Study of composite polymer degradation for high pressure hydrogen vessel by machine learning approach

K. Kadri[a], A. Kallel[c], G. Guerard[b], A. Ben Abdallah[a], S. Ballut[a], J. Fitoussi[c], M. Shirinbayan[c,*]

[a]*R&D Departement, Aptiskills, 28 Rue Edouard Vaillant, 92300 Levallois-Perret, France*
[b]*De Vinci Research Center, La Défense, Léonard de Vinci Pôle Universitaire, Paris, France*
[c]*Laboratoire PIMM, Arts et Métiers, CNRS, Hesam Université, 151 boulevard de l'Hôpital, Paris, France*

Abstract

The aim of this paper is to study the degradation of a composite material under static pressure. The high pressure condition is similar to the one encountered inside hydrogen tanks. Damage modeling was used to evaluate the behavior of hydrogen tanks to high pressure. A practical approach, coupling a finite element method (FEM) simulation and machine learning algorithm (ML), is suggested. The representative volume element (RVE) was used in association with a choice of a behavior law and a damage law as an input data. Algorithms for ML classification such as K-nearest neighbors (k-NN) and a special k-NN with a dynamic time warping (DTW) metric were used. The hierarchical clustering through dendrograms visualizations allowed to exhibit the impact of composite parameters in relation to fiber, matrix properties and fiber volume fraction on the strain degradation under external static pressure. Continuing this, the optimum RVE which shows a low degradation value will be exhibited.

Keywords: Hydrogen pressure vessel, RVE, Damage law, Behaviour law, Feature classification, Time series, Machine learning, Classification, Dendrogram.

*Corresponding author, Mohammadali.SHIRINBAYAN@ensam.eu

1. Introduction

Composite materials made of a polymeric matrix and a fiber as a reinforcement can offer remarkable physical and chemical properties, such as high strength, low specific weight, fatigue and corrosion resistance, and low thermal expansion. These physical properties enable composites to develop new engineering applications. In fact, they are in a pole position candidate for the type 4 hydrogen pressure vessels. They are promising materials for this application due to the broad scope engineering architecture possibilities. By designing the micro-structure, the composite properties can be controlled and functionally upgraded. Homogenization method [1, 2, 3] will be used to study the structure of the composite material and its properties. This simplification approach [4, 5] of Mean-field calculation is a prediction of global behaviour when the microstructure is heterogeneous. This method offers advantages like evaluation of composites such as structure, electromagnetic, and thermal properties! as well as the study flow in porous media [6]. In a nutshell, this method can be described as the calculus of the stresses and strains at any point of the homogenized medium. They are noted S and E and are assumed to be the over all average of any RVE of the corresponding local microscopic quantities σ and ϵ. We use the same notation as in this article [7], and we are reminded how S and E are calculated within the homogenization method.

$$S = \frac{1}{|V|} \int_{\partial V} dS x \otimes \underline{\underline{\sigma}} \cdot \underline{n} \tag{1}$$

$$E = \frac{1}{|V|} \int_{\partial V} dS \frac{1}{2} \left(\underline{u} \otimes \underline{n} + \underline{n} \otimes \underline{u} \right) \tag{2}$$

Here, x is a point of the RVE and its boundary, u the microscopic displacement field, n the normal vector to the closure ∂V and the term $\underline{\underline{\sigma}} \cdot \underline{n}$ is the boundary traction. The symbol $\otimes$ denotes dyadic tensor product. The previous equations 2 allow a simplified computation of the macroscopic variables in comparison with volume integrals over the RVE. For a periodic media, the microscopic fields must have suitable periodicity conditions ensuring continuity of boundary tractions and displacements across two opposite faces of the cell. $\underline{\underline{\sigma}} \cdot \underline{n}$ must take equal and opposite values at two corresponding points of ∂V. For this reason we will use periodic boundary conditions

(PBC) that will be explained in the following.

Metals as well as composites have been studied through the scope of ML and deep learning (DL) in term of properties and fatigue life. Studies have been conducted to explore the contribution of ML to solve the issue of properties prediction [8] and defect location, size, morphology in life cycle of a metal Ti-6Al-4 V alloy [9]. Support Vector Machines ($SVMs$) algorithm was used to accelerate the optimization process. The grid search approach with cross validation was selected for fitting the model parameters. The predicted fatigue life can reach up to 0.99 precision only by learning from defects. An other major axe of research was to employ ML and DL to detect cracks in pressure vessels at an early stage [10]. Crack identification techniques for pressure vessels that used genetic algorithm (GA)-based on feature selection, deep neural network (DNN) on acoustic emission (AE) examination. Heterogeneous features like time domain, frequency domain, and time-frequency were selected by GA to represent the most discriminant features that are highly effective for the DNN classifier when identifying crack types.
Experimental analyses of composite behaviour are costly and often involve high nonlinearity. ML techniques provide effective models that can capture high nonlinearity [11]. In this previous study, using a finite element modeling as a starting point, different ML techniques like Random Forest Regressor, Decision Tree algorithm, and Multiple Linear Regression were employed to establish the causal relation between the fiber angle and buckling capacity of a cylinder under bending-induced loads.

The present study attempts to expand the scope of ML algorithms to the issue of composite damage material as a response to external static pressure. In fact, the importance of understanding the degradation of a composite material under static pressure is of utmost importance to guarantee the safety of hydrogen storage vessels [12].
The coupling between FEM simulation and ML will provide numerical support to develop a well suited composite. Optimal ML algorithms aim to improve the predictive performance of the material resistance to high pressure hydrogen tank. It will also shed light on degradation of a composite under different solicitations. The subsequent sections of this article are structured as follows: the first section exhibits the central problem of ductile damage and rupture criteria under high static pressure. The second section presents the modeling methodology. The result of FEM simulation will be exposed

in term of ductile damage where the effect of RVE 's size will be addressed. The third section exposes the use of ML modeling to evaluate the damage behavior. It will introduce the description of the used database and the pre-processing treatment. In this supervised learning for classification, the model evaluation metrics will be discussed. Deployment of both k-NN, and dynamical time warping k-NN models are described with the performances and comparison of proposed models. The results in term of material degradation classification and performance will be given. Finally, the last section draws conclusions and outlines promising directions for future work.

2. Problem position

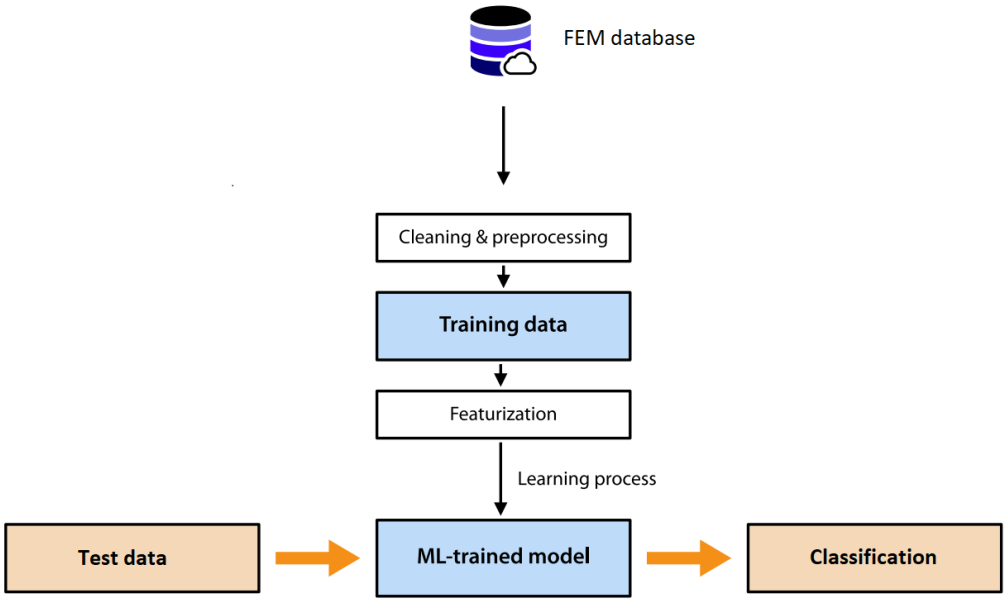

Figure 1: A schematic representation of the key steps of the ML workflow. The composite material at microscopic scale is simulated by FEM. The post-processing of the ABAQUS simulation allows to extract the dataset for ML classification. Classification performance is measured over the test dataset (*adapted from [13]*).

The general workflow for the problem studied in this article is illustrated in Figure 1. The study aims to predict patterns of degradation in composite materials by leveraging a hybrid approach that combines Finite Element Method (FEM) simulations and Machine Learning (ML) techniques. The

research problem is addressed by conducting FEM simulations of Representative Volume Elements (RVEs) of square unit cell unidirectional composites (UD). A body-centered cubic (bcc) unit cell configuration is employed as it is suitable for representing the fiber-matrix interaction. The aim is to use ML algorithms, specifically K-Nearest Neighbors (KNN) and KNN with Dynamical Time Warping (DTW), to classify and visualize the degradation patterns of different composites based on their matrix type, fiber type, and volume fraction. By analyzing the hierarchical clustering tree derived from these algorithms, insights into the relative impacts of matrix composition, fiber type, and volume fraction on the degradation behavior of the composites can be gained. The overarching aim is to develop a predictive framework for understanding and comparing composite degradation mechanisms, facilitating advancements in material design and engineering.

2.1. Damage parameter for uniaxial loading conditions

Damage is characterized by the decline of material stiffness. It plays a central role in the analysis of fiber-reinforced composite materials. Several of these materials exhibit elastic-brittle behavior. Damage in these materials is initiated without significant plastic deformation. The fibers are assumed to be parallel. Material properties are specified in a local coordinate system 1. However, the material response can also be defined in terms of the engineering constants and by specifying the elastic stiffness matrix directly as mentioned in the following section.

2.2. Ductile damage

Investigating the damage process in a composite material under constant loading can be challenging. Many studies on the laws of ruptures have been carried out [14, 15, 16, 17].

The ductile damage initiation criterion is a model used to predict the triggering of damage due to nucleation, growth, and coalescence of voids in ductile metals [18, 19, 20, 21]. It has been used also for composites [22, 23, 24]. The model assumes that the equivalent plastic strain at the onset of damage is a function of stress triaxiality and strain rate [25]. The ductile criterion can be used in conjunction with the Mises, Johnson-Cook, Hill, and Drucker-Prager plasticity models [26]. More precisely, the ductile damage model that we use and adapt to our need is inspired from this article [27]. It has its justification

from the anisotropic behaviour due to the specific orientation of the fiber inside the matrix.

2.3. Rupture criteria under high pressure

The damage initiation criterion in this model is defined in terms of the effective plastic strain $\bar{\epsilon_{pl}}$. The damage propagate when the scalar ω reaches 1, where ω is defined as:

$$\omega = \int_0^t dt \frac{\dot{\bar{\epsilon_{pl}}}}{\bar{\epsilon_{pl}^{cr}}(\eta, \bar{\epsilon_{pl}})} \qquad with \qquad \dot{\bar{\epsilon_{pl}}} = \sqrt{\frac{2}{3}\dot{\epsilon_{pl}} : \dot{\epsilon_{pl}}} \tag{3}$$

$\bar{\epsilon_{pl}}$ is defined as the plastic deformation. It increases monotonically as the state variable ω increases. Here, $\bar{\epsilon_{pl}^{cr}}(\eta, \bar{\epsilon_{pl}})$ is the equivalent plastic strain at the onset of damage, and η is the triaxiality. $\dot{\bar{\epsilon_{pl}}}$ is the strain rate.
The damage initiation occurs when the equation 3 (satisfies $\omega = 1$). There here a function of triaxiality and strain-rate to take into account the triaxiality and rate dependent damage found in thermoplastic matrices. Numerical values used are reported in Table 2.
The damaged stress tensor σ can be expressed as following:

$$\sigma = \bar{\sigma}(1 - D) \tag{4}$$

where $\bar{\sigma}$ is the undamaged stress tensor. D, is the damage variable that increases from 0 to 1. The damage evaluation [26] can be defined using displacement at failure in two cases: linear form and exponential form. The governing equation for equivalent plastic displacement is:

$$D = D(u^{pl}) \tag{5}$$

Where, u^{pl} is the effective plastic displacement. In this study linear form is used. The software used in this study for FEM simulation, is ABAQUS. It offers a damage model which allow to predict the onset and the damage evolution for elasto-plastic materials with anisotropic behavior. The ABAQUS anisotropic damage model is based on the work of Hashin, Matzenmiller and Camanho, [20, 19, 21].

3. Material and method

The shape of the RVE is chosen according to the microstructure, which can be an accurate reflection of reality or an idealization [28, 4]. In this study, in the case of unidirectional fibers distributed at the center of the RVE, the body-centered cubic packing is chosen as it is an ideal representation of the composites. The choice of the RVE type is motivated by the microstructure according to these studies [29, 30, 31].

The numerical definition of each phase of the composite corresponds to the matrix and fiber. In this study, we numerically explored five different polymers for the matrix since they are the most common [32] as shown in table 3. Two types of fibers are used; glass and aramid one. The matrix and fiber which constitute the composite are separately defined in a local coordinate system. The material properties to be inserted in the model are defined in terms of engineering constants: Young modulus E and Poisson ratio ν. The mechanical properties of the RVEs are defined within the ABAQUS module named elasticity. This subroutine module uses the previous values of the constant of Young modulus and Poisson ratio to reproduce the elastic properties. We add tabular values (see Table 3) for the elasto-plastic properties of both fiber [33] and matrix. In order to describe the behaviour of the material at a displacement stress, the values of plastic strain according to the yield stress are implemented manually (see next Table 1).

Table 1: Explored range of the input parameters of the plastic behaviour Ps and the yield value Ys of each matrix used in FEM simulation. Extracted from this reference [34]

PA6		PA12		PP		LDPE		HDPE	
$Ps_{(MPa)}$	Ys(%)	$Ps_{(MPa)}$	Ys(%)	$Ps_{(MPa)}$	Ys(%)	$Ps_{(MPa)}$	Ys(%)	$Ps_{(MPa)}$	Ys(%)
0.1	0	0.16	0.01	0.0	0.0	13.0	0.0	0.1	0.0
5.50	0.16	7.71	0.5	14.38	0.10	13.2	0.01	11.32	0.96
6.42	0.19	14.8	1	17.54	0.20	13.3	0.02	16.91	1.84
7.34	0.22	21.3	1.5	20.34	0.35	13.4	0.1	21.1	3.11
8.26	0.25	26.9	1.99	22.43	0.54	13.5	0.15	23.82	4.82
25.0	0.77	32	2.5	24.70	0.77	13.6	0.35	25.14	6.79
42.4	1.32	36	3.0	26.8	1.14	13.7	1.0	25.55	9.608
53.4	1.70	38.3	3.5	28.8	1.57	13.8	3.0	25.17	13.31
63.9	2.09	39.7	4.0	30.2	2.10			22.70	24.15
73.2	2.49	40.3	4.5	31.1	2.72				
80.7	2.92	40.5	5.0	31.5	3.42				
85.6	3.39			30.1	4.71				
87.4	3.95			30.32	5.98				

In this study, the ductile damage subroutine involves the initiation criterion with the following variables as presented in Table 2. The overall damage variable is D. The default setting of D depends on whether elements should be deleted or not after reaching the maximum degradation. Output variable proper to ABAQUS is SDEG and it contains the value of D. SDEG is defined as overall scalar stiffness degradation. It is used when damage evolution is specified:

Table 2: Damage parameters used to implement numerically the ductile damage of the composite RVE [35]

Fracture strain	Stress triaxiality	Strain rate
0.05	-0.333	0.1

4. Modeling methodology : Finite element modeling

By employing those input parameters and subroutine of elasticity, plasticity, ductile damage, and elasto-plastic law, the response of the RVE subbmited to different type of solicitation from tensile traction to compression and shear will be determined. The constitutive modeling approach is used to create a model that reflects comprehensively the behaviour of the test material. To achieve this goal, FEM simulation will follow the different steps as precised in the next section. This section explains the basic steps of the Python codes. The code is called in ABAQUS/CAE through the menu command File, then the script is run

Developing a constitutive models requires to consider the relationship between stress and strain [36, 37]. This relationships gives the most reliable insight window into the mechanics of material behaviour. Therefore, linear elasticity models are the simplest and most common material models in use. They are the key (practical) material models used for solids mechanics and structural analyses. This law has to be generalized into a triaxial stress and strain state where the proportionality constant is a fourth-order tensor. The most general vectorized expression of linear elasticity, implemented in the FEM solvers is given in the following equation:

$$\sigma = D : \epsilon \tag{6}$$

Where σ and ϵ are second-order stress and strain tensors, which can be presented in the Voigt contracted notation, and D is a fourth-order tensor called the stiffness tensor. In practice, the set of 81 terms of the D tensor is always reduced out of presence of symmetry. In the following, we consider special cases of isotropic transverse symmetry for the stiffness tensor. This means that the material possesses a plane of symmetry. A longitudinal axis is oriented in, the Z-axis as presented in Figure 2(a). All properties in the X and Y-directions will be similar. A fibre reinforced unidirectional composite shows this property with the modulus, in the main fibre direction, E_{33} for the Z or number 3 axis. The properties in the two transverse axes (E_{22} and E_{33}, for example) will be the same. In short, the unidirectional ply material is stiffer in the direction of the fiber than in the others directions. A symmetry is observed in the longitudinal direction. This model will be expressed in an orthonormal basis as following:

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \gamma_{23} \\ \gamma_{13} \\ \gamma_{12} \end{bmatrix} = \begin{bmatrix} \frac{1}{E_L} & -\frac{\nu_{LT}}{E_L} & -\frac{\nu_{LT}}{E_L} & 0 & 0 & 0 \\ -\frac{\nu_{LT}}{E_L} & \frac{1}{E_T} & -\frac{\nu_{LT}}{E_L} & 0 & 0 & 0 \\ -\frac{\nu_{LT}}{E_L} & -\frac{\nu_{TT}}{E_T} & \frac{1}{E_T} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2(1+\nu_{TT})}{E_T} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_{LT}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{LT}} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{23} \\ \tau_{13} \\ \tau_{12} \end{bmatrix} \tag{7}$$

Where the Es are the Young modulus and the ν_{LT}, ν_{TT} correspond to Poisson ratio in the "longi/trans" and "trans/trans" direction. The G_{LT} and G_{TT} are the shear modulus.

4.1. (RVE) modeling of unidirectional composite in ABAQUS

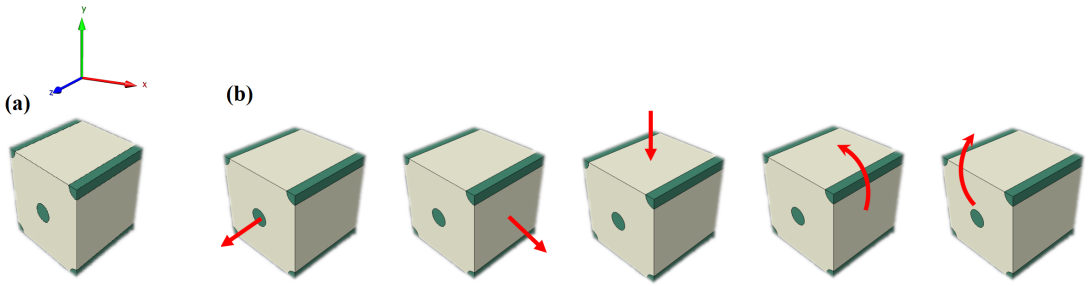

Figure 2: *In (a)* 3D RVE representing in *green* the fiber and in *white* the matrix with the (bcc) symmetry generated by ABAQUS. *In (b)* the 5 solicitations realized by the homogenisation method from *left* to *right* (X-tensile, Y-tensile, Z-compression, Shear in plan (XY), shear in plan (YZ)).

The unit cell employed in the present study uses a (bcc) symmetry of the RVE. Aligned reinforcements according to a Z-axe direction are generated at the center and four corners of the cubic volume as we can see in Figure 2(a). In order to allow a direct assessment of fiber-matrix effects, 100 different RVE were generated sequentially. Cylindrical fibers occupy the same positions and have the same orientations. Aspect ratio and reinforcement volume fraction were looped over in a range between $[35\% - 40\%]$ according to the following formula [38] that relates volume fraction (F_v) to the radius of the fiber $R_{fiber} = \left(\frac{F_v WHL}{\pi}^{1/3}\right)$. Where W, L, H are the weight, length, and height of the unit cell. The geometric parameters of the RVE are defined into unite

value $W = L = H = 1$ mm as shown in Figure 2(b). In the simulation, we use a python script to set the main variables that drive numerical simulation on ABAQUS engine solver. We loop over the fiber properties (E, ν) as well as the matrix's characteristics and the volume fraction F_v as it is exposed in the following Table 3:

Table 3: Detailed range values of the input parameters used in the python script: Matrix and fiber properties, Volume fraction percentage, and solicitation direction.

Matrix name	PA12	PA6	PP	LDPE	HDPE
$(E_{Young}(MPa),\nu)$	(1250, 0.39)	(800, 0.4)	(1550, 0.43)	(1080,0.45)	(565, 0.46)
Fiber name	glass	aramid			
$(E_{Young}(MPa),\nu)$	(8000, 0.34)	(3500, 0.35)			
Volume fraction	35%	40%			
Solicitation	Tensile:X, Z	Compression:Y	Shear:XY,YZ		

The used value for fibers are taken from the following references: (cited in foote notes [1] [2]). In the table 4 , values of the isotropic plastic data for both glass and aramid fibers are presented.

[1] Campus datasheet DUPONT engineering Plastic Zytel®G35HSLRA4BK267-PA66GF35

[2] Campus datasheet DUPONT engineering Plastic Zytel® 70K20HSL BK284-PA66AF20

Table 4: Explored range of the input parameters of the plastic behaviour of each fiber used in FEM simulation. The datas are taken from the following website [34]

Glass fiber		Aramid fiber	
Plastic strain (MPa)	Yield stress (%)	Plastic strain (MPa)	Yield stress (%)
0	0	0	0
27.1	0.22	26.02	0.54
60.99	0.52	49.28	1.08
78.57	0.68	68.01	1.62
95.32	0.85	81.52	2.17
109.02	1	89.97	2.71
121.29	1.14	95.14	3.25
133.38	1.29	98.62	3.79
145.26	1.45	101.39	4.32
157.62	1.64	103.83	4.86
169.36	1.83	105.34	5.4
180.78	2.05		
191.22	2.31		
200.01	2.6		
204.36	3.01		

4.2. Finite element mesh generation

A graphical illustration of the convergence principle is shown in the Figure 3**(a)**, where a material strain in Z direction S_{33}, is calculated following an FEM process while changing mesh number (and consequently density) from a few hundreds of meshes to thousand of meshes.

As the mesh density increases, the FEM solution will continue to change until an optimal solution is reached (**red dashed line**) at that moment there is very few changes that occur in the predicted parameter (e.g. strain in Z direction, etc.). The choice of hexahedral mesh is made based on two reasons. At first it allows to reach more precision on the homogenization of deformation in comparison with tetrahedral mesh. The second reason is the fact that the longer time it takes the higher precision in learning will be reached.

Therefore, the mesh value fixed is defined in ABAQUS as fine mesh. This corresponds to: 740 nodes, *ie* 1400 elements. The fibre part is finely meshed

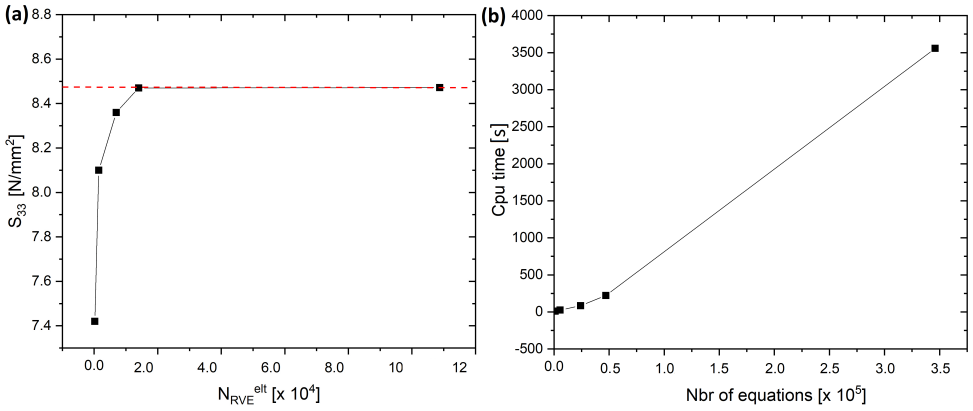

Figure 3: *(a)* The dependence of FEM solutions on mesh density for strain values S_{33} for an RVE. *(b)* Evolution of CPU time according to number of equations.

as well. The Figure 3 illustrates the meshing result on RVE on both matrix and fiber parts. The reference point (RP) is settled to indicate the rigid body reference point. Constraints that were applied to the reference point are then applied to the entire rigid part. In FEM study, in order to reach convergence the solution must be mesh size independent. For this reason a balance between fine mesh and time of calculus 3**(b)** was fixed *(corresponding to 1400 elements).*

4.3. Implementation of periodic boundary conditions:

Periodic boundary conditions (PBC) were applied to the RVEs. Those PBC are used for Homogenization method [39, 40]. First, the resolution of any FEM problem simulation is dependent on the implemented boundary conditions [39]. PBCs are widely used since it has been shown in literature that they can quickly arrive at the convergent solution of an FEM problem [41, 42]. The PBC stipulates that opposite pairs of edges or surfaces on the boundary of an RVE should deform identically under a given loading history [43]. The PBC essentially precludes the constraint of enforcing edges to remain planar after deformation. In ABAQUS, for most RVEs, boundary surfaces always appear in parallel pair. The displacement of a pair of opposite parallel boundary surfaces as proposed by Li and Wongsto [4], and the distortion for each pair of nodes on the parallel boundary surface of the RVE

are written according to the following formula:

$$\begin{cases} u_i^A = \bar{\epsilon_{ik}} x_k^A + u_i^* \\ u_i^B = \bar{\epsilon_{ik}} x_k^B + u_i^* \end{cases} \tag{8}$$

A and B are the name of parallel surface contacts where i is the i^{th} number of the pair nodes. the paired nodes of surfaces in ABAQUS are taken when geometry of study is initially undeformed.

To ensure the symmetry of mesh nodes on the periodic plane, three partitions according to three symmetry plans $(XY), (YZ), (ZX)$ are made.

Here $\bar{\epsilon_{ik}}$ is the average strain of the RVE and u_i^* is the displacement components of the periodic part. Therefore, it is possible to write the relative displacement between two nodes as follows [38]:

$$u_i^A - u_i^B = \bar{\epsilon_{ik}}(x_k^A - x_k^B) = \bar{\epsilon_{ik}} \Delta x_k \tag{9}$$

Where Δx_k is the constant for each pair of boundary surfaces parallel. The previous equation is implemented by setting the linear constraint of the displacement between each pair of nodes on the parallel boundary surface. These boundary conditions guarantees continuous displacement but also tensile continuity on the parallel surfaces of the RVE. Each node of a given surface, must correspond to a node on the opposite surface. The six independent boundaries are calculated independently for each RVE. In each case, an applied PBC separately according to the equation 9. The value of $\bar{\epsilon}$ is given by six cases, and the matrix $\bar{C}$ can be obtained. The average stress and strain in RVE are calculated by the following formula:

$$\begin{cases} \bar{\sigma_{ij}} = \dfrac{1}{V} \displaystyle\int_V dV \sigma_{ij} \\ \bar{\epsilon_{ij}} = \dfrac{1}{V} \displaystyle\int_V dV \epsilon_{ij} \end{cases} \tag{10}$$

Where V is the volume of the RVE. The relationship between the stress matrix for six boundary conditions and stiffness matrix can be written as follows:

$$\{\bar{\sigma}\} = \bar{C} \{\bar{\epsilon}\} \tag{11}$$

Where $\bar{\sigma_{ij}}$ and $\bar{\epsilon_{ij}}$ are the macro stress and strain tensor for each boundary conditions. $\bar{C}$ is the elasticity response matrix.

The determination of the mechanical properties of heterogeneous systems is based on the virtual work principle. This approach uses only the model variables (nodal forces and displacements) at the retained (or corner nodes) to compute the work done within the virtual domain. This will be used subsequently to determine the stresses, strains tensors.
In the present ABAQUS simulation work, we consider a 3D RVE window enclosing a heterogeneous material by boundary edges identified by Ω_k where $k = [L, T, R, B]$, for left, top, right and bottom. Nodal reaction k forces and displacements for 3 retained nodes are given as f_i and x_i where $i = 1, 2, 3$ on retained nodes N1, N2 and N3. Only nodal displacement and force values for 3 retained nodes will be needed to generate the constitutive behaviour (i.e. the stress strain response) of the heterogeneous material.
In the Figure 2 (b) is represented the 5 solicitations that allow to calculated numerically the $\bar{C}$ matrix.

The $\bar{C}$ tensor is the elasticity response of the heterogeneous material obtained numerically. In details, it dependants on different inputs variable of the RVE and can be written by explicit argument as follows :

$$\bar{C}(F_v, E_{matrix}, \nu_{matrix}, E_{fiber}, \nu_{fiber}) \tag{12}$$

Where F_v is the volume fraction of the fiber inside the matrix. This form 12 with arguments resulting from input parameter related to mechanical properties of the RVE , reveals the diversity that can be obtained. Each time one argument is changed a whole new $\bar{C}$ is obtained. This is an other reason why ML approach is appropriate in this work.

4.4. Field output extraction: measures of stress and strain

The FEM process is a predictive modeling, which needs the use of computational methods to determine the behaviour of a given material. The principles of continuum mechanics is to define the response of a material body. Specifically, the finite deformation of a material body. The numerical measures of strains and stresses will concludes the practical formulations of stresses. Such practical stress formulations include: principal stresses, von Mises stresses, Tresca stresses etc...
Numerically, we scripted an editor to select the tensor variable to be included in a field output request. ABAQUS automatically writes all components of that variable to the output database during the step. In the present three-dimensional model, ABAQUS outputs the three displacement components

U_1, U_2, and U_3 in addition to 36 other measurable variables that are needed to characterize the damage process. The main important ones are obtained after a process of feature selection (see next section) and are summed up in the following table 5.

Table 5: Output variables used to monitor the ductile damage of the RVE's composite.

S	PE	LE	RF	SDEG	CFAILURE
stress	plastic strain	logarithmic strain	reaction forces	Stiffness degradation	failure

4.5. Results of FEM simulation

The visualization module of ABAQUS can be used to output SDEG variable as shown in the next Figure 4. It is observed that the degradation is generally dependant on the kind of solicitation. Here there is three type of applayed external forces. For the same amount of load for every RVE we can see that for a tensile in Z direction, degradation field is distributed along the direction of fiber and inside only the matrix that took all the damage. This result is quite logical because of the role of the isotropic fiber. Lower damage in encountered in fiber (center and corners in green colormap). At contrary, when damage is obtained during tensile X solicitation, which is perpendicular to the fiber direction, we can observe that fibers took all the damage all along (red and yellow colormap) elsewhere in the matrix low degradation field appears in green and blue.

Compressing solicitation along Y direction seems to highly contribute to the global degradation of the RVE. In fact, it shows high value of degradation in addition to broad spreading across the volume of the RVE. The degradation under compression affects more specifically the fiber than the matrix. Concerning shear solicitation in both (XY) and (YZ) plan, it is showing the lowest damage on RVE, this is mainly due to the isotropic resistance of the fiber along the Z direction.

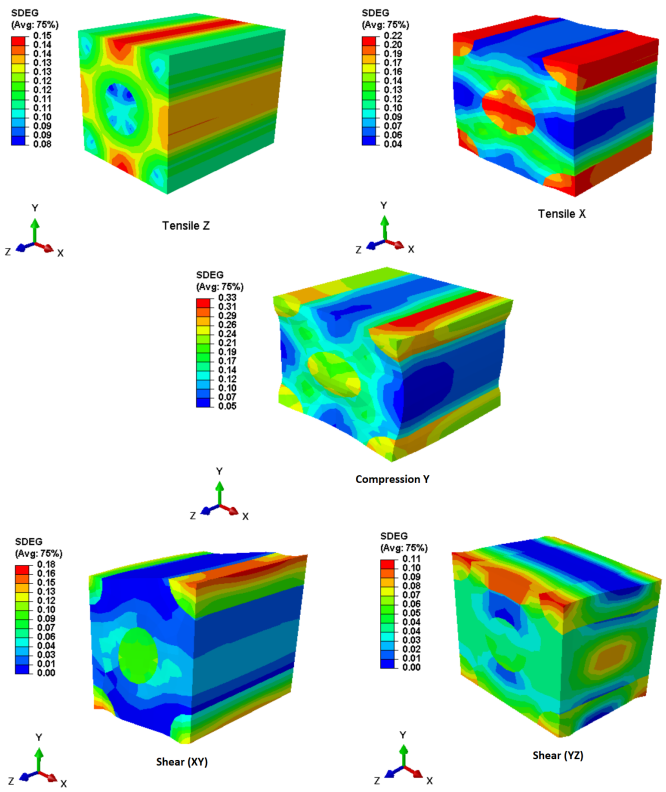

Figure 4: The 5 solicitations effect on degradation realized by the homogenisation method on one RVE ($F_v = 35\%, Fiber = aramid, Matrix = PP$). *On top* is presented the effect of tensile strength in X, Z direction. *On middle* is the effect of compression along Y direction. Finally, *on bottom* is the shear effect that is represented. The overall scalar stiffness degradation is represented in color map.

Obtained results are compared to numerical outcomes presented in this study [35], which, investigates the parameters that controlled damage behaviour of continuous fiber-reinforced thermoplastic composite. The RVE was submitted to transverse tensile loading and showed good accordance with experimental tensile tests on $[90]_8$ laminated model reported in[44]. For our FEM study, qualitative confirmation can be deduced. In fact, our results are in harmony with outcomes results reported in [45]. When comparing with the four cases treated in the previous study, for transverse tension, transverse compression, in-plane shear there is similarities. in fact, in the epoxy carbon fiber reinforced polymer (CFRP) unidirectional lamina with a random distri-

bution of fibers, as for our compression cases, the transverse compression is the most damaged one. Then in Figure 4 in [45] for transverse compression the matrix is more affected. This is confirmed in the present simulation by the strain degradation field that is more spatially extended.

4.6. Effect of RVE's size variation

The effect of RVEs size variation in FEM modeling has been discussed and some authors [46, 47], have proposed different methodologies to determine the RVE size. In the present study the choice of the microscopic scale for the homogenization method bring the question of the effect of the size on damage evolution [28]. In the present work, was investigated numerically this effect. Practically, three RVEs sizes have been tested through homogenization method with (PBC) conditions. The following sizes of the RVEs unite celle in micrometer : $(8, 12, 16)$, was implemented on ABAQUS software. The stress, strain curve was extracted to compare. In fact, the strain σ_{33} according to the tensile strength in Z direction was extracted. This approach uses only nodal forces and displacements at corner nodes to determine stress and strain of a RVE. The mathematical formulation for deriving the homogenized (volume-averaged) stresses as in [38] is given as:

$$< \sigma_{33} >= \frac{1}{V} \left[x_1 \otimes f_{N1} + x_2 \otimes f_{N2} + x_3 \otimes f_{N3} \right] \tag{13}$$

Where, f_i and x_i are Nodal reaction forces and displacements for 3 retained nodes where $i = 1, 2, 3$.

In the previous Figure 5 stress values are extracted according to Equation 13 for the different RVE sizes. The classical curve obtained enables to distinguish the initial linear domain followed by a narrow plastic domain than the inflexion of the curve corresponding to degraded stiffness. The superposition of the three stress strain curves allows the affirmation that the effect of size variation is not observed. That means that the microscopic approach used in this study is validated.

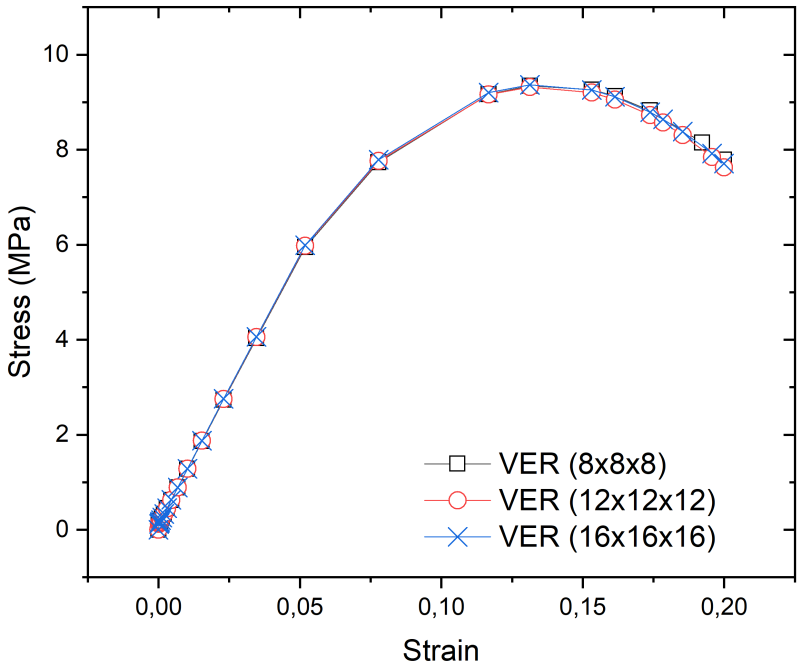

Figure 5: Superposition of Stress-strain curve for three different RVEs size $8\mu m, 12\mu m, 16\mu m$.

5. Implementation of machine learning model for classification

In the last decades, many studies in mechanical material have been using ML models. It was used to treat data from experiments or from numerical simulations in a new way that gives a deeper insight and predictions [13]. However, it is important to keep in mind the limitation of (ML) which is the high dependence on data of quality with implemented models. The recent advancements in deep learning are intriguing; however, this approach comes with a significant drawback: the model is essentially a black box. Given that our study focuses on examining various polymer degradations under pressure, the imperative to comprehend the underlying processes is a crucial necessity. Therefore, this disqualifies the use of deep learning models in our context.

Figure 6 presents the flowchart of the pipeline:

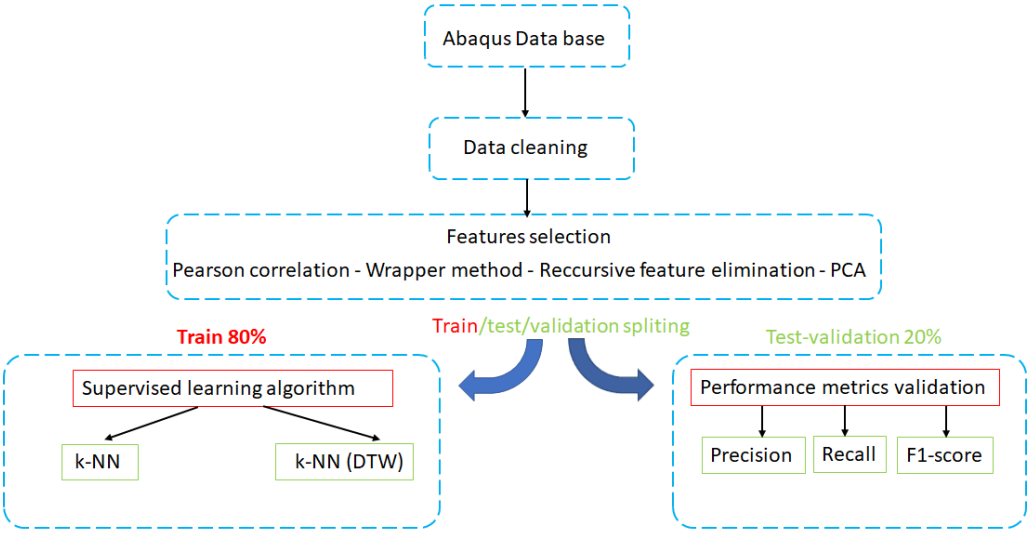

Figure 6: The flowchart of the different steps used for the supervised ML classification. Three major steps are exposed: data cleaning, feature selection, train-test splitting and finally pattern damage classification.

1. **Database**: to provide valuable data to analyse, ABAQUS is used to implement and to simulate various RVE's with FEM simulation.

2. **Data Preprocessing**:

   - **Data Cleaning**: the goal of this step is to consolidate the database (type of data, type of frame, type of file, how to read and save the data, etc.). Then the data analysis provides useful tips about the data, like correlation or multivariate analysis.
   - **Features Selection**: since all the features will not be useful for the prediction, one have to analyze them to determine the relevant features. All features have to be independent to the others, must have a good correlation to the predicted features and must have a dense data (no NULL inputs in the data).

3. **Training**: since the goal of our method is a prediction, the data are splitted into a train-test set respectively 80%-20%. A simple k-NN and variants algorithm named k-NN with dynamic time warping are used

to classify degradation.

4. **Validation**: the common metrics for classification are accuracy, precision, recall, F1-score are used to evaluate the accuracy of the model to build classes.

5.1. Data Collection and Database Creation

Having a dataset of large size significantly enhances the learning process, leading to improved predictions. Utilizing FEM simulation with Python scripting, we generated 100 elasto-plastic models for our study. Parameterization of the simulation through loops enables the exploration of a diverse range of Representative Volume Elements (RVEs), making the incorporation of Machine Learning (ML) highly relevant. For this investigation, the data set is made of 200 RVEs generated through FEM simulation in ABAQUS. Each of the 200 RVEs models includes output fields as responses to five distinct solicitations, involving two tensile strains, one compression, and two shear strains (see Figure 2(b)). Each RVE comes with an output database containing 39 output variables (see Table 5) that are columns and 1400 raws that corresponds to the element of discretization.

5.2. Features Selection

This phase involved the critical process of feature selection, aiming to pinpoint the highly correlated features in the model related to the strain degradation variable. To achieve this, we employed the Pearson correlation matrix, a statistical measure gauging the linear relationship between multiple variables. The resulting correlation matrix provides insights into the interrelationships among the various features in our dataset. The Pearson correlation coefficient ranges from -1 to 1, reflecting the strength and direction of a linear relationship.

Examining Figure 7, features with Pearson correlation values greater than 0.70 are considered the most positively correlated. To refine our feature selection, we applied the Recursive Feature Selection algorithm (RFS) [48], specifically the recursive feature elimination (RFE) variant. This preprocessing step is crucial for preventing over-fitting, facilitating faster prediction and training, minimizing storage requirements for both model and dataset, and ultimately producing a more interpretable model.

In our study, the RFE algorithm successfully reduced the features to a set of twelve columns, namely: [LE.LE23, LE.LE22, LE.LE12, PE.PE33,

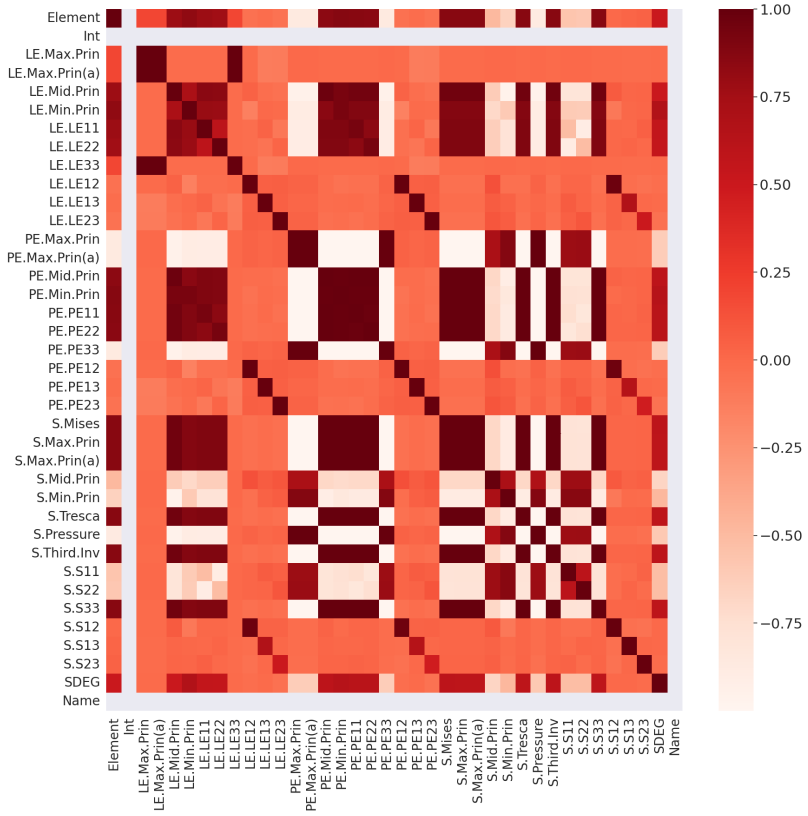

Figure 7: Pearson matrix of correlation of the model features with the SDEG column values.

PE.PE11, PE.Min.Prin, PE.PE23, LE.LE33, PE.PE22, LE.Min.Prin, PE.PE12, SDEG]. This selection, as depicted in Figure 8, represents the most crucial features obtained from the ABAQUS engine, streamlining the dataset and enhancing the efficiency and interpretability of our model.

5.3. Classification Method

The approach throughout using k-NN algorithm had been many times used to classify and detect the existence of patterns and to rank the most important features. In the after mentioned study [49] welding defects caused by laser process was studied to extract the most important features. In the second study [50], It was possible to extract several feature sets to allow multi-dimensional classification of damage modes. The presented research of principal component analysis and pattern recognition methods (k-NN,

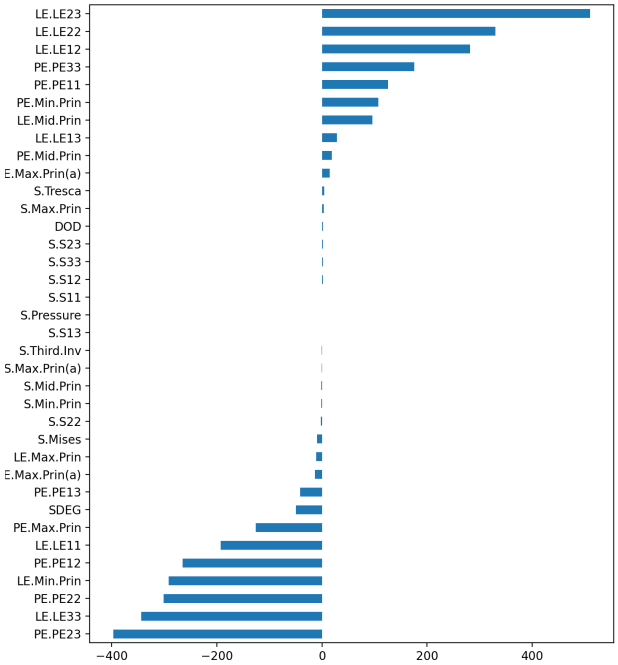

Figure 8: Bar charts plot representing the ordered importance of each feature according the RFE algorithm. Twelve column features are the most important.

decision tree, neural networks), has the goal of providing the presence, type, and severity of damage with a high degree of confidence. The last point is this review article [51] that focus on damage assessment on composite materials (detection, quantification and localization). They point out the fact that the success of the machine learning paradigm for damage assessment depends on the representational capability of the discriminative features for the problems of interest.

In the subsequent phase of our investigation, we aim to employ a classification approach to distinguish between different degrees of degradation. It exists several type of classification methods mostly based on statistical, probability or pattern recognition. Since the features of the dataset is heterogeneous with mixed data (numerical and categorical), only non-parametric methods like k-NN [52] fit with the problem. Operating on the assumption that similar objects in the data are close to each other based on a specified metric, the k-NN algorithm is particularly suited for our task.

The process of k-NN algorithms is as follows. Given an observation X representing the degradation feature of the RVE we aim to predict, the k-NN algorithm identifies similar objects in the dataset and looks for the k points closest to the given observation. Subsequently, it uses the output variable associated with these nearest neighbors to predict the value of X (i.e., degradation). Noteworthy benefits of k-NN include its simplicity, ease of implementation, and non-parametric nature, eliminating the need for tuning multiple hyperparameters. However, its drawback lies in a significant slowdown as the volume of data increases.

The k-NN model classifies different RVEs according to the previously defined degradation law. This approach provides valuable output indications about the characteristics of the composite. It enables insights into identifying RVEs similarity in term of behaviors and degradation under the five different solicitations. This information, in turn, allows us to deduce the optimal compositions of fiber and matrix, as well as the best volume fraction F_v.

Our approach begins by defining our features (X) and target (y), where the target corresponds to the label of the simulation (referred to as Name in the dataset). Subsequently, we perform a train-test split. Scaling is then executed using the MinMaxScaler function, which fits and transforms the train-test set. The k-NN model is implemented, fitted to our test set, and the training and test sets are scored separately. In this case, the number of k neighbors is set to the default value of 5.

5.4. Performance Metrics

To assess the effectiveness of our model, various performance metrics are employed to evaluate the classification prowess of the k-NN algorithm. The selection of these metrics is pivotal as it provides insight into how well ML algorithms perform and facilitates meaningful comparisons.

Currently, our focus lies on the most common metrics utilized for classification problems: accuracy, precision, recall, and F1-score. The Confusion Matrix serves as a valuable tool for visualizing the accuracy of the classification [53].

A **confusion matrix** is a table that describes the performance of a classification model. It presents a summary of the model's predictions and the actual outcomes. **Accuracy** represents the ratio of correctly predicted instances to the total instances in the dataset. **Precision** is the ratio of correctly predicted positive observations to the total predicted positives. It assesses the accuracy of the positive predictions. **Recall** is the ratio of correctly predicted positive observations to the all observations in the actual class. It measures the model's ability to capture all relevant instances. **F1-Score** is the weighted average of precision and recall. It provides a balance between precision and recall.

5.5. Hierarchical Representation

To gain a deeper understanding of the k-NN classification, a valuable approach is to employ an agglomerative clustering algorithm. This algorithm operates in a "bottom-up" manner, initially treating each object as a single-element cluster (leaf). At each step, the algorithm combines the two clusters that are most similar according to Euclidean distance, creating a larger cluster (nodes). The agglomerative hierarchical clustering involves the following steps:

1. **Similarity Matrix:** Calculate the similarity distance between every pair of objects in the dataset.

2. **Linkage:** Utilize a linkage function to group objects into a hierarchical cluster tree based on the distance information generated in step 1. Objects or clusters in close proximity are linked together using the linkage function (in occurence, the ward function).

3. **Cut:** Determine where to cut the hierarchical tree into clusters. This step results in a partition of the data. In this study, the cut is done

when the difference of distance between two levels is more than the total distance of the lower levels.

5.6. Considering Dataset as Time Series and Unsupervised

Dynamic Time Warping (DTW) [54] is an algorithm designed for aligning temporal sequences and measuring their similarities. It enables local shifts, contractions, and stretches in temporal sequences and computes a globally optimal alignment path between two given sequences, subject to certain restrictions [55]. The DTW distance, often used to quantify similarity, is the sum of point-wise distances along the alignment path, expressed as:

$$D(P, Q) = \sum_{(i,j)\in\Pi} d(i, j) \tag{14}$$

Here, Π denotes the path between two sequences P and Q, and (i, j) represents a pair of matched points on the alignment path. The distance function $d(i, j)$ is the squared Euclidean distance between points i and j [55]. Figure 9 illustrates the extraction of the optimal path between two sequences, represented by a red line.

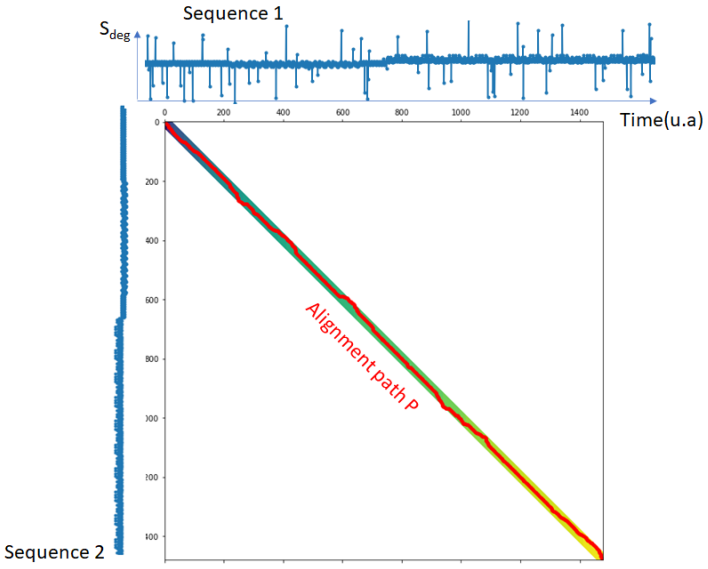

Figure 9: Illustration of the optimal warping path extracted from two sequences of strain degradation SDEG. (*red line*).

The focus shifts to training a model to accurately predict the class of degradation using a dataset with labeled time sequences. To address this challenge, the K-Nearest Neighbors algorithm is once again employed, but this time, instead of relying on the classic Euclidean distance, the implementation utilizes the DTW distance. Practically, the Python package Dtaidistance [56] is employed, and the k-NN classification leverages a precomputed distance to group time series into different clusters.

To demonstrate the quality of results of the DTW distance, we use unsupervised hierarchical clustering (thus without the knowledge of the kind of stess, neither the kind of composite polymer).

6. Results

6.1. k-NN Classification Results

In Figure 10, the classification outcomes, based on the degradation feature $SDEG$, are illustrated. We utilized the established confusion matrix method from the scikit-learn library [57]. This tabular summary offers an overview of both correct and incorrect predictions made by the k-NN classifier.

In this visual representation, each RVE, identified by its unique composition (matrix, fiber, volume fraction), is labeled (represented by numbers). The k-NN classification model's performance hinges on the comparison between the actual labels of the RVEs and the predicted labels acquired during the learning step. The model's accuracy is readily discerned by examining the diagonal of the matrix. Dark blue diagonal values indicate proficient model performance, while off-diagonal values suggest lower accuracy in classification. Notably, this distinction is more apparent for labeled models in the intervals $[120, 135]$ and $[160, 180]$. These cases pertain to two distinct sets of models under shear solicitations in the (XY) and (YZ) planes, respectively. This lower distinguishability in classification is likely influenced by the fiber's orientation along the z direction. Additionally, this reduced classification distinction corresponds particularly to models with the lowest degradation under shear solicitation, as depicted in Figure 4. After the learning phase through training and validation on the validation set, evaluating the model's performance becomes feasible. The performance metrics used are summarized in Table 6. On average over 200 RVEs, the precision values are quite satisfactory for k-NN classification. A complete Grid search cross validation to fine tune the k-NN parameters has been run. Table 6 summed up the best parameters that had been obtained. This represents 72 models

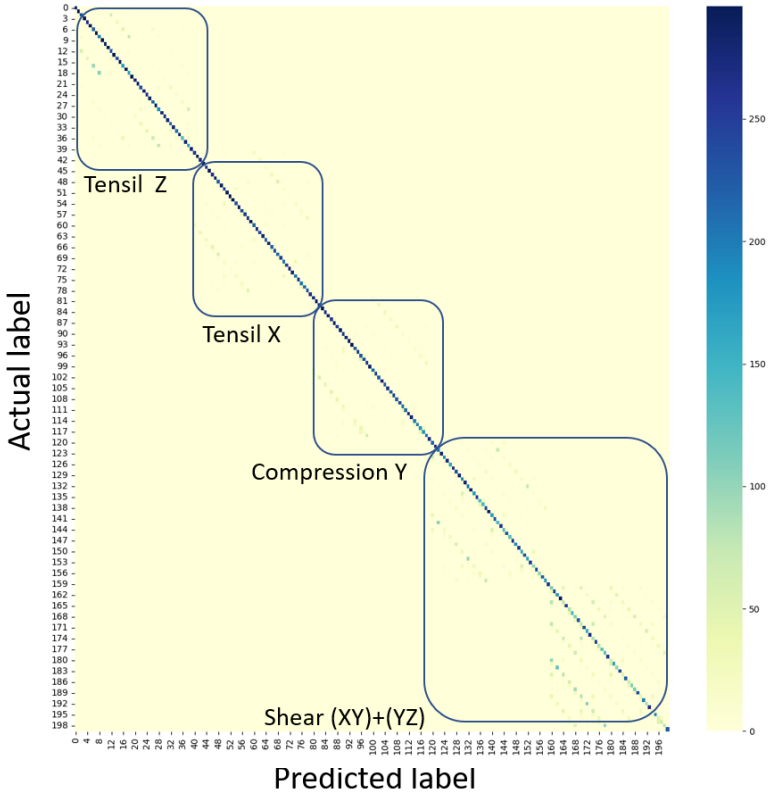

Figure 10: Confusion matrix of the k-NN classification for the different 200 RVEs models after data augmentation and recursively features elimination. RVEs corresponding to shear (XY) and (YZ) are less well classified because of the out of diagonal values. Whereas the other solicitations: tensile strength in X,Z direction and compression in Y direction are fully located on the diagonal.

of k-NN to identify the best parameters according to a metric of accuracy. Those 72 models had been trained then fitted. The confusion matrix used for visualization reveals that the initial RVEs labels closely approach unity for precision, recall, and f1-score. This implies that the k-NN algorithm effectively classified the three type of solicitations of tensile stress in (X) and (Z) and compression stress, whereas the shear solicitation metrics exhibit lower values in comparison.

Table 6: Averaged metric values overs the whole FEM models for the classification of the 200 augmented RVEs.

Best model	k	Weights	metric	
K neighbors classifier	4	distance	Manhattan	
Metrics	precision	recall	f1-score	support
accuracy			0.81	57200
macro avg	0.82	0.82	0.81	57200
weighted avg	0.82	0.81	0.81	5730

6.2. Hierarchical classification from k-NN results

Applying a hierarchical clustering on k-NN classification clusters approach was the principal variation to identify different patterns of degradation similar to this article [58]. In Figure 11(a), the hierarchical representation results are presented. The k-NN classification applied on the test set had predicted four clusters. The dendrogram's in Y direction named distance sensitivity revealed that C1 and C2 clusters *(orange color)* are associated with tensile stress in X and Z directions, while C3 encompasses all compressions in the Y direction. The most populated cluster, C4, groups RVEs subjected to shear in both (XY) and (YZ) planes. Figure 11(b) is a two-dimensional plot of nodes of each RVE affected by degradation. It illustrates how clusters were distributed by the k-NN classification. It allowed to conclude that cluster C4 of RVE submitted to shear has more nodes affected by degradation in comparison with C2, C3 clusters. At contrary to cluster C1 (RVEs submitted to tensile Z), the number of affected nodes is more restricted. k-NN punctual analysis enables concluding how far a material will be damaged based on the cluster assignment

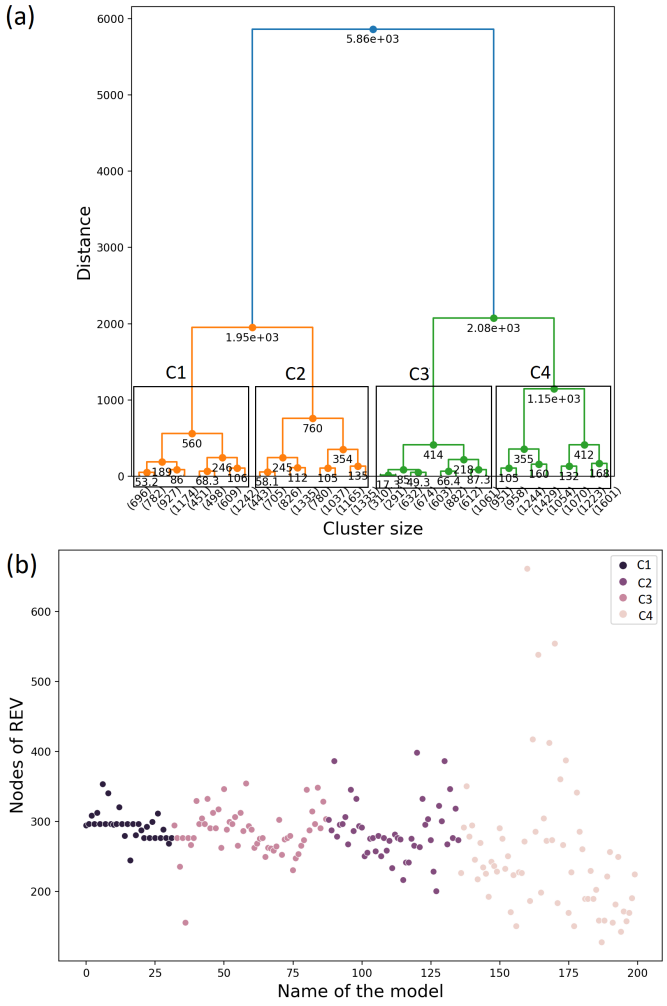

Figure 11: (a) Hierarchical agglomerative clustering of 200 (RVEs) showing 4 main separable clusters. In y direction the euclidean distance report how far the clusters are to one another.(b) Scatter plot graph of the predicted four main clusters resulting from the k-NN classification.

Figure 12 compiles box-plots summarizing statistical values of the main features selected by the RFE algorithm (twelve columns in the dataset). It depicts feature characteristics through mean values, standard deviation, and quantiles, grouping values of each feature of the train set according to the four clusters obtained previously. Consistent patterns emerge for each cluster, particularly for PE.PE33, PE.PE11, PE.PE22, and SDEG, showcasing similar shapes of inter-quartile ranges.

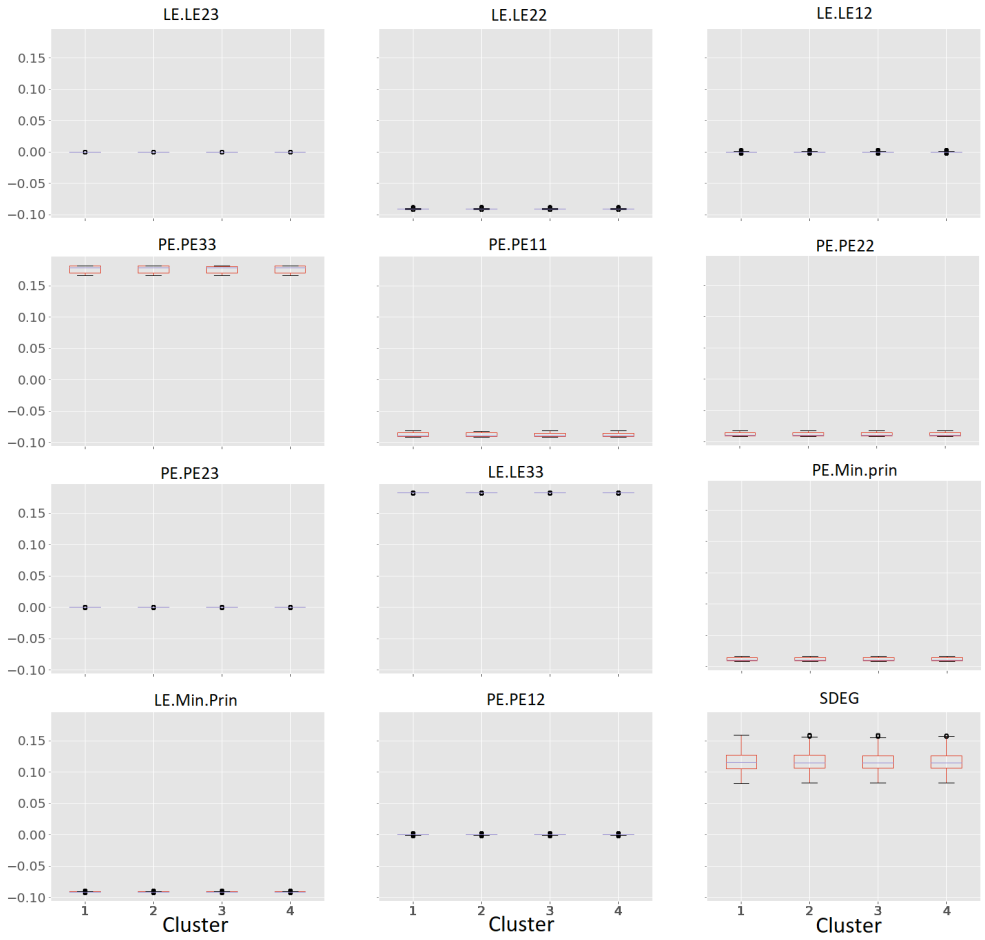

Figure 12: The statistical representation of the extracted features for the four clusters. Boxplot grouping the twelve features selected previously by cluster name C1, C2, C3, C4. Same pattern of boxes according to component of plastic strain (PE11, PE22, PE33) is observed.

The results of the k-NN model are compared with those of other algorithm models in the next table, where the results are summarizes for random forest (RF), support vector classifier (SVC), and decision tree (DT) across 6 consecutive trials with different splits (see Table 7).

Table 7: Sum up table of accuracy of classification between models of random forest, decision tree and support vector classifier.

Models	precision	cross validation	CPU time (min)
k-NN	0.82	6	1
Random forest	0.91	6	37
Decision tree	0.86	6	45
Support vector classifier	0.65	6	120

From the table 7, it is evident that RF outperforms the other models in classification accuracy. However, it is worth noting that RFs can be challenging to interpret and act as a "black box" algorithm. The DT classifier shows slightly better performance than the k-NN algorithm but with a longer execution time, and its results are also more challenging to interpret. SVC is notably less suitable and less efficient compared to the other models.

The performance of k-NN is closely tied to the distance metric used. While the Euclidean distance is simple and effective in some cases, extensive research [59, 60] suggests that k-NN performance can be enhanced by learning another distance metric (e.g., Mahalanobis distance). Let's see the results with DTW distance.

6.3. Results with DTW distance

It is important to notice that the dataset used during this clustering had no input data or features to inform of the external solicitation applied during (FEM) simulation.

Figure 13(a) illustrates a complete hierarchical tree using the hierarchical Tree Wrapper class algorithm. The tree provides an order in dissimilarity, ranging from the cluster with the least dissimilar time series, C1 (representing tensile strength along the Z-direction), to the most distinct time series cluster, C3, corresponding to compression in the Y-direction sequences. Notably, there is a substantial disparity in the tree distance index value (7.14).

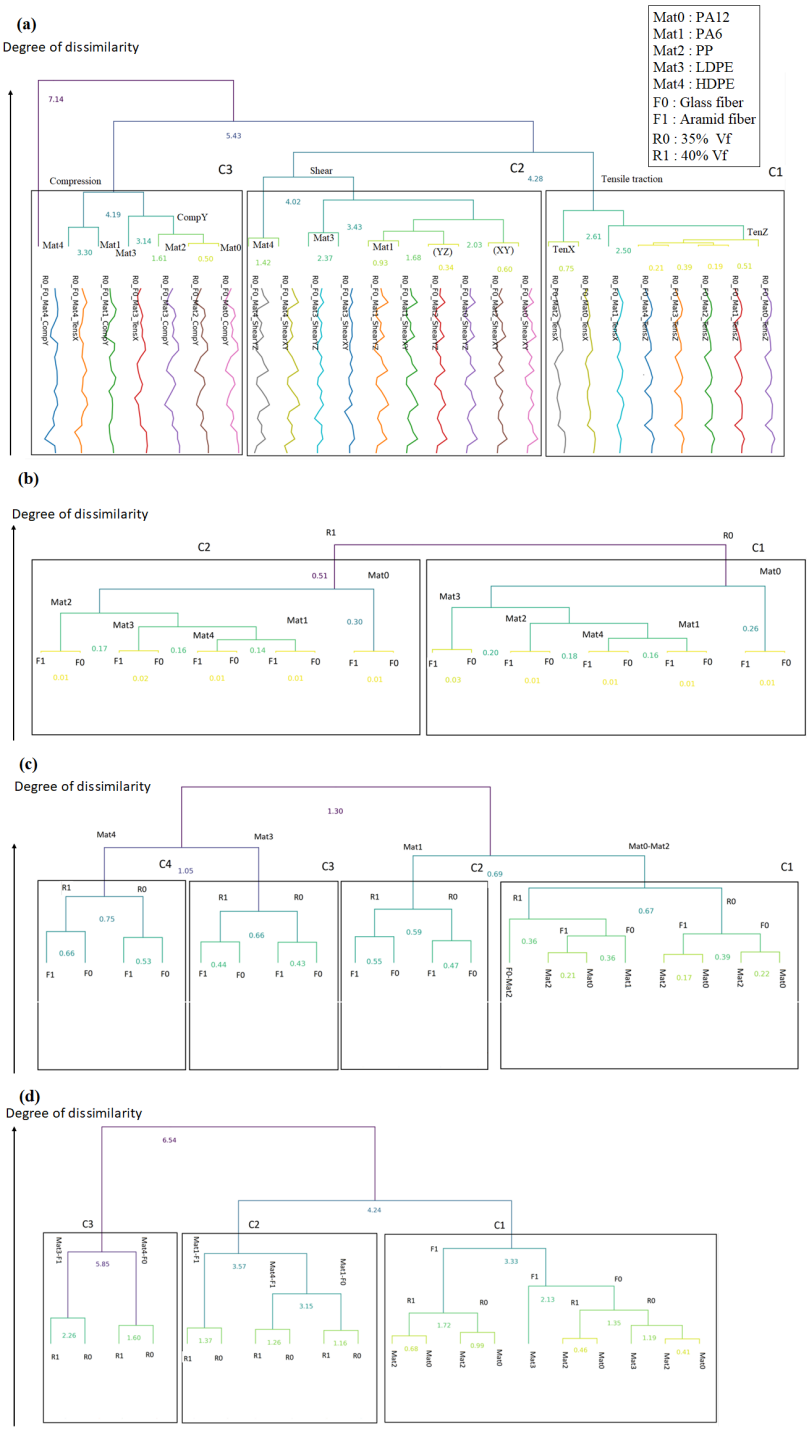

Figure 13: **(a)** Hierarchical clustering tree for 25 different (RVEs) for temporal degradation under different solicitations. Three clusters groups the fives different composites were fiber type $F0$ and volume's fraction $R0$ are fixed. **(b)** Hierarchical clustering tree for 20 different (RVEs) temporal degradation during tensile Z solicitation. **(c)** Hierarchical clustering tree for 20 different (RVEs) temporal degradation during shearing (YZ) solicitation. The level of node in tree structure informs about the impact in degradation. **(d)** Hierarchical clustering tree for 20 different (RVEs) temporal degradation during compression solicitation in Y direction. The optimal number of clusters is equal to three.

The cluster that aggregates RVEs experiencing degradation due to compression exhibits the highest dissimilarity compared to the other two clusters. This classification based on DTW distance appears to blend sequences of degradation resulting from shear in the (XY) and (YZ) planes. This lack of dissimilarity in shear solicitation was previously observed with the simple k-NN approach.

6.3.1. Tensile Z

In this section, a more systematic study was undertaken to evaluate the DTW k-NN classification. For a fixed external solicitation involving Z-tensile strength, shear in the (YZ) plane, and compression along the Y direction, all other parameters (Matrix type, fiber type, volume fraction) were systematically explored. As reminder, those parameters are not included as inputs in the dataset during the classification.

The dendrogram in Figure 13(b) illustrates a hierarchical tree showcasing the shortest distances between 20 RVEs. The present time series depict the evolution of the degradation feature $SDEG$ under Z-tensile strength solicitation over time. Five different materials $[Mat0 - Mat4]$, two fibers $(F0, F1)$, and two volume fractions $(R0, R1)$ are classified, as indicated in (3).

Reading the tree from the leaf nodes (graph of the shortest path) to the root nodes reveals that the difference in fibers $(F0 - F1)$ has minimal dissimilarity in degradation (index tree path equal to 0.01). However, the nature of the matrix material induces dissimilarities significant enough to categorize each fiber pair $(F0 - F1)$ into separate nodes (Mat0, Mat1, Mat2, Mat3, Mat4). The primary feature differentiating in terms of degradation is the volume fraction determined by the radius of the fiber. It is evident that radius $R0$ and $R1$ are the two nodes splitting the tree into two clusters, each with an equal population of RVEs.

From this agglomerative hierarchical tree, we can deduce that when applying tensile strength in the direction of the fiber, the radius of the fiber (i.e., volume fraction) plays a crucial role in degradation.

6.3.2. Shear in (YZ) plan

In the scenario of a constant shear (YZ) solicitation, the Figure 13(c) illustrates how the (DTW) k-NN classification delineates the impact of input features on the final degradation, ranging from the least affected RVEs to

the most affected ones. The dendrogram exhibits a distinct tree structure in this case.

In Cluster C1, representing minimal dissimilarities, the nature of the matrix material for $Mat0$ and $Mat2$ (PA12 and PP, respectively) exhibits negligible impact, possibly due to their closely matched Young's modulus values ($E_{PA12} = 1250\ MPa$ and $E_{PP} = 1550\ MPa$, see Table 3).

However, in Clusters C2, C3, and C4, where the (XY) shear induces a more pronounced response, the nature of the matrix material plays a more significant role. Higher dissimilarity levels of 1.05 are observed with matrix materials $Mat1, Mat3, Mat4$, indicating a distinct tree structure within these clusters.

6.3.3. Compression in Y-direction

In the context of a fixed compression in the Y direction, a hierarchical tree dendrogram for degradation is obtained through (DTW) k-NN classification applied to 20 RVEs. Figure 13(d) reveals that this mode induces significant dissimilarities in degradation behavior.

The tree can be divided into three clusters based on large values of the degree of dissimilarity index: 4.24 between clusters $C1$ and $C2$, and 6.54 for cluster $C3$. Cluster $C1$ encompasses $Mat0, Mat2, Mat3$, corresponding to $PA12$, PP, $LDPE$, with both glass fiber and aramid glass. Clusters $C2$ and $C3$ share the same tree structure in terms of nodes.

Interestingly, the radius of the fiber ($R0$, $R1$), situated low in the tree, imparts less difference in degradation compared to the nature of the fiber (glass or aramid). For instance, cluster C3 groups $Mat3 = LDPE$ when reinforced with Aramid fiber $F1$, and $Mat4 = HDPE$ with glass fiber. This underscores the significance of the nature of the reinforcement fiber, ultimately altering the material's response to compression in the direction perpendicular to the fiber's orientation.

7. Discussions and perspectives

7.1. Damage mode detection by k-NN classification

In this study, two classification methods, simple k-NN and (DTW) k-NN, were employed to compare the accuracy of classification. The (DTW) classification provided more nuanced insights through clustering. It became evident that tensile forces in the direction of the Z fiber were the least destructive in terms of ductile damage to the matrix-fiber material. The anisotropic nature

of the fiber played a protective role, with the radius and nature of the fiber further enhancing this effect. The dendrogram resulting from the (DTW) k-NN classifier for shear in the (YZ) plane solicitation revealed a tree structure similar to the Z-tensile tree. This observation underscores the importance of reinforcing the polymer matrix with multi directional fibers, facilitating strain transfer along the fiber directions. Additionally, the (DTW) k-NN classifier highlighted the higher damage susceptibility of Y-compression, especially in matrices like HDPE and LDPE, which have lower Young modulus values regardless of the fiber's nature. Lower Young's modulus values imply that these matrices are more compliant, making them more susceptible to deformation and damage under loading conditions. The susceptibility to damage under Y-compression primarily depends on the properties of the matrix material. Even if the reinforcing fibers possess higher strength and stiffness, the overall performance of the composite material can be compromised if the matrix material is unable to effectively transfer and distribute loads [61]. Which is the present case.

7.2. Comparison of degradation between composite

For a more comprehensive understanding, a comparison of the five composites in terms of ductile damage was conducted through agglomerative trees. The analysis revealed two distinct sets of matrices.

Set 1: This group comprises materials with high Young modulus values, namely Mat0 and Mat2 (PA12, PP). During tensile strength in the Z direction, they exhibited the highest dissimilarity, occupying opposite positions in the dendrogram (see Figure 13(b)). For shear (YZ) solicitation in both matrix materials PP and PA12, with either glass or aramid fiber, they have been demonstrating similar degradation, forming cluster C1 13(c)). This can be related to the identical value of shear modulus (400 MPa). The shear modulus provides insights into the material's behavior when subjected to shear forces perpendicular to the (YZ) plane. Therefore, while the fibers in glass or aramid composites may be relatively insensitive to shear external solicitation, the response of the matrix is important for ensuring the overall mechanical performance of the composite material. In composite materials where the fibers are glass or aramid, it's generally true [62, 45] that these fibers are relatively insensitive to shear external solicitation compared to the matrix. This is because glass and aramid fibers typically have high tensile strength and stiffness, making them resistant to deformation under shear stress. In such composites, the fibers provide strength and stiffness to the

composite structure, while the matrix material serves to transfer loads between the fibers. However, while the fibers themselves may be less sensitive to shear stress, the response of the matrix under shear loading is still important

Set 2: This set includes Mat1, Mat3, and Mat4, representing matrices with lower Young modulus values (800, 1080, 565 MPa, respectively). These matrices occupied central clusters with lower dissimilarities between each other in each dendrogram for tensile, shear, or compression loads.

This initial exploration of composite material degradation through the hybrid (FEM-ML) method provides valuable insights for application in real data obtained through experimental means. It is crucial to note that this innovative combination of FEM and ML should be further assessed through an experimental study investigating other types of damage laws.

8. Conclusion

The study presents a new method that integrates Finite Element Method (FEM) with Machine Learning (ML) to analyze composite degradation. The method uses Dynamical Time Warping (DTW) for KNN classification and visualizes the degree of damage using agglomerative clustering trees. The study suggests that the hybrid FEM-ML approach has potential in predicting damage patterns under high external pressure. The primary mode of degradation identified is perpendicular compression to the fiber direction. The study's main contributions are automated multi-class clustering of composite degradation and recurrent feature elimination analysis, which identified twelve essential features related to degradation. The k-NN algorithm was implemented and achieved an 82% accuracy in classifying composite degradation under various external solicitations. Future research could focus on increasing the complexity of numerical simulations, such as incorporating multiple fibers with more intricate loading paths. Furthermore, the hybrid FEM-ML approach could potentially benefit from the implementation of deep learning techniques. This is particularly relevant due to the large number of input parameters involved, which could enhance the model's predictive capabilities and provide valuable insights into composite degradation mechanisms.

Data Availability Statements

The data that support the findings of this study are available from the authors upon reasonable request and with permission of [De Vinci Research

Center].

Acknowledgements

The authors would like to thank the DeVinci Research center (DVRC) for their implication.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] M. H. Nemat-Nasser, M. Micromechanics: overall properties of heterogeneous materials, volume 37, North-Holland, 1993.

[2] P. Kanout, D. P. Boso, J. Chaboche, B. Schrefler, Multiscale Methods for Composites : A Review (2009) 31–75.

[3] V. P. Nguyen, M. Stroeven, L. Sluys, Multiscale continuous and discontinuous modeling of heterogeneous materials 3 (2011) 229–270.

[4] S. Li, A. Wongsto, Unit cells for micromechanical analyses of particle-reinforced composites, Mechanics of Materials 36 (2004) 543–572.

[5] H. Ahmadi, M. Hajikazemi, W. Van Paepegem, Closed-form formulae for prediction of homogenized ply-properties and laminate thermoelastic constants in symmetric laminates containing ply cracks in multiple orientations, Composite Structures 241 (2020) 112061.

[6] K. Daly, T. Roose, Homogenization of two fluid flow in porous media, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 471 (2015) 20140564–20140564.

[7] A. Taliercio, R. Coruzzi, Mechanical behaviour of brittle matrix composites: A homogenization approach, International Journal of Solids and Structures 36 (1999) 3591–3615.

[8] G. X. Gu, C. T. Chen, M. J. Buehler, De novo composite design based on machine learning algorithm, Extreme Mechanics Letters 18 (2018) 19–28.

[9] H. Bao, S. Wu, Z. Wu, G. Kang, X. Peng, P. J. Withers, A machine-learning fatigue life prediction approach of additively manufactured metals, Engineering Fracture Mechanics 242 (2021) 107508.

[10] M. Islam, M. Sohaib, J. Kim, J. M. Kim, Crack classification of a pressure vessel using feature selection and deep learning methods, Sensors (Switzerland) 18 (2018).

[11] A. Kaveh, A. Dadras Eslamlou, S. M. Javadi, N. Geran Malek, Machine learning regression approaches for predicting the ultimate buckling load of variable-stiffness composite cylinders 232 (2021) 921–931.

[12] Regulation no 134 of the economic commission for europe of the united nations (un/ece) — uniform provisions concerning the approval of motor vehicles and their components with regard to the safety-related performance of hydrogen-fuelled vehicles (hfcv) [2019/795], 2019. URL: http://data.europa.eu/eli/reg/2019/795/oj.

[13] S. Chibani, F. X. Coudert, Machine learning approaches for the prediction of materials properties, APL Materials 8 (2020).

[14] S. Skovsgaard, S. Heide-Jørgensen, Three-dimensional mechanical behavior of composite with fibre-matrix delamination through homogenization of micro-structure, Composite Structures 275 (2021) 114–418.

[15] Z. Ullah, L. Kaczmarczyk, C. J. Pearce, Three-dimensional nonlinear micro meso mechanical response of the fibre-reinforced polymer composites, Composite Structures 161 (2016) 204–214.

[16] F. de Francqueville, P. Gilormini, J. Diani, Representative volume elements for the simulation of isotropic composites highly filled with monosized spheres, International Journal of Solids and Structures 158 (2019) 277–286.

[17] J. Correia, N. Apetre, A. Arcari, A. De Jesus, M. Muñiz-Calvente, R. Calçada, F. Berto, A. Fernández-Canteli, Generalized probabilistic model allowing for various fatigue damage variables, International Journal of Fatigue 100 (2017) 187–194.

[18] Z. Hashin, A. Rotem, A fatigue criterion for fiber-reinforced materials, Journal of. Composite Materials 7 (1973) 448–464.

[19] Z. Hashin, Failure criteria for unidirectional fiber composites, Journal of Applied Mechanics 47 (1980) 329–334.

[20] A. Matzenmiller, J. Lubliner, R. Taylor, A constitutive model for anisotropic damage in fiber-composites, Mechanics of Materials 20 (1995) 125–152.

[21] P. Camanho, C. Dávila, Mixed-mode decohesion finite elements for the simulation of delamination in composite materials, NTRS - NASA Technical Reports Server (2002).

[22] T. Drabek, H. J. Böhm, Damage models for studying ductile matrix failure in composites, Computational Materials Science 32 (2005) 329–336.

[23] A. I. Mel'ker, A. V. Ivanov, Computer modeling of dynamics of damage cumulation in the ductile matrix of a fiber-reinforced composite, Journal of Mechanics of Composite Materials 23 (1988) 682–685.

[24] I. Lapczyk, J. A. Hurtado, Progressive damage modeling in fiber-reinforced materials, Composites Part A: Applied Science and Manufacturing 38 (2007) 2333–2341.

[25] Z. Marciniak, K. Kuczyński, Limit strains in the processes of stretch-forming sheet metal, International Journal of Mechanical Sciences 9 (1967) 609–620.

[26] H. Hooputra, H. Gese, H. Dell, H. Werner, A comprehensive failure model for crash worthiness simulation of aluminium extrusions, International Journal of Crashworthiness 9 (2004) 449–464.

[27] R. K. Abu Al-Rub, G. Z. Voyiadjis, On the coupling of anisotropic damage and plasticity models for ductile materials, International Journal of Solids and Structures 40 (2003) 2611–2643.

[28] K. Terada, M. Hori, T. Kyoya, N. Kikuchi, Simulation of the multi-scale convergence in computational homogenization approaches, International Journal of Solids and Structures 37 (2000) 2285–2311.

[29] C. Sun, R. Vaidya, Prediction of composite properties from a representative volume element, Composites Science and Technology 56 (1996) 171–179.

[30] M. Arif, N. Saintier, F. Meraghni, J. Fitoussi, Y. Chemisky, G. Robert, Multiscale fatigue damage characterization in short glass fiber reinforced polyamide-66, Composites Part B: Engineering 61 (2014) 55–65.

[31] J. Fitoussi, F. Meraghni, Z. Jendli, G. Hug, D. Baptiste, Experimental methodology for high strain-rates tensile behaviour analysis of polymer matrix composites, Composites Science and Technology 65 (2005) 2174–2188.

[32] H. Fujiwara, H. Ono, K. Ohyama, M. Kasai, F. Kaneko, S. Nishimura, Hydrogen permeation under high pressure conditions and the destruction of exposed polyethylene-property of polymeric materials for high-pressure hydrogen devices, International Journal of Hydrogen Energy 46 (2021) 11832–11848.

[33] H. Mehdipour, P. P. Camanho, G. Belingardi, Elasto-plastic constitutive equations for short fiber reinforced polymers, Composites Part B 165 (2019) 199–214.

[34] C. mbH, Campus® - a material information system for the plastics industry, 2022. URL: `https://www.campusplastics.com/`.

[35] D. Pulungan, G. Lubineau, A. Yudhanto, R. Yaldiz, W. Schijve, Identifying design parameters controlling damage behaviors of continuous fiber-reinforced thermoplastic composites using micromechanics as a virtual testing tool, International Journal of Solids and Structures 117 (2017) 177–190.

[36] A. Ali, M. Hosseini, B. B. Sahari, A review of constitutive models for rubber-like materials, American Journal of Engineering and Applied Sciences 3 (2010) 232–239.

[37] T. Beda, An approach for hyperelastic model-building and parameters estimation a review of constitutive models, European Polymer Journal 50 (2014) 97–108.

[38] M. Okereke, S. Keates, Finite element applications: A Practical Guide to the FEM Process, Springer International Publishing, 2018.

[39] Q.-C. He, Effects of size and boundary conditions on the yield strength of heterogeneous materials, Journal of the Mechanics and Physics of Solids 49 (2001) 2557–2575.

[40] V. Kouznetsova, W. Brekelmans, F. Baaijens, An approach to micro-macro modeling of heterogeneous materials, Computational mechanics 27 (2001) 37–48.

[41] A. Akpoyomare, M. Okereke, M. Bingley, Virtual testing of composites: Imposing periodic boundary conditions on general finite element meshes, Composite Structures 160 (2017) 983–994.

[42] J. Tyrus, M. Gosz, E. D. Santiago, A local finite element implementation for imposing periodic boundary conditions on composite micromechanical modrabekdels, International Journal of Solids and Structures 44 (2007) 2972–2989.

[43] I. Gitman, H. Askes, L. Sluys, Representative volume: Existence and size determination, Engineering Fracture Mechanics 74 (2007) 2518–2534.

[44] M. Jerabek, D. Tscharnuter, Z. Major, K. Ravi-Chandar, R. Lang, Multiaxial yield behaviour of polypropylene, EPJ Web of Conferences 6 (2010).

[45] L. Wan, Y. Ismail, C. Zhu, P. Zhu, Y. Sheng, J. Liu, D. Yang, Computational micromechanics-based prediction of the failure of unidirectional composite lamina subjected to transverse and in-plane shear stress states, Journal of Composite Materials 54 (2020) 3637–3654.

[46] D. Savvas, G. Stefanou, M. Papadrakakis, Determination of RVE size for random composites with local volume fraction variation, Computer Methods in Applied Mechanics and Engineering 305 (2016) 340–358.

[47] S. M. Mirkhalaf, F. M. Andrade Pires, R. Simoes, Determination of the size of the Representative Volume Element (RVE) for the simulation of heterogeneous polymers at finite strains, Finite Elements in Analysis and Design 119 (2016) 30–44.

[48] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, Journal of machine learning research 3 (2003) 1157–1182.

[49] S. H. Lee, J. Mazumder, J. Park, S. Kim, Ranked Feature-Based Laser Material Processing Monitoring and Defect Diagnosis Using k-NN and SVM, Journal of Manufacturing Processes 55 (2020) 307–316.

[50] S. S. Kessler, P. Rani, Pattern recognition for damage characterization in composite materials, Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 8 (2007) 8507–8517.

[51] A. Khan, N. Kim, J. K. Shin, H. S. Kim, B. D. Youn, Damage assessment of smart composite structures via machine learning: a review, JMST Advances 1 (2019) 107–124.

[52] G. Guo, H. Wang, D. Bell, Y. Bi, K. Greer, KNN model-based approach in classification, 2003. URL: https://api.semanticscholar.org/CorpusID:1350330.

[53] M. Hossin, M. N. Sulaiman, A review on evaluation metrics for data classification evaluations, International journal of data mining & knowledge management process 5 (2015) 1.

[54] M. Müller, Dynamic time warping, Information retrieval for music and motion 2 (2007) 69–84. doi:10.1007/978-3-540-74048-3_4.

[55] S. Kyaagba, Dynamic time warping with time series, 2022. URL: https://medium.com/@shachiakyaagba_41915/dynamic-time-warping-with-time-series-1f5c05fb8950.

[56] W.Meert, Knn dtw, 2022. URL: https://dtaidistance.readthedocs.io/en/latest/usage/dtw.html.

[57] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, the Journal of machine Learning research 12 (2011) 2825–2830.

[58] A. Maugeri, M. Barchitta, G. Basile, A. Agodi, Applying a hierarchical clustering on principal components approach to identify different patterns of the SARS-CoV-2 epidemic across Italian regions, Scientific Reports 11 (2021) 1–9.

[59] A. Bellet, A. Habrard, M. Sebban, A survey on metric learning for feature vectors and structured data, CoRR abs/1306.6709 (2013). URL: http://arxiv.org/abs/1306.6709. arXiv:1306.6709.

[60] K. Q. Weinberger, L. K. Saul, Distance metric learning for large margin nearest neighbor classification, Journal of Machine Learning Research 10 (2009) 207–244.

[61] W. Grellmann, B. Langer, Deformation and Fracture Behaviour of Polymer Materials, volume 247, 2017. doi:10.1007/978-3-319-41879-7.

[62] Fracture behaviour of fibre-reinforced composite materials subjected to shear loading: An experimental and numerical study, International Journal of Lightweight Materials and Manufacture 6 (2023) 108–116.