Delft University of Technology

Qualification of distributed optical fiber sensors using probability of detection curves for delamination in composite laminates

Falcetelli, Francesco; Cristiani, Demetrio; Yue, Nan; Sbarufatti, Claudio; Troiani, Enrico; Di Sante, Raffaella; Zarouchas, Dimitrios

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

Qualification of distributed optical fiber sensors using probability of detection curves for delamination in composite laminates

Francesco Falcetelli[1,2] , **Demetrio Cristiani[2,3]** , **Nan Yue[2,4]** , **Claudio Sbarufatti[3], Enrico Troiani[1], Raffaella Di Sante[1] and Dimitrios Zarouchas[2,4]**

Abstract
Despite the promising application of Distributed Optical Fiber Sensors (DOFS) in monitoring damage in composite structures, their implementation outside academia is still unsatisfactory due to the lack of a systematic approach to assessing their damage detection performance. The existing tool developed for non-destructive evaluation, Probability of Detection (POD) curves, needs to be adapted for structural health monitoring applications to account for spatial and temporal dependency. Damage detection performance with DOFS is deeply related to the inherent variability sources of the system, the strain transfer properties of the optical fiber, and the loading conditions, which determine the damage-induced strain on the structure. This paper establishes a systematic approach based on the Length at Detection (LaD) method to qualify DOFS for damage detection in composites under different scenarios. Specifically, this study considers two DOFS with different strain transfer properties for monitoring delamination in carbon fiber reinforced polymers double-cantilever beam specimens under mode I quasi-static and fatigue loading. The POD curves derived from the LaD method confirm that this methodology can quantify the change in the detection performance due to the DOFS type and the loading conditions. The study also proposes a practical solution to compare POD curves obtained with different sample sizes, by introducing the concept of virtual specimens to simulate the lower confidence bound convergence.

Introduction

Over the last decades, composite laminates have become the predominant structural material in various engineering applications. Nowadays, the quest to develop safer and lighter structures still fosters the scientific community to investigate different damage mechanisms in composite materials and their reciprocal interaction. However, despite the impressive amount of research, open questions are still present, and the understanding of the physics behind failure modes in composites is limited.[1] Moreover, composite structures are particularly susceptible to flaws arising from the manufacturing process and service and exhibit complex failure modes as opposed to metals. Among them, delamination constitutes one of the most common damage

mechanisms and can also occur in adhesive bonds.[2] Using large safety factors mitigates the risks of

[1]Department of Industrial Engineering, University of Bologna, Bologna, Italy
[2]Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands
[3]Dipartimento di Meccanica, Politecnico di Milano, Milano, Italy
[4]Center of Excellence in Artificial Intelligence for Structures, Prognostics and Health Management, Faculty of Aerospace Engineering, Delft, The Netherlands

Corresponding author:
Francesco Falcetelli, Department of Industrial Engineering, University of Bologna, Via Fontanelle 40, Forlì, 47121, Italy.
Email: francesco.falcetelli@unibo.it

catastrophic structural failure but leads to heavier designs and might not be deemed sufficient to guarantee safety.

Consequently, delamination growth represents a severe threat to structural integrity in carbon fiber reinforced polymer (CFRP) structures, and it becomes necessary to implement Structural Health Monitoring (SHM) strategies. SHM can provide essential information about delamination existence, location, and size. Moreover, it can also deepen the understanding of other correlated damage mechanisms and thus promote the introduction of innovative composite materials and structures.[3] SHM offers a wide range of techniques, each with its strengths and weaknesses depending on the application.

Among them, Optical Fiber Sensors (OFS) provide numerous advantages over traditional strain sensing techniques. For example, they are intrinsically immune to electromagnetic interference; they have large bandwidth, which enables multiplexing solutions; and can survive harsh environmental conditions if protected with appropriate coatings and cable sheaths. Alj et al. provide further details about the durability of OFS.[4] Moreover, their lightweight and small size allow them to be embedded in composites[5] as well as 3D printed structures,[6] and they have recently been shown to be a viable alternative to accelerometers for modal analysis.[7] Recent advances in optical fiber technology fostered the use of Distributed Optical Fiber Sensors (DOFS) based on Raman, Brillouin,[8] and Rayleigh backscattering. DOFS based on Rayleigh backscattering are particularly promising for monitoring damage, such as delamination growth, in composites since they provide millimeter resolution along the fiber length within several meters of range.[9]

However, assessing the damage detection performance of optical fibers is not a straightforward process. The optical fiber datasheet specifies the geometrical, mechanical, and optical properties. On the other hand, the interrogator datasheet provides the resolution, wavelength range, wavelength stability, maximum sensor length, measurement uncertainty, and sampling rate. Nevertheless, these metrics do not directly assess strain-based damage detection performance since damage is not a physical quantity that can be directly measured. Indeed, Axiom IVa of SHM states that sensors cannot measure damage and that a feature extraction process is needed to obtain damage-related information.[10] For example, considering delamination monitoring, the detection performance is expected to change depending on the loading conditions because they affect the damage-induced strain in the structure. In addition, depending on the strain transfer occurring from the structure to the fiber core,[11] DOFS may exhibit different detection performances. The current

literature lacks well-established methodologies for certification and performance evaluation for damage detection, preventing the adoption of this technology in many applications.

The performance of Non-Destructive Evaluation (NDE) methods, widely accepted in many industries (aerospace, automotive, oil and gas, medical, and marine, to name a few), is evaluated following the guidelines provided in the MIL-HKBK-1823A.[12] First, damage detection performance is quantified using Probability of Detection (POD) curves and Probability of False Alarm (PFA). Furthermore, varying the threshold value makes it possible to evaluate the POD against the PFA and obtain the so-called Receiver Operating Characteristic curve.[13,14]

It is legitimate to ask whether these NDE reliability metrics can be applied to SHM. The naive application of POD curves in SHM would lead to inconsistent results. One of the most critical differences between NDE and SHM is their variability sources. The human factor represents the highest variability contribution in NDE, whereas SHM is affected by both temporal and spatial sources of variability. Moreover, SHM is typically characterized by repeated measurements over time, implying that the independent measurement assumption used in NDE does not hold.[15]

Meeker et al.[12] reviewed and proposed statistical methods for SHM,[16] extending the theory described in the MIL-HKBK-1823A. The authors demonstrate that the Length at Detection (LaD),[17] and the Repeated Measures Random Effects Model (REM²),[18] are valid statistical methods to handle SHM data. However, in both cases, the lack of data often hinders their applications since it is challenging to manufacture and test many structures equipped with identical sensing systems. Model-Assisted POD (MAPOD) curves can reduce the amount of requested experimental data. They allow the modeling of many types of variability sources, but the computational cost can be prohibitive due to the curse of dimensionality. Surrogate modeling can mitigate this problem and is already available in software such as CIVA.[19,20]

A systematic literature review by the authors[21] shows that in SHM, POD curves were mainly applied to Guided Waves[14,15,22–29] and occasionally to other techniques such as Comparative Vacuum Monitoring,[17,30] Acoustic Emissions,[31,32] and Carbon Nanotubes.[18] Supplemental Table A1, in the Appendix of Falcetelli et al.,[21] highlights that only a few POD studies on OFS are present, and no POD studies on DOFS are available. Grooteman developed a numerical model of a three-stringer thermoplastic composite panel installed with Fiber Bragg Gratings (FBGs) and computed the frequency shift in the eigenmodes. Then, using the modal strain energy as a damage indicator,

they generated a POD curve using the hit/miss approach.[33] Sbarufatti and Giglio[34] developed POD curves to quantify the performance of FBGs bonded onto an aluminum stiffened panel in terms of minimum detectable crack length. In this work, the authors compared the confidence interval for a population proportion method[35] with the one-sided tolerance interval (OSTI) for a normal distribution.[36]

To the best of the authors' knowledge, no study presents a rigorous methodology for qualifying DOFS in different scenarios using POD curves. A first attempt at generating POD curves for DOFS was explored by Falcetelli et al.[37] This study profoundly extends the preliminary research activity by establishing a systematic methodology based on the LaD method to qualify the damage detection performance of DOFS. The method is validated considering the use of DOFS for monitoring delamination in composite structures as a case study. Specifically, Single-Mode (SM) DOFS with ORMOCER® coating and Graded-Index Multimode (GIM) DOFS with a dual acrylate coating are surface mounted onto CFRP double cantilever beam (DCB) specimens under Mode I quasi-static and fatigue loading.

POD curves developed with the LaD method are used to evaluate the performance of the monitoring system of the two DOFS types in the two loading conditions. The results confirm that both strain transfer and loading conditions affect POD curves and prove that the proposed methodology can quantify the damage detection performance of DOFS in different scenarios.

Moreover, the authors introduce a practical approach to evaluating the required number of specimens based on the expected level of uncertainty. This method is two-fold since it can also serve for comparing POD curves generated from different sample sizes introducing the concept of virtual specimens. This dual functionality might be of great help in real applications where it is rare having homogeneous datasets.

The final aim of this article is not to promote the use of specific DOFS for delamination detection, rather is to develop a comprehensive methodology to assess the performance of DOSF and show the implications of making a POD study in SHM using DOFS. The proposed methodology aims to provide the SHM community with a reference procedure required to deploy DOFS in composite aircraft structures.

The article is organized as follows: Section "Materials and methods" presents the methodologies of this work, including the basic notions about POD in SHM, the method to estimate the required number of specimens and to compare POD curves produced from various sample sizes, the working principle of DOFS with Rayleigh backscattering, and the experimental methodology employed in the study; Section "Results" presents the results of the research; Section "Discussion" discusses the results and highlights their implications; Section "Conclusions" retraces the main stages of the article and suggests potential future research activities.

Materials and methods

POD for SHM: LaD method

The MIL-HKBK-1823A[12] describes how to derive POD curves using the famous â versus a method. As an example, Figure 1 shows the output of this methodology using synthetic data.
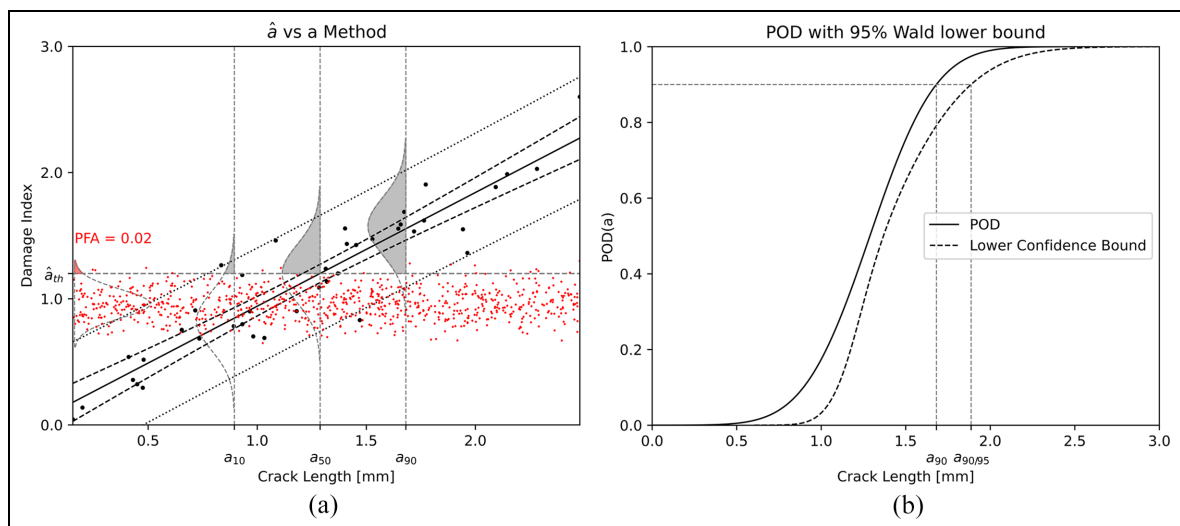
Figure 1. Example of the â versus a method (a), and the corresponding POD curve (b). (adapted from Falcetelli et al.[21]).
POD: Probability of Detection.

The critical a_{90} value represents the crack length whose POD equals 90%. On the other hand, $a_{90/95}$ represents the same concept but is referred to as the POD lower bound. Hence, it can be thought of as the crack length which is detected with 95% confidence with a POD of 90%. The difference between $a_{90/95}$ and a_{90} reflects the uncertainty associated with the experiments.

Different from NDE, where all the observations can be considered independent, in SHM, engineers may have to deal with repeated measurements. In this case, the hypothesis of statistical independence of the observations does not hold anymore, making it impossible to apply the standard linear regression model used for POD in NDE studies.[21]

The LaD method avoids the issue of dealing with repeated measurements by considering only one observation among the time series of data. Similar to the standard method proposed in the MIL-HKBK-1823A,[12] detection is positive if a certain measurement is above a certain threshold. However, only the measurements at which damage (delamination in this study) is detected for the first time are considered. The crack lengths at detection lay on the threshold, potentially following different statistical distributions. Assuming that the population is normally distributed, the POD curve can be determined as a function of the crack length using Equation (1):

$$POD(a) = Pr(X < a) = \Phi_{norm}\left(\frac{a - \bar{x}}{s}\right) \qquad (1)$$

The symbol Φ_{norm} denotes the standard normal cumulative distribution function; $\bar{x}$ and s represent the sample mean and variance, respectively, which differ from the true mean and the true variance, both of which are unknown. One of the available statistical methods to produce confidence intervals is the OSTI approach.[17] The tolerance bound T for a certain quantile of a normally distributed population can be obtained by exploiting Equation (2):

$$T = \bar{x} + k(n, \gamma, \alpha) \cdot s \qquad (2)$$

The *tolerance factor*, k, is the key term in Equation (2) and depends on three parameters: n, γ, and α. The first parameter n represents the number of samples. In this study, as outlined in Section "Experimental setup," the number of samples is equivalent to the total number of DOFS segments bonded in the specimens. It can be demonstrated that k decreases as n increases.[38] The second parameter γ controls the desired confidence level and is set to 95% in most cases. Finally, the third parameter α is used to define the coverage level. In POD studies, the value of α is usually equal to 90%.

Table 1. Pseudo-code for estimated sample size.

1	$n = n_i$	Define the number of samples used n the pilot test study
2	$\gamma = 0.95$	Confidence level on the lower bound of the POD
3	While $\Delta >$ tol	Imposing Guard Condition, start while loop
4	$\delta = z_{90}\sqrt{n}$	Non-centrality parameter of the non-central t distribution
5	$n = n + 1$	Increase the number of specimens by 1
6	$k = t_{n-1, \gamma, \delta}/\sqrt{n}$	Computation of the tolerance factor k
7	$a_{90/95} = \bar{x} + k \cdot s$	Updating $a_{90/95}$
8	$\Delta = a_{90/95} - a_{90}$	Updating Δ
9	End	End of the While Loop

It should be noted that the LaD is not the only method to derive POD curves. For instance, the REM[2] method[16,18] is a valid alternative that allows more efficient data use. Nevertheless, when less than ten observations are available, issues arise to fit a five-parameters model such as the REM[2].[21,39] In these kinds of studies, where testing many specimens or even real structures becomes costly and time-consuming, the LaD approach seems more appropriate.

Required number of specimens

Here the authors propose a practical scheme to assess the required number of samples, n, required for the experimental campaign. The strategy is to perform first a pilot study with few samples required to compute $\bar{x}$ and s. Then, $\Delta = a_{90/95} - a_{90}$ is iteratively computed, leveraging the properties of the non-central t-distribution,[40] and increasing the number of samples, n, in each cycle. Once Δ exceeds the imposed tolerance value, the algorithm exits the while loop and returns the required number of samples. The corresponding pseudo-code is represented in Table 1:

Where $t_{n-1, \gamma, \delta}$ is the inverse of the γ percentile of the non-central t distribution with n-1 degrees of freedom and non-centrality parameter δ.

Simulating the effect of virtual specimens on the lower bound

As the $a_{90/95}$ value reduces with n, in principle it would only be possible to compare the POD curves obtained from equal sample sizes. This aspect can be a limiting factor when acquiring data is particularly expensive and time-consuming. Therefore, in real applications, there is the need to consistently compare POD curves generated from a different number of test structures. The problem can be tackled by virtually augmenting the number of samples to a common value.

The first step is to use the LaD method to compute the value of $\bar{x}$ and s. Then, assuming that the tested specimens properly captured the main variability sources affecting the experimental setup, it is possible to simulate the effect of an increasing number of specimens. Using the same equations described in Table 1, one can define the tolerance factor k at different n values but keeping the values of $\bar{x}$ and s fixed. Introducing these virtual specimens shrinks the lower bound toward the POD curve, potentially allowing for comparison of POD obtained from small datasets with others obtained from greater sample sizes. The results of the proposed simulation must be taken cautiously, and based on the user expertise and the available previous knowledge, one can judge if the initial hypothesis that the original specimens properly capture the inherent variability of the experimental setup holds to be true.

Specimens manufacturing

The DCB coupons were produced following the guidelines described in the ASTM D5528 standard.[41] The AS4 HexPly 8552® unidirectional carbon prepreg[42] was employed to fabricate a 300 mm square panel with $[0_{24}]$ stacking sequence by hand layup. A 12 μm Teflon[TM] film was placed during lamination at the panel mid-plane. This non-adhesive insert served as an initiation site for the delamination, providing an initial crack length of 50 mm. The specimens were cut from the panel utilizing an automated Proth® cutting machine such that 25 mm strips were obtained. Ad hoc loading bocks were machined, matching the specimens width of 25 mm. Before bonding, the loading block surface was sandblasted, whereas the bonding surface of the specimens was slightly scrubbed with traditional sandpaper. Impurities were removed with an alcoholic solution, and the 3M™ Scotch-Weld™ EC-9323 structural epoxy adhesive[43] was used for bonding.

Optical fiber sensors

Two types of DOFS were used in this study: SM OFSs with ORMOCER®[44] coating, produced by FBGS Technologies GmbH (Jena, Germany), and GIM DOFS, produced by Plasma Optical Fibre (Eindhoven, The Netherlands). These were connected via LC/APC connectors to an ODiSI-B,[45] developed by LUNA Innovations Inc. (Roanoke, VA, USA). The interrogator uses swept-wavelength coherent interferometry to measure Rayleigh backscattering,[9,46,47] which originates as a result of non-propagating material-density fluctuations.[48] The scattered light exhibits a repeatable profile that is sensitive to longitudinal strain, ε, and temperature variation, ΔT. By correlating the scattered light profile before (baseline) and after (testing) a
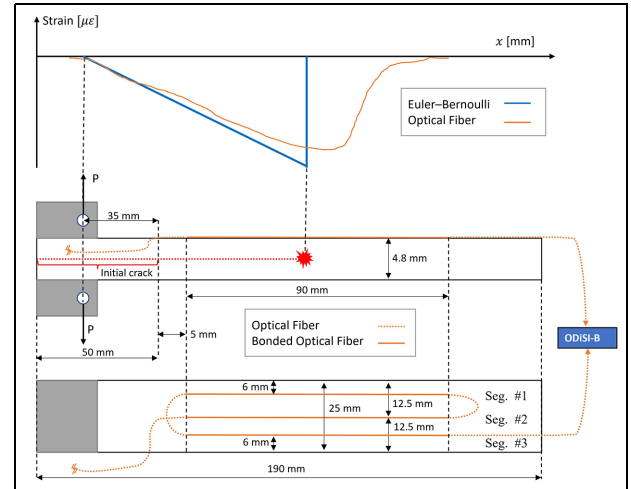
Figure 2. DCB specimen geometry and the optical fiber layout. DCB: double cantilever beam.

certain perturbation, it is possible to compute the spectral shift, $\Delta \nu$, or the variation in the resonance wavelength, $\Delta \lambda$, of the scattered light according to Equation (3):

$$\frac{\Delta \lambda}{\lambda} = \frac{\Delta \nu}{\nu} = K_T \Delta T + K_\varepsilon \varepsilon \qquad (3)$$

Where K_T and K_ε are the temperature and strain calibration constants. Equation (3) resembles the response of an FBG sensor. However, in this case, strain and temperature changes can be computed as a function of the fiber length with a certain spatial resolution, Δx, rather than just at the grating location. In this study, the interrogator was set up with a sampling frequency of 23.8 Hz and Δx equal to 0.65 mm.

Experimental setup

There are a large number of studies proposing analytical solutions for DCB specimens. The simplest analytical solution considers the DCB arms as cantilever beams clamped at the crack tip.[49] Both the Euler–Bernoulli beam theory and the Timoshenko beam theory can be used, with the latter providing more accurate results (Euler–Bernoulli-based solutions are a special case of Timoshenko-based solutions if the shear stiffness becomes infinite).[50]

The plot in Figure 2(a) shows a qualitative representation of the theoretical (Euler–Bernoulli solution) and expected measured experimental strain profiles along the longitudinal direction (*x*-axis) of the DCB specimen. The scheme shown in Figure 2(b) illustrates the specimen dimensions and the positioning of the DOFS along the top surface. The labels Seg. #1, Seg. #2, and
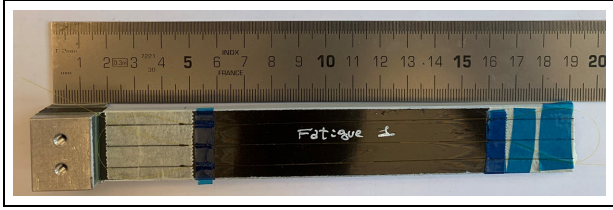
Figure 3. Specimen example used in the fatigue test.

Figure 4. Crack length estimation in the DCB specimen.
DCB: double cantilever beam.

Seg. #3, denote the three bonded segments present in each specimen. The configuration was chosen to minimize the bending radii of the DOFS.

Figure 3 shows an example of a DCB specimen used in the fatigue test and the DOFS positioned above its top surface. The bonding of the DOFS was achieved using ThreeBond 1742® cyanoacrylate adhesive.[51]

Before testing, one side of the DCB coupons was coated with a thin layer of white spray paint. After drying, 1 mm spaced vertical lines were used as a reference for visually estimating the crack length from the camera. An extra vertical mark is placed at the crack tip after the pre-cracking procedure explained in the D5528 standard.[41] Figure 4 shows a picture captured from a 9-Megapixel camera positioned in front of the specimen.
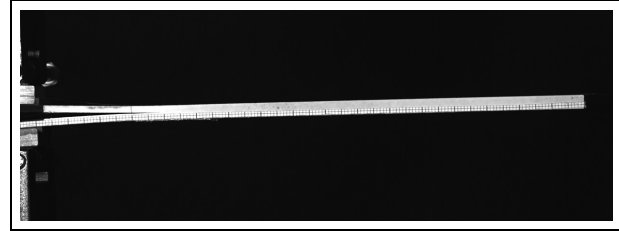
The true crack length is estimated by exploiting its relationship with the compliance C, which can be defined as the ratio between the load point displacement δ and the applied load in the DCB specimen P. As explained in the D5528 standard,[41] and shown in Sans et al.,[52] there is a linear relationship between the cube root of C and the crack length a:

$$\sqrt[3]{C} = \sqrt[3]{\delta/P} = c_1 a + c_2 \qquad (4)$$

Where c_1 and c_2 are the fitting parameters of the linear model. Therefore, once a sufficient number of observations is available, the linear model can be fitted, allowing the assessment of future crack length estimations from the C values (available at each time step) without visualizing hundreds of images. For example, Figure 5 shows the linear regression performed on specimen number 4, along with confidence and prediction intervals. As shown in the zoomed view, data fall inside the prediction intervals.

Data structure

Acquired strain data during static and fatigue tests of the i-th DOFS segment and j-th specimen are organized in a matrix S_i^j as follows:
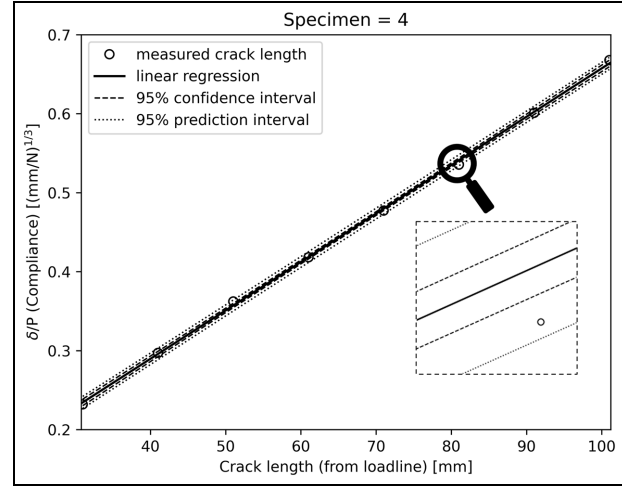
Figure 5. Compliance calibration.

$$\boldsymbol{S}_i^j = \begin{bmatrix} \varepsilon_{t=0}^{x=0} & \cdots & \varepsilon_{t=0}^{x=n} \\ \vdots & \ddots & \vdots \\ \varepsilon_{t=T}^{x=0} & \cdots & \varepsilon_{t=T}^{x=n} \end{bmatrix} \qquad (5)$$

Where t and x represent the time at which the measurement was taken and the location along the x-coordinate, respectively. The columns of $\boldsymbol{S}_i^j$ can be interpreted as the time history, t, of a single sensing element, whereas each row shows the strain profile along the fiber segment at a certain moment.

Similarly, the crack length is organized in a vector $\boldsymbol{a}_c^j$, where the lower script c indicates that the crack has been estimated with Equation (3) leveraging on the compliance value:

$$\boldsymbol{a}_c^j = \begin{bmatrix} a_{c,\,t=0} \\ \vdots \\ a_{c,\,t=T} \end{bmatrix} \qquad (6)$$

Static tests

A Zwick—20 kN tensile test machine was used for static testing, as shown in Figure 6.
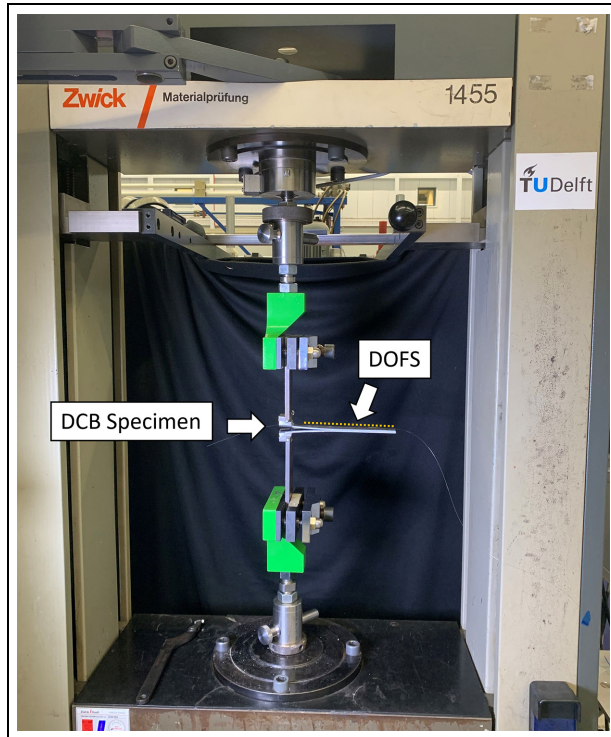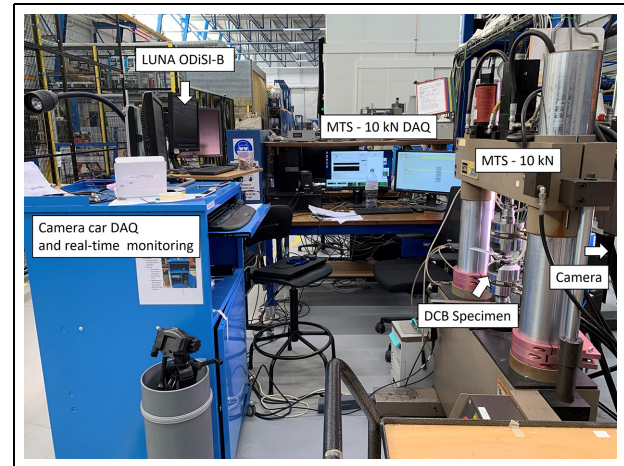
Figure 6. DCB specimen installed in the Zwick—20 kN tensile test machine.
DCB: double cantilever beam.

Figure 7. Fatigue test setup.

Figure 8. Fatigue test loading and measurement scheme.

The Zwick software was set up to synchronize the LUNA system and the camera. The tensile load is applied at a displacement rate of 1 mm/min. A sampling frequency of 0.5 Hz was used to collect data. The first experimental campaign used a total of five DCB specimens equipped with SM OFSs with ORMOCER® coating. Since three optical fiber segments are bonded onto each specimen, the number of linear regressions used to build POD curves can be multiplied by three.
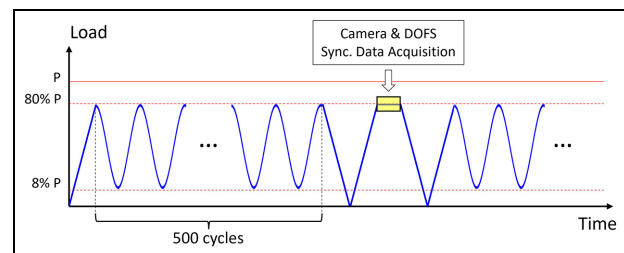
The same methodology was applied in a second experimental campaign, where six specimens equipped with GIM DOFS were tested. Preliminary results revealed that GIM optical fibers are more sensitive to small bending radii. As a result of the repeated bending of the optical fiber, the configuration shown in Figure 2 would have resulted in an unsatisfactory signal-to-noise ratio. Therefore, in this case, only one central optical fiber segment was bonded in the specimen.

Fatigue test

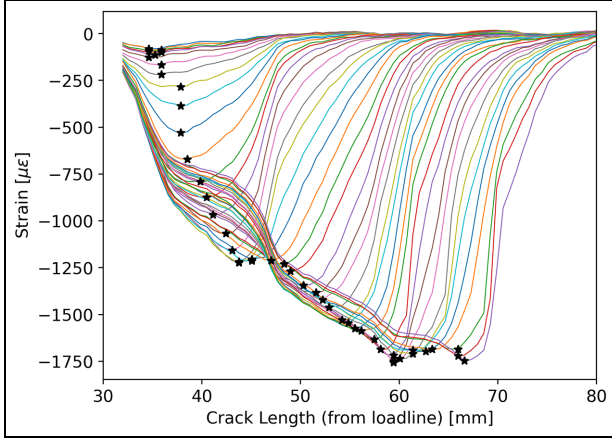An experimental fatigue test campaign was carried out on three specimens, where SM DOFS with ORMOCER® coating were surface bonded using the scheme previously shown in Figure 2. The DCB specimens were mounted in an MTS—10 kN Elastomer hydraulic test machine equipped with a 10 kN load cell. The whole experimental setup is shown in Figure 7.

The fatigue tests were performed in load control. Figure 8 shows a schematic overview of how the cycling loading was applied to the DCB specimens. Preliminary fatigue tests using DCB specimens manufactured from the same CFRP laminate were performed to assess the optimal load level to be used during fatigue testing. This preliminary study found that 80% of the pre-cracking load was the optimum load level for delamination growth. Lower loads would have led to very slow delamination growth, whereas higher loads would have resulted in an unstable delamination growth, which is not suitable for developing POD curves.

The MTS software was programmed to reach 80% of the pre-cracking load, P, with a ramp. Then, after every 500 cycles at 5 Hz with a loading ratio, R, equal to 0.1, the test is paused, and a trigger signal is sent to both the camera and the LUNA system to allow synchronized DOFS measurement and crack length estimation, respectively. This scheme was necessary

Figure 9. Strain profiles generated from DOFS analysis.

Figure 10. LaD method applied to SM DOFS data for crack detection.
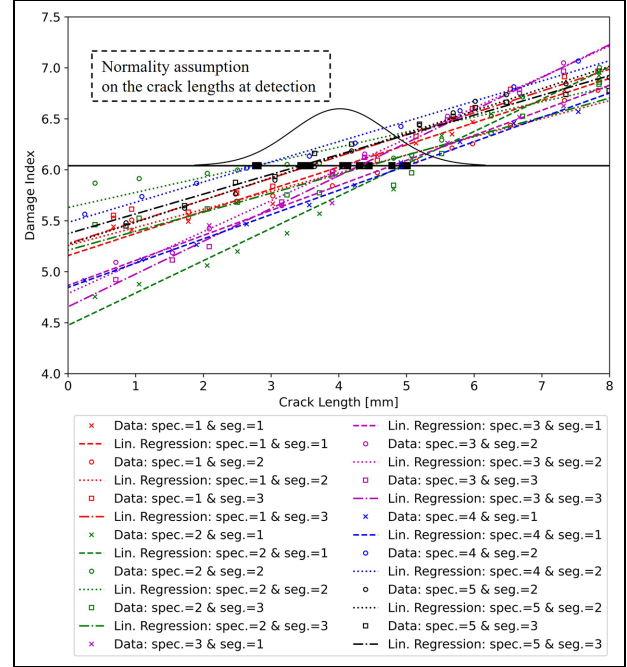LaD: Length at Detection; SM: Single-Mode.

because the DOFS signal-to-noise ratio degrades with vibrations, and acquiring clean data without interrupting the test is difficult. Moreover, this acquisition configuration guarantees that DOFS measurements are acquired under the same applied load on the specimens during the fatigue test, which is desirable since the damage index (defined in the next Section "Fatigue test") is load-dependent. Even if this choice brings some difficulties because the crack propagation may become unstable, the damage index will depend only on the crack propagation and not on the applied load.

Damage index definition

The first step in developing POD curves is to identify a proper damage-sensitive feature. From theory, it is possible to predict that the stress field reaches its maximum compressive value at the crack tip. Therefore, the strain value at the crack tip is a potential damage-sensitive feature. For a generic delamination value, and thus a generic time value t, it is possible to define a damage index DI_t as:

$$DI_t = \left| \min\left(\varepsilon_t^{x=0}, \ldots, \varepsilon_t^{x=n} \right) \right| \tag{7}$$

Figure 9 shows an example of the strain profiles obtained using DOFS at different times in the static test profile. The black stars, placed in correspondence with the lower peak of each strain profile, highlight the crack tip position and its relative propagation as delamination grows. Due to the non-linear strain transfer occurring between the specimen and the optical fiber,[11] and the distortion in the measured strain due to the interrogator resolution, the strain does not decrease linearly with the delamination length. The linearity is restored by applying a logarithmic transformation to

the damage index. The new definition of the damage index at generic time value, t, is given as:

$$DI_t = \ln \left| \min\left(\varepsilon_t^{x=0}, \ldots, \varepsilon_t^{x=n} \right) \right| \tag{8}$$

Then, it is possible to define a damage index vector $\mathbf{DI}_i^j$ denoting the evolution over time of the damage index related to the i-th DOFS segment and j-th specimen:

$$\mathbf{DI}_i^j = \ln \begin{bmatrix} \left| \min\left(\varepsilon_{t=0}^{x=0}, \ldots, \varepsilon_{t=0}^{x=n} \right) \right| \\ \vdots \\ \left| \min\left(\varepsilon_{t=T}^{x=0}, \ldots, \varepsilon_{t=T}^{x=n} \right) \right| \end{bmatrix} = \begin{bmatrix} DI_{t=0} \\ \vdots \\ DI_{t=T} \end{bmatrix} \tag{9}$$

Results

Static test

SM optical fibers. Figure 10 shows the application of the LaD method to SM DOFS with ORMOCER® coating. The abovementioned damage index behaves linearly with respect to the crack length, and linear regression is performed for every damage index vector $\mathbf{DI}_i^j$. Since every regression line has its own intercept and slope, the between-segment and between-specimen variability is considered in the model.

In Figure 10 the abscissa assumes zero value at the onset of the bonding length of each DOFS segment.
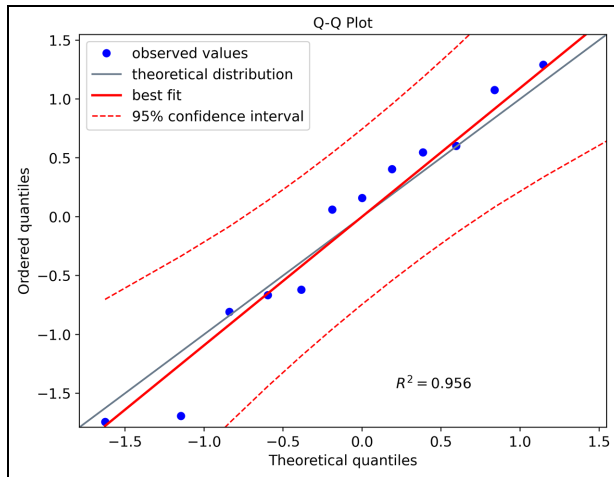
Figure 11. Darling–Anderson test.

Figure 13. LaD method applied to GIM DOFS data for crack detection.
LaD: Length at Detection; GIM: Graded-Index Multimode.
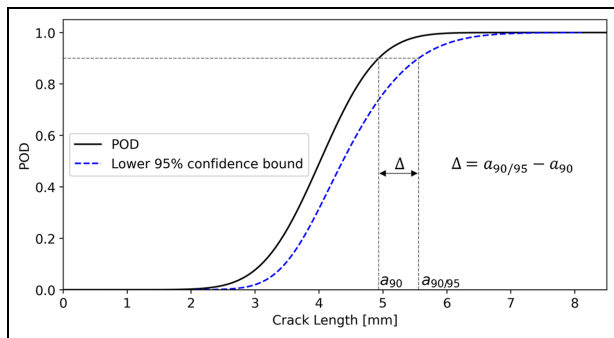

Figure 12. POD and its lower 95% confidence bound of SM DOFS.
POD: Probability of Detection; SM: Single-Mode.

The threshold was chosen by quantifying the noise level in preliminary experiments. Precisely, three standard deviations related to noise data were summed to the highest intercept of the regression lines. This procedure avoids negative lengths at detection, which would be the equivalent of saying that the crack was detected before it reached the bonded region of the DOFS, which should not be possible in principle.

The normality assumption of the lengths at detection can be verified using the Anderson Darling test (Figure 11). The null hypothesis, H_0, states that the data follow a normal distribution. The null hypothesis can be rejected if, for a certain significance level, α, the Anderson Darling statistics, A^2, is greater than the critical value. In the present case, considering a significance level $\alpha = 0.05$, and with a sample size $N = 13$, the critical value is equal to 0.679. The Anderson Darling statistics resulted to be $A^2 = 0.298 < 0.679$. Hence, H_0
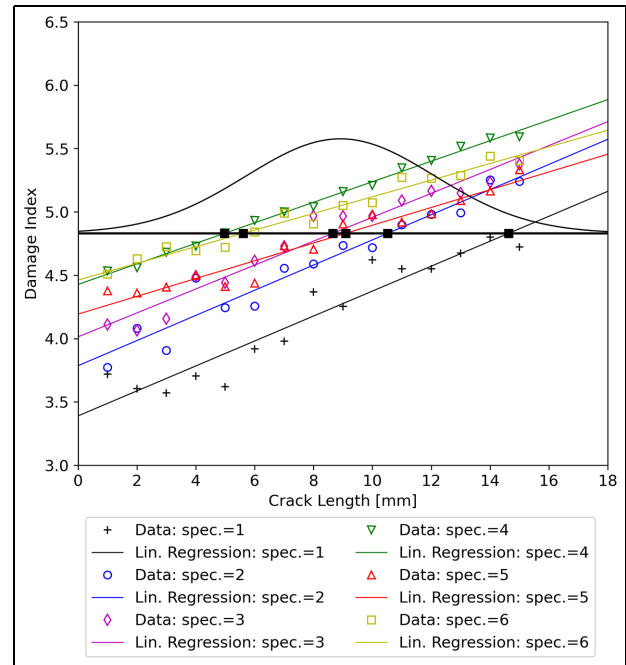
cannot be rejected, and the normality assumption holds.

Under the assumption that the crack lengths at detection, denoted as black squares in Figure 10, follow a normal distribution, it is possible to build a POD and its relative lower bound by applying Equations (1) and (2), respectively. Figure 12 shows the POD that was obtained using this methodology.

The identified values for a_{90} and $a_{90/95}$ in Figure 12 were 4.93 and 5.56 mm, respectively. The symbol Δ denotes their difference.

GIM fibers. The same methodology used for SM DOFS in Section "SM Optical Fibers" is now applied to the static test data obtained with GIM DOFS. The LaD results are shown in Figure 13. Although it is difficult to verify the normality assumption using the Anderson–Darling due to the low number of samples, the collected data are enough to show how a different strain transfer performance affects the resulting POD curve. The GIM DOFS has a dual acrylate coating whose stiffness is lower than the ORMOCER® coating of SM DOFS. This results in a lower strain transfer performance and a higher discrepancy between the real strain (the one present on the specimen surface) and the measured strain (strain present in the fiber core).
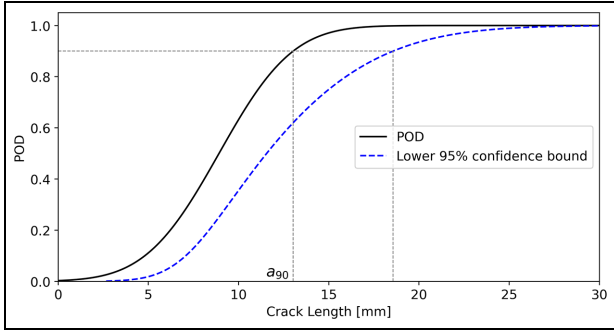
Figure 14. POD and its lower 95% confidence bound of GIM DOFS.
POD: Probability of Detection; GIM: Graded-Index Multimode.
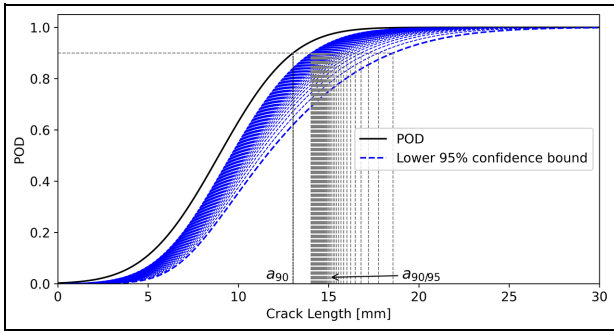
Figure 15. Lower bound convergence as the number of specimens increases.

Figure 16. LaD method applied to SM DOFS data in fatigue loading conditions for crack detection.
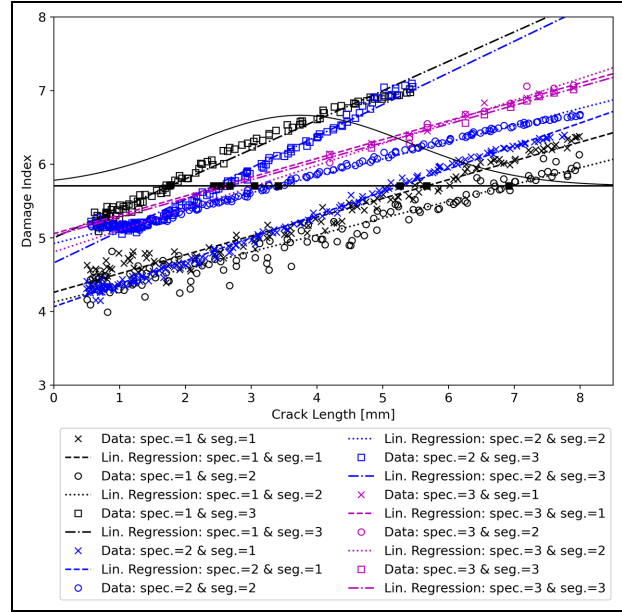LaD: Length at Detection; SM: Single-Mode.
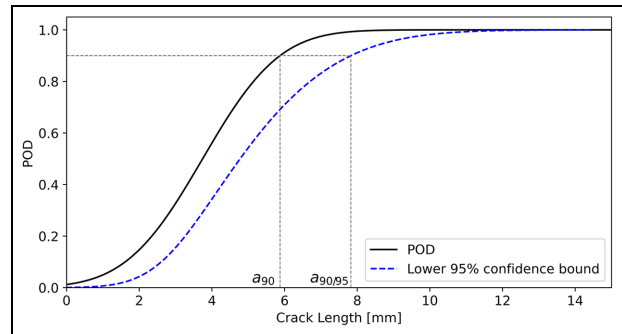
Figure 17. POD and its lower 95% confidence bound of SM DOFS in fatigue loading conditions.
POD: Probability of Detection; SM: Single-Mode.

Figure 14 displays the corresponding POD curve with a_{90} and $a_{90/95}$ equal to 13.03 and 18.56 mm, respectively. The poor performance in terms of strain transfer is reflected in a_{90} and $a_{90/95}$ whose values are significantly higher than the previous static case.

Applying the method proposed in Section "Required number of specimens," it is possible to show the convergence of the lower bound as the number of specimens increases (Figure 15). For example, when the number of specimens is equal to 97, the $a_{90/95}$ value reaches 14.02 mm. This procedure is useful for comparing experimental data collected from different samples. As will be shown in Section "Discussion," to compare the $a_{90/95}$ values of different experimental setups, the number of samples is virtually augmented to 30 in all cases.

Fatigue test

Figure 16 shows that the LaD method was applied to fatigue test data. The variability within segments of the same specimens and between different specimens is more pronounced than for the static case, even if the

same type of SM DOFS was used (with ORMOCER® coating).

The corresponding POD curve is shown in Figure 17, with a_{90} and $a_{90/95}$ equal to 5.88 and 7.82 mm, respectively.

The data highlight that both variability sources due to between-specimens and within-specimen heterogeneity are present. The first two specimens (black and blue color in Figure 16) have more data points with respect to the static case because of the large number of samples acquired every 500 cycles. On the other hand, the third specimen (magenta color in Figure 16) has few data points because the crack propagated beyond the bonded region of the DOFS right after the application

Table 2. Summary of a_{90} and $a_{90/95}$ values for the different case studies.

	Static testing		Fatigue testing
Optical fiber	Step Index SM	Graded Index Multi-Mode	Step Index SM
Coating	ORMOCER®	Dual acrylate	ORMOCER®
Number of samples (*n*)	13	6	9
a_{90}	4.93 mm	13.03 mm	5.88 mm
$a_{90/95}$	5.56 mm	18.56 mm	7.82 mm
$\Delta = a_{90/95} - a_{90}$	0.63 mm	5.53 mm	1.94 mm

SM: Single-Mode.

Table 3. Comparison of a_{90} and $a_{90/95}$ values for the different case studies after the application of the pseudo-code given in Table 1 (number of virtual samples equal to 30).

	Static testing		Fatigue testing
Optical fiber	Step Index SM	Graded Index Multi-Mode	Step Index SM
Coating	ORMOCER®	Dual acrylate	ORMOCER®
Number of *virtual* samples (*n*)	30	30	30
$a_{90;\ n=30}$	4.93 mm	13.03 mm	5.88 mm
$a_{90/95;\ n=30}$	5.29 mm	14.62 mm	6.70 mm
$\Delta_{n=30}$	0.35 mm	1.59 mm	0.82 mm

SM: Single-Mode.

of the pre-cracking load and propagated faster than in the previous two cases.

Compared to the static case, the measured strain is lower because the specimens were fatigue loaded at 80% of the pre-cracking load, *P*. This translates into a lower signal-to-noise ratio and consequently lower DI values, producing higher a_{90} and $a_{90/95}$.

Discussion

Comparison of POD curves

Table 2 summarizes the results in terms of a_{90} and $a_{90/95}$ for the different case studies.

Different optical fibers in the same loading configuration exhibit different a_{90} values, as shown in Table 2. GIM DOFS with the dual acrylate coating have lower strain transfer performance than SM fibers with ORMOCER® coating, which is reflected in the higher a_{90} value. Note that even the load plays a significant role in these metrics even if the fiber is the same. This is outlined in Table 2, comparing the first and the third columns. In fatigue loading, the POD curves are worse than the static case with higher values of a_{90}.

Comparison of POD lower bounds with virtual samples

The difference between a_{90} and $a_{90/95}$, Δ, can be considered a measure of the variability sources involved in the experiments. Indeed, as shown in Equation (2), the position of the lower confidence bound is a function of the standard deviation, *s*, of the lengths at detection.

In this study, the number of DOFS segments (samples) in each case is different. This situation is likely to occur in real applications due to the availability of different DOFS or, for example, a limited amount of time to perform fatigue tests compared to static tests.

As described in Section "Required number of specimens," since the tolerance factor *k* decreases as *n* increases, a higher number of samples, *n*, shrinks the difference between the POD curve and its lower bound, decreasing $a_{90/95}$ and correspondingly Δ. For example, Figure 15 shows how the sample size affects the $a_{90/95}$ value. The result was obtained using the pseudo-code developed in Table 1 (Section "Required number of specimens"). However, performing long and expensive experimental activities is not always possible. In such a case, the only solution to lower $a_{90/95}$ is redesigning the experimental setup and diminishing the associated variability sources to achieve a lower sample standard deviation.

Nevertheless, it would be interesting to compare the results obtained in this research by having the same number of samples for each case study. Referring to the procedure outlined in Section "Simulating the effect of virtual specimens on the lower bound," the authors virtually augmented the number of samples of the different case studies to 30 units. Under the assumption that the experimental data correctly captured the

variability sources involved in the experimental setup, this methodology allows a fair comparison between the different cases, eliminating a potential bias error due to the different sample sizes. Applying this procedure to Table 2, one obtains the results in Table 3.

The results confirm what is already seen in Table 2, even if the differences within the case studies are less accentuated.

Interpretation and implications of the results

The DOFS type proved to be a determinant factor in the POD analysis, which can be directly correlated to different strain transfer properties. On the other hand, the loading type is also shown to be a key variable. This is not surprising since DOSFs are sensitive to strain which depends on the applied load. The higher scattering in the fatigue data can be attributed to a lower signal-to-noise ratio. First, the test itself involves a higher amount of noise due to vibrations. Second, the crack propagates at a lower load, thus further reducing the signal-to-noise ratio. Moreover, the mechanisms involved in delamination growth are different in fatigue loading compared to quasi-static loading.[2] For example, a different amount of fiber bridging can affect the strain field in the process zone,[53] thus affecting the damage index and the POD parameters.

This result suggests that also the loading mode could potentially lead to different POD curves. Indeed, different mode mixites of Mode I and Mode II would affect the process zone and the strain profile, thus affecting the damage index. In such a case, a novel and more appropriate damage index should be developed because the strain at the crack tip may no longer be the best damage-sensitive feature.

Temperature variations are not considered in this study but are expected to be determinant in the POD analysis due to the relation between ΔT and $\Delta \lambda$ given in Equation (3). More in general, variation of Environmental and Operational Conditions (EOCs), damage morphology, sensor drift due to degradation (sensor and coupling), and additional variability sources dependent on the specific application will certainly affect POD curves. Therefore, it is essential to raise awareness about the limitations of the results and perform sensitivity studies to address the influence of the most determinant parameters.

Upscaling POD curves

In real applications, it could be inconceivable to test a sufficiently high number of structures to perform a statistically consistent POD study for DOFS. Indeed, one should be able to produce and replicate a large number of identical complex structures, each equipped with an identical DOFS setup. Even though the proposed methodology was developed considering DOFS in laboratory case studies, it offers a framework for assessing POD curves in real applications in two different ways.

First, it is possible to use the same methodology as a basis to derive MAPOD for DOFS. This could be achieved by simulating the outcome of the LaD method given the noise level, the loading conditions, and the strain transfer properties of the DOFS-structure mechanical system. The variability sources can be modeled assigning a certain probability distribution to the most critical parameters.

Second, POD curves obtained at a coupon level could be transferred at a structure level to monitor a specific damage type. The objective is to use the proposed methodology and build an experimental setup that mimics the local perturbation caused by damage in the strain field of a real structure. For example, in a hot spot monitoring scenario, where the structure is expected to fail due to mode-I delamination, the POD curves obtained from equivalent DCB specimens can provide an acceptable estimate of the damage detection performance of the system in the real application.

Conclusions

To the best of the authors' knowledge, this is the first time an experimental POD study has been performed for DOFS based on the Rayleigh backscattering. The study proposed a methodology to develop POD curves using the LaD method focusing on delamination, which is one of the major causes of failure for composites. Mode I static and fatigue loading experiments were performed on DCB specimens with two types of DOFS (SM fibers with ORMOCER® coating and GIM fibers with dual acrylate coating).

Probably, better POD curves could be obtained by using stiffer adhesives, redesigning the experimental setup to have lower noise, or using DOFS with higher strain transfer properties. However, the case studies that have been shown only serve as examples to show the implications of performing a POD study in SHM using DOFS. The goal is to develop an easily reproducible methodology to assess the performance of DOSF and to bring the attention of the SHM community to this topic which is often underestimated.

The following bullet points summarize the main finding of this research:

- Both loading conditions and DOFS type affect the performance in delamination detection
- POD curves for DOFS can also be sensitive to different loading modes, damage types, and laminate

stacking sequences, dramatically increasing the problem complexity compared to classical NDE applications.

- The LaD model proved effective in producing POD curves for DOFS, but the normality assumption is difficult to verify as the sample size decreases.
- Other POD models, such as the REM, do not require any normality assumption but are difficult to fit with small sample sizes.
- In many cases, the only feasible solution is to derive a MAPOD. The proposed framework, combined with preliminary knowledge regarding the most frequent damage modes in the structure, could be used to develop MAPOD for DOFS.
- The study provides a practical approach to estimating the required number of samples for the POD study.
- The same approach can be used to simulate the lower bound convergence, imposing a certain number of virtual samples to compare POD curves obtained from different sample sizes. Caution must be taken in interpreting the results since the underlying assumption is that the available samples properly captured the variability.
- The presence of unexpected variability sources, which are not captured in the experiments, such as varying EOCs, leads to unconservative results.

Based on the finding of this work, further research is needed and should be devoted to the following aspects:

- development of multi-dimensional POD curves varying the mode mixites between Mode I and Mode II for delamination;
- development of a MAPOD framework for DOFS;
- link the concepts of strain transfer and POD curves;
- development of compensation strategies for varying EOCs, sensor drift, and other variability sources potentially affecting POD curves.
- analysis of upscaling potentialities and limitations of such methodology, from both structural complexity and loading complexity aspects;

The final aim of this work is to spark a constructive debate in the SHM community about developing the most appropriate methodologies to certify DOFS for damage detection using POD curves.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

Francesco Falcetelli https://orcid.org/0000-0003-2820-776X
Demetrio Cristiani https://orcid.org/0000-0001-9054-0366

References

1. Saeedifar M and Zarouchas D. Damage characterization of laminated composites using acoustic emission: a review. *Compos Part B Eng* 2020; 195: 108039.
2. Pascoe JA, Alderliesten RC and Benedictus R. Methods for the prediction of fatigue delamination growth in composites and adhesive bonds – a critical review. *Eng Fract Mech* 2013; 112–113: 72–96.
3. Giurgiutiu V. *Structural health monitoring of aerospace composites*. Amsterdam: Elsevier, 2016.
4. Alj I, Quiertant M, Khadour A, et al. Environmental durability of an optical fiber cable intended for distributed strain measurements in concrete structures. *Sensors* 2021; 22: 141.
5. Di Sante R. Fibre optic sensors for structural health monitoring of aircraft composite structures: recent advances and applications. *Sensors* 2015; 15: 18666–18713.
6. Falcetelli F, Di Sante R and Troiani E. Strategies for embedding optical fiber sensors in additive manufacturing structures. In: Rizzo P and Milazzo A (eds) *European workshop on structural health monitoring. Lecture notes in civil engineering*. Cham: Springer International Publishing, 2021, vol. 128, pp. 362–371.
7. Falcetelli F, Martini A, Di Sante R, et al. Strain modal testing with fiber bragg gratings for automotive applications. *Sensors* 2022; 22: 946.
8. Bastianini F, Di Sante R, Falcetelli F, et al. Optical fiber sensing cables for brillouin-based distributed measurements. *Sensors* 2019; 19: 5172.
9. Gifford DK, Kreger ST, Sang AK, et al. Swept-wavelength interferometric interrogation of fiber Rayleigh scatter for distributed sensing applications. In: *Proceedings of the fiber optic sensors and applications V*, 12 October 2007, vol. 6770, p. 67700F. Boston, MA, United States: International Society for Optics and Photonics.
10. Farrar CR and Worden K. *Structural health monitoring: a machine learning perspective*. Chichester, West Sussex, UK; Hoboken, NJ, USA: Wiley, 2012.
11. Falcetelli F, Rossi L, Di Sante R, et al. Strain transfer in surface-bonded optical fiber sensors. *Sensors* 2020; 20: 3100.
12. *Department of defense handbook: nondestructive evaluation system reliability assessment*. 2009. https://statistical-engineering.com/mh1823/
13. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006; 27: 861–874.
14. Haynes C, Todd MD, Flynn E, et al. Statistically-based damage detection in geometrically-complex structures

using ultrasonic interrogation. *Struct Health Monit* 2013; 12: 141–152.

15. Schubert Kabban CM, Greenwell BM, DeSimio MP, et al. The Probability of detection for structural health monitoring systems: repeated measures data. *Struct Health Monit Int J* 2015; 14: 252–264.

16. Meeker WQ, Roach D and Kessler SS. Statistical methods for probability of detection in structural health monitoring. In: *Proceedings of the structural health monitoring 2019*, 15 November 2019. Stanford, California, USA: DEStech Publications, Inc.

17. Roach D. Real time crack detection using mountable comparative vacuum monitoring sensors. *Smart Struct Syst* 2009; 5: 317–328.

18. Kessler SS, Dunn CT, Swindell P, et al. Detection sensitivity analysis for a potential drop (PD) structural health monitoring (SHM) system. In: *Proceedings of the proceedings of the 12th international workshop on structural health monitoring*, Stanford, CA, USA, 10 September 2019. DEStech Publications, Inc.

19. Foucher F, Fernandez R, Leberre S, et al. *New tools in CIVA for model assisted probability of detection (MAPOD) to support NDE reliability studies*, In: NDE of Aerospace Materials & Structures, Missouri, United States: ASNT. June 5 2018, p. 12.

20. Foucher MF, Lonne S, Toullelan G, et al. *An overview of validation campaigns of the CIVA Simulation Software 9*. In: 12th European Conference on Non-Destructive Testing (ECNDT 2018), Gothenburg, Sweden: NDT.net.

21. Falcetelli F, Yue N, Di Sante R, et al. Probability of detection, localization, and sizing: the evolution of reliability metrics in structural health monitoring. *Struct Health Monit*. Epub ahead of print 27 December 27 2021. DOI: 10.1177/14759217211060780.

22. Mariani S, Rendu Q, Urbani M, et al. Causal dilated convolutional neural networks for automatic inspection of ultrasonic signals in non-destructive evaluation and structural health monitoring. *Mech Syst Signal Process* 2021; 157: 107748.

23. Tschoke K, Mueller I, Memmolo V, et al. Feasibility of model-assisted probability of detection principles for structural health monitoring systems based on guided waves for fibre-reinforced composites. *IEEE Trans Ultrason Ferroelectr Freq Control* 2021; 68: 3156–3173.

24. Mariani S and Cawley P. Change detection using the generalized likelihood ratio method to improve the sensitivity of guided wave structural health monitoring systems. *Struct Health Monit*. Epub ahead of print 31 December 2020. DOI: 10.1177/1475921720981831.

25. Yue N and Aliabadi MH. Hierarchical approach for uncertainty quantification and reliability assessment of guided wave-based structural health monitoring. *Struct Health Monit*. Epub ahead of 28 August 2020. DOI: 10.1177/1475921720940642.

26. Liu C, Dobson J and Cawley P. Efficient generation of receiver operating characteristics for the evaluation of damage detection in practical structural health monitoring applications. *Proc R Soc Math Phys Eng Sci* 2017; 473: 20160736.

27. Moriot J, Quaegebeur N, Le Duff A, et al. Model-based approach for statistical assessment of detection and localization performance of guided wave–based imaging techniques. *Struct Health Monit* 2018; 17: 1460–1472.

28. Gianneo A, Carboni M and Giglio M. Feasibility study of a multi-parameter probability of detection formulation for a lamb waves–based structural health monitoring approach to light alloy aeronautical plates. *Struct Health Monit Int J* 2017; 16: 225–249.

29. Leung MSH and Corcoran J. Evaluating the probability of detection capability of permanently installed sensors using a structural integrity informed approach. *J Nondestruct Eval* 2021; 40: 82.

30. Roach D. Use of comparative vacuum monitoring sensors for automated, wireless health monitoring of bridges and infrastructure. In: *Proceedings of the proceedings of the 9th international conference on bridge maintenance, safety and management*, Melbourne, Australia: CRC Press. July 2018.

31. Pollock AA, Thompson DO and Chimenti DE. *A pod model for acoustic emission—discussion and status*. Rhode Island: Kingston, 2010, pp. 1927–1933.

32. Sause MGR, Linscheid FF and Wiehler M. An experimentally accessible probability of detection model for acoustic emission measurements. *J Nondestruct Eval* 2018; 37: 17.

33. Grooteman FP. Damage detection and probability of detection for a SHM system based on optical fibres applied to a stiffened composite panel. In: *Proceedings of the proceedings of the 25th international conference on noise and vibration engineering, ISMA2012 in conjunction with the 4th international conference on uncertainty in structural dynamics, USD 2012*, Leuven, Belgium, 17 September 2012, vol. 4, pp. 3317–3330. Katholieke Universiteit Leuven.

34. Sbarufatti C and Giglio M. Performance qualification of an on-board model-based diagnostic system for fatigue crack monitoring. *J Am Helicopter Soc* 2017; 62: 1–10.

35. Rummel WD. Recommended practice for demonstration of nondestructive evaluation (NDE) reliability in aircraft parts. *Mater Eval* 1982; 40: 922–932.

36. SAE International. *Aerospace recommended practice: guidelines for implementation of structural health monitoring on fixed wing aircraft* ARP6461, 2013.

37. Falcetelli F, Cristiani D, Yue N, et al. Probability of delamination detection for CFRP DCB specimens using rayleigh distributed optical fiber sensors. In: Rizzo P and Milazzo A (eds) *European workshop on structural health monitoring*. *Lecture Notes in Civil Engineering*. Cham: Springer International Publishing, 2023, vol. 254, pp. 917–926.

38. Meeker WQ, Hahn GJ and Escobar LA. *Statistical intervals: a guide for practitioners and researchers*; Wiley series in probability and statistics. 2nd ed. Hoboken, New Jersey: Wiley, 2017.

39. O'Connor E. Quantifying method differences in predicting the probability of detection for structural health monitoring applications. master of science (MS) in statistics. Ames, Iowa, United States: Iowa State University, 2019.

40. Owen DB. A survey of properties and applications of the noncentral T-distribution. *Technometrics* 1968; 10: 445.

41. D30 Committee, *Test Method for Mode I Interlaminar Fracture Toughness of Unidirectional Fiber-Reinforced Polymer Matrix Composites*. Pennsylvania, United States: ASTM International.

42. Hexcel HexPly® 8552 UD Carbon Prepregs, https://www.hexcel.com/user_area/content_media/raw/HexPly_8552_eu_DataSheet.pdf (accessed on 3 February 2022).

43. 3M^TM Scotch-Weld^TM Structural Epoxy Adhesive EC-9323 B/A, 1 L, Redorange, https://www.3m.co.uk/3M/en_GB/p/d/b40066445/ (accessed on 4 February 2022).

44. FBGS Technologies GmbH ORMOCER® Coating, https://fbgs.com/technology/ormocer-coating/ (accessed on 4 February 2022).

45. Luna Innovations Inc. *Optical Distributed Sensor Interrogator Model ODiSI-B: User's Guide*. Roanoke, VA, USA, 2017.

46. Soller BJ, Wolfe MS and Froggatt ME. Polarization resolved measurement of Rayleigh backscatter in fiber-optic components. In: *Proceedings of the OFC technical digest*, 2005. Anaheim, California United States: OSA Publishing.

47. Kreger ST, Gifford DK, Froggatt ME, et al. High resolution distributed strain or temperature measurements in single- and multi-mode fiber using swept-wavelength interferometry. In: *Proceedings of the optical fiber sensors*, Cancún, Mexico: OSA, 2006, p. ThE42.

48. Lu P, Lalam N, Badar M, et al. Distributed optical fiber sensing: review and perspective. *Appl Phys Rev* 2019; 6: 041302.

49. Ripling EJ, Mostovoy S and Corten HT. Fracture mechanics: a tool for evaluating structural adhesives. *J Adhes* 1971; 3: 107–123.

50. Škec L, Alfano G and Jelenić G. Complete analytical solutions for double cantilever beam specimens with bi-linear quasi-brittle and brittle interfaces. *Int J Fract* 2019; 215: 1–37.

51. Product Information – TB1742, https://threebond-europe.com/products/tb1742/ (accessed on 4 February 2022).

52. Sans D, Renart J, Costa J, et al. Assessment of the influence of the crack monitoring method in interlaminar fatigue tests using fiber bragg grating sensors. *Compos Sci Technol* 2013; 84: 44–50.

53. Stutz S, Cugnoni J and Botsis J. Studies of mode I delamination in monotonic and fatigue loading using FBG wavelength multiplexing and numerical analysis. *Compos Sci Technol* 2011; 71: 443–449.