Submitted: December 4, 2025 **Revised: February 16, 2026** **Accepted: March 18, 2026**

Topological data analysis and graph signal processing: quantitative defect assessment and localization of structural inhomogeneities in composites from nondestructive testing data

A.I. Borovkov [1] ✉, Kh.M. Vafaeva [1] ✉, N.I. Vatin [1], Zh.S. Nuguzhinov [2]

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

[2] Kazakhstan Multidisciplinary Institute of Reconstruction and Development Republican State Enterprise on the Right of Economic Use, Karaganda, Republic of Kazakhstan

✉ vafaeva.khm@gmail.com

ABSTRACT

The aim of the study is to develop and perform a preliminary evaluation of a method for diagnosing and locating internal defects in multilayer polymer composite materials using non-destructive testing data. A method combining topological data analysis and graph-signal processing is proposed. To demonstrate the potential feasibility of this method, an analysis of synthetic ultrasonic signals was conducted. A set of 81 ultrasonic signals, representing various locations of internal defects in the composite material, was used. Topological data analysis enabled the identification of informative topological features and distinguished defective from defect-free cases with a silhouette score of 0.471. The results show the potential of using topological data analysis to automate the structural monitoring of internal defects in multilayer polymer composite materials.

KEYWORDS

topological data analysis • graph signal processing • nondestructive testing • composite materials structural inhomogeneities • defect localization • signal interpretation • structural health monitoring

Funding. *This work has been supported by the grant of the Russian Science Foundation (grant No. 24-19-00691).*

Citation: Borovkov AI, Vafaeva KhM, Vatin NI, Nuguzhinov ZhS. Topological data analysis and graph signal processing: quantitative defect assessment and localization of structural inhomogeneities in composites from nondestructive testing data. *Materials Physics and Mechanics.* 2026;54(2): 140–166. http://dx.doi.org/10.18149/MPM.5422026_11

Introduction

Multilayer polymer composite materials and products made from them are used in the aerospace industry, energy sector, civil construction, and other fields [1–3]. The wide use of multilayer polymer composite materials is because of their unique mechanical properties. The manufacturing processes for these materials and products can create hidden defects and structural inconsistencies, which need to be monitored during operation [4,5]. Quantitative evaluation of hidden defects in multilayer polymer composite materials and their spatial distribution remains important, especially for enhancing the reliability of engineering structures [6,7].

To identify hidden structural defects in multilayer polymer composite materials, non-destructive testing methods – including ultrasonic, acoustic emission, thermographic, and radiographic techniques – are used, enabling the collection of large data volumes [8–10]. The listed methods indicate the presence of defects, but traditional analysis of

measurement results often relies on simplified physical models of wave or heat propagation in the material, empirical correlations between signal features and defect types, and the operator's subjective interpretation of the data obtained. Consequently, the accuracy of quantitative assessment of defect parameters (such as size, shape, and morphology) and their spatial location can be significantly diminished, especially when dealing with complex defects or environments with high noise levels [11–13].

Recent advances in the micromechanics of heterogeneous crystalline nanostructures have highlighted the role of misfit-stress relaxation via misfit dislocation formation, providing critical insights into defect generation in composite-like systems, including spherical and cylindrical nanostructures with core-shell or Janus configurations [14]. Complementing this theoretical framework, experimental and numerical investigations of the mechanical properties of porous thermoplastics, including polyacrylonitrile-co-butadiene-co-styrene, polycarbonate, and polyetherimide produced via 3D printing, have revealed a gradual accumulation of plastic strain under cyclic loading. Finite element modeling has validated multilinear isotropic hardening laws that account for porosity-induced variations in stress-strain behavior [15]. These findings collectively underscore the need for integrated analytical approaches, such as topological data analysis combined with graph signal processing, to quantitatively assess and localize structural inhomogeneities in advanced composites materials.

Topological data analysis

Advancements in topological data analysis (TDA) provide mathematical principles for extracting shape and connectivity information that is invariant to continuous deformations and resilient to noise [16,17]. The use of TDA extends beyond mechanical properties. TDA has also been effectively applied to solve heat transfer problems. This method improved temperature uniformity, reducing its variation by 57 %, and successfully eliminated local overheating without increasing energy costs for coolant pumping [18].

In [19], a multi-scale topology optimization method was developed to reduce the vibration response of cellular composites within a specific frequency range. To lower computational costs at the microlevel, Kriging metamodels were employed, enabling quick prediction of the properties of various cell types. At the macro level, the SOMMG method was used to accelerate the calculation of the entire structure's dynamic response. The results facilitate the design of complex 2D and 3D composites with improved dynamic properties and lower vibration levels compared to traditional solutions [19].

An existing topology optimization method was successfully extended and adapted to handle complex multilayer and multicomponent structures in [20]. The results of the study confirmed that the TDA method can be used to create a rigid structure or design uniformly stressed parts [20].

Graph signal processing

Graph signal processing (GSP) continues to evolve. GSP tools facilitate the use of relational information and are effective for describing both local and global dependencies within the structure of materials. In [21], a comprehensive review is presented on how graph signal processing methods are used to analyze geometric data, such as 3D point clouds,

demonstrating that GSP is a useful tool for dealing with irregularly structured data.

Building on the same principles of GSP, in [22], it was presented the "Grid-GSP" framework for data analysis in electrical grids, where the authors model the power system as a graph and show that voltage data behaves as a low-frequency signal on the graph. GSP has been successfully applied to practical problems, such as anomaly detection, data compression, and missing measurement recovery, as demonstrated using real-world power system models [22]. In [23], GSP is applied to develop a computationally efficient structural health monitoring (SHM) algorithm, where data from acceleration sensors are treated as signals on a graph, and damage in the bridge structure causes changes in the smoothness of these signals.

Problem definition

Although TDA and GSP have individually shown potential for solving various scientific and engineering problems [23–25], their combined use for the quantitative analysis and high-precision localization of defects in composite materials based on non-destructive testing data remains an underexplored area. This study aims to develop and preliminarily validate a method that combines topological data analysis and graph signal processing for the quantitative assessment and localization of internal defects in multilayer polymer composite materials using non-destructive testing data.

This goal is achieved by tackling the following tasks, which are designed based on the identified research gap:

1. To analyze the principles of data collection and processing used in non-destructive testing procedures for multilayer polymer composite materials and to identify characteristic defect signatures using synthetic ultrasonic signals as an example.
2. To provide a detailed explanation of the theoretical foundations of topological data analysis and graph-signal processing, including the development of graph models for diagnosing multilayer polymer composite materials.
3. Develop a framework that combines topological data analysis with signal-graph processing to extract meaningful topological features of defects and interpret them within the context of structural monitoring.
4. Demonstrate the capabilities of the proposed method using synthetic ultrasound data modeling different delamination positions, followed by an evaluation of how effectively topological features can be extracted and clustered.
5. Identify areas for further research, such as expanding analysis methods to other defect types and different non-destructive testing data formats.

Materials and Methods

The general flow chart of the methods used in this study is shown in Fig. 1. The third phase (Conceptual Stage, Steps 6–8) represents the planned integration of graph signal processing and nonlinear post-processing; these steps have not been implemented in the present study and are explicitly labelled "future work" in Fig. 1. The research methods are divided into six steps, organized into three main phases. A detailed description of these phases is provided below.

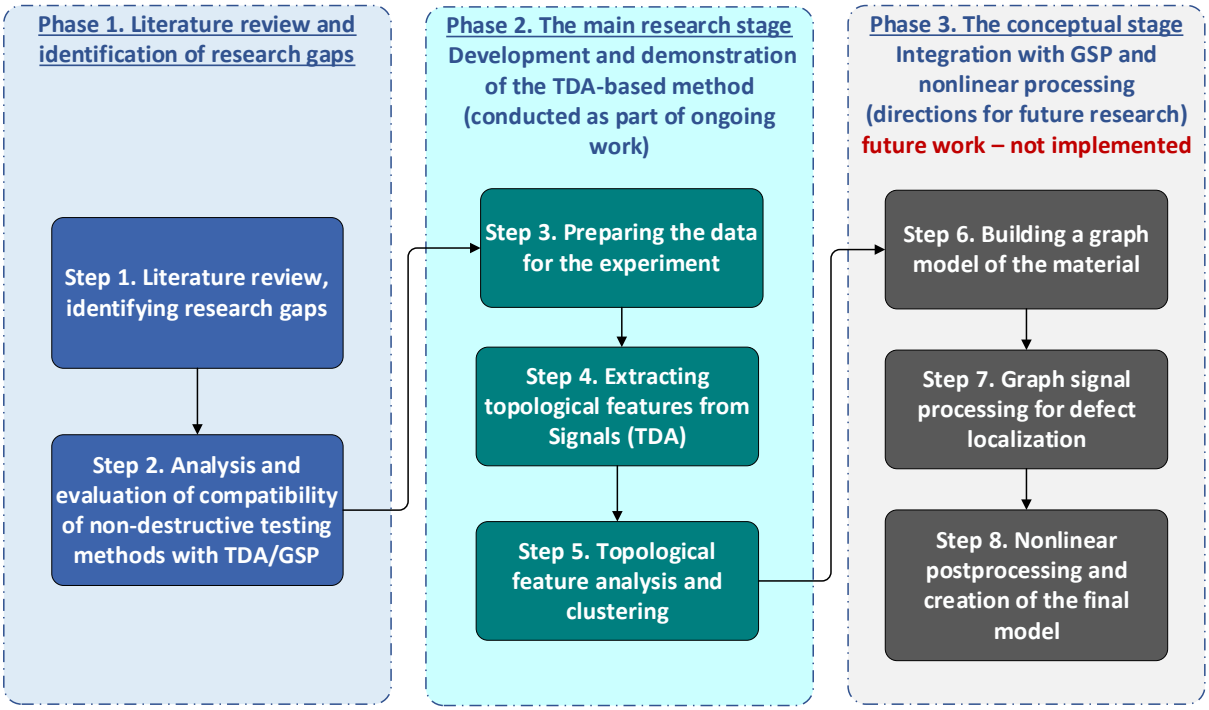

Fig. 1. Research methods flowchart

Preparatory and analytical phase

The preparatory and analytical stage was dedicated to a thorough analysis of the research gap and a justification for the need to develop a new method based on an examination of existing materials publications.

Step 1: Publication analysis and identification of research gaps. This step involved analyzing publications and identifying gaps in prior research. A systematic search and review of scientific literature in the Scopus database was conducted from 2020 to July 2025. Search queries related to topological data analysis (TDA), graph-signal processing (GSP), and non-destructive testing of composites were used. Analyzing 4 072 publications revealed that the combined use of TDA and GSP for the quantitative assessment and localization of defects in composites is an area that is not well understood, confirming the relevance and scientific novelty of the study.

Step 2: Analysis and compatibility evaluation of non-destructive testing methods with TDA/GSP. A comparative analysis of the main non-destructive testing methods (ultrasonic, eddy current, acoustic emission, etc.) was performed. Their operating principles, data formats, and suitability for integration with TDA and GSP were evaluated. It was found that ultrasonic testing (UT), acoustic emission, and infrared thermography are the most promising for integration with TDA/GSP because of their data formats (time series, C-scans, thermal images). These findings are detailed in the following sections.

Main research phase

The primary objective of the main research phase was to demonstrate the functionality of the key component of the proposed method (TDA) by using an example of classifying

defective and non-defective signals.

Step 3: Data preparation for the numerical experiment. Synthetic ultrasound data from the KU Leuven RDR dataset [26,27] were used to demonstrate the method. The dataset includes 81 ultrasound signals that simulate various delamination locations within a composite material.

Step 4. Topological feature extraction from signals (TDA). Each one-dimensional time series was analyzed using TDA with a custom Python script. A detailed description of the topological feature extraction procedure is provided in the "Feature extraction with topological data analysis" section. The process involved the following substeps: the signals were converted into multidimensional point clouds using the Takens method; vietoris-rips complexes were built from the point clouds; stability diagrams showing the "lifetime" of topological features (connected components H0 and cycles H1) were calculated; quantitative topological descriptors (number of points, entropy, and overall stability) were extracted from the diagrams.

Step 5. Analysis and clustering of topological features. The topological feature vectors obtained from all 81 signals were analyzed using the K-means clustering algorithm. The optimal number of clusters was determined through silhouette analysis, achieving a maximum value of 0.471 at $K = 2$. The clustering results were visualized in the principal component analysis (PCA) space.

The successful separation of the signals into two distinct clusters with a silhouette score of 0.471 demonstrates that TDA can extract meaningful topological features suitable for automatic classification of different material states.

Conceptual stage. Future research directions

The goal of this stage is to describe how the developed TDA module will be integrated into a broader system for precise spatial localization of defects. These steps have not been implemented but serve as a plan for future work.

Step 6. Construction of a graph model of the material. Representation of the composite material structure as a graph G=(V,E,W), where vertices are measurement points and edges represent the spatial proximity and/or similarity of signals.

Step 7. Graph-signal processing for defect localization. The topological features obtained in Step 4 will be treated as signals at the graph vertices. Next, GSP tools are used to filter and accurately locate anomaly vertices that correspond to defects.

Step 8. Nonlinear postprocessing and creation of the final model. To enhance accuracy, it is suggested to implement nonlinear processing methods such as wavelets and autoencoders. All features – including topological, graph, and nonlinear – will be combined into a single vector for the final classification and localization of defects.

Results and Discussion

Limitations in quantitative assessment of defects and localization of structural heterogeneities

To identify current limitations in the quantitative assessment of defects and the localization of structural inhomogeneities in composite materials based on non-destructive testing

data, a comprehensive review of existing publications was conducted. The potential application of topological data analysis and graph signal processing for diagnostics and structural integrity monitoring of composite products was also examined analyzed.

A publication search was conducted using the international scientific and analytical database Scopus for the period from 2020 to July 2025. The initial selection of publications was based on the criteria presented in Table 1.

Table 1. Criteria for Selecting Scientific Publications

Selection criteria description	Selection criteria value
Time range	Published between 2020 and 2025 (inclusive)
Publication language	Unrestricted
Source type	Peer-reviewed articles, conference proceedings, reviews
Metadata	Export to RIS format

Using logical operators, search queries were created and applied to the TITLE-ABS-KEY fields. The main search queries and the corresponding number of relevant publications are shown in Table 2. The downloaded data were subsequently combined and analyzed with VOSviewer software.

Table 2. Search results

No.	Search query	Number of documents found
1	Topological data analysis composite	69
2	Graph signal processing composite	30
3	Nondestructive testing composite	1 913
4	Defect localization composite	263
5	Quantitative assessment composite	1 639
6	Structural inhomogeneities composite	158
Total:		**4 072**

Existing non-destructive testing methods

This section offers a brief analysis of the most common non-destructive testing methods for composite materials, including their operating principles, types of detectable defects, data formats, limitations, and an evaluation of their compatibility with topological data analysis and signal-graph processing.

Ultrasonic testing relies on the reflection and scattering of ultrasonic waves as they pass through a material. It allows the detection of internal flaws, such as delaminations, voids, and cracks, by examining variations in the amplitude and timing of reflected signals [28,29]. The data formats for ultrasonic testing results are time series and two-dimensional C-scans. Penetration depth can reach 1 m under ideal conditions (low-attenuating materials), but in composites, it usually does not exceed 50–200 mm.

Eddy current testing employs induced electromagnetic currents to identify surface and subsurface flaws in electrically conductive materials [30–32]. Effective for detecting cracks, corrosion, and other defects at depths of up to 1–10 mm. Eddy current testing data is shown as frequency and amplitude complexes. The method is limited to electrically conductive materials and needs signal transformation to integrate with TDA/GSP.

Acoustic emission (AE) records elastic waves produced during the initiation and

growth of defects under load. It allows for the detection of active cracks, microdefects, and delamination processes throughout the entire thickness of a component; however, precise localization requires a sensor array and triangulation [33–35]. The AE data format consists of a time series of events.

Infrared thermography is a non-destructive testing technique that analyzes the distribution of thermal radiation emitted from a material's surface during localized or uniform heating [36,37]. The technology enables the detection of hidden defects that impact thermal conductivity, such as delamination, pores, voids, and areas of loose contact at depths of 10–20 mm. The data collected is interpreted as two-dimensional or three-dimensional thermograms. The effectiveness of the method depends on external factors like ambient temperature, air flow, and the thermophysical properties of the tested material, especially its thermal characteristics inertia.

Table 3. Brief overview of non-destructive testing methods

NDT method	Operating principle	Defect types	Penetration depth	Data format	Main limitations	TDA/GSP compatibility
Ultrasonic inspection (UT)	Reflection/ scattering of ultrasonic waves	Delamina tions, cracks, voids	Medium to high (up to 1 m, material dependent)	Time series, C-scans	Dependent on input angle, difficulties with curved surfaces	High, as signals lend themselves well to topological and graph analysis
Eddy current Inspection (ET)	Changes in induced current in conduc tors	Surface and subsurface cracks, corrosion	Low (1–10 mm, in conductors)	Frequency/ amplitude complexes	Limited to electrically conductive materials	Limited and requires signal transforma tion into a suitable format
Acoustic emission (AE)	Recording of acoustic waves at defect initiation	Active cracks, delamination growth, microdefects	High (full thickness, but without location without triangulation)	Time series (events)	Difficulty with localization without a sensor array	Spatio-temporal signals are applicable in a graph model
Infrared thermography (IRT)	Thermal response to external excitation	Delamina tions, voids, leaks	Low and medium (10-20 mm)	2D/3D thermal images	Dependent on thermal inertia, noise	Can be analyzed as images through cubic complexes
X-ray/CT	X-ray absorption and trans mission	Internal inclusions, pores, cracks	Very high (up to 100 mm+)	2D/3D voxel images	Expensive, slow, radiation hazardous	Possible through voxel processing and grid graphing

X-ray and computed tomography produce detailed images of a material's internal structure by analyzing how X-rays transmit and are absorbed. They can identify internal inclusions, pores, and cracks at depths of up to 100 mm or more, depending on the material's density and source power [38–40]. The data is in 2D/3D voxel format, which requires significant computing resources and compliance with radiation safety measures.

Despite their widespread use, existing non-destructive testing methods have several significant limitations. They often depend on subjective operator interpretation, struggle to detect small or complex defects, require expensive equipment and highly skilled personnel, and are not always universally suitable for different types of composite materials. To address these issues, a method combining topological data analysis and graph-signal processing is proposed, offering more accurate and objective defect detection.

To evaluate the suitability of existing non-destructive testing methods for TDA/GSP applications, a comparative review was compiled and shown in Table 3. It highlights the key parameters of each method, including defect types detected, penetration depth, data format, limitations, and the capability to integrate with TDA and GSP. Concerning Table 3, it is important to note that the maximum penetration depth is reported under optimal conditions and may vary depending on composite properties such as thickness, density, filler, and binder type).

Table 3 shows that ultrasonic testing, acoustic emission, and infrared thermography are most suitable for integration with TDA and GSP methods because these techniques produce data with the required spatiotemporal features, a regular grid layout, or a continuous intensity scale. This allows for both topological filtering and spectral decomposition on graphs. Meanwhile, radiography and eddy current testing need extra processing steps, like signal transformation, voxel data segmentation, or connected component construction, which make their direct inclusion in a graph model of the material more complicated. However, with proper transformations and graph construction algorithms on a voxel grid, they can also be adapted for the TDA-GSP method.

Collection of non-destructive testing data

Data types and formats. This section provides an analysis of various types of non-destructive testing data that reflect both the temporal and spatial structures of the measured processes, which could be applied with further development of the proposed method. The data types and formats are summarized in Table 4, including the respective measurement types, data presentation, and brief explanations.

Table 4. Data types and formats

Data type	Data format	Description
Ultrasound (UT)	Time series, 2D C-scans	1D: signal amplitude over time (A-scan); 2D: amplitude map over the surface (C-scan)
Eddy current (ET)	Time series, frequency spectra	Induced current signals over time and/or spectral domain
Acoustic emission (AE)	Time series of events	Bursts of pulses (time, amplitude, coordinates); event-based format with detection threshold
Infrared thermography (IRT)	2D/3D thermal images	Spatial temperature distribution on a surface or in a volume
X-ray/CT	2D/3D voxel images	2D X-ray or 3D tomography voxel data

To ensure compatibility with graph-topology processing, time series should be in .csv, .tdms, or .mat formats and contain numerical tables with time and amplitude data. Raster images should be in .png, .tif, or .dcm formats and include two-dimensional and three-dimensional visual data suitable for analysis with cubic complexes. Hybrid datasets can include time signals synchronized with the spatial coordinates of sensors. The use of these formats will provide flexibility, allowing TDA and GSP to be applied to a wide range of signals and images typical for structural applications diagnostics.

Feature extraction with topological data analysis. Topological data analysis offers a formal way to extract invariant structural features that represent the geometry and shape of data across various scales. The main tool of TDA is persistent homology, which provides a quantitative way to describe how topological objects appear and disappear as the filtering scale varies. It examines features that are resilient to noise and allows for comparing complex structures based on their overall topology (Table 5).

Table 5. Topological features

Topological feature	Brief description
0D – dimension 0 (β_0)	Individual clusters or regions that may correspond to local defects, such as pores or inclusions. In the context of Takens embedding of a 1D time series, β_0 reflects the density and fragmentation of points in the reconstructed phase space. High NumPoints_H0 and Entropy_H0 indicate noise-induced fragmentation or amplitude anomalies caused by defects.
1D – dimension 1 (β_1)	Closed contours or loops reflecting structures such as cracks or delamination boundaries.
2D – dimension 2 (β_2)	Spatial voids or inclusions characteristic of large defects in composites.

Applying H0 to Takens-embedded time series allows for a direct assessment of point cloud density and connectivity. As the filtration radius ϵ expands, H0 tracks the transition from isolated points to clusters, effectively capturing noise-induced fragmentation. While a point cloud of length L starts with NumPoints_H0 near L, it is the Entropy_H0 that reveals the underlying uniformity of the distribution. Although H1 remains the primary tool for detecting delamination via wave interference, H0 provides critical context regarding density anomalies. Combining both descriptors ensures the feature vector is robust enough for reliable clustering. The way of extracting topological features depends on the data source (time series or image).

Time series processing. For one-dimensional signals such as ultrasound, eddy currents, and acoustic emission, the following steps are used:

1. Embedding into phase space. The Takens embedding method is used, in which the signal $x(t)$ is transformed into a point cloud: $X_t = [x(t), x(t+t), x(t+2\cdot t), \ldots, x(t+(d-1)t)]$, where $x(t)$ is the original time signal, τ is the time delay, d is the embedding dimension, X_t is the vector in phase space corresponding to time t.

2. After transforming the time signal into phase space (using Takens' method), we obtain a set of points X_t in a dimensional space d. Next, a Vietoris-Rips complex is constructed, which reflects the topological structure of these points to understand which points are connected and how they form geometric structures (line segments, triangles, etc.). For each pair of points X_t and X_j in phase space, the Euclidean distance is calculated:

$d(X_i, X_j) = \sqrt{\sum_{k=1}^{d}(X_i^{(k)} - X_j^{(k)})^2}$, where d is the dimension of the phase space, $X_i^{(k)}$ is k-th coordinate of the point X_i . Two points X_i and X_j are connected by an edge if the distance between them does not exceed the radius ε: $d(X_i, X_j) \leq \varepsilon$.

A set of $k+1$ points forms a k-simplex if all pairs between them are pairwise connected, that is, the distance between each pair does not exceed ε: $d(X_i, X_j) \leq \varepsilon \quad for\ all \quad i < j \quad from \ \{0,1,\ldots,k\}$. The entire complex at a fixed level ε can be denoted as: $VR(X, \varepsilon)$, where X is a set of points, ε is the filtration radius (threshold distance).

Although H1 (cyclicity) is the primary indicator of defects (stable loops appear due to wave interference at delamination), H0 is not redundant. In the Takens-embedded point cloud H0 quantifies sample density: defects change the local amplitude distribution, increasing fragmentation (higher NumPoints_H0) or altering uniformity (Entropy_H0). Thus, the full feature vector {NumPoints_H0, Entropy_H0, NumPoints_H1, Entropy_H1, Amplitude_H0, Amplitude_H1} provides complementary information and improves clustering robustness.

3. Calculating persistence diagrams is a graph that displays all topological objects (components, cycles, etc.) with their birth and death times (depending on ε). Each object corresponds to a point (b, d), where b is the value of ε, at which the object appeared, and d is the value of ε, at which it disappeared. If $d - b$ is large, then the object is considered significant, and if $d \approx b$, then it is most likely noise. That is, the further the point is from the diagonal $b = d$, the more significant the topological structure.

4. After constructing the persistence diagram, a quantitative assessment of the topological features is performed. These numerical characteristics are called topological descriptors and serve as features for analysis or machine learning (Table 6).

Table 6. Topological descriptors

No.	Descriptor	Designation	Description
1	Number of components	β_0	Number of distinct regions in the data
2	Number of cycles	β_1	Number of cycles (loops)
3	Object stability	$Persistence_i = d_i - b_i$	"Lifetime" of a topological object
4	Average stability	$Persistence_{avg} = \frac{1}{N}\sum_{i=1}^{N}(d_i - b_i)$	Average importance of all objects
5	Diagram entropy	$H = -\sum_{i=1}^{N}(p_i \cdot log(p_i))$	Complexity/diversity of topological structures
6	Object share	$p_i = \frac{d_i - b_i}{\sum_{j=1}^{N}(d_j - b_j)}$	Used in entropy calculations
7	Maximum stability	$Persistence_{max} = max_i(d_i - b_i)$	The longest-lived object
8	Total stability	$Persistence_{total} = \sum_{i=1}^{N}(d_i - b_i)$	The sum of all lifetimes

Image processing with TDA. For 2D and 3D images (e.g., ultrasound C-scans, thermography heat maps, CT tomographic data), topological feature extraction is carried out by constructing cubic complexes.

1. Image preparation (binarization). Before constructing topological objects, the image is converted into a binary mask, where each pixel (or voxel) takes the value 1 (object) or 0 (background): $I(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq T \\ 0, & \text{if } f(x,y) < T \end{cases}$, where $f(x,y)$ is the intensity at an image point, T is the selected threshold, $I(x,y)$ is the binarized image.

2. Construct a cubic complex. Each pixel (in 2D) or voxel (in 3D) with a unit value corresponds to a vertex (0-simplex). Vertices are connected by edges (1-simplexes) if they are adjacent on the grid (e.g., 4-neighborhood or 8-neighborhood in 2D): $K = CubicalComplex(I)$, where K is the set of all simplexes constructed from a binary image I.

3. Filtering by intensity level. To analyze inhomogeneities, filtering is used, i.e., successively lowering the T threshold, which binarizes the image, creating a hierarchy of nested complexes: $K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n$, where K_i is the cubic complex obtained by binarization with the T_i threshold. Each subsequent T_{i+1} level is lower than the previous one: $T_1 > T_2 > \ldots > T_n$.

4. Calculating stability diagrams. For each filtering level, the creation and disappearance of topological objects (clusters, holes, cavities) is tracked. The resulting stability diagram records: β_0 is the number of connected components; β_1 is the number of closed contours (loops); β_3 is the number of voids (for 3D).

5. Similar to time series (Table 6), descriptors are extracted from the image stability diagrams. Image processing using TDA enables the quantitative evaluation of the shape, connectivity, and heterogeneity of the structure, which is especially important for analyzing microporosity, delamination, and cracks observed in thermal images or tomograms.

Construction of a graph model of the material. The construction of the graph model and GSP are described as conceptual steps for analyzing spatial dependencies. Nonlinear processing involves using wavelets and autoencoders but is not implemented in the current study. These steps will be tested in future work (the Conclusion section).

The graph model enables formalizing the structure of a material or system as a mathematical graph that considers spatial and/or functional relationships between measured points, microstructure grains, sensors, or pixels clusters.

1. Each graph vertex corresponds to a specific spatial element: a sensor coordinate, a grain or phase center, an image pixel cluster, or a measurement point (e.g., in ultrasound or CT). The set of vertices is denoted as $V = \{v_1, v_2, \ldots, v_N\}$, where N is the total number of points used to construct the graph.

2. Definition of graph edges. Edges are established between vertices based on spatial proximity (e.g., connecting nearest neighbors) and signal similarity (e.g., correlating signals from different sensors). Distance-based connection condition: $e_{ij} \in E, \quad if\ d(v_i, v_j) \leq \varepsilon$, where e_{ij} is the edge between vertices $v_i$and v_j, $d(v_i, v_j)$ is the Euclidean distance between vertices, ε is the distance threshold.

3. Defining edge weights. An edge weight is a number reflecting the degree of "connection" between two vertices. It can be determined in different ways:

(a) by distance (exponential decay): $w_{ij} = exp\left(-\frac{d(v_i,v_j)^2}{\sigma^2}\right)$, where σ is the scaling parameter (the width of the Gaussian);

(b) by correlation between signals: $w_{ij} = corr(s_i, s_j)$, where s_i, s_j are the signals measured at the vertices i, j;

(c) by topological similarity: $w_{ij} = exp\left(-\frac{||T_i - T_j||^2}{\gamma^2}\right)$, where T_i, T_j are the vectors of topological descriptors (e.g., β_0, β_1, entropy) and γ is the scaling parameter.

4. A weighted undirected graph $G = (V, E, W)$ is formed, where V is the set of vertices, E is the set of edges, and W is the edge weight matrix.

Graph-signal processing. The GSP method enables analysis of data not on a regular grid, like in traditional signal processing, but on graph vertices. In non-destructive testing, this could include signal amplitudes from sensors, topological descriptors, or energy parameters derived from images or signals over time series. A graph signal is a function defined on the vertices of a graph $f = [f(v_1), f(v_2), \dots, f(v_N)]^T$, where $f(v_i)$ is the signal value at a vertex v_i, and N is the number of vertices.

The graph Laplacian is a key matrix describing the structure of the graph $L = D - W$, where D is the diagonal degree matrix of $D_{ii} = \sum_j w_{ij}$; W is the edge weight matrix of $W_{i_j} = w_{ij}$.

1. Spectral decomposition of the signal. The graph signal f is decomposed in terms of the graph's eigenvectors (analogous to the Fourier transform): $Lu_k = \lambda_k u_k$ $f = \sum_{k=1}^{N} \hat{f}(\lambda_k) u_k$, where λ_k are the eigenvalues (graph "frequencies"), u_k are the eigenvectors (graph "harmonics"), $\hat{f}(\lambda_k)$ are the spectral coefficients of the signal.

2. Graph filtering. Filtering is used to suppress noise or highlight abrupt changes (anomalies). A low-pass filter (signal smoothing) is defined by: $f_{low} = \sum_{\lambda_k \le \lambda_c} \hat{f}(\lambda_k) u_k$, and a high-pass filter (amplifying abrupt transitions): $f_{high} = \sum_{\lambda_k > \lambda_c} \hat{f}(\lambda_k) u_k$, where λ_c is the cutoff threshold.

3. Anomaly detection. Localization by graph derivative. Nodes with abrupt signal changes relative to their neighbors can be found using the graph gradient: $\nabla_G f = Lf$, where $(Lf)_i$ indicates how much the value at node v_i differs from its neighbors. High values indicate anomalies or defects.

In practice, GSP is implemented using open-source libraries. PyGSP handles spectral filtering and graph analysis, NetworkX allows for graph construction and manipulation, and SciPy and NumPy are used for linear algebra and eigenfactorization. Scikit-learn is employed for clustering and feature-based machine learning tasks.

Nonlinear post-processing. The final stage of data processing involves using a combination of nonlinear methods to extract meaningful features, reduce noise, and organize signal information. One fundamental method is the wavelet transform, which enables signal decomposition across the time-frequency spectrum, preserving details about both overall and local features. After transformation, insignificant coefficients are set to zero, and the signal is reconstructed with less noise.

A method utilizing autoencoders (specialized symmetric neural networks) trained to reconstruct input data is also proposed. After training, the central layer (bottleneck) contains a compressed representation of features, free from noise and redundant information, which can be used as input for subsequent processes analysis.

The final step involves combining all the extracted features into a single vector, normalizing them, and inputting them into machine learning or visualization algorithms. This helps ensure compatibility across different data types (signals, images, graphs) and enhances the interpretability of the results. An overview of the proposed methods and key steps for their practical implementation are provided in Table 7.

Table 7. Stages of nonlinear post-processing of signals and images

No.	Method	Objective	Main steps	Tools
1	Wavelet transform	Extracting local signal features	1. DWT transform of the signal 2. Zeroing out small coefficients 3. Inverse DWT transform	PyWavelets, SciPy
2	Local PCA	Dimensionality reduction in a local neighborhood	1. Forming a local window around each point 2. PCA in this region 3. Using 1–2 principal components as new features	scikit-learn, NumPy
3	Autoencoders	Noise removal and compact coding	1. Training a network to reconstruct input features 2. Extracting features from the bottleneck layer 3. Using them as output	Keras, PyTorch, TensorFlow
4	Feature fusion	Preparing data for ML/clustering	1. Normalizing all features 2. Combining into a single vector 3. Clustering or classification	NumPy, Scikit-learn

Results of the literature review. A review and analysis of the literature in the Scopus database (2020 – July 2025) identified key trends and current limitations in the non-destructive testing of composite materials. The total number of relevant publications was 4 072, as shown in Table 2, based on the search queries.

Figure 2 displays word clouds generated from publication abstracts and keywords. The figure on the left shows the most common keywords, such as "Non-destructive testing", "Non-destructive examination", "Carbon fiber", "Composite material", "Ultrasonic testing", "Fiber reinforced", and "Damage detection". This highlights the main research areas in the field of non-destructive testing of composite materials, focusing on methods and research objects.

Fig. 2. Word clouds for annotations and keywords

The figure on the right, a word cloud of abstracts, displays terms from the abstracts, including "method", "material", "structure", "composite", and "inhomogeneity", highlighting the focus on methods, material properties, structural analysis, and inhomogeneity issues composites.

Figure 3 illustrates the distribution of publications across different fields of knowledge. The highest number of publications is in "Engineering" (over 2,000), followed by "Materials Science" (around 1 500), and "Physics and Astronomy" (around 1,000), highlighting the interdisciplinary nature of research in nondestructive testing of composites, with a strong focus on engineering and materials science.

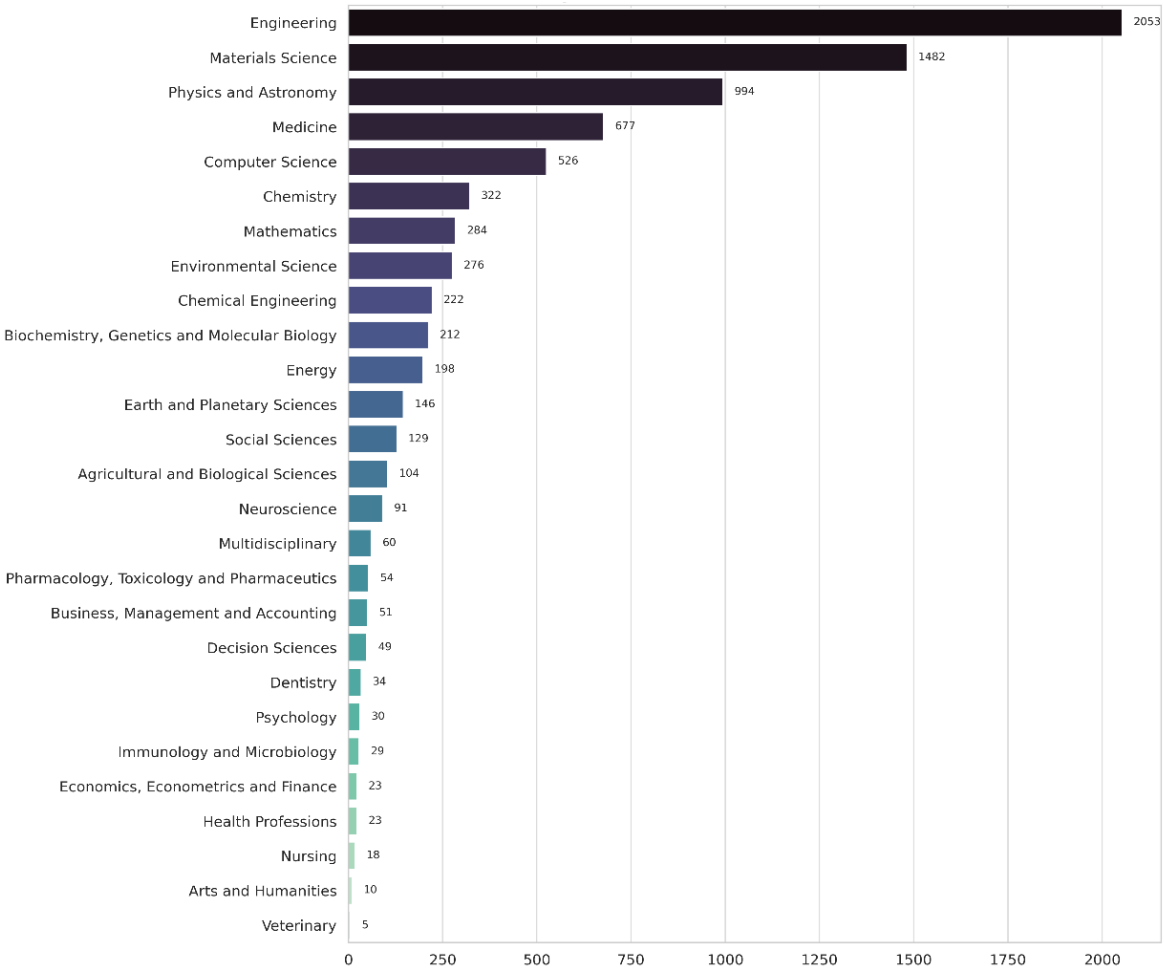

Fig. 3. Distribution of publications by industry

Figure 4 illustrates the trends in the top 15 keywords from 2020 to 2025. A significant rise in publications is observed from 2021 to 2023, followed by a decline in 2024 and 2025 (due to incomplete data for 2025). Among the most common keywords are "Carbon composites", "Carbon fiber reinforced plastics", "Composites material", "Damage detection", "Deep learning", "Defect detection", "Delamination", "Non-destructive", "Non-destructive testing", "Non-destructive examination", "Thermography (imaging)", "Ultrasonic testing", and "Ultrasonic waves", showing steady interest in various aspects of non-destructive testing and defect diagnostics in composites, as well as the increasing importance of machine learning in this field.

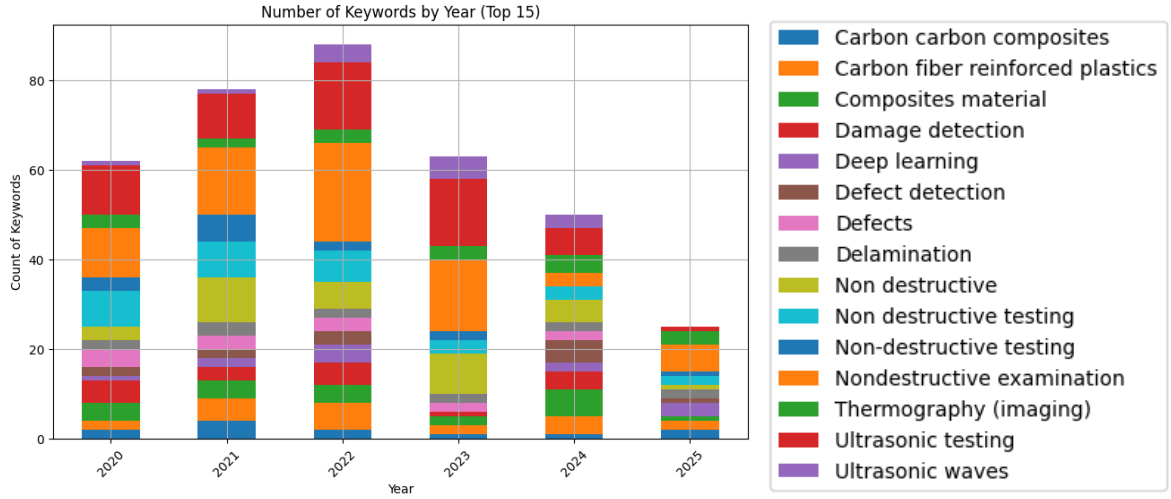

Fig. 4. Top 15 most frequently used keywords

Figure 5 shows the total number of publications by year in more detail. It is clear that the number of publications rose sharply from 2020 to 2023, peaking in 2023 with over 750 publications. A decline appears in 2024 and 2025, which may be due to incomplete data for 2025.

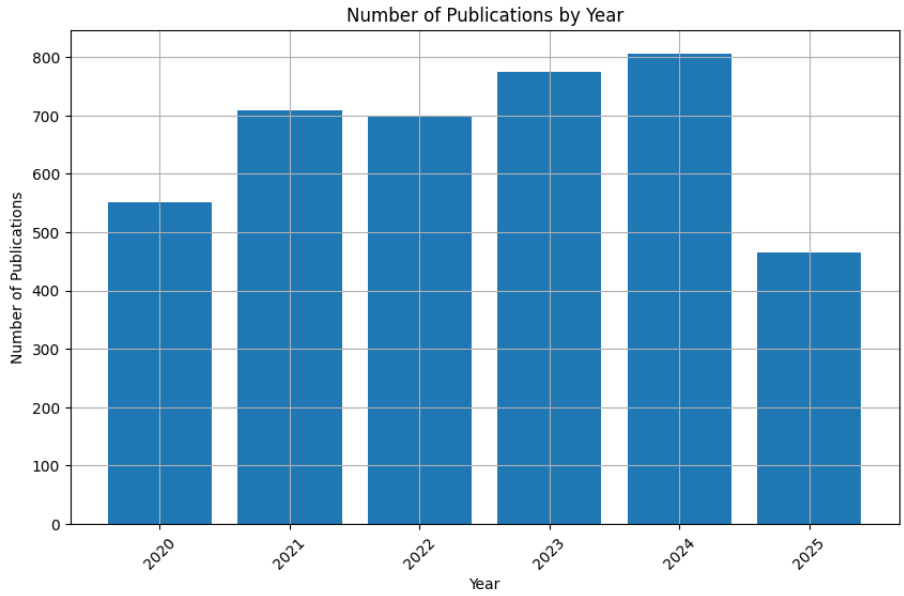

Fig. 5. Dynamics of the number of publications by year

Figure 6 displays the results of a keyword cluster analysis conducted with VOSviewer. Three main clusters (Cluster 1, Cluster 2, Cluster 3) are visible in the image. The analysis identified four major thematic clusters, representing research areas in non-destructive testing and composite material analysis. The first cluster, including terms like "composite materials", "carbon fiber", "reinforced plastics", and "delamination", focuses on composite properties and defects. The second cluster, with keywords such as "ultrasonic testing", "acoustic emission", and "defect detection", highlights non-destructive testing techniques. The third cluster, containing "deep learning", "feature extraction", and

"convolutional neural networks", shows the use of artificial intelligence in data analysis. The fourth cluster, with terms like "compression testing", "compressive strength", and "mechanical failure", concentrates on the mechanical properties and testing of materials.

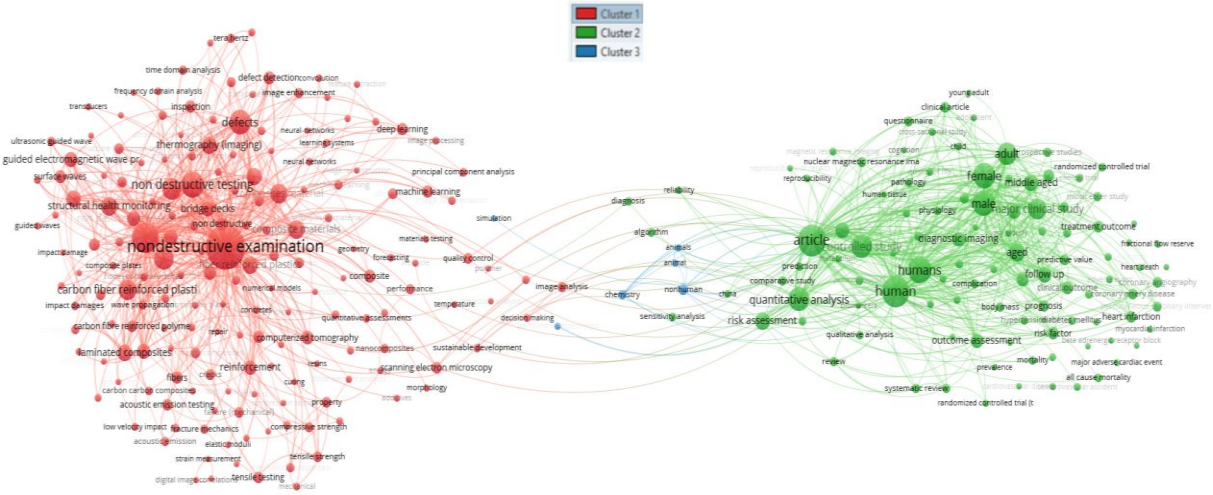

Fig. 6. General view of analytics results in VOSviewer

The relationship between the clusters illustrates how testing methods, data analysis, and composite property studies work together to enhance reliability and efficiency in various industries. An analysis of the selected publications identified some limitations of existing nondestructive testing and data processing methods:

1. Topological data analysis and graph signal processing methods are virtually absent in current research on nondestructive testing of composites. Only 69 papers on TDA and 30 papers on GSP in the context of composites have been published in the past five years. This highlights a significant research gap and untapped potential for combining TDA and GSP methods to address diagnostics and structural integrity monitoring challenges. Most of this work remains primarily theoretical nature.
2. Most work on nondestructive testing data analysis depends on traditional signal processing techniques (e.g., spectral analysis, PCA, standard filters) and machine learning methods. While these approaches yield certain results, they often lack interpretability, particularly when dealing with high-dimensional and noisy data typical of nondestructive testing.
3. Quantitative defect assessment methods based on traditional approaches show low noise immunity and high sensitivity to model parameters, which diminishes their reproducibility and reliability in real-world applications where non-destructive testing data always contains some level of noise.

Demo example on synthetic data

To partly demonstrate the effectiveness and advantages of the developed method, a series of experiments was conducted on synthetic and real non-destructive testing data. The data used to showcase the performance of the proposed method is shown in Table 8. The experiments involved synthetic ultrasound data ("Averaged 10dB normalized delamination damage position at 0 to 1.csv", KU Leuven RDR), including 81 signals with

normalized delamination positions [0, 1] and an SNR of 10 dB [26]. Each signal consists of 1122 points normalized to the amplitude range [-1, 1]. Graph signal processing and nonlinear post-processing were not implemented in the present study.

Table 8. Data sets used for demonstration

Dataset	Type	Source
KU Leuven RDR – delamination [26]	Ultrasound	Lu, Houyu; Cantero Chinchilla, Sergio; Yang, Xin; Gryllias, Konstantinos; Chronopoulos, Dimitrios, 2024, «Multiple ultrasonic-guided wave signals generated by wave propagation model: focused on delamination damage location variations in cross-ply composite beams», *https://doi.org/10.48804/JAIG58*, KU Leuven RDR, V1

Topological data analysis. Visualization and results. Topological data analysis was used to extract invariant structural features that reflect the geometry and shape of the non-destructive testing data across multiple scales, as described in the "Feature extraction with topological data analysis" section.

This section provides diagnostic plots for several of the studied signals, illustrating their main characteristics. Each set of plots includes a signal time series that shows how the signal amplitude changes over time; an amplitude histogram that displays the distribution of amplitude values to estimate frequently occurring levels; and a power spectrum that illustrates how the signal's power is distributed across different frequencies. Peaks in the power spectrum highlight dominant frequency components, which are important for identifying hidden periodicities and evaluating the overall frequency content.

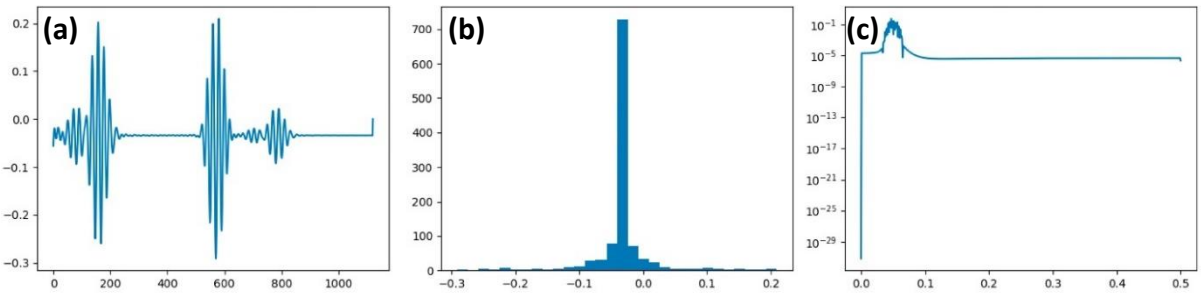

Fig. 7. Signal 1 diagnostics: (a) time series; (b) histogram; (c) power spectrum

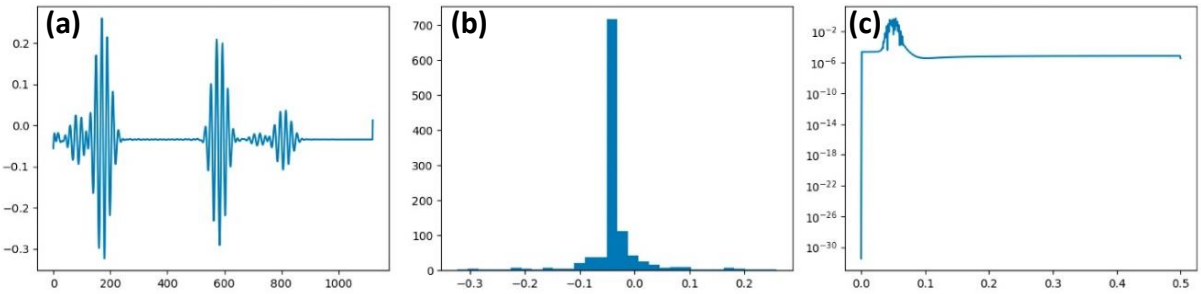

Fig. 8. Signal 2 diagnostics: (a) time series; (b) histogram; (c) power spectrum

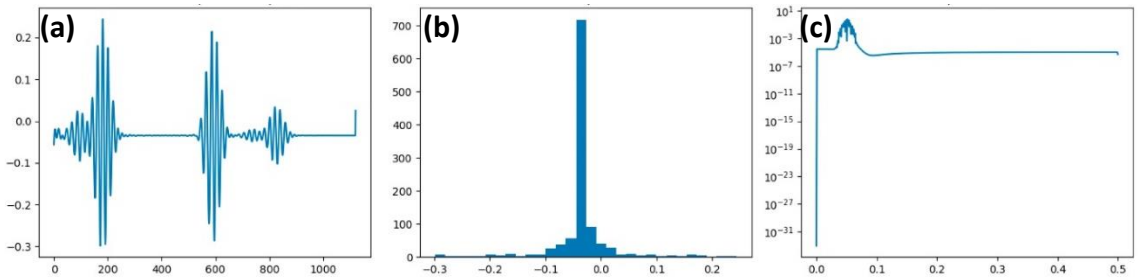

Fig. 9. Signal 3 diagnostics: (a) time series; (b) histogram; (c) power spectrum

All input data for TDA had the shape (81, 1122), meaning 81 individual time series, each consisting of 1122 data points. Diagnostic plots (time series, amplitude histogram and power spectrum) for three representative ultrasonic signals are shown in Fig. 7–9. The corresponding diagnostic descriptions are given in Tables 9–11.

A preliminary analysis of the time series, histograms, and power spectra shows that the signals 2 and 3 have similar characteristics, indicating a more uniform state, while the signal 1 displays greater high-frequency content activity.

Table 9. Signal 1 diagnostics

Time series	Signal 1 shows an oscillating, noise-like pattern without sharp anomalies at this scale. The amplitude varies roughly from -0.25 to 0.2. The lack of strong peaks or dips suggests a relatively stable material or background noise in the absence of a defect.
Histogram	The signal amplitude distribution is almost symmetrical, with most values near zero, which is typical of random noise or signals oscillating around a mean value. The histogram shape shows no bimodality or strongly shifted peaks, which could indicate obvious defects or consistent signal offsets.
Power spectrum	The power spectrum shows a fairly broad frequency range with several noticeable but not dominant peaks at lower frequencies. The lack of very high peaks at specific frequencies suggests the signal does not have strong periodic components, which might be typical of background noise or a complex mix of weak oscillations.

Table 10. Signal diagnostics 2

Time series	The time series of Signal 2 also shows an oscillatory pattern, but with a slightly lower amplitude compared to Signal 1 (roughly from -0.2 to 0.15). Visually, it looks smoother or has less high frequency than Signal 1.
Histogram	The histogram, like Signal 1, is centered around zero and has a symmetrical distribution, but with a more prominent peak, suggesting a smaller spread of amplitude values and possibly less random noise.
Power spectrum	The power spectrum of Signal 2 shows significantly less overall power than Signal 1, especially at higher frequencies, confirming the visual observation of a smoother signal and the lack of pronounced high-frequency components often linked to sudden changes or impulsive events.

Table 11. Signal 3 diagnostics

Time series	Signal 3 is an oscillatory time series, similar in amplitude and overall appearance to Signal 2 (ranging from approximately -0.2 to 0.15). It also shows no obvious sharp anomalies or high-amplitude pulses.
Histogram	The histogram of Signal 3 shows a peak near zero and looks symmetrical, indicating a similar statistical distribution of amplitudes.
Power spectrum	The power spectrum of Signal 3 shows relatively low overall power and no distinct peaks at high frequencies, similar to Signal 2, which may suggest similar underlying processes.

For each data time series, a phase-space embedding step was performed using the Takens embedding method. The embedding dimension and time delay $\tau = 10$ were chosen as the embedding parameters, allowing the one-dimensional signals to be transformed into multidimensional point clouds. Vietoris-Rips complexes were then constructed for the resulting point clouds, which serve as the basis for calculating topological characteristics. The corresponding stability diagrams (persistence diagrams) for these three signals are presented in Fig. 10.

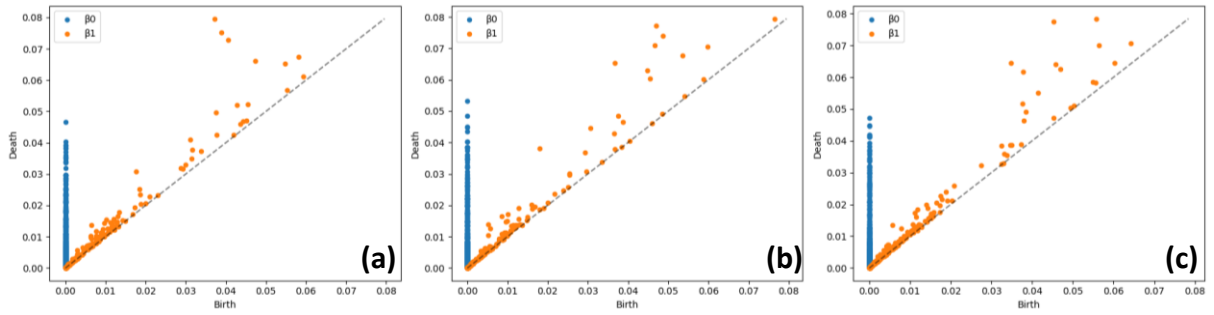

Fig. 10. Stability diagrams of three signals: (a) signal 1; (b) signal 2; (c) signal 3

Table 12. Stability diagrams

Signal stability diagram number	**β_0 (connected components, blue dots)**	**β_1 (cycles, orange dots)**
Signal stability diagram 1	A large number of blue dots are visible, located close to b = d. This indicates numerous short-lived connected components that quickly form and disappear with a slight increase in the epsilon radius. Such dots typically correspond to noise fluctuations or minimal, transient data features that are not structurally significant. However, there are also individual β_0 dots with longer lifetimes (far from the diagonal), which may indicate stable, well-separated clusters inherent in the signal structure.	The presence of many orange dots with relatively long lifetimes (far from the diagonal) is a key feature of this diagram. These dots depict stable closed loops in the reconstructed phase space of the signal. Their presence suggests complex topological structures that could indicate defects such as cracks or delaminations, which form loops in the signal's phase space due to nonlinear interaction dynamics. One of these dots has considerably greater stability (a large distance from the diagonal) highlighting a highly significant topological cycle.
Signal stability diagram 2	Similar to signal 1, short-lived β_0 components close to the diagonal dominate. The number of more stable β_0 points is presumably lower than for signal 1, which may reflect less clustering or a more uniform distribution of points in phase space.	The number of orange β_1 points in this diagram is significantly smaller compared to signal 1, and their distance from the diagonal is also smaller. This indicates that fewer stable cycles are formed in the phase space of signal 2. The reduced number and significance of cycles are consistent with the hypothesis that this signal arises from a less damaged region.
Signal stability diagram 3	The distribution of β_0 points is similar to signal 2, indicating a predominance of noise-like or short-term connected components.	As with signal 2, the number and lifetime of β_1 points in this diagram are low, confirming that signal 3 also contains fewer stable topological cycles.

A comparative analysis of the stability diagrams of the three signals demonstrates that signals related to defective regions are characterized by the presence of stable first order β_1 cycles with a long lifetime. These cycles correspond to closed anomalies in the reconstructed signal dynamics, which is a powerful indicator of damage. A detailed description of the stability diagrams is given in Table 12.

Signals from defect-free zones (represented by Signals 2 and 3) predominantly contain short-lived zero-order β_0, components reflecting noise fluctuations and significantly fewer or no stable β_1 cycles, providing direct evidence of TDA's ability to effectively differentiate different material states based on their internal topological structure.

Based on the calculated stability diagrams, quantitative topological descriptors were derived. These descriptors provide a numerical representation of the topological properties of the data and are used as features for further analysis or machine learning (Table 6). The first five rows of the extracted topological descriptor matrix are shown in Table 13.

Table 13. Topological descriptors (first five rows)

NumPoints_H0	NumPoints_H1	Entropy_H0	Entropy_H1	Amplitude_H0	Amplitude_H1
0	1101	307	8.597069	0.135013	0.038029
1	1101	299	8.546403	0.147386	0.036210
2	1101	308	8.539554	0.145726	0.035309
3	1101	289	8.491680	0.131108	0.033008
4	1101	268	8.560095	0.129314	0.038186

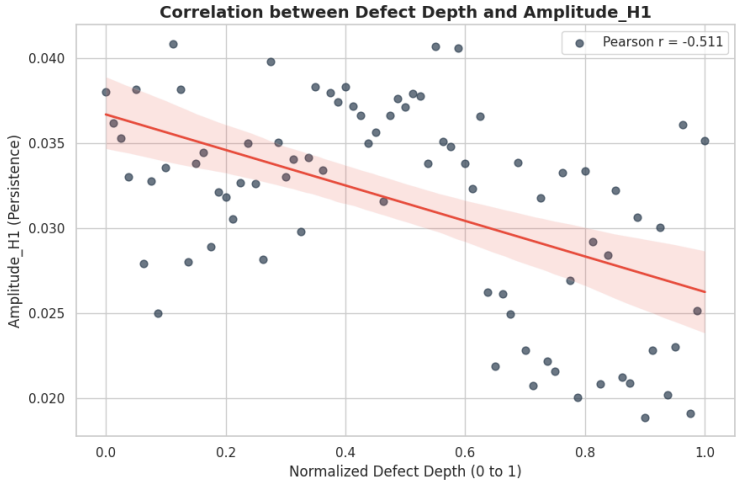

Fig. 11. Correlation between normalized delamination position (defect depth) and Amplitude_H1 topological descriptor (Pearson r ≈ -0.51, p < 0.001). Blue points – individual signals; red line – linear regression trend

The dependence is visualized in Fig. 11. Analysis of the KU Leuven RDR dataset showed a moderate negative correlation (Pearson $r \approx -0.51$, $p < 0.001$) between the normalized delamination position (0–1, interpreted as relative defect depth along the beam) and the topological descriptor Amplitude_H1 (total lifetime of persistent H1 cycles). This indicates that shallower delaminations (lower position values) produce longer-lived topological cycles in the reconstructed phase space. Physically, this is explained by stronger nonlinear scattering and mode conversion at the defect interface when the delamination is closer to the excitation/sensor surface; deeper positions lead

to greater attenuation and weaker phase-space distortion, shortening the lifetime of the detected H1 loops. Thus, the Amplitude_H1 metric not only distinguishes defective from defect-free signals but also carries quantitative information about defect depth, which will be crucial for future spatial localization with graph signal processing. A detailed interpretation of all topological descriptors used in the analysis is provided in Table 14.

Table 14. Topological descriptors

Descriptor	**Brief description of the topological descriptor**
NumPoints_H0	The number of connected components detected in 0-dimensional homology. In the context of Takens embedding applied to time series, this descriptor primarily reflects the density of points in the reconstructed phase space. It begins with a value close to the number of embedded points L (as each point initially forms its own component) and thereby provides a measure of noise-induced fragmentation versus underlying signal structure. Higher values typically indicate finer-scale clustering attributable to noise or detailed signal features, while lower values after merging point to more coherent topological organization.
NumPoints_H1	The number of cycles (holes) detected in 1-dimensional homology is a key feature for identifying defects like cracks or delaminations, which appear as stable closed structures in phase space. The higher this value, the more such structures there are.
Entropy_H0 Entropy_H1	Entropy of the stability diagram for 0th and 1st homology, respectively. High entropy indicates a broad range of «lives» (stability) for topological objects, which may result from the complexity or chaos of the signal's internal structure. For defective regions, increased entropy is expected in H_1, reflecting the diversity and complexity of the defects.
Amplitude_H0 Amplitude_H1	Stability amplitude (or total stability) for 0th and 1st homology, respectively. These values reflect the overall significance of all topological objects in a given measurement. High amplitude, especially for H_1, indicates the presence of strong, long-lasting topological features characteristic of pronounced defects.

Table 14 shows that NumPoints_H0 for the first five signals remains constant (1101), likely representing the number of data points in each time series. More important for classification are variations in NumPoints_H1, Entropy_H0, Entropy_H1, Amplitude_H0, and Amplitude_H1. For instance, signal 0 has NumPoints_H1=307 and Amplitude_H1=0.038029, whereas signal 4 has NumPoints_H1=268 and Amplitude_H1=0.038186. These detailed differences in topological features can be used to accurately distinguish material states.

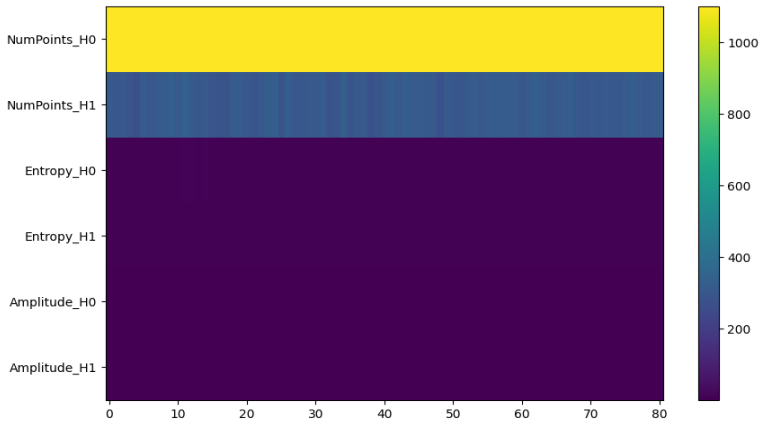

Fig. 12. Heat map of topological features

The topological feature heatmap, which visualizes the values of each of the six extracted descriptors for all 81 signals under study, is shown in Fig. 12. The x-axis represents the signal indices, and the y-axis displays the names of the topological features. The color scale (ranging from blue to yellow/light green) indicates the normalized feature values. Brighter colors (yellow, light green) represent high values, while darker colors (violet, blue) signify low values.

The heatmap shows horizontal bands and vertical patterns. Each row represents a specific topological feature, and each column represents a specific signal. Visually, groups of signals (columns) with similar topological characteristics can be identified. For example, the NumPoints_H1, Entropy_H1, and Amplitude_H1 features (bottom three rows) display more noticeable variations and contrasting areas compared to NumPoints_H0 and Entropy_H0. Although H1 features dominate the separation (see stability diagrams), H0 descriptors also contribute: defective signals show higher variation in NumPoints_H0 and Entropy_H0 (visible as brighter vertical bands in the upper rows of the heatmap).

If we assume that defective signals are characterized, for example, by high NumPoints_H1 values (more cycles) and Amplitude_H1 (more stable cycles), then such signals will be highlighted on the heatmap by brighter vertical columns in the corresponding rows. The image indeed shows areas of increased and decreased feature values, forming clear boundaries. This confirms that topological descriptors effectively generate distinguishable patterns for defective and defect-free areas. Such visually distinguishable clusters on the heat map provide direct evidence of the suitability of these features for automatic defect classification or segmentation.

To further identify hidden structures in the data and automatically segment defective and defect-free areas, K-Means clustering was applied based on the extracted topological feature vectors. The clustering results are summarized in Table 15.

Table 15. Clustering results

Clustering results
K = 2, silhouette score=0.471
K = 3, silhouette score=0.334
K = 4, silhouette score=0.336
K = 5, silhouette score=0.300
Optimal number of clusters: 2 with a silhouette score of 0.471

The silhouette score is a metric used to assess clustering quality. It measures how similar an object is to its own cluster compared to other clusters. Silhouette score values range from -1 to 1, with values close to 1 indicating well-separated and dense clusters, and values close to 0 or negative values indicating overlapping or poorly formed clusters.

In this case, the highest silhouette score of 0.471 was achieved with K = 2. This strongly indicates that the dataset has the most natural and well-separated division into two main signal categories. This binary division is ideal for non-destructive testing tasks, where it is essential to distinguish between defects and the absence of defects.

The graph in Fig. 13 shows the results of topological feature clustering after reducing the dimensions to two using principal component analysis (PCA) for easier visualization. Each point on the graph represents a single signal, and the color of each point indicates its cluster membership, as determined by K-means algorithm.

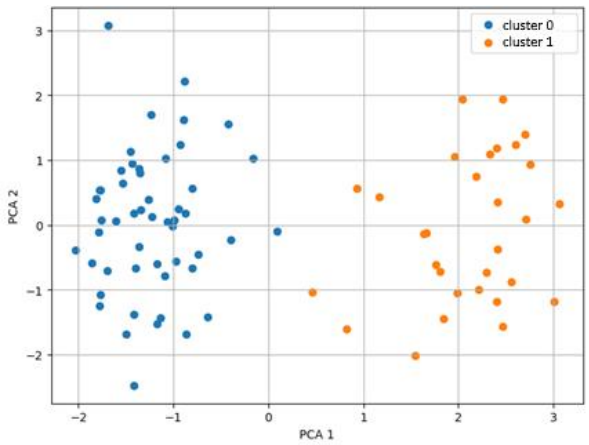

Fig. 13. Clusters by topological features (PCA)

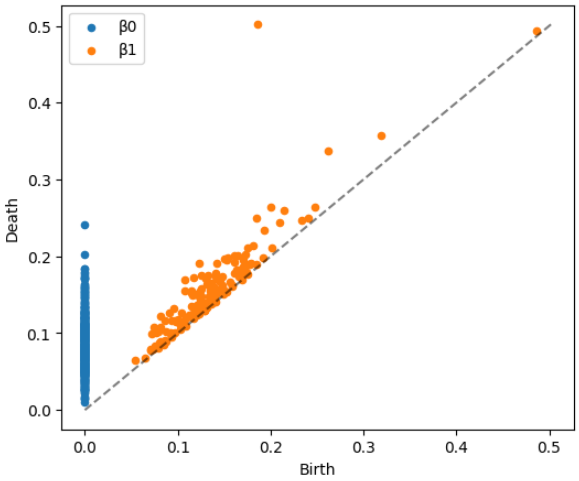

Fig. 14. Synthetic signal stability diagram (points on the synthetic signal diagram: 637)

The graph shows two distinct, well-defined groups of points, colored differently (blue and orange), representing two optimal clusters based on the silhouette score. These clusters have minimal overlap in the two-dimensional principal component space, indicating significant and statistically meaningful differences in the topological features between signals from different clusters.

Since topological features were extracted to distinguish defective from defect-free regions, this clear separation into two clusters confirms that the TDA method successfully detects internal structural differences in the data. This allows for automatic classification of signals based on their topological characteristics, such as those from defective regions (characterized by the presence of stable cycles, as shown by stability diagrams) and those from healthy regions (with a predominance of noise-like components).

To benchmark the proposed TDA approach against conventional linear techniques, the same 81 ultrasonic signals were processed using two standard baseline methods: Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT, Daubechies 4, level 5) followed by an identical K-Means clustering and silhouette evaluation routine. While the FFT feature vectors comprised total power, dominant frequency, and sub-band power distribution, and the DWT approach relied on the energy of detail coefficients, both methods yielded significantly lower silhouette scores than the 0.471 achieved through TDA. This performance gap is clearly explained by the diagnostic plots in Fig. 7–9, which show that at a 10 dB noise level, the power spectra of defective and healthy regions overlap so heavily that frequency- and wavelet-domain features cannot reliably differentiate between them. In contrast, TDA identifies persistent H1 topological cycles that remain stable despite the noise, effectively capturing the physical delamination interface within the reconstructed phase space and proving far more robust than classical signal processing in such challenging signal-to-noise conditions.

To fundamentally validate the TDA's ability to reliably identify known topological structures, a test was performed on a synthetic signal (Fig. 14). For this, a simple sinusoidal signal with some random noise was generated. It is known that such a signal, when embedded in phase space using the Takens method, should form a clearly defined topological structure loop.

The synthetic signal plot (Fig. 14) clearly shows a dominant orange dot β_1, indicating a very stable topological cycle. Its location far from the diagonal confirms the cycle's significant lifetime. This fact demonstrates that the TDA algorithm accurately finds and describes closed structures, or loops, in the signal's phase space, even in the presence of noise.

TDA was applied to 81 signals using the Takens embedding (m = 3, τ = 1) and Vietoris-Rips complexes. Stability diagrams (Fig. 10) show that signals with delaminations (e.g., signal 1) contain stable H1 cycles (orange dots far from the diagonal), corresponding to anomalies caused by delaminations in the composite beam. Signals without defects (signals 2 and 3) have more short-lived H0 components, reflecting noise. Table 14 shows the descriptors (NumPoints_H1, Amplitude_H1) varying depending on the delamination position. Clustering (Table 15, Fig. 13) divided the signals into defective and defect-free clusters (silhouette score 0.471).

The mathematical logic of the TDA pipeline is as follows:

1. For each ultrasonic signal s_i ∈ $\mathbb{R}^{1122}$: P_i = TakensEmbedding(s_i, *m* = 3, τ = 1).
2. Construct Vietoris-Rips complex VR (P_i, ε) for ε ∈ [0, ε_max].
3. Compute persistence diagram D_i = Pers(VR) in dimensions 0 and 1.
4. Extract feature vector F_i = {NumPoints_H0, NumPoints_H1, Entropy_H0, Entropy_H1, Amplitude_H0, Amplitude_H1}.
5. Form matrix F ∈ ℝ^{81×6}, standardize, and apply K-Means for k = 2…10.
6. Select optimal k by maximum silhouette score.
7. Validation on synthetic sinusoidal signal with known H1 cycle.

The results showed that TDA can identify basic topological features and that these features remain stable even with background noise, making them distinguishable from random fluctuations. These early findings can lay the groundwork for applying the method to real-world nondestructive testing signals, where defects may appear as similar topological features in the signal's phase space. Detecting these with TDA allows for accurate damage assessment diagnosis.

Conclusions

A non-destructive testing method for composite materials has been developed that combines topological data analysis and graph signal processing. This study emphasizes applying topological data analysis to synthetic ultrasonic data, marking a first step toward integrating these approaches processing.

The results obtained allow the following conclusions to be drawn:

1. Topological data analysis allows the extraction of noise-resistant features, such as the number of first-order cycles (H1), the entropy of stability diagrams, and the stability amplitude from synthetic ultrasonic data (Table 14). First-order cycles indicate structural anomalies associated with delaminations in composite materials, thereby helping distinguish defective from defect-free signals.
2. The K-means algorithm divided the signals into two clusters, corresponding to defective (containing delaminations) and defect-free regions, with a silhouette score of 0.471, demonstrating topological data analysis's ability to identify differences in the signal structure related to physical aspects of defects.

3. A test on a noisy sinusoidal signal demonstrated that TDA reliably detects topological structures (stable cycles), even amidst noise, highlighting its robustness.

In this study, the analysis is limited to applying topological data analysis to synthetic data, without employing graph signal processing. Validation on real non-destructive testing data was not conducted, which limits the conclusions about the method's practical applicability. The computational complexity of topological data analysis necessitates optimization to enable real-time use. The results obtained provide a foundation for future integration with graph-signal processing and will enhance the accuracy of quantifying and localizing defects in complex composite structures.

Future research directions include:

1. Application of graph signal processing for spatial dependency analysis and defect localization in synthetic and real non-destructive testing data.
2. Conducting tests on diverse datasets, including different types of composite materials and defects (such as delaminations, pores, and cracks), to verify the scalability of the method.
3. Refinement of computational algorithms for topological data analysis and graph signal processing using parallel computing to accelerate processing and adapt to industrial environment conditions.

Integration of the proposed method with graph neural networks to automate defect classification and enhance the interpretability of results.

CRediT authorship contribution statement

Alexey I. Borovkov: writing – review & editing, supervision; **Khristina M. Vafaeva**: writing – original draft, conceptualization; investigation; data curation; **Nikolay I. Vatin**: writing – review & editing, supervision; **Zhmagul S. Nuguzhinov**: writing – review & editing, supervision.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Li Y, Cheng S, Wang S, Yuan C, Luo Z, Zhu Y, Hu J, He J, Li Q. Multilayered ferroelectric polymer composites with high energy density at elevated temperature. *Composites Science and Technology.* 2021;202: 108594.
2. Bastovansky R, Smetanka L, Kohar R, Mishra RK, Petru M. Comparison of mechanical property simulations with results of limited flexural tests of different multi-layer carbon fiber-reinforced polymer composites. *Polymers.* 2024;16(11): 1588.
3. Gilys L, Griškonis E, Griškevičius P, Adlienė D. Lead-free multilayered polymer composites for radiation shielding. *Polymers.* 2022;14(9): 1696.
4. Berretti S, Thomas J-B, Hayat K, Abdollahi-Mamoudan F, Ibarra-Castanedo C, Maldague XPV. Non-destructive testing and evaluation of hybrid and advanced structures: a comprehensive review of methods, applications, and emerging trends. *Sensors.* 2025;25(12): 3635.
5. Wiener J, Arbeiter F, Kolednik O, Pinter G. Influence of layer architecture on fracture toughness and specimen stiffness in polymer multilayer composites. *Materials & Design.* 2022;219: 110828.
6. Feng M, Feng Y, Zhang T, Li J, Chen Q, Chi Q, Lei Q. Recent advances in multilayer-structure dielectrics for energy storage application. *Advanced Science.* 2021;8(23): 2102221.
7. Huang X, Su S, Xu Z, Miao Q, Li W, Wang L. Advanced composite materials for structure strengthening and resilience improvement. *Buildings.* 2023;13(10): 2406.

8. Kumpati R, Skarka W, Ontipuli SK. Current trends in integration of nondestructive testing methods for engineered materials testing. *Sensors.* 2021;21(18): 6175.
9. Mortada H, El Mousharrafie S, Mahfoud E, Harb M. Noncontact nondestructive ultrasonic techniques for manufacturing defects monitoring in composites: a review. *Structural Health Monitoring.* 2024;23(3): 1969–1997.
10. Gupta R, Mitchell D, Blanche J, Harper S, Tang W, Pancholi K, Baines L, Bucknall DG, Flynn D. A review of sensing technologies for non-destructive evaluation of structural composite materials. *Journal of Composites Science.* 2021;5(12): 319.
11. Sabry AH, Ungku Amirulddin UAB. A review on fault detection and diagnosis of industrial robots and multi-axis machines. *Results in Engineering.* 2024;23: 102397.
12. Zhang J, Peng L, Wen S, Huang S. A review on concrete structural properties and damage evolution monitoring techniques. *Sensors.* 2024;24(2): 620.
13. Kosova F, Altay Ö, Ünver HÖ. Structural health monitoring in aviation: a comprehensive review and future directions for machine learning. *Nondestructive Testing and Evaluation.* 2025;40(1): 1–60.
14. Gutkin MYu, Kolesnikova AL, Krasnitckii SA, Mikaelyan KN, Petrov DA, Romanov AE, Smirnov AM. Micromechanics of misfit stress relaxation in heterogeneous crystalline nanostructures: a review. *Materials Physics and Mechanics.* 2025;53(5): 1–34.
15. Keresten IA, Pirozhnikov PB, Suranov IS, Erofeev DA, Titov AG. Experimental and numerical determination of mechanical properties of porous thermoplastic. *Materials Physics and Mechanics.* 2025;53(4): 66–75.
16. Wang Z, Jia Z, Ren J, Wang L, Lan D, Zhang S, Shi X, Liu X, Gao Z, Wu G. Multi-topological network engineering of Co/MnO composites for electromagnetic wave absorption. *Journal of Materials Science & Technology.* 2025;235: 81–90.
17. Al-Dayel I, Nadeem MF, Khan MA. Topological analysis of tetracyanobenzene metal–organic framework. *Scientific Reports.* 2024;14: 1789.
18. Sun C, Wang W, Tian XW, Zeng X, Qian SH, Cai YZ, Wang XH. Thermal design of composite cold plates by topology optimization. *International Journal of Mechanical Sciences.* 2023;259: 108594.
19. Liu X, Gao L, Xiao M. An efficient multiscale topology optimization method for frequency response minimization of cellular composites. *Engineering with Computers.* 2025;41: 267–291.
20. Alfouneh M, Hoang VN, Luo Z, Luo Q. Topology optimization for multi-layer multi-material composite structures. *Engineering Optimization.* 2023;55(5): 773–790.
21. Hu W, Pang J, Liu X, Tian D, Lin CW, Vetro A. Graph signal processing for geometric data and beyond: theory and applications. *IEEE Transactions on Multimedia.* 2022;24: 3961–3977.
22. Ramakrishna R, Scaglione A. Grid-graph signal processing (grid-gsp): a graph signal processing framework for the power grid. *IEEE Transactions on Signal Processing.* 2021;69: 2725–2739.
23. Cheema MA, Sarwar MZ, Gogineni VC, Cantero D, Rossi PS. Computationally efficient structural health monitoring using graph signal processing. *IEEE Sensors Journal.* 2024;24(7): 11895–11905.
24. Zhao Z, Chen NZ. Spatial-temporal graph convolutional networks (STGCN)-based method for localizing acoustic emission sources in composite panels. *Composite Structures.* 2023;323: 117496.
25. Vasudevan A, Prieto JZ, Zorkaltsev S, Haranczyk M. TDA-segmentor: a tool to extract and analyze local structure and porosity features in porous materials. *Computer Physics Communications.* 2024;305: 109344.
26. Lu H, Cantero-Chinchilla S, Yang X, Gryllias K, Chronopoulos D. Multiple ultrasonic-guided wave signals generated by wave propagation model: focused on delamination damage location variations in cross-ply composite beams. *KU Leuven RDR.* 2024; V1.
27. Lu H, Cantero-Chinchilla S, Yang X, Gryllias K, Chronopoulos D. Deep learning uncertainty quantification for ultrasonic damage identification in composite structures. *Composite Structures.* 2024;338: 118087.
28. Lin Y, Liu H, Yu J, Cheng Z, Song Y, Zeng L, Ji X. Non-destructive testing of metal/CFRP composite defects using continuous-wave laser ultrasonic technique. *Optics & Laser Technology.* 2025;189: 113113.
29. Zarei A, Pilla S. Laser ultrasonics for nondestructive testing of composite materials and structures: a review. *Ultrasonics.* 2024;136: 107163.
30. Machado MA. Eddy currents probe design for NDT applications: a review. *Sensors.* 2024;24(17): 5819.
31. Cheng J, Zhu Y, Wang B, Liu M, Xu D, Qiu J, Takagi T. Noncontact visualization of multiscale defects in CFRP composites using eddy current testing with T-R probe. *NDT & E International.* 2024;145: 103138.
32. Guo Z, Lee KM, Yu H, Xiong Z. Magnetic field-based eddy-current probe design, modeling, and computing methods for edge defect detection. *IEEE Sensors Journal.* 2024;24(12): 18889–18902.

33. Sharma S, Vishnu VG, Srivastava V. Detection of failure in materials under tensile testing using acoustic emission technique. *e-Journal of Nondestructive Testing*. 2025;30(6).
34. van de Velde M, Vandecruys E, Verstrynge E, Reynders E, Lombaert G. Vibration monitoring and acoustic emission sensing during progressive load tests of corroded reinforced concrete beams. *Engineering Structures*. 2024;306: 117851.
35. Melchiorre J, D'Amato L, Agostini F, Rizzo AM. Acoustic emission onset time detection for structural monitoring with U-net neural network architecture. *Developments in the Built Environment*. 2024;18: 100449.
36. Hsiao TY, Sfarra S, Liu Y, Yao Y. Two-dimensional Hilbert-Huang transform-based thermographic data processing for non-destructive material defect detection. *Quantitative InfraRed Thermography Journal*. 2025;22(4): 297–312.
37. Tang S, Gao X, Tian K, Zhang Q, Zhang X, Peng J, Guo J. Non-destructive evaluation of weld defect with coating using electromagnetic induction thermography. *Nondestructive Testing and Evaluation*. 2024;39(2): 347–365.
38. Evans EE, Brooks RA, Liu J, Hall ZEC, Liu H, Lowe TJE, Withers PJ, Kinloch AJ, Dear JP. Comparison of X-ray computed tomography and ultrasonic C-scan techniques and numerical modelling of impact damage in a CFRP composite laminate. *Applied Composite Materials*. 2024;31: 249–264.
39. Brewer CE, Poovathingal SJ. Pyrolysis front detection in carbon phenolic composites using X-ray computed tomography. *Composites Part A: Applied Science and Manufacturing*. 2024;187: 108444.
40. Li C, Xia T, Gao S, Yan M, Chen Y, Wan B, Li J, Wan F. Microstructure analysis of quartz fiber reinforced SiO_2 matrix composites by X-ray computed tomography. *Materials Characterization*. 2024;209: 113745.