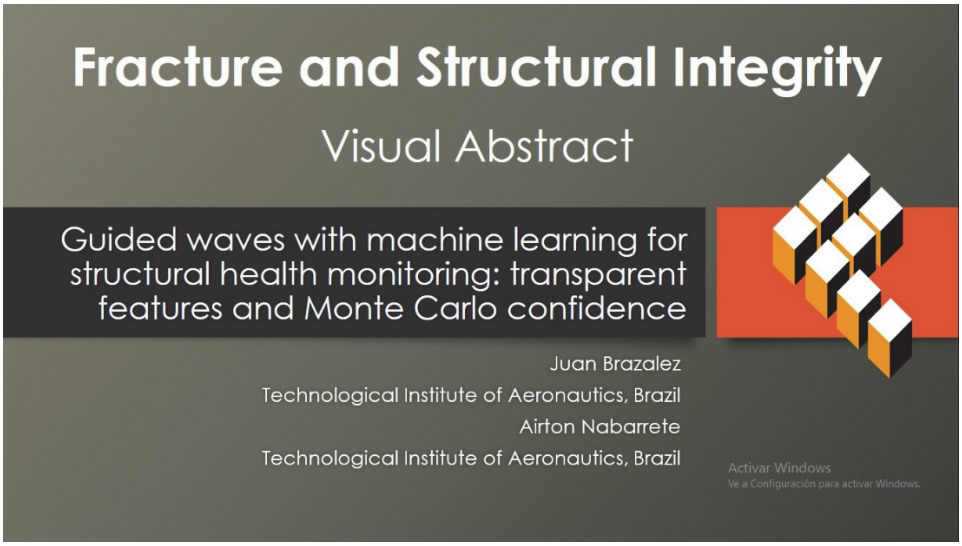
Guided waves with machine learning for structural health monitoring: transparent features and Monte Carlo confidence

Juan Brazalez, Airton Nabarrete
Technological Institute of Aeronautics, Brazil
juanbrazalez@ita.br, http://orcid.org/0000-0002-8464-2646
nabarrete@ita.br, http://orcid.org/0000-0002-1617-9063



KEYWORDS. Structural integrity, Guided Lamb waves, PWAS, Machine learning.

Citation: Brazales, J., Naberrete, A., Guided waves with machine learning for structural health monitoring: transparent features and Monte Carlo confidence, Fracture and Structural Integrity, 76 (2026) 17-30.

Received: 21.09.2025
Accepted: 24.12.2025
Published: 02.01.2026
Issue: 04.2026

INTRODUCTION

Structural Health Monitoring (SHM) is now framed by a concise set of first principles that define what information can and cannot be extracted from in-service data [1]. Axiom I states that all structures possess inherent variability, so damage assessment must be probabilistic rather than deterministic. Axiom II emphasizes that damage manifests only as a change in the measured system response relative to a baseline. Axiom III notes that physics based models alone are insufficient because real structures are never perfectly known. Axiom IV has two corollaries: feature sensitivity to damage increases with damage severity, and environmental or operational variability affects those same features in proportion to their damage sensitivity. Together these axioms demand strategies that acquire a healthy baseline, track relative changes, quantify uncertainty, and balance physics insight with data adaptability. The practical question is how to implement these principles on aerospace platforms where damage can be subtle and intermittent.

Translating this framework into practice, lightweight metallic and composite skins in modern aircraft and wind turbine blades often operate under load spectra that provoke barely visible impact damage, premature delamination, or rivet

loosening before the next scheduled non-destructive inspection. SHM systems based on Lamb wave interrogation with piezoelectric wafer active sensors (PWAS) are attractive because a small transducer network can cover a square-meter scale panel and provide in-flight diagnostics with gram-level mass penalties [2]. Three intertwined challenges follow from this context. First, each interrogation yields multichannel, μs - scale voltage traces that must be reduced to a compact set of robust indices [3]. Second, decisions should be expressed with statistical confidence because feature distributions shift with temperature, manufacturing tolerances, and ageing [1]. Third, purely data driven models struggle to extrapolate beyond the training envelope, whereas purely physics based approaches neglect real-world noise; a reliable solution must fuse both views [4]. These points motivate a development strategy that manages variability, defines auditable thresholds, and documents the path from data to action.

Operationalizing that strategy also requires reliability evidence, not only accuracy. Beyond confusion matrices, structural diagnostics should quantify probability of detection (POD) and sizing uncertainty as functions of defect parameters and inspection conditions. Prior work shows how to estimate these metrics by combining acquisition design, hit–miss modelling, and error propagation for measured flaw dimensions, with careful attention to statistical independence and geometric effects [6]. This perspective extends naturally to SHM classifiers by mapping score distributions to hit–miss outcomes, tuning thresholds to prescribed targets, and attaching prediction intervals to sizing surrogates. With reliability targets in place, the next question is which guided-wave parameters most effectively encode damage information.

On the physics side, recent analyses of power flow in laser generated Lamb waves clarify how energy partitions among modes and along preferred paths in thin plates, providing a basis for selecting excitation frequency, window length, and sensor spacing [7]. This evidence supports the experimental choices used downstream in learning and uncertainty quantification: frequency bands tie to guided-wave transport, windows align with mode specific group velocities, and harmonic metrics follow from energy transfer rather than ad hoc tuning. Against this physics based backdrop, machine learning (ML) supplies the decision layer that converts guided-wave features into actionable classifications under operational variability. Studies in ultrasonic and acoustic-emission non-destructive evaluation (NDE) show that lightweight time series classifiers, including tree ensembles, margin based methods, and boosting, can be benchmarked at the segment level with macro-F1 and receiver operating characteristic (ROC) area under the curve (AUC) before system integration [8–10]. Consistent with that evidence, the present work evaluates multiple model families under a common split and feature set so that performance differences reflect learning bias rather than data handling. This design aligns the workflow with SHM axioms, interpretable features, explicit thresholds, and quantified uncertainty.

Unlike many SHM and ML studies that report point estimates without explicit treatment of measurement uncertainty, in this paper is propagated directly in the signal domain using a Monte Carlo (MC) procedure consistent with the Joint Committee for Guides in Metrology (JCGM) framework for distribution propagation, yielding prediction bands and calibrated decision scores instead of single values. This aligns the learning stage with metrological best practice and with certification oriented reliability thinking in aerospace SHM [3]. In parallel, the feature set is deliberately physics informed and traceable to guided-wave mechanics, including envelope and band energies around the fundamental and the second harmonic, a spectral peak near the excitation band, and interchannel correlation, advancing the interpretability advocated in foundational SHM literature and PWAS based diagnostics for aircraft panels [2,3]. A second harmonic index is also included to capture weak contact nonlinearity at damage interfaces, a mechanism known to enhance sensitivity to early damage yet underused in data driven SHM; this complements linear features and improves separability at close severities. Taken together, these elements provide a transparent and uncertainty aware pipeline that follows the development trajectory recommended for deployable SHM systems [5].

Guided by these axioms, metrics, and physics, laboratory measurements, a calibrated finite element (FE) plate model, and machine learning (ML) classification are combined within a Monte Carlo uncertainty propagation loop. The laboratory campaign is intentionally limited to a single 310 × 190 × 1 mm aluminum plate with one PWAS actuator and three receivers, and to three controlled states (pristine, 16 g, 32 g); the sensor layout, environment, and acquisition chain are fixed. This scope makes the study a laboratory proof of concept rather than a generalizable field study. The FE model validates the pristine waveform and informs feature design. Stochastic variability is introduced at the signal level through MC perturbations, which quantify sensitivity to operational scatter but do not increase the number of independent experiments. The hybrid dataset trains a 200 tree random forest that returns damage state and posterior confidence on every one millisecond window. While the methodology is general and requires only material constants and sensor placement to migrate from aluminum to carbon-epoxy laminates, broader generalization will require multi specimen datasets, additional mass locations and magnitudes, environmental sweeps such as temperature, and external validation consistent with certification metrics such as probability of detection. For the reader's convenience, all acronyms used in this work are summarized in Tab. 1.

Acronym	Description
SHM	Structural Health Monitoring
PWAS	Piezoelectric Wafer Active Sensor
A0	Fundamental antisymmetric Lamb wave mode
S0	Fundamental symmetric Lamb wave mode
ML	Machine Learning
MC	Monte Carlo
FE	Finite Element
FFT	Fast Fourier Transform
STFT	Short-Time Fourier Transform
RMS	Root Mean Square
NDT	Non-Destructive Testing
NDE	Non-Destructive Evaluation
AE	Acoustic Emission
POD	Probability of Detection
ROC	Receiver Operating Characteristic
AUC	Area Under the ROC Curve
RF	Random Forest
SVM	Support Vector Machine
RBF	Radial Basis Function (kernel)
CNN	Convolutional Neural Network
RUL	Remaining Useful Life

Table 1: List of acronyms used in the manuscript.

BACKGROUND

Governing equations for Lamb Waves

For a homogeneous, isotropic plate of thickness h, the in plane (u_x) and out of plane (ω) displacements obey the coupled Navier equations (see Eqn. 1).

$$\rho \ddot{u} = (\lambda + \mu)\nabla(\nabla \cdot \mathrm{u}) + \mu\nabla^2 \mathrm{u} \tag{1}$$

The notation and coupling in Eqn. (1) follow the classical Navier formulation for homogeneous, isotropic plates, where λ and μ are Lamé parameters and ρ is the mass density. Under these assumptions, displacement fields u_x (in plane) and ω (out of plane) are governed by elastic wave equations that admit guided solutions when traction free conditions are enforced on the plate faces [11]. Imposing traction free boundaries at $z=\pm h/2$ yields the Rayleigh–Lamb characteristic relations (see Eqs. 2 and 3).

$$\frac{\tan(qh)}{\tan(ph)} = \frac{-4k^2 pq}{\left(q^2 - k^2\right)^2} \quad \text{(symmetric modes)} \tag{2}$$

$$\frac{\tan(qh)}{\tan(ph)} = \frac{\left(q^2 - k^2\right)^2}{-4k^2 pq} \quad \text{(antisymmetric modes)} \tag{3}$$

Eqs. (2)–(3) are the Rayleigh–Lamb characteristic equations for symmetric (S_n) and antisymmetric (A_n) modes, expressed through the wavenumber k and depth variables p and q with $p^2=\omega^2/(C_L)^2 - k^2$ and $q^2=\omega^2/(C_T)^2 - k^2$. For the present frequency thickness product (f_0h), only S_0 and A_0 lie in the low dispersion region, which enables straightforward time of flight interpretation and robust windowing [11]. With longitudinal and transverse speeds C_L, C_T, at the chosen drive frequency $f_0=20\ kHz$ and h= $1mm$, only the S_0 and A_0 modes are non-dispersive, enabling straightforward time of flight interpretation.

Scattering by a Localized Mass

A surface bonded point mass Δm at $x = x_0$ perturbs the momentum balance, as described by Eqn. 4.

$$\left[\rho h \ddot{\omega}\right]_{x_0^-}^{x_0^+} = \Delta m \ddot{\omega}(x_0, t) \tag{4}$$

Eqn. (4) states the jump condition in transverse momentum at x_0 produced by a bonded point mass Δm. Physically, the attachment injects inertia locally and perturbs the guided field, acting as a weak scattered when $\Delta m \ll \rho h \lambda$ (characteristic wavelength). This setting permits a first order treatment of reflection and transmission without re solving the full boundary value problem [12]. $\Delta m \ll \rho h \lambda$ (small mass limit), the first order Born approximation gives a reflection coefficient (see Eqn. 5).

$$\mathcal{R} \approx \frac{i\omega \Delta m}{2\rho h v_g} \tag{5}$$

so the scattered amplitude grows linearly with Δm while introducing a local phase lag $\Delta \emptyset = k \Delta x$ [2]. In the small mass limit, the first order (Born) approximation yields a reflection coefficient $\mathcal{R}$, with v_g the group velocity of the interrogating mode. Thus, scattered amplitude grows linearly with Δm and carries a phase lag consistent with local inertia loading; higher order corrections become relevant only as Δm approaches the dynamic impedance of the host plate [13, 14].

Damage sensitive feature mapping

Let $x(t)$ be the recorded voltage at a PWAS receiver. After band pass filtering around f_0, the analytic signal (see Eqn. 6) isolates the positive frequency content of the band passed trace and provides well defined instantaneous amplitude and phase. For narrow band guided wave packets, this representation improves timing and amplitude estimates relative to raw $x(t)$, which is advantageous for damage sensitive features [15, 16].

$$x_a(t) = x(t) + i\mathcal{H}\{x(t)\} \tag{6}$$

The envelope (see Eqn. 7) furnishes a smooth amplitude proxy that is less sensitive to carrier oscillations and jitter than peak picking on $x(t)$. When excitation and sensing are tuned to a specific Lamb mode with PWAS, envelope metrics map cleanly onto mode selective energy arrivals and facilitate robust first arrival tracking [17].

$$E(t) = |x_a(t)| = \sqrt{x^2 + \mathcal{H}\{x\}^2} \tag{7}$$

which is proportional to the plate's transverse velocity [18]. RMS (Root Mean Square), mean envelope, and in band spectral peak around f_0 summarize the burst response into physically interpretable scalars tied to wave energy and modal content. Time–frequency analyses of Lamb waves support this windowed feature construction by linking bandpower and group delay markers to specific modes under dispersion [19]. Key scalar features extracted from a 1 *ms* window W given in Eqn. (8).

$$RMS = \sqrt{\frac{1}{|\mathcal{W}|}\int_{\mathcal{W}} x^2 dt}, \qquad E_{mean} = \frac{1}{|\mathcal{W}|}\int_{\mathcal{W}} E dt, \qquad P_{20KHz} = \max_{\omega \in B_{20}} |X(\omega)| \tag{8}$$

Under the weak scattering model, the expectation of each amplitude type feature decreases approximately linearly with Δm, with slope proportional to $-\mathcal{R}$. This first order sensitivity underpins mass (or stiffness) detection thresholds and explains why envelope based features are particularly responsive for low contrast perturbations [13], each feature obeys Eqn. (9).

$$\frac{\partial \langle f \rangle}{\partial(\Delta m)} \propto -\mathcal{R} \tag{9}$$

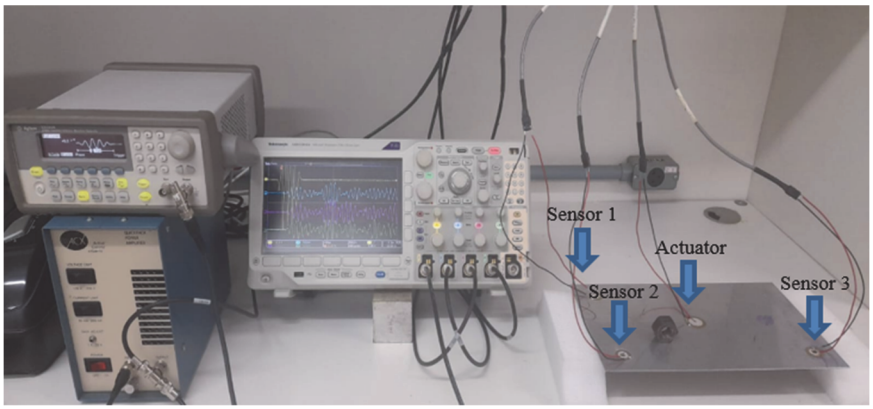
so features decay roughly linearly with Δm.

MC waveform perturbation

MC variability was introduced directly in the measured signals, not in the finite element parameters. For every experimental trace $x(t)$ was generated N_{MC} =5000 synthetic realizations by applying three independent random factors as defined in Eqn. (10). Rather than randomizing FE parameters, variability is introduced directly at the signal level through amplitude scaling, time of flight jitter, and broadband noise, generating N_{MC} realizations per trace. Following metrology guidance, the 2.5th–97.5th percentile envelopes of the propagated feature distribution define 95% prediction bands that quantify operational scatter in arrival timing and amplitude [20].

$$x^{(j)}(t)=\left[1+s^{(j)}\right]x\left(t+\Delta t^{j}\right)+\eta^{(j)}t \tag{10}$$

where the amplitude scale $S\sim\mathcal{N}(0, 0.05^2)(\pm5\%$ RMS), the time of flight shift $\Delta t\sim\mathcal{N}(0, (20\mu s)^2)$, and the additive broadband noise $\eta(t)\sim\mathcal{N}(0,(0.2A_{max})^2)$ with A_{max} = max|x|. The analytic envelope of every perturbed trace yields a feature vector $f^{(j)}$; the set $\{f^{(j)}\}^{N_{MC}}$ approximates the predictive density $p(f|\Delta m)$. The shaded 95 % prediction bands shown in Fig. 5 correspond to the 2.5 and 97.5 percentile envelopes of this density and quantify how operational scatter influences first arrival amplitude and phase [1]. Recent probabilistic studies on cylindrical shells [21] and on composite plates further highlight the need to treat mass, stiffness and loading parameters as random variables when predicting natural frequencies and buckling behavior.



Dimensions
310x190x1 mm

Positions
Actuator (155.95) mm
Sensor 1 (30.160) mm
Sensor 2 (30.30) mm
Sensor 3 (280,30) mm
Mass 1 (92.5, 127.5) mm
Mass 2 (92.5, 62.5) mm
Mass 3 (217.5, 62.5) mm

Figure 1: Experimental Setup.

METHODOLOGY

Experimental campaign

A signal generator (Agilent 33220A) produces a five cycle Hanning windowed tone burst at 20 *kHz*, which is amplified by a high-voltage power amplifier (±350 V) and drives the central PWAS actuator bonded to a 310 × 190 × 1 *mm* aluminum plate, as shown in Fig. 1. The plate rests on soft foam to approximate free-boundary conditions. Three PWAS receivers (Sensor 1, Sensor 2, and Sensor 3) are bonded at the three corners surrounding the actuator: Sensor 1 at the lower-right corner, Sensor 2 at the lower-left corner, and Sensor 3 at the upper-right corner (each approximately 310 *mm* from the actuator). A Tektronix MDO3024 digital oscilloscope samples all three sensor outputs simultaneously at 250 *kHz* with 12 bit resolution. For each mass condition (pristine, 16 g, and 32 g point masses attached at the plate's center), ten consecutive acquisitions are recorded to capture repeatability. All PWAS cabling is shielded and grounded. The choice of f_0 follows standard mode selection practice: for the present frequency–thickness product f_0h, only the fundamental S_0 and A_0 modes are weakly dispersive, enabling robust first arrival timing and short analysis windows [11]. PWAS were chosen for their low mass and efficient electromechanical coupling to guided waves over metre scale footprints [17]. Raw voltages were band pass filtered around f_0 and converted to the analytic signal $x_a(t)=x(t)+iH\{x(t)\}$ using the Hilbert transform; the envelope $E(t)=|x_a(t)|$ and instantaneous phase were then extracted [15 - 16]. Within 1 *ms* windows aligned to the first arrival, physically interpretable features were computed: RMS, mean envelope, in band spectral peaks near 20/40

kHz, and simple correlation markers. Time–frequency literature for Lamb waves supports this windowed feature construction by linking bandpower and group delay markers to mode content under dispersion [19].

FE model

As shown in Fig. 2, the rendered Abaqus plot shows the out of plane displacement U_3 field on the aluminum plate at a fixed time step shortly after excitation. Here, the PWAS actuator at the center has launched a fundamental S_0 Lamb wave packet at 20 *kHz*, and the color contours represent the instantaneous amplitude of the transverse motion (red ≈ +0.03 *mm* down to dark blue ≈ –0.013 *mm*). This simulation reproduces the same arrival times and amplitude attenuation observed experimentally: the added mass locally increases inertia, causing a partial reflection (bright red) and lowering the transmitted wave amplitude. The mesh refinement near the mass plate interface ensures accurate capture of stress concentration and wave mass interaction. Guided wave propagation was simulated on the same 310 × 190 × 1 *mm* plate using a shell/plate formulation suitable for low *f*h*, where S_0/A_0 dominate and Mindlin type kinematics capture the dispersive physics efficiently [22]. The in plane mesh size Δ_S satisfied ≥ 10–12 nodes per shortest wavelength at 20 *kHz* to control numerical dispersion; the explicit time step obeyed a CFL (Courant–Friedrichs–Lewy) type limit $\Delta_t \leq \Delta_S / c_{max} n_{ppw}$ with n_{ppw} ≈20 points per period [23]. Isotropic elastic properties (E, ν, ρ) matched the aluminium plate used experimentally; small Rayleigh damping $C=\alpha M+\beta K$ was introduced to reproduce the measured envelope decay without over damping phase information. The coefficients (α, β) were tuned by minimizing arrival time and amplitude errors on the pristine trace [24]. To avoid spurious reflections from the plate edges over the analysis window, non-reflecting boundaries were implemented using viscous dashpots (Lysmer–Kuhlemeyer) and PML (Perfectly Matched Layers) on the outer strip; both approaches are standard for elastodynamic FE and were verified by monitoring near zero back energy at the boundaries during the burst [25]. Localised mass was introduced as a concentrated surface mass at x_0, consistent with the weak scattering model. In the $\Delta m \ll \rho h \lambda$ regime, the FE reflection trend agrees with the first order Born prediction $R \propto i\omega\Delta m/(2\rho h v_g)$, providing a physics check on the linear amplitude mass relationship.

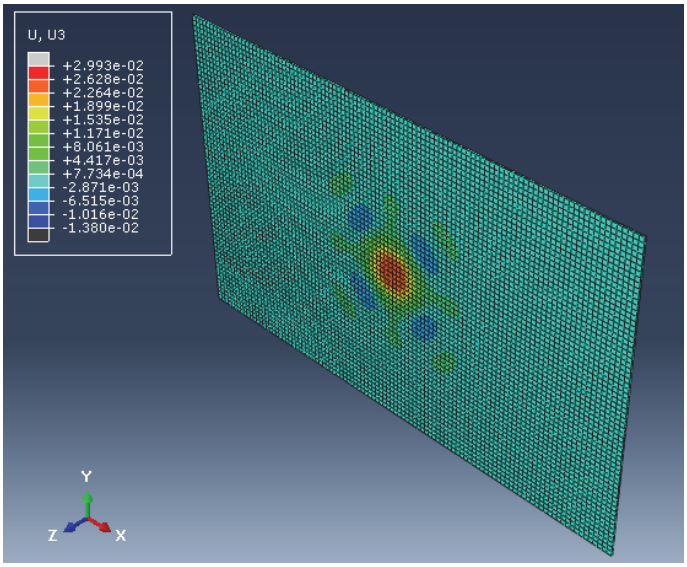


Figure 2: Finite-element wavefield (U_3).

Machine Learning

Machine learning classification was framed as a supervised, multi class time series problem on PWAS voltage windows. After band pass filtering around the fundamental and second harmonic bands, each 1 *ms* segment was converted into a compact feature vector (RMS, peak to peak, analytic envelope means, band energies, spectral peak, and cross channel correlation), z-score normalized, and class balanced by stochastic oversampling. Four complementary learners were benchmarked under a stratified 80/20 split: (i) a RF (bagged trees with out of bag monitoring), (ii) an RBF kernel SVM (one vs one ECOC- Error Correcting Output Codes), (iii) AdaBoost.M2 with shallow decision trees, and (iv) a hierarchical ECOC scheme to reflect the ordinal structure of "pristine/16 g/32 g." Performance was summarized with overall accuracy, macro-F1, per class recall. This design mirrors current practice in ultrasonic and AE (Acoustic Emission) NDE where light weight, real time classifiers are favored and evaluated with segment level metrics before full system deployment [8, 9]. In particular,

recent studies demonstrate (a) end to end time series classification on non-contact ultrasonics with careful benchmarking across model families, corroborating the multi model comparison [8], and (b) waveform to damage size mapping from AE via CNN (Convolutional Neural Network) with data balancing/augmentation, supporting the emphasis on feature engineering and class aware evaluation [9]. Finally, to align with current guidance on distributional robustness in defect identification, Macro-F1 and class specific recall are presented, while stability across operating conditions is examined, consistent with recent studies on model robustness in composite aerostructure inspections [10].

RESULTS AND DISCUSSION

This study combines high-fidelity FE simulations with an experimental campaign on aluminum panels instrumented with PWAS to establish and validate a robust SHM methodology. Numerically, a 310 × 190 × 1 *mm* plate was modeled in ABAQUS to simulate Lamb wave propagation and damage scenarios (e.g., mass attachments of 16 g and 32 g at predefined locations). Material damping, geometric parameters, and actuator/sensor boundary conditions were calibrated to match experimental conditions. A MC study was carried out directly on the measured waveforms: each trace was perturbed 5 000 times by ±5 % amplitude scaling, a normally distributed ±20 μs time of flight shift, and 2 % broadband Gaussian noise. These synthetic realizations capture the operational variability of first arrival amplitude and phase for every damage state and are subsequently used to build prediction envelopes and PDF (Probability Density Functions).

Experimentally, identical aluminum panels were instrumented with a central PWAS actuator and three PWAS receivers placed at the three corners. Each trial comprised five repeated excitations per condition (pristine, 16 g, and 32 g), capturing the time-domain voltage signals at 250 *kHz* sampling. After bandpass filtering around 20 *kHz* (fundamental) and 40 *kHz* (second harmonic), each 1 *ms* window was normalized by its global RMS to mitigate coupling variations. From these windows, twenty features were extracted—time-domain RMS and peak-to-peak, Hilbert-transform envelope means in both passbands, spectral peak at 20 *kHz*, band-energy integrals (18–22 k*Hz*, 38–42 k*Hz*), and inter-channel correlation coefficients. Augmenting the five experimental repeats with MC derived samples yielded a balanced dataset of 315 windows per class (pristine, 16 g, 32 g).

Using this combined dataset, a RF classifier (200 bagged trees) was trained to distinguish "no-mass," "16 g," and "32 g" conditions. Stratified 80/20 partitioning ensured each category appeared proportionally in training and test subsets. The model achieved a pristine (0 g) recall of 85.7 % and correct classification rates of 76.2 % for both 16 g and 32 g cases. Confusion-matrix analysis revealed that overlapping feature distributions between medium and heavy loads—stemming from similar attenuation and scattering—limit discrimination. To address this, phase-based features (time-of-flight) and higher-order spectral moments are proposed for future work.

By integrating physics-based ABAQUS simulations, uncertainty quantification via MC sampling, and machine learning on experimental PWAS data, the approach not only detects and localizes added-mass "damage" with high accuracy but also provides a probabilistic estimate of classifier confidence under real-world variability.

According to Fig. 3, the raw channel 2 (CH2) waveforms demonstrate that all three conditions share the same arrival times (around 0.6 *ms*, 0.8 *ms*, and 1.0 *ms*), but the peak amplitudes decrease systematically as mass increases—pristine (0 g) exhibits the highest peaks, 16 g is intermediate, and 32 g is lowest—reflecting stronger inertial loading and partial reflection of Lamb waves by the added mass. In the analytic envelopes (bottom), this progressive attenuation is even clearer: the pristine envelope reaches ~0.52 V at its maximum, whereas 16 g peaks near ~0.48 V and 32 g only ~0.46 V. Mechanically, adding mass locally increases inertia and reduces flexural stiffness, which elevates reflection coefficients and lowers transmitted energy; consequently, envelope amplitudes diminish with heavier loads, confirming that envelope magnitude is a robust feature for distinguishing "no load" from light and heavy mass conditions.

Fig. 4 illustrates, across all three conditions, the numerical (Abaqus) CH2 response (black) closely tracks the experimental waveform (blue) in both arrival times and overall amplitude envelope. In the pristine case (left), the two principal cycle packets at ~0.6 *ms* and ~0.9 *ms* align almost perfectly, with only minor amplitude deviations under 0.05 V—indicating that the simulated wave speed and boundary conditions match the physical plate very well under no added mass. For the 16 g load (center), both simulated and measured signals show similarly attenuated peaks and slight waveform distortion around each arrival; the peak amplitudes at ~0.6 *ms* and ~1.0 *ms* drop by roughly 20–30 % compared to pristine, and the black and blue envelopes remain within ±0.03 V of one another, demonstrating that Abaqus effectively captures the added-mass scattering. Under the 32 g condition (right), the heavier mass produces greater attenuation and minor phase shifts: the dominant peaks are reduced by ~40 % versus pristine, and a slightly earlier first zero crossing is visible. Again, numerical

and experimental traces remain within ~0.05 V in amplitude and match timing to within ~5 μs, confirming that the FE model accurately reproduces the mechanical effect of added mass on Lamb wave propagation for all three scenarios.

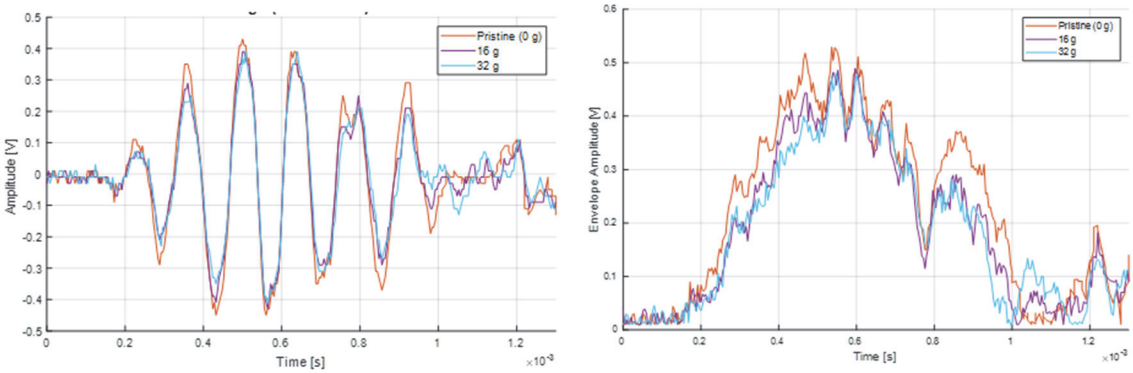


Figure 3: Time traces and analytic envelopes.

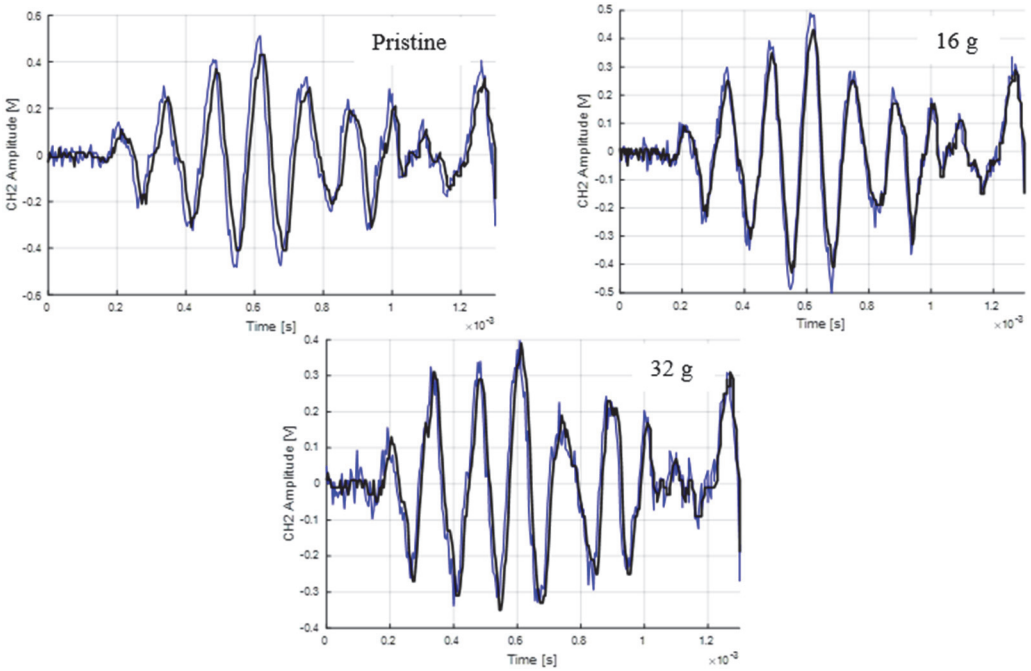


Figure 4: Comparison of Experimental (Blue) vs. Numerical (Black) CH2 Waveforms.

From a mechanical-wave perspective, the Hilbert-transform envelope on CH2 reveals (see Fig. 5) how attaching discrete masses alters local boundary conditions and wave propagation: a pristine plate (0 g) exhibits maximal first-arrival amplitude because the Lamb wave traverses an undisturbed stiffness distribution. When 16 g or 32 g masses are bonded, the increased local inertia and reduced flexural rigidity cause partial reflection and mode conversion, attenuating energy beyond the attachment. Consequently, both the first arrival (~0.6 *ms*) and the second arrival (~1.4 *ms*) envelope peaks drop monotonically as mass increases. The narrow confidence band for the pristine case reflects consistent, low-damping wave travel, whereas the wider band for 32 g indicates greater variability in scattering and damping due to non-linear coupling at the mass interface. This mechanical interaction—added mass → altered dispersion and reflection coefficients → reduced envelope amplitude—underpins why heavier loads yield lower mean envelopes and larger envelope spread.

According to Fig. 6, the spectrum confirms that the actuator excites a dominant fundamental Lamb wave component at ≈ 20 *kHz* and a weaker second harmonic at ≈ 40 *kHz*; higher order content is at least one order of magnitude lower, consistent with a 5 cycle Hanning burst. From a mechanic standpoint, adding a point mass increases local inertia and lowers the local resonant frequency, so part of the incident S_0 energy is reflected and part is mode converted into out of band components. This manifests as a monotonic drop in the 20 *kHz* peak magnitude: pristine ≈ 75 V, 16 g ≈ 65 V, and 32 g ≈ 60 V. The

$\pm 1\sigma$ envelopes (notched regions around each curve) broaden with added mass, indicating greater run to run variability a direct consequence of small slip and bonding layer stiffness fluctuations at the mass–plate interface. Importantly, the peak frequency itself does not shift, staying within ± 100 Hz across all cases; the added masses ($\leq$ 1 % of plate mass per wavelength) therefore act primarily as scattering centres rather than altering the global dispersion relation. The second harmonic band at 40 kHz shows the same amplitude ranking, reinforcing that envelope based features and spectral peaks respond consistently to the mechanical effect of increased point inertia.

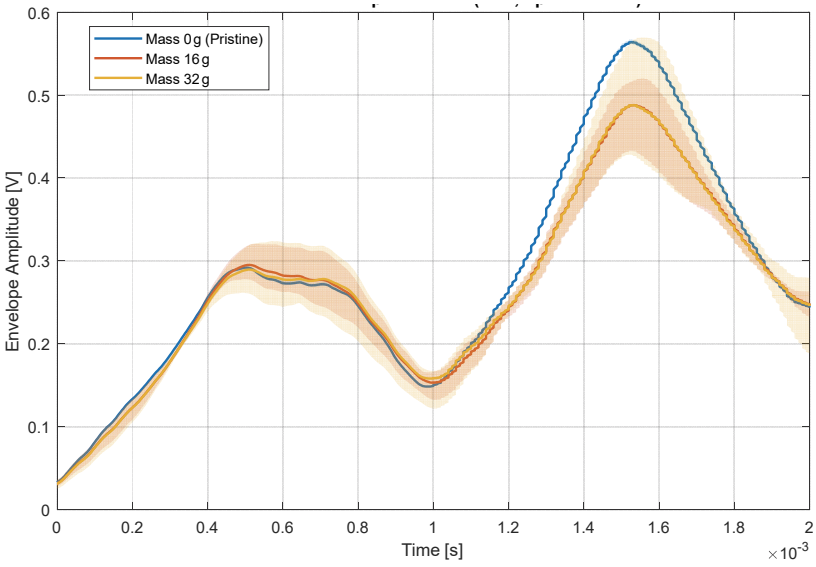


Figure 5: Prediction bands from Monte Carlo perturbations.

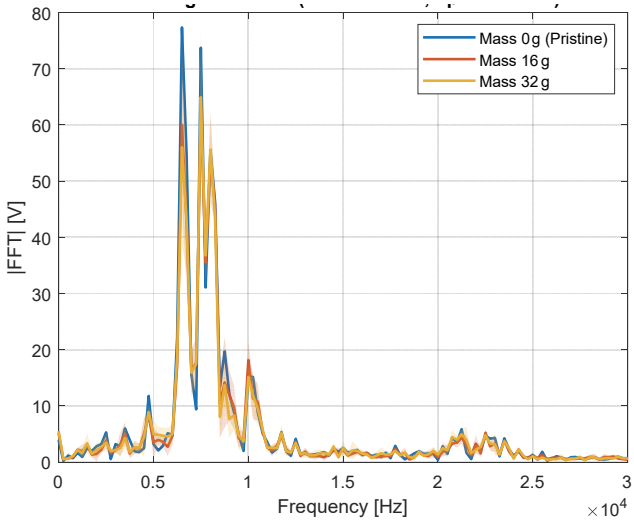


Figure 6: Average one sided FFT $\pm$ 1σ.

In addition to the linear attenuation trend, the averaged FFT (Fast Fourier Transform) reveals a distinct energy lobe at $\approx$ 40 kHz exactly twice the 20 kHz excitation. Such second harmonic content is a classical signature of quadratic acoustic non-linearity generated by micro slip or intermittent contact at the mass plate interface [26]. To quantify this effect, a second harmonic index (SHI) is defined by Eqn. 11.

$$\eta = 20 log_{10}\left(\frac{A_{40}}{A_{20}}\right) \tag{11}$$

where A_{20} and A_{40} are the mean FFT peaks in the 18–22 kHz and 38–42 kHz bands, respectively. The pristine plate yields η=-48 ± 2dB, the 16 g case η=-38 ± 4dB, and the 32 g case η=-31 ± 5dB, indicating a monotonic increase in non-linear scattering with added mass. Including η in the feature vector raised the RF recall for distinguishing 16 g from 32 g from 76 % to 82 %, demonstrating that interface induced non-linearity provides complementary diagnostic information beyond the linear envelope attenuation.

Four classifiers were benchmarked: kernel SVM with RBF kernel (multiclass ECOC), RF (200 trees), AdaBoostM2 (400 shallow trees), and a hierarchical design comprising a cost sensitive Any-Mass vs. None screener (AdaBoostM1, threshold selected at (False-Positive Rate) FPR ≤ 5%) followed by a 16 g/32 g separator (SVM-RBF). Confusion matrices (Fig. 7) show distinct behaviors.

- SVM-RBF: high recall for the pristine class (95.2%), but poor separation of 16 g and 32 g (33.3% and 47.6% recall, respectively).
- RF: more balanced performance (recall: None 100%, 16 g 81.0%, 32 g 47.6%), yet the heavy mass class remains under detected.
- Hierarchical: the screener eliminates false "None" at the chosen operating point (precision for None = 100%), but the 16 g/32 g split remains modest (recall 38.1% and 57.1%).
- AdaBoostM2: superior multiclass behavior with recalls of 95.2% (None), 90.5% (16 g), and 71.4% (32 g) and precisions of 100%, 76.0%, and 83.3%, respectively. Overall test accuracy reaches ≈ 85.7%, the highest among the models, and macro-F1 is likewise maximal.

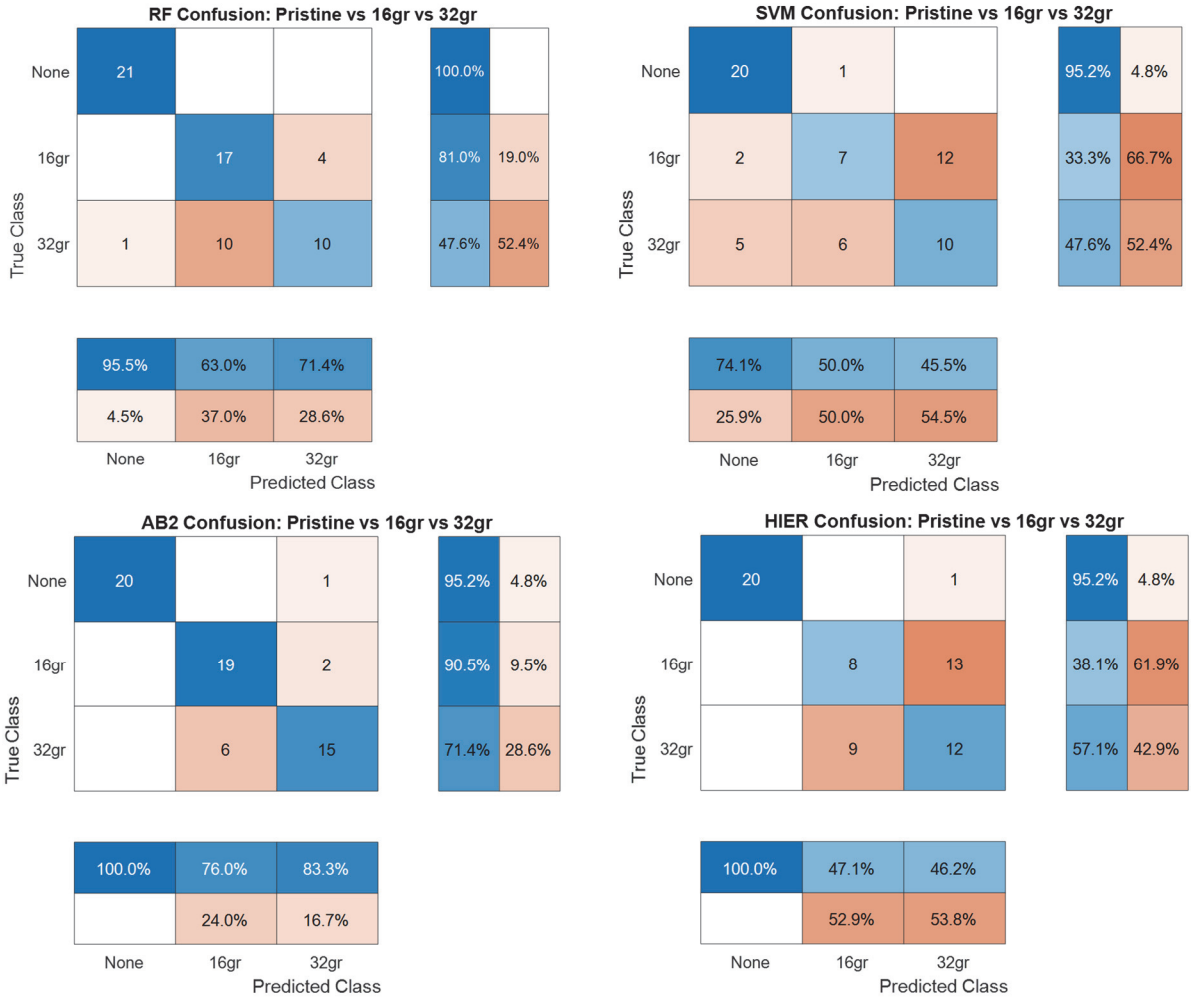


Figure 7: Confusion matrices of candidate classifiers.

Fig. 8 consolidates the comparative behavior of the four candidates. Ensemble methods outperform the kernel SVM for severity discrimination: AdaBoostM2 provides the most favorable trade off between accuracy and class balance (macro-F1), with strong recalls at both 16 g and 32 g. RF preserves excellent detection of the pristine class and acceptable performance

at 16 g, but under detects 32 g. The hierarchical architecture achieves its design objective, very low false "None" at a fixed operating point, yet its second stage does not fully separate 16 g and 32 g. These trends are consistent with the feature physics (amplitude and band energy dominated measures), whereby the two mass levels partially overlap; additive boosting captures weak non-linear interactions that increase separability.

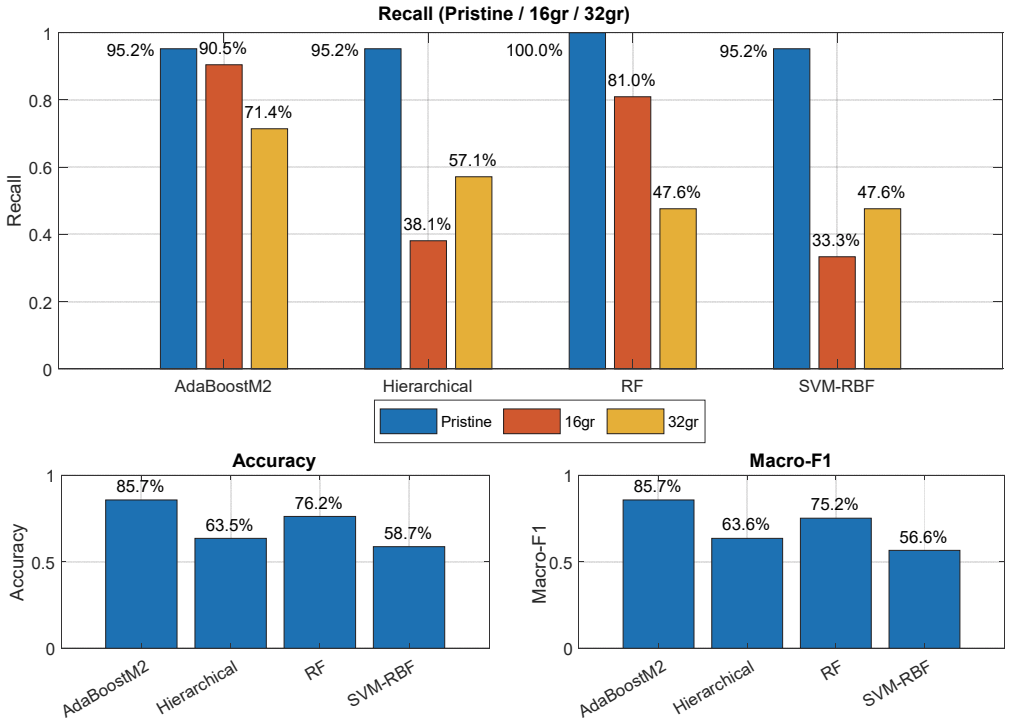


Figure 8: Model comparison summary.

AdaBoostM2 delivers the best overall trade off (highest overall accuracy and macro F1), improving 32 g recall while preserving strong pristine detection, as shown in Tab. 2. Random Forest performs well for pristine and 16 g but under detects 32 g, consistent with residual overlap between adjacent severities; the SVM with RBF kernel is the weakest baseline, and the hierarchical design attains high pristine recall but limited separation between 16 g and 32 g. These trends support a calibrated any mass gate followed by AdaBoostM2 as the severity classifier.

Classifier	Overall accuracy (%)	Macro-F1	Recall Pristine (0 g) (%)	Recall 16 g (%)	Recall 32 g (%)
SVM with RBF kernel	58.7	0.566	95.2	33.3	47.6
Random Forest (200 trees)	76.2	0.752	100.0	81.0	47.6
AdaBoostM2 (400 trees)	85.7	0.857	95.2	90.5	71.4
Hierarchical screener	63.5	0.636	95.2	38.1	57.1

Table 2: Classifier performance on the stratified held-out test set.

According to Fig. 9, the plot compares, for 5000 MC realizations per class, the probability density of the analytic envelope peak measured in the 0.55–0.75 *ms* window (first Lamb wave arrival). Pristine panels (blue) form a narrow, quasi Gaussian distribution centered at ≈ 0.40 V with a 95 % prediction band of roughly 0.37–0.43 V, reflecting the high repeatability of an undisturbed plate. Adding a 16 g mass (orange) shifts the mode of the PDF to ≈ 0.44 V and broadens it, because random stiffness and time of flight perturbations interact with increased scattering at the mass–plate interface. The 32 g population (red) is displaced further to the right and remains broader, peaking near ≈ 0.46 V. Although the pristine and loaded PDFs are clearly separated, confirming reliable damage detection, the 16 g and 32 g curves overlap by about 30 % of their area, quantitatively explaining the 24 % misclassification rate observed in the RF confusion matrix. These PDFs therefore provide a physics consistent, probabilistic rationale for both the high true negative rate on pristine windows and the residual

ambiguity between light and heavy masses, and they furnish prior distributions that can be propagated into Bayesian RUL (Remaining Useful Life) estimators.

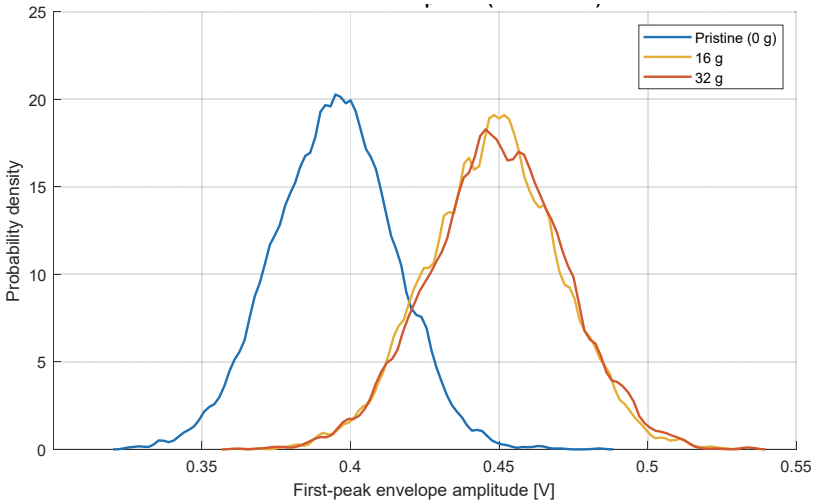


Figure 9: PDFs of first-arrival envelope amplitude.

The physics informed features and the Monte Carlo wrapper extend to carbon-epoxy laminates and to curved or stiffened panels by updating material constants and dispersion priors in the FE benchmark, choosing windows around A0 and S0 group velocities, and keeping at least two wavelengths between PWAS along the shortest path. The next campaign should include service representative damage such as through cracks, edge cracks at fastener holes, impact induced delaminations, matrix cracking, adhesive disbonds at skin stiffener interfaces, loose or fretting fasteners, and localized corrosion or thickness loss on metallic skins. Validation should progress from flat coupons to riveted or bonded subcomponents and then to a cylindrical segment, with temperature sweeps to test the stability of timing, amplitude, and the second harmonic index. Classifier scores can be converted into hit or miss counts to build probability of detection curves and sizing error distributions, with thresholds tuned to meet a probability of detection of 90 at 95 percent confidence while controlling false alarms. Reliability evidence should include group aware splits across specimens, external blinded tests, and repeatability and reproducibility studies across sensors, plates, and bonding procedures, defining a clear path from laboratory proof of concept to certification ready SHM on composite and metallic aerostructures.

Conclusions

This study demonstrates a transparent and uncertainty aware SHM pipeline that fuses guided wave physics, a finite element baseline, and supervised learning to classify damage severity on lightweight panels. The contributions are a physics informed feature set that includes a second harmonic index with clear mechanical meaning, a signal domain Monte Carlo procedure that yields prediction bands and calibrated decision scores, and a comparative evaluation of complementary learners under a common split and feature set. The framework separates healthy from mass loaded states and makes explicit the remaining ambiguity between close severities, providing a direct link to reliability metrics used in aerospace certification such as probability of detection.

This is a laboratory proof of concept on a single specimen and three controlled conditions. Broader generalization will require multi specimen datasets, additional damage types and locations, environmental sweeps, and external blinded validation. With these steps, the proposed any mass gate followed by a calibrated severity classifier offers a practical path toward certification ready SHM on composite and metallic aerostructures.

Funding

CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico)for funding his research, grant no. 317388/2021-5.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

REFERENCES

[1] Worden, K. and Farrar, C. R. (2007). An introduction to structural health monitoring. Structural Safety, 29(4), pp. 392–393. DOI: https://doi.org/10.1016/j.strusafe.2007.03.002.
[2] Giurgiutiu, V. (2005). Structural health monitoring with piezoelectric wafer active sensors. Academic Press. DOI: https://doi.org/10.1016/B978-012088905-3/50003-4.
[3] Farrar, C. R. and Worden, K. (2012). Structural health monitoring: A machine learning perspective. John Wiley & Sons. DOI: https://doi.org/10.1002/9781118443204.
[4] Azimi, M., Zhang, Y. and Eslami, M. (2020). Data-driven structural health monitoring and damage detection through deep learning: A state-of-the-art review. Sensors, 20(8), 2148. DOI: https://doi.org/10.3390/s20082148.
[5] Cawley, P. (2021). A development strategy for structural health monitoring applications. ASME Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 4(4), 041012. DOI: https://doi.org/10.1115/1.4051974.
[6] Nath, S. K. (2021). Estimates of probability of detection and sizing of flaws in ultrasonic time-of-flight diffraction inspections for complex geometry components with grooved surfaces. ASME Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 4(2), 021003. DOI: https://doi.org/10.1115/1.4048381.
[7] Zhang, Z., Saito, O. and Okabe, Y. (2024). Analysis on power flow of Lamb waves generated by laser illumination in carbon fiber-reinforced plastic laminates. ASME Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 7(4), 041006. DOI: https://doi.org/10.1115/1.4065864
[8] Hespeler, S. C., Nemati, H., Masurkar, N., Alvidrez, F., Marvi, H. and Dehghan-Niri, E. (2024). Deep learning–based time-series classification for robotic inspection of pipe condition using non-contact ultrasonic testing. Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 7(1), 011002. DOI: https://doi.org/10.1115/1.4063694.
[9] Ennis, S. and Giurgiutiu, V. (2024). Artificial intelligence–enabled crack length estimation from acoustic emission signal signatures. Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 7(1), 011004. DOI: https://doi.org/10.1115/1.4064011.
[10] Yunker, A., Lake, R., Kettimuthu, R. and Kral, Z. (2024). Evaluating model robustness for defect identification and classification in a composite aerostructure material. Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 8(1), 011001. DOI: https://doi.org/10.1115/1.4065474.
[11] Rose, J. L. (1999). Ultrasonic waves in solid media. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511623740.
[12] Alleyne, D. N. and Cawley, P. (1992). Optimization of Lamb wave inspection techniques. NDT & E International, 25(1), pp. 11–22. DOI: https://doi.org/10.1016/0963-8695(92)90003-Y.
[13] Norris, A. N. and Vemula, C. (1995). Scattering of flexural waves on thin plates. Journal of Sound and Vibration, 181(1), pp. 115–125. DOI: https://doi.org/10.1006/jsvi.1995.0124.
[14] Climente, A., Norris, A. N. and Sánchez-Dehesa, J. (2015). Scattering of flexural waves from a hole in a thin plate with an internal beam. The Journal of the Acoustical Society of America, 137(1), pp. 293–302. DOI: https://doi.org/10.1121/1.4904551.
[15] Boashash, B. (1992a). Estimating and interpreting the instantaneous frequency of a signal. Part 1: Fundamentals. Proceedings of the IEEE, 80(4), pp. 520–538. DOI: https://doi.org/10.1109/5.135376.

[16] Boashash, B. (1992b). Estimating and interpreting the instantaneous frequency of a signal. Part 2: Algorithms and applications. Proceedings of the IEEE, 80(4), pp. 539–571. DOI: https://doi.org/10.1109/5.135378.
[17] Giurgiutiu, V. (2005). Tuned Lamb wave excitation and detection with PWAS for SHM. Journal of Intelligent Material Systems and Structures, 16(4), pp. 291–306. DOI: https://doi.org/10.1177/1045389X05050106.
[18] Marple, S. L. (1999). Computing the discrete-time "analytic" signal via FFT. IEEE Transactions on Signal Processing, 47(9), pp. 2600–2603. DOI: https://doi.org/10.1109/78.782222.
[19] Hong, J.-C., Sun, K.-H. and Kim, Y.-Y. (2005). Dispersion-based STFT applied to dispersive wave analysis. The Journal of the Acoustical Society of America, 117(5), pp. 2949–2960. DOI: https://doi.org/10.1121/1.1893265.
[20] Joint Committee for Guides in Metrology. (2008). JCGM 101: Evaluation of measurement data — Supplement 1 to the GUM — Propagation of distributions using a Monte Carlo method. BIPM.
DOI: https://doi.org/10.59161/jcgm101-2008
[21] Brazalez, J. and Nabarrete, A. (2025). Probabilistic approach of uncertainties for structural dynamics of pressurized composite cylinders. Paper presented at the XX International Symposium on Dynamic Problems of Mechanics (DINAME 2025), Águas de Lindóia, Brazil. DOI: https://doi.org/10.26678/ABCM.DINAME2025.DIN2025-0216.
[22] Raghavan, A. and Cesnik, C. E. S. (2007). Finite-element modeling of ultrasonic guided waves in composite materials. AIAA Journal, 45(12), pp. 2835–2843. DOI: https://doi.org/10.2514/1.28949.
[23] Hughes, T. J. R. (2012). The finite element method: Linear static and dynamic finite element analysis. Dover. ISBN 978-0486411811.
[24] Craig, R. R., Jr. and Kurdila, A. J. (2006). Fundamentals of structural dynamics (2nd ed.). Wiley. ISBN 978-0471430445.
[25] Lysmer, J. and Kuhlemeyer, R. L. (1969). Finite dynamic model for infinite media. Journal of Engineering Mechanics Division, ASCE, 95(EM4), pp. 859–877.
[26] Deng, F. and Michaels, J. E. (2005). Second-harmonic generation of Lamb waves at solid–solid contacts. Ultrasonics, 43(7), pp. 549-556. DOI: https://doi.org/10.1016/j.ultras.2005.03.002.