Integration of Multiphysic and Machine Learning Techniques for Structural Health Monitoring in CFRP Composites

Amanda Aryda Silva Rodrigues de Sousa[1][0000-0002-0907-2989], Marcela Rodrigues Machado[2][0000-0002-7658-5514], and Rafael de Oliveira Teloli[3][0000-0002-7488-7201]

[1] Department of Mechanical Engineering, University of Brasilia, Brasília, Brazil. eng.amandaaryda@gmail.com
[2] Department of Mechanical Engineering, University of Brasilia, Brasília, Brazil. marcelam@unb.br
[3] Department of Applied Mechanics, D-SMART, Supmicrotech - ENSMM, Institut FEMTO-ST, France. rafael.teloli@femto-st.fr

Abstract. Composite materials have been employed in many engineering structures, such as aircraft, wind turbines, and aerospace structures. These materials are renowned for their superior stiffness, corrosion resistance, fatigue resistance, and wear resistance, as well as their enhanced thermal properties and reduced weight. However, identifying damage in composite structures is still a challenge, given the influence of external factors on the behaviour of these materials, such as temperature variation and their heterogeneity on the meso/micro-scale. This study focuses on the Structural Health Monitoring (SHM) of carbon fibre-reinforced polymer (CFRP) plates under varying temperature conditions. Strategically positioned piezoelectric sensors are employed to generate Lamb waves and monitor their propagation through the structure, enabling the collection of detailed information about potential damage. Damage is simulated by applying test masses with varying diameters, representing different levels of severity. Temperature variation is included in the study to assess its influence on structural behaviour and sensor responses. The data collected by the sensors are processed using the PyMLDA code, a machine-learning tool designed to implement and train models that efficiently classify the structural state. This work demonstrates the potential of integrated and intelligent solutions to advance SHM in composite materials, contributing to the safety and efficiency of critical structures under challenging conditions.

Keywords: Damage Detection· SHM· Machine Learning· Multiphysic· PyMLDA code.

1 Introduction

Composite materials are vital in advanced engineering sectors like aerospace and wind energy due to their superior stiffness-to-weight ratio, corrosion and fatigue resistance, wear performance, and thermal properties. Their design versatility also allows for complex structural designs [1,2]. Nevertheless, these structures often endure harsh mechanical and environmental conditions during service, leading to faster degradation and higher maintenance. Typical defects include matrix cracks, fiber-matrix debonding, interlaminar voids, uneven resin, and poor consolidation [3–6]. Delamination is considered the most critical failure, severely reducing load-bearing capacity and potentially causing sudden catastrophic failure [7].

To reduce costs and improve operational efficiency, a real structural health monitoring method is vital for detecting defects. The successful integration of Machine Learning with SHM in various fields has led to the development of precise and automated procedures. Vibration analysis-based systems are highly promising for structural monitoring and fault diagnosis. Studies by Liu et al. [8] and Jakkamputi et al. [9] demonstrated the effectiveness of Gradient-Boosted Decision Trees (GBDT) and Decision Trees (DT) in damage identification. Additionally, Viotti et al. [10] showed that Support Vector Machine (SVM) with autoregressive features significantly improved autonomous defect detection in thermoplastic composites. Furthermore, research by Pashmforoush et al. [11], Hamdi et al. [12], and Diaz-Escobar et al. [13] utilised K-means and K-nearest neighbours (KNN) to evaluate delamination and classify damage in composites, proving their robustness across different contexts and data. In a related study, Sousa et al. [14] analysed a CFRP plate, focusing on the impact of thickness, temperature, and damage index on structural integrity, classifying 23 damage states and finding KNN and RF to be effective machine learning models. These results underscore the variety and capability of ML techniques for accurate and effective damage detection and classification in composite structures within SHM frameworks.

This study investigates the integrity of a composite structure subjected to temperature changes. It employs an SHM-ML approach utilising the PyMLDA code, developed [15–17], for damage assessment. The input data consists of ultrasonic-guided elastic waves obtained from a carbon fiber-reinforced polymer (CFRP) plate. The PyMLDA code analyses the temporal or frequency responses of the structure to detect, categorise, identify patterns in, and quantify damage, enabling accurate and real-time structural monitoring. Applying this approach in practice can help reduce operational expenses and increase the durability of composite structures, thereby strengthening their sustainability and economic viability.

2 Methodology

The proposed methodology builds upon and enhances the functionalities of the open-source PyMLDA framework [17, 18] , integrating an improved multiclass classification strategy to assess structural damage under varying thermal and mechanical conditions. To monitor the CFRP polymer plate, we adopt the model illustrated in Fig. 1, which follows the original PyMLDA classification pipeline, with specific modifications to increase robustness and predictive accuracy. The workflow is structured into five core stages, each of which is vital in enabling accurate and interpretable damage assessments.

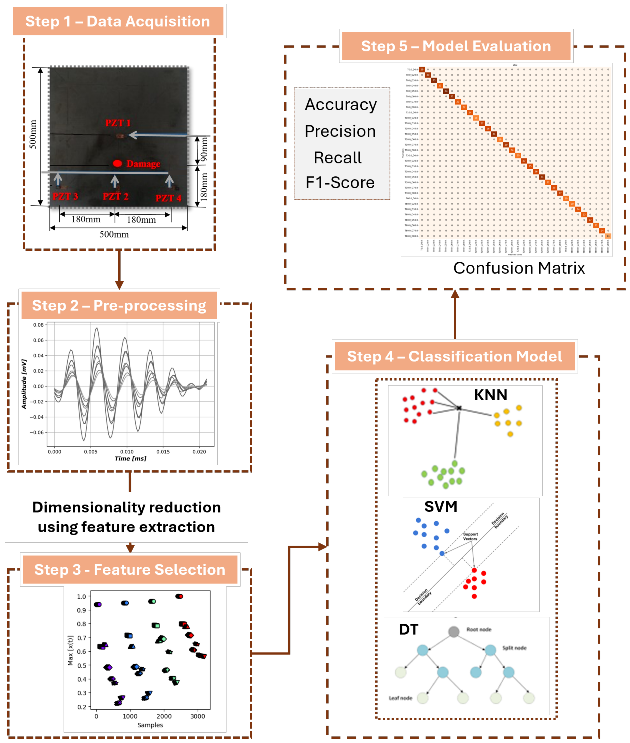


Fig. 1: Schematic functionality of the PyMLDA software, illustrating each step of the proposed methodology.

- **Step 1 – Data Acquisition:** Dynamic response signals are collected from the CFRP structure using embedded sensors, forming the basis for subsequent analyses.
- **Step 2 – Pre-processing:** The acquired signals undergo spectral analysis, where sensor outputs are validated, and relevant components are selected. Further methodological details are presented in Section 3.
- *[1]**Step 3 – Feature Selection and Normalization:** Key statistical descriptors, such as root mean square (RMS), amplitude range, and peak values, are extracted to reduce dimensionality while preserving critical information. Subsequently, normalisation is applied to ensure all features are on

[1] The asterisk (*) indicates a novelty introduced in this work.

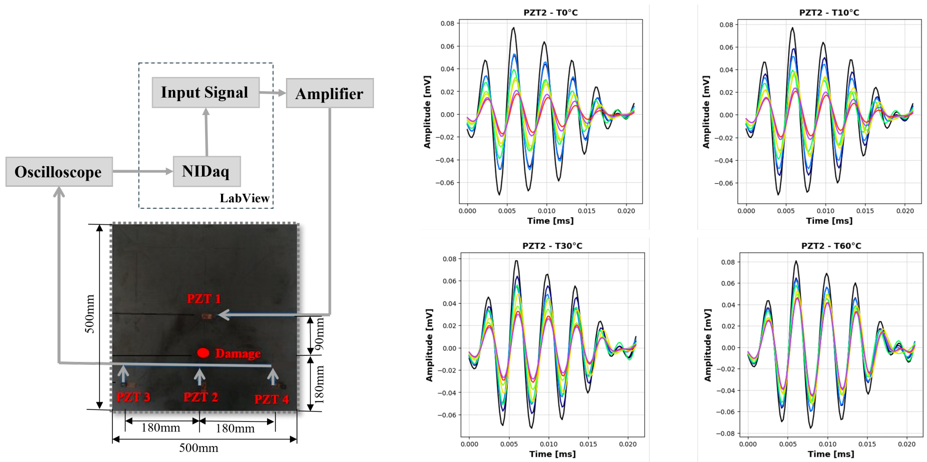
comparable scales, which is crucial for enhancing model convergence and performance.

- **Step 4 – Classification Model:** The dataset is split into 75% for training/validation and 25% for testing. Three supervised multiclass classifiers Support Vector Machine (SVM), k-Nearest Neighbors (KNN), and Decision Tree (DT)—are trained using the Scikit-learn library.
- **Step 5 – Model Evaluation:** The trained models are evaluated on the test dataset. Performance is assessed using standard metrics such as Accuracy, Precision, Recall, and F1-Score. Additionally, a confusion matrix is analysed, and a five-fold cross-validation scheme is implemented to ensure model generalisation and mitigate overfitting.

3 Benchmark Description: CarbON-epoxy CompositE PlacaTe47 Dataset

The dataset used in this study is publicly available through the GitHub repository *CONCEPT: CarbON-epoxy CompositE PlacaTe47* [19]. The experimental setup consists of a CFRP polymer plate positioned under free-free boundary conditions to minimise operational variability, as shown in Figure 2a. The plate measures 500 × 500 × 2 mm and comprises 10 layers of unidirectional woven fibres aligned along its edges. To monitor the plate's behaviour, four SMART layers of PbZrTi (PZT) with a diameter of 6.35 mm and 0.25 mm in thickness are affixed using epoxy resin, with PZT 1 serving as an actuator and the others (PZT 2, PZT 3, and PZT 4) as sensors to acquire the output signal by using a sampling frequency of 5MHz. An excitation signal comprising a 5-cycle sine wave with a 35v amplitude and a 250 kHz frequency is employed, recorded within a 100 μs window utilising suitable acquisition equipment. Industrial adhesive putty is applied to the plate's surface to simulate damage. This additional localised mass simulates local changes in the plate's damping, similar to delamination in composite structures. Data acquisition is conducted across varied controlled temperatures (0°C, 10°C, 30°C, and 60°C) employing a Thermotron-manufactured therma chamber. The dataset is statistically characterised through 100 repeated measurements for each temperature setting.

Figure 2b shows PZT2 sensor data (time-domain responses at 0-60°C) illustrating the combined influence of simulated damage and temperature on Lamb wave propagation in a carbon-epoxy plate. Proximity to damage makes PZT2 highly sensitive, showing increased signal attenuation with added mass due to localised damping. Temperature variations also cause propagation delays, especially at lower temperatures, complicating damage identification due to this multiphysics interaction. Consequently, subsequent analysis focuses solely on PZT2 signals.



(a) (b)

Fig. 2: Experimental setup: (a) data acquisition strategy; (b) results for progressive damaged conditions from 20 mm to 80 mm at each temperature(0°C, 10°C, 30°C, and 60°C)

3.1 Features extraction and normalisation

As part of the proposed modelling framework, time-domain features are extracted to improve the accuracy of classification tasks. To this end, three signal processing techniques are employed to derive informative descriptors from the raw time-domain signal $x(t)$, as outlined in [20]:

- **Maximum value:** $\max[x(t)]$
- **Amplitude range:** $\max[x(t)] - \min[x(t)]$
- **Root Mean Square (RMS):** $\sqrt{\frac{1}{N}\sum_{k=1}^{N} x(t)_k^2}$

These statistical features are selected to reduce the dataset's dimensionality while preserving key signal characteristics representative of the structural dynamics under varying operational and environmental conditions. All numerical variables are normalised to enhance feature comparability further and promote stability during model training. Specifically, each feature column is scaled according to:

$$x_{\text{norm},i} = \frac{x_i}{\max(x_i)} \tag{1}$$

where x_i denotes the original feature values, $\max(x_i)$ is the maximum value within the respective feature column, and $x_{\text{norm},i}$ is the resulting normalised value. This normalisation procedure maps each feature to the interval $[0, 1]$, ensuring consistent magnitudes across variables and facilitating the convergence of machine learning algorithms.

The effectiveness of the extracted and normalised features is illustrated in Fig. 3, which presents the maximum value (left), amplitude range (centre), and RMS (right) de criptors. The top row displays the variation of each feature across the sample indices, while the bottom row depicts their sensitivity to temperature changes. Damage levels—simulated by the progressive addition of mass (in mm)—are represented by distinct markers: $\circ$ = 0 mm (undamaged), $\triangle$ = 20 mm, = 30mm, $\diamond$ = 40 mm, $\times$ = 50 mm, $+$ = 60 mm, $*$ = 70 mm, $\bullet$ = 80 mm.

These visual encodings enable the intuitive identification of damage progression across varying thermal scenarios. A consistent trend is observed across all features: Normalised values approaching unity (i.e., $x_{\text{norm},i} \geq 0.90$) are strongly associated with undamaged structural states. These high-magnitude values are particularly evident at lower temperatures, suggesting that structurally sound systems retain stable signal responses despite moderate environmental variation.

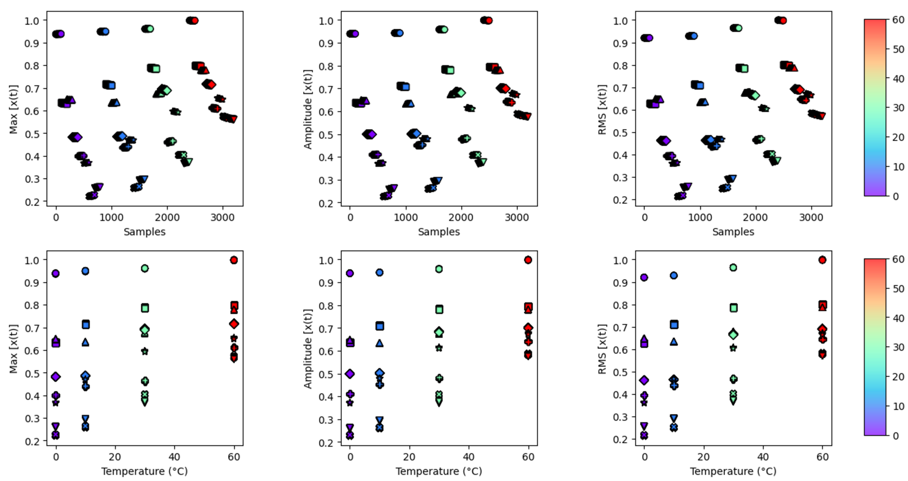


Fig. 3: Normalised time-domain features from the signal $x(t)$: maximum value (left), amplitude range (centre), and RM (right). The top row shows feature values by sample index; the bottom row shows variation with temperature.

4 Results and discussion

Following data preprocessing, a multiclass classification of 800 samples across 32 classes of undamaged and damaged states was performed. Damage, simulated by 30-80 unit mass increments, was assessed at 10°C, 30°C, and 60°C, capturing temporal signal responses under temperature variations. Feature vectors (max, amplitude, RMS) from time-domain signals trained SVM, KNN, and DT classifiers. Cross-validation ensured robustness, averaging performance across folds for unbiased generalisation and overfitting mitigation.

Figure 4 scatter plots compare PZT2 sensor features (Max, Amp, RMS split into two signal divisions) across 0- 60°c. Points below the 0.9 damage threshold dominate at all temperatures, indicating damage. As temperature rises, points approach the threshold, with one healthy reading at 60°C. This suggests consistent damage detection by the descriptors despite temperature variations, which influence PZT2 measurements.

Table 1: **Performance Metrics of Different Classification Models.**

Performance Metric	**SVM**	**KNN**	**DT**
Cross-validation (%)	76.55	99.94	99.61
Accuracy (%)	73.87	100.00	99.75
Precision (%)	65.27	100.00	99.77
Recall (%)	77.26	100.00	99.71
F1-Score (%)	68.94	100.00	99.73

The classification analysis was evaluated using accuracy, precision, recall, F1-score, and the confusion matrix, providing a comprehensive assessment of each model's ability to classify damage states under varying conditions accurately. Cross-validation reinforced the statistical validity and generalizability for real-world SHM. Table 1 shows SVM underperformed (73.87% accuracy, 68.94% F1-score, 76.55% cross-validation), suggesting limited generalisation and sensitivity (77.26% recall) compared to KNN and DT.

Conversely, KNN achieved perfect classification (100% across metrics), and DT performed near-perfectly (99.75% accuracy, 99.73% 1-score). These results highlight the effectiveness of instance-based (KNN) and partitioning-based (DT) learning for the extracted features, significantly outperforming SVM, likely due to the dataset's non-linear separability. KNN and DT are promising approaches for this application, offering superior accuracy and robustness.

As shown in Table 1, KNN outperformed all other classifiers across every metric, followed closely by the DT model. While SVM delivered moderately acceptable results, its comparatively lower precision and F1-score suggest limitations in distinguishing between closely related classes within the dataset.

The confusion matrix for the KNN classifier, shown in Figure 5, visually confirms its flawless classification capability across all 32 class combinations of damage levels (0.0 to 80.0) and temperature settings (T10°C to T60°C). The complete diagonal dominance and the absence of off-diagonal elements indicate that each class was correctly identified without misclassification. This outcome further validates the reported 100% performance metrics and confirms the model's capacity to discern even subtle variations in structural responses under complex operational conditions.

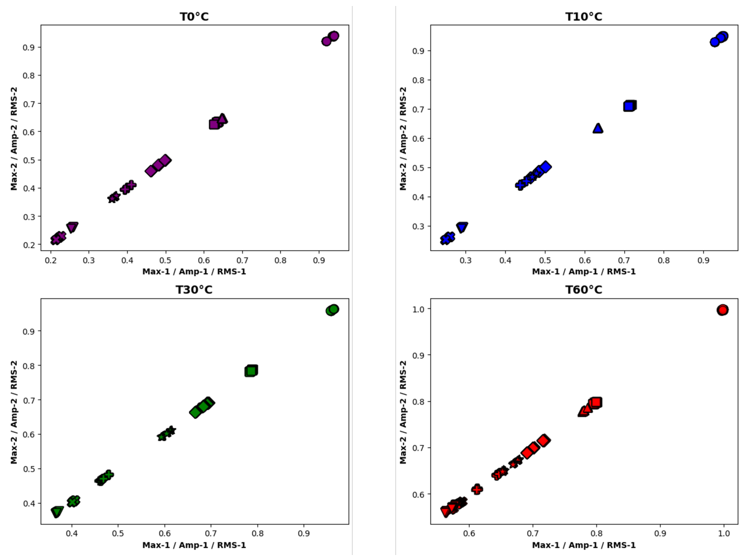


Fig. 4: Feature comparison between temperature conditions using Max, Amplitude, and RMS values.

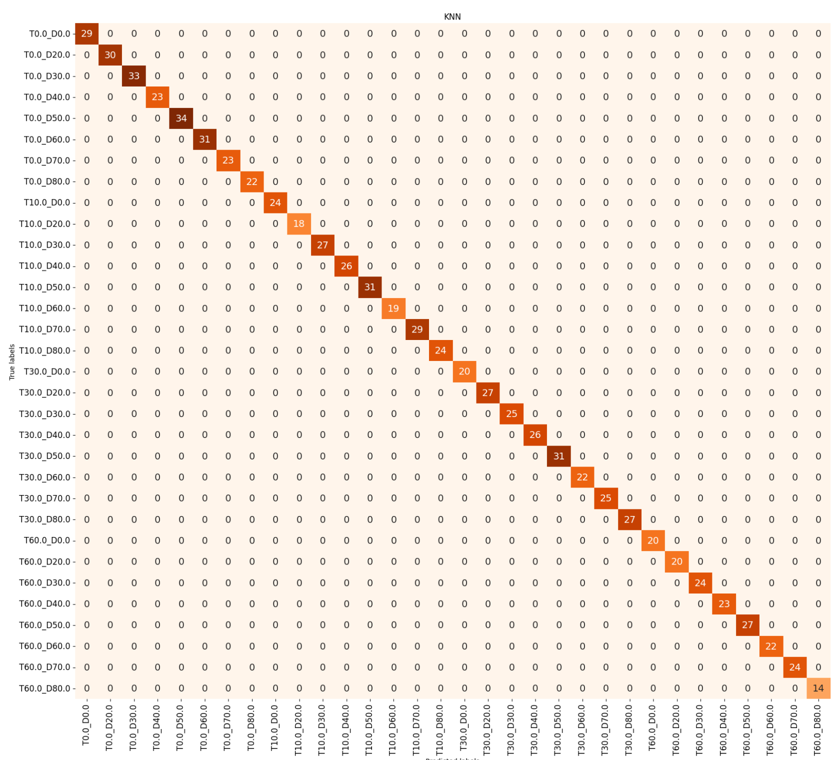


Fig. 5: Confusion matrix for the KNN classifier across all damage and temperature classes.

5 Conclusion

This study proposed a data-driven framework for structural health monitoring that integrates time-domain feature extraction, normalisation, and supervised machine learning techniques to identify damage in composite structures under varying temperature conditions.

A multiclass classification task involving 800 labelled samples spanning healthy and damaged scenarios across three temperature settings was performed using SVM, KNN, and DT classifiers. Evaluation metrics revealed that KNN achieved perfect classification performance, while DT followed closely with near-perfect accuracy and an F1 Score. In contrast, SVM exhibited limited generalisation capability, likely due to the dataset's non-linear separability.

Overall, the results confirm the relevance of the selected time-domain features and validate the effectiveness of instance-based (KNN) and partition-based (DT) learning models in distinguishing damage states under temperature variability. These findings support using lightweight, interpretable features with robust classification algorithms for reliable SHM applications, offering strong potential for real-world deployment in temperature-sensitive environments.

6 Acknowledgements

This research is part of project No. 444595/2024-4, funded by the National Council for Scientific and Technological Development-CNPq, call CNPq/MCTI/FNDCT Nº 22/2024 - Brazil Knowledge Program - Support for Collaborative Projects with Brazilian Researchers Overseas.

Reference

1. Ghrib M, Berthe L, Mechbal N, et al. (2017) Generation of controlled delaminations in composites using symmetrical laser shock configuration. Composite Structures 171: (https://doi.org/286–297. 10.1016/j.compstruct.2017.03.039).
2. Hassani, S., Mousavi, M., Gandomi, A.H.: Structural health monitoring in composite structures: a comprehensive review. Sensors 22(1), 153 (2021). (https://doi.org/10.3390/s22010153).
3. Mei, H.; Giurgiutiu, V. Guided wave excitation and propagation in damped composite plates. Struct. Health Monit.2019,18, 690–714. (https://doi.org/10.1177/147592171876595).
4. Talreja, R. Manufacturing defects in composites and their effects on performance. In Polymer Composites in the Aerospace Industry; Irving, P., Ed.; Woodhead Publishing: Cambridge, UK, 2020; pp. 83–97. (https://doi.org/10.1016/B978-0-08-102679-3.00004-6)
5. Shi, Y.; Swait, T.; Soutis, C. Modelling damage evolution in composite laminates subjected to low velocity impact. Composite Struct.2012,94, 2902–2913. (https://doi.org/10.1016/j.compstruct.2012.03.039).

6. Khan, A.; Kim, H.S.; Youn, B.D. Modeling and assessment of partially debonded piezoelectric sensor in smart composite laminates. Int. J. Mech. Sci. 2017,131–132, 26–36. (https://doi.org/10.1016/j.ijmecsci.2017.06.031).
7. Khan A, Kim HS. A brief overview of delamination localization in laminated composites. Multiscale Sci Eng. 2022;4(3):102–110. (https://doi.org/10.1007/s42493-022-00085-w).
8. Liu P, Xu D, Li J, et al. Damage mode identification of composite wind turbine blade under accelerated fatigue loads using acoustic emission and machine learning. Struct Heal Monit. 2020;19(4):1092–1103. (https://doi.org/10.1177/1475921719878259).
9. Jakkamputi L, Devaraj S, Marikkannan S, et al. Experimental and computational vibration analysis for diagnosing the defects in high performance composite structures using machine learning approach. Appl Sci. 2022;12(23):12100. (https://doi.org/10.3390/app122312100).
10. Viotti ID, Gomes GF. Delamination identification in sandwich composite structures using machine learning techniques. Comput Struct. 2023;280:106990 (https://doi.org/10.1016/j.compstruc.2023.106990)
11. Pashmforoush F, Khamedi R, Fotouhi M, et al. Damage classification of sandwich composites using acoustic emission technique and k-means genetic algorithm. J Nondestruct Eval. 2014;33(4):481–492. (https://doi.org/10.1007/s10921-014-0243-y)
12. Hamdi K, Moreau G, Aboura Z. Digital image correlation, acoustic emission and in-situ microscopy in order to understand composite compression damage behavior. Compos Struct. 2021;258:113424. (https://doi.org/10.1016/j.compstruct.2020.113424)
13. Diaz-Escobar J, Díaz-Montiel P, Venkataraman S, et al. Classification and characterization of damage in composite laminates using electrical resistance tomography and supervised machine learning. Struct Control Heal Monit. 2023;2023:1–19. (https://doi.org/10.1155/2023/1675867).
14. SOUSA, A. A. S. R.; MACHADO, M. R. ; TELOLI, R. O. ; COELHO, J. S. . STRUCTURAL MONITORING OF COMPOSITE CFRP PLATE UNDER TEMPERATURE VARIATION USING THE MACHINE LEARNING-PYMLDA CODE. In: 7th Brazilian Conference on Composite Materials, 2024, Brasília-DF. 7th Brazilian Conference on Composite Materials, 2024. v. 1.
15. Coelho, Jefferson S. ; Machado, Marcela R. ; Dutkiewicz, Maciej ; O. Teloli, Rafael . Data-driven machine learning for pattern recognition and detection of loosening torque in bolted joints. Journal of the Brazilian Society of Mechanical Sciences and Engineering, v. 46, p. 75, 2024. (https://doi.org/10.1007/s40430-023-04628-6).
16. Sousa, A. A. S. R.; Coelho, J. S. ; Machado, M.R. ; Dutkiewicz, M. . Multiclass Supervised Machine Learning Algorithms Applied to Damage and Assessment Using Beam Dynamic Response. Journal of Vibration Engineering e Technologies, p. 1-20, 2023. (https://doi.org/10.1007/s42417-023-01072-7).
17. Coelho, Jefferson S. ; Machado, Marcela R. Sousa, A. A. S. R., PyMLDA: A Python open-source code for Machine Learning Damage Assessment, Software Impacts, Volume 19, 100628.2024. (https://doi.org/10.1016/j.simpa.2024.10062).
18. MACHADO, M.R., and SOUSA, A. A. S. R., and COELHO, J. S. Structural Integrity Monitoring Based on Machine Learning Techniques. INPI BR Patent BR10202401528, July 2024.
19. Silva S and Paixão J. Unesp-concept: Carbon-epoxy composite plate, 2020. (https://doi.org/10.13140/RG.2.2.35767.34722).
20. D. Knittel, and H. Makich, and M. Nouari. Milling diagnosis using artificial intelligence approaches. Software Impacts. 809. 20. 2019.