. Fotouhi, M. Hajikhani, and M. Ahmadi, "Damage Categorization of Sandwich Composites Using Acoustic Emission Technique and K-Means Genetic Algorithm," *Journal of Nondestructive Evaluation* 33, no. 4 (2014): 481–492, https://doi.org/10.1007/s10921-014-0243-y.

26. A. Mahdian, J. Yousefi, M. Nazmdar, N. Karimi, M. Ahmadi, and G. Minak, "Damage Evaluation of Laminated Composites Under Low-Velocity Impact Tests Using Acoustic Emission Method," *Journal of Composite Materials* 51, no. 4 (2016): 479–490, https://doi.org/10.1177/0021998316648228.

27. R. Mohammadi, M. Najafabadi, H. S. Saghafi, and D. Zarouchas, "A Quantitative Assessment of the Damage Mechanisms of CFRP Laminates Interleaved by PA66 Electrospun Nanofibers Using Acoustic Emission," *Composite Structures* 258 (2021): 113395, https://doi.org/10.1016/j.compstruct.2020.113395.

28. A. Oskouei, H. Heidary, M. Ahmadi, and M. Farajpur, "Unsupervised Acoustic Emission Data Clustering for the Analysis of Damage Mechanisms in Glass/Polyester Composites," *Materials and Design* 37 (2012): 416–422.

29. M. Saeedifar, M. A. Najafabadi, D. Zarouchas, H. H. Toudeshky, and M. Jalalvand, "Clustering of Interlaminar and Intralaminar Damages in Laminated Composites Under Indentation Loading Using Acoustic Emission," *Composites, Part B: Engineering* 144 (2018): 206–219, https://doi.org/10.1016/j.compositesb.2018.02.028.

30. D. Francis, R. P. Tatam, and R. M. Groves, "Shearography Technology and Applications: A Review," *Measurement Science and Technology* 21, no. 10 (2010): 102001, https://doi.org/10.1088/0957-0233/21/10/102001.

31. W. Steinchen and L. Yang, in *Digital Shearography: Theory and Application of Digital Speckle Pattern Shearing Interferometry* (SPIE, 2003).

32. N. Tao, A. G. Anisimov, and R. M. Groves, "Towards Safe Shearography Inspection of Thick Composites With Controlled Surface Temperature Heating," *NDT and E International* 139 (2023): 102907, https://doi.org/10.1016/j.ndteint.2023.102907.

33. N. Tao, A. Anisimov, and R. M. Groves, "Shearography Non-Destructive Testing of Thick GFRP Laminates: Numerical and Experimental Study on Defect Detection With Thermal Loading," *Composite Structures* 282 (2022): 115008, https://doi.org/10.1016/j.compstruct.2021.115008.

34. A. Anisimov, M. Serikova, and R. M. Groves, "3D Shape Shearography Technique for Surface Strain Measurement of Free-Form Objects," *Applied Optics* 58, no. 3 (2019): 498–508, https://doi.org/10.1364/ao.58.000498.

35. V. Placet, E. Ramasso, L. Boubakar, and N. Zerhouni, "Online Segmentation of Acoustic Emission Data Streams for Detection of Damages in Composite Structures in Unconstrained Environments," in *11th International Conference on Structural Safety and Reliability* (New York (NY), June 2013), 1–8.

36. E. Pomponi and A. Vinogradov, "A Real-Time Approach to Acoustic Emission Clustering," *Mechanical Systems and Signal Processing* 40, no. 2 (2013): 791–804, https://doi.org/10.1016/j.ymssp.2013.03.017.

37. H. Jiang, J. Jang, and O. Nachum, "Robustness Guarantees for Density Clustering," in *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics* (Naha, Okinawa, Japan, April 2019).

38. E. N. Nasibov and G. Ulutagay, "Robustness of Density-Based Clustering Methods With Various Neighborhood Relations," *Fuzzy Sets and Systems* 160, no. 24 (2009): 3601–3615, https://doi.org/10.1016/j.fss.2009.06.012.

39. Y. Xie and S. Shekhar, "Significant DBSCAN Towards Statistically Robust Clustering," in *Proceedings of 16th International Symposium on Spatial and Temporal Databases* (Vienna, Austria, Augest 2019).

40. P. Rajpurkar, J. Irvin, K. Zhu, et al., "Chexnet: Radiologist-Level Pneumonia Detection on Chest X-Rays With Deep Learning," *Computer Vision and Pattern Recognition* 1, 2017, https://arxiv.org/abs/1711.05225.

41. AEwin Software-User's Manual, *In Review* 4 (2017).

42. P. Cawley and R. D. Adams, "The Mechanics of the Coin-Tap Method of Non-Destructive Testing," *Journal of Sound and Vibration* 122, no. 2 (1988): 299–316, https://doi.org/10.1016/s0022-460x(88)80356-0.

43. M. E. Ibrahim, "Nondestructive Evaluation of Thick-Section Composites and Sandwich Structures: A Review," *Composites Part A* 64 (2014): 36–48, https://doi.org/10.1016/j.compositesa.2014.04.010.

44. S. J. Kim, "Damage Detection in Composite Under in-Plane Load Using Tap Test," *Journal of Mechanical Science and Technology* 29, no. 1 (2015): 199–207, https://doi.org/10.1007/s12206-014-1103-5.

45. J. Gryzagoridis and D. Findeis, "Tap Testing of Composites Benchmarked With Digital Shearography," *Insight–Non-Destructive Testing and Condition Monitoring* 56, no. 1 (2014): 35–38, https://doi.org/10.1784/insi.2014.56.1.35.

46. J. Gryzagoridis and D. Findeis, "Tap Testing vs. Thermography," *E-Journal of Nondestructive Testing* 22, no. 1 (2017).