TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Moradi, M., & Zarouchas, D. (2026). Multi-Sensor Fusion for Aerospace Structural Health Indicators: Passive and Active Sensory Networks. In *Proceedings of the 12th European Workshop on Structural Health Monitoring (EWSHM 2026): July 7-10, 2026 in Toulouse, France* Article 1499 (e-Journal of Nondestructive Testing; Vol. 2026, No. 08). NDT.net. https://doi.org/10.58286/33832

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

More info about this article: https://www.ndt.net/?id=33832

Multi-Sensor Fusion for Aerospace Structural Health Indicators: Passive and Active Sensory Networks

Morteza MORADI[1], Dimitrios ZAROUCHAS[1]
[1]Center of Excellence in Artificial Intelligence for structures, prognostics & health management, Aerospace Engineering Faculty, Delft University of Technology, the Netherlands, M.Moradi-1@tudelft.nl

Abstract Developing reliable health indicators (HIs) for aeronautical composite structures is challenging because damage evolution is complex, stochastic, and affected by uncertainties such as impact events and manufactured defects. In this context, passive and active structural health monitoring (SHM) techniques provide complementary information: acoustic emission (AE) captures temporally dense signatures of damage activity, whereas guided waves (GW) provide state-sensitive interrogation of the structure at discrete inspection times. This work therefore investigates heterogeneous sensor fusion for HI construction in composite T-stiffener panels subjected to run-to-failure compression–compression fatigue loading. Modality-specific AE- and GW-based HI frameworks are first developed independently using signal processing and AI. The resulting HIs are then synchronized in the fatigue-cycle domain and combined through inter-modality fusion to obtain a single fused HI. The study addresses key challenges including the absence of true HI labels, synchronization of passive and active measurements, and learning from limited multimodal overlap. The results show the potential of combining AE and GW to construct more informative and robust HIs for prognostics of composite aerostructures.

Keywords: Prognostics and health management, Structural health monitoring, Health indicator, Heterogeneous sensor fusion, Semi-supervised deep neural network, Composite aerostructures

Introduction

Awareness of the health state of aeronautical composite structures is essential for safe operation, maintenance planning, and weight-efficient design. In prognostics and health management, this information is commonly summarized through a health indicator (HI), i.e., a scalar trajectory intended to reflect structural degradation from the healthy state to end of life. For composite structures, however, constructing such an HI remains difficult because damage evolves through interacting and stochastic mechanisms across scales, while the measured

This work is licensed under CCBY 4.0

Media and Publishing Partner NDT.net
https://doi.org/10.58286/33832

response depends on material architecture, manufacturing variability, loading history, and sensing modality [1]. As a result, directly measurable physical HIs are generally unavailable, and data-driven HIs must instead be designed to retain prognostic relevance.

Among structural health monitoring (SHM) techniques, passive acoustic emission (AE) and active guided waves (GW) provide complementary information on damage evolution. AE continuously records transient signals associated with damage initiation and progression, and is therefore rich in temporal information, but it does not directly quantify the instantaneous structural state. GW interrogations, in contrast, are more informative of the current condition of the structure and wave–damage interaction, yet they are available only at discrete inspection times. A single modality is therefore unlikely to capture both degradation history and current structural state with equal fidelity. This motivates multimodal HI construction, where AE contributes temporal sensitivity and GW contributes state-sensitive active interrogation [2, 3].

Because true HI labels are unavailable, HI quality is assessed through prognostic criteria. Monotonicity (Mo) measures whether the HI evolves consistently in one direction over life, prognosability (Pr) measures the consistency of HI values at end of life across units, and trendability (Tr) measures the similarity of degradation trends among units; their sum defines the Fitness score. In PHM, these criteria should be evaluated over all run-to-failure units available in the dataset, regardless of whether they belong to the training, validation, or test split; these scores are reported as "-All". However, cohort-level scores may still obscure poor generalization to a held-out unit. For this reason, "-Test" scores are additionally reported based on the more rigorous test-oriented definitions introduced in [4].

This paper presents a heterogeneous multisensor late-fusion framework for constructing HIs for composite T-stiffener panels subjected to run-to-failure compression–compression (C-C) fatigue loading, including representative uncertainty sources such as impact damage and manufactured disbonds. Modality-specific AE- and GW-based HI frameworks are first developed independently using all data available for each sensing modality [4, 5]. Within each framework, intra-modality fusion is used to combine complementary information and improve the robustness of the unimodal HI. The resulting AE- and GW-based HIs are then synchronized and combined through inter-modality fusion. The aim is to determine whether integrating passive and active sensory networks improves HI robustness and prognostic quality, particularly on unseen units, relative to the corresponding unimodal indicators.

1. Experimental Setup and Data Partitioning

Since the detailed unimodal AE- and GW-based HI frameworks have been reported in [5, 4], this section only outlines the experimental campaign and multimodal setup adopted for fusion. The data were obtained from run-to-failure C–C fatigue tests on composite T-stiffener panels within the H2020 ReMAP project [6]. The panels featured an Embraer-inspired design with an IM7/8552 carbon/epoxy skin and a single T-stiffener. Fatigue tests were conducted on a hydraulic MTS machine at 2 Hz with a load ratio of $R = 10$. To create representative damage scenarios, 10 J impacts were introduced at several locations, and some specimens also contained manufactured disbond defects.

Structural degradation was monitored using both AE and GW systems, as illustrated in Fig. 1. AE data were acquired continuously during fatigue loading using four VS900-M broadband sensors covering 100–900 kHz. GW measurements were collected every 5000 cycles using eight surface-bonded piezoelectric transducers (PZTs), with each transducer acting in turn as the actuator and the others as receivers. Excitation frequencies of 50, 100, 125, 150, 200, and 250 kHz were used. Additional details of the sensing setups are given in [4, 5].

A total of 14 panels were tested to failure, with detailed information provided in Ta-

This work is licensed under CCBY 4.0

Media and Publishing Partner NDTnet

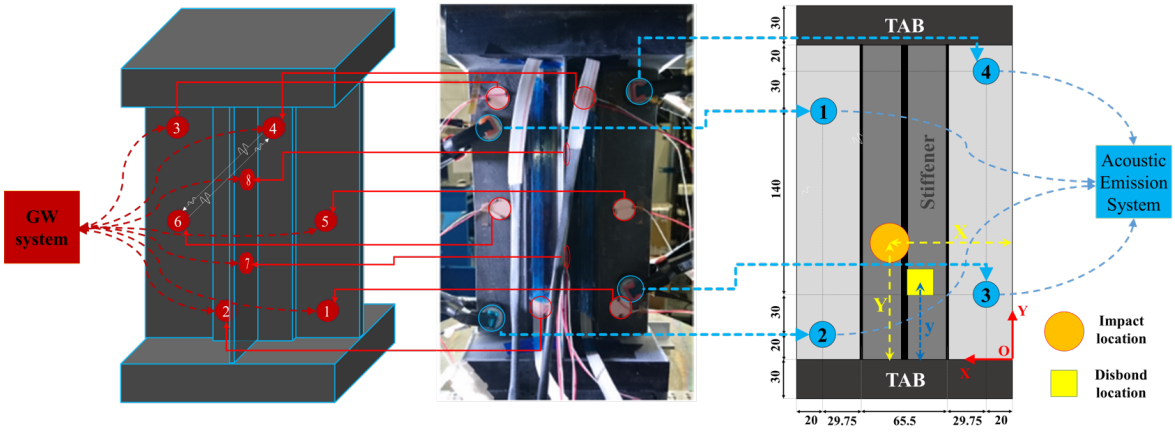

Fig. 1. *Instrumentation of the composite T-stiffener panel with guided-wave (GW) and acoustic emission (AE) sensors, together with the representative impact and disbond locations.*

ble 1. Valid AE data were available for 12 units, GW data for 5 units, and only 3 units were instrumented with both sensing systems. Because of this limited overlap, direct multimodal learning from raw signals is not feasible. Instead, a distributed late-fusion architecture is adopted, as shown in Fig. 2. First, modality-specific AE- and GW-based HI models are pre-trained independently using the full training datasets available for each sensing modality, with the overlapping units included in their respective unimodal cohorts. The resulting unimodal models are then transferred to the overlapping units to provide AE- and GW-based HI sequences, which are subsequently aligned on the cyclic fatigue timeline, resampled, and used as paired inputs for inter-modality fusion to estimate a single fused HI.

Table 1. Summary of the composite T-stiffener panels tested under C–C fatigue loading. The index column is given as AE/GW/Fusion; "–" indicates that the corresponding data were not available.

Unit	**Index** (AE/GW/Fusion)	**Impact** (x, y, cycle)	**Disbond**	**Load (min/max)** $[kN]$	**Lifetime** (cycles)
L1-03	1/1/1	$(50, 80, 0)$	None	$-6.5/-65$	152,458
L1-04	–/2/–	$(25, 80, 0)$	None	$-6.5/-65$	280,098
L1-05	2/3/2	$(115, 160, 0)$	None	$-6.5/-65$	144,969
L1-09	3/4/3	$(82.5, 140, 0)$	None	$-6.5/-65$	133,281
L1-23	–/5/–	None	30×30 @ $y = 60$	$-5/-50$, $-6/-60$	438,000
L1-49	4/–/–	$(50, 160, 5000)$	15×20 @ $y = 60$	$-6.5/-65$	48,702
L1-50	5/–/–	$(50, 160, 5000)$	20×20 @ $y = 60$	$-6.5/-65$	65,500
L1-51	6/–/–	$(50, 160, 5000)$	20×25 @ $y = 60$	$-6.5/-65$	94,431
L1-52	7/–/–	$(50, 160, 0)$	None	$-6.5/-65$	368,558
L1-54	8/–/–	$(50, 160, 0)$	None	$-6.5/-65$	510,961
L1-55	9/–/–	$(115, 80, 5000)$	None	$-6.5/-65$	226,356
L1-56	10/–/–	$(50, 160, 5000)$	None	$-6.5/-65$	756,226
L1-59	11/–/–	$(50, 160, 5000)$	None	$-6.5/-65$	110,137
L1-60	12/–/–	$(115, 80, 5000)$	None	$-6.5/-65$	170,884

Across all stages, model development follows an inductive semi-supervised setting in which true HI labels are unavailable. Instead, proxy targets and prognostic criteria are used to guide training, while validation on held-out units is used to prevent overfitting to the proxy scale [2]. To improve robustness, each base model is trained from multiple random initializations and the resulting predictions are aggregated through lightweight ensemble models. All evaluations are performed with strictly unit-wise splits to avoid leakage across specimens, while the fusion stage is assessed on the three overlapping units in a leave-one-unit-out manner.

This work is licensed under CCBY 4.0

Media and Publishing Partner NDTnet

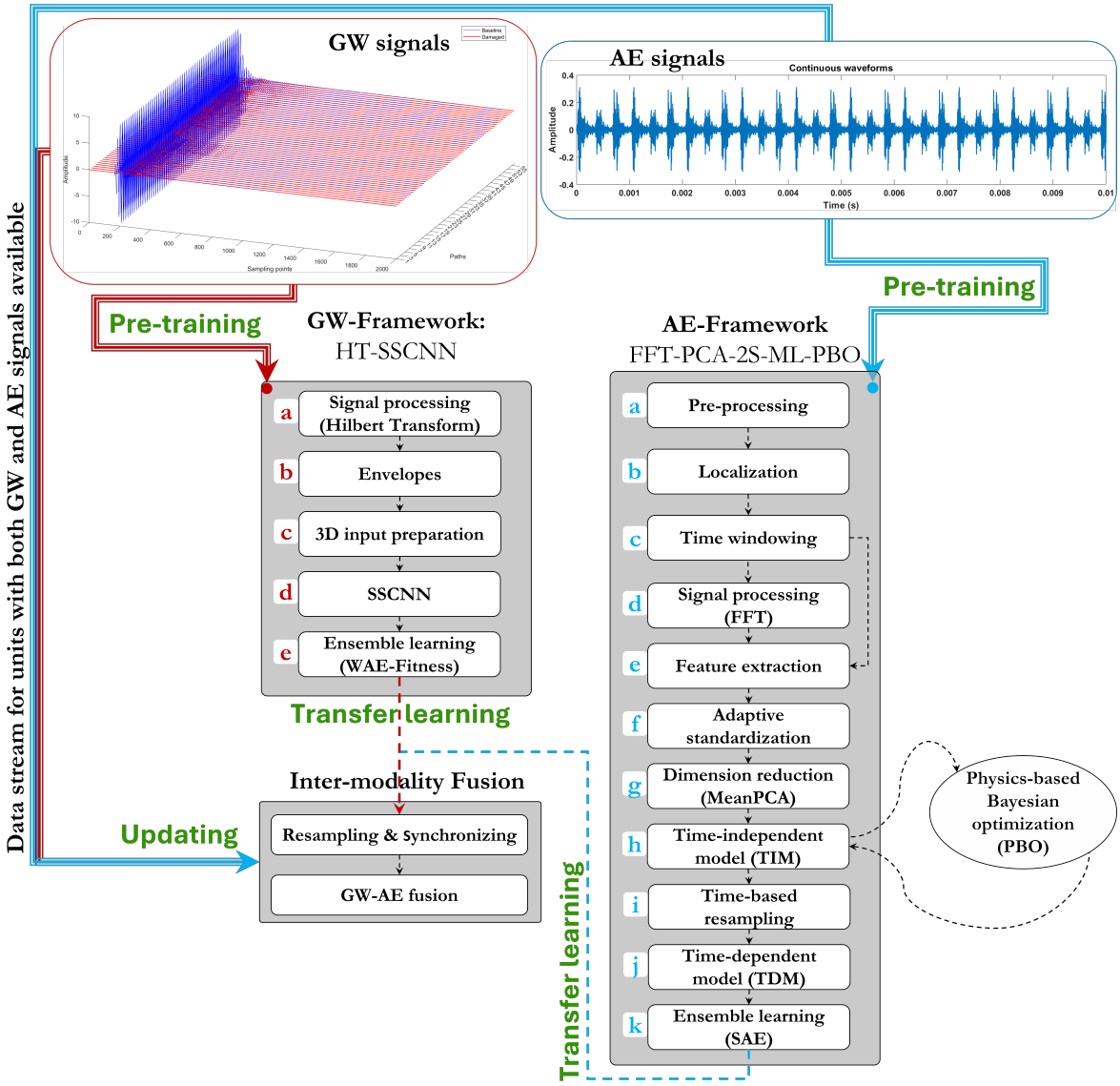

Fig. 2. *Inter-modality late-fusion framework in which pretrained AE- and GW-based health indicators are synchronized and fused into a single indicator.*

2. AE-based HI Development

The AE-based HI is constructed using the framework described in [5]. The workflow combines frequency-domain feature extraction, adaptive standardization, dimensionality reduction, and a two-stage learning strategy to derive an HI from continuously acquired AE data during fatigue loading. AE hits are first pre-processed and localized, and events with large position uncertainty are discarded. To handle the non-stationary nature of fatigue monitoring, the remaining events are grouped into consecutive time windows, from which standard AE descriptors and FFT-based spectral features are extracted. Because the statistical properties of AE signals evolve during degradation, feature scaling is performed online on a per-unit basis, using only the current and past measurements of that unit. This avoids data leakage from future observations and improves robustness when monitoring previously unseen specimens.

The resulting high-dimensional feature space of 201 is then reduced to 10 features using principal component analysis (PCA). In particular, four variants were examined: ConcatPCA, MeanPCA, FullPCA, and RPCA [5]. MeanPCA, in which eigenvectors are obtained at the unit level and then averaged across training units, was retained because it provided the

This work is licensed under CCBY 4.0

Media and Publishing Partner

most stable representation across specimens while keeping the downstream model compact. The reduced representation is passed to a time-independent model (TIM), which maps the window-level AE features to a first-level HI. Since true HI labels are unavailable, training relies on physics-inspired proxy targets together with a validation-driven optimization procedure that favors desirable prognostic behavior. A second, time-dependent model (TDM) is then used to refine the first-stage output by learning temporal dependencies along the degradation trajectory. To accommodate the variable run-to-failure lengths of different units, the first-level HI sequences are resampled on a normalized lifetime scale before temporal modeling. This second stage improves smoothness and consistency by restoring the sequential information that is not explicitly captured by the TIM. To reduce sensitivity to stochastic initialization, the AE framework is trained multiple times and the resulting predictions are combined through a lightweight ensemble model. The final AE-based HI is therefore more stable and better aligned with the expected monotonic degradation trend, making it suitable for subsequent fusion with the GW-based HI.

3. GW-based HI Development

The GW branch follows the history-independent framework introduced in [4], where each active interrogation is used to estimate the current structural state of the composite panel. In contrast to AE, this branch does not rely on continuous monitoring, but instead converts each inspection snapshot into a scalar HI that reflects the condition of the structure at that cycle. For each acquisition, the measured waveforms are processed using the Hilbert transform (HT) to extract upper and lower signal envelopes. For every actuator-sensor path, the envelopes from the baseline and current states are stacked into a three-dimensional input tensor, preserving the multi-path GW information in a structured format suitable for convolutional learning. A convolutional neural network (CNN) is then used to map the input tensor directly to a GW-based HI. As in the AE framework, training is performed in an inductive semi-supervised setting because ground-truth HI labels are not available. Proxy targets are therefore used during learning, while a mandatory validation subset and prognostic criteria are used to guide model selection.

To improve robustness, the SSCNN is trained multiple times with different random initializations, and the resulting HI predictions are combined through ensemble learning. The GW-HI retained here is obtained by weighted averaging, where larger weights are assigned to models with better fitness score [4]. Multi-frequency information is incorporated at the same stage by combining the HI predictions obtained from the different excitation frequencies, with each frequency weighted according to its performance. This frequency-level fusion improves robustness because different excitation frequencies are sensitive to different wave–damage interactions and may not be equally informative throughout the entire degradation process. An additional advantage is that it removes the need to preselect a single “best” interrogation frequency. This framework transforms high-dimensional GW inspections into a compact HI while preserving the spatial sensitivity of the active sensing network. The resulting GW-based HI is particularly useful for capturing the structural state at discrete inspection times and complements the temporally dense AE-based HI used in the fusion stage.

4. Fusion of AE- and GW-based HIs

The overall intra- and inter-modality workflow is summarized in Fig. 2. Fusion is performed at the HI level rather than directly on the raw sensor data. This choice is motivated not only by the limited number of units instrumented with both AE and GW systems, but also by the different

This work is licensed under CCBY 4.0

Media and Publishing Partner NDTnet

acquisition regimes and data structures of the two modalities. AE yields dense windowed sequences from continuous monitoring, whereas GW provides sparse inspection snapshots derived from multi-path wave responses. Direct data-level fusion would therefore require strong alignment and representation-harmonization assumptions, increase model complexity, and restrict training to the small overlapping subset. By contrast, HI-level fusion allows the unimodal AE and GW models to be developed independently using all available data for each modality, after which their HI outputs are aligned and combined on the overlapping units. This strategy also facilitates a clear comparison between individual SHM techniques and their fusion, making it possible to assess whether the multimodal combination provides a meaningful improvement over the unimodal indicators.

Before fusion, the AE- and GW-based HIs are aligned on a common fatigue-cycle timeline, with the denser AE sequence used as the temporal reference. Because GW inspections are available only at discrete intervals, the GW-based HI is mapped onto the AE timeline using a causal zero-order-hold scheme, whereby each AE window is assigned the most recent GW estimate available at or before that cycle. This preserves causality, avoids the use of future information, and is consistent with an online deployment scenario.

After synchronization, the AE- and GW-based HIs are used as paired inputs to four regression models, namely Gaussian process regression (GPR), support vector regression (SVM), multilayer perceptron (MLP), and long short-term memory network (LSTM). Fusion performance is evaluated in a leave-one-unit-out manner over the three units instrumented with both sensing systems. In each fold, two units are used for model retraining, and the remaining unit is reserved exclusively for testing. Unless otherwise specified, Bayesian optimization (BO) is used to adjust the hyperparameters in 5-fold cross-validation applied to data of the two training units. The concatenated measurements of these two units are divided and shuffled to create the five folds. BO uses mean squared error (MSE) as the loss between predictions and simulated labels, with an exploration ratio of 0.5, four starting points, and a maximum of 30 trials. The calibrated hyperparameters and model configurations for each fusion model are presented below.

- GPR is implemented using a squared-exponential kernel with a constant basis function. Within the BO framework, the noise standard deviation is optimized and the best value is found to be $\sigma = 0.6750$. This model provides a flexible nonparametric baseline for HI fusion while remaining suitable for the limited amount of synchronized multimodal data.
- SVM uses a linear kernel with $\epsilon = 3.727$ and is trained by sequential minimal optimization. The linear kernel is retained to provide a compact and well-regularized mapping from the paired AE- and GW-based HIs to the fused HI.
- MLP is used as a nonlinear pointwise fusion model (Table 2) and trained with Levenberg–Marquardt backpropagation and MSE loss, with 30% of synchronized samples from the two training units reserved for validation.
- LSTM is used to capture temporal dependencies in the synchronized HI sequences (see Table 2 for the architecture). The network is trained using Adam with an initial learning rate of 0.1 and L2 regularization with weight decay 0.001. A mini-batch size of 20, a maximum of 2000 epochs, validation every 30 iterations, and early stopping with a patience of 50 validation checks are used, and the model is selected based on the best validation loss.

5. Results and Discussion

The proposed fusion framework was evaluated on the three units instrumented with both AE and GW sensing using a leave-one-unit-out protocol. In each fold, one unit was held out for testing and the remaining two were used for training. Figure 3 compares the unimodal and fusion-based HIs, while Table 3 summarizes the corresponding Mo, Pr, Tr, and Fitness scores.

This work is licensed under CCBY 4.0

Media and Publishing Partner NDTnet

Table 2. Architecture of the neural network fusion models.

Fusion model	Layer type	Layer size	Activation function
MLP	Input layer	2	–
	Fully connected layer 1	5	Log-sigmoid
	Fully connected layer 2	3	Log-sigmoid
	Fully connected layer 3	1	Log-sigmoid
	Output layer	1	Linear
LSTM	Sequence input layer	2	–
	LSTM layer	10	LeakyReLU (0.01)
	Fully connected layer 1	3	LeakyReLU (0.01)
	Fully connected layer 2	1	Swish
	Output layer	1	Linear

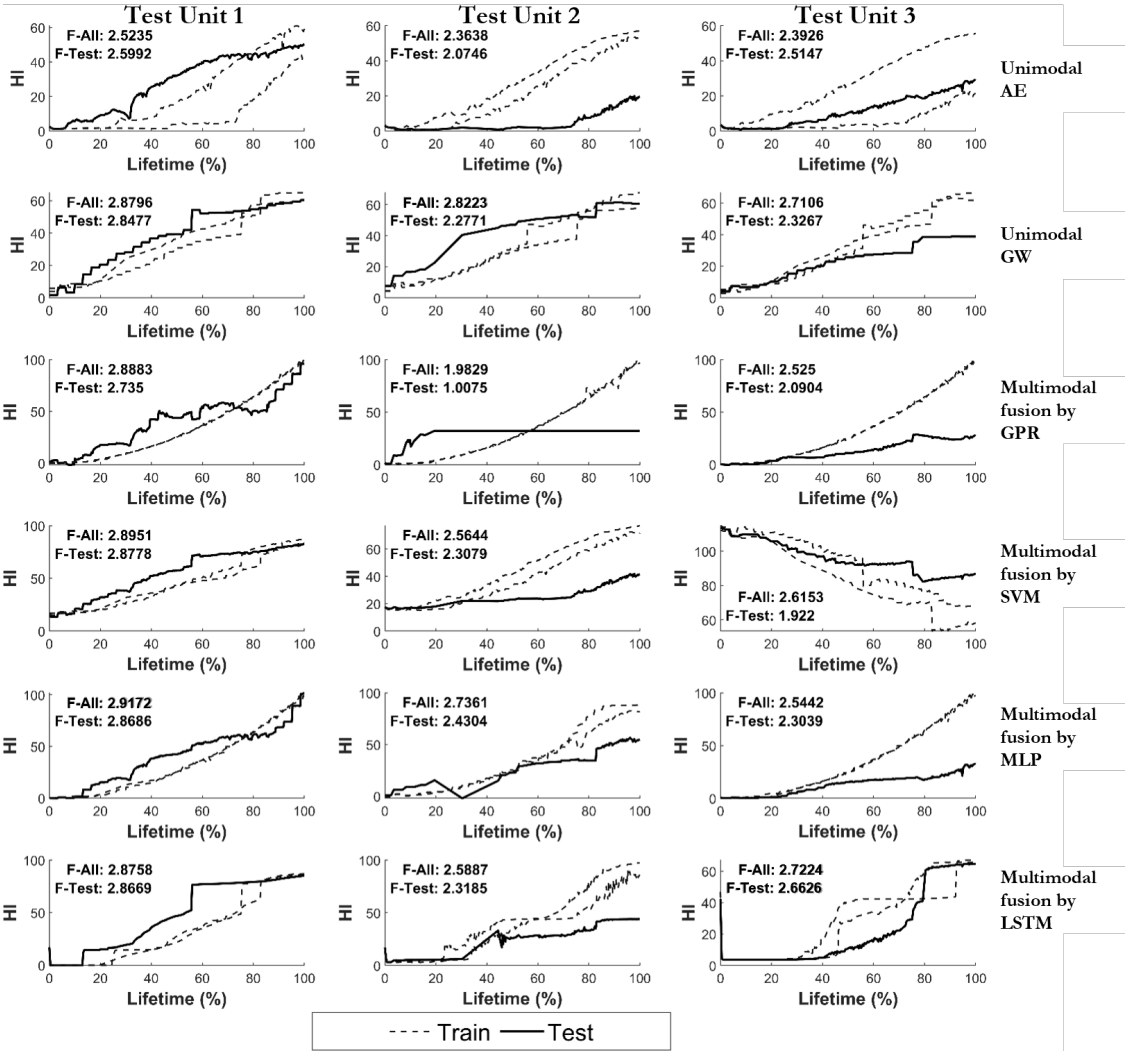

Fig. 3. *HIs obtained for the unimodal (AE or GW) and multimodal fusion models under leave-one-unit-out evaluation, where columns correspond to the held-out test units. Fitness scores calculated for all units (F-All) and for the test unit only (F-Test) are reported.*

As a standalone indicator, the GW-based HI outperformed the AE-based HI, with average Fitness values of 2.80 (out of a maximum of 3.00) versus 2.43 for all units and 2.48 versus 2.40 for the held-out units. This likely reflects the strong state sensitivity of active interrogation, further supported by the denser GW sensing network (eight PZTs versus four AE sensors). However, GW monitoring remains intermittent, since measurements are acquired only at scheduled inspection intervals and require temporary interruption of fatigue loading.

Among the fusion models, the neural network fusion models were more competitive:

Table 3. HIs' criteria scores computed for all units ("-All") and for the test units only ("-Test"), reported across the three folds as mean values (±standard deviation).

Criteria	Unimodal models (after resampling)		Multimodal fusion models			
	AE	GW	GPR	SVM	MLP	LSTM
Mo-All	1(±0)	1(±0)	0.98(±0.02)	0.78(±0.30)	0.91(±0.08)	0.95(±0.05)
Mo-Test	0.93(±0.04)	0.70(±0.28)	0.59(±0.37)	0.78(±0.30)	0.91(±0.08)	0.95(±0.05)
Pr-All	0.67(±0.16)	0.87(±0.11)	0.72(±0.23)	0.78(±0.17)	0.79(±0.19)	0.85(±0.20)
Pr-Test	0.71(±0.28)	0.85(±0.20)	0.59(±0.33)	0.68(±0.25)	0.68(±0.27)	0.79(±0.29)
Tr	0.76(±0.08)	0.94(±0.03)	0.76(±0.30)	0.92(±0.01)	0.94(±0)	0.88(±0.10)
F-All	2.43(±0.09)	**2.80(±0.09)**	2.47(±0.46)	2.69(±0.18)	2.73(±0.19)	2.73(±0.14)
F-Test	2.40(±0.28)	2.48(±0.32)	1.94(±0.87)	2.37(±0.48)	2.53(±0.30)	**2.62(±0.28)**

MLP reached F-Test $= 2.53$, while LSTM achieved the highest average test-unit Fitness at 2.62, corresponding to improvements of about 9% over AE and 6% over GW. Both MLP and LSTM reached F-All $= 2.73$, slightly below the unimodal GW result. However, LSTM provided the better balance between test performance and consistency, whereas the larger gap between F-All and F-Test for MLP suggests greater sensitivity to overfitting.

6. Conclusions

This paper presented a heterogeneous multisensor late-fusion framework for constructing health indicators for composite structures. Modality-specific frameworks were first developed independently and then combined at the HI level. The GW-based HI was the strongest unimodal indicator, while LSTM-based fusion achieved the highest test-unit Fitness, improving by about 9% over AE and 6% over GW. These results show that HI-level fusion improves robustness on unseen units by combining the complementary strengths of passive and active sensing. Future work will consider larger multimodal cohorts and uncertainty quantification.

References

[1] Peter WR Beaumont. The structural integrity of composite materials and long-life implementation of composite structures. *Applied Composite Materials*, 27(5):449–478, 2020.

[2] M. Moradi. *Designing Health Indicators for Aerospace Structures by Intelligent Information Fusion*. Dissertation (tu delft), Delft University of Technology, 2024.

[3] Sahar Hassani, Ulrike Dackermann, Mohsen Mousavi, and Jianchun Li. A systematic review of data fusion techniques for optimized structural health monitoring. *Information Fusion*, 103:102136, 2024.

[4] Morteza Moradi, Ferda C. Gul, and Dimitrios Zarouchas. A novel machine learning model to design historical-independent health indicators for composite structures. *Composites Part B: Engineering*, 275:111328, 2024.

[5] Morteza Moradi, Juan Chiachío, and Dimitrios Zarouchas. Health indicator modeling leveraging time-independent and time-dependent subtasks with adaptive standardization and physics-based bayesian optimization for aeronautical structures. *Engineering Applications of Artificial Intelligence*, 163:112767, 2026.

[6] Dimitrios Zarouchas, Agnes Broer, George Galanopoulos, Theodoros Loutas, Nazih Mechbal, and Marc Rebillat. Compression Compression fatigue tests on single stiffener aerospace structures 2, 2023.

This work is licensed under CCBY 4.0

Media and Publishing Partner NDTnet