

Void detection in structural adhesive joints using a k-Nearest Neighbors model with features from Electromechanical Impedance Spectroscopy

Pedro A. O. Vilela

Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal (up202007668@edu.fe.up.pt) ORCID 0009-0007-2082-1972

A. Francisco G. Tenreiro

Advanced Joining Processes Unit (AJPU), Institute of Science and Innovation in Mechanical Engineering (INEGI), Rua Dr. Roberto Frias, 4200-465 Porto, Portugal (up201304383@edu.fe.up.pt) ORCID 0000-0002-6410-0196

António M. F. Mendes Lopes

Department of Mechanical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal (aml@fe.up.pt) ORCID 0000-0001-7359-4370

Lucas F. M. da Silva

Department of Mechanical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal (lucas@fe.up.pt) ORCID 0000-0003-3272-4591

Author Keywords	Abstract					
Structural health monitoring, electromechanical impedance spectroscopy, adhesive joints, machine learning, k-Nearest Neighbors.	Electromechanical Impedance Spectroscopy (EMIS) is a promising Structural Health Monitoring (SHM) method which allows the early detection of defects by analyzing the structure response to an AC electrical signal swept through a range of high frequencies. In this work, EMIS measurements					
Type: Research article	of pristine and damaged adhesive joints were performed and features were extracted from the experimenta					
ට් Open Access	measurements. These features were inputted to a k-Nearest					
Peer Reviewed	that only one type of features is enough for damage					
CC BY	detection. Furthermore, the use of the Manhattan Distance in the kNN enables a better classification.					

1. Introduction

Structural adhesive bonding has been adopted in the manufacture of vehicular structures, since other conventional mechanical connections require the addition material, increasing weight and fuel consumption (da Silva et al., 2018). However, adhesives are prone to thermal or aging effects (Machado et al., 2018; Brandão et al., 2022), diverse loading conditions (Wei et al., 2022; Sousa et al., 2022) and the accidental introduction of contaminants (Brandão et al., 2022). These phenomena negatively impact the mechanical performance of structural joints, thus requiring adequate damage detection methods, such as Structural Health Monitoring (SHM) that employ various sensors and computer algorithms to determine information on possible damage foci while the vehicle is in operation (Na et al., 2018; Güemes et al., 2020; Tenreiro et al., 2021).

Electromechanical Impedance Spectroscopy (EMIS) is a Structural Health Monitoring (SHM) method that employs the coupled electromechanical behavior of piezoelectric sensors, which

can work simultaneously as actuator, converting an electrical signal to mechanical stress, and as sensor, converting mechanical stress into an electrical signal. In this manner, the electric current, $I(j\omega)$, changes with excitation frequency, ω , for a specific supply voltage, V, and, consequently, the measured electrical impedance, $Z(j\omega)$, is influenced by the local mechanical behavior, thus allowing the detection of voids in structural adhesives through the use of algorithms, such as Machine Learning (ML) models (Na et al., 2018; Tenreiro et al., 2021). This early detection can, in certain cases, avoid defect propagation or, in an advanced stage, failure.

Since aeronautical structures are some of the most promising applications of this technology, which mainly use aluminum components connected by adhesive materials, due to their inherent light weight, aluminum single lap joints will be used to test this technology (Abbas et al., 2019). The obtained experimental measurements are then used to train a k-Nearest Neighbors (kNN) model for void detection. This ML algorithm was used for its simplicity and lower computational cost, unlike what occurs with commonly used artificial neural networks (Rautela and Bijudas, 2019).

2. Experimental Details

2.1. Materials

The manufactured joints have adherends of Aluminum Al 6082, a commonly used aluminum in the aeronautical industry, due to its reduced weight and its relatively high stiffness and resistance.

The structural adhesive selected was the Nagase T-836/R-810, which is produced by Nagase ChemteX (Osaka, Japan). It is a modified epoxy adhesive with high resistance, which means it is stiff, like normal epoxy adhesives. However, being a modified epoxy, it also has some ductility and tenacity.

To ensure contact between the piezoelectric sensor and the aluminum substrate, another adhesive is used, the Plexus MA 422, produced by ITW Performance Polymers (Chicago, Illinois, United States of America). This is a two-part methacrylate adhesive, selected due to the necessity of having an easy deposition and being able to cure at room temperature, during 24 hours, with the application of a low pressure load, due to the sensor fragility (Tenreiro et al., 2023a). The mechanical properties of both adhesives and the aluminum alloy are shown in Table 1 (Tenreiro et al., 2023a). Please note that, as shown in previous work (Tenreiro et al., 2023a, Tenreiro et al., 2023b), the excited mechanical vibrations are of low amplitudes and occur in the elastic domain.

Material	E [GPa]	υ [–]	η [–]	ρ [kg m ⁻³]
Aluminum Al 6082	70.0	0.33	5×10 ⁻⁴	2710
Nagase T-836/R-810	2.55	0.41	0.1	1200
Plexus MA 422	0.994	0.31	0.45	960

 Table 1: Mechanical properties of the used materials (Tenreiro et al., 2023a).

The substrates used to manufacture single lap joints have a length of 97.5mm, a width of 25mm and a thickness of 2mm, as shown in Figure 1.

The manufactured single lap joints were instrumented with piezoelectric sensors PRYY + 1119, fabricated by PI Ceramics (Lederhose, Germany), to allow the impedance spectrum measurements. The piezoelectric sensor mechanical properties are evidenced in Table 2.



Figure 1: Single Lap Joint Dimensions

Density [3]	Elastic Complia	Coupling Factor		
cm ^o	$S_{11}^{E} \ [m^2 \cdot N^{-1}]$	$S_{33}^{E} [m^2 \cdot N^{-1}]$	k _p	k _t
7.80	16.1×10^{-12}	20.7×10^{-12}	0.62	0.47

Table 2: Piezoelectric sensor mechanical properties (Tenreiro et al., 2023a).

2.2. Specimen Manufacturing

- Substrate Surface Treatment

To ensure good adhesion in the substrates, superficial treatments are required before applying the adhesive. Anodizing has been revealed fundamental in the adhesive joint manufacture with aluminum substrates. This process allows to create a thin oxide layer, in a controlled manner, also creating a surface with pores and whiskers (da Silva et al., 2018).

Initially, the substrates are sand blasted, allowing the removal of existing oxide layers and impurities in the surface and increasing the rugosity of the surface. To perform the anodization, a phosphoric acid aqueous solution, with 10% mass of phosphoric acid was used. Following suggested values by the literature, an electric supply with a voltage of 16V was used, for 25 minutes, with the lowest electric current possible (ASTM, 2004). Afterwards, the surfaces were cleaned using acetone and distilled water, being then left to dry for 30 minutes.

- Manufacture of Single Lap Joint (SLJ)

To ensure the creation of a void in the structural adhesive, it is necessary to use a Polytetrafluoroethylene (PTFE) infill, glued onto the desired position, so that the adhesive does not occupy that volume during the cure process.

After the joints are placed in the mold, the curing process began with pressure and temperature being applied in a hot press. The recommended cure cycle for the adhesive requires a curing temperature of 160°C for 3 hours, as illustrated in Figure 2.



- Instrumentation

Electrical impedance measurements are performed to detect possible structural dynamic behavior changes, thus requiring a connection between the piezoelectric sensor and the structure, which is obtained using the Plexus MA 422 adhesive. It is also necessary to guarantee a good connection of the copper wires to the piezoelectric sensor.

The connection of the copper wire to the top of the piezoelectric sensor is obtained through a soldering process. However, the connection between the other copper wire to the bottom of the piezoelectric sensor is more complicated, due to the necessity of assuring the simultaneous contact of the copper wire to the bottom of the piezoelectric sensor and the aluminum substrate. For that, a copper wire spiral is made, as shown in Figure 3, and during the cure process, a small force is applied to ensure contact between the sensor, the copper spiral, and the substrate.



Figure 3: Scheme of the connection of the piezoelectric sensor to the aluminum substrate.

2.3. Impedance Measurements

To perform impedance measurements of the instrumented joints, the negative and positive poles of the impedance analyzer were connected to the copper wires, creating an electric circuit. A supply voltage of 1V was defined.

With the goal of evaluating the influence of boundary conditions in the spectra, measurements were performed in two different setups. In the first condition, a joint is free and without any boundary conditions, being only supported by the table. In the second condition, shims and clamps are used to fix the joint. The setups are shown in Figure 4.



Figure 4: Boundary conditions for impedance measurements.

Furthermore, for each joint and for each condition, three distinct measurements were performed, varying the frequency range and the scale, allowing one to better analyze different parts of the spectrum. Each measurement contains a total of 801 points, which is the

maximum number of obtainable sampling points with the impedance analyzer. The measurements are listed below:

- 1kHz to 1MHz with linear scale (801 points);
- 1kHz to 1MHz with logarithmic scale (801 points);
- 100kHz to 1MHz with linear scale (801 points);

Measurements obtained through the impedance analyzer were saved in .csv files containing the 801 sampled frequency points and the respective phase and magnitude values of the impedance. A Python script reads these files and uses the phase and magnitude of the impedance to calculate the real and imaginary components of the impedance, saving new .csv files with this additional information.

3. Algorithm Description

3.1. K-Nearest Neighbors (kNN)

The k-Nearest Neighbors (kNN) is a supervised learning algorithm able to perform either classification or regression, depending on the task at hand. In the case of damage detection, a classification kNN model is used. This algorithm, based on the analysis of introduced features, will classify each instance, in this case, pristine or defective joint. Figure 5 presents a flowchart describing the functioning of the classification algorithm.



Figure 5: kNN algorithm flowchart.

For this classification task, an integer value, k, must be given, for the algorithm to determine the k nearest neighbors, and then classify the instances as the most common class in the k nearest neighbors. In this case, of a binary classification, k should be odd, to avoid ties. However, when a tie occurs, the classification to one of the two classes is random.

3.2. Feature Extraction

Between the various possible features that can be extracted from the spectra to train the detection algorithm, the following were selected, with the goal of comparing which of the selected feature combination is more effective in damage detection:

- Frequency of the Re(Z) peaks (n features);
- Real value of the Re(Z) peaks (n features);
- Frequency + Real component value of the Re(Z) peaks (2n features);

• Frequency + Real component value + Imaginary component value of the Re(Z) peaks (3n features);

In all of the considered options, the number of features depends on the number of extracted peaks, *n*, which will also affect the accuracy of the algorithm.

Since, during this research, there was only measurements from a total of 54 SLJs (39 with defect, which were codified as being of class 0, and 15 pristine, which were codified as being of class 1), spectra were joined to obtain a higher number of instances, so that the kNN model could be used without underfitting. Here is the list of combinations used:

- 1kHz to 1MHz Linear (54 instances);
- 1kHz to 1MHz Logarithmic (54 instances);
- (1kHz-1MHz-Lin) + (100kHz-1MHz Lin) (54 instances);
- (1kHz-1MHz Log) + (100kHz-1MHz Lin) (54 instances);
- (1kHz-1MHz-Lin) + (1kHz-1MHz-Log) + (100kHz-1MHz-Lin) (54 instances);



Figure 6: Noise detected as peaks.

Once all the spectrum fusions were performed, the first attempt of feature extraction was done using the *find_peaks* function from the *SciPy* module (Virtanen, 2020). However, noise inherent to the measurements was being detected as peaks, as shown in Figure 6, where peaks are identified with yellow marks. To remove these incorrectly detected peaks, a peak width of at least 2 was required, which was effective and allowed a correct peak detection.

Next, a peak selection was required, since each spectrum has a different number of detected peaks, and all instances should have the same number of features. For that, a criterion to select which peaks to keep and which to remove was necessary. It was decided to keep the peaks with the highest prominence values. However, one still needed to find an optimal number of peaks to keep.

After analyzing the number of peaks of each instance, it was verified that the spectra with the minimum number of peaks was of approximately 30 peaks. Therefore, the number of selected had to be lower, to guarantee that the selected peaks had significant prominence values. In this manner, all instances have 25 peaks.

3.3. Detection Algorithm Results

To evaluate the influence of each parameter in the accuracy of the detection algorithm, default parameters were defined, so that each parameter could be tested separately, keeping the remaining parameters with default values, which were:

- Test_size = 0.2;
- k = 5;
- n = 25;
- Euclidian Distance.

All comparisons were performed with this baseline case, where a kNN model was trained with frequency peaks only, which were extracted from measurements of Fixed SLJs.

3.4. Influence of boundary conditions

To ascertain the influence of boundary conditions of the specimens in the results, ten iterations of the algorithm were performed, presenting the values of the accuracy in Table 3. In Figure 7, a box plot of the accuracy for each boundary condition is presented.

i		1	2	3	4	5	6	7	8	9	10	
Random	_state	13996	21698	29852	15151	64553	19130	63599	14447	40453	65233	Avg
racy 6)	Free	55.56	64.81	64.81	61.11	62.96	68.52	74.07	72.22	62.96	72.22	65.92
Accu (%	Fixed	75.93	77.78	75.93	85.19	79.63	74.07	81.48	72.22	81.48	66.67	77.04



 Table 3: Influence of boundary conditions in the kNN model accuracy.

Figure 7: Accuracy boxplot for each boundary condition.

It was detected that the extracted features from spectra of fixed SLJs allowed the kNN model to obtain a significantly higher accuracy, when compared with trained kNN models of adhesive joints in a free condition.

3.5. Feature influence

In this case, all default parameters were considered, and all four possible feature combinations were extracted from spectra of SLJs with fixed boundary conditions. For each of the possible features, 10 iterations of the detection algorithm were performed, with the values of the accuracy being presented in Table 4. A box plot is also presented in Figure 8.

With the boxplot visualization, it is noticed that the accuracy variation with the selected feature is small. Therefore, using the last two options, where a comparatively high number of

features is used, no significant gain in accuracy is obtained, and, conversely, the use of computational resources has increased.

	i	1	2	3	4	5	6	7	8	9	10	
R	andom_state	50113	30317	10442	11778	37525	29348	51974	25682	41512	11671	Avg
	Freq	70.37	74.07	75.93	72.22	88.89	85.19	77.78	64.81	70.37	74.07	75.37
cy (%)	Re	77.78	68.52	74.07	70.37	75.93	81.48	77.78	72.22	79.63	68.52	75.18
Accura	Freq + Re	70.37	74.07	75.93	72.22	88.89	85.19	77.78	64.81	70.37	74.07	75.37
	Freq + Re + Im	70.37	75.93	75.93	72.22	88.89	85.19	77.78	64.81	70.37	74.07	75.56

 Table 4: Influence of selected feature in the kNN model accuracy.



Figure 8: Accuracy boxplot for each feature selected.

	i	1	2	3	4	5	6	7	8	9	10	
Rand	om_state	61786	15899	2484	55317	46880	83086	52587	47256	22890	59705	Avg
racy ()	Euclidian	79.63	68.52	83.33	72.22	70.37	77.78	68.52	77.78	70.37	79.63	74.82
Accu (%	Manhattan	92.59	70.37	75.93	83.33	83.33	83.33	72.22	83.33	83.33	88.89	81.67

Table 5: Influence of the adopted distance in the kNN model accuracy.

3.6. Influence of the adopted type of distance for the kNN model

In this scenario, only two distances - Euclidian and Manhattan distances -are evaluated. For each distance, 10 iterations were done, with the obtained accuracy values listed in Table 5, and the respective boxplot shown in Figure 9.



Figure 9: Accuracy boxplot for each type of distance.

It was noticed that the accuracy of the kNN model was significantly superior when using the Manhattan distance, when compared to the Euclidian distance, which indicates that the Manhattan distance is more suitable in kNN damage detection algorithms.



Figure 10: Accuracy variation with value of k.

3.7. Influence of k value

To evaluate the influence of k, the number of nearest neighbors for classification, this value was changed from 3 to 50, and ten iterations of the detection algorithm were performed. In this study, the minimum, median and maximum obtained accuracy were compared. In Figure 10, a plot of the variation of these parameters with k is shown, and in Figure 11, examples of confusion matrixes for k=5 and k=50 are presented.



It was expected that, from the representation of the variation of kNN accuracy with the value of k, it would be possible to determine a value, or range of values, for k, to maximize the accuracy of the algorithm. Furthermore, for higher values of k, a decrease in the accuracy was

expected. However, that does not happen because the majority of the SLJs used in this project were defective (195 instances of a total of 270 instances, 72%) and, as the value of k increases, the algorithm tends to classify most of the instances as defective, as shown in Figure 11, thus keeping accuracy values close to the percentage of defective instances, 72%.

3.8. Influence of the dataset division in train and test subsets

To evaluate the influence of the database split into train and test subsets, the test size was increased from 10% to 30%, with a step of 1%, and, for each value, 10 iterations were made, estimating the minimum, median and maximum of the accuracy. In Figure 12, there is a representation of the obtained accuracy with the test dataset size.



Figure 12: Accuracy variation with test size.

It is noticeable that, when a large portion of the dataset becomes part of the test subset, there is a decrease in the maximum and median values of the algorithm accuracy, beginning at 22% and 20%, respectively. However, the minimum value of accuracy does not follow the same pattern.

4. Conclusions

Based on the accuracy of the predictions performed by the classification algorithm, it is clear that boundary conditions have a significant influence on the results, and that measurements with the SLJ fixed are more suitable to the use in damage detection algorithms.

From the analysis of the obtained accuracy with each feature, none of the feature combinations stood out. Therefore, using a higher number of features involves an unnecessary use of computational resources, without an increase of classification accuracy.

It is also possible to conclude that the Manhattan distance is more appropriate to use in kNN damage detection, since the average accuracy was 7% higher than the average accuracy of the kNN model with the Euclidian distance.

One was expecting to find a value, or range of values, for *k*, to optimize the classification performance of the kNN model, and an accuracy decrease for higher values of k. However, it was not possible to find an optimal value for k, and the accuracy drop did not happen because the algorithm tends to classify most of the instances as defective, which corresponds, in a similar proportion, to the portion of defective instances in the database.

By analyzing the variation of accuracy with the test size, it was evidenced that, for values above 20%, the maximum and the median of the accuracy start to decrease. As such, it is not advisable to have a test subset bigger than 20% of the overall dataset size.

References

- Abbas, S., F. Li and J. Qiu. 2018. "A Review on SHM Techniques and Current Challenges for Characteristic Investigation of Damage in Composite Material Components of Aviation Industry". *Materials Performance and Characterization* 7, no. 1: 20170167, https://doi.org/10.1520/MPC20170167.
- American Society for Testing Materials ASTM. 2004. ASTM D3933-98 "Standard Guide for Preparation of Aluminum Surfaces for Structural Adhesives Bonding (Phosphoric Acid Anodizing)". West Conshohocken, PA, United States of America. https://www.astm.org/d3933-98r17.html
- Brandão R., C. S. P. Borges, E. A. S. Marques, R. J. C. Carbas, A. Akhavan-Safar, F. Schmid, C. Ueffing, P. Weißgraeber and L.F.M. da Silva. 2022. "The influence of humidity and immersion temperature on the properties and failure mode of PBT-GF30/silicone bonded joints". *Composite Structures* 289: 115421. https://doi.org/10.1016/j.compstruct.2022.115421.
- Brandão, R., C. S. P. Borges, E. A. S. Marques, R. J. C. Carbas, A. Akhavan-Safar, P. D. P. Nunes, C. Ueffing, P. Weißgraeber, F. Schmid and L. F. M. da Silva. 2023. "Effect of surfactant contamination on the properties of aluminum/silicone adhesive joints". *Mechanics of Advanced Materials and Structures* 30, no. 9: 1875-1888. https://doi.org/10.1080/15376494.2022.2046217.
- Sousa, F., A. Akhavan-Safar, G. Rakesh and L.F.M da Silva. 2022. "Fatigue life estimation of adhesive joints at different mode mixities". *The Journal of Adhesion* 98, no. 1: 1-23. https://doi.org/10.1080/00218464.2020.1804376.
- da Silva, L. F. M., A. Öchsner and R. D. Adams. 2018. "Handbook of Adhesion Technology". Springer Science & Business Media, Heidelberg. https://doi.org/10.1007/978-3-642-01169-6.
- Güemes, A., A. Fernandez-Lopez, A. R. Pozo and J. Sierra-Pérez. 2020. "Structural health monitoring for advanced composite structures: A review". *Journal of Composite Science* 4, no. 1: 13. https://doi.org/10.3390/jcs4010013.
- Machado, J., E. Marques E and L. F. M. da Silva. 2018. "Influence of low and high temperature on mixed adhesive joints under quasi-static and impact conditions". *Composite Structures* 194: 68–79. https://doi.org/10.1016/j.compstruct.2018.03.093.
- Na, W. and J. Baek. 2018. "A review of the piezoelectric electromechanical impedance based structural health monitoring technique for engineering structures". Sensors 18, no. 5: 1307. https://doi.org/10.3390/s18051307.
- Rautela, M., and C. R. Bijudas. 2019. "Electromechanical Admittance Based Integrated Health Monitoring of Adhesive Bonded Beams Using Surface Bonded Piezoelectric Transducers". International Journal of Adhesion and Adhesives 94, no. 10: 84–98, https://doi.org/10.1016/j.ijadhadh.2019.05.002.
- Tenreiro, A. F. G., A. M. Lopes and L. F. M. da Silva. 2021. "A review of structural health monitoring of bonded structures using electromechanical impedance spectroscopy". *Structural Health Monitoring* 21, no. 2: 228–249. https://doi.org/10.1177/1475921721993419.
- Tenreiro, A. F. G., A. M. Lopes, L. F. M. da Silva and J. D. P. Amorim. 2023a. "Effect of Mechanical Properties and Geometric Dimensions on Electromechanical Impedance Signatures for Adhesive Joint Integrity Monitoring". *Mechanics of Advanced Materials* and Structures 30, no.7: 1437–52. https://doi.org/10.1080/15376494.2022.2033891.

- Tenreiro, A. F. G., A. M. Lopes, L. F.M. Da Silva and R. J. C. Carbas. 2023b. "Influence of Void Damage on the Electromechanical Impedance Spectra of Single Lap Joints". NDT & E International 138, no.9: 102865. https://doi.org/10.1016/j.ndteint.2023.102865.
- Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. 2020. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". Nature Methods 17, no. 3: 261-272. https://doi.org/10.1038/s41592-019-0686-2.
- Wei, T., N. Jingxin, M. Wenlong, W. Guangbin and F. Yao. 2022. "Effects of hygrothermal aging on the mechanical properties of aluminum alloy adhesive joints for highspeed train applications" *The Journal of Adhesion* 98, no.3: 227-256. https://doi.org/10.1080/00218464.2020.1828878.