instruments

MDPI

Article

A Procedure for Performing Reproducibility Assessment of the Accuracy of Impact Area Classification for Structural Health Monitoring in Aerospace Structures †

Luciano Chiominto [1,*], Giulio D'Emilia [1], Antonella Gaspari [2], Emanuela Natale [1], Francesco Nicassio [3] and Gennaro Scarselli [4]

[1] Department of Industrial and Information Engineering and Economics, University of L'Aquila, 67100 L'Aquila, Italy; giulio.demilia@univaq.it (G.D.); emanuela.natale@univaq.it (E.N.)
[2] Department of Mechanics Mathematics and Management, Polytechnic of Bari, 70125 Bari, Italy; antonella.gaspari@poliba.it
[3] Department of Engineering for Innovation, University of Salento, Via per Monteroni, 73100 Lecce, Italy; francesco.nicassio@unisalento.it
[4] Department of Aeronautics and Astronautics, University of Southampton, Building 176, Boldrewood Innovation Campus, Burgess Road, Southampton SO16 7QF, UK; g.scarselli@soton.ac.uk
* Correspondence: luciano.chiominto@graduate.univaq.it

† This study is an extension of the authors' paper in Proceedings of the 2025 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0 & IoT), Castelldefels, Spain, 1–3 July 2025.

Abstract

The principal objective of this work is to develop an optimized procedure that guarantees the reproducibility of results across different applications and laboratories, facilitating potential field applications of methodologies for Structural Health Monitoring in aerospace structures. The focus is to accurately detect and localize impact areas on planar structures using in situ transducers and Machine Learning (ML) techniques. The research concentrates on an aluminum plate where impacts are generated by metal spheres of different masses dropped from a fixed height. The resulting Lamb waves are detected by PZT sensors glued on the surface. Various data processing and feature extraction algorithms are implemented and compared to extract the differences in Time of Flight (ΔToF). The obtained features are used for training ML classification models. Then, the influence of various parameters in signal acquisition and data processing are assessed along with the reproducibility of the results. For this reason, an interlaboratory comparison is conducted in which the trained models are applied to data collected under varying conditions. The experimental results show that the most influencing factors for impact area classification are the algorithm for ΔToF estimation, the number of training points used in ML models, the type of classification model, the distribution of the impact points on the component, and their balance in the classification area. This evidence suggests approaches for reducing both issues, therefore improving the reproducibility of results.

Keywords: measurement; validation; SHM; Lamb waves; piezoelectric sensors; machine learning; diagnostic; prognostic; classification; impact localization

Academic Editor: Mauro Serpelloni

Received: 5 November 2025
Revised: 20 January 2026
Accepted: 22 January 2026
Published: 26 January 2026

1. Introduction

In the aerospace industry, ensuring structural integrity and early detection of defects is of paramount importance for several reasons. For example, structure monitoring improves safety, reduces unnecessary inspection and maintenance costs, and minimizes the weight of

https://doi.org/10.3390/instruments10010006

aircraft components during the design phase. To achieve these aims, non-destructive testing (NDT) techniques are commonly employed for structural assessments, often alongside theoretical models and simulations [1–6].

In recent decades, Structural Health Monitoring (SHM) has gained a lot of attention as its significance has become increasingly apparent. SHM involves a range of techniques that use sensor networks and sophisticated algorithms to provide real-time evaluations of structural reliability [7]. An SHM system is designed to detect structural damage in order to formulate a potential intervention plan. The damage identification process can be delineated into three primary stages: detection, diagnosis, and prognosis [8].

The first SHM phase is damage detection. This includes changes to the material and/or geometric properties of these structures which adversely affect the system's performance [9].

Diagnosis is an ongoing process throughout the structure's lifecycle. This stage aims to monitor its operational condition and identify any damage resulting from various factors such as material stress, collisions, or aging. This phase is further divided into passive and active diagnosis [10]. Passive diagnosis relies on sensors that capture any changes in the structure during normal use. In contrast, active diagnosis uses an actuator to apply stress, generating a response that is measured by receiving sensors, thus enabling assessment of the structure's condition.

The prognosis phase focuses on assessing the severity of any identified damage and estimating the remaining lifespan of the structure.

One of the key concepts of SHM is online monitoring, allowing all of the previously described phases to be carried out in embedded systems. These "smart" structures overcome several problems of traditional inspection methods. In fact, manual inspections, NDT, and model-based techniques are often labor-intensive, time-consuming, and may be inadequate for detecting hidden or developing damage [11]. This approach has several potential benefits:

- Facilitates condition-based maintenance.
- Prevents catastrophic damage.
- Reduces machine downtime.
- Eliminates human error.
- Monitors inaccessible components.

In the aerospace sector, the integrity and longevity of structures are significantly influenced by the frequency and intensity of impacts from orbiting objects, space debris, or, for aircraft, bird strikes [12].

One subset of SHM methods used for examining the impacts of external objects on a thin planar plate uses a passive diagnosis system in conjunction with the analysis of Lamb waves. Lamb waves represent an acoustic emission (AE) phenomenon and are a combination of longitudinal and transverse modes. The propagation properties of these waves depend on the excitation source and the structural geometry. Another important aspect of Lamb waves is their dispersive behavior. This means that propagation speed is frequency-dependent. A significant advantage of these waves is their ability to travel over long distances, thereby eliminating the need to access the structure using local detection methods [1,13].

To identify the impact damage, several studies have used Machine Learning (ML) techniques applied on AE signals. In [14], the authors developed an SHM system that identifies the source of an AE, allowing them to study the A0 mode of Lamb waves in a composite structure using PZT ($\mathrm{Pb[Zr_xTi_{1-x}]O_3}$) sensors. The signals are analyzed by means of Continuous Wavelet Transform (CWT) and then the impact coordinates are determined by solving a set of non-linear equations using a combination of a local Newton iterative method and global unconstrained optimization. Other impact point localization algorithms are presented in the literature [15,16]. These methods do not use ML techniques and focus on the use of analytical techniques like FEM models combined with specific experimental setups. In all of these cases, the authors use closed-form mathematical solutions. Other approaches used for localizing impacts take advantage of ML techniques, as in [17]. Here, the authors used pre-trained stacked autoencoders in a two-step approach that first localizes AE sources and then characterizes them. Another method is presented in [18] and is based on the time reversal technique and ML approaches. This algorithm consists of two steps: The first involves a training process, where a baseline of structural responses from impact tests is computed. The second step assesses the impact location by utilizing the highest cross-correlation coefficient, derived from interpolating the impact response baseline. Other approaches are based on different models, as presented in [19]. In that study, automatic impact detection and localization use Random Forest and Stacked Autoencoder (SAE) techniques. A steel ball is dropped on a composite aircraft elevator to collect AE signals. The Random Forest algorithm is trained on 3600 AE samples with 15 features to identify impact zones, while the SAE utilizes both raw AE signals and their fast Fourier Transform (FFT) for impact localization. All of these studies demonstrate the potential of applying signal processing and ML techniques in SHM based on AEs with reference to specific cases and applications. Nevertheless, an important requirement to be discussed concerns the reproducibility of each approach, as experimental methods must be able to achieve accurate results even when used in different situations. Accurate bias and uncertainty analysis mitigates the effect of systematic and random errors. It is essential to quantitively estimate those effects to support interlaboratory comparisons, which are useful for assessing the effects of the such improvements [20]. This is more relevant when ML methods are used.

Based on the above considerations, this paper aims to assess the effects of many aspects of the procedure, from the measurement phase to data processing, on the final reproducibility of the data, including a comparison of results with measurements carried out in different laboratories. This paper is an extension of a preliminary work presented by the authors at the IEEE MetroInd 2025 conference [21]. The present study aims to evaluate the accuracy of different methods that use ML algorithms for identifying the area of impact of metal spheres of different masses. These are dropped on the surface of an aluminum plate in different locations, therefore generating elastic waves that propagate through the component and are measured by Piezo Wafer Active sensors (PWASs). The feature of interest for impact localization is the difference between the Times of Flight (ΔToFs) of two vibration signals. The content presented in [21] is expanded here, as we assess the influence of different parameters on the variability of the results and the possible systematic effects that influence the localization capabilities and variability of the trained models. In particular, the aspects examined are the number of impact points of each area of interest, its layout, and the strategy for generating and analyzing the impacts. With reference to this, the parameters evaluated are the type and material of spheres, the number of repeated impacts on the same point used for building the training dataset for the ML models, the algorithm used for ΔToF evaluation, and the frequency of acquisition of the acoustic signal. Moreover, different classification algorithms are compared for localizing the impacts areas. To assess the reproducibility of the methods and results, an interlaboratory comparison is performed where the trained models are used on data acquired in different test and laboratory conditions.

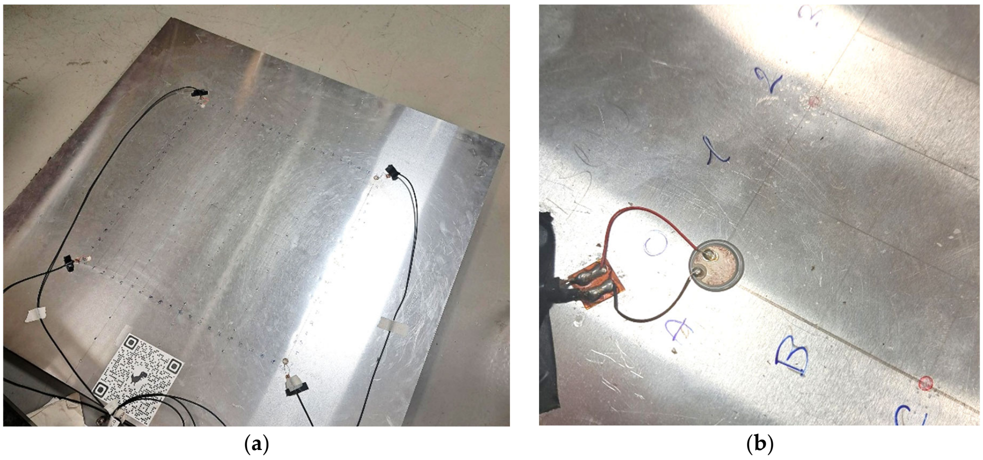
In Section 2, the test bench and the data acquisition and processing techniques are described, highlighting the parameters that influence the variability of results, which are theoretically and experimentally analyzed. In Section 3, the results of the parametric analysis are described and discussed to evaluate the most important parameters to be kept under control. Finally, an interlaboratory comparison is presented to assess the reproducibility of the experimental information.

The final section of this paper presents the conclusions.

2. Materials and Methods

2.1. Experimental Setup and Measurement Chain

The experimental setup centers on an aluminum alloy plate, which is a component of a complex structure, such as fuselage skin or wing surfaces, in a real-world context. The experimental setup used in this work is presented in Figure 1.



(**a**) (**b**)

Figure 1. Aluminum plate (**a**) under analysis and PZT sensor used to sense impacts (**b**).

As shown in Figure 1a, the aluminum alloy plate utilized in the experiment is representative of a portion of a panel used in the aerospace sector. It is 1000 mm in both length and width, with a thickness of 1.5 mm. It has an elastic modulus of 72 GPa, a Poisson's ratio of 0.33, and a density of 2700 $\mathrm{kg/m^3}$. This plate is positioned on a foam rubber mattress that dampens its movements. This setup avoids unwanted vibrations and facilitates the detection of acoustic waves that propagate solely within the plate's thickness as Lamb waves.

As shown in Figure 1b, for detecting the impacts, four PZT ($\mathrm{Pb[Zr_xTi_{1-x}]O_3}$) piezoelectric ceramic sensors are secured to the plate surface using bicomponent epoxy glue. The sensor model is PIC255 by Physik Instrumente (Physik Instrumente, Karlsruhe, Germany) [22], characterized by a diameter of 10 mm. The PZT sensors are positioned at the corners of a square with a side length of 500 mm. The sensors are placed 240 mm from the edges of the plate. This placement helps to ensure a cleaner initial impact signal. The reverberations due to the edges are delayed since the impact wave travels 240 mm + 240 mm before being registered by the sensor.

As depicted in Figure 2, a regular grid of 21 × 21 points defines the impact locations on the aluminum plate in the experimental tests, where the four positions (positions 1, 2, 3, and 4) at the vertex of the squared grid are occupied by the sensors. In total, $441 - 4$ (sensor positions) = 437 impact points are realized on the component surface, at 25.0 mm from each other.

For generating the impacts, two spheres, namely S1 and S2, with different masses, are dropped on the plate surface. S1 is an 8.3 g stainless-steel sphere with a diameter of 12.6 mm, while S2 is a 3.6 g lead sphere with a diameter of 8.5 mm. As shown in Figure 3b, these are dropped from a fixed height of 330 mm using a tube for guiding the fall of the spheres at each impact point.

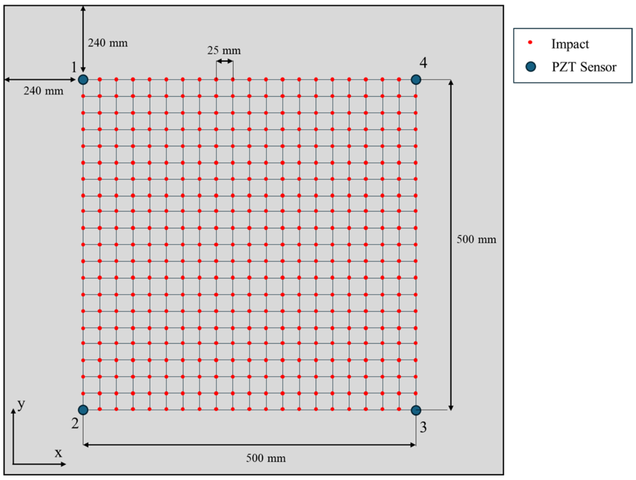


Figure 2. A scheme of the impact locations defined on the aluminum plate and placement of the PZT sensors.

(**a**) (**b**)

Figure 3. The spheres used for generating the impact on the plate surface (**a**) and a detailed image of the process (**b**).

The data from the four PZT sensors are acquired using a Siglent (Siglent, Shenzhen China) SDS824X HD 4-channel oscilloscope, connected to a PC using an Ethernet cable to download the data. The sampling frequency of the signals is 5 MHz, and the resolution is 12 bit. The vertical scale of the oscilloscope is set to 10.0 V/div, allowing signals in range ± 40.0 V. To start the sampling, the acquisition channels of the oscilloscope are triggered with a threshold equal to 1.0 V. The total recording length is set to 200 ms including a pre-trigger of 20 ms. In this manner, in each signal, the whole impact phenomenon is acquired. It represents the impact event itself and the subsequent vibrations on the plate until it returns to a steady state. Even though the above-mentioned reverberations are not the object of this study, some considerations are added in the following paragraphs.

The test campaign includes 437 impacts repeated two times using sphere S1 and 437 impacts repeated two times using sphere S2. In total, 1748 impact signals are acquired from each PZT sensor.

2.2. Data Processing

In thin aluminum plates, only the zero-modes of Lamb waves, specifically the symmetric (S0) and antisymmetric (A0) mode, can be activated, and consequently, the acoustic emission frequency spectrum is predominantly concentrated below 100 kHz [21].

Previous research by the authors revealed that the signal's arrival time, or Time of Flight (ToF), is primarily attributed to the A0 antisymmetric mode [10]. This is because the symmetric mode S0 exhibits negligible amplitudes within this frequency range. For this reason, the following analysis considers the Lamb waves in A0 mode at the specific frequency of 40 kHz.

Since the exact time instant of impact cannot be determined, the ToF is assessed from the start of acquisition. To determine the impact location, the key parameter is the difference in ToFs (ΔToF) recorded by the PZT sensors. The ΔToFs are estimated using Cross Correlation (CC), Continuous Wavelet Transform (CWT), and Short-Time Fourier Transform (STFT). All of the data processing algorithms are implemented in MATLAB 2023b.

Figure 4 shows the impact signals measured by two sensors, demonstrating the ΔToF concept when sensor 1 (blue) and sensor 3 (orange) are considered. In this case, for instance, the two vertical lines indicate the temporal instance when the signals start changing from a steady state. It can also be noted that a pre-trigger condition is set on the oscilloscope, as already stated in Section 2.1. Measuring the difference in Time of Flight at the very beginning of the signal was found to be the most repeatable way; setting a threshold caused higher variability, probably since the starting slope at the beginning is not uniform depending on position.

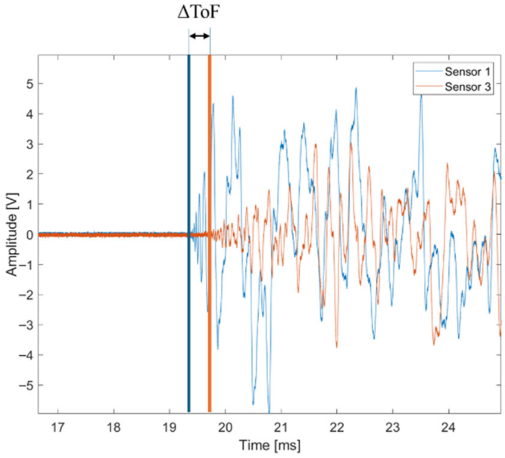


Figure 4. Example of ΔToF of impact measured by two sensors.

2.2.1. Cross Correlation

The oscilloscope signal is filtered using a bandpass Butterworth's filter of order 4. The selected cutoff frequencies are 35 kHz and 45 kHz, enabling us to analyze only the 40 kHz signal component. This is the frequency at which a 1.5 mm thick aluminum plate has the maximum A0 mode amplitude [10]. Then, using the MATLAB built-in function findchangepts() [23], the data index where the signal changes most significantly is identified. The option used in this function is "std", detecting changes in the standard deviation. The first instant of arrival of the wave at the four sensors is identified, and starting from this time point, considering the duration of the phenomenon, a time interval of 0.6 ms for all four acquired signals is selected for analysis. Considering an impact close to a sensor, 0.6 ms is the minimum time that it takes for the impact wave to travel from that point to the most distant sensor without introducing many reverberations in the first signal. The maximum distance found on the plate is 700 mm and the corresponding wave velocity is about 1135 m/s [10].

Successive waves are due to reflections and reverberations of the impact inside the plate. However, the selected signal length reduces unwanted vibrations. These signals are compared in pairs by calculating the CC coefficient [24]. The Cross Correlation coefficient peak corresponds to the associated delay between the two signals, which is the ΔToF. This procedure is repeated for all six sensor pairs.

The flowchart of this procedure is represented in Figure 5.

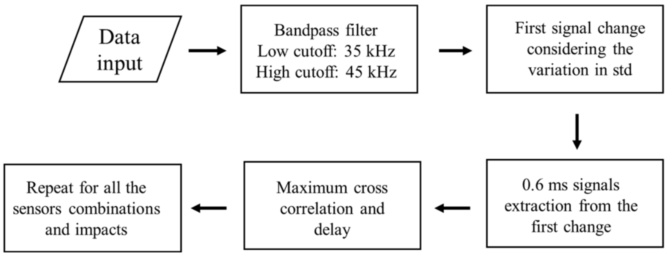


Figure 5. Flowchart of ΔToF estimation using CC.

2.2.2. Continuous Wavelet Transform

The impact signal for each sensor is analyzed using CWT to extract both time and frequency information. CWT uses a mother wavelet that is scaled and shifted to provide different window sizes for different frequencies. In this case, the CWT of the signal is computed in frequencies starting from 20 kHz to 60 kHz using Morse wavelets with a symmetry parameter equal to 3 and a time-bandwidth product equal to 60. This type of wavelet is considered useful in localizing discontinuities in signals [25]. To better represent the signal, the wavelet is computed using 16 voices per octave. The amplitude of the scale closer to the target frequency is considered, and its squared modulus is calculated. The impact instant is the time that corresponds to the first significant peak in the scalogram. The ΔToFs are the differences between the time instants identified for each sensor. All of these steps are summarized in Figure 6.

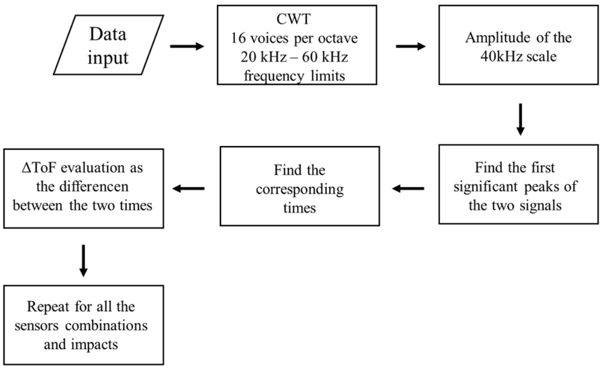


Figure 6. Flowchart of ΔToF estimation using CWT.

2.2.3. Short-Time Fourier Transform

Calculating STFT, the signal is analyzed both in time and frequency domains. In this case, the window length is 0.1 ms and the corresponding frequency resolution is equal to 10 kHz. In STFT, the window length is fixed for all frequencies. Selection of the window length must balance the necessity of frequency and time resolution. To reduce leakage, hamming windows are used with 90% overlapping. The ToF is calculated considering the amplitude of the signal corresponding to 40 kHz. Then, the local peaks of the energy at this frequency are identified using the built-in Matlab function findpeaks() [26]. This function identifies local maxima in a signal vector depending on the option ‘MinPeakHeight’, which is employed for refining the findings of the peaks. In this case, this option is activated to reduce the effect of initial noise on the measured signal. The first notable peak in the signal is recognized as the impact instant. Therefore, its time of occurrence indicates the ToF for each sensor, and the ΔToFs are determined considering all sensor pairs for each impact point.

The variability in each ΔToF estimation is calculated with reference to each sensor pair considering four repeated impacts and assuming a uniform distribution [27].

The calculated ΔToFs considering all points and all repetitions represent the predictors for localizing impacts using Machine Learning approaches.

Figure 7 illustrates ΔToF estimation using the STFT methodology.

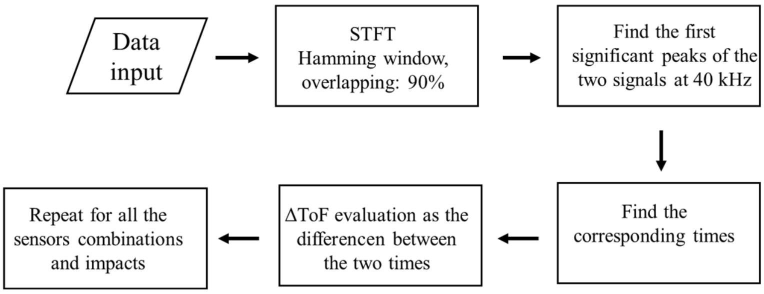


Figure 7. Flowchart of ΔToF estimation using STFT.

2.3. Machine Learning Models

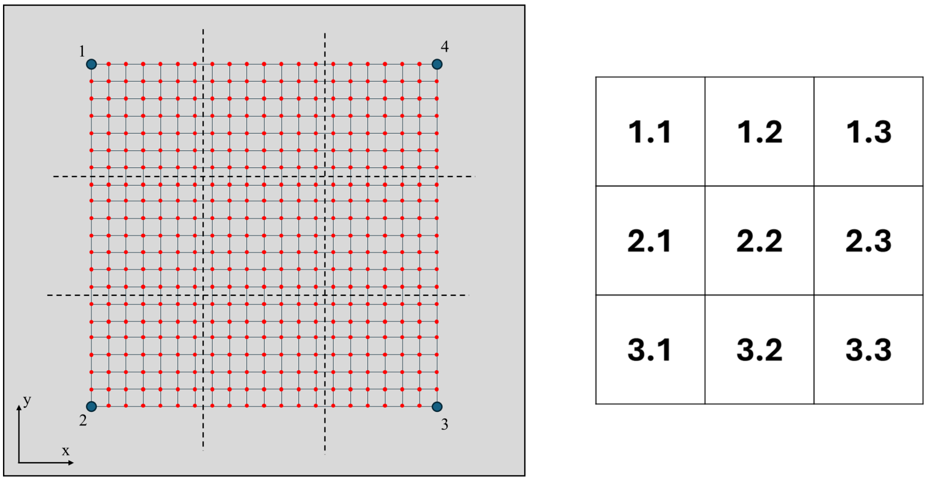
This work focuses on using classification algorithms applied to the impact ΔToFs for identifying the regions of the plate where impacts occurred.

Since there are many possible classification models for this task and it is not possible to identify the best one a priori, multiple models are trained and compared in terms of classification accuracy.

2.3.1. Definition of Classification Zones

The definition of the plate zones is based on geometrical considerations to avoid biases in classification using ML algorithms. In this case, the impact point distribution is uniform on the surface; so, in each sub-area, there is the same number of points. This aspect is important for generating a training dataset that does not present class imbalances.

The aluminum plate surface is subdivided considering a matrix of 3 × 3 elements, resulting in 9 impact zones. Figure 8 shows the subdivision of components.



(**a**) (**b**)

Figure 8. Impact zones (**a**) and naming for each defined area (**b**). The red dots represent the impact points on the plate while the blue ones are the sensors. The black dot lines are the ideal subdivision of the plate surface.

The impact points are in a square measuring 150 × 150 mm. Using this subdivision, the number of impact points is balanced for each zone and equal to 49, or 48 for areas positioned at the edges, where a position is occupied by the PZT sensor. A representation of the zone nomenclature is reported in Figure 8b.

2.3.2. Effect of Impact Points Configuration

The points near the borders are the most critical. To evaluate their effect, a different point configuration is considered that excludes them. Each zone is redefined as shown in Figure 9.

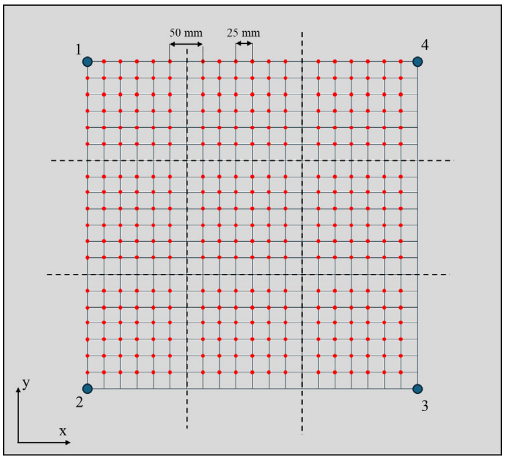


Figure 9. A scheme of the aluminum plate used for evaluating the effect of impact point layout. The red dots indicate the impact points on the plate, while the blue dots represent the sensors. The black dashed lines illustrate the subdivision of the plate's surface.

In this configuration, points that are located less than 25 mm from the border are not considered. Compared to the previous case, the dataset is 70% of the original one.

2.3.3. Effect of Number of Points

The effect of the amount of data used for training on classification performance is assessed by considering half impact points. Figure 10 represents the distribution of the impact points used in this analysis.

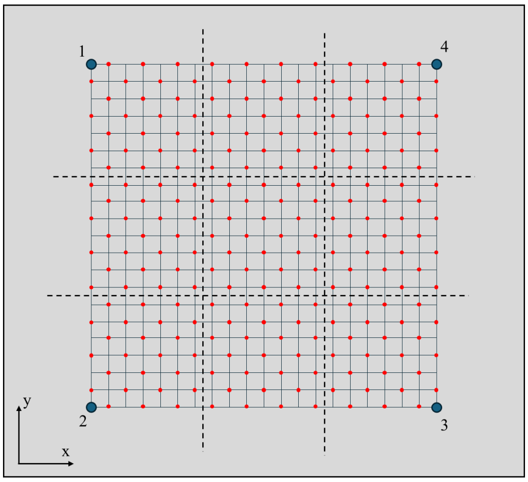


Figure 10. A representation of the impact point distribution for evaluating the effect of the amount of data used for training. The sensors are blue dots while the impacts are represented by the red dots. The dashed lines are the borders of each zone.

With the impact points distributed as shown in Figure 10, along with data reduction, the minimum distance between two points is equal to ~35.4 mm, i.e., the diagonal of the square with dimensions of 25.0 mm.

2.3.4. Effect of Signal Sampling Frequency

Signal sampling frequency is important for several aspects of this analysis. First, it defines the maximum frequency that could be analyzed, according to the Nyquist theorem. Secondly, it influences the number of data points recorded and the further computational cost for ΔToF estimation. To evaluate the effect of the sampling frequency, the original 5.00 MHz signal is resampled to 1.00 MHz and 500 kHz. To perform signal resampling, the MATLAB resample() [28] function is used, allowing the input signal to be resampled with a pre-defined ratio between positive integers. This means that not only the input signal but also the constants 'p' and 'q' are provided; so, the function resamples the input sequence at p/q times the original sample rate.

2.3.5. Reproducibility of Results

To verify the robustness of the classification methods developed in this work with the aim of assessing the reproducibility of the results obtained, the impact signals acquired using a different experimental setup are analyzed. In particular, the same aluminum plate along with the PZT sensors is used for recording 167 impact signals. In this case, the data are sampled at 3.96 MHz using a different instrument [10] and the impacts are generated with bodies of different masses. To maintain the vibrational behavior of the plate, it is put on the same foam material.

The provided impact signals are then analyzed using the three developed methods and the corresponding ΔToFs are classified using the trained models.

Figure 11 shows a reference scheme of the general flowchart described in this paper. Regarding the ΔToF estimation methods, these are better described in Appendix A, where the pseudocodes are provided for each technique.

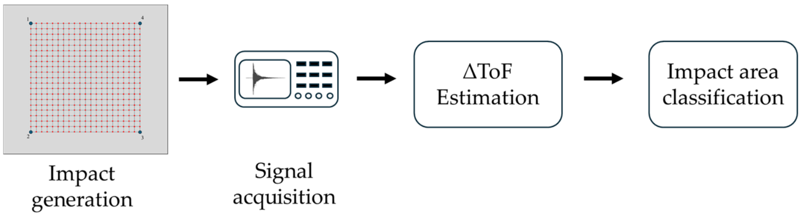


Figure 11. Scheme of processing chain for impact classification. The impacts depicted as red dots are generated on the aluminum plate. The signals are acquired using an oscilloscope and then are analyzed for estimating the ΔToFs and for classifying the impact area.

3. Results

Using the experimental setup outlined in the previous section, an example of the temporal signals acquired during a single impact is presented in Figure 12.

In this section, the results of the ΔToF estimation methods and the training of classification models are presented. As stated in Section 2, the influencing factors analyzed are as follows:

- ΔToF estimation method.
- Classification model.
- Layout of impact points.
- Sampling frequency of signals.

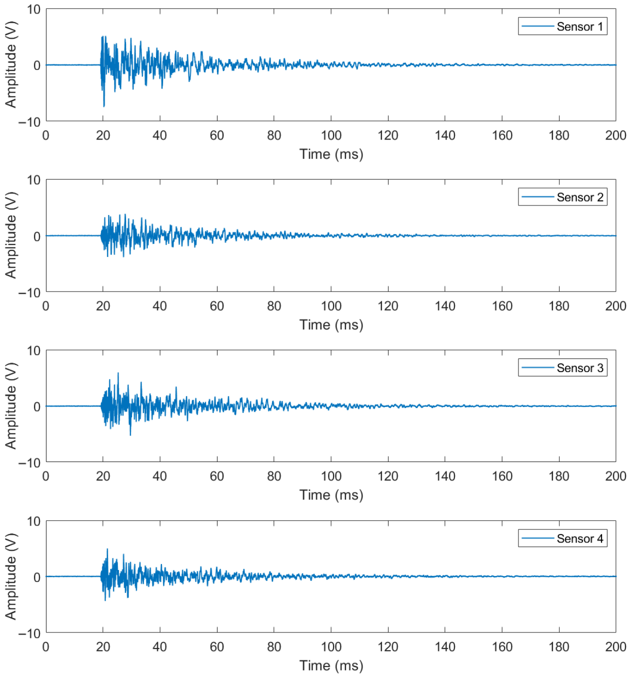


Figure 12. Temporal signal during impact using PZT sensors.

3.1. *Differences Among ΔToF Estimation Methods*

Figure 13 shows the estimations of the ΔToFs of a pair of sensors using the three different methods and considering both spheres for generating the impacts. The impact point coordinates as defined on the aluminum plate are represented on the X and Y axes, while the ΔToFs are reported on the Z axis.

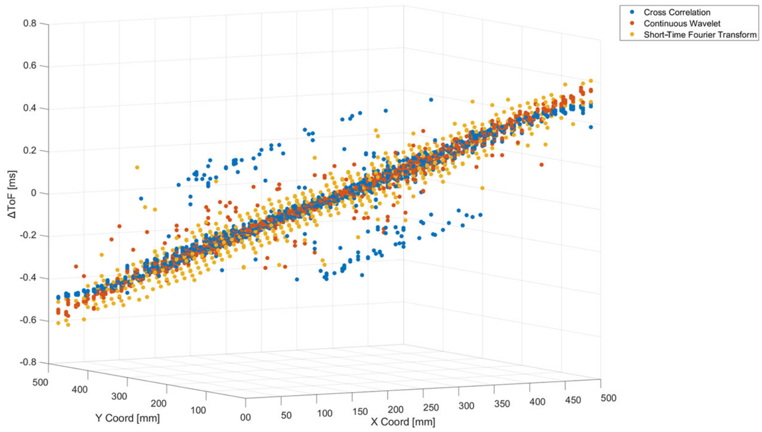
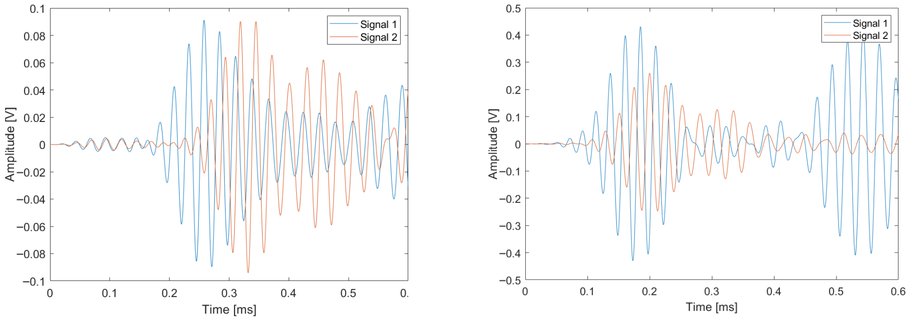


Figure 13. Resulting ΔToFs using all estimation methods.

In this case, the ΔToFs of sensor 1 and sensor 3 in a pair are reported, placed on the diagonal of the regular grid. Sensor 1 is the reference and has coordinates (0, 500) while sensor 3 is placed at (500, 0). Since the plate material is homogeneous with no discontinuities, the wave velocity inside it is constant. Consequently, the ΔToFs of the sensors placed on the diagonal are distributed on a plane, as can be seen in Figure 10.

CC is the method that presents more points that deviate more from the ideal behavior, clearly visible in Figure 10. This result is due to the method used for extracting the initial part of the signal. In some cases, in fact, the selected signal length also includes reverberations that affect the analysis. Figure 14 reports examples of different signals analyzed using the CC method. Figure 14a shows a signal where only the initial part of the impact is extracted, while Figure 14b represents a case where there are also some reverberations. It has to be noted that the represented signals are from two different impacts.



(**a**) (**b**)

Figure 14. Examples of signals analyzed using CC. (**a**) only the initial impact is present. (**b**) some reverberations are present. Signal 1 and signal 2 refer to two different impacts.

Considering the ΔToFs estimated using STFT, the values are more diffused along the ideal fitting plane of the points shown in Figure 13. This is due to the temporal resolution linked to the selected window length.

For CWT, the points are closer and less diffused near the ideal plane due to the better temporal resolution of this method.

The variabilities of the ΔToFs are evaluated across four repeated impacts in all locations. In Figure 13, as an example, the ΔToF variability of sensor pair 1–3 is reported for each estimation method.

For CC in Figure 15a, ΔToF variability values are in the range from 0.0004 ms to 0.20 ms. The greater variability is acknowledged in the direction of the line connecting the two sensors, along the diagonal of the grid. The same behavior occurs considering the other pairs of sensors. This appears to be a consequence of the signal extraction method used for selecting which part to analyze, as described in Section 2.

The ΔToFs estimated using the CWT method have a minimum variability of 0.0002 ms and a maximum value of 0.0974 ms. The impacts with greater variability appear to be randomly distributed on the grid.

Finally, the minimum variability of STFT is equal to zero while the maximum value is 0.14 ms. Also, in this case, point variability distribution does not present a regular pattern.

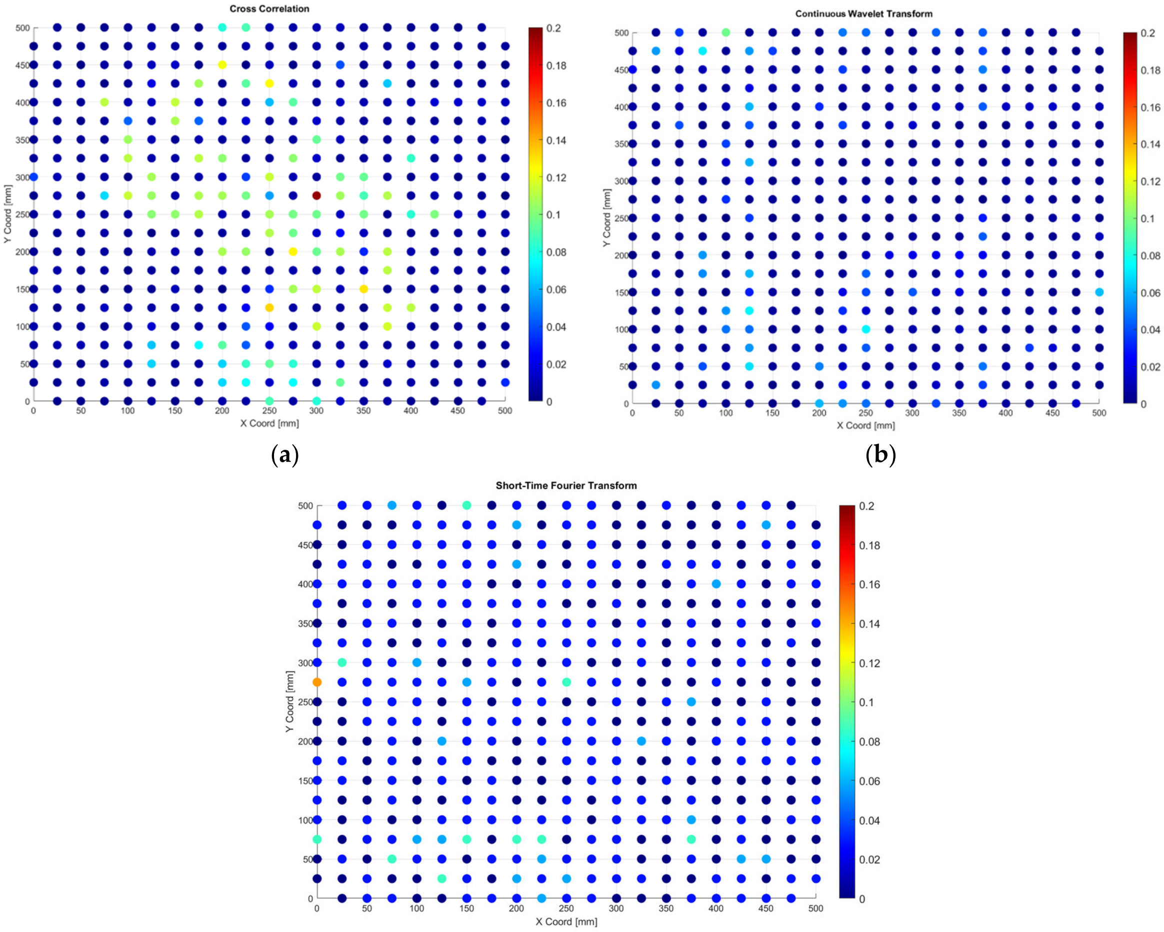
The mean ΔToF variability of the different estimation methods is summarized in Table 1.

Table 1. Mean variability of different ΔToF estimation methods.

ΔToF Estimation Method	Mean Variability [ms]
CC	0.012
CWT	0.008
STFT	0.019

As reported in Table 1, CWT presents less variability compared to the other two approaches.

A repeatability test is performed, repeating 20 impacts on one location and considering the signals from sensor 1 and sensor 3. The ΔToFs are calculated using the three estimation methods and the repeatability is calculated considering the standard deviation. The repeatability of the ΔToFs estimated using CC is equal to 0.009 ms, while with CWT it is 0.002 ms and using STFT it is 0.030 ms. These results are satisfactory and prove that the handheld tube with careful positioning ensures the repeatability of the impacts within a maximum variability of 4–6%.



(**a**) (**b**)

(**c**)

Figure 15. Variability of ΔToFs using Cross Correlation (**a**), Continuous Wavelet Transform (**b**), and Short-Time Fourier Transform (**c**).

3.2. Classification Models

As described in the previous section, the dataset is composed of 1748 impact signals, of which 80% are used for training while 20% are for model testing. To avoid overfitting, a k-fold cross-validation scheme is employed, partitioning the training dataset into five subfolders.

The features used as input of the classification models are the six ΔToFs from all sensor pairs, while the output is the impact zone.

A classification model is trained for each ΔToF estimation method, and different ML models are trained to find the best one. Testing accuracy is used as the first comparison metric of the different classifiers. Table 2 summarizes the performance of the trained classification models.

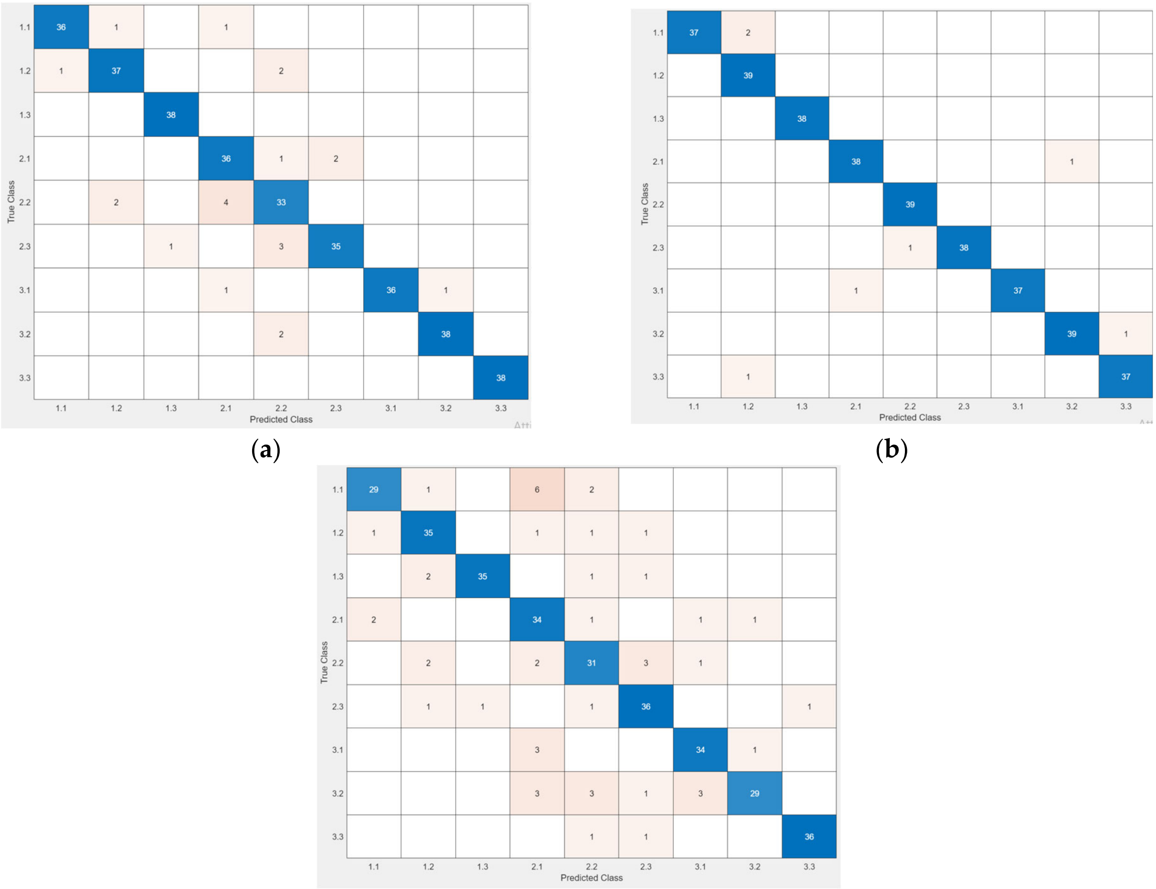
Table 2. Performance of trained classification models.

ΔToF Data	Model	Validation Accuracy [%]	Testing Accuracy [%]
CC	Ensemble Subspace KNN	94.3	93.7
CWT	Ensemble Boosted Trees	97.2	98.0
STFT	Weighted KNN	85.6	85.7

The performance of the models varies depending on the ΔToF estimation method. For CC data, the best performing model is an Ensemble Subspace KNN, which utilizes K-Nearest Neighbors (KNNs) enhanced by subspace and ensemble learning. It trains 30 KNN classifiers on different subsets of data with fewer features to improve performance. Considering CWT data, an Ensemble Boosted Trees model is used. This is composed of 30 decision trees trained using the AdaBoost algorithm. Finally, for STFT ΔToFs, the selected model is a Weighted KNN trained with 10 nearest neighbors. The weights for each of them are the squared inverse of the Euclidean distance.

Comparing the different ΔToF estimation methods, the model trained on the CWT data provides the best results with a classification accuracy of 98%. This is a consequence of the reduced variability of the ΔToFs, as shown in the previous section. For all trained models, the 5-fold cross-validation scheme avoids overfitting since the validation and testing accuracies are similar.

Figure 16 represents the confusion matrixes of the trained models in the testing phase, enabling us to better understand the classification capabilities of each trained model.



(**a**) (**b**)

(**c**)

Figure 16. The confusion matrix in the testing phase for the classification models trained using CC (**a**), CWT (**b**), and STFT (**c**) ΔToF estimation algorithms. The blue squares on the diagonal and the corresponding numbers represent the correctly classified impacts for each zone. On the contrary, the red ones outside the diagonal are the number of impacts classified incorrectly for each zone.

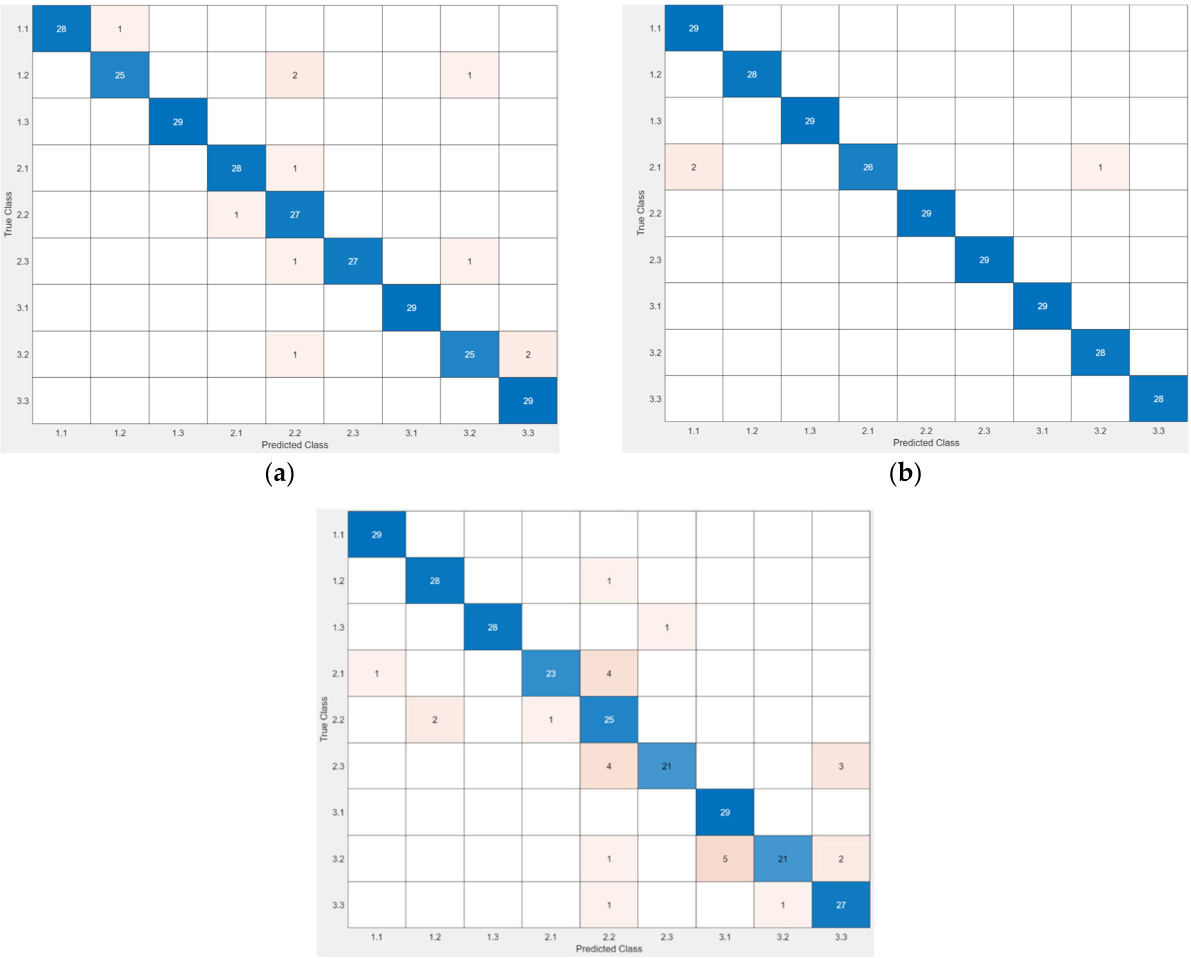
From the confusion matrixes reported in Figure 16, for CC- and CWT-trained models, wrongly classified impacts are mostly localized in zones adjacent to the real one. The zone that is affected the most by this effect is 2.2, since it borders all the others. Considering the model trained using ΔToFs estimated with STFT (Figure 16c), there are several classification errors and many of them are not localized in nearby zones due to STFT frequency and time resolution.

3.2.1. Effect of Impact Point Layout

The effect of impact point layout is evaluated by removing the points on the borders of each zone. With this arrangement, errors due to the misclassification of impacts in zones close to each other are reduced. Table 3 summarizes the accuracy of the trained models using all impact points and the dataset without data on the borders.

As reported in Table 3, by removing the points on the borders, an increment in validation and testing accuracy is observed for all classification models. This is due to the reduction in misclassification errors in contiguous zones.

In Figure 17, the confusion matrixes of the models are reported.



(**a**)

(**b**)

(**c**)

Figure 17. The confusion matrix of the classification models trained using CC (**a**), CWT (**b**), and STFT (**c**) ΔToF estimation algorithms after removing the impacts on the borders. The blue squares along the diagonal and the associated numbers represent the correctly classified impacts for each zone. In contrast, the red squares outside the diagonal indicate the number of impacts that were incorrectly classified for each zone.

As shown in Figure 17, there is a reduction in errors in contiguous zones for all models. However, as shown in Figure 17a, zone 2.2 still has different impacts classified in zones next to it. As shown in Figure 17c, eliminating the points on the border also improves the performance of the STFT model.

Table 3. A summary of the accuracy of the classification models trained using all impact points and removing the points on borders.

	All Points		**No Points on Borders**	
ΔToF Data	**Validation Accuracy [%]**	**Testing Accuracy [%]**	**Validation Accuracy [%]**	**Testing Accuracy [%]**
CC	94.3	93.7	97.0	95.7
CWT	97.2	98.0	97.6	98.8
STFT	85.6	85.7	90.8	89.5

3.2.2. Effect of Number of Impact Points

The effect of the number of points used for training is evaluated by considering half impacts and a minimum distance between points of 35.35 mm.

Table 4 reports the comparison between the accuracies of the classification models trained using all impact points and using only half of the dataset.

Table 4. A summary of the accuracy of the classification models trained using all impact points and only half of the dataset.

	ALL POINTS		**HALF POINTS**	
ΔToF Data	**Validation Accuracy [%]**	**Testing Accuracy [%]**	**Validation Accuracy [%]**	**Testing Accuracy [%]**
CC	94.3	93.7	93.0	94.3
CWT	97.2	98.0	95.9	96.6
STFT	85.6	85.7	83.9	88.1

As shown in Table 4, no significative effects are observed due to the reduction in impact points.

3.2.3. Effect of Sampling Frequency

In Table 5, a summary of the accuracies of the models trained using subsampled signals is reported.

Table 5. A summary of the accuracy of the classification models trained using a dataset with different sampling frequencies.

	5 MHz		**1 MHz**		**0.5 MHz**	
ΔToF Data	**Validation Accuracy [%]**	**Testing Accuracy [%]**	**Validation Accuracy [%]**	**Testing Accuracy [%]**	**Validation Accuracy [%]**	**Testing Accuracy [%]**
CC	94.3	93.7	94.1	93.4	94.0	93.1
CWT	97.2	98.0	96.9	95.7	97.7	97.7
STFT	85.6	85.7	86.2	85.7	86.8	82.5

For all of the trained classification models, no significant variations are observed in the accuracy of the classification models. This is aspect is of great interest in terms of reducing the computational cost of the analysis, which required further analysis compared to previous works [10]. A previous study focused on a regression analysis for impact localization, which exploited the maximum sampling frequency of 5.0 MHz provided by the oscilloscope. Identification of the impacted area with lower acquisition and computational costs was confirmed. This represents a promising result for the optimization of the overall procedure, since the accuracy of classification is a global indicator, accounting for sampling frequency effects as well as the other effects analyzed.

3.2.4. Reproducibility Assessment

To verify the reproducibility of the analysis and further test the classification models, the above trained classification models are used with a different ΔToF dataset, obtained by means of tests carried out in a different laboratory under similar but not identical conditions; the models were trained with three datasets: one considering the total number of points, one excluding the points near the borders, and one using half of the points.

The results in terms of the accuracy of this analysis are summarized in Table 6.

Table 6. Classification accuracies using trained models on different datasets.

Model	TRAINING DATASET		
	All Points	**No Borders**	**Half Points**
CC	94.6	91.0	79.6
CWT	94.6	95.2	86.2
STFT	79.6	77.2	80.2

The models trained with the ΔToFs estimated with CWT proved to be the most accurate with respect to changes in the experimental setup. They showed satisfactory reproducibility even when challenged by a reduction in the number of training points. In the case of CC, a high sensitivity to the reduction in the number of points was highlighted. On the other hand, the STFT model showed less reproducibility when different experimental setup datasets were used as input. This outcome is probably due to the worse temporal resolution of the method.

Our results show that the models trained using CWT as the ΔToF estimator are more robust to reductions in training data and can classify impact signals acquired with a different setup while providing satisfactory accuracy. This promising outcome regarding the reproducibility of data processing and analysis is a relevant result for in-field implementation.

Therefore, as a final consideration, the aspects with the greatest impact on classification accuracy are the algorithm used to estimate the ToFs, the number of points used to train the ML models, the classification model chosen, the geometry of impact point distribution, and the balance of points across the classification areas; instead, a negligible effect was observed for acquisition frequency.

4. Conclusions

In this paper, the effects of different parameters were studied for classifying the localization of an impact on an aerospace component. The analysis of variability in different aspects of the methodology provides suggestions for improving the reproducibility of results. The main aspects that influence variability and bias errors in the results have been studied. These are related to the experimental procedure, such as the layout of the impact grid and its density, the shape of the impact areas to be classified, and the frequency of acquisition of data, and to the data processing technique, such as algorithms for estimating the differences in Times of Flight (ΔToFs) as well as ML classifiers and their parameters. Tests were carried out on an aluminum plate with four PZT sensors to acquire impact signals. Among CC, CWT, and STFT, it was found that CWT estimations of ΔToFs have the lowest mean variability, which is equal to 0.008 ms, and can be used as training features for different classification models. The optimized classification method shows a testing classification accuracy of 98%. Finally, these trained and improved models have been applied to datasets obtained in a different laboratory with the aim of studying the reproducibility of the results as well as the potential for an in-field transfer of the developed techniques. Our results show that the trained models exhibited an accuracy in the order of 95%, even though many testing conditions were changed such as the data acquisition system and the algorithm used to estimate ΔToFs, the data acquisition frequency, the operators, the impact spheres, and the laboratory environmental conditions. These results demonstrate the potential of the procedure not only for use with homogeneous materials but also for cases where more complex materials are such as in aerospace applications. Having quantitively assessed a possible benchmark for a data-driven ML approach for SHM application, the next step is to integrate it with advanced analytical models. This step is necessary for analyzing more complex components, such as composite plates made of carbon fiber or wing-shaped parts and reinforced components with screw holes. Indeed, the application of this procedure to more complex materials will be a further step in this project's work.

Author Contributions: Conceptualization, G.D., E.N., A.G., and F.N.; methodology, L.C., E.N., and A.G.; software, L.C. and F.N.; validation, L.C., A.G., and E.N.; formal analysis, L.C., E.N., A.G., and F.N.; investigation, L.C., E.N., A.G., and F.N.; resources, G.D. and G.S.; data curation, L.C. and F.N.; writing—original draft preparation, L.C. and E.N.; writing—review and editing, L.C., E.N., A.G., and F.N.; visualization, L.C.; supervision, G.D.; funding acquisition, G.D. and G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SHM	Structural Health Monitoring
ML	Machine Learning
ToF	Time of Flight
NDT	Non-Destructive Testing
AE	Acoustic Emission
CC	Cross Correlation
CWT	Continuous Wavelet Transform
STFT	Short-Time Fourier Transform
SAE	Stacked Autoencoder
PWAS	Piezo Wafer Active sensor

Appendix A

ΔToF Estimation Method Pseudocodes

In this section, the pseudocodes for the ΔToF estimation methods are reported. All the algorithms consider only a pair of sensors.

Algorithm 1: CrossCorrelationToF

1.

Input:
impactData

Output:
signalDelay: maxLagms

Algorithm:

```
BEGIN
// LOAD DATA
LOAD impactData

// TIME AXIS SETUP
SET sf TO $5 \times 10^6$// Sampling frequency in Hz
SET t TO (0 TO size(impactData, 1) - 1) × (1000 / sf)  // Time axis in ms

// DETECT CHANGE POINTS
FOR i FROM 1 TO 4 DO
    SET changePt[i] TO findchangepts(impactData[:, i], “Statistic”, “std”)

SET minChangePt TO min(changePt)

// CONSIDER 0.6 ms OF SIGNAL
sensorData = impactData[minChangePt TO minChangePt + (0.6 / 1000 × sf), :]

// CREATE SIGNALS
SET signal1 TO sensorData[:, 1] // SENSOR 1 AND SENSOR 2 ARE CONSIDERED
SET signal2 TO sensorData[:, 2]

// DESIGN BANDPASS FILTER
SET lowCutoff TO 35,000
SET highCutoff TO 45,000
SET order TO 4
[b, a] = butter(order, [lowCutoff, highCutoff] / (sf / 2), ‘bandpass’)

// FILTER SIGNALS
signal1 = filter(b, a, signal1)
signal2 = filter(b, a, signal2)
// CALCULATE NORMALIZED CROSS-CORRELATION
[cross, lags] = xcorr(signal1, signal2, [], “normalized”)
lagsms = lags * 1000 / sf //LAGS IN ms
// FIND MAXIMUM CROSS-CORRELATION AND LAG
[maxCross, maxIdx] = max(abs(cross))
SET maxLagms TO lagsms[maxIdx]

END
```

Figure A1. Pseudocode for ΔToF estimation using CC method.

Algorithm 2: ContinuousWaveletToF

1.

Input:
ImpactData

Output:
signalDelay: deltaTof

Algorithm:

```
BEGIN
// LOAD DATA
LOAD impactData

// TIME AXIS SETUP
SET sf TO $5 \times 10^6$ // Sampling frequency in Hz
SET t TO (0 TO size(impactData, 1) - 1) × (1000 / sf) // Time axis in ms

// CREATE SIGNALS
SET signal1 TO impactData[:, 1] // SENSOR 1 AND SENSOR 2 ARE CON-SIDERED
SET signal2 TO impactData[:, 2]
// CALCULATE CONTINUOUS WAVELET TRANSFORM (CWT)
[cfs1, frq1] = cwt(signal1, “morse”, fs, ‘VoicesPerOctave’, 16, ‘FrequencyLimits’, [20,000, 60,000])
[cfs2, frq2] = cwt(signal2, “morse”, fs, ‘VoicesPerOctave’, 16, ‘FrequencyLimits’, [20,000, 60,000])

// FIND PEAKS IN CWT AT 40 kHz
[peaks1, peaks1Idx] = findpeaks(abs(cfs1(10, :)$^2$), “MinPeak-Height”, 0.002)
[peaks2, peaks2Idx] = findpeaks(abs(cfs2(10, :)$^2$), “MinPeakHeight”, 0.002)

// EXTRACT TIME OF FIRST PEAK
SET firstSignal1 TO t[peaks1Idx(1)]
SET firstSignal2 TO t[peaks2Idx(1)]

// CALCULATE DELTA TIME OF FLIGHT
SET deltaTof TO (firstSignal1 - firstSignal2)
END
```

Figure A2. Pseudocode for ΔToF estimation using CWT method.

Algorithm 3: ShortTimeFourierTransformToF

1.

Input:
impactData

Output:
signalDelay: deltaTof

Algorithm:

```
BEGIN

// LOAD DATA
LOAD impactData

// TIME AXIS SETUP
SET sf TO $5 \times 10^6$  // Sampling frequency in Hz
SET t TO (0 TO size(impactData, 1) - 1) × (1000 / sf)  // Time axis in ms

// SET PARAMETERS FOR STFT
SET M TO 25  // Window size of 0.1 ms
SET g TO hamming(m) //Humming window of size M
SET L TO 22  // Number of points for overlap
SET freq TO 40,000  // Frequency to extract in Hz

// CREATE SIGNALS
SET signal1 TO impactData[:, 1] // SENSOR 1 AND SENSOR 2 ARE CON-SIDERED
SET signal2 TO impactData[:, 2]

// CALCULATE STFT
[s1, f1, t1, ps1] = spectrogram(signal1, g, 0, M, fs)
[s2, f2, t2, ps2] = spectrogram(signal2, g, 0, M, fs)

// FIND PEAKS IN STFT
[peaks1, peaks1Idx] = findpeaks(abs(s1(f1 == freq, :)$^2$), "MinPeakHeight", 0.1)
[peaks2, peaks2Idx] = findpeaks(abs(s2(f1 == freq, :)$^2$), "MinPeakHeight", 0.1)

// EXTRACT TIMES OF FIRST PEAK
SET firstSignal1 TO t1(peaks1Idx(1)) × 1000  // Convert to ms
SET firstSignal2 TO t2(peaks2Idx(1)) × 1000  // Convert to ms

// CALCULATE  TIME DIFFERENCE
SET deltaTof TO (firstSignal1 - firstSignal2)

END
```

Figure A3. Pseudocode for ΔToF estimation using STFT method.

References

1. Ricci, F.; Mal, A.K.; Monaco, E.; Maio, L.; Boffa, N.D.; Di Palma, M.; Lecce, L. Guided Waves in Layered Plate with Delaminations. In *Proceedings of the EWSHM—7th European Workshop on Structural Health Monitoring*; Le Cam, V., Mevel, L., Schoefs, F., Eds.; IFFSTTAR, Inria, Université de Nantes: Nantes, France, 2014.
2. Monaco, E.; Boffa, N.D.; Memmolo, V.; Ricci, F.; Maio, L. Detecting Delaminations and Disbondings on Full-Scale Wing Composite Panel by Guided Waves Based SHM System. In *Health Monitoring of Structural and Biological Systems 2016*; SPIE: Bellingham, WA, USA, 2016; Volume 9805, pp. 255–262.
3. Boccardi, S.; Boffa, N.D.; Carlomagno, G.M.; Maio, L.; Meola, C.; Ricci, F. Infrared Thermography and Ultrasonics to Evaluate Composite Materials for Aeronautical Applications. *J. Phys. Conf. Ser.* **2015**, *658*, 012007. [CrossRef]
4. Boccardi, S.; Boffa, N.D.; Carlomagno, G.M.; Del Core, G.; Meola, C.; Russo, P.; Simeoli, G. Inline Monitoring of Basalt-Based Composites under Impact Tests. *Compos. Struct.* **2019**, *210*, 152–158. [CrossRef]
5. Natale, E.; Gaspari, A.; Chiominto, L.; D'Emilia, G.; Stamopoulos, A.G. Morphological Analysis of As-Manufactured Filament Wound Composite Cylinders Using Contact and Non-Contact Inspection Methods. *Eng. Fail. Anal.* **2024**, *158*, 108011. [CrossRef]
6. D'Emilia, G.; Gaspari, A.; Natale, E.; Stamopoulos, A.G.; Di Ilio, A. Experimental and Numerical Analysis of the Defects Induced by the Thermoforming Process on Woven Textile Thermoplastic Composites. *Eng. Fail. Anal.* **2022**, *135*, 106093. [CrossRef]

7. Beard, S.J.; Kumar, A.; Qing, X.; Chan, H.L.; Zhang, C.; Ooi, T.K. Practical Issues in Real-World Implementation of Structural Health Monitoring Systems. In *Smart Structures and Materials 2005: Industrial and Commercial Applications of Smart Structures Technologies*; SPIE: Bellingham, WA, USA, 2005; Volume 5762, pp. 196–203.
8. Malekloo, A.; Ozer, E.; AlHamaydeh, M.; Girolami, M. Machine Learning and Structural Health Monitoring Overview with Emerging Technology and High-Dimensional Data Source Highlights. *Struct. Health Monit.* **2022**, *21*, 1906–1955. [CrossRef]
9. Ciang, C.C.; Lee, J.-R.; Bang, H.-J. Structural Health Monitoring for a Wind Turbine System: A Review of Damage Detection Methods. *Meas. Sci. Technol.* **2008**, *19*, 122001. [CrossRef]
10. Dipietrangelo, F.; Nicassio, F.; Scarselli, G. Structural Health Monitoring for Impact Localisation via Machine Learning. *Mech. Syst. Signal Process.* **2023**, *183*, 109621. [CrossRef]
11. Kosova, F.; Altay, Ö.; Ünver, H.Ö. Structural Health Monitoring in Aviation: A Comprehensive Review and Future Directions for Machine Learning. *Nondestruct. Test. Eval.* **2025**, *40*, 1–60. [CrossRef]
12. Cha, J.-H.; Kim, Y.; Sathish Kumar, S.K.; Choi, C.; Kim, C.-G. Ultra-High-Molecular-Weight Polyethylene as a Hypervelocity Impact Shielding Material for Space Structures. *Acta Astronaut.* **2020**, *168*, 182–190. [CrossRef]
13. Jaussaud, G.; Rebufa, J.; Fournier, M.; Logeais, M.; Bencheikh, N.; Rébillat, M.; Guskov, M. Improving Lamb Wave Detection for SHM Using a Dedicated LWDS Electronics. In *11th International Symposium on NDT in Aerospace*; COFREND: Paris, France, 2019; Volume 25, pp. 1–7.
14. Ciampa, F.; Meo, M. A New Algorithm for Acoustic Emission Localization and Flexural Group Velocity Determination in Anisotropic Structures. *Compos. Part A Appl. Sci. Manuf.* **2010**, *41*, 1777–1786. [CrossRef]
15. Kundu, T. Acoustic Source Localization. *Ultrasonics* **2014**, *54*, 25–38. [CrossRef] [PubMed]
16. Marino-Merlo, E.; Bulletti, A.; Giannelli, P.; Calzolai, M.; Capineri, L.; Marino-Merlo, E.; Bulletti, A.; Giannelli, P.; Calzolai, M.; Capineri, L. Analysis of Errors in the Estimation of Impact Positions in Plate-Like Structure through the Triangulation Formula by Piezoelectric Sensors Monitoring. *Sensors* **2018**, *18*, 3426. [CrossRef] [PubMed]
17. Ebrahimkhanlou, A.; Dubuc, B.; Salamone, S. A Generalizable Deep Learning Framework for Localizing and Characterizing Acoustic Emission Sources in Riveted Metallic Panels. *Mech. Syst. Signal Process.* **2019**, *130*, 248–272. [CrossRef]
18. Cuomo, S.; De Simone, M.E.; Andreades, C.; Ciampa, F.; Meo, M. Machine Learning for Impact Detection on Composite Structures. *Mater. Today Proc.* **2021**, *34*, 93–98. [CrossRef]
19. Ai, L.; Soltangharaei, V.; Bayat, M.; Van Tooren, M.; Ziehl, P. Detection of Impact on Aircraft Composite Structure Using Machine Learning Techniques. *Meas. Sci. Technol.* **2021**, *32*, 084013. [CrossRef]
20. Chiominto, L.; Natale, E.; D'Emilia, G.; Grieco, S.A.; Prato, A.; Facello, A.; Schiavi, A. Responsiveness and Precision of Digital IMUs under Linear and Curvilinear Motion Conditions for Local Navigation and Positioning in Advanced Smart Mobility. *Micromachines* **2024**, *15*, 727. [CrossRef] [PubMed]
21. Chiominto, L.; D'Emilia, G.; Dipietrangelo, F.; Gaspari, A.; Natale, E.; Nicassio, F.; Scarselli, G. Integrated Sensors for Structural Health Monitoring: Groundwork for an Interlaboratory Comparison. In *2025 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0 & IoT)*; IEEE: New York, NY, USA, 2025; pp. 45–50.
22. Piezoceramics Disks. Available online: https://www.physikinstrumente.com/en/products/piezoelectric-transducers-actuators/disks-rods-and-cylinders (accessed on 31 October 2025).
23. findchangepts. MATLAB Documentation. Available online: https://it.mathworks.com/help/signal/ref/findchangepts.html (accessed on 31 October 2025).
24. xcorr. MATLAB Documentation. Available online: https://it.mathworks.com/help/matlab/ref/xcorr.html (accessed on 31 October 2025).
25. Martinez-Ríos, E.A.; Bustamante-Bello, R.; Navarro-Tuch, S.; Perez-Meana, H. Applications of the Generalized Morse Wavelets: A Review. *IEEE Access* **2023**, *11*, 667–688. [CrossRef]
26. findpeaks. MATLAB Documentation. Available online: https://it.mathworks.com/help/signal/ref/findpeaks.html?s_tid=srchtitle_support_results_1_findpeaks (accessed on 31 October 2025).
27. *ISO/IEC, Guide 98-3:2008*; Uncertainty of Measurement—Part 3: Guide to the Expression of Uncertainty in Measurement. ISO: Geneva, Switzerland, 2018.
28. resample. MATLAB Documentation. Available online: https://it.mathworks.com/help/signal/ref/resample.html?s_tid=srchtitle_support_results_1_resample (accessed on 31 October 2025).