

Data-driven constitutive models for brittle solids displaying progressive anisotropic damage

Weijian Ge, Vito L Tagarielli *

Department of Aeronautics, Imperial College London, SW72AZ, London, UK

ARTICLE INFO

Keywords:

Anisotropic
Damage
Data-driven
Constitutive model

ABSTRACT

We propose and demonstrate a computational framework to obtain data-driven surrogate constitutive models capturing the mechanical response of anisotropic brittle solids displaying progressive anisotropic damage. We train the constitutive models on data obtained from the analysis of a volume element of a material of interest; the data is generated by a constitutive model for braided composites, displaying a complex anisotropic damage evolution progressively transitioning from transversely isotropic to orthotropic. Training involves imposing six-dimensional random strain histories on the physical model and recording the histories of stress, strain and homogenised stiffness matrix of the material, obtained by a set of linear perturbation analyses. Supervised machine learning and dimensionality reduction are applied to the data and a structure for a surrogate model is proposed. The surrogate predicts the evolution of the stiffness of the solid consequent to an arbitrary imposed six-dimensional strain increment, thereby calculating the corresponding increment in stress. The model displays high accuracy and is able to reproduce the homogenised material's response via simple neural networks.

1. Introduction

Progressive anisotropic damage is observed in many quasi-brittle materials, such as composites, wood, bone, geomaterials, concrete, and more [1]. Anisotropic damage is typically modelled, analytically, through the introduction of damage tensors of second [2–4], or fourth order [5,6], although some attempts have been made using tensors of higher order [7]. Different functional forms are typically assumed for these tensors and their evolution with deformation, based on material-specific considerations which too often have only weak connections with quantitative observations. Due to this, several authors [8–10], opt to determine constitutive descriptions using detailed numerical simulations of the response of volume elements of the materials of interest, homogenising such response; apart from the high computational cost and the problems with representativeness of this approach [11], the information learned from these detailed simulations is recorded and used in practice in very simplified ways.

In this work we address this problem by a data-driven approach. Data-driven solutions to constitutive modelling problems hold great promise to improve the accuracy of numerical simulations of material failure, but no studies have focused on data-driven constitutive models for materials with *arbitrarily anisotropic progressive damage*; we propose

here how to fill this gap. Several researchers have built multiscale anisotropic damage models using data-driven approaches. In [12] and [13] efficient ANN models were developed to predict the strength of CFRPs, however only in specific monotonic loading conditions. In [14] a considerable advance is introduced in the form of a surrogate model for composite damage at two distinct scales, however assuming a very specific form of damage initiation and evolution.

We aim at capturing all details of the initiation and evolution of damage in brittle materials, by applying machine learning and dimensionality reduction to an appropriately designed training dataset. The training dataset can be obtained via simulations of volume elements explicitly modelling the details of the material's microstructure, or can in principle directly be obtained from measurements, employing appropriate experimental techniques (see for example [15–17]). In this study the training dataset is obtained from a phenomenological model inspired by simulations of representative volume elements (RVEs) of braided composites [18], based on a modified Hashin damage model. The choice to use a phenomenological model rather than data from the analysis of material RVEs is made to save on computational time, but also to have full control over the structure of the data produced. The training data is obtained by simulating the response of a material element to pseudo-random, six-dimensional histories of

* Corresponding author.

E-mail address: v.tagarielli@imperial.ac.uk (V.L. Tagarielli).

<https://doi.org/10.1016/j.jcomc.2024.100501>

non-proportional and non-monotonic imposed strains, similar to what we demonstrated in previous studies ([19]). The necessary information on the loss of cohesion of the material is obtained by extracting the material's current stiffness matrix during deformation.

The aim of the paper is to propose a technique to effectively 'copy' the material response predicted by numerical analyses of representative or stochastic volume elements, and to reproduce such response at negligible computational cost in multiscale simulations of the response of composite materials.

$$F_{MC} = \left[\left(\frac{Y_C}{2S_t} \right)^2 - 1 \right] \times \left(\frac{\sigma_{22} + \sigma_{33}}{Y_C} \right) + \frac{(\sigma_{22} + \sigma_{33})^2}{4S_t^2} + \frac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{S_t^2} + \frac{\sigma_{12}^2 + \sigma_{13}^2}{S_t^2} \geq 1 \quad (5)$$

(Matrix Compression)

In the next section we introduce the phenomenological model used to generate the data and illustrate the strategy to produce the training dataset. In Section 3 we describe the proposed data-driven modelling strategy, and results are presented and discussed in Section 4.

2. Training dataset

In this section we summarise the features of the phenomenological constitutive model used to produce the training dataset, and we illustrate the numerical details of the simulations conducted to this effect.

2.1. Anisotropic constitutive model for data generation

The phenomenological model used in this study is based on a modification of the Hashin damage model, which is commonly used to model the response of UD fibre-reinforced composites. The Hashin model [20] considers the effects of fibre tearing and microbuckling, matrix cracking, and fibre-matrix interface debonding, which are the primary damage mechanisms that can occur in UD FRPs. Prior to the initiation of a damage mechanism, the material's response is assumed to be orthotropic and elastic, i.e.

$$\sigma_{ij} = C_{ijkl} \varepsilon_{kl} \quad (1)$$

where σ_{ij} , ε_{kl} , C_{ijkl} ($i, j, k, l = 1, 2, 3$) are true stress and strain, and elastic stiffness matrix. The indices 1, 2, 3 refer to the fibre direction, in-plane transverse direction and out-of-plane direction as usual. In [18], Hashin's model is modified to represent braided composites. The choice of this model was driven by the fact that it was readily available in the public domain; in addition, the material stiffness in [18] is assumed to transition from transversely isotropic to orthotropic through a complex damage evolution, governed by four parameters through a complex set of rules, and capturing such complexity is a good test for the proposed machine learning algorithm.

Four different failure modes are introduced in [18]. In presence of tensile stress in direction 1 ($\sigma_{11} \geq 0$), the onset of damage is given by

$$F_{FT} = \left(\frac{\sigma_{11}}{X_T} \right)^2 + \frac{\sigma_{12}^2 + \sigma_{13}^2}{S_t^2} \geq 1 (\text{Fibre Tension}). \quad (2)$$

When the material is compressed in direction 1 ($\sigma_{11} < 0$), damage initiates when

$$F_{FC} = \left(\frac{\sigma_{11}}{X_C} \right)^2 \geq 1 (\text{Fibre Compression}). \quad (3)$$

For tensile loading in the 2–3 plane ($\sigma_{22} + \sigma_{33} \geq 0$), failure initiates when

$$F_{MT} = \left(\frac{\sigma_{22} + \sigma_{33}}{Y_T} \right)^2 + \left(\frac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{S_t^2} \right) + \left(\frac{\sigma_{12}^2 + \sigma_{13}^2}{S_t^2} \right) \geq 1 (\text{Matrix Tension}) \quad (4)$$

while in compression ($\sigma_{22} + \sigma_{33} < 0$) the onset of damage occurs at

In Eqs. (2)–(5) F_{FT} , F_{FC} , F_{MT} , F_{MC} are damage initiation variables; X_T , X_C , Y_T , Y_C are axial tensile strength, axial compressive strength, transverse tensile strength, transverse compressive strength of the material, respectively; S_t , S_l are longitudinal and transverse shear strength, respectively.

Once one of the damage initiation criteria above is satisfied, further straining will result in the degradation of the material's stiffness. Following the original paper [18], the general damage initiation criterion is denoted by F_I , where I is intended as an index taking values FT , FC , MT , MC .

Introduce a characteristic length l_c of the material element under consideration, as the cubic root of the volume of the finite element used to represent such volume in numerical simulations; further, introduce the dissipated energy in this finite element G_I , as

$$G_I = \frac{1}{2} \sigma_{eq}^f \varepsilon_{eq}^f l_c \quad (6)$$

in which again $I = FT, FC, MT, MC$ refers to fibre tension, fibre compression, matrix tension, and matrix compression, where σ_{eq}^f is an equivalent strength and ε_{eq}^f an effective failure strain.

The evolution of damage associated to each of the four failure modes is specified by introducing four damage parameters d_I as

$$d_I = \frac{\delta_{I,eq}^f (\delta_{I,eq} - \delta_{I,eq}^0)}{\delta_{I,eq}^f (\delta_{I,eq}^f - \delta_{I,eq}^0)}, \quad I = FT, FC, MT, MC \quad (7)$$

where $\delta_{I,eq}^f$ is the equivalent displacement at complete failure for a certain failure mode, and $\delta_{I,eq}^0$ is the equivalent displacement at which a certain failure mode is initiated. Such displacements are computed as

$$\delta_{I,eq}^f = \frac{2G_I}{\sigma_{I,eq}^0} \quad (8)$$

$$\delta_{I,eq}^0 = \frac{\delta_{I,eq}}{\sqrt{F_I}} \quad (9)$$

where $\sigma_{I,eq}^0$ denotes the stress at damage initiation for a certain failure mode and F_I are the numerical values of the damage initiation variables defined in Eqs. (2)–(5). The equivalent displacements $\delta_{I,eq}$ and stresses σ for each failure mode can be computed as follows:

For fibre tension:

$$\delta_{FT,eq} = l_c \sqrt{(\varepsilon_{11})^2 + \varepsilon_{12}^2 + \varepsilon_{13}^2} \quad (10)$$

$$\sigma_{FT,eq} = \frac{l_c ((\sigma_{11}) \langle \varepsilon_{11} \rangle + \sigma_{12} \varepsilon_{12} + \sigma_{13} \varepsilon_{13})}{\delta_{FT,eq}} \quad (11)$$

For fibre compression:

$$\delta_{FC,eq} = l_c \langle -\epsilon_{11} \rangle \quad (12)$$

$$\sigma_{FC,eq} = \frac{l_c \langle -\sigma_{11} \rangle \langle -\epsilon_{11} \rangle}{\delta_{FC,eq}}. \quad (13)$$

For matrix tension :

$$\delta_{MT,eq} = l_c \sqrt{\langle \epsilon_{22} \rangle^2 + \langle \epsilon_{33} \rangle^2 + \epsilon_{12}^2 + \epsilon_{13}^2 + \epsilon_{23}^2} \quad (14)$$

$$\sigma_{MT,eq} = \frac{l_c (\langle \sigma_{22} \rangle \langle \epsilon_{22} \rangle + \langle \sigma_{33} \rangle \langle \epsilon_{33} \rangle + \sigma_{12} \epsilon_{12} + \sigma_{23} \epsilon_{23} + \sigma_{13} \epsilon_{13})}{\delta_{MT,eq}}. \quad (15)$$

For matrix compression:

$$\delta_{MC,eq} = l_c \sqrt{\langle -\epsilon_{22} \rangle^2 + \langle -\epsilon_{33} \rangle^2 + \epsilon_{12}^2 + \epsilon_{13}^2 + \epsilon_{23}^2} \quad (16)$$

$$\sigma_{MT,eq} = \frac{l_c (\langle -\sigma_{22} \rangle \langle -\epsilon_{22} \rangle + \sigma_{12} \epsilon_{12})}{\delta_{MT,eq}}. \quad (17)$$

The 3D damaged compliance matrix is taken to have the form [18].

$$S = \begin{bmatrix} \frac{1}{d_f E_{11}} & \frac{\nu_{21}}{E_{22}} & \frac{\nu_{31}}{E_{33}} & & & \\ \frac{\nu_{12}}{E_{11}} & \frac{1}{d_m E_{22}} & \frac{\nu_{32}}{E_{33}} & & & \\ \frac{\nu_{13}}{E_{11}} & \frac{\nu_{23}}{E_{22}} & \frac{1}{E_{33}} & & & \\ & & & \frac{1}{d_f d_m G_{12}} & & \\ & & & & \frac{1}{d_f G_{13}} & \\ & & & & & \frac{1}{d_m G_{23}} \end{bmatrix} \quad (18)$$

and its inverse C is obtained as

$$C = \frac{1}{D} \begin{bmatrix} d_f E_{11} (1 - d_m \nu_{23} \nu_{32}) & d_f d_m E_{11} (\nu_{21} + \nu_{23} \nu_{31}) & d_f E_{11} (\nu_{31} + \nu_{21} \nu_{32}) & & & \\ & d_m E_{22} (1 - d_f \nu_{13} \nu_{31}) & d_m E_{22} (\nu_{32} + d_f \nu_{12} \nu_{31}) & & & \\ & & E_{33} (1 - d_f d_m \nu_{12} \nu_{21}) & & & \\ & & & D d_f d_m G_{12} & & \\ & & & & D d_f G_{13} & \\ & & & & & D d_m G_{23} \end{bmatrix} \quad (19)$$

$$d_f = (1 - d_{FT})(1 - d_{FC}), d_m = (1 - d_{MT})(1 - d_{MC}),$$

$$D = \frac{1}{1 - d_f d_m \nu_{12} \nu_{21} - d_m \nu_{23} \nu_{32} - d_f \nu_{13} \nu_{31} - 2 d_f d_m \nu_{21} \nu_{32} \nu_{13}}.$$

In summary, for an arbitrary 6-dimensional state of strain and stress, 4 pairs of critical (equivalent) damage initiation stress and damage initiation displacements can be calculated, making use of Eqs. (10)–(17). Eq. (7) implies a linear evolution of damage parameters with equivalent displacements; this choice is made for simplicity in the original model. The four values of the displacements at complete failure are specified by Eq. (8), which ensures that the integral of equivalent stress versus displacement curve in each mode is equal to the value specified by the fracture energies G_I . Material properties enter the model via the expressions of the parameters F_I (which contain characteristic strengths) and these fracture energies, as well as via the material's initial stiffness matrix.

The constitutive model above was publically available in the form of a VUMAT subroutine for Abaqus [21]. This subroutine was used with mechanical properties taken from [18] and presented in Table 1.

2.2. Finite element simulations

To generate a training dataset for the surrogate model we use Abaqus Explicit [21] to perform FE simulations of the response of a single FE of type C3D8R. Simulations considered geometrical nonlinearity and used adaptive time stepping. The mass was rescaled to allow larger timesteps, and the magnitude of time was stretched to ensure that the kinetic energy in the simulations was a negligible fraction of the internal energy (considerably lower than 1 %) at all times.

To generate simulation data, we impose on the FEs pseudo-random histories of macroscopic deformation. These are intended, in true strain space, as the histories of the component of the vector $\epsilon = [\epsilon_{11}, \epsilon_{22}, \epsilon_{33}, \gamma_{12}, \gamma_{13}, \gamma_{23}]$. The histories are terminated upon meeting the condition $\sqrt{\epsilon_{11}^2 + \epsilon_{22}^2 + \epsilon_{33}^2 + \gamma_{12}^2 + \gamma_{13}^2 + \gamma_{23}^2} \geq \epsilon_{\max}$ and comprise multiple loading steps, of fixed time duration Δt . In each time step, each component of macroscopic strain is updated by an amount

$$\Delta \epsilon_i = \text{sgn}(r_1 - \theta) [\Delta \epsilon_{\min i} + r_2 (\Delta \epsilon_{\max i} - \Delta \epsilon_{\min i})] \quad (20)$$

where $\Delta \epsilon_{\min i}$, $\Delta \epsilon_{\max i}$ are the minimum and maximum allowable amplitudes of the increments in each strain component over the loading steps, r_1 and r_2 are uniform random variables between 0 and 1, and θ defines the probability that the amplitude of each strain step is negative (equal

Table.1

Mechanical properties for the phenomenological model.

(Gpa)	E_{22} (Gpa)	E_{33} (Gpa)	ν_{12}	ν_{13}	ν_{23}	G_{12} (GPa)	G_{13} (GPa)	G_{23} (GPa)
198.2	11.2	11.2	0.29	0.29	0.5	8.6	8.6	3.7
X_T (Mpa)	X_C (Mpa)	Y_T (Mpa)		Y_C (Mpa)		S_f (Mpa)	S_f (Mpa)	
4222	3800	122		218		81	81	
G_{FT} (J/mm ²)	G_{FC} (J/mm ²)					G_{MT} (J/mm ²)	G_{MC} (J/mm ²)	
12.5	12.5					1	1	

probability of positive and negative steps is assumed in this study, $\theta = 0.5$), and i is an index indicating the six components of the vector ϵ .

The strain histories obtained as above are then smoothed by a smoothing function. The time/strain pairs obtained via Eq. (20) are used as seeding points, and intermediate points between two consecutive seeding points (t_j, ϵ_j) and $(t_{j+1}, \epsilon_{j+1})$, are interpolated as $(t, A_j + (A_{j+1} - A_j)\zeta^3(10 - 15\zeta + 6\zeta^2))$, with $\zeta = (t - t_j)/(t_{j+1} - t_j)$.

Here we set $\epsilon_{\max} = 0.05$, $\Delta\epsilon_{\max} = 0.01$, $\Delta\epsilon_{\max} = 0.005$ based on the results of a preliminary study. In this study, we employ a rate-insensitive constitutive models, and hence the numerical value of time t does not influence the stress-strain histories produced. We set the value of Δt to be 10. We note in passing that the smoothing function above can improve the distribution (uniformity) of strain rate data, which is important when rate-dependant constitutive models are used.

A video illustrating the random history of deformation of boundary conditions is provided at the link indicated in the Data availability section.

2.3. Current material stiffness

Quantitative data on the evolution (degradation) of the stiffness of the material with continued random deformation is needed to inform our data-driven model. For brittle materials, absent any plastic deformation, knowledge of the stiffness matrix of the material at any stage of the deformation is equivalent to knowledge of the constitutive response. We therefore extract data on the current homogenised stiffness of the material of interest; this was successfully demonstrated in [22] for isotropic RVEs of ductile materials with isotropic damage, and is applied here to a brittle solid with complex anisotropic damage. The procedure is generally applicable to any form of anisotropic damage initiation and evolution.

For a material model comprising m integration points, at a general time of interest t , we can record the deformed coordinates of the nodes, the initial stiffness matrices and the local damage variables of all the m integration points. Then, a secondary, linear elastic FE model is built, with nodes and elements corresponding to the deformed configuration of the primary simulation at time t , and damaged elastic stiffness distributed across the integration points according to the microscopic damage law used. Then, similar to [23], the secondary FE model is subjected to a set of 6 simulations, each prescribing a small variation in one macroscopic strain component while keeping the other 5 components equal to zero; this allows extracting the columns of the homogenised stiffness matrix of the material volume element. For the purpose of this demonstration, the RVE consists of a single FE obeying the constitutive model illustrated in Section 2.1.

The primary FE simulations are non-linear in geometry and material response, therefore the FE solutions are incremental. A datapoint of our training database is represented by a general increment in the primary simulations, and contains data on the homogenised stiffness matrix at the beginning and end of the increment. Multiple random strain histories are applied the model, and the corresponding histories of stress, strain and current stiffness matrix are recorded. To develop a data-driven model of suitable accuracy, we simulated the response of the model to a set of 550 different random strain histories, 500 of which were used for training, while the remaining 50 were used for testing.

3. Data-driven surrogate model

In this section we present the data-driven model in detail. As the material displays no plasticity, its response is simply given by

$$\sigma = C\epsilon \quad (21)$$

where C is the current overall stiffness, ϵ , σ are the macroscopic strain and stress vectors, corresponding to tensors but written in Voigt notation. In this study, due to the particular model used, the material's

behaviour is at most orthotropic and the current macroscopic elastic stiffness tensor C comprises 9 independent components, namely

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} & & & \\ & C_{22} & C_{23} & & & \\ & & C_{33} & & & \\ & & & C_{44} & & \\ & & & & C_{55} & \\ & & & & & C_{66} \end{bmatrix} \quad (22)$$

We manipulate this tensor in the vector form

$$\mathbf{c} = [C_{11}, C_{12}, C_{13}, C_{22}, C_{23}, C_{33}, C_{44}, C_{55}, C_{66}]^T. \quad (23)$$

The high dimensionality of the stiffness vector $\mathbf{c}$ can pose a practical challenge for the machine learning exercise to conduct. To overcome this challenge we will employ data compression to the stiffness vector (23), using a principal component analysis (PCA), as detailed below.

3.1. Dimensionality reduction

PCA allows bijective mapping of data of high-dimensionality D into a lower-dimensional space of dimension d , as

$$f: \mathbb{R}^D \rightarrow \mathbb{R}^d \quad (24)$$

Following [24], we apply a singular value decomposition to find a new set of orthogonal axes that are uncorrelated and ranked according to the variance of the reduced dataset. The coordinates of the low-dimensionality data, denoted as principal components $\mathbf{v} = (pc_1, pc_2, \dots, pc_d)^T = f(\mathbf{c})$, are ordered in such a way that the first principal component is associated with the greatest variance, the second principal component with the second greatest variance, and so on. The D -dimensional data can be reconstructed from the d -dimensional vector $\mathbf{v} = (pc_1, pc_2, \dots, pc_d)^T$ by applying the inverse PCA mapping as

$$\hat{\mathbf{c}} = f^{-1}[f(\mathbf{c})] \quad (25)$$

To assess the effectiveness of the data compression, we conducted an investigation varying d from 1 to 9. An additional $\varphi = 50000$ datapoints, unseen while performing the dimensionality reduction, were used to evaluate the accuracy of the low-dimensional representation, as quantified by the root mean squared error E_{rmse} , defined as

$$E_{rmse} = \sqrt{\sum_{j=1}^{\varphi} \|\hat{\mathbf{c}}^j - \mathbf{c}^j\|^2} / \varphi \quad (26)$$

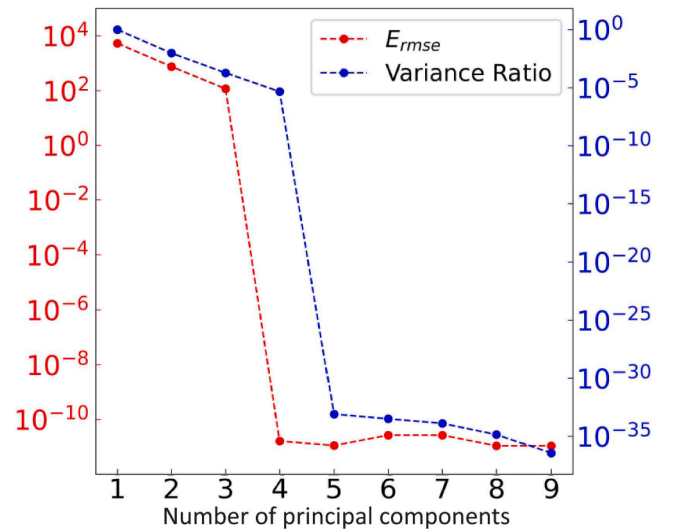


Fig. 1. Root mean squared error E_{rmse} and explained variance ratio as a function of the chosen value of the reduced dimensionality d .

The dependence of results of E_{rmse} upon d is presented in Fig. 1, which includes the explained variance ratio [25] for each value of d .

Clearly, the use of $d \geq 4$ results in a significant decrease in E_{rmse} and variance ratio. Further increasing d beyond 4 gives negligible additional advantage. This is unsurprising here, since the evolution of the stiffness matrix depends only on 4 independent damage variables, as shown in Eqs. (2)–(20). This exercise indicates that the PCA can effectively extract key information from the dataset, to facilitate the machine learning application (in this case reducing the dimensionality of the $\mathbf{c}$ vector from 9 to 4).

3.2. Data-driven surrogate model

The proposed model is composed of two different sub-models: a classification model and a regression model. These models predict increments in the current material stiffness, from which the macroscopic stress can be calculated for any given increment in strain. In all sub-models supervised machine learning is applied, by employing simple multilayer feedforward neural networks, with a single output y , two hidden layers having 100 neurons each, both using the activation function $\text{ReLU}(\alpha) = \max(0, \alpha)$; these networks $NN(\mathbf{x})$ therefore have the structure

$$y = NN(\mathbf{x}) = \mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 (\text{ReLU}(\mathbf{W}_0 \mathbf{x} + \mathbf{b}_0)) + \mathbf{b}_1) + \mathbf{b}_2 \quad (27)$$

where $\mathbf{W}_0, \mathbf{W}_1, \mathbf{W}_2$ are weight matrices and $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$ are column vectors of biases, $\mathbf{x}$ is the input vector. The choice of these hyperparameters for all NNs was based on results of previous studies. In this paper we do not focus on optimising these hyperparameters. Similar surrogate models to that proposed below can be obtained by applying different machine learning techniques, and here NNs are chosen for their simplicity.

A 90 % fraction of the training datapoints were randomly selected to train the NNs and the remaining 10 % served as validation data. The training data coordinates were normalized by subtracting their average and dividing by their variance. The training was carried out by back-propagation algorithm, using stochastic gradient descent with adaptive moment estimation. The batch size was set to 50, the learning rate was 0.001, and early stopping was employed to halt the training if the validation loss did not decrease for 100 epochs.

3.2.1. Classification

This model was initially applied to all deformation increments to determine whether or not a change in material stiffness will occur during the increments. A subset of ϕ training datapoints was randomly selected amongst all the increments, such to have $\phi/2$ increments in which the stiffness matrix did not change and $\phi/2$ increments in which the matrix did change. The subset was categorized into these two groups by a label y

$$y = \begin{cases} 0 & \text{if } \Delta \mathbf{C} = \mathbf{0} \quad (\text{elastic}) \\ 1 & \text{if } \Delta \mathbf{C} \neq \mathbf{0} \quad (\text{inelastic}) \end{cases} \quad (28)$$

In this classification we assume that current strain, PC vector (quantifying the stiffness at the beginning of the increment) and strain increment are sufficient to determine the value of the label, i.e.

$$y = NN_{clas}(\boldsymbol{\varepsilon}, \mathbf{v}, \Delta \boldsymbol{\varepsilon}) \quad (29)$$

The loss function was chosen as the binary cross entropy

$$H(\mathbf{x}) = \frac{-1}{\phi} \sum_{j=1}^{\phi} \left(y_{label}^j \ln(\hat{y}^j) + (1 - y_{label}^j) \ln(1 - \hat{y}^j) \right) \quad (30)$$

where $\hat{y}^j$ are the predicted values of the label by the neural network and y^j are the corresponding true values.

3.2.2. Regression

This model predicts the increment in the low-dimensional stiffness vector $\Delta \mathbf{v}$. We assume that this increment depends on the initial strain and stiffness and on the strain increment, i.e.

$$\Delta \mathbf{v} = NN_{reg}^v(\boldsymbol{\varepsilon}, \mathbf{v}, \Delta \boldsymbol{\varepsilon}) \quad (31)$$

The loss function used to train this NN is the mean absolute error (MAE)

$$MAE = \frac{1}{\phi} \sum_{j=1}^{\phi} \sum_{k=1}^b |y_k^j - \hat{y}_k^j| \quad (32)$$

where $\hat{y}_k^j$ are predictions of the NN and y_k^j the true values of the outputs. b is the length of the output vector and ϕ is the number of training datapoints, in this case equal to all increments in the training dataset classified as inelastic, i.e. $NN_{clas} = 1$.

3.2.3. Computational framework

The data-driven constitutive model provides the increment in macroscopic stress as a function of an arbitrary imposed strain increment. The computational framework is summarised in Fig. 2. Consider an arbitrary strain history imposed to the model, starting from vanishing stress and strain and with an initial “undamaged” stiffness vector $\mathbf{c}_0$, compressed to d dimensions and denoted as $\mathbf{v}_0$

$$\mathbf{v}_0 = f(\mathbf{c}_0) \quad (33)$$

For each imposed strain increment, first the classification model is applied to determine whether the increment is elastic or inelastic. In both cases the strain is updated as $\boldsymbol{\varepsilon}_{n+1} = \boldsymbol{\varepsilon}_n + \Delta \boldsymbol{\varepsilon}_n$. If an increment is classified as elastic the stiffness matrix at the beginning of the increment is used to update the stress. If the increment is classified as inelastic, the increment in the low-dimensional stiffness vector is determined by the regression model as

$$\Delta \mathbf{v}_n = NN_{reg}^v(\boldsymbol{\varepsilon}_n, \mathbf{v}_n, \Delta \boldsymbol{\varepsilon}_n) \quad (34)$$

where the mean absolute error (MAE) is used as a loss function. At the

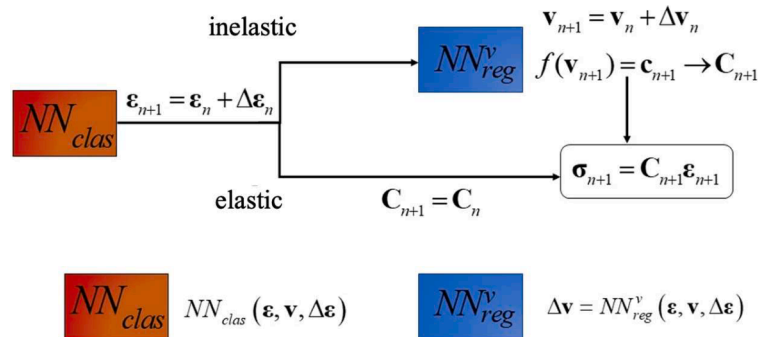


Fig. 2. Illustration of the proposed computational framework.

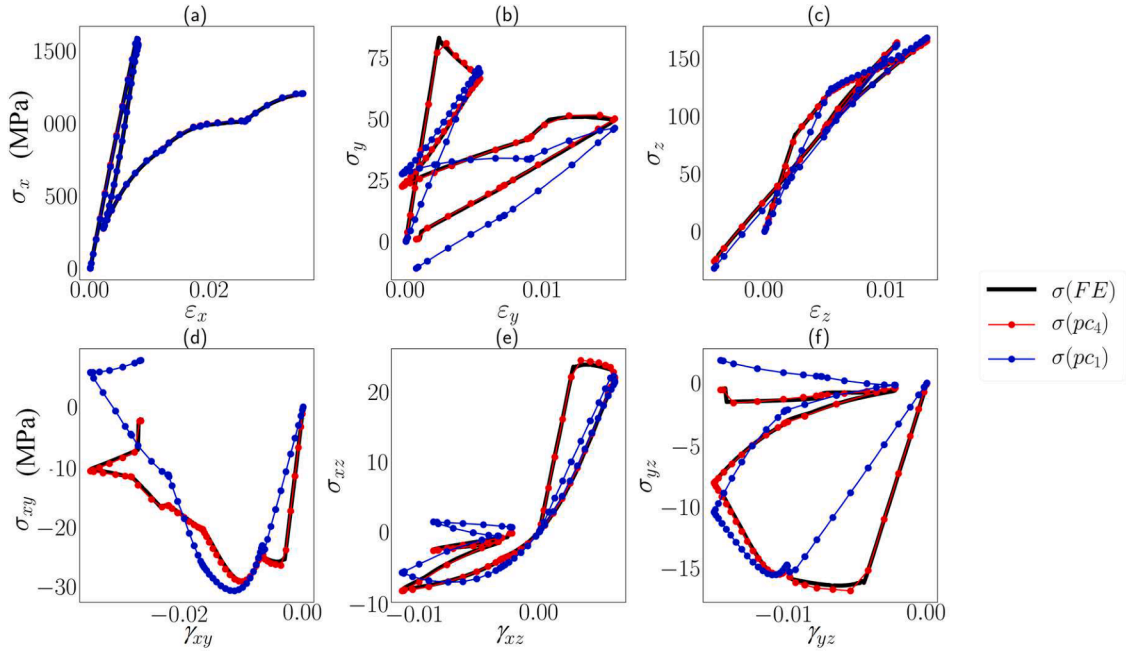


Fig. 3. Comparison of strain-stress curves generated by FE simulations and 2 surrogate models with $d = 1$ and $d = 4$.

end of the increment, the low-dimensional stiffness vector is updated as

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \Delta \mathbf{v}_n \quad (35)$$

The inverse PCA mapping is then applied to obtain the stiffness vector at the end of the increment

$$\mathbf{c}_{n+1} = \mathbf{f}^{-1}(\mathbf{v}_{n+1}), \quad (36)$$

following which the final stiffness matrix $\mathbf{C}_{n+1}$ is calculated and the stress $\boldsymbol{\sigma}_{n+1}$ is updated as

$$\boldsymbol{\sigma}_{n+1} = \mathbf{C}_{n+1} \boldsymbol{\epsilon}_{n+1}. \quad (37)$$

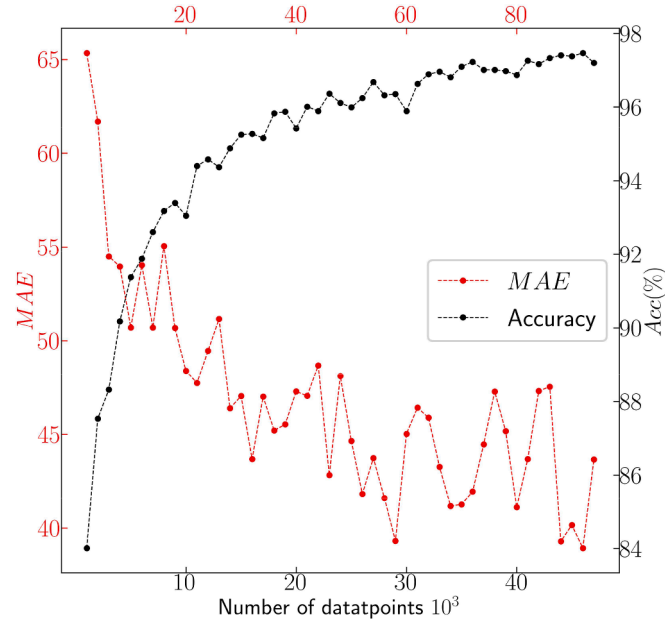


Fig. 4. Sensitivity of the loss function of the regression model (MAE) and of the accuracy of the classification model to the size of the training dataset. Note that the number of datapoints indicated refers to the regression model, while the classification model uses double that number.

4. Results and discussion

To illustrate the accuracy of the surrogate model, as well as the effect of the dimensionality reduction technique, we test the model by predicting the stress histories corresponding to 6-dimensional random strain histories not included in the training or validation datasets. An example of the predictions is shown in Fig. 3 which includes the ground truth (as determined by FE simulations) as well as two predictions computed using 1 PC and 4 PCs to compress the information associated to the material stiffness. As expected compressing the data using one single PC ($d = 1$) gives predictions which deviate considerably from the ground truth with the exception of the response in direction x , which is predicted accurately. This is due to the anisotropic nature of the damage, which cannot be captured by one single variable. On the contrary, when using 4 PCs ($d = 4$), the stress versus strain histories are predicted very accurately, such that the predictions are nearly indistinguishable from the ground truth. We recall that the damage evolution depends in the FE simulations on only 4 parameters; it is therefore unsurprising that an increase of the number of PCs beyond 4 does not lead to tangible improvement.

Based on the results of Fig. 3 we proceed with the choice $d = 4$. To further quantify the fidelity of the surrogate model, in Fig. 4 we present the dependence of the accuracy of the classification (intended as the ratio of the total correctly classified increments over the total number of increments examined) and of the regression (quantified by its loss function, MAE) on the size of the training dataset employed. The figure indicates that MAE decreases and the classification accuracy increases monotonically with the size of the training dataset.

Given the monotonic trends observed in Fig. 4, we proceed with the maximum number of training datapoints of 50,000 to train the regression model and 100,000 for the classification model. We note that in this paper data generation was computationally inexpensive, due to the use of a phenomenological model. When a large RVE is used (10^6 nodes or more), the generation of training data can take some computational efforts; however a number of simulation increments of 10^5 , found adequate to train a surrogate model in this study, is within the reach of ordinary computer hardware, with calculation times of order of hours or days. Such calculations necessary for training are performed only one time.

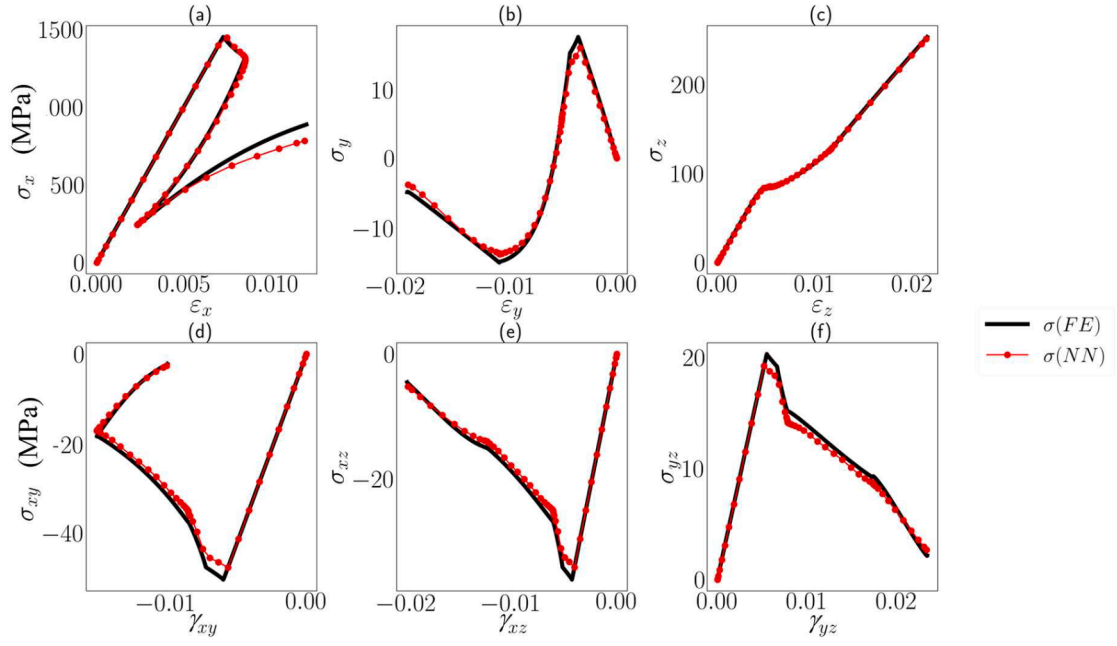


Fig. 5. Accuracy of the surrogate model in predicting unseen 6D stress-strain curves.

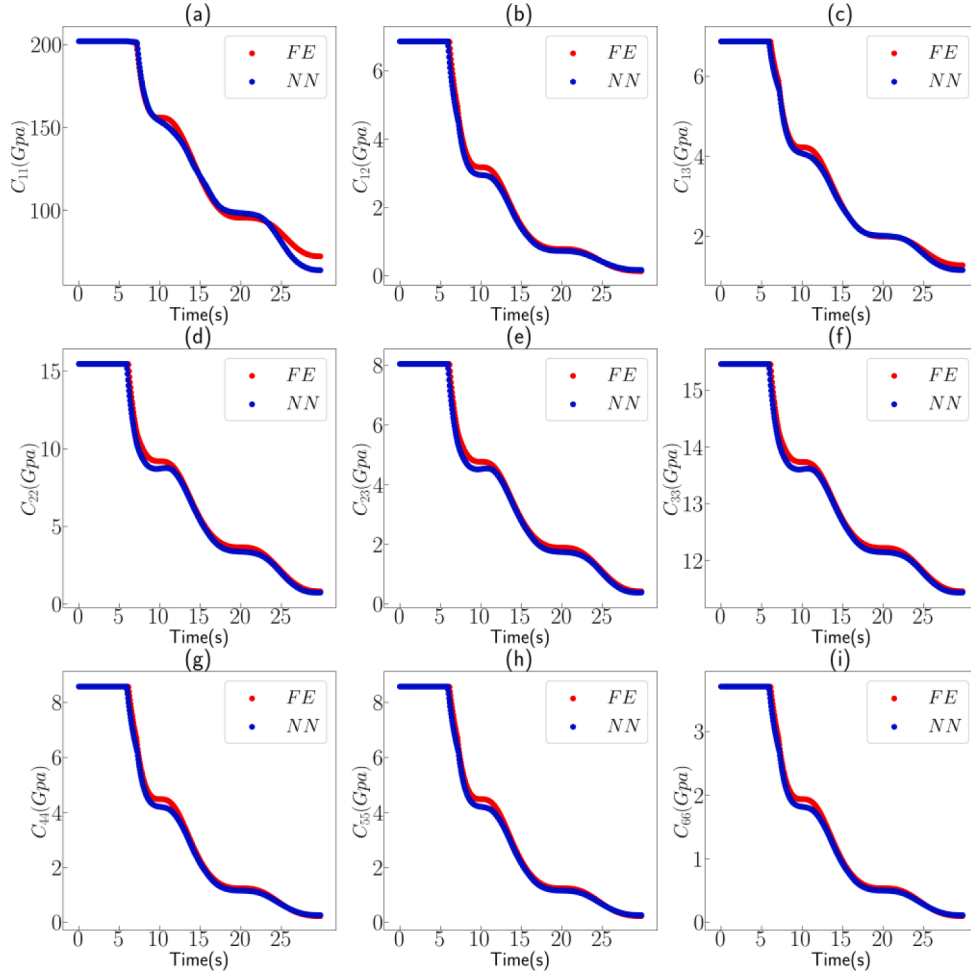


Fig. 6. Accuracy of the surrogate model in predicting the evolution of the components of the stiffness matrix with time.

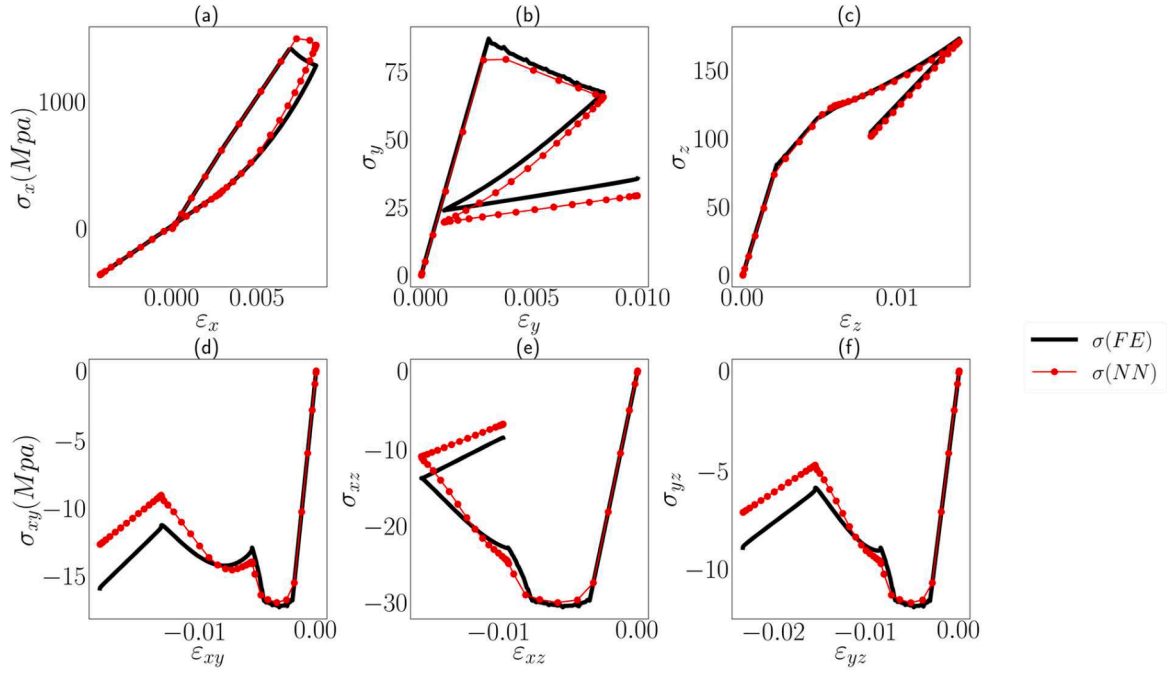


Fig. 7. Accuracy of the surrogate model in predicting unseen 6D stress-strain curves, for a loading case of higher complexity.

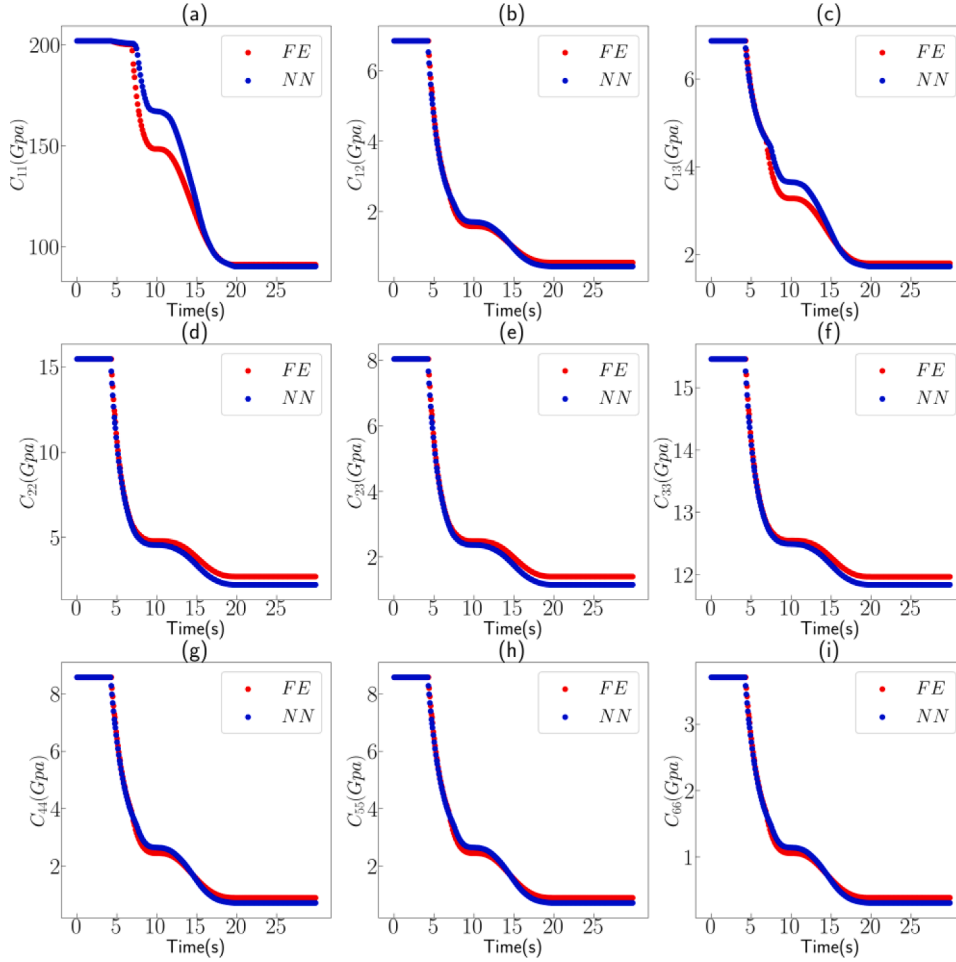


Fig. 8. Accuracy of the surrogate model in predicting the evolution of the components of the stiffness matrix with time, for a loading case of higher complexity.

The MAE and accuracy examined in Fig. 4 quantify the accuracy of the surrogate in predicting the evolution of stiffness over one single simulation increment. If the surrogate is employed in predicting the material's response to an imposed history of strains, the errors made in successive increments accumulate. To illustrate visually the predictive capabilities of the surrogate model, we present predictions of the stress histories corresponding to unseen imposed strain histories. We present two randomly selected loading cases in Figs. 5–8, corresponding to two random strain histories not seen in the training process. We show selected stress-strain curves as well the evolution of the 9 terms of the stiffness matrix with time. The accuracy of the surrogate model's predictions is observed to persist along most of these 2 very complex, non-monotonic, and non-proportional strain histories. Minor errors are observed in the second selected loading case, of relatively higher complexity, in the final parts of the random strain paths; this is likely due to accumulation of the error after a large number of increments, but also to the relative scarcity of training datapoints for the NNs in those final parts of the random strain paths. These minor deviations can be corrected by optimising the structure of the hyperparameters of the NNs employed, the probability distribution of the training data over the input space, and the weighing of the loss function over the input space. This is not pursued here for the sake of brevity, but also to demonstrate that very simple and suboptimal machine learning algorithms can achieve high accuracy and efficiency in learning the macroscopic response of arbitrarily anisotropic elastic solids.

We also tested an alternative model that performs the simple regression $\Delta\sigma = NN(\epsilon, \sigma, \Delta\epsilon)$ and it therefore does not include the information of the current material's stiffness; this model, with a NN of the same size, and trained over the same dataset used in this study, gave an error between one and two orders of magnitude higher than that in the examples of Figs. 5–8. It follows that it is essential to include information on the current material stiffness to achieve efficient and accurate surrogate predictions. The surrogate model can be easily incorporated in any FE or other type of numerical solver, as evaluated in e.g. [16].

No efforts are made in this study to enforce that the model satisfies the first and second law of thermodynamics; the training data does satisfy these principles and the trained model learns quickly not to violate them; this can be enforced by a constraint during the training, possibly leading to saving in the computational time for training; this will be pursued in future studies.

5. Limitations of the study

The computational procedure presented in this study can be implemented as a constitutive model for any FE solver, for example in the form of a UMAT subroutine for Abaqus standard; this was not pursued here as it was demonstrated in previous studies [16], where we also showed that a similar data-driven model can offer additional computational advantages over classical FE models, as it performs a direct prediction of the stress increment without the need for iterative procedures.

As the phenomenological model that we try to reproduce is rate-independent, we do not include time variables in the proposed regressions; we have however previously shown in [19] that rate-dependant data-driven models can be obtained by the simple addition of the physical time increment as an input of the regression and by appropriate design of the input data space, in particular along the strain rate axis.

6. Conclusions

Given a numerically generated volume element of a material of interest, and physical simulations of the response of such volume element to arbitrary histories of applied strain, we propose a computational framework to effectively learn the homogenised material response in all its details, to encode it in two simple neural networks, and to be able to reproduce it at negligible computational cost. The computational

framework is demonstrated for the case of a brittle material with complex anisotropic damage evolution, as shown by most composite materials and brittle anisotropic solids. We show that surrogate predictions of high accuracy can be achieved by including as inputs and outputs of the model the current stiffness of the material. To facilitate the application of machine learning, the stiffness-related information can be effectively compressed using dimensionality reduction techniques as the principal component analysis. The proposed framework can aid multiscale simulations and facilitate the encoding, application and sharing of material models.

CRedit authorship contribution statement

Weijian Ge: Writing – original draft, Methodology, Investigation, Formal analysis. **Vito L Tagarielli:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The full training dataset used to produce the surrogate model is available at this permanent link: <https://zenodo.org/records/12191930>.

References

- [1] J. Vilppo, R. Kouhia, J. Hartikainen, K. Kolari, A. Fedoroff, K. Caloni, Anisotropic damage model for concrete and other quasi-brittle materials, *Int. J. Solids Struct.* 225 (2021), <https://doi.org/10.1016/j.ijsolstr.2021.111048>.
- [2] M. Fassin, R. Eggersmann, S. Wulfinghoff, S. Reese, An anisotropic brittle damage model with a damage tensor of second order using a micromorphic approach, *PAMM* 19 (1) (2019), <https://doi.org/10.1002/pamm.201900138>.
- [3] H. Holthusen, T. Brepols, S. Reese, J.W. Simon, A two-surface gradient-extended anisotropic damage model using a second order damage tensor coupled to additive plasticity in the logarithmic strain space, *J. Mech. Phys. Solids* 163 (2022), <https://doi.org/10.1016/j.jmps.2022.104833>. Mar.
- [4] G.Z. Voyiadis, Y. Zhou, P.I. Kattan, A new anisotropic elasto-plastic-damage model for quasi-brittle materials using strain energy equivalence, *Mech. Mater.* 165 (2022), <https://doi.org/10.1016/j.mechmat.2021.104163>.
- [5] C. Oliver-Leblond, R. Desmorat, B. Koley, Continuous anisotropic damage as a twin modelling of discrete bi-dimensional fracture, *Eur. J. Mech. A* 89 (2021), <https://doi.org/10.1016/j.euromechsol.2021.104285>.
- [6] L. Olsen-Kettle, Bridging the macro to mesoscale: evaluating the fourth-order anisotropic damage tensor parameters from ultrasonic measurements of an isotropic solid under triaxial stress loading, *Int. J. Damage Mech.* 28 (2) (2019), <https://doi.org/10.1177/1056789518757293>.
- [7] J.W. Ju, On isotropic and anisotropic damage variables in continuum damage mechanics, in *Mech. Comput. Beyond* (1991), [https://doi.org/10.1061/\(asce\)0733-9399\(1990\)116:12\(2764](https://doi.org/10.1061/(asce)0733-9399(1990)116:12(2764).
- [8] S. Ghosh, J. Bai, D. Paquet, Homogenization-based continuum plasticity-damage model for ductile failure of materials containing heterogeneities, *J. Mech. Phys. Solids* 57 (7) (2009) 1017–1044, <https://doi.org/10.1016/J.JMPS.2009.04.002>. Mar.
- [9] K. Dorhmi, L. Morin, K. Derrien, Z. Hadjem-Hamouche, J.P. Chevalier, A homogenization-based damage model for stiffness loss in ductile metal-matrix composites, *J. Mech. Phys. Solids* 137 (2020), <https://doi.org/10.1016/J.JMPS.2019.103812>. Mar.
- [10] T. Lesičar, J. Sorić, Z. Tonković, Ductile damage modelling of heterogeneous materials using a two-scale computational approach, *Comput. Methods Appl. Mech. Eng.* 355 (2019) 113–134, <https://doi.org/10.1016/J.CMA.2019.06.013>. Mar.
- [11] K. Matouš, M.G.D. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, *J. Comput. Phys.* 330 (2017) 192–220, <https://doi.org/10.1016/j.jcp.2016.10.070>.
- [12] K. Zhang, et al., Strength prediction and progressive damage analysis of carbon fiber reinforced polymer-laminate with circular holes by an efficient Artificial Neural Network, *Compos. Struct.* 296 (2022) 115835, <https://doi.org/10.1016/J.COMPOSTRUCT.2022.115835>. Sep.

- [13] J. Reiner, R. Vaziri, N. Zobeiry, Machine learning assisted characterisation and simulation of compressive damage in composite laminates, *Compos. Struct.* 273 (2021) 114290, <https://doi.org/10.1016/J.COMPSTRUCT.2021.114290>. Oct.
- [14] S. Yan, X. Zou, M. Ilkhani, A. Jones, An efficient multiscale surrogate modelling framework for composite materials considering progressive damage based on artificial neural networks, *Compos. B Eng.* 194 (2020), <https://doi.org/10.1016/j.compositesb.2020.108014>. Aug.
- [15] Y. Song, A. Schiffer, V.L. Tagarielli, The effects of heterogeneous mechanical properties on the response of a ductile material, *Sci. Rep.* 11 (1) (2021) 1–17, <https://doi.org/10.1038/s41598-021-97495-x>. 2021 11:1Sep.
- [16] B. Tasdemir, A. Pellegrino, V. Tagarielli, A strategy to formulate data-driven constitutive models from random multiaxial experiments, *Sci. Rep.* 12 (1) (2022), <https://doi.org/10.1038/S41598-022-26051-Y>. Dec.
- [17] B. Tasdemir, V. Tagarielli, A. Pellegrino, A data-driven model of the yield and strain hardening response of commercially pure titanium in uniaxial stress, *Mater. Des.* 229 (2023) 111878, <https://doi.org/10.1016/J.MATDES.2023.111878>. May.
- [18] C. Zhang, N. Li, W. Wang, W.K. Binienda, H. Fang, Progressive damage simulation of triaxially braided composite using a 3D meso-scale finite element model, *Compos. Struct.* 125 (2015) 104–116, <https://doi.org/10.1016/J.COMPSTRUCT.2015.01.034>. Jul.
- [19] W. Ge and V.L. Tagarielli, “A computational framework to establish data-driven constitutive models for time-or path-dependent heterogeneous solids”, [doi: 10.1038/s41598-021-94957-0](https://doi.org/10.1038/s41598-021-94957-0).
- [20] Z. Hashin, Failure criteria for unidirectional fiber composites, *J. Appl. Mech.* 47 (2) (1980) 329–334, <https://doi.org/10.1115/1.3153664>. Jun.
- [21] A. Duval et al., “Abaqus/CAE 6.14 user’s manual,” Dassault Systèmes Inc. Providence, RI, USA, vol. IV, no. June 2014.
- [22] W. Ge, V.L. Tagarielli, Data-driven homogenisation of the response of heterogeneous ductile solids with isotropic damage, *Mater. Des.* 238 (2024) 112738, <https://doi.org/10.1016/J.MATDES.2024.112738>. Feb.
- [23] E. Andreassen, C.S. Andreasen, How to determine composite material properties using numerical homogenization, *Comput. Mater. Sci.* 83 (2014) 488–495, <https://doi.org/10.1016/J.COMMATSCI.2013.09.006>. Feb.
- [24] M.V. Pathan, S.A. Ponnusami, J. Pathan, R. Pitongsawat, B. Erice, V.L. Tagarielli, predictions of the mechanical properties of unidirectional fibre composites by supervised machine learning, *Sci. Rep.* 9 (2019) 13964, <https://doi.org/10.1038/s41598-019-50144-w>.
- [25] T. Kurita, “P principal component analysis (PCA),” 2020, [doi: 10.1007/978-3-030-03243-2_649-1](https://doi.org/10.1007/978-3-030-03243-2_649-1).