

# Acoustic emission-based remaining useful life prognosis of aeronautical structures subjected to compressive fatigue loading



Georgios Galanopoulos<sup>a</sup>, Dimitrios Milanoski<sup>a</sup>, Nick Eleftheroglou<sup>a,b</sup>, Agnes Broer<sup>b,c</sup>, Dimitrios Zarouchas<sup>b,c</sup>, Theodoros Loutas<sup>a,\*</sup>

<sup>a</sup> Applied Mechanics Laboratory, Department of Mechanical Engineering and Aeronautics, University of Patras, Rio University Campus, 26504 Rio, Greece

<sup>b</sup> Structural Integrity and Composites Group, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629HS Delft, the Netherlands

<sup>c</sup> Center of Excellence in Artificial Intelligence for Structures, Aerospace Engineering Faculty, Delft University of Technology, the Netherlands

## ARTICLE INFO

### Keywords:

Structural health monitoring  
Composite structures  
Stiffened panels  
Acoustic emission  
Remaining useful life  
Fatigue

## ABSTRACT

An increasing interest for Structural Health Monitoring has emerged in the last decades. Acoustic emission (AE) is one of the most popular and widely studied methodologies employed for monitoring, due to its capabilities of detecting, locating and capturing the evolution of damage. Most literature so far, has employed AE for characterizing damage mechanisms and monitoring propagation, while only a few employ it for real time monitoring and even fewer for Remaining Useful Life (RUL) prognosis. In the present work, we demonstrate a methodology for leveraging AE recordings for prognostics of composite aerospace structures. Single stiffened CFRP panels are subjected to a variety of compressive fatigue loadings, while AE sensors monitor the panels' degradation in real time. Several AE features, both from the time and frequency domains, are utilized to identify features capable of capturing the degradation and used as Health Indicators for RUL prognosis. The choice of Health Indicators is predominantly made based on three prognostic attributes, i.e. monotonicity, trend and prognosability, which can overall affect the prognostic performance. RUL prediction of the panels is performed by employing two prominent machine learning algorithms, i.e. Gaussian Process Regression and Artificial Neural Networks. It is evidenced that the proposed AE-based methodology is highly capable to be utilized for RUL prediction of composite structures under variable loading conditions.

## 1. Introduction

With composite materials being used in increasingly more industries such as aeronautics and aerospace, it is important that their structural integrity is ensured. Though these materials display high strength to weight ratios, as well as improved fatigue and corrosion resistance compared to traditional metallic structures [1–3], their complex mechanical behavior, due to their anisotropic and inhomogeneous nature, make them more difficult to comprehend and predict their performance [4,5]. During fatigue, the gradual degradation of the composites is largely caused by the stiffness change and redistributions of stresses and strains [6] and if combined with pre-existing or new damage, e.g. in the form of an impact, it further affects and complicates the materials' degradation.

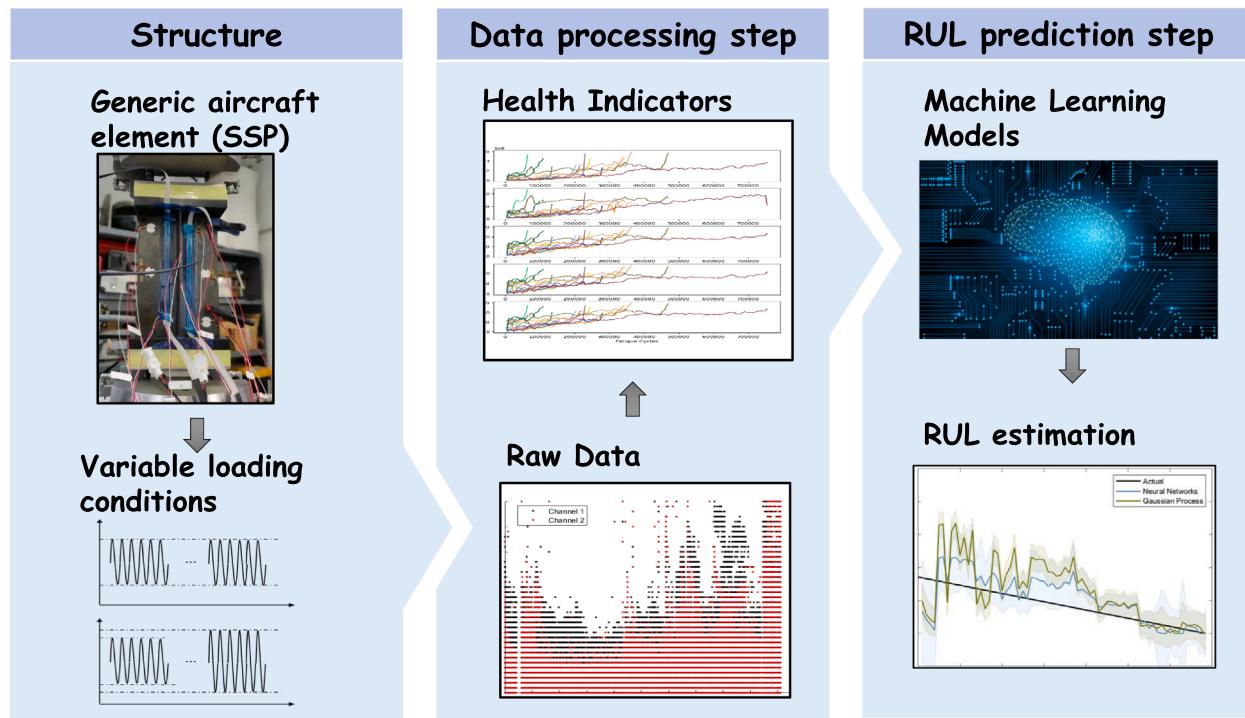
To tackle such issues, Structural Health Monitoring (SHM) has been introduced as an extension of non-destructive testing [7]. SHM can be

defined as an automated method for determining adverse changes in the integrity of mechanical systems utilizing permanently installed sensors. Methodologically, the main tasks of SHM come from the Prognostics and Health Management discipline and consist of data acquisition, data processing, identification of degradation features/ health indicator construction, fault detection diagnostics and on the final level end-of-life (EoL) prognostics and decision making [8,9]. Real-time SHM technologies are essential in assessing the integrity of composites structures [10] providing the ability to detect, locate, identify the type of damage and sometimes even quantify its extent.

Various SHM technologies for non-destructive evaluations are being studied, with Acoustic Emission (AE) being among the most popular, due to its capability to detect the progression of damage during the composite's lifetime [11]. Generated signals from the various damage mechanisms can be captured by strategically placed AE sensors, providing information on the degradation. An inherent advantage of AE

\* Corresponding author at: Applied Mechanics Laboratory, Department of Mechanical Engineering and Aeronautics, University of Patras, Rio University Campus, 26504 Rio, Greece.

E-mail address: [thloutas@upatras.gr](mailto:thloutas@upatras.gr) (T. Loutas).



**Fig. 1.** Graphical representation of the proposed methodology.

over other non-destructive methods, is the ability to distinguish the various damage mechanisms by studying the different AE features [12]. However, AE is prone to external noise requiring carefully tuned acquisition parameters as well as efficient post-processing to distinguish useful degradation information from the external noise [13].

AE has been widely employed to monitor composite specimens [14], though most researches have been focused on diagnostics. De Groot et al. [15] used peak frequency to identify different damage mechanisms in composite laminates. Four main mechanisms were identified including matrix cracking, fiber/ matrix debonding, fiber pullout and fiber cracking. Gutkin et al. [16], also used frequency to distinguish the different damage mechanisms in various different loadings, like tension, compression and other. In [17] AE was used to monitor the compressive behavior of composite coupons. Amplitude, duration and energy were used to identify the propagation of different damage mechanisms. De Oliveira et al. [11] performed SHM on GFRP laminates using AE. Artificial Neural Networks (ANN) were used to identify and classify the different damage mechanisms. Liu et al. [18], used AE on composite coupons with centrally located open-hole to monitor damage evolution and identify the failure mechanisms. Saeedifar et al. [19] studied the effect of Barely Visible Impact Damage (BVID) in composite laminates using AE. The evolution of the BVID was investigated by analyzing the AE signals from the different damage mechanisms using wavelet packet transform. Broer et al. [20], used AE sensors to identify, locate and assess the severity of damage in single stiffened panels. The AE data were also fused with strain data to provide improved diagnostics for the assessment of damage.

Though there is extensive literature on diagnostics on composites utilizing AE, very few studies have focused on prognostics. Rajendrabooopathy et al. [21], performed tensile tests on GFRP specimens aiming to eventually predict their maximum strength. A portion of the specimens was used to train an ANN from AE data at various stages during the test, while the remaining specimens were left to test the capabilities of the ANN in predicting the failure load. In [22] impacted composite laminates were subjected to tensile tests. An ANN was trained using AE data collected from different failure load percentages in order to predict the failure load of new specimens. The previous studies

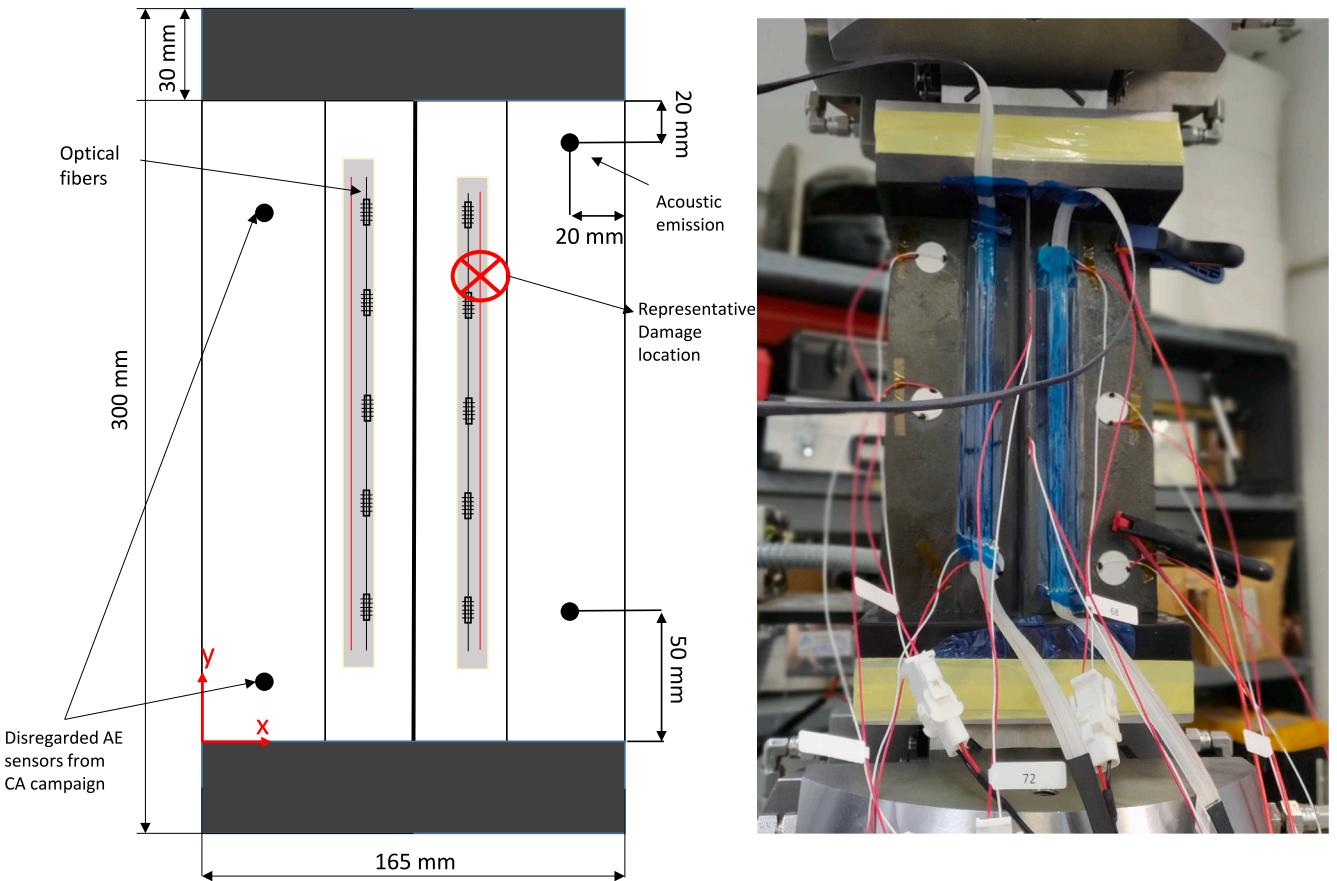
focused only on predicting failure loads. Only few researches have tackled the task of predicting the RUL. Liu et al. [23], subjected composite beams to fatigue loading and used AE to monitor the degradation. A normalized damage index is proposed as a feature to capture the degradation. A Gaussian process (GP) model was then used to predict the RUL of these composite beams. Eleftheroglou and Loutas [24] employed AE to monitor open-hole CFRP coupon subjected to tensile fatigue. The rise-time to amplitude ratio (RA) was proposed as a Health Indicator to train a nonhomogeneous hidden semi Markov model (NHHSM) to predict the RUL of the specimens. In [25] the NHHSM was compared with a Bayesian NN for the prediction of the RUL of open-hole composite coupons. The NHHSM displayed better performance compared to the Bayesian NN, with was further solidified by the existence of prediction intervals. Eleftheroglou et al. in [26] fused AE data with strain data from Digital image correlation (DIC) to monitor composite coupons. The fused feature displayed higher monotonicity than each separate monitoring technique and provided improved RUL predictions. A modified Markov model was used for RUL predictions of composite coupons which had been subjected to tension fatigue. Recently in [27] an adaptive NHHSM was developed for RUL prediction of composite coupons. The adaptive model's parameters are estimated from diagnostic features derived from both the training and testing data, and the model is able to adapt to the new training data. The model was trained with AE data collected from open-hole composite coupons tested under tension fatigue, while for evaluating the methodology, AE data collected from composite coupons subjected to fatigue loading with in-situ impacts; which are considered unseen events; were used. The model displayed high performance and was proven capable to adapt to the unexpected events.

The rather limited literature on prognostics of composite structures using acoustic emission is mostly focused on coupon-level specimens under constant loading and to the best of our knowledge no research is available for more complex structures such as stiffened panels subjected to variable amplitude fatigue loading, as this involves more complex test campaigns. In this paper, we use AE sensors to monitor the compression fatigue behavior of representative aeronautical panels, i.e., a composite single-stiffened panel, and propose a methodology to estimate the RUL

**Table 1**

Table of SSPs loadings and fatigue life.

Specimen #	Impact/disbond	disbond/impact location [x, y] [mm]	Loading type	Max Load range [kN]	# of cycles to failure
SSP 01	Impact (10 J)	[140, 80]	Constant	-65.0	280,000
SSP 02	Impact (10 J)	[50, 160]	Constant	-65.0	145,000
SSP 03	Impact (7.37 J)	[60, 75]	Variable	[-40.0, -60.0]	202,300
SSP 04	Impact (10 J)	[82.5, 140]	Constant	-65.0	133,300
SSP 05	Impact (10 J)	[105, 60]	Variable	[-40.0, -65.0]	243,000
SSP 06	Disbond ( $30 \times 20 \text{ mm}^2$ )	[82.5, 45]	Variable	[-35.0, -60.0]	345,000
SSP 07	Impact (7.37 J)	[115, 165]	Variable	[-40.0, -60.0]	242,000
SSP 08	Disbond and Impact( $20 \times 20 \text{ mm}^2$ , 10 J)	[82.5, 45] and [115, 160]	Constant	-65.0	65,500
SSP 09	Impact (10 J)	[115, 160]	Constant	-65.0	368,600
SSP 10	Impact (10 J)	[115, 160]	Constant	-65.0	511,000
SSP 11	Impact (10 J)	[50, 80]	Constant	-65.0	226,400
SSP 12	Impact (10 J)	[115, 160]	Constant	-65.0	756,200
SSP 13	Impact (10 J)	[115, 160]	Constant	-65.0	110,100

**Fig. 2.** (a) Schematic representation of the SSP and (b) SSP on the hydraulic test machine fully sensorized.

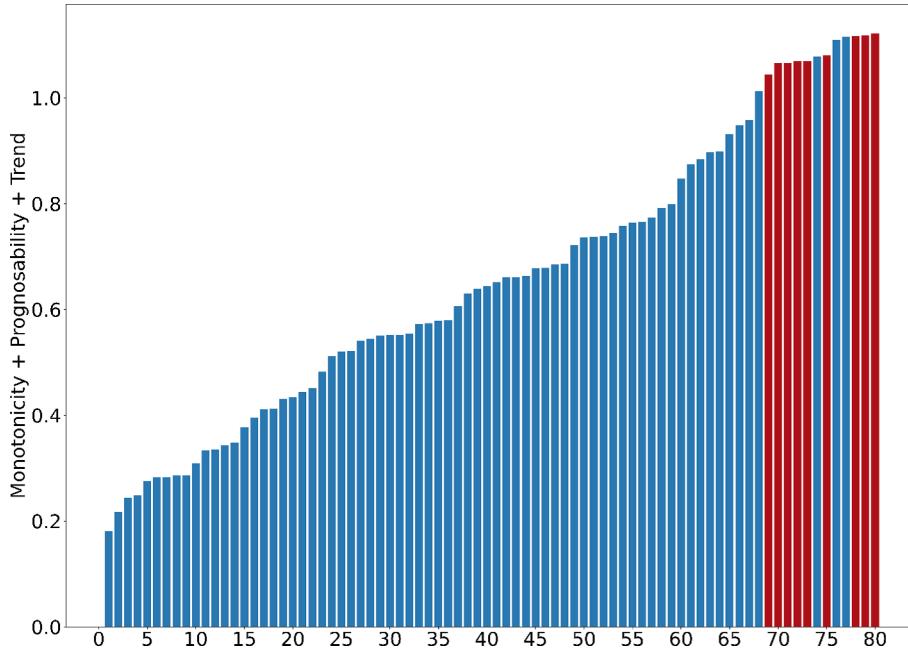
of these panels under different loading conditions. An initial damage is first introduced to the panels in the form of barely visible impact or artificial disbond, and AE is employed to monitor the panel under cyclic loading of different amplitudes. The experiments are also performed in different locations and under different conditions further increasing the complexity of the prognostic task. Health Indicators are proposed by exploring AE features in both time and frequency domains, which are able to capture the degradation and enable a more accurate RUL prognosis. The concept of the paper is graphically presented in Fig. 1.

The remainder of the paper is organized as follows. Section 2 presents the case study, comprising of the experimental campaign and the AE Health Indicator development, while Section 3 describes in short the

ML algorithms utilized for the prognostic task. Section 4 shows the main results and discusses them and finally, the paper is concluded in Section 5.

## 2. Case study

Single-stiffened composite panels (SSPs) were manufactured from IM7/8552 unidirectional pre-preg. The skin's layup is [45/-45/0/45/90/-45/0]<sub>S</sub> and the single T-shaped stiffener's layup is [45/-45/0/45/-45]<sub>S</sub>. The total length of the panels is 300 mm while the free length is 240 mm, since 30 mm resin cast tabs are placed on the edges of the panel to ensure proper and uniform load introduction. Quasi-static



**Fig. 3.** Ranked sum of monotonicity, prognosability and trend for the 82 features extracted from time and frequency domain.

compression tests were first performed to determine the collapse load, with a rate of 0.5 mm/min. The collapse load, as averaged from 3 panels, was determined at  $-100kN$  and was used to guide the fatigue load selection.

### 2.1. Experimental campaign

Two fatigue test campaigns were performed, one in the Delft University of Technology (faculty of Aerospace engineering) (TUD) and the other at the University of Patras (Department of Mechanical and Aeronautical engineering) (UPAT) consisting of constant amplitude compression fatigue and variable amplitude compression fatigue respectively. The fatigue parameters remained constant in both campaigns with a load ratio of 10 and a frequency of 2 Hz. To allow for some additional measurements, like strain measurements (DIC, FBG strain sensing) every 500 cycles quasi-static loadings were performed. These additional data are not considered in the present research.

For the constant amplitude fatigue, the maximum fatigue load was set at 65% of the collapse load, namely  $65kN$  (absolute value), while the variable amplitude fatigue load started from 35% up to 65% of the collapse load. The variable loading was applied in constant blocks and was increased arbitrarily (usually by  $5kN$ ) after inspecting the extent of the damage using a portable phased array camera [28].

An initial damage is introduced before subjecting the stiffened panels to fatigue, either via a low-energy impact (to create BVID) or by a manufacturing defect; in the form of an artificial disbond via a Teflon insert; at the interface between the skin and the stiffener's foot. The impact damage is induced by drop-tower with an approximate energy of 10 J and the goal was to create a damage at the skin/stiffener interface (usually near the  $\frac{3}{4}$  or  $\frac{1}{4}$  of the panel length), similar to the disbond. However, due to the nature of the composite materials, the response to the impact damage is not always the same, creating a wide range of different damages, increasing the complexity of our task. More information on the panels is provided in Table 1.

Among the other SHM systems, strategically placed AE sensors, are employed to monitor the fatigue life of the SSPs. An AMSY-6 Vallen acquisition system is paired with four (4) Vallen VS900-M 100–900kHz broadband sensors and an external pre-amplification of 34 dB was employed to monitor the constant amplitude fatigue, while in the other campaign; the variable amplitude in UPAT; two (2) Micro200HF

500–4500kHz wideband sensors, with a Micro-II acquisition system, from Mistras Group, and an external pre-amplification with a gain of 40 dB were used. The different number of sensors is due to hardware limitations. For consistency purposes, only the AE recordings from two out of the four sensors are considered in the upcoming study (constant amplitude campaign only). The channels that were ignored were chosen depending on the damage location, and channels located closer to the damage were selected (vertical direction only). This was also the case for positioning the two AE sensors in the UPAT campaign (at the side of the damage). The sensor positioning, with respect to the damage is shown in Fig. 2. The sensors were clamped to the back side of the skin; where the stiffener is located; and a coupling agent was applied between the sensor and the skin. It should also be noted that sensor locations are side dependent, i.e., the y location is dependent on whether the sensors are on the left or right side. Were the damage on the left side, the sensors to the right are disregarded and the sensors and positions on the left side are used.

### 2.2. Acoustic emission based health indicators

Features capable of capturing the degradation of the structure are essential in data-driven prognostics. In the present research, we refer to such features as Health indicators (HIs). The HI's quality plays a crucial role, affecting the performance and accuracy of the prognostic algorithms [29]. Some of the most sought attributes for an HI, in order to have enhanced prognostic potential, are monotonicity and prognosability [30,31]. Monotonicity refers to the HI having an increasing or decreasing behavior with damage/ time evolution. The most common mathematical expression for monotonicity is:

$$\text{Monotonicity} = \frac{\text{No.} \frac{dF}{dt} < 0 - \text{No.} \frac{dF}{dt} \geq 0}{n - 1} \quad (1)$$

where  $n$  the length of the HI time series and  $F$  an AE feature.

Prognosability is a measure of variance in the EoL values of a population of similar specimens. It is expressed by Eq. (2).  $F_{fail}$  and  $F_{start}$  the values of feature  $F$  at the failure and starting time respectively.

$$\text{Prognosability} = \exp \left( -\frac{\text{std}(F_{fail})}{\text{mean}(|F_{start} - F_{fail}|)} \right) \quad (2)$$

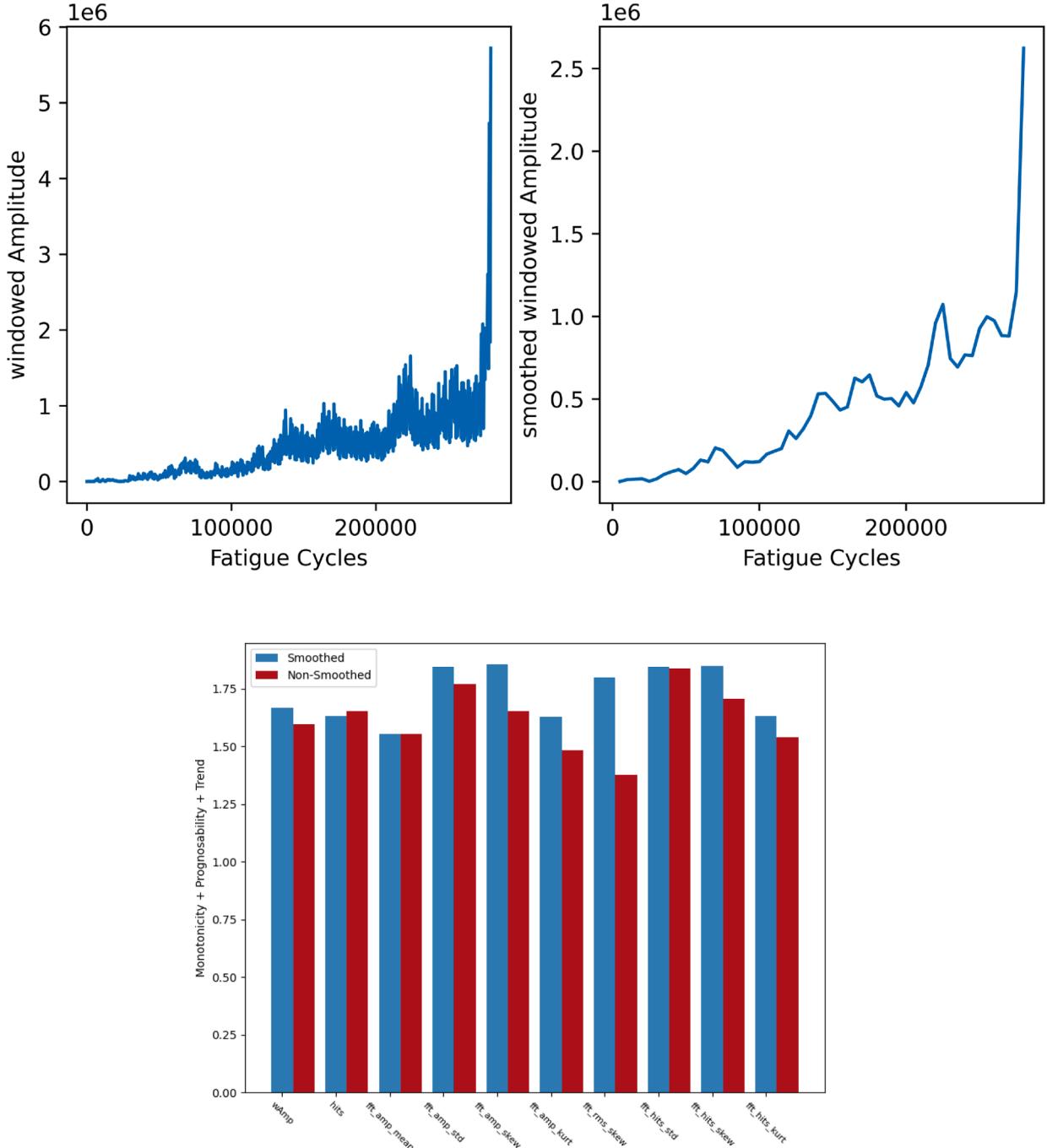


Fig. 4. (a) Original feature (left), (b) Smoothed Feature (right), (c) Improvement of monotonicity, prognosability and trend (bottom).

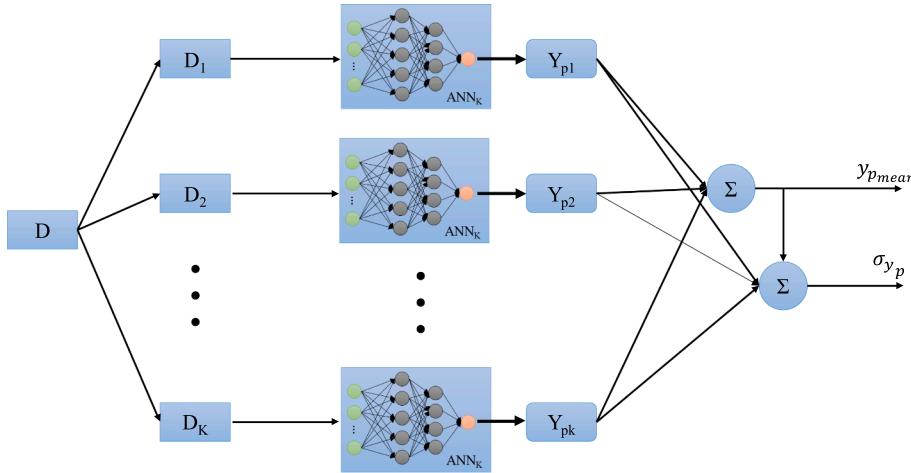
Finally, a third metric proposed in [32] is trend, which measures the temporal correlation between each feature and time. Trend is described in Eq. (3).  $F_{i,spec_j}$  is the  $i^{th}$  feature of specimen  $j$  and  $t_{spec_j}$  is the operational time of specimen  $j$ .

$$Trend(F_i) = \left| \frac{1}{N} \sum_{i=1}^N corr(F_{i,spec_j}, t_{spec_j}) \right| \quad (3)$$

AE provides an abundance of features as possible HI, such as the number of hits, hit amplitude, energy etc. The rise time to amplitude ratio (RA) has also been employed as a prognostic feature [24]. Time domain and frequency domain analysis can also provide new features for the purpose of degradation monitoring. Some common features defined in the time and frequency domains can be found in Appendix A. In this

paper we utilized 8 common AE features, namely amplitude, rise time, duration, counts to peak, RMS, energy, hits and RA as these are commonly associated with degradation. Peak frequency, which is also commonly associated with damage, were not recorded due to the limitations of the acquisition systems (recording waveforms in real-time fatigue requires a lot of processing power). Inspired by [24], these features are windowed so as to reveal a possibly hidden degradation trend. After some experimentation, a 500-cycle window was proven appropriate. Though windowing, especially for this many features and data, can be a bit intensive, for the training data it can be done offline and hence not affect the applicability of the methodology in real-time. On the other hand, new data can be processed as soon as each time window become available and only the required features can be calculated.

By discarding features whose change in consecutive time windows is



**Fig. 5.** Bootstrap methodology using ANN.

less than  $10^{-2}$ , we were left with 81 features as possible His (Appendix B). The sum of average monotonicity, prognosability and trend was measured as a quantitative metric to evaluate possible prognostic features. The selected features and the value of the summed attributes are shown in Fig. 3. Visual inspection of these features was also performed to validate the results of the feature selection. The three top-ranked features not selected through visual inspection, i.e. amplitude median, amplitude mean and amplitude RMS, had high monotonicity (due to many constant values), though trend and prognosability were lacking, and hence were not considered further. For this reason, the automatic selection procedure should be carefully set up, to avoid such setbacks.

As can be seen in Fig. 4a (also see Appendix B), the developed HIs are rather noisy throughout the test. To smoothen them and enhance the underline trend as well as reduce the training time of the prognostic algorithms in the next step, a 5000 cycle sliding moving average window with no overlap data reduction is applied. The values of the features are averaged in these 5000 cycles windows and they are assigned time values of  $x^* \cdot 5000$  cycles, where  $x$  is the increasing window number. Fig. 4b depicts the smoothed windowed feature. It is evident that the overall trend of the HI remains similar and at the same time, the noise and volatile trend are reduced, providing a more monotonic HI, while the overall attributes have also been improved (Fig. 4c).

Data collected from multiple sensors have different range of values. To address the heterogeneity of the input data, it is important to normalize them to a similar range. A normalization to the range  $[0,1]$  is performed using Eq. (4) before employing them to train the ML algorithms. No prior failure threshold is needed to be set with the regression approach that we follow and the normalization parameters of the training set are used to normalize the test set as well.

$$\bar{F} = \frac{F_i - \min(F_i)}{\max(F_i) - \min(F_i)} \quad (4)$$

### 3. Machine learning algorithms for prognostics

Several machine learning algorithms may be utilized for addressing the challenging task of RUL prognosis. In this study two commonly used regression algorithms from the literature are employed to take on this challenge. Gaussian process regression is selected for its ability to model complex trends with the use of the kernel function while also providing a distribution for the RUL, while Artificial Neural Networks were preferred because they can easily model non-linear behaviors.

#### 3.1. Gaussian processes for regression (GPR)

Gaussian processes have been widely employed for RUL prognosis of a variety of systems [33–36] and structures [23,37,38]. A GPs is a collection of random variables with a joint Gaussian distribution, and are a function of  $f(\mathbf{x})$  at  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ . GP can be completely specified [39] by its mean function:

$$m(\mathbf{x}) = E[f(\mathbf{x})] \quad (5)$$

And its covariance function:

$$k(\mathbf{x}, \mathbf{x}') = E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \quad (6)$$

Then the GP can be written as:

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x})) \quad (7)$$

The mean function  $m(\mathbf{x})$  is usually set to be zero. In our case though we selected a linear function to characterize the mean (see Section 4). As it is noted in [39] different covariance functions yield different regression results, so this function should be considered carefully depending on the data.

Assuming a degradation history  $\mathbf{H} = [x_i, y_i]_{i=1}^N$ , where  $x_i$  the input variables and  $y_i = f(x_i) + \epsilon_i$  the noisy target variables, with  $\epsilon_i$  is an i.i.d with 0 mean and  $\sigma_n^2 (\epsilon_i \sim i.i.d N(0, \sigma_n^2))$  the joint distribution of observed target values  $\mathbf{y} = [y_i]_{i=1}^N$  and unobserved target values  $f^*$  at new input locations  $x^*$  can be denoted as:

$$\begin{bmatrix} \mathbf{y} \\ f^* \end{bmatrix} \sim N \left( 0, \begin{bmatrix} \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I} & \mathbf{K}(\mathbf{x}, \mathbf{x}^*) \\ \mathbf{K}(\mathbf{x}^*, \mathbf{x}) & \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) \end{bmatrix} \right) \quad (8)$$

where  $\mathbf{I}$  the identity matrix and  $\mathbf{K}$  a matrix of all the covariance pairs  $k(x_i, x_j)$ .

The predictive (posterior) distribution for GPR, given the new inputs  $x^*$ , the historic input data  $\mathbf{x}$  and targets  $\mathbf{y}$  is defined by:

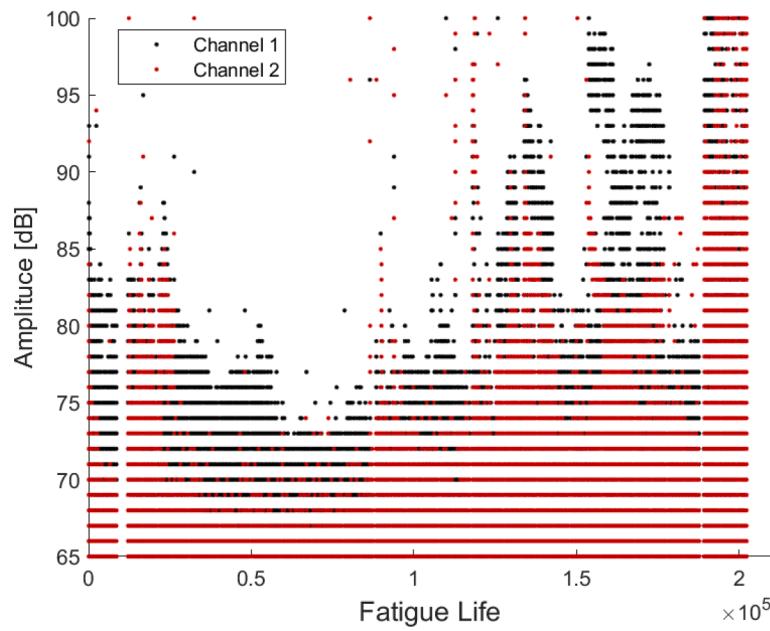
$$p(f^* | \mathbf{x}, \mathbf{y}, x^*) \sim N(\bar{f}^*, cov(f^*)) \quad (9)$$

$$\bar{f}^* = E[f^* | \mathbf{x}, \mathbf{y}, x^*] = \mathbf{K}(x^*, \mathbf{x}) [\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{y}, \quad (10)$$

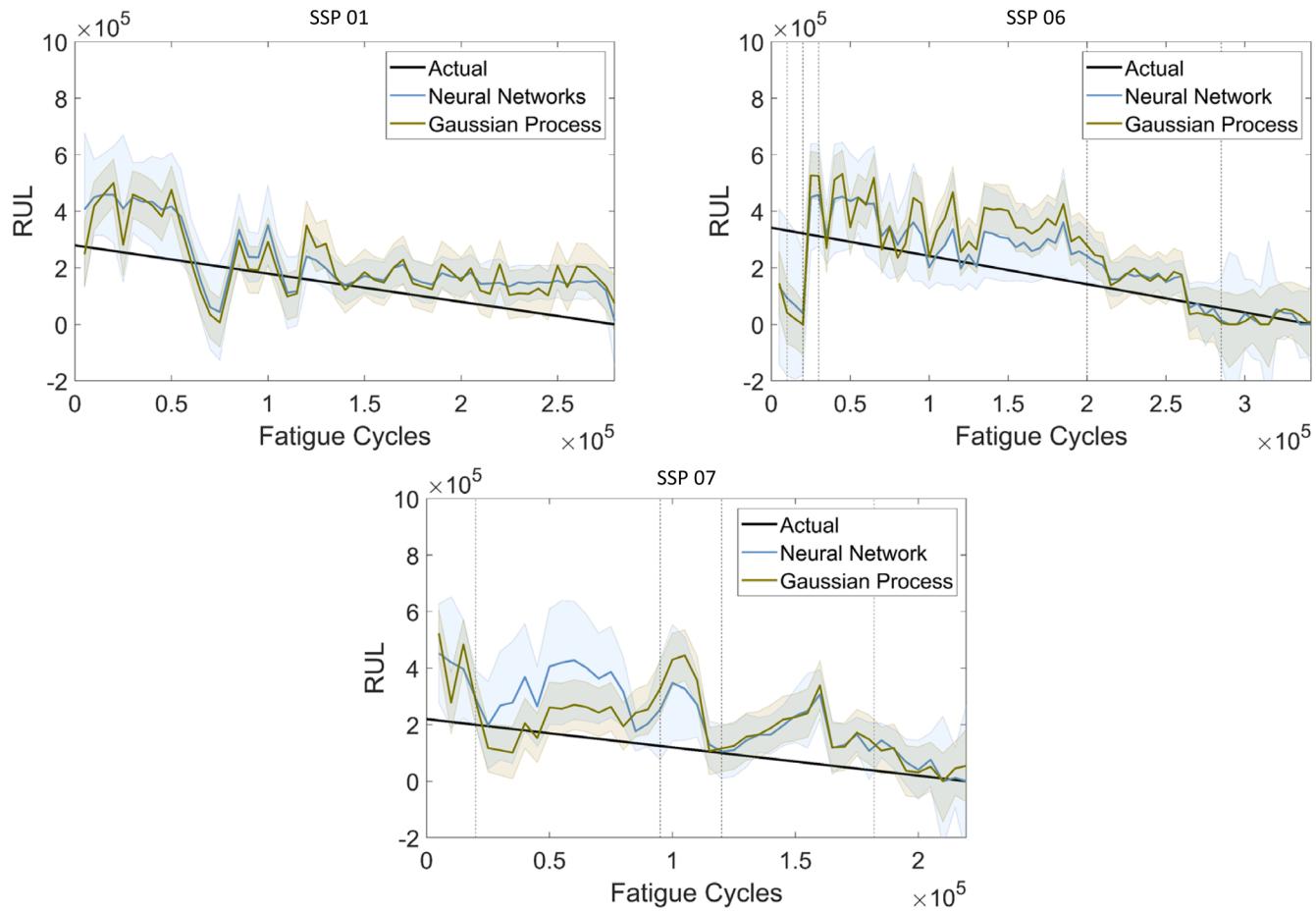
$$cov(f^*) = \mathbf{K}(x^*, x^*) - \mathbf{K}(x^*, \mathbf{x}) [\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{K}(\mathbf{x}, x^*) \quad (11)$$

#### 3.2. Bootstrapped Neural Networks (BNN)

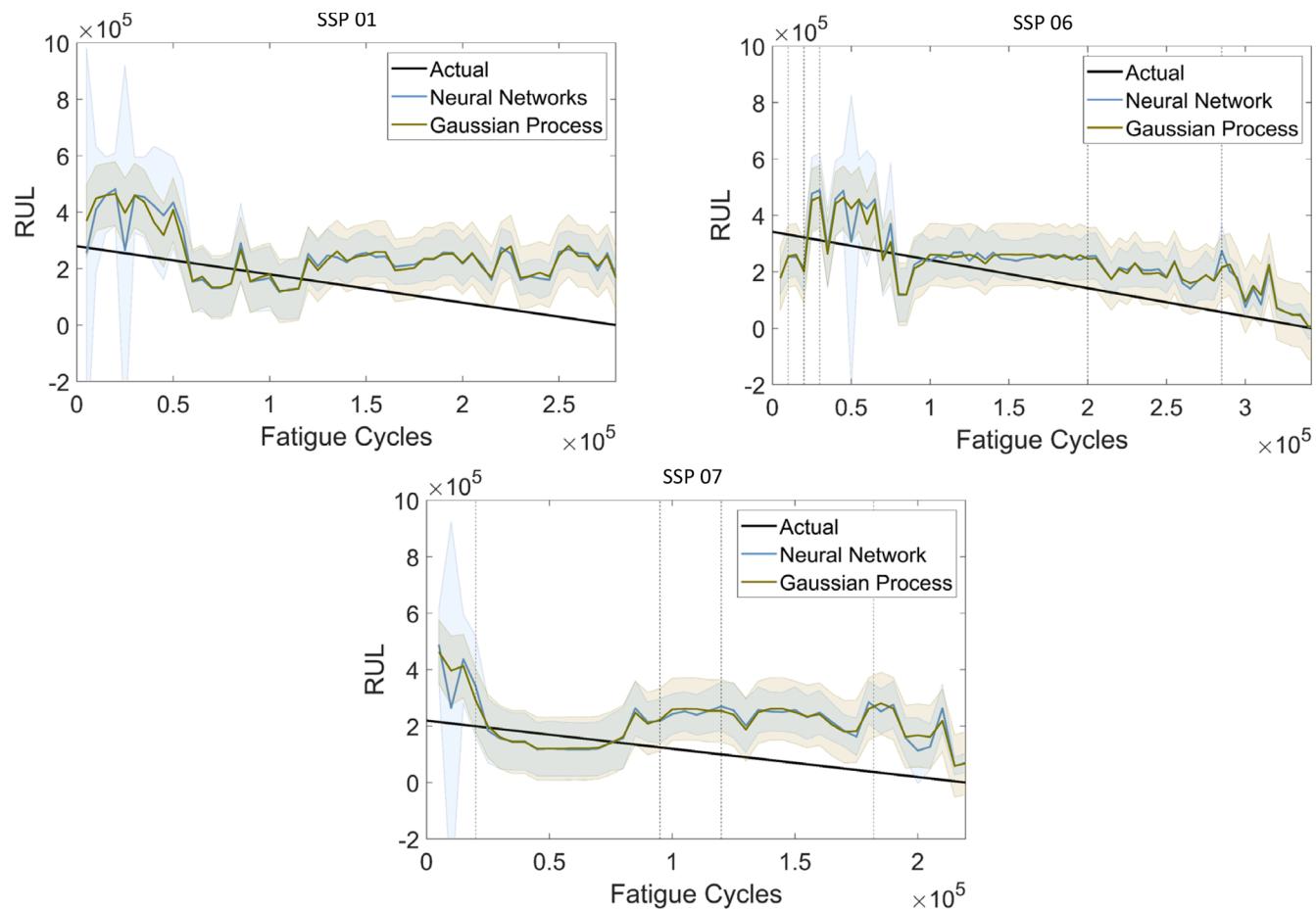
ANNs are a popular machine learning model that has found uses in



**Fig. 6.** A representative Acoustic Emission amplitude vs fatigue life after pre-processing.



**Fig. 7.** RUL prediction using multivariate models (dashed lines represent the load changes).



**Fig. 8.** RUL predictions using single-Feature models (dashed lines represent load changes).

**Table 2**

Prognostic performance metrics for the different methodologies and specimens calculated at 100% and 50% of lifetime (lower is better, bold denotes the best metric value).

Specimen Name			SSP 01		SSP 06		SSP 07		Average	
Metric	Algorithm	Method	100%	50%	100%	50%	100%	50%	100%	50%
MAE (kcycles)	GPR	Single	<b>69.86</b>	118.88	<b>70.96</b>	78.94	<b>81.28</b>	127.56	<b>74.03</b>	108.46
		Multi	106.50	<b>66.12</b>	133.01	72.67	103.43	76.40	114.31	71.73
	BNN	Single	131.09	157.49	89.19	94.13	121.75	161.21	114.01	137.61
		Multi	100.34	79.74	86.54	<b>58.50</b>	119.93	<b>75.12</b>	102.27	<b>71.12</b>
MAPE (%)	GPR	Single	264.10	511.12	96.80	170.12	272.80	514.03	211.23	398.42
		Multi	<b>183.60</b>	<b>294.28</b>	90.10	101.96	136.40	185.10	136.70	193.78
	BNN	Single	323.70	590.10	102.60	173.05	282.20	504.74	236.17	422.63
		Multi	191.80	322.71	<b>61.70</b>	<b>78.64</b>	<b>134.90</b>	<b>161.09</b>	<b>129.47</b>	<b>187.48</b>
RMSE (kcycles)	GPR	Single	<b>96.49</b>	132.44	<b>86.82</b>	96.24	<b>104.15</b>	139.27	<b>95.82</b>	122.65
		Multi	140.91	<b>84.55</b>	169.65	101.57	150.29	101.90	153.62	96.01
	BNN	Single	145.62	165.37	102.99	104.67	142.95	169.47	130.52	146.50
		Multi	116.05	89.20	109.38	<b>75.35</b>	148.00	<b>98.03</b>	124.48	<b>87.52</b>
wMAE (kcycles)	GPR	Single	121.36	156.33	88.85	73.54	127.50	152.02	112.57	127.30
		Multi	<b>92.82</b>	<b>90.07</b>	79.47	48.40	<b>77.80</b>	<b>53.66</b>	83.36	<b>64.04</b>
	BNN	Single	157.94	174.70	88.12	89.39	140.50	158.82	128.85	140.97
		Multi	101.32	94.91	<b>56.23</b>	<b>37.42</b>	85.32	60.56	<b>80.96</b>	64.30

**Table A1**  
Time domain statistical features.

Feature Name	Equation
Mean Value	$X_m = \frac{1}{N} \sum_{t=1}^N x(t)$
Standard Deviation	$X_{sd} = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (x(t) - X_m)^2}$
Skewness	$X_{sk} = \frac{\sum_{t=1}^N (x(t) - X_m)^3}{(N-1)X_{sd}^3}$
Kurtosis	$X_{kur} = \frac{\sum_{t=1}^N (x(t) - X_m)^4}{(N-1)X_{sd}^4}$
Root Mean Square (RMS)	$X_{rms} = \sqrt{\frac{1}{N} \sum_{t=1}^N (X(t))^2}$
Median	$X_{med} = \frac{1}{N} \sum_{t=1}^N \bar{x}(t)$

$x(n)$  denotes the  $n^{\text{th}}$  window where  $n = 1, \dots, N$ .

$N$  number of windows.

$\bar{x}(t)$  denotes the momens of occurrence of  $x(t)$ .

**Table A2**  
Frequency domain features.

Feature Name	Equation
Mean Frequency	$X_{mf} = \frac{1}{N} \sum_{k=1}^N s(k)$
Spectral Standard deviation	$Xf_{sd} = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (s(k) - X_{mf})^2}$
Spectral Skewness	$Xf_{sk} = \frac{\sum_{k=1}^N (s(k) - X_{mf})^3}{(N-1)X_{sd}^3}$
Spectral Kurtosis	$Xf_{kur} = \frac{\sum_{k=1}^N (s(k) - X_{mf})^4}{(N-1)X_{sd}^4}$

$s(k)$  denotes spectrum the  $k^{\text{th}}$  window where  $k = 1, \dots, N$ .

$K$  denotes the number of windows.

both regression and classification tasks. They are capable of modeling non-linearities in addition to linear correlations between the input and target data. The general equation governing these types of models at an  $i^{\text{th}}$  time point is:

$$t_i = y(x_i) + \varepsilon_i = f(x_i, w) + \varepsilon_i \quad (12)$$

where  $y(x_i)$  is the real regression mean and  $\varepsilon_i$  a noise term modeled as an  $i.i.d.$  with zero mean. The function  $f(x_i, w)$  is used to correlate the input  $x_i$  with the regression output  $y(x_i)$  given a set of parameters  $w$ .

The main disadvantage of ANNs is their inability to incorporate uncertainty in their predictions. To this end, bootstrapping is applied to tackle this shortcoming. Bootstrapping is an algorithmic method for constructing confidence and prediction intervals using the outputs of multiple Neural Networks. This technique is used to approximate a statistical distribution of the error term [40]. The idea behind bootstrapping is that given a dataset  $D$ ,  $K$  ANNs are randomly initialized and dataset  $D$  is split into  $K$  training sets by resampling with replacement. If a sufficiently large enough  $K$  is selected ( $K > 100$ ) then it can be assumed that the predictions follow a normal distribution [41–43]. The mean and variance of the predictions can be calculated as:

$$y_{p\text{mean}}(x_i) = \frac{1}{K} \sum_{k=1}^K y_p^k(x_i) \quad (13)$$

$$\sigma_{y_p}^2(x_i) = \frac{1}{K-1} \sum_{k=1}^K \left( y_{p\text{mean}}(x_i) - y_p^k(x_i) \right)^2 \quad (14)$$

where  $y_p^k(x_i)$  is the estimate of the  $k^{\text{th}}$  NN for the true value  $y(x_i)$  as estimated by  $f(x_i, w)$ . A schematic representation of the bootstrap

methodology for ANN can be seen in Fig. 5.

Once the mean and the variance of the ANN are calculated, the confidence interval at every point  $x_i$  can be constructed for the  $(1-a)\%$  confidence level as:

$$CI(x_i) = y_{p\text{mean}}(x_i) \pm t_{1-\frac{a}{2}} \sqrt{\sigma_{y_p}^2(x_i)} \quad (15)$$

#### 4. Results and discussion

Due to the complexity of the loading conditions as well as the noise from the hydraulic machine, AE produces millions of hits with no evident degradation information at a first glance. Significant data pre-processing steps are required in order to separate the useful degradation information from the external and internal noise. The first step is to filter the constant noise produced by the hydraulic machine. An amplitude threshold is set at a high 65 db in order to exclude AE hits related to noise from the machine which can contaminate and mask degradation information. At the second stage, hits recorded during low load (due to some extra SHM measurements made) are discarded. Some extra post-processing is applied (study RA, energy, duration) to try and separate degradation hits from other events captured during the cyclic loading, such as sensor cables touching the panels. Fig. 6 shows the retained hits for a representative panel.

Out of the 13 SSPs tested in this research, 3 are randomly chosen as the test set for evaluating the proposed methodology. The rest of the SSPs (10 SSPs) are used to learn the machine learning model parameters.

The two main tunable parameters of the GPR are the mean and covariance functions. Due to the nature of the output, a linear mean function is used as shown in Eq. (16). For the covariance function, a Matern class kernel (Matern 5/2) is selected since it was found to fit the data better (Eq. (17)).  $r$  is the Euclidean distance between two points  $x_i$ ,  $x_j$ . The hyperparameters to be optimized during the training procedure are  $[a, b, \sigma_f, \sigma_l]$ .

$$m(X) = aX + b \quad (16)$$

$$k(r) = \sigma_f^2 \left( 1 + \frac{\sqrt{5} r^2}{\sigma_l^2} + \frac{5r^2}{3\sigma_l^2} \right) \exp \left( -\frac{\sqrt{5} r}{\sigma_l} \right) \quad (17)$$

For the BNN a simple feedforward ANN is employed with 2 hidden layers consisting of 25 and 12 Neurons respectively. This layer combination is chosen by trial and error and no extensive investigation of optimal configuration is performed. The activation function in the two hidden layers is a sigmoid function while in the output layer a linear activation function is applied. The Levenberg-Marquardt algorithm is used to train the networks and each network uses 80% of the available data for the training process to prevent overfitting. For the bootstrap,  $K$  is set to 200 ANNs. The training time for all 200 Networks was 320 s, while for the GPR model only 6 s.

Fig. 7 depicts the RUL prediction results accompanied by their 95% Confidence intervals. We can qualitatively say that for SSP 01 the predictions using GPR are always close to the true RUL, though at the EoL it does not manage to converge satisfactorily to the ground truth. Unlike SSP 01, the GPR model for SSP 06 shows overall better RUL estimations. Though at the beginning the predictions are not closing into the true RUL, after 50% of the lifetime they start converging to the truth. For specimen SSP 07, up until approximately 75% of the lifetime, the predicted RUL is rather volatile, and does not have a clear trend. At the 75% mark, the predictions start to converge to the true RUL, though at the final prediction points the predicted RUL increases, diverging from the truth.

For SSP 01 the BNN predicted RUL at the early stages is close to the truth, though at the middle stages it starts to diverge, until near the EoL, when it rapidly converges to the true RUL. BNN prediction for SSP 06 shows a similar trend with the GPR, where after 50% of the lifetime the predictions start to converge to the true RUL. Something similar is

**Table B1**

Extracted Feature Names. Bold defines the selected 10 features.

Feature No.	Feature Name	Feature No.	Feature Name	Feature No.	Feature Name
1	RMS Kurtosis	28	Spectral RA std	55	Spectral Rise Time skewness
2	RA Kurtosis	29	Spectral Duration Mean	56	Spectral RMS std
3	Rise Time Kurtosis	30	Duration RMS	57	Duration Mean
4	RMS skewness	31	Energy skewness	58	Counts
5	Counts Kurtosis	32	Energy Mean	59	Spectral RA skewness
6	Spectral Energy Kurtosis	33	Counts std	60	RMS
7	Duration Kurtosis	34	Energy	61	RMS Median
8	Rise Time Skewness	35	Rise Time Mean	62	Spectral RMS Mean
9	Rise Time std	36	Spectral Counts Mean	63	Spectral Duration Skewness
10	RA Skewness	37	RA Mean	64	RMS Mean
11	Spectral Energy Skewness	38	Duration Median	65	RMS RMS
12	Counts Skewness	39	Counts RMS	66	Duration
13	Rise Time RMS	40	Rise Time Median	67	Spectral RMS Kurtosis
14	Duration Skewness	41	Amplitude Kurtosis	68	Amplitude std
15	RA std	42	RA Median	69	Spectral RMS Skewness
16	Spectral Rise Time Mean	43	Rise Time	70	Spectral Amplitude Mean
17	Energy Median	44	Spectral Duration std	71	Amplitude
18	RA RMS	45	Counts Mean	72	Hits
19	Spectral RA Mean	46	Spectral Counts Kurtosis	73	Spectral Hits Kurtosis
20	Energy Kurtosis	47	Spectral Rise Time Kurtosis	74	Amplitude Median
21	Spectral Rise Time std	48	RMS std	75	Spectral Amplitude Kurtosis
22	Energy std	49	Amplitude Skewness	76	Amplitude RMS
23	Energy RMS	50	RA	77	Amplitude Mean
24	Spectral Energy Mean	51	Spectral Counts std	78	Spectral Amplitude std
25	Counts Median	52	Spectral Counts Skewness	79	Spectral Hits std
26	Spectral Energy std	53	Spectral RA Kurtosis	80	Spectral Amplitude Skewness
27	Duration std	54	Spectral Duration Kurtosis	81	Spectral Hits Skewness

observed for SSP 07. Unlike the GPR after 75% of the lifetime, the BNN's predictions approach the true RUL with a very good approximation, especially near the EoL. What is worth highlighting is that the points of load changes the trend of the predicted RUL tends to change. This is mostly visible in SSP 06 at the early stages and at SSP 07 at both the early and the middle stages.

To demonstrate the need for the multivariate approach, i.e. using 10 features instead of 1 to estimate the RUL, we compare the RUL estimations of the multivariate approach with the predictions of the best feature in terms of prognostic attributes, i.e. spectral skewness of hits. It can be observed (Fig. 8) that the single feature-based predictions display a more smooth and less volatile (fewer spikes) behavior compared to the multivariate one, which is a result of the higher overall monotonicity of this particular feature. However, it is evident that the EoL of two out of the three panels is not predicted accurately, which can be attributed to overall lower prognosability. Only for SSP 06 we observed very good RUL estimations, though the multivariate approach's estimations are still superior.

To quantify the prediction performance, some common prognostic metrics are employed such as the MAE (Mean Absolute Error), MAPE (Mean Absolute Prediction Error) and RMSE (Root Mean Squared Error) [44]. Their formulations are presented in eq. (18)-(20) respectively:

$$MAE = \frac{1}{N} \sum_{i=1}^N |RUL_i - RUL_i^*| \quad (18)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|RUL_i - RUL_i^*|}{RUL_i} \times 100 \quad (19)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (RUL_i - RUL_i^*)^2}{N}} \quad (20)$$

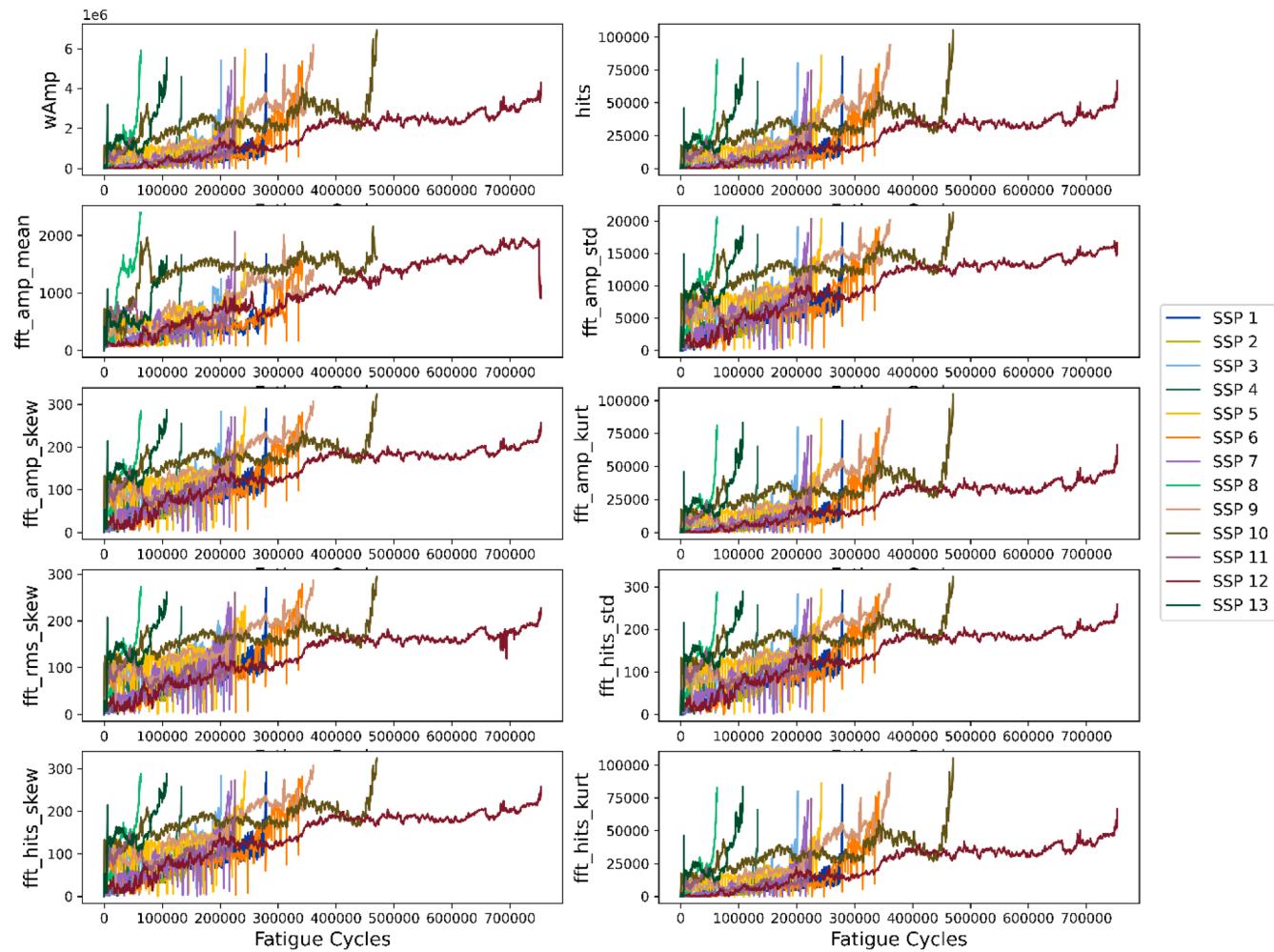
$RUL_i$  is the actual RUL and  $RUL_i^*$  is the predicted RUL at time  $t = i$ . Table 2 summarizes the aforementioned metrics for the different RUL estimation methodologies. The metrics are calculated at both 50% and 100% of the lifetime. When the metrics are calculated at 100% of the lifetime, we can observe that though visually the multivariate

predictions appear superior, metrics like MAE and RMSE do not verify this observation. This can be explained by looking at the overall trend of the predicted RULs. The single feature observations demonstrate more smooth and steady behaviors, while the multivariate RULs show a more erratic trend. Since the metrics measure the average error, it is reasonable that the single feature estimations display better results since at the beginning of the panels' lifetimes, they remain close to the true RUL. However, near the EoL they are not able to converge as effectively and this can be validated by the metrics calculated at 50% of the lifetime. When we observe the error metrics for 50% of the lifetime it can be seen that the multivariate approach greatly outperforms the single feature approach by up to 55% (see MAPE). For this reason, we also propose a prognostic performance metric that gives more significance to predictions closer to the EoL. We denote this metric as weighted MAE (wMAE) and it is expressed by eq. (21).  $w_i$  is a weight vector which gives higher weight (and hence a larger penalty) to the error near the EoL.

$$wMAE = \frac{1}{N} \sum_{i=1}^N |(RUL_i - RUL_p^i) w_i| \quad (21)$$

where  $0 \leq w_i \leq 1$  and  $\sum w_i = 1$

For the metrics at 100% of the lifetime, unlike the MAE and the RMSE, MAPE for the single feature-based predictions is significantly higher, which occurs since the metric measures the relative error to the true RUL. This can be also seen in the wMAE where the increased penalty at the later stages of prediction has a significant negative effect on the single feature estimations. In the case where the calculations are made only for 50% of the lifetime, the multivariate approach outperforms almost every time, with an improvement of up to 55%. Though training the single feature BNN is much more efficient (170 s training time) it is believed that the benefits to the performance of the multivariate approach are significant. However, the single feature GPR requires 5 s to train which is only a 1 s improvement over the multivariate model. When comparing the performance of the two algorithms, we can see that it is comparable with no algorithm significantly outperforming the other, except for the training time.



**Fig. B1.** Selected AE features for prognostics before smoothing.

## 5. Conclusions

We present a methodology for RUL prediction of composite single-stiffened panels subjected to fatigue compression loading based on AE data. Two different fatigue loadings were applied, namely, constant amplitude and variable amplitude. Acoustic emission sensors were employed to monitor the panels' degradation aiming towards an SHM framework for RUL prediction. To this end, several acoustic emission features from the time and frequency domains were extracted and evaluated for their suitability to perform the RUL estimation task. The discovered features were visually inspected to select the 10 most promising to perform this task. Three common attributes were utilized to validate the selected features, i.e. monotonicity, prognosability and trend. The selected features are normalized and smoothed to create suitable Health Indicators to perform the RUL prognosis.

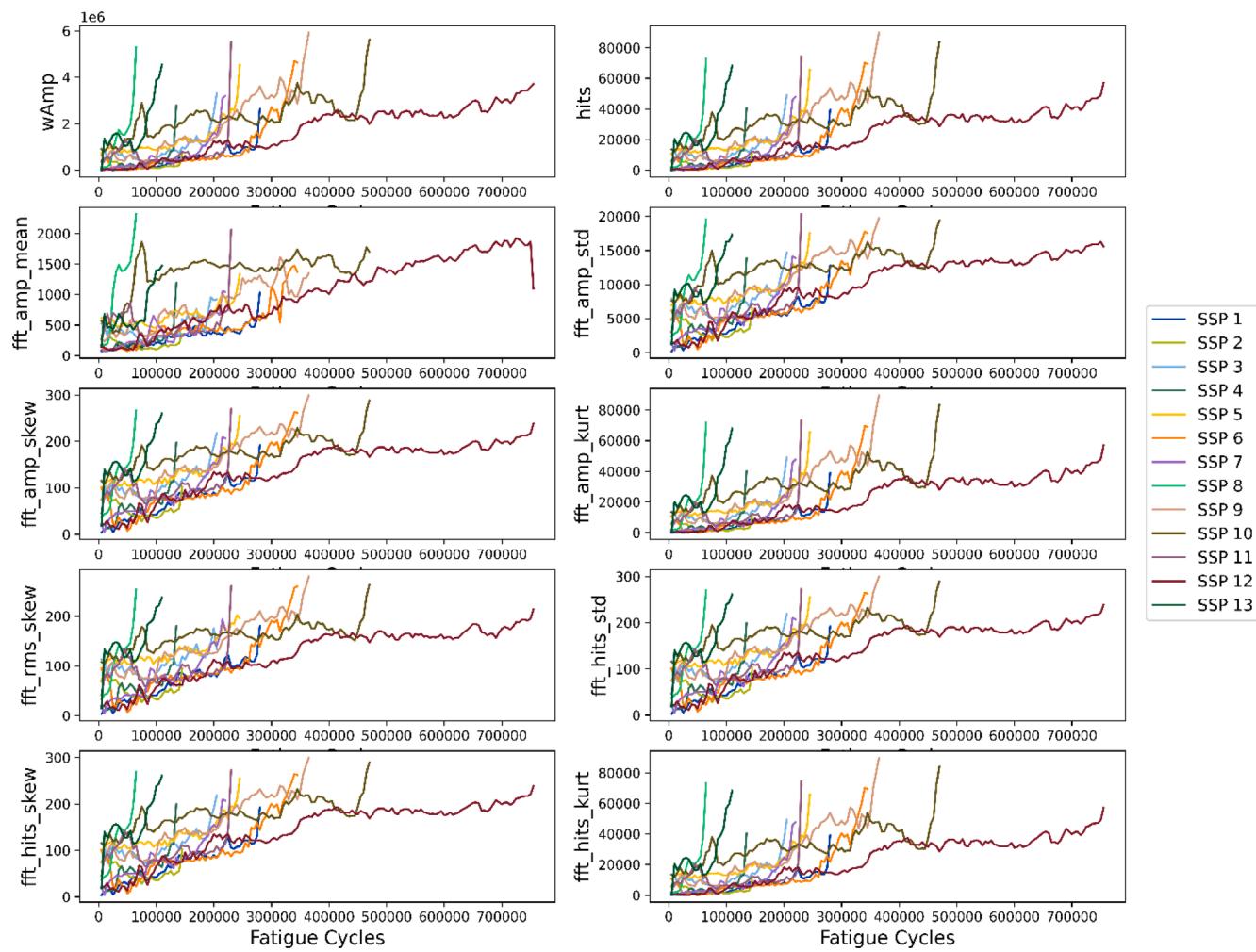
For the challenging task of RUL prognosis, two well established machine learning algorithms were employed, Gaussian process regression and Artificial Neural Networks. Out of the 13 available degradation histories 3 are selected at random to be used as a test set, while the remaining 10 are used to train the ML models and tune their parameters. The RUL estimations of both models show good results managing to estimate the true RUL with great accuracy. However, though the prediction performance is comparable, training time for the GPR is significantly lower. The multivariate RUL estimations are then compared with the prediction of the best feature (in terms of prognostic attributes), which showed good estimations, though at the EoL it failed to capture the true RUL. The predictions were then evaluated with some common

prognostic performance metrics from the literature. It was observed that though visual inspection displayed the superior performance of the multivariate method, some metrics did not depict that when calculated for the entire lifetime, due to the smoother behavior the single feature predictions displayed. However, when looking only at prediction from the middle up until the EoL, the multivariate approach significantly outperforms the single feature one with an improvement in metrics of up to 55%. This is happening because the multivariate approach is able to estimate the RUL close to the EoL much better. A new metric was also proposed, which gives more weight to the predictions closer to the EoL. This weighted prognostic metric (weighted MAE) along with the metrics calculated at 50% of the lifetime demonstrated the superiority of the multivariate methodology which significantly outperformed the single-feature methodology.

Though, the challenges of employing AE in more complex structures are not addressed, we demonstrate the feasibility of using a data-driven methodology for this prognostic task. This methodology is not proposed as an in-situ framework for RUL estimation, however, due to the relatively fast feature extraction (especially since only the discovered features can be calculated) and algorithm training process (which can also be done offline a priori) it can serve as a first step towards real-time monitoring of composite structures.

## CRediT authorship contribution statement

**Georgios Galanopoulos:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft. **Dimitrios**



**Fig. B2.** Selected AE features after applying the 5000 cycle window.

**Milanoski:** Conceptualization, Investigation. **Nick Eleftheroglou:** Conceptualization, Writing – review & editing. **Agnes Broer:** Investigation, Data curation. **Dimitrios Zarouchas:** Writing – review & editing, Funding acquisition. **Theodoros Loutas:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition.

#### Declaration of Competing 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.

#### Acknowledgement

The authors would like to acknowledge the colleagues at Delft University of technology (Aerospace Structures and Materials Laboratory and Structural Integrity and Composites group) and University of Patras (Applied Mechanics Laboratory) for the technical support.

#### Funding

The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant” (Project Number: 2573). A. Broer acknowledges funding from the European Union’s Horizon 2020 research and innovation program (grant no.769288, ReMAP project).

#### Appendix A

Tables A1 and A2 present some frequently used features extracted from AE signals in the time and frequency domain respectively [45].

#### Appendix B

In this appendix we present the feature names and curves (only selected) of the different features. Table B1 displays the feature names. Fig. B1 displays the selected windowed features, while Fig. B2 shows the selected smoothed features for all 13 panels studied in this paper.

#### References

- [1] Mourtzis AP, Townsend C, Shah Khan MZ. Non-destructive detection of fatigue damage in thick composites by pulse-echo ultrasonics. Compos. Sci. Technol., Jan. 2000;60(1):23–32. [https://doi.org/10.1016/S0266-3538\(99\)00094-9](https://doi.org/10.1016/S0266-3538(99)00094-9).
- [2] Wei B-S, Johnson S, Haj-Ali R. A stochastic fatigue damage method for composite materials based on Markov chains and infrared thermography. Int J Fatigue 2010; 32(2):350–60.
- [3] Saxena A, Goebel K, Larrosa CC, Janapati V, Roy S, Chang F-K. Accelerated aging experiments for prognostics of damage growth in composite materials. National Aeronautics and Space Administration Moffett Field CA Ames Research ..., 2011.
- [4] Philippidis TP, Vassilopoulos AP. Fatigue strength prediction under multiaxial stress. J Compos Mater 1999;33(17):1578–99.
- [5] Roundi W, El Mahi A, El Gharad A, Rebiere JL. Acoustic emission monitoring of damage progression in Glass/Epoxy composites during static and fatigue tensile tests. Appl Acoust Mar. 2018;132:124–34. <https://doi.org/10.1016/J.APACOUST.2017.11.017>.

- [6] Degrieck J, Van Paepegem W. Fatigue damage modeling of fibre-reinforced composite materials: review. *Appl Mech Rev* Jul. 2001;54(4):279–300. <https://doi.org/10.1115/1.1381395>.
- [7] Kralovec C, Schagerl M. Review of structural health monitoring methods regarding a multi-sensor approach for damage assessment of metal and composite structures. doi: 10.3390/s20030826.
- [8] Gouriveau R, Medjaher K, Zerhouni N. From prognostics and health systems management to predictive maintenance 1: Monitoring and prognostics 2016; 4: 1–163. 10.1002/9781119371052.
- [9] Nguyen KTP, Medjaher K, Tran DT. A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines. *Artif Intell Rev* 2022. <https://doi.org/10.1007/s10462-022-10260-y>.
- [10] Aggelis DG, Barkoula NM, Matikas TE, Paipetis AS. Acoustic structural health monitoring of composite materials : damage identification and evaluation in cross ply laminates using acoustic emission and ultrasonics. *Compos Sci Technol* Jun. 2012;72(10):1127–33. <https://doi.org/10.1016/J.COMPSCITECH.2011.10.011>.
- [11] de Oliveira R, Marques AT. Health monitoring of FRP using acoustic emission and artificial neural networks. *Comput Struct* 2008;86(3–5):367–73. <https://doi.org/10.1016/J.COMPSTRUC.2007.02.015>.
- [12] Loutas TH, Kostopoulos V, Ramirez-Jimenez C, Pharaoh M. Damage evolution in center-holed glass/polyester composites under quasi-static loading using time/frequency analysis of acoustic emission monitored waveforms. *Compos Sci Technol* 2006;66(10):1366–75. <https://doi.org/10.1016/J.COMPSCITECH.2005.09.011>.
- [13] Lima RAA, Drobizak M, Bernasconi A, Carboni M. On crack tip localisation in quasi-statically loaded, adhesively bonded double cantilever beam specimens by acoustic emission. *Theor Appl Fract Mech* 2022;118:103286. <https://doi.org/10.1016/J.TAFMEC.2022.103286>.
- [14] Saeedifar M, Zarouchas D. Damage characterization of laminated composites using acoustic emission: a review. *Compos Part B Eng* 2020;195:108039. <https://doi.org/10.1016/J.COMPOSITEB.2020.108039>.
- [15] de Groot PJ, Wijnen PAM, Janssen RBF. Real-time frequency determination of acoustic emission for different fracture mechanisms in carbon/epoxy composites. *Compos Sci Technol* 1995;55(4):405–12. [https://doi.org/10.1016/0266-3538\(95\)00121-2](https://doi.org/10.1016/0266-3538(95)00121-2).
- [16] Gutkin R, Green CJ, Vangrattanachai S, Pinho ST, Robinson P, Curtis PT. On acoustic emission for failure investigation in CFRP: Pattern recognition and peak frequency analyses. *Mech Syst Signal Process* 2011;25(4):1393–407.
- [17] Zhou W, Lv ZH, Li ZY, Song X. Acoustic emission response and micro-deformation behavior for compressive buckling failure of multi-delaminated composites 2016; 51(6): 397–407. doi: 10.1177/0309324716645244.
- [18] Liu PF, Chu JK, Liu YL, Zheng JY. A study on the failure mechanisms of carbon fiber/epoxy composite laminates using acoustic emission. *Mater Des* 2012;37: 228–35. <https://doi.org/10.1016/j.matdes.2011.12.015>.
- [19] Saeedifar M, Najafabadi MA, Zarouchas D, Toudestsky HH, Jalalvand M. Clustering of interlaminar and intralaminar damages in laminated composites under indentation loading using Acoustic Emission. *Compos Part B Eng* 2018;144: 206–19. <https://doi.org/10.1016/j.compositesb.2018.02.028>.
- [20] Broer A, Galanopoulos G, Benedictus R, Loutas T, Zarouchas D. Fusion-based damage diagnostics for stiffened composite panels. *Struct Heal Monit* 2021. p. 14759217211007128.
- [21] Rajendrabooopathy S, Sasikumar T, Usha KM, Vasudev ES. Artificial neural network a tool for predicting failure strength of composite tensile coupons using acoustic emission technique. *Int J Adv Manuf Technol* 2009;44(3–4):399–404. <https://doi.org/10.1007/S00170-008-1874-X>.
- [22] Arumugam V, Shankar RN, Sridhar BTN, Stanley AJ. Ultimate strength prediction of carbon/epoxy tensile specimens from acoustic emission data. *J Mater Sci Technol* 2010;26(8):725–9. [https://doi.org/10.1016/S1005-0302\(10\)60114-4](https://doi.org/10.1016/S1005-0302(10)60114-4).
- [23] Liu Y, Mohanty S, Chattopadhyay A. A Gaussian process based prognostics framework for composite structures. Modeling, signal processing, and control for smart structures 2009 2009;7286:72860.
- [24] Eleftheroglou N, Loutas T. Fatigue damage diagnostics and prognostics of composites utilizing structural health monitoring data and stochastic processes. *Struct Heal Monit* 2016;15(4):473–88.
- [25] Loutas T, Eleftheroglou N, Zarouchas D. A data-driven probabilistic framework towards the in-situ prognostics of fatigue life of composites based on acoustic emission data. *Compos Struct* 2017;161:522–9. <https://doi.org/10.1016/j.comstruct.2016.10.109>.
- [26] Eleftheroglou N, Zarouchas D, Loutas T, Alderliesten R, Benedictus R. Structural health monitoring data fusion for in-situ life prognosis of composite structures. *Reliab Eng Syst Saf* 2018;178:40–54. <https://doi.org/10.1016/j.ress.2018.04.031>.
- [27] Eleftheroglou N, Zarouchas D, Benedictus R. An adaptive probabilistic data-driven methodology for prognosis of the fatigue life of composite structures. *Compos Struct* 2020;245:112386.
- [28] Galanopoulos G, Milanoski D, Broer A, Zarouchas D, Loutas T. Health monitoring of aerospace structures utilizing novel health indicators extracted from complex strain and acoustic emission data. *Sensors* 2021;21(17):5701.
- [29] Loutas T, Eleftheroglou N, Georgoulas G, Loukopoulos P, Mba D, Bennett I. Valve failure prognostics in reciprocating compressors utilizing temperature measurements, PCA-based data fusion, and probabilistic algorithms. *IEEE Trans Ind Electron* 2019;67(6):5022–9.
- [30] Coble J, Hines JW. Identifying optimal prognostic parameters from data: a genetic algorithm approach. In: Annual conference of the PHM society. vol. 1, no. 1; 2009.
- [31] Medjaher K, Zerhouni N, Baklouti J. Data-driven prognostics based on health indicator construction: application to PRONOSTIA's data. *European Control Conference (ECC) 2013;2013:1451–6*.
- [32] Yan J, He Z, He S. A deep learning framework for sensor-equipped machine health indicator construction and remaining useful life prediction. *Computers & Industrial Engineering*, Volume 172, Part A, 2022, <https://doi.org/10.1016/j.cie.2022.108559>.
- [33] Liu K, Gebrael NZ, Shi J, Stewart HM. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis the authors are with. *IEEE Trans Autom Sci Eng* 2013; 10(3). doi: 10.1109/TASE.2013.2250282.
- [34] Li M, Sadoughi M, Shen S, Hu C. Remaining useful life prediction of lithium-ion batteries using multi-model Gaussian process. 2019 IEEE Int Conf Progn Heal Manag ICOPHM 2019;2019. <https://doi.org/10.1109/ICOPHM.2019.8819384>.
- [35] Benker M, Bliznyuk A, Zaeh MF. A Gaussian process based method for data-efficient remaining useful life estimation. *IEEE Access* 2021;9:137470–82. <https://doi.org/10.1109/ACCESS.2021.3116813>.
- [36] Roberts S, Osborne M, Ebdon M, Reece S, Gibson N, Aigrain S. Gaussian processes for time-series modelling. *Gaussian Process Time-Series Model* 2013. <https://doi.org/10.1098/rsta.2011.0550>.
- [37] Liu Y, Mohanty S, Chattopadhyay A. Condition based structural health monitoring and prognosis of composite structures under uniaxial and biaxial loading. *J Nondestruct Eval* 2010;29(3):181–8. <https://doi.org/10.1007/s10921-010-0076-2>.
- [38] Galanopoulos G, Eleftheroglou N, Milanoski D, Broer A, Zarouchas D, Loutas T. An SHM data-driven methodology for the remaining useful life prognosis of aeronautical subcomponents; 2023. p. 244–253. doi: 10.1007/978-3-031-07254-3\_24.
- [39] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, vol. 2, no. 3. MIT press Cambridge, MA, 2006.
- [40] Tibshirani RJ, Efron B. An introduction to the bootstrap. *Monogr Stat Appl Probab* 1993;57:1–436.
- [41] Heskes T. Practical confidence and prediction intervals. *Advances in neural information processing systems* 1996;9.
- [42] Zio E. A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. *IEEE Trans Nucl Sci* 2006;53(3):1460–78. <https://doi.org/10.1109/TNS.2006.871662>.
- [43] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on neural networks* 2011;22(9):1341–56.
- [44] Saxena A, et al. Metrics for evaluating performance of prognostic techniques. In: 2008 international conference on prognostics and health management; 2008. p. 1–17.
- [45] Lei Y. Intelligent fault diagnosis and remaining useful life prediction of rotating machinery. *Intell Fault Diagnosis Remain Useful Life Predict Rotating Mach* 2016: 1–366. <https://doi.org/10.1016/C2016-0-00367-4>.