

Acoustic emission-based damage identification in CFRP

A novel machine learning method to cluster and identify damage modes in CFRP in compression-after-impact tests

Jesse del Canho¹, José Cunha¹, Luka Distelbrink¹, Mihai Fetecau¹, Jakub Korycki¹, Bo Lee¹, Dominik Stiller¹, Tomás Teixeira Pijpers¹, Roland van Dijk¹, Thomas van Kuik¹ (Group D5)

Abstract

The use of composite materials in the aerospace industry, specifically carbon fiber-reinforced polymers, is increasing rapidly. While these materials provide many benefits, their brittle nature makes them vulnerable to impact damage. Acoustic emission analysis is a promising method for non-destructive evaluation of damage growth in impacted composites. This paper investigates which damage modes can be differentiated and identified in compression-after-impact tests using acoustic emission data. Following a novel analysis pipeline, features were extracted from filtered acoustic emission data and clustered using k-means and fuzzy c-means. From the clusters, fiber fracture could be differentiated and identified with high confidence. Fiber–matrix debonding and matrix cracking, however, could not be definitively differentiated. Additionally, delamination could not be identified, likely due to its physical similarity to matrix cracking and debonding. Further research on the characteristics of matrix cracking and fiber–matrix debonding may allow these damage modes to be differentiated using acoustic emissions. This is suggested by the presence of distinct clusters, despite their characteristic features conflicting with literature. We also recommend to improve validation and identification by utilizing a labeled dataset, accompanied by an internal damage simulation to better discern similar damage modes.

Nomenclature

AE acoustic emission

BVID barely visible impact damage

CAI compression after impact

CFRP carbon fiber-reinforced polymer

NDE non-destructive evaluation

PCA principal component analysis

QI quasi-isotropic

1 Introduction

In recent decades, composite materials such as carbon fiber-reinforced polymer (CFRP) have seen increased usage in aircraft designs [1]. This is mainly due to their high strength and low weight, which helps increase aircraft efficiency. These composite airframes are subject to many impacts of different magnitudes over their lifetime. Therefore, a good understanding of the effects of impact damage on CFRP is highly valuable for modern aircraft design.

Barely visible impact damage (BVID) is considered the most dangerous impact scenario in composites [2]. BVID is hard to detect by visual inspection and it degrades the integrity of the composite aircraft structure. This type of damage can grow under fatigue loading, especially under compression [3]. Since little is known about the exact growth patterns of compression after impact (CAI) damage, a no-growth policy is currently used in aircraft design [4]. This means that the structure needs to be designed with greater safety factors than conventional metal-based designs to ensure impact damage will not grow. This comes at the cost of structural mass. If more were known about this type of damage propagation, a slow-growth [5] principle could be employed, where damage

is allowed to undergo some controlled growth before repair. This would lead to significant mass savings in the aircraft structure.

Various destructive and non-destructive evaluation (NDE) devices exist, yet these methods have shortcomings. Ultrasonic c-scans are widely used, but can only show delaminations, as other damage types are too small to be detected [6]. Micro-graphic observation allows for the identification of all damage types, but is destructive, and can only show damage in the exposed cross-section [7]. X-ray computed tomography can construct very detailed images of damage, but only in a small scan volume and requiring specialized equipment [8]. Acoustic emission (AE) has the potential to solve a number of these issues, making it an interesting alternative NDE method for future testing [9].

AE uses piezoelectric sensors to detect the elastic vibrations emitted when damage occurs in CFRP. This technique has been investigated, among others, by Saeedifar *et al.* [10], Shateri *et al.* [11], Khamedi *et al.* [12], and Groot *et al.* [13]. AE has many advantages over other NDE techniques [9], the most important being the ability to record over time, lending itself to the detection of damage propagation. Additionally, AE can be used to identify specific damage modes by clustering the data based on various signal components. This enables AE to analyze damage in real-time and rapidly locate the damage source when combining multiple sensors. Future aircraft equipped with AE sensors will benefit from monitoring of both BVID and the propagation and localization of larger damage. However, a drawback of AE compared to other NDE techniques is the large amount (>50 000 examples) of signal data

¹B.Sc. student, Faculty of Aerospace Engineering, TU Delft, Netherlands

requiring processing and clustering to enable damage mode identification.

By completing this research, we answer the following question: *Which damage modes of impact-damaged CFRP panels under static compression can be accurately differentiated and identified using AE analysis?* This question is twofold, with differentiation being the prerequisite for identification of damage modes. The answer to this question gives insight into the reliability and clustering possibilities of AE methods as a passive monitoring technique for future studies and industry applications.

2 Methods

To separate and identify the different damage modes in composites, we needed to process and analyze the AE waveforms. For this purpose, we developed a data processing pipeline, a diagram of which can be seen in Figure 1. The consecutive stages of the pipeline are described in Sections 2.3 to 2.7. Starting from the unprocessed AE signals, the first three stages of the pipeline filter the signals to reduce noise, extract the features that characterize them, and reduce them to their most important components. Based on these principal components, two clustering techniques group the signals. Each of the clusters is considered a different damage mode. In the final evaluation stage, data from tests with certain fiber orientations were used to link each cluster to a damage mode. The principal component analysis (PCA) and clustering stages require training, after which damage mode prediction can be performed on new datasets. The source code for this pipeline is publicly available on GitHub [14].

2.1 Test setup and data acquisition

The experimental data was provided by D. Biagini, a Ph.D. candidate in the Structural Integrity and Composites research group at Delft University of Technology. The experimental setup and standardized tests employed by him are described in detail in Appendix A. Each AE measurement had a duration of 1 ms and contained 2048 samples.

Three datasets with different CFRP specimens were collected:

- DS-QI: Specimen is 100 x 150 x 5.15 mm with a 32-layer quasi-isotropic [45, 0, -45, 90]_{4,s} layup. A CAI test was performed on this specimen. Dataset contains 14612 measurements.
- DS-0: Specimen is 110 x 85 x 1.8 mm with a unidirectional [0]₁₂ layup. The unimpacted specimen was compressed. Dataset contains 14324 measurements.
- DS-90: Specimen is 135 x 100 x 3.8 mm with a unidirectional [90]₂₄ layup and is shown in Figure 10. The unimpacted specimen was compressed. Dataset contains 101791 measurements.

DS-0 is loaded in the direction of the fibers, fiber fracture is expected to be prevalent. Since DS-90 is compressed perpendicular to the fiber direction, most of the load will be on the matrix itself. Therefore we expect to see mostly matrix cracking and fiber–matrix debonding. Since DS-QI is the only non-unidirectional specimen, delamination is expected to occur exclusively in this dataset, in addition to the three other damage modes.

2.2 Dataset selection

Four different combinations of datasets were used for training of the pipeline and prediction, each with varying numbers of clusters:

- Training on DS-0 + DS-90 (limited to 100 000 randomized measurements), prediction on DS-0;
- Training on DS-0 + DS-90 (limited to 100 000 randomized measurements), prediction on DS-90;
- Training on DS-0 + DS-90 (limited to 100 000 randomized measurements), prediction on DS-QI;
- Training on DS-QI (full dataset), prediction on DS-QI.

Since different damage modes are expected to occur more frequently in different datasets, combining them yields a more comprehensive training set. The size of this training set was limited due to memory constraints. Prediction was then performed on each dataset individually.

2.3 Pre-processing

Loading data from one or more datasets is the first step in the pipeline. These are shuffled to ensure a representative distribution in case only a subset of the data is used, then measurements for which the sensor was saturated are discarded. Saturation is detected by checking if the peak amplitude of a signal equals the saturation threshold.

To reduce the effects of noise, the data is then filtered (Figure 2). Following the specifications in the data sheet of the AE sensors, a bandpass filter of 100 kHz to 900 kHz is applied to the signals. Subsequent wavelet filtering [15] removes noise within the frequencies detected by the sensors. Then, signals containing multiple hits are split into individual signals, a process referred to as peak splitting (Figure 3). Details can be found in Appendix B.

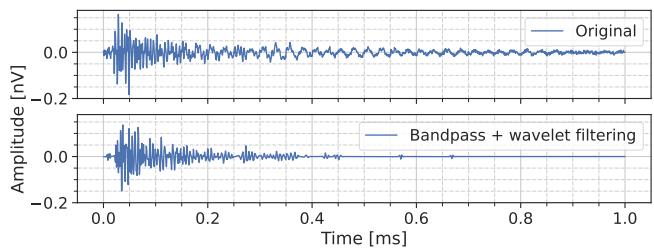


Figure 2. Signal before and after bandpass filtering and wavelet filtering. Noise is reduced especially in the tail.

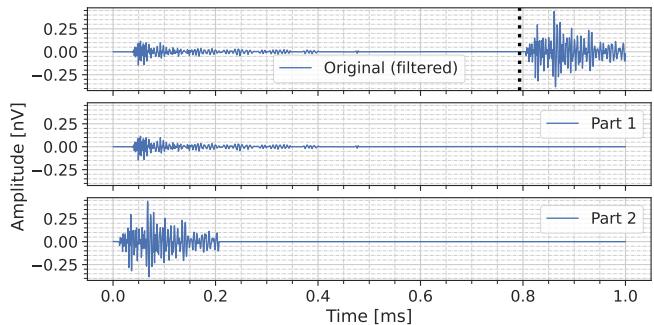


Figure 3. Signal before and after peak splitting at the black line, applied after filtering. The partial signals are padded to the original length.

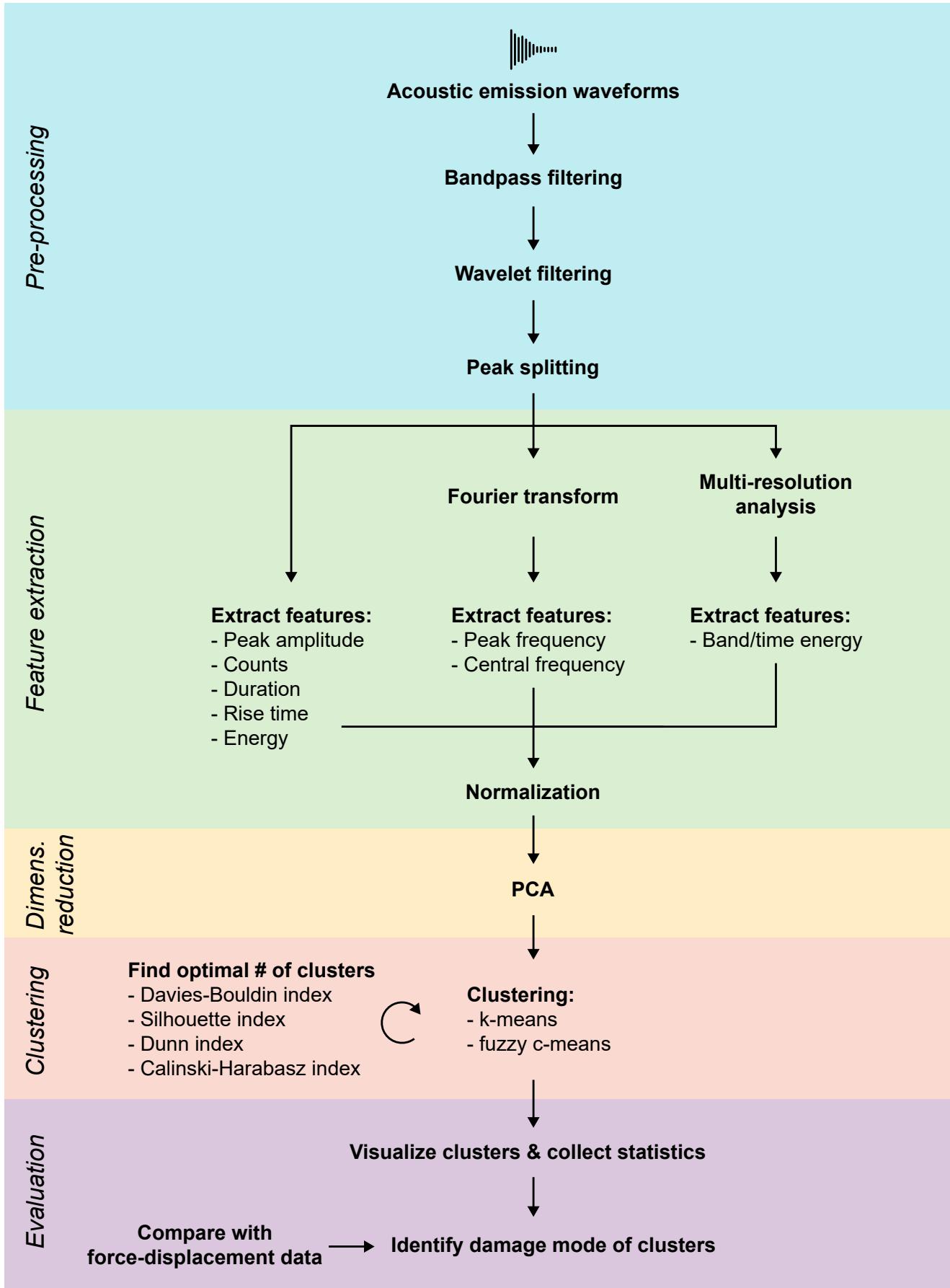


Figure 1. An overview of the data analysis pipeline. The pipeline consists of five stages, extracting features from acoustic emission waveforms to cluster them in dimensionality-reduced feature space. Cluster indexes quantify differentiation, while cluster visualization and statistics inform damage mode identification.

2.4 Feature extraction

Characteristics of the waveforms, called features, are used for clustering instead of the waveforms themselves. The three feature extractors in the pipeline lead to 23 features in total, encompassing time domain and frequency domain information. To ensure that certain features with inherently higher values do not disproportionately affect clustering, each feature was normalized to a range from -1 to 1.

The first feature extractor obtains features such as the peak amplitude directly from the signal. Counts, duration and rise time were found based on a threshold amplitude. Additionally, the energy was found from the integrated squared signal. Note that while the physical unit is J/Ω , energy is reported in attojoules ($aJ = J \times 10^{-18}$) or yoctojoules ($yJ = J \times 10^{-24}$).

Next to direct feature extraction, frequency domain information is obtained through the Fourier transform. This gives the average (or central) and peak frequencies.

Multi-resolution analysis based on the wavelet packet transform [16] is the third feature extraction method used. This technique decomposes signals in both time and frequency, as opposed to frequency-only decomposition like the Fourier transform. The energy in certain time/frequency bands is then used as feature, allowing damage modes with characteristic frequencies at specific points in time to be distinguished. Four time bands and four frequency bands are used, giving 16 features.

2.5 Dimensionality reduction

PCA is used to reduce the dimension of the 23-dimensional feature space. This is necessary since the distance metrics used in the clustering algorithms lose meaning with high dimensionality. PCA finds the orthogonal components that best capture the variance in large quantities of data [17]. This removes components that do not meaningfully divide data and would otherwise interfere with clustering. The principal components are chosen such that they collectively explain at least 95% of the variance in the dataset features.

2.6 Clustering

Clustering is then performed in the reduced-dimension feature space. Each cluster should represent a different damage mode. Two clustering methods, k-means and fuzzy c-means, were compared to determine the optimal one. While k-means assigns a measurement to a single cluster, fuzzy c-means allows measurements to belong to multiple clusters to varying degrees. Their application to AE data analysis is documented by Saeedifar and Zarouchas [18].

The number of clusters must be set a priori for both methods. Since noise is filtered, which may result in an additional cluster, we expected to find 3 clusters corresponding to the expected damage modes for DS-0 + DS-90 (matrix cracking, matrix–fiber debonding, fiber fracture). However, we trained with 3 to 5 clusters to confirm this assumption. As we additionally expected delamination in the DS-QI data, the training was performed for 3 to 6 clusters.

The centroids of each cluster are saved after training and used to predict the cluster membership of a new measurement. The performance of a certain number of clusters is evaluated using the Davies-Bouldin index, the Silhouette index, the

Dunn index, and the Calinski-Harabasz index. These indices can answer the differentiation part of the research question. We used an approximation of the indices based on 15 000 measurements due to memory limitations.

2.7 Evaluation

Once the clusters had been found, the damage mode belonging to each cluster could be identified, answering the second part of the research question. Identification was done manually based on:

- Characteristic features: the average features of a cluster, such as the mean peak frequency, can be compared to characteristic values found in the literature.
- The size of clusters: some damage modes are expected to occur more in certain datasets. For example, fiber breakage occurs more often in the DS-0 dataset, compared to the DS-90 dataset. If a cluster is very small in the latter, and much larger in the former, this may indicate that this cluster corresponds to fiber breakage.
- Cumulative energy–displacement relationship: the cumulative energy of a cluster shows how much energy is released by the corresponding damage mode. When related to the displacement, this shows when this type of damage occurred in the loading process.
- Force–displacement graphs: these graphs can be compared with the cumulative energy to see how the load-bearing capability of the specimen is affected when a certain energy release occurs.

Note that the ultimate aim was the identification of damage modes in the impacted specimen of DS-QI. The analysis of DS-0 and DS-90 only served as a means of labeling clusters and finding common characteristics.

3 Results

3.1 Force and displacement

The force–displacement relationship for all three datasets is shown in Figures 4a to 4c. Both force and displacement are shown relative to their maximum values. The DS-90 graph (Figure 4b) shows a much larger decrease in force after major failure compared to the DS-0 specimen (Figure 4a).

3.2 Clustering indices

Clustering was performed based on 9 principal components, explaining 96 % of variance for all datasets. As explained in Section 2.6, the clustering performance of the two clustering algorithms can be quantified using the four different indices that we selected. These indices have different scales:

- The Davies-Bouldin index has a minimum score of 0. Smaller numbers indicate better performance.
- The Silhouette index ranges from -1 to 1. Values close to 1 indicate that points are correctly attributed to their cluster. Values around -1 mean that a lot of points are incorrectly clustered due to too few or too many clusters.
- The Dunn index is between 0 and 1, with values closer to 1 indicating better clustering performance.
- The Calinski-Harabasz index is not bounded. High values imply dense clusters that are well separated, while low values imply the opposite.

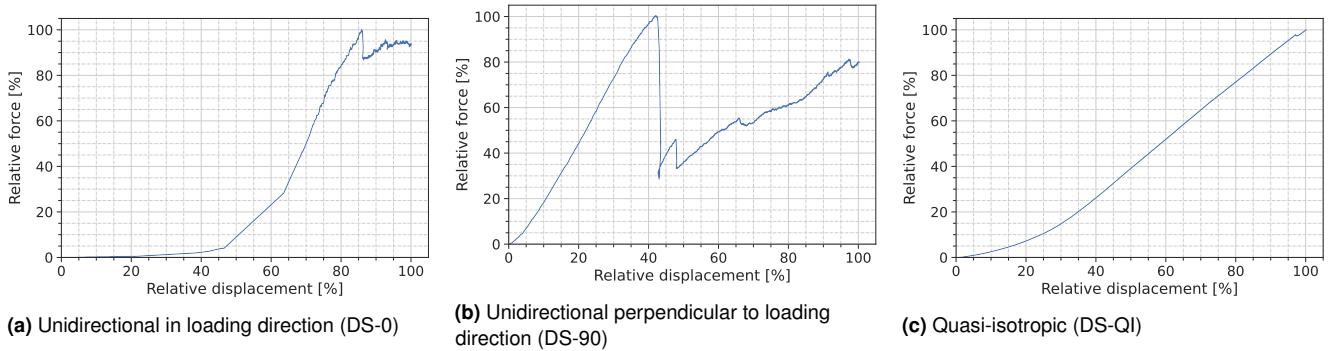


Figure 4. Force–displacement graphs for the three datasets, collected on specimens with different fiber directions.

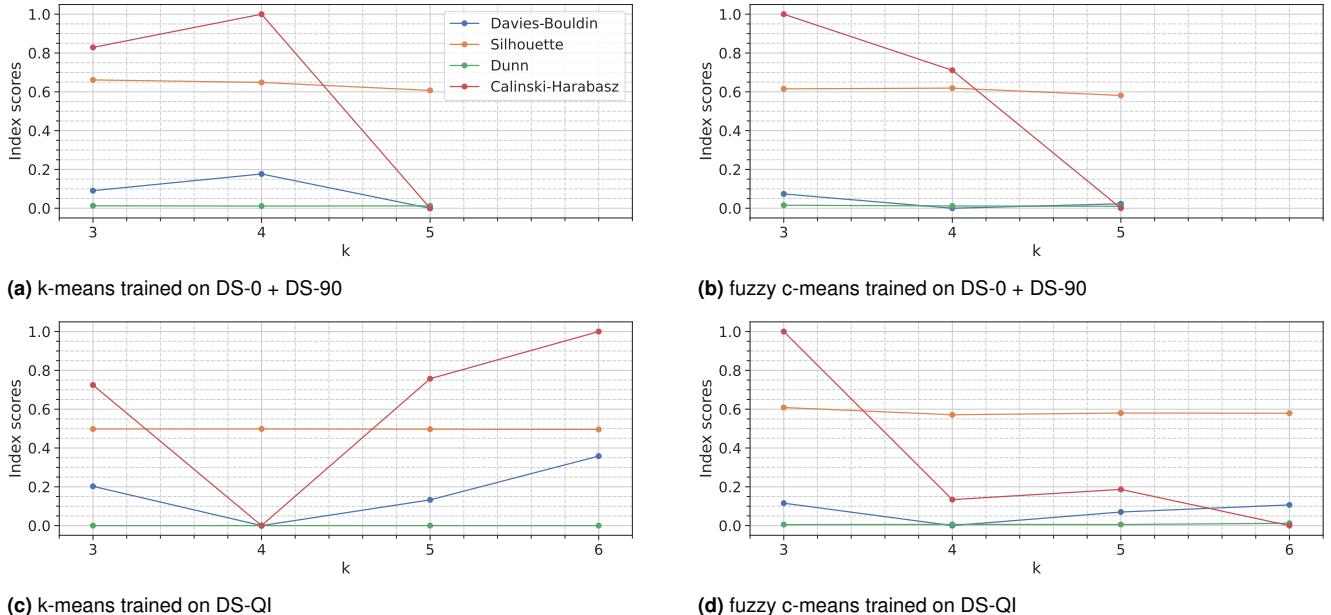


Figure 5. Normalized clustering performance indices of clustering algorithm for 3 to 6 clusters on different training datasets. High values indicate better clustering performance.

Table 1. Comparison of clustering algorithms based on performance indices for 3 clusters with a training dataset of DS-0 + DS-90. The higher normalized average for k-means suggests that this algorithm is better overall.

	Cluster algorithm	Davies-Bouldin	Silhouette	Dunn	Calinski-Harabasz	Average
Absolute	k-means	1.3132	0.3227	0.0128	6894	—
	fuzzy c-means	1.5256	0.2306	0.0155	6593	—
Normalized	k-means	1.000	0.323	0.013	1.000	0.584
	fuzzy c-means	0.861	0.231	0.016	0.956	0.516

To facilitate comparison of these indices, the values were normalized to a range of 0 to 1, with high values indicating better performance for all indices.

The graphs in Figure 5 are used to compare the performance of the different clustering algorithms across different numbers of clusters. Figures 5a and 5b show the indices after training for 3 to 5 clusters on the DS-0 and DS-90 datasets combined. Figures 5c and 5d show the indices after training for 3 to 6 clusters on the DS-QI dataset.

Table 1 presents the absolute and normalized clustering indices, applied to both clustering algorithms with 3 clusters. These final values are used to compare the overall performance of the clustering algorithm to determine the most optimal for damage mode identification.

3.3 Characteristics of clusters

Tables 2 to 4 showcase the cluster-averaged features along with the number of instances of every cluster for the two clustering algorithms. Results for 3 and 4 clusters are shown for training on DS-0 + DS-90 and DS-QI, respectively. Furthermore, the specimen from which these features were extracted can be seen in Figure 11 in Appendix A, showing the physical damage after compression.

Energy is another feature which helps with identifying the clusters as a particular damage mode. As such, the cumulative energy over relative displacement per cluster was plotted in Figures 6 to 8. From these plots, it is possible to see which clusters accumulate small amounts of energy per measurement, and which clusters accumulate energy in more sudden jumps.

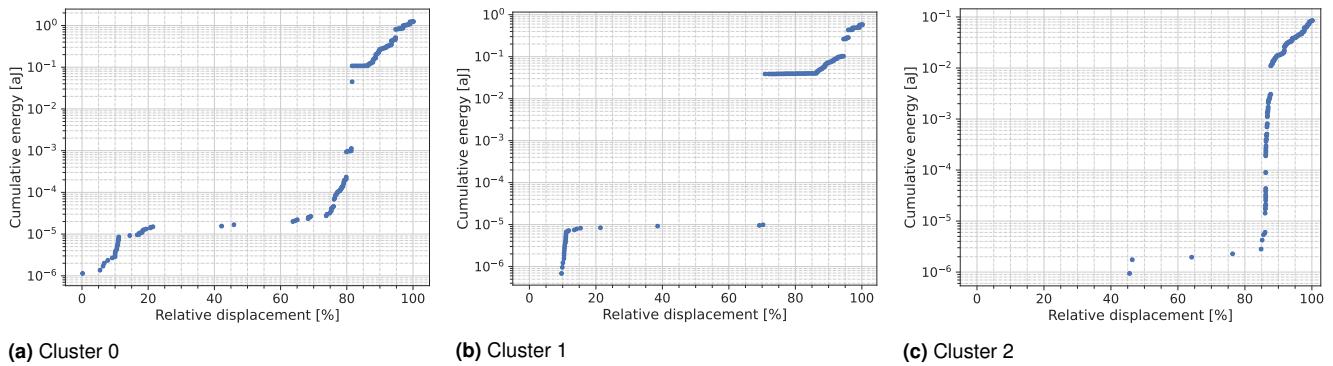


Figure 6. Cumulative energy–displacement graphs for the k-means clusters of the DS-0 dataset.

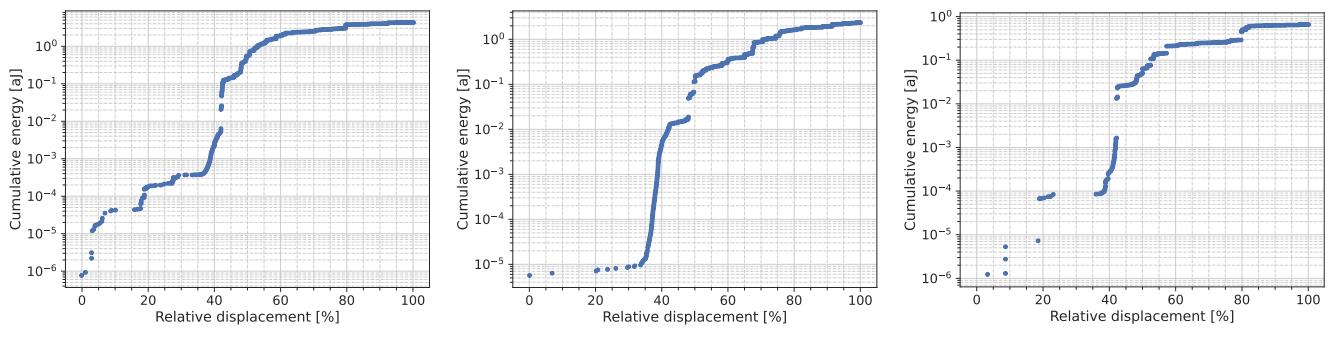


Figure 7. Cumulative energy–displacement graphs for the k-means clusters of the DS-90 dataset.

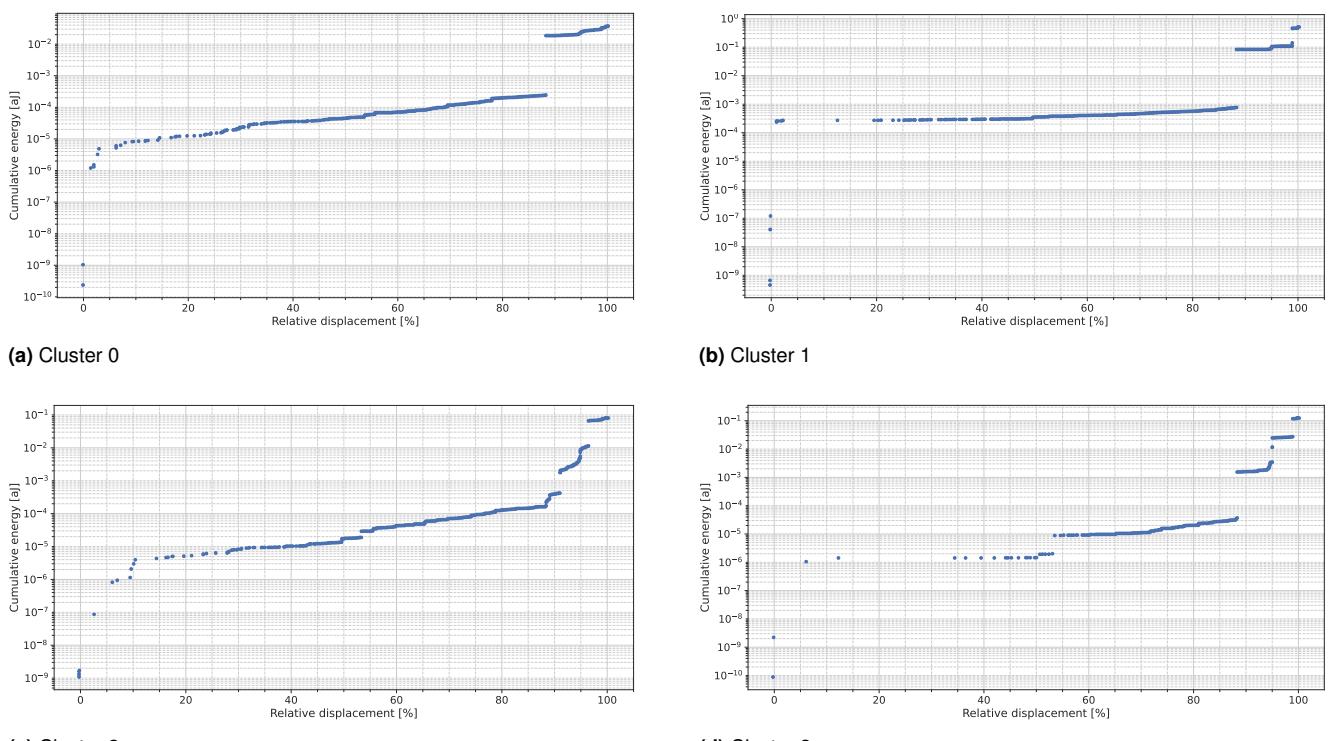


Figure 8. Cumulative energy–displacement graphs for the k-means clusters of the DS-QI dataset.

Table 2. DS-0 cluster statistics for k-means clustering with means of features.

Method	Cluster	# Instances	Peak amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	12335	691	32.1	159	35.8	89.0	148	291
	1	4165	640	30.1	109	21.4	59.0	266	353
	2	2593	641	53.6	319	69.1	214.0	143	298

Table 3. DS-90 cluster statistics for k-means clustering with means of features.

Method	Cluster	# Instances	Peak amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	84243	406	28.1	152	35.8	54.0	142	301
	1	68359	324	24.6	113	23.2	12.0	243	357
	2	20998	263	42.5	384	77.6	77.0	149	319

Table 4. DS-QI cluster statistics for k-means clustering with means of features.

Method	Cluster	# Instances	Peak amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	7598	186	19.2	116	33.4	19.0	152	351
	1	8592	163	20.9	123	35.7	55.0	150	289
	2	7612	136	17.6	110	26.7	10.0	363	405
	3	3252	156	39.4	408	132.9	12.0	164	334

4 Discussion

4.1 Cluster count and algorithms

The two clustering algorithms were scored using the indices presented in Figure 5. It was observed that the Dunn index was close to 0 for every combination of clusterer and number of clusters, which indicates poor clustering. However, Dunn is a comparatively simple metric defined as the minimum inter-cluster distance divided by the maximum intra-cluster distance [19]. The clusters from the pipeline are all close together (see Figures 12 to 14) and contain outliers, leading to a low Dunn index. Due to its sensitivity to outliers, the Dunn index is a poor clustering performance index for our data.

Figures 5a and 5b show that k-means and fuzzy c-means clustering perform best for 4 and 3 clusters, respectively, for training on DS-0 + DS-90. When trained on DS-QI, Figures 5c and 5d show that the best number of clusters are 6 for k-means and 3 for fuzzy c-means. This shows that 3 clusters seem to be optimal for the fuzzy c-means algorithm. For k-means, having 4 clusters is the best when trained on DS-0 + DS-90 but it is also the worst when trained on the DS-QI dataset, for which 6 was optimal. This is inconsistent with our expectation to find a single additional damage mode in the DS-QI dataset that is not present in DS-0 + DS-90.

Table 1 shows a higher average for the normalized indices of k-means, indicating that k-means outperforms fuzzy c-means. When using k-means with 4 clusters for DS-0 + DS-90 (optimal according to indices), two clusters presented very similar physical features. Therefore, only k-means with 3 clusters will be used for the subsequent analysis of the damage modes.

While fuzzy c-means is disregarded for now, Tables 5 to 7 in Appendix C show well-separated clusters with features that can indicate certain damage modes. This demonstrates fuzzy c-means could still be a valid method for damage mode analysis.

4.2 DS-0 test

In the compression test along the fiber direction with an unimpacted specimen, fiber fracture and matrix cracking were

expected to occur predominantly. Note that no separate cluster of noise was expected due to the filtering and peak splitting during the pre-processing. The following two clusters were identified using the cluster-averaged features shown in Table 2, based on the specimen presented in Figure 11a.

Literature suggests that the fiber fracture damage mode is associated with the highest peak frequency [18], which would mean that cluster 1 can be associated with fiber fracture. This is in accordance with the result of the energy accumulation in Figure 6b, as there is a clear energy jump present before the final maximum relative force around 70 % displacement is reached as presented in Figure 4a. Moreover, it can be seen that only a few instances are recorded during the initial displacement, while most are present in the final part before failure. This makes sense, as fiber fracture is the expected failure mode for compression along the fibers. Thus, we identified cluster 1 as fiber fracture, with an average peak frequency of 266 kHz.

Cluster 2 has a duration of 319 μs , which is more than twice as long as the duration of the other clusters in Table 2. Thus, according to literature, cluster 2 would belong to matrix cracking [18]. However, this is not in accordance with the energy accumulation plot of Figure 6c, in which the damage mode occurs most after some of the fibers have failed around 85 % displacement. This contradicts expectations since matrix cracking occurs throughout the whole test. Cluster 0 in Figure 6a would correspond more to this physical behaviour. However, the short duration of 159 μs of cluster 0 undermines it from being identified as matrix cracking. Additionally, both clusters could be identified as debonding, since both have an average amplitude and peak frequency [18]. Concluding, both cluster 0 or cluster 2 could be identified as either matrix cracking or debonding, due to the conflicts between literature and the physical behaviour.

The similarity of fiber-matrix debonding and matrix cracking are likely due to their similar and often simultaneous physical mechanism [20]. Cracks at the interface between fiber and matrix, which debonding comprises, mainly stem from cracks present in the matrix itself aligned with the fiber direction. Often, these damage modes ultimately combine.

Discerning these two damage modes is therefore complex. Further research should investigate dedicated models to more accurately determine what signals are produced for each specific mode.

4.3 DS-90 test

In the compression test perpendicular to the fiber direction with an unimpacted specimen, matrix cracking and fiber-matrix debonding were expected to be prevalent. As mentioned in Section 4.2, it is known that high peak frequency can be used to identify fiber fracture, while duration and rise time can be used to single out matrix cracking [18].

Looking at the force-displacement plot in Figure 4b, the maximum relative force is reached around 40 % relative displacement. This peak is followed by a drop in force, indicating that the specimen's load bearing capabilities greatly decreased. This can be attributed to fibers being locally misaligned due to the large displacement, resulting in high tensile stresses and ultimately fracture. Such behaviour can be correlated with high occurrence of hits around 40 % relative displacement, see Figure 7. The following region in the force-displacement plot displays a gradual increase with a lower slope than the previous region. This can be interpreted as the matrix becoming more integral to the force bearing capabilities of the specimen, given that the matrix elastic modulus is lower than that of the fiber. This suggests that the drop in relative force is due to the fibers, the main load bearing material, being highly damaged. Using this argument together with Figure 7b, it can be concluded that the cluster which presents the sharpest increase in cumulative energy around 40 % relative displacement (cluster 1) can be associated with fiber fracture. This is supported by the high peak frequency of cluster 1, which is known to correspond with fiber fracture as mentioned before.

Similarly to the DS-0 test, we can identify a cluster that has a 2 to 3 times higher rise time and duration than the other clusters. Cluster 2 can be attributed to matrix cracking based on the aforementioned features and previous research [18]. Yet, Figure 7c shows that the damage type in this cluster occurs most predominantly towards the end of the test. This is in contradiction with the physical expectation, as matrix cracking was expected to happen throughout the entire test. Additionally, given that the stronger fiber is oriented at a right angle with the force applied, we would expect that matrix cracking would happen more often closer to the beginning of the test. Indeed, this behaviour can be observed within cluster 0, as shown in Figure 7a. However, based on literature [18], this cluster could also be associated with debonding, as it has a similar medium frequency and amplitude. This uncertainty again makes the distinction and identification of debonding and matrix cracking unreliable.

4.4 DS-QI test

The quasi-isotropic specimen subject to the CAI test introduces the delamination damage mode. Unlike the previous two tests, this test was performed on an impacted specimen. The appearance of delamination stems from the fact that layers are now present in different orientations.

This dataset was additionally analyzed by training on DS-QI (see Table 8), as a damage mode that is not present during training on DS-0 + DS-90 will not reliably appear during prediction. However, in order to be able to compare to the damage identification characteristics obtained from the previous two tests (which mainly isolate fiber fracture and matrix cracking), only the data trained on DS-0 and DS-90 will be presented. Combining the damage modes from the unidirectional specimens and delamination, 4 damage modes are expected.

In this case, a distinct delamination cluster could not be identified for either training dataset. In literature, not enough characterizing elements could be found to identify delamination from the cluster statistics. The elements that were present, were often too similar to matrix cracking and debonding to be distinguishable. Moreover, no significant identifying features can be found in the associated energy plots. Therefore, only the previously discussed damage modes will be discussed for this dataset.

As the damage modes from previous tests are combined, it is possible to use the selection criteria from the two previous tests in order to better identify the damage modes of each cluster in the DS-QI test. The similarities can be found in Table 4. A high peak frequency of 363 kHz can be seen in cluster 2. This is comparable to cluster 1 of the DS-90 and DS-0 sets. Using the conclusions from the DS-0 and DS-90 tests, cluster 2 of DS-QI will contain fiber fracture with high confidence.

As before, the remaining clusters are inconclusive. Cluster 3 has a high duration and rise time, and can be linked to cluster 2 of both DS-0 and DS-90. However, as stated before, these clusters could not be identified with a certain damage mode due to conflicting information. The same holds for this cluster. The remaining two clusters (0 and 1) are also inconclusive due to the similarities of their features, but are similar to cluster 0 of both DS-0 and DS-90.

4.5 Validation

The best way to validate our findings would be comparison with a labeled dataset comprising all analyzed damage modes, with the labels obtained through other NDE methods. Currently there is no such dataset available, so validation of our results is not possible. Our only means of validation is comparison with literature, even though the interpretation of our results is partially based on that same literature. For future research, we recommend incorporating a labeled dataset into the testing procedure for validation purposes.

5 Conclusions

The purpose of this study was to determine which damage modes of CFRP panels in CAI tests could be differentiated and identified based on acoustic emission analysis. Fiber fracture could be identified with high confidence in all datasets by its high cluster-averaged peak frequency, ranging from 243 kHz to 281 kHz. Throughout all tests, matrix cracking and debonding could not be clearly identified. Clusters corresponding to the long duration associated with matrix cracking, which is expected to occur early in the test, actually accumulated energy only in the final region. Delamination was also not identifiable in the DS-QI tests.

Clustering with 4 clusters in anticipation of delamination resulted in two very similar clusters. This could be due to the fact that delamination, debonding and matrix cracking are physically very similar, and may not be distinguishable by acoustic emission.

While not every damage mode could be identified clearly, acoustic emission analysis shows good results for identifying fiber fracture. Therefore, AE currently is not a replacement for other NDE methods, but can be used complementarily. With further research of the characteristics of matrix cracking and fiber–matrix debonding using a computational model, potentially other unique characteristics can be found. With such unique characteristics, these remaining two damage modes may also be identified using AE since Tables 2 to 4 show two distinct clusters, presumably corresponding to these damage modes.

Finally, this research was constrained by hardware limitations, limiting the amount of data the model could be trained on. This also prohibited the analysis of fatigue tests. In subsequent research, more powerful hardware and more efficient code may be used to increase the amount of data for training and prediction. This would also allow for fatigue analysis, from which we expect similar results to their static test counterparts.

Acknowledgements

We would like to thank D. Biagini from the Structural Integrity and Composites (SI&C) research group at the Faculty of Aerospace Engineering at the Delft University of Technology for providing us with the experimental data used in this study, his guidance and his meaningful critiques throughout the writing of this article.

References

1. A. Ghobadi, “Common Type of Damages in Composites and Their Inspections,” *World Journal of Mechanics*, vol. 07, no. 02, pp. 24–33, Feb. 2017. DOI: [10.4236/WJM.2017.72003](https://doi.org/10.4236/WJM.2017.72003).
2. J. C. Prichard and P. J. Hogg, “The role of impact damage in post-impact compression testing,” *Composites*, vol. 21, no. 6, pp. 503–511, Nov. 1990. DOI: [10.1016/0010-4361\(90\)90423-T](https://doi.org/10.1016/0010-4361(90)90423-T).
3. M. Mitrovic, H. T. Hahn, G. P. Carman, and P. Shyprykevich, “Effect of loading parameters on the fatigue behavior of impact damaged composite laminates,” *Composites Science and Technology*, vol. 59, no. 14, pp. 2059–2078, 1999. DOI: [10.1016/s0266-3538\(99\)00061-5](https://doi.org/10.1016/s0266-3538(99)00061-5).
4. J. Tomblin and W. Seneviratne, *Determining the fatigue life of composite aircraft structures using life and load-enhancement factors*, <https://www.tc.faa.gov/its/worldpac/techrpt/ar10-6.pdf>, 2011.
5. J.-A. Pascoe, “Slow-growth damage tolerance for fatigue after impact in FRP composites: Why current research won’t get us there,” *Theoretical and Applied Fracture Mechanics*, vol. 28, no. 1, pp. 726–733, Dec. 2020. DOI: [10.1016/j.tafmec.2021.103127](https://doi.org/10.1016/j.tafmec.2021.103127).
6. S. Gholizadeh, “A review of non-destructive testing methods of composite materials,” *Procedia Structural Integrity*, vol. 1, no. 1, pp. 50–57, Jan. 2016. DOI: [10.1016/J.PROSTR.2016.02.008](https://doi.org/10.1016/J.PROSTR.2016.02.008).
7. A. R. Chambers and N. O. Heinje, “Damage characterization in CFRP using acoustic emission, X-ray tomography and FBG sensors,” in *17th International Conference on Composite Materials*, 2009, pp. 27–31.
8. D. J. Bull, S. M. Spearing, and I. Sinclair, “Observations of damage development from compression-after-impact experiments using ex situ micro-focus computed tomography,” *Composites Science and Technology*, vol. 97, no. 1, pp. 106–114, Jun. 2014. DOI: [10.1016/j.compscitech.2014.04.008](https://doi.org/10.1016/j.compscitech.2014.04.008).
9. C. U. Grossé and M. Ohtsu, *Acoustic emission testing: Basics for Research-Applications in Civil Engineering*. Springer Berlin Heidelberg, 2008, pp. 1–404. DOI: [10.1007/978-3-540-69972-9](https://doi.org/10.1007/978-3-540-69972-9).
10. M. Saeedifar, M. N. Saleh, H. M. El-Dessouky, S. Teixeira De Freitas, and D. Zarouchas, “Damage assessment of NCF, 2D and 3D woven composites under compression after multiple-impact using acoustic emission,” *Composites Part A: Applied Science and Manufacturing*, vol. 132, no. 1, p. 105 833, May 2020. DOI: [10.1016/j.compositesa.2020.105833](https://doi.org/10.1016/j.compositesa.2020.105833).
11. M. Shateri, M. Ghaib, D. Svecova, and D. Thomson, “On acoustic emission for damage detection and failure prediction in fiber reinforced polymer rods using pattern recognition analysis,” *Smart Materials and Structures*, vol. 26, no. 6, p. 065 023, May 2017. DOI: [10.1088/1361-665X/aa6e43](https://doi.org/10.1088/1361-665X/aa6e43).
12. R. Khamedi, S. Abdi, A. Ghorbani, A. Ghiami, and S. Erden, “Damage characterization of carbon/epoxy composites using acoustic emission signals wavelet analysis,” *Composite Interfaces*, vol. 27, no. 1, pp. 111–124, Jan. 2020. DOI: [10.1080/09276440.2019.1601939](https://doi.org/10.1080/09276440.2019.1601939).
13. P. J. de Groot, P. A. Wijnen, and R. B. Janssen, “Real-time frequency determination of acoustic emission for different fracture mechanisms in carbon/epoxy composites,” *Composites Science and Technology*, vol. 55, no. 4, pp. 405–415, 1995.
14. D. Stiller *et al.*, *Damage mode identification for composites*, <https://github.com/DominikStiller/tudelft-damage-identification>, 2022.
15. G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O’Leary, “Pywavelets: A python package for wavelet analysis,” *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, Apr. 2019. DOI: [10.21105/JOSS.01237](https://doi.org/10.21105/JOSS.01237).
16. S. Mallat, “Wavelet packet and local cosine bases,” in *A Wavelet Tour of Signal Processing: The Sparse Way*, Elsevier, 2009, pp. 377–395. DOI: [10.1016/B978-0-12-374370-1.X0001-8](https://doi.org/10.1016/B978-0-12-374370-1.X0001-8).
17. S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering*. Cambridge University Press, Jan. 2019, pp. 1–46. DOI: [10.1017/9781108380690](https://doi.org/10.1017/9781108380690).

18. M. Saeedifar and D. Zarouchas, "Damage characterization of laminated composites using acoustic emission: A review," *Composites Part B: Engineering*, vol. 195, no. 1, p. 108039, Aug. 2020. DOI: [10.1016/j.compositesb.2020.108039](https://doi.org/10.1016/j.compositesb.2020.108039).
19. J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 28, no. 3, pp. 301–315, 1998. DOI: [10.1109/3477.678624](https://doi.org/10.1109/3477.678624).
20. M. F. S. Ferreira De Moura, "Interaction of matrix cracking and delamination," in *Delamination Behaviour of Composites*, S. Sridharan, Ed., 1st ed., Woodhead Publishing, Oct. 2008, pp. 327–343. DOI: [10.1533/9781845694821.3.327](https://doi.org/10.1533/9781845694821.3.327).
21. Delta-tech S.p.A, *Matrix TDS-Technical Data Sheet DT120 Versatile High Toughness Epoxy Matrix*, Feb. 2015.
22. ASTM International, *Standard Test Method for Measuring the Damage Resistance of a Fiber-Reinforced Polymer Matrix Composite to a Drop-Weight Impact Event 1*, 2021. DOI: [10.1520/D7136_D7136M-20](https://doi.org/10.1520/D7136_D7136M-20).
23. ASTM International, *Standard Test Method for Compressive Residual Strength Properties of Damaged Polymer Matrix Composite Plates*, 2021. DOI: [10.1520/D7137_D7137M-17](https://doi.org/10.1520/D7137_D7137M-17).
24. Vallen Systeme GmbH, *AE-Sensor Data Sheet VS900-M*, <https://www.vallen.de/sensors/broad-band-sensors/vs900-m/>, 2019.
25. Vallen Systems GmbH, *AE-Preamplifier Data Sheet AEPH5*, https://www.vallen.de/wp-content/uploads/2019/03/Datasheet_AEP5_1602.pdf, 2019.
26. MTS Systems, *Servohydraulic System Load Cells*, <https://www.mts.com/en/products/materials/load-cells/servohydraulic-system-load-cells>, 2022.
27. M. G. R. Sause, "Investigation of Pencil-lead Breaks as Acoustic Emission Sources," *J. Acoustic Emission*, vol. 29, no. 1, pp. 184–196, 2011.
28. R. Green, "Basic Wave Analysis of Acoustic Emission," in *Mechanics of Nondestructive Testing*, Stinchcomb W. W., Duke J. C., Henneke E. G., and Reifsnider K. L., Eds., 1st ed., Springer, 1980, pp. 55–76. DOI: [10.1007/978-1-4684-3857-4{_}2](https://doi.org/10.1007/978-1-4684-3857-4{_}2).
29. S. Mallat, "Sparse representations," in *A Wavelet Tour of Signal Processing: The Sparse Way*, Elsevier, 2009, pp. 1–22. DOI: [10.1016/B978-0-12-374370-1.X0001-8](https://doi.org/10.1016/B978-0-12-374370-1.X0001-8).
30. F. Adamo, G. Andria, F. Attivissimo, A. M. L. Lanzolla, and M. Spadavecchia, "A comparative study on mother wavelet selection in ultrasound image denoising," *Measurement*, vol. 46, no. 8, pp. 2447–2456, Oct. 2013. DOI: [10.1016/J.MEASUREMENT.2013.04.064](https://doi.org/10.1016/J.MEASUREMENT.2013.04.064).

A Test setup and specimens

This appendix contains the description of the impact and compression tests performed by D. Biagini to obtain the data analyzed in this paper, introduced in Section 2.1.

A.1 Test setup

The CFRP specimens used for the tests were made of 200 g/m² unidirectional intermediate modulus carbon fiber in a matrix of DT120 thermoset resin [21].

Initially, the DS-QI specimen was damaged using the D7136/D7136M-20 standardized impact test method [22], modified with a somewhat smaller impactor mass of 4.8 kg instead of 5.5 kg. To maintain the correct impact energy according to [22], a modified drop height of 0.733 m was used. The specimen was then subjected to quasi-static compression loading at a rate of 1 mm per minute, following the D7137/D7137M-17 compression residual strength test standard for damaged composite plates [23]. The impact setup and impactor are shown in Figure 9.

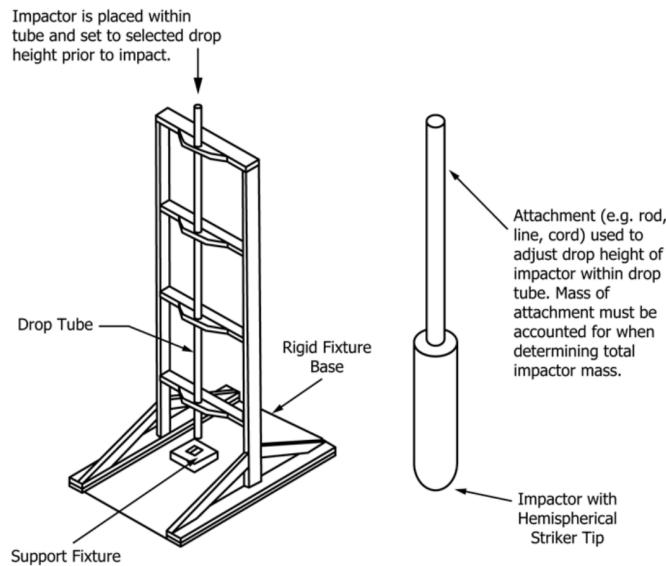


Figure 9. Diagram of the impact test setup, the impact device is shown on the left and the impactor is shown on the right [22, p. 5]

During the compression test, recordings of the AE waveforms were made using two VS900-M piezoelectric AE sensors [24], held in place by holders bonded to the specimen, indicated by the arrows in Figure 10. The signals from the AE sensors were amplified using Vallen AEP5 34 dB pre-amplifiers [25]. AE measurements were triggered when the AE signal crossed a threshold. Each AE measurement had a duration of 1 ms and contained 2048 samples. Force measurements were taken with a 250 kN MTS hydraulic load cell [26]. A standardized pencil-lead breakage test [27] was performed prior to every experiment to assess if the AE system functioned correctly.

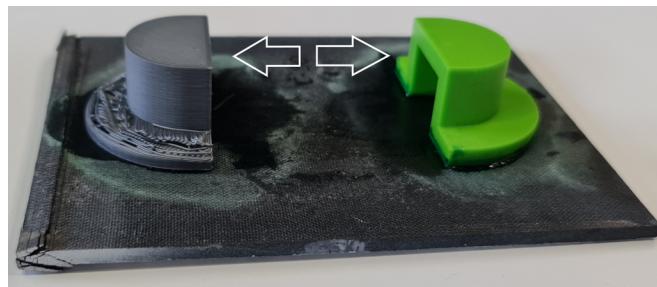
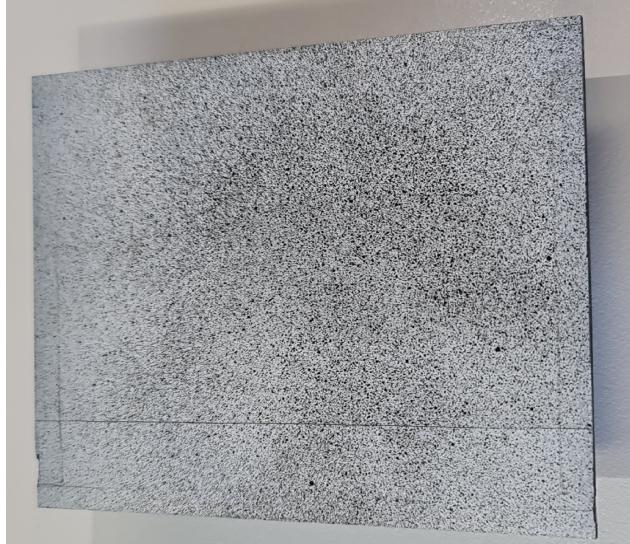


Figure 10. DS-90 specimen with sensor holders indicated by the arrows.

A.2 Specimens

Figure 11 displays all the specimens used after the tests were performed. The DS-0 specimen got the least amount of visible damage out of the three. As shown in Figure 11a, only slight matrix deforming and cracking are present predominantly along the length of the specimen. On the other hand, the DS-90 specimen shows the most damage, with the important mention that this damage occurred mostly toward the end of the test, after the buckling threshold was surpassed. Thus, as presented in Figure 11c, all three expected damage modes could be identified, namely angled matrix cracks, debonding of the matrix along the fiber interface and clear fiber fracture. The third and final specimen, displayed in Figure 11d and Figure 11b, shows clear signs of matrix cracking and a mixture of debonding and delamination between layers with different orientation.



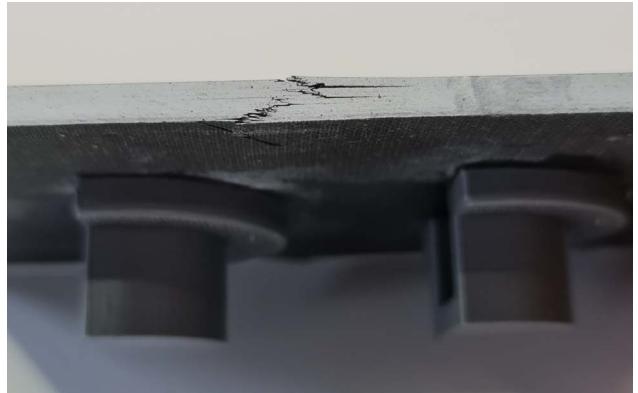
(a) DS-0 specimen top-view after compression



(b) DS-QI specimen top-view after compression



(c) DS-90 specimen side-view after compression



(d) DS-QI specimen side-view after compression

Figure 11. DS-0, DS-90 and DS-QI specimens after compression

B Pre-processing

Pre-processing is the first stage in the pipeline, preparing the examples for feature extraction. This appendix contains details on the implementation of filtering and peak splitting.

B.1 Filtering

Bandpass filtering was done using a digital Butterworth filter between 100 kHz and 900 kHz, since this is the expected range for AE information and known sensor noise is filtered out. An order of 25 was chosen, since the steepness was insufficient at lower orders.

Wavelet filtering was chosen as it allowed for the reduction of signal noise while retaining the primary characteristics of the original wave. Moreover, denoising using wavelets is especially applicable to AE signals, as they are burst-type signals [28], where energy is released in relatively short time bands and specific frequency bands when damage is recorded. This in turn leads to a sparse representation of the wavelet coefficients, where wavelet filtering is particularly powerful and efficient [29]. Coiflet and Daubechies wavelets were used because their orthogonality allowed signal energy to be preserved [30].

B.2 Peak splitting

An issue that needed to be solved in the pre-processing step was that some signals contain multiple hits. These hits can be classified as either noise or damage modes. This was considered a problem for the clustering process, as a signal with multiple damage modes or sections of noise would yield different features than one with a single isolated damage mode. Therefore, the signal was split when another peak was detected in the same signal. In order to accurately and reliably split these waveforms into multiple signals based on their contents, we devised an algorithm that uses a moving average of 160 consecutive points of the waveform. Using this moving average, the standard deviation σ was calculated. If more than five consecutive amplitudes were detected to be higher than $4 \cdot \sigma$, the waveform was split around that point with a margin, in order to get the entirety of the initial hit. The signals were then padded with zeros in order to have the same length as the original waveform.

It is important to note that after the signal is filtered, residual noise can be picked up as an individual hit. Thus clusters that solely contain noise could be generated. However, to limit the amount of pure noise signals, the standard deviations had to be greater than 10% of the peak amplitude of the absolute waveform to be counted. Additionally, incomplete hits, which are cut off due to the end of the signal, were removed by imposing a minimal duration of 20% for a valid signal.

C Additional clustering results

This appendix includes additional results from the pipeline. These include the cluster visualizations for each dataset and the clustering data from the fuzzy c-means clustering method. Note that the clustering visualizations in Figures 12 to 14 only show a small subset of the dimensions (3 instead of 9 in PCA space, 3 instead of 23 in feature space). Therefore, clusters look visually less separated than they are in reality. For PCA, these 3 most principal components account for about 75 % of the variance.

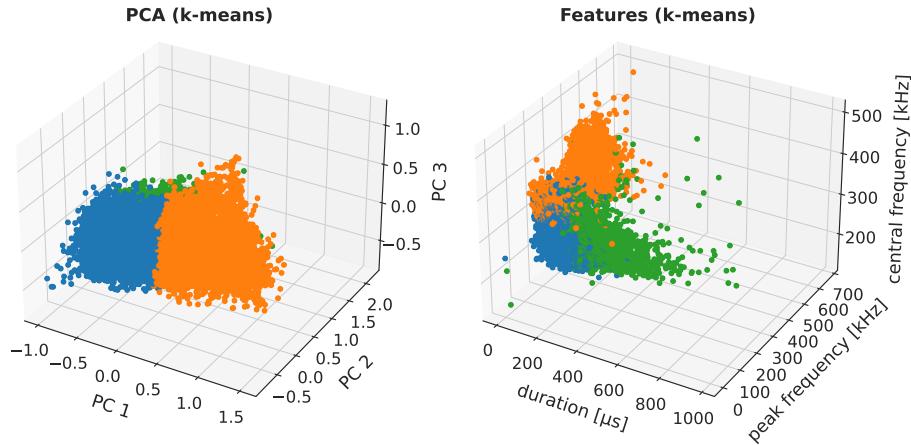


Figure 12. DS-0 clusters for k-means in the PCA (3 most principal components) and original feature space (3 out of 23 features).

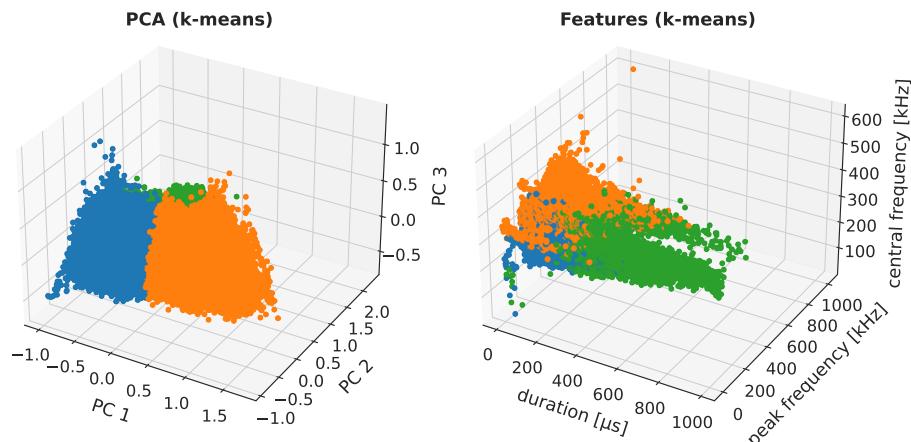


Figure 13. DS-90 clusters for k-means in the PCA (3 most principal components) and original feature space (3 out of 23 features).

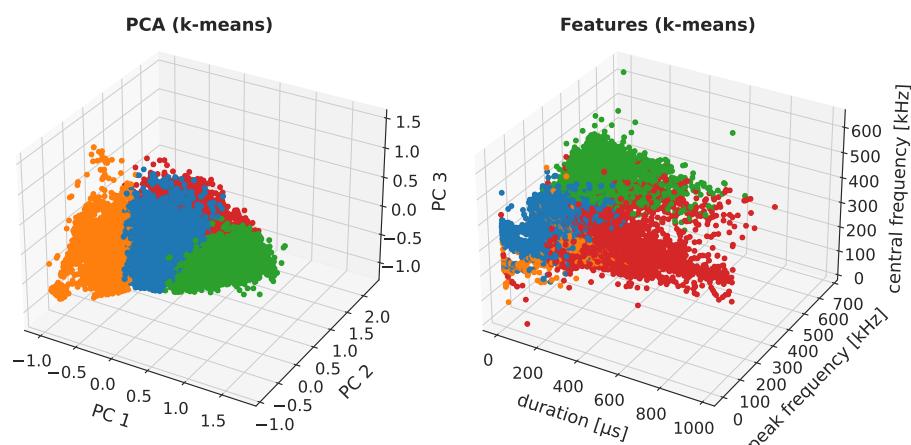


Figure 14. DS-QI clusters for k-means in the PCA (3 most principal components) and original feature space (3 out of 23 features).

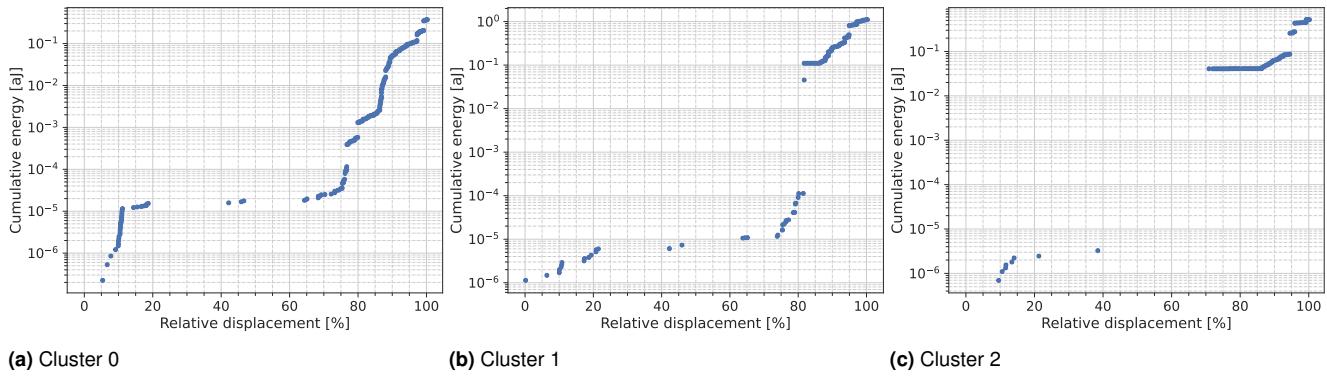


Figure 15. Cumulative energy–displacement graphs for the fuzzy c-means clusters of the DS-0 dataset.

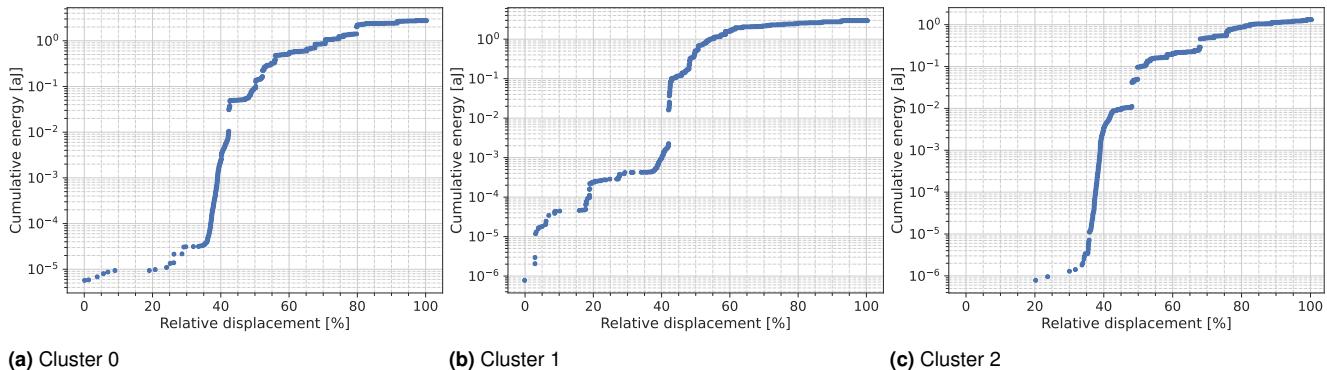


Figure 16. Cumulative energy–displacement graphs for the fuzzy c-means clusters of the DS-90 dataset.

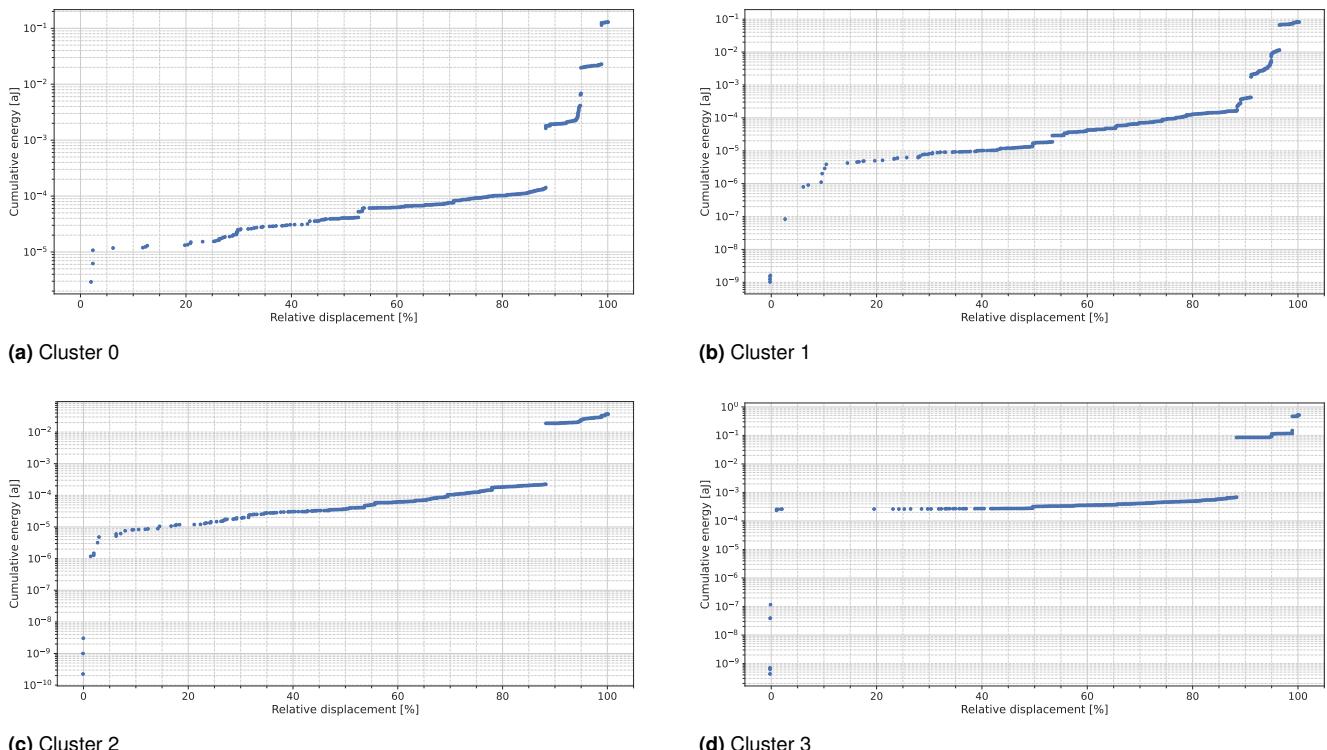


Figure 17. Cumulative energy–displacement graphs for the fuzzy c-means clusters of the DS-QI dataset.

Table 5. Clusterer statistics for DS-0 with means of features.

Algorithm	Cluster	# Members	Peak Amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	12335	691	32.1	159	35.8	89.0	148	291
	1	4165	640	30.1	109	21.4	59.0	266	353
	2	2593	641	53.6	319	69.1	214.0	143	298
fuzzy c-means	0	5299	725	32.1	151	29.7	110	153	323
	1	11037	670	36.7	194	44.7	114	146	283
	2	2757	584	31.1	97.1	21.6	24	322	364

Table 6. Clusterer statistics for DS-90 with means of features.

Algorithm	Cluster	# Members	Peak Amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	84243	406	28.1	152	35.8	54.0	142	351
	1	68359	324	24.6	113	23.2	12.0	243	357
	2	20998	263	42.5	384	77.6	77.0	149	319
fuzzy c-means	0	65432	325	28.4	174	37.9	26	148	334
	1	61108	418	31.7	197	44.3	79	140	287
	2	47060	321	24.2	111	22.1	9	287	363

Table 7. Clusterer statistics for DS-QI with means of features.

Algorithm	Cluster	# Members	Peak Amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	3391	189	41.4	405	132	101	159	332
	1	10448	159	17.3	104	30.5	14	153	333
	2	5927	177	22.9	139	39.0	27	155	294
fuzzy c-means	0	7288	138	17.9	112	27.2	11	369	408
	1	5271	114	13.8	94	30.4	2	157	311
	2	7747	198	32.2	225	66.5	63	145	297
	3	6842	180	19.7	148	43.5	23	163	357
	4	7194	137	18.1	119	30.9	11	372	409

Table 8. Clusterer statistics for DS-QI trained on DS-QI with means of features.

Algorithm	Cluster	# Members	Peak Amplitude [μV]	Counts	Duration [μs]	Rise time [μs]	Energy [yJ]	Peak frequency [kHz]	Central frequency [kHz]
k-means	0	7598	186	19.2	116	33.3	19	152	351
	1	8592	163	20.9	123	35.7	55	150	288
	2	7612	136	17.6	110	26.7	10	363	405
fuzzy c-means	0	3252	156	39.4	408	133	12	164	333
	1	4832	185	35.1	293	92.8	27	146	329
	2	7884	134	17.8	118	30.3	10	362	405
	3	6897	187	18.4	120	34.4	21	154	354
	4	7441	150	20.2	125	36.5	50	151	281

D Task distribution

This appendix contains the writing task distribution for this paper in Table 9.

Table 9. Writing task distribution, X's indicate what was worked on.

Name	Abstract	Introduction	Methods	Results	Discussion	Conclusion
J.H.W. del Canho		X	X			
D. Stiller			X	X		
L.N. Distelbrink	X	X	X			X
M. Fetecau		X		X	X	
J. Korycki		X	X	X		
B.R.M.K. Lee	X	X	X			
J.B.P.M.L. Cunha		X	X			
T.T. Pijpers			X	X	X	X
R.J. van Dijk				X	X	
T.F. van Kuik				X	X	