A multiscale deep learning model for elastic properties of woven composites

Citation for the original published paper (version of record):
Ghane, E., Fagerström, M., Mirkhalaf, S. (2023). A multiscale deep learning model for elastic properties of woven composites. International Journal of Solids and Structures, 282.
http://dx.doi.org/10.1016/j.ijsolstr.2023.112452

(article starts on next page)

International Journal of Solids and Structures 282 (2023) 112452

Contents lists available at ScienceDirect

International Journal of Solids and Structures

journal homepage: www.elsevier.com/locate/ijsolstr

A multiscale deep learning model for elastic properties of woven composites

E. Ghane [a], M. Fagerström [b], S.M. Mirkhalaf [a,*]

[a] Department of Physics, University of Gothenburg, Gothenburg, Sweden
[b] Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Dataset link: MultiscaleElasticWovenNeuralNet (Original data)

Keywords:
Woven composites
Multiscale analysis
Artificial neural networks
Elastic properties

ABSTRACT

Time-consuming and costly computational analysis expresses the need for new methods for generalizing multiscale analysis of composite materials. Combining neural networks and multiscale modeling is favorable for bypassing expensive lower-scale material modeling, and accelerating coupled multi-scale analyses (FE^2). In this work, neural networks are used to replace the time-consuming micromechanical finite element analysis of unidirectional composites, representing the local material properties of yarns in woven fabric composites in a multiscale framework. Leveraging the fast multiscale data generation procedure, we presented a second neural networks model to estimate the elastic engineering coefficients of a particular weave architecture based on a broad range of dry resin and fiber properties and yarn fiber volume fraction. As an outcome, this paper provides the user with a generalized, neural network-based approach to tackle the balance of computational efficiency and accuracy in the multiscale analysis of elastic woven composites.

1. Introduction

Efficient, simulation-driven design of woven fabric composites requires a model capable of considering different fiber types and matrix systems such that one can tailor materials and find the optimized trade-off between the properties and cost. An absolute requirement for such a model is that it should efficiently and accurately predict the elastic behavior from constituent properties and microscale information. Multiscale computational methods, often based on computational homogenization, have therefore been established to address the complexity of modeling woven composites (Ivanov and Lomov, 2020). Nevertheless, a generic model with low computational cost is needed to observe the possible combination of microscopic features based on different loading cases and mesoscale weave architectures.

The elastic behavior of woven composites has been successfully assessed using both analytical procedures, cf. Shokrieh et al. (2017) and Adumitroaie and Barbero (2011) and numerical methods, such as mesh-free methods (Wen and Aliabadi, 2009) and finite element analysis (Melro et al., 2012; Varandas et al., 2020). In all cases, it is assumed that a material unit cell, representing the repetitive pattern of the reinforcement weave, can represent the overall behavior of the composite. For this unit cell, the yarns properties are usually computed based on a micromechanical analysis of an equivalent unidirectional (UD) composite at each local point. In this regard, governing micromechanical features, including fiber volume fraction, fiber dimension and orientation distribution, constitutive properties of the matrix and fiber are considered. On the one hand, analytical micromechanical models have been developed for fast estimation through applying some simplifying assumptions and accepting a rather high level of error with respect to the actual physics (Heidari-Rarani et al., 2018; Aboudi et al., 2013). On the other hand, high-fidelity models based on the finite-element method (Melro et al., 2012; Varandas et al., 2020; Bai et al., 2015) provide higher accuracy, although at the expense of a higher computational cost.

Data-driven computational analysis is beneficial for exploring a multiscale, high-dimensional design space taking advantage of data to find the best design. Among different data-driven analyses, machine learning-enhanced frameworks utilize large data sets to find optimal designs or constitutive relations based on microscale quantities (Thomas et al., 2022). Among others, the model-free approach solves the boundary value problem directly using data, instead of relying on constitutive relations for the lower scale constituents in a multiscale framework (Karapiperis et al., 2021; Wu et al., 2021; Peng et al., 2020).

Artificial neural networks (ANN), as a category of machine learning algorithms, have recently shown the promising capability to integrate multiscale mechanical features and predict mechanical behavior (Liu et al., 2021) by minimizing a loss (cost) function which essentially is the difference between experimental observations and network predictions. Recently, physics-informed neural networks have been introduced to avoid the nonphysical behavior of the networks, e.g. models that violate thermodynamic principles (Bischof and Kraus, 2021). In this

* Corresponding author.
E-mail address: mohsen.mirkhalaf@physics.gu.se (S.M. Mirkhalaf).

https://doi.org/10.1016/j.ijsolstr.2023.112452
Received 27 October 2022; Received in revised form 25 July 2023; Accepted 10 August 2023
Available online 21 August 2023

approach an ANN is trained with an extended cost function that also includes physical balance laws (Linka et al., 2021; Vlassis and Sun, 2021; Haghighat et al., 2021). However, Problems arise from multi-phase materials like composites when the differentiation of variables is direction-dependent, making physics-informed neural networks demanding for anisotropic materials. Instead, surrogate modeling with ANNs is demonstrated for multiscale, inhomogeneous, or anisotropic applications (Mozaffar et al., 2018). The motivation is to use good generalization properties of neural networks for unseen data. In addition to the experimental data available, different developed techniques in computational mechanics can generate the necessary data set for training a neural network. For example, mean-field models can be applied to short fiber composites (Friemann et al., 2023), and full-field finite elements or Fast Fourier Transform models for crystal plasticity (Bonatti et al., 2022).

Surrogate modeling is employed to accelerate multiscale modeling and replace the finite-element-based lower-scale model with a data-based surrogate model for a particular material system (Yan et al., 2020; Furtado et al., 2021). Following the study by Mentges et al. (2021) on short fiber reinforced composites, we herein present a machine-learning framework to estimate elastic properties of woven composites with the aim to replace the time-consuming multiscale FE simulations. As a starting point, a series of statistical analyses have been conducted to obtain the required representative volume element size. Fiber volume fraction, fiber and matrix stiffness and Poisson's ratios are considered input parameters. Further, to develop the training set for a micromechanical ANN, an automatic process is arranged to generate 3D unidirectional representative volume elements (RVE)s, then analyze these FE simulations to obtain homogenized stiffness quantities. Accordingly, an ANN model is trained and validated for the microscopic FE simulations. The micro-ANN model, which is a feed-forward multi-layer perceptron neural network, is then used at the mesoscale level to compute the elastic properties of a plain-weave fabric composite in a high-fidelity mesoscale model. Based on mesoscale predictions, a second feed-forward ANN is trained for a given weave architecture to estimate the resulting elastic behavior of woven composite based on the different combinations of fiber and matrix materials. The two ANN models are applicable for a vast combination of fiber volume fractions and constitutive properties. A schematic representation of the data flow and the two micro- and meso-ANNs are presented in Fig. 1.

The remaining of the paper is structured as described in the following. Design of experiments is explained in Section 2. Computational analysis of each design point through the automatic process of data generation and analysis is detailed in Section 3 for the micro-mechanical and the mesoscale model. The artificial neural network model is explained thoroughly in Section 4. A comprehensive study is done in Section 5 on the interpolated results and algorithms operation. Finally, the concluding remarks are provided in Section 6.

2. Design of computational experiments

It is necessary to construct a large and reliable design set for training an ANN model. One approach is to generate such a data set through physical experiments. However, in many cases, it is infeasible or even impossible to generate or have access to enough experimental data. An example of such a case is the prediction of homogenized stiffness properties of unidirectional composites for a very wide range of combinations of fibers and matrix materials at various volume fractions. An alternative is then to generate data by conducting a large number of micromechanical simulations based on a relevant range of input features. In this study, homogenized elastic properties are obtained via computational homogenization (3.1) of microscopic RVEs, subjected to unit strains through periodic boundary conditions.

Table 1
Microstructural parameters space for micromechanical simulations (Barbero, 2011; Herakovich, 1998).

Parameter	Minimum	Maximum
Matrix stiffness E_M (MPa)	3500	10,000
Matrix Poisson's ratio ν_M (−)	0.25	0.49
Fiber stiffness E_F (MPa)	69,000	786,000
Fiber Poisson's ratio ν_F (−)	0.20	0.40
Fiber volume fraction V_f (%)	30	60

2.1. Input features

An important step is to do feature engineering to find out which input features should be considered in the data set generation process. Three different descriptors are considered: (I) micro-and/or meso-structure geometry, which relates to the description of the representative volume element, (II) material properties of both matrix and fibers, and (III) the state of applied deformation. Then, the following properties are defined in order to characterize a unidirectional composite micro-structure: elastic properties of matrix material (E_m, ν_m); elastic properties of fiber (E_f, ν_f); fiber diameter (d_f) and fiber volume fraction (V_f). These features are also considered as the input variables in the mesoscale model representing the unit cell of the woven composite. For this level, we only consider a constant mesoscale geometry. Although more realistic material properties can easily be employed in the framework, in this study, we have used isotropic elasticity for both matrix and fibers (Sections 3.2 and 3.3). Considering that the present study aims to determine only the elastic properties of composites, no debonding between matrix and fibers is considered.

The RVE size plays an important role in the statistical representation of heterogeneous material behavior. Different RVE sizes are suggested for UD composites in the elastic domain (see, e.g. Melro et al., 2012; Bai et al., 2015). In the present study, a rectangular cuboid RVE with equal width and height is chosen. Then, RVE width is increased with respect to fiber diameter until convergence in elastic properties is achieved. Fig. 2 shows the convergence of homogenized engineering stiffness components while the RVE width and height (perpendicular to fiber direction) are seven times larger than the fiber diameter. Similarly, the RVE length is studied. A constant value of 70 μm is adopted, as in an earlier study (Melro et al., 2012) it has been determined to be sufficient to ensure that the stiffness components are independent of the RVE length. It should be mentioned that an RVE with a depth of one element though the thickness (along the fiber) has the same homogenized properties. CPU time is approximately seven times faster for the one-element sample simulations.

For generating the training data, different approaches are available. Here, we pick the high-fidelity 3D FE model to capture the properties of RVE while assuming some simplification (e.g., ignoring the effect of fiber misalignment on longitudinal compressive behavior) with the aim of low computational cost. Table 1 summarizes the defined parameter space. In Section 3 a detailed simulation process for both micromechanical and mesoscale RVE analyses is discussed. The ranges of parameters are extracted in a sense to embody composite components ranging from carbon to glass to natural fibers (Barbero, 2011).

2.2. Sampling technique

Having uniformly distributed input features is a critical factor for proper training of an ANN model (Géron, 2019). It is also important to avoid a regular grid of sample points, since in that case there would be multiple coincident point projections in the different hyperplanes, which may very well deteriorate the machine learning process on high-dimensional input features (Bessa et al., 2017). Effectively exploring the design space requires a consistent data set to reach a good correlation between regular grid and random distribution. Hence, a

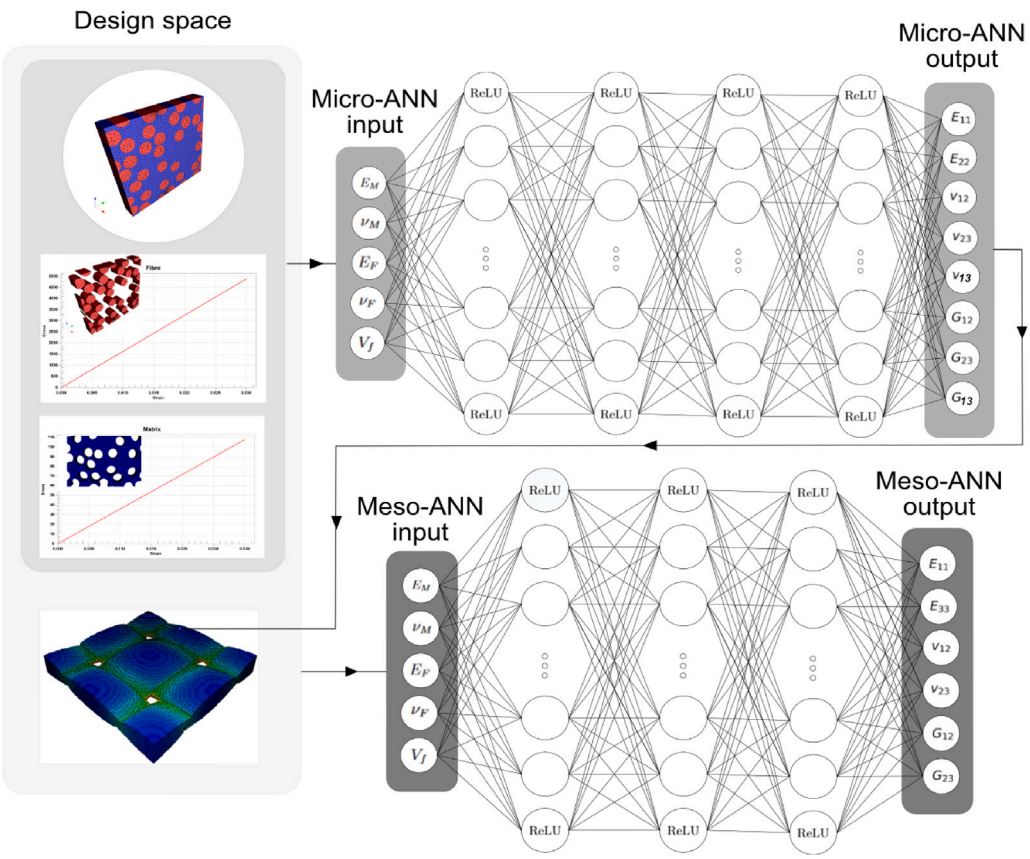


Fig. 1. A schematic representation of data flow and a summary of networks' architectures for both the micro-ANN and meso-ANN models.

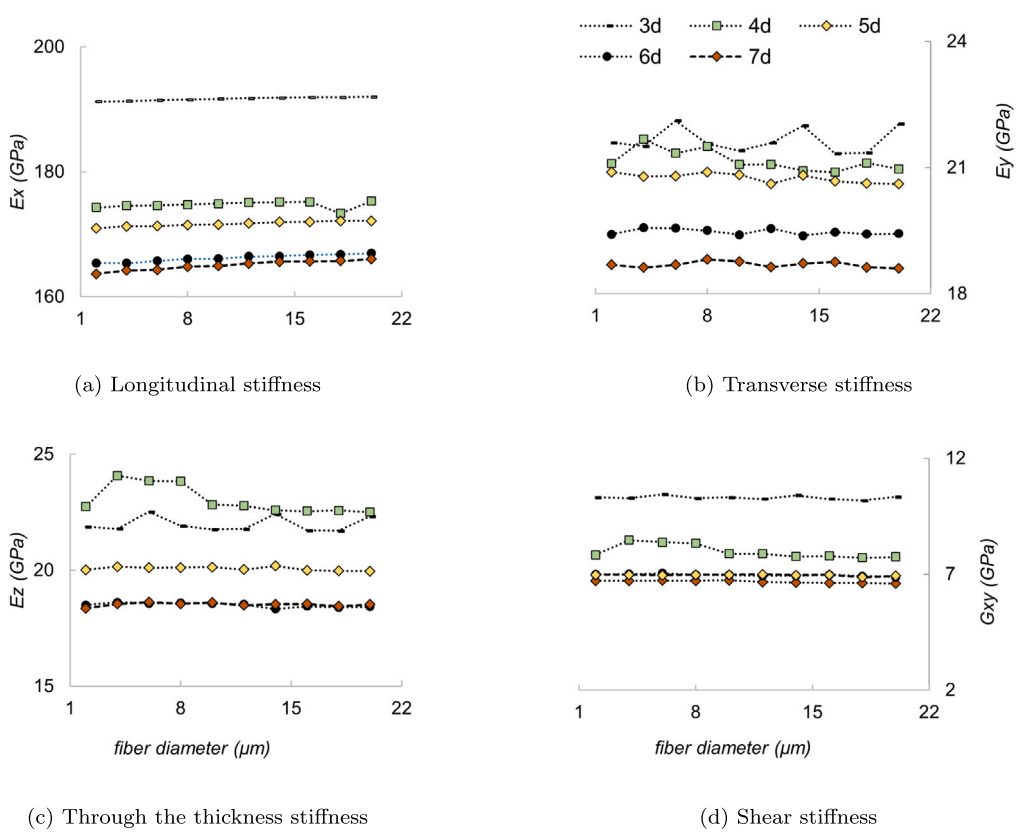


Fig. 2. Dependency study of engineering stiffness components to RVE dimensions: (a) longitudinal stiffness, (b) transverse stiffness, (c) through the thickness stiffness, and (d) shear stiffness.

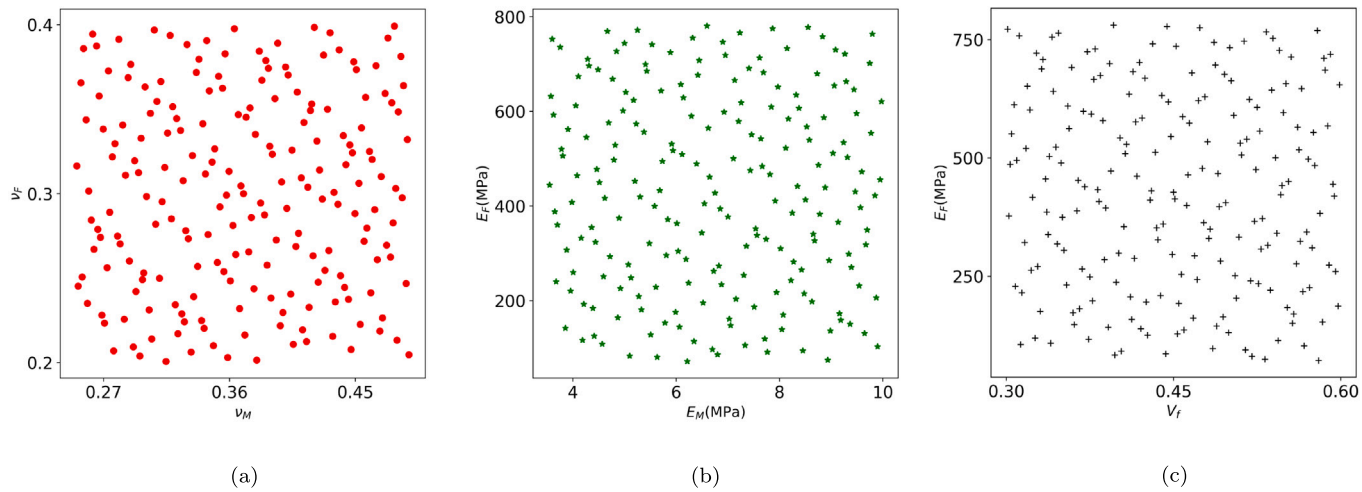


Fig. 3. Scatter plot for different pairs of parameters: (a) matrix and fiber Poisson's ratio; (b) fiber and matrix Young's modulus; (c) fiber stiffness and fiber volume fraction.

proper sampling technique facilitates achieving a random and uniform distribution with a reduced amount of data, enabling us to decrease the cost of simulations.

Random sampling and stratified sampling usually contains some clusters and gaps (Burhenne et al., 2011). To address it, Latin hypercube sampling is introduced to divide the design space dependently between features and generate samples in the new sub-intervals. This method can guarantee a uniform and random distribution if and only if enough sample points exist. Nevertheless, such sampling encounters difficulties with small amount of data. In this regard, *Sobol sequence* sampling (Saltelli et al., 2010) belongs to the family of quasi-random sampling (also known as low-discrepancy sequences) commonly used to perform uncertainty and sensitivity analyses. Sobol sequence sampling method is a class of Latin hypercube sampling with a key distinguishing point. The sample values are chosen under consideration of the previously sampled points and thus avoiding the occurrence of clusters and gaps even with a portion of the data (Burhenne et al., 2011). It generates multiple parameters as uniformly as possible over the multi-dimensional parameter space (Renardy et al., 2021; Saltelli et al., 2010; Dige and Diwekar, 2018). In fact, Sobol sequence sampling is the only one among other pseudo-random space-filling algorithms which consider predetermined sample points in the design space. A good distribution of parameters without clusters or gaps is obtained for the micromechanical analysis, as explained in Fig. 3 for three pairs of parameters. A convergence study has performed on the tasks' required data points, and 400 data points were found appropriate.

3. Data generation with finite element simulation

For generating the required data sets for both microscale and mesoscale training, computational homogenization is used on both scales. Thereby, it is considered that the homogenized results obtained for a unidirectional composite locally characterize the homogenized properties of the yarns. The large size of the modeling space (due to the range of different microstructural parameters) and complexity in mesoscale geometry make RVE generation an cumbersome process if the aim is to cover the whole parameter space on both scales. Therefore, in this work only micromechanical features are allowed to vary, while the geometrical features on the mesoscale are constant. The homogenized yarn properties are then considered as the input for the mesoscale homogenization.

3.1. Computational homogenization

Full-field homogenization can be used to predict the macroscopic response of micro-structurally heterogeneous materials by simulating a numerical RVE of the material micro-structure (Geers et al., 2010). Initially, a micro-structural RVE is defined, based on the assumption that constituents have a known constitutive behavior. The macroscopic response is obtained by volume averaging of the response of micro-structural points over the domain of the RVE (Geers et al., 2010).

In this study, computational homogenization is used to find the constitutive relationship between the macroscopic strain ($\overline{\epsilon}$) and the macroscopic stress ($\overline{\sigma}$). Elastic material properties can be obtained from the stiffness tensor ($\overline{C}$), where $\overline{\sigma} = \overline{\mathbf{C}} : \overline{\epsilon}$. To guarantee deformation equivalence across the scales, periodic boundary conditions are imposed on the boundaries of the analyzed RVEs. For every coupled node pair, constraints are used in the finite element simulation to link the displacements of each pair of the opposite cube faces as

$$\mathbf{u}^{+} - \mathbf{u}^{-} = \overline{\epsilon} \cdot \left(\mathbf{x}^{+} - \mathbf{x}^{-}\right). \tag{1}$$

here, superscripts + and − indicate the image and mirror (opposite) faces of the RVE, with node-pair coordinate x^{+} and x^{-}. The homogenized stress $\bar{\sigma}_{ij}$ is then obtained as the volume-averaged microscopic stress σ_{ij}

$$\bar{\sigma}_{ij} = \frac{1}{V}\int_{V} \sigma_{ij} dV \tag{2}$$

The volume averaged strain (over the RVE domain) has to be equal to the macro-strain Svenning et al. (2016). Elastic properties is then obtained from the homogenized stiffness tensor (relating homogenized stress to homogenized strain). In practice, all elastic properties can be determined from the Voigt matrix representation of the elastic stiffness tensor. This Voigt stiffness matrix is straightforwardly obtained by subjecting the RVE to six individual load cases, with one non-zero strain component per load case. The resulting stresses from each load case then corresponds to the corresponding column in the Voigt stiffness matrix.

3.2. Micromechanical analysis

In this work, Digimat-FE (2020) is used to create the geometry and spatial discretization of UD RVEs for imposing periodic boundary conditions, and for obtaining the homogenized properties. In the micromechanical model, prism elements with triangular cross-section are used throughout the RVE. The geometry of a UD RVE and its mesh is shown in Fig. 4.

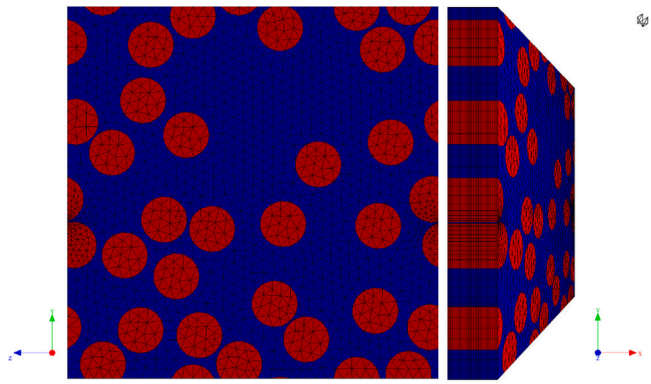


Fig. 4. Spatial discretization of a unidirectional representative volume element (RVE): Reinforcements (red) are randomly distributed in the matrix (blue).

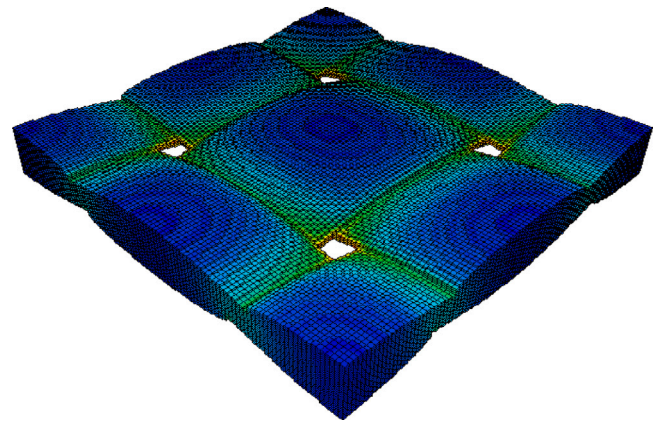


Fig. 5. Equivalent strain distribution in yarns of the woven composite RVE.

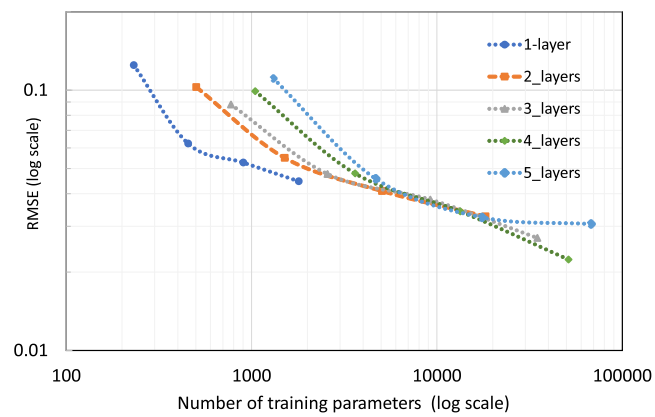


Fig. 6. Micro-ANN RMSE of the validation set vs. the number of training parameters for different numbers of hidden layers.

Compared to the diameter of the fiber, the element size is set as six times smaller in micro-RVE samples. The effective element size can be smaller in some locations (e.g., close to curved edges Digimat-FE, 2020.) No element smaller than 3% of the fiber diameter is allowed to be generated (the minimum element size). The internal coarsening algorithm in Digimat-FE mesher increases the element size inside the volume away from the bounding faces and edges. Additionally, the curvature control algorithm reduces the size of elements on curved edges and faces (Digimat-FE, 2020). Furthermore, as a result of sharing the nodes algorithm at interfaces, a continuous mesh will be maintained across phase interfaces, resulting in a reduction in the number of nodes required. As an example, for a volume fraction of 40%, the thick and one-element thickness RVEs contain 2944 and 85 376 conforming (hex-dominated) triangle elements with 3100 and 46 500 nodes, respectively. The total CPU time required to evaluate all elastic parameters for thin and thick microscale samples (mentioned in Section 2.1) is around 40 and 290 s.

To automate the data generation, a Python script is developed to run Digimat in batch mode and thereby perform UD RVE generations and simulations automatically (considering different fiber volume fractions and different constitutive properties of the fibers and matrix). Simulations are conducted with applied unit strains in normal and shear directions to obtain all elastic properties. The script then reads the simulation results and stores them in a file. Each data point refers to a unidirectional RVE generation, meshing and finite-element simulation, which takes on average 10 min in an Intel(R) Core(TM) i7-10610U CPU system.

3.3. Mesoscale analyses

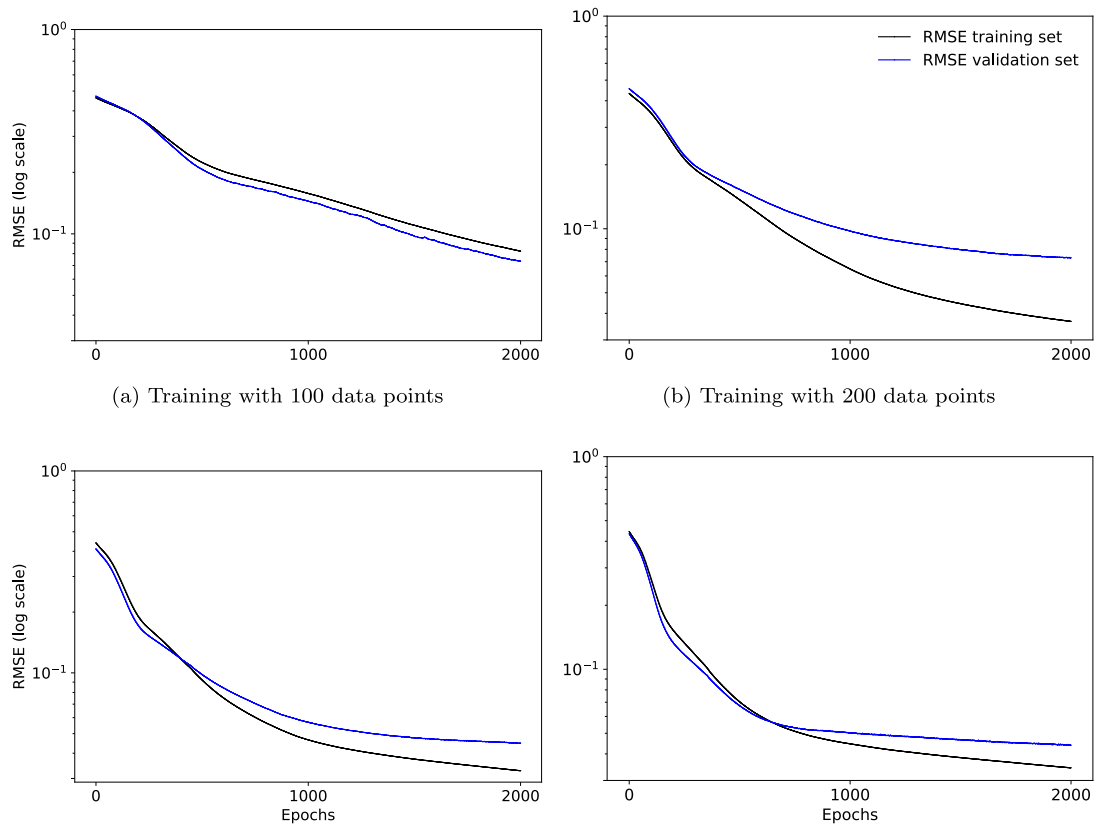
In the next step, a mesoscale plain woven RVE is generated using the geometric textile modeling software package TexGen (Brown and Long, 2021). A textile weave model is created with elliptical yarn cross-sections and with a yarn spacing of 10 mm, a yarn width of 8 mm, and a fabric thickness of 2 mm. Furthermore the reduced-order C3D8R Abaqus hexahedral continuum element is used to mesh the mesoscale RVE. A mesh sensitivity analysis is performed to ensure the output is independent of the mesh size. There are 100*100 elements in the plane and 50 elements along the thickness of the mesoscale RVE.

As mentioned above, the homogenized micromechanical properties from the micromechanical analysis are used as the equivalent material properties of yarns in the mesoscale analyses. As the unit cell of a plain woven composite is in-plane periodic, the same homogenization technique as that discussed in Section 3.1 is applicable also to these mesoscale simulations. For the implementation, TexGen has the capability to automatically define periodic boundary conditions and the finite element discretization, thereby creating a ready-to-run simulation model. We used this option in TexGen to create a model including PBCs, which was thereafter exported to an Abaqus FE solver format.

Simulations are conducted with applied unit strains periodic boundary conditions to obtain all elastic properties. The total CPU time for the mesoscale sample is 2930 s in an Intel(R) Core(TM) i5-6500 CPU @ 3.20 GHz system. A sample of the von Mises equivalent strain distribution $\left(\sqrt{\frac{2}{3}\epsilon^{dev} : \epsilon^{dev}}\right)$ where ϵ^{dev} is the deviatoric strain tensor, on the mesoscale yarns is shown in Fig. 5.

4. ANN model

The micro- and meso-scale databases prepare the ground to use supervised learning to train two feed-forward neural networks to predict both yarn and composite ply properties. The motivation to spend extra efforts to use somewhat complex neural network models rather than more straightforward machine learning methods, like logistic regression, is that the number of input and outputs in material engineering studies can expand considerably. A simple regression would not be able to handle a large number of features (Géron, 2019), while no matter how many features govern the problem, ANN can draw reasonable relations between them. The critical steps of pre-processing the data, training, and utilizing the network are presented in what follows.



(a) Training with 100 data points (b) Training with 200 data points

(c) Training with 300 data points (d) Training with 400 data points

Fig. 7. The effect of extending the data set on the RMSE convergence of micro-ANN.

4.1. Feed forward neural networks

Feed-forward neural networks are the most common ANNs in which neurons in one particular layer are connected in one direction to all neurons in the next layer. The connection strengths between layers $l-1$ and l are collected in a connectivity matrix $\mathbf{W}^{(l)}$, where $W_{ij}^{(l)}$ is the connection from neuron j in the preceding layer to neuron i in the current layer. The total input to neuron i is the weighted sum of the output $V_j^{(l-1)}$ from all N neurons in the previous layer. The input is then subtracted by a threshold value $\theta_i^{(l)}$, which models the threshold of real neurons that the input signal must exceed for the neuron to fire. The input to the neuron is thus given by

$$b_i^{(l)} = \sum_{j=1}^{N} W_{ij}^{(l)} V_j^{(l-1)} - \theta_i^{(l)}. \tag{3}$$

This quantity is referred to as the local field of neuron i in layer l. The input is then passed through a nonlinear function $g(\cdot)$ referred to as the activation function, which models the response the neuron has to the received input. To tackle over-fitting problem the rectified linear unit is chosen as activation functions (Mehlig, 2021):

$$\mathrm{ReLU}(x) = \begin{cases} x & \text{if} \quad x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Applying the activation function, we obtain the output of neuron i in layer l as

$$V_i^{(l)} = g(b_i^{(l)}) \quad \text{where} \quad l = 1, 2, \dots, L. \tag{5}$$

The training process is basically a minimization problem of cost function C, where $(t_i - V_i^{(L)})^2$ is the mean square error of target values and the neurons of the network's last layer. The summation is over all the n sample points in the test set.

$$C = \frac{1}{n} \sum_{i=1}^{n} (t_i - V_i^{(L)})^2 \tag{6}$$

The output of last layer $\mathrm{V}^{(L)}$ is given by activation functions. For the sake of notation simplicity, the neurons indices, (i, j), are not written in the following. Backpropagation is used in ANN to update weights and thresholds in forward direction and errors in backward direction during the training process. The network first computes the gradients of neurons in output layer with respect to all neurons in all preceding hidden layers as in Eq. (7).

$$\frac{\partial V^{(L)}}{\partial V^{(l)}} = \prod_{k=L}^{l+1} [w^{(k)} g'(b^{(k)})] \tag{7}$$

Then the network computes the error $\delta_i^{(l)}$ using Eq. (8) as

$$\delta^{(l)} = [t - V^{(L)}(x)] g'(b^{(L)}) \prod_{k=L}^{l+1} [w^{(k)} g'(b^{(k-1)})]. \tag{8}$$

The error in each layer (l) is used to update weights and thresholds between each neurons by adding $\delta w^{(l)}$ and $\delta\theta^{(l)}$, defined in Eqs. (9) and (10), to the previous weights and thresholds considering the learning rate (μ):

$$\delta w^{(l)} = \mu \delta^{(l)} V^{(l-1)}, \tag{9}$$

$$\delta\theta^{(l)} = \mu \delta^{(l)}. \tag{10}$$

Weights and thresholds are usually initialized as Gaussian random variables with mean zero and variance one over the number of neurons in the preceding layer (Mehlig, 2021), that give rise to large local field $|b|$ at the beginning of training. Large local fields and the form of

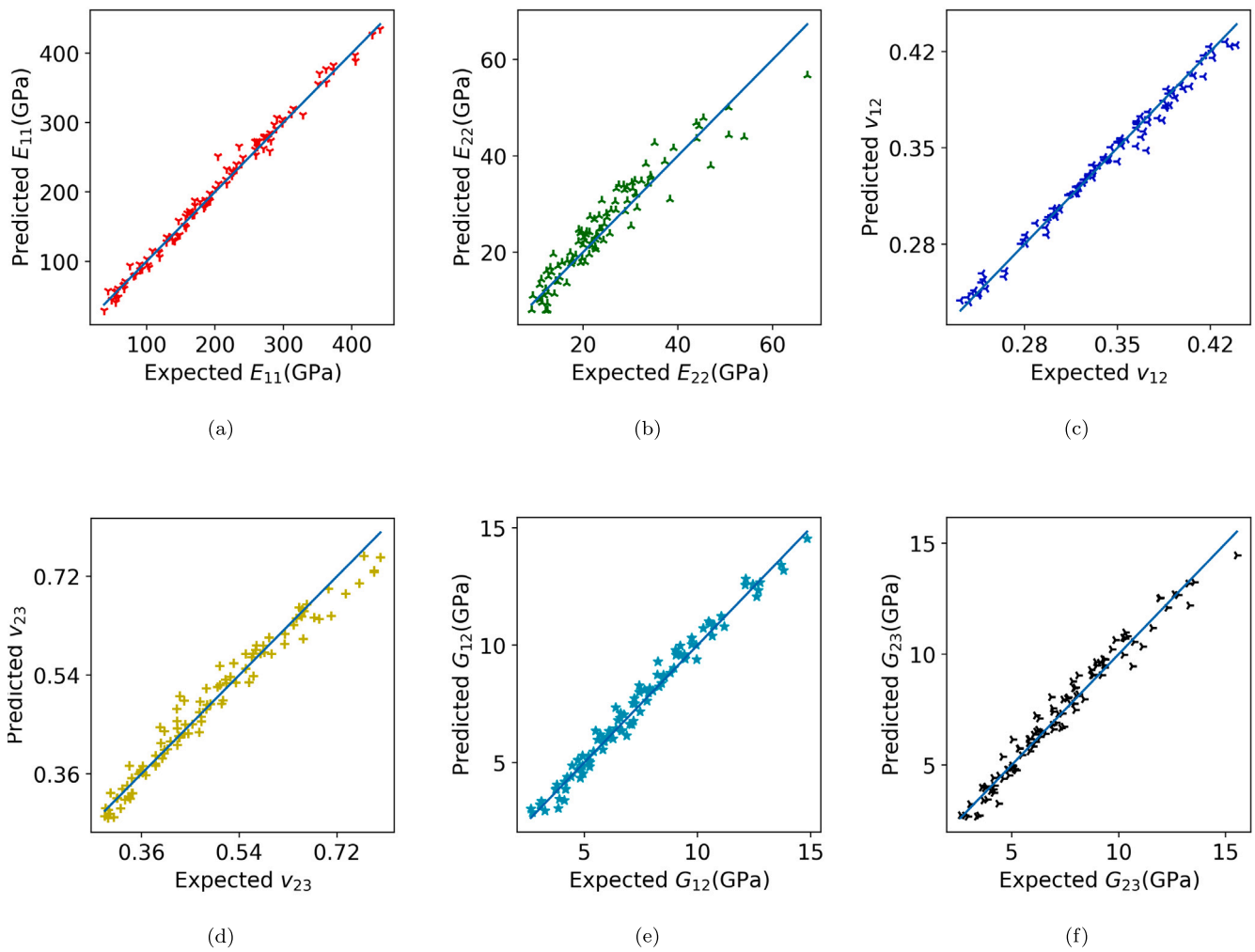


Fig. 8. Micro-ANN prediction on unseen data vs. (ground truth) expected UD RVE engineering stiffness coefficients.

activation functions cause exponentially small derivation of activation function $g'(b)$. Therefore, the error $\delta^{(l)}$ in Eq. (8) or gradients ($\delta w^{(l)}$ and $\delta\theta^{(l)}$) vanishes quickly for earliest hidden layers (as l decreases). The result will be a slow training known as vanishing gradients decent. To address this issue, one needs to feed the network with the whole training data-set several times, called an *epoch* each time, and exploit the data set. On the other hand, more epochs may cause over-fitting problem for the network. The difficulty of designing and training an ANN is usually related to minimization of the cost function (corresponding to convergence in the cost function) without facing over-fitting.

4.2. Scaling

Differences in the magnitudes across input variables may increase the difficulty of the modeled problem. Large input values may lead to an ANN model with large weight values. The larger the weights are, the more unstable the model is, resulting in a worse training performance (Géron, 2019). Large input values also causes higher error values and weak convergence. Therefore, it is essential to scale the input parameters. The two most used methods are normalization and standardization. Normalization maps data within the range of 0 to 1, based on each feature's minimum and maximum values. Standardization re-scales the distribution of values so that the mean of observed values is 0 and the standard deviation is 1 (Géron, 2019). Without normalization, the objective function gives a higher priority to higher values. So with Young's moduli being several orders of magnitude larger than other quantities, the Poisson's ratios are probably neglected in the ANN learning process.

In this work, Scikit-learn (Buitinck et al., 2013), an ML library for the Python programming language, is used to normalize both input and output data. All the features were transformed using the respective mean and variance from the training set. Furthermore, the test set is scaled with the mean and variance of the training set, since we aim to hide the exact feature of the test set from the model and estimate the model performance on the test (unseen) data. Finally, each parameter in the training data set is normalized based on its own range of values. So the highest value is then equal to 1, and the lowest is equivalent to 0.

4.3. Neural networks design

Using the open source machine learning library, Keras (Chollet, 2021), a feed-forward neural network is considered sufficient due to the problem linearity (Mentges et al., 2021). As discussed in Section 2.1, both the microscopic and mesoscopic ANN models contain 5 independent input variables, the fiber and matrix constitutive properties, and fiber volume fraction. In the output layer, 6 and 8 independent variables for the micro- and meso-ANN, respectively, are considered for describing the composites' engineering stiffness variables (see Fig. 1.) The data set is then split randomly into 80% for training and the rest for testing the ANN model. After each epoch (feeding once the whole training set) the Root Mean Squared Error (RMSE) is calculated

according to Eq. (11) between the output values in the database $y_{\text{data},i}$ and the ANN model predictions $y_{\text{model},i}$.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\text{data},i} - y_{\text{model},i}\right)^2}. \tag{11}$$

Using excessively complex configurations can adversely affect the model's efficiency due to vanishing and exploding gradients in the back-propagation process (Mehlig, 2021; Zobeiry et al., 2020; Géron, 2019). Some tuning techniques, e.g. the grid-search hyper-parameter, are suggested to find the optimal number of neurons and hidden layers (Mentges et al., 2021). In this study, for the purpose of reaching an accurate ANN model, a variety of architectures, from very simple to rather complex networks, are examined. In all ANN architectures, an equal number of hidden neurons are assumed for hidden layers. Starting from networks consisting of 1 hidden layer and 16 neurons, we systematically increased them to 5 hidden layers with 128 neurons until the ANNs accurately captured the expected response at both scales. Fig. 6 illustrates the effect of adding more hidden layers and neurons to the micro-ANN model. As can be seen, there is a slight increase in the generalization error (RMSE on the test set in this case) in the network with five hidden layers and 128 hidden neurons in each layer (when the number of training parameters is increased from 17 544 to 67 848). Thus, the networks with the lowest RMSE before observing over-fitting are selected. In this regard, The microscopic and mesoscopic ANN models consist of 4 and 3 densely connected layers, with 64 nodes using the max(0,x) (ReLu) activation function. Furthermore, the Adam algorithm (Kingma and Lei, 2015) is adopted for parameters update with the learning rate of 1e−5 for both the microscopic and mesoscopic ANN models after comparing the convergence time and minimum RMSE. The minimum RMSE on the validation set with this approach is found as 0.043978 and 0.030215 for micro- and meso-ANN, respectively.

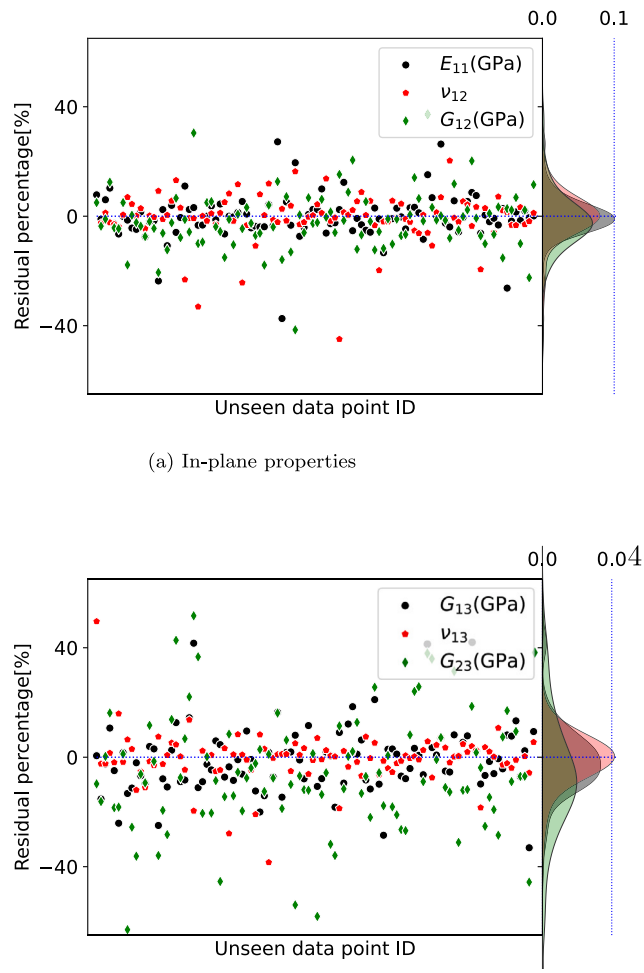
5. Results and discussion

This section compares the obtained results from microscale and mesoscale ANNs with finite element simulations. A further discussion on the influence of hyperparameters and features is also provided.

5.1. Micro-ANN results

We utilize the Sobol sequence technique as an optimized way to explore the design space with a limited number of training data. Therefore, we started by generating a limited number of data points and performed an initial training of the network. Checking the residual errors, we performed more FE simulations and supplied the network with additional data points. Finally, fine-tuning the hyperparameters (i.e., the number of hidden layers and neurons) were performed. Fig. 7 shows the effect of extending the data set size from 100 to 400 samples on the training convergence of the network.

Results show that around 300 data points are enough for the network to converge in a reasonable number of epochs without over-fitting (early-stopping technique guarantees that the loss value for the validation set is converged to a value close to zero and is not increasing for a certain number of epochs) to predict elastic properties of the UD composite. Moreover, The RMSE of the trained network on the validation set is 0.041 using 300 data point. As it can be seen in Fig. 7 the curves of the training and validation intersect on graph (c). It is not impossible that the RMSE of the test set will be slightly lower than that of the training set. Small databases are particularly susceptible to such a rare event. While the Sobol sampling technique avoids an unbalanced data set, and there is no dropout or augmentation in the training process (Mehlig, 2021), it is possible that with only 300 data points the validation data points were selected by chance in such a way that they are too easy to predict by the network. By performing k-fold



(a) In-plane properties

(b) Out-of-plane properties

Fig. 9. Micro-ANN residuals of predictions vs. expected values of mesoscale properties for the test (unseen data) set. The marginal graph also indicates that the distribution of residuals is close to normal.

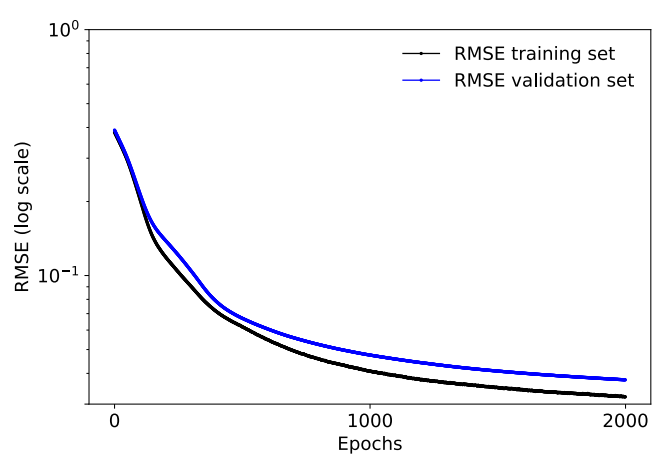


Fig. 10. Meso-ANN convergence on loss and root mean square error for both validation and training sets.

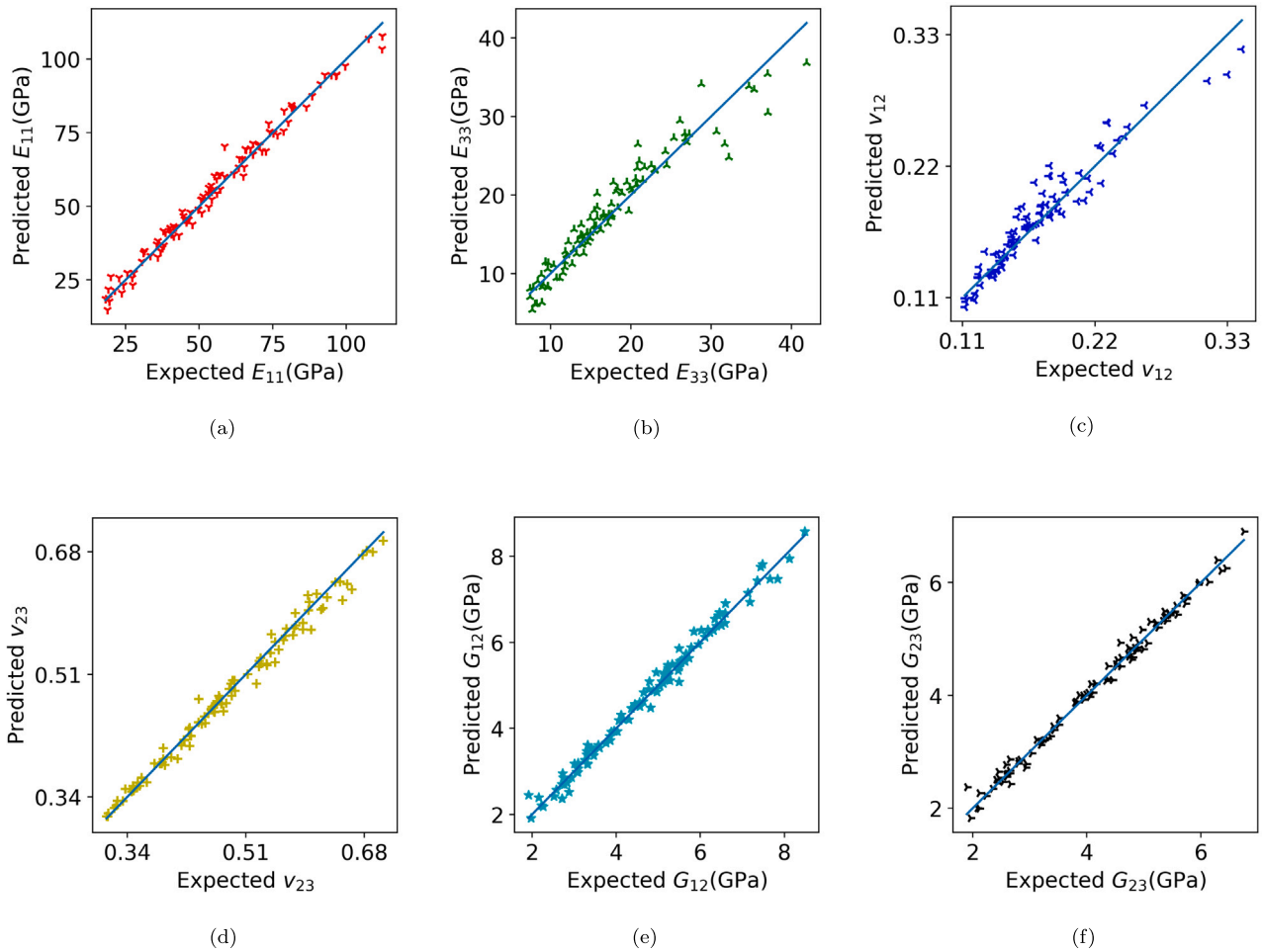


Fig. 11. Mesoscale ANN prediction vs. ground truth (expected) engineering stiffness coefficients for the test (unseen data) set.

cross-validation (Géron, 2019) in the training process, we ensured that there was no systematic discrepancy in the loss or RMSE values. Graph (a) with only 100 data points, however, is systematically observed to have a lower RMSE on the test set, indicating an insufficient data set. To reduce the deviation of predictions from the test set and reach smaller RMSE on the validation set, 100 more data points are generated and fed into the network. The validation RMSE of the trained micro-ANN on the final data set with 400 data points is 0.034. The rest of the results are provided by using 400 data points for training the micro- and mesoscale neural networks.

The obtained engineering stiffness coefficients of the micromechanical model together with the results obtained from the micro-ANN model are shown in Fig. 8. Assuming "1" as the fiber direction and "2–3" as two directions perpendicular to the fibers, longitudinal stiffness (E_{11}), transverse (E_{22}), shear moduli (G_{12} and G_{23}), and Poisson's ratios (ν_{12} and ν_{23}) are compared against the micromechanical FE predictions, for the test set. It is seen that the ANN model provides good predictions for the elastic modulus.

For the case with a single element through the thickness, the FEM simulation of the microscopic RVE takes more than 40 s (approximately 7 s for each homogenized elastic parameter), compared to the trained ANN model which predicts all engineering material properties simultaneously in less than a second. Although the FEM simulations depend on many parameters but for our setup, we could observe approximately a 40-fold decrease in the computational time required to evaluate elastic engineering properties at the microscale.

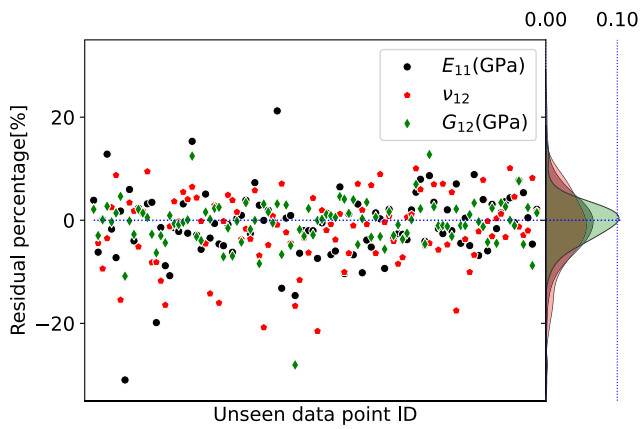
Errors scatter plot of residuals of the micro-ANN model with respect to the ground-truth values are presented in Fig. 9. Less than 20% error is observed with 95% confidence interval.

5.2. Meso-ANN results

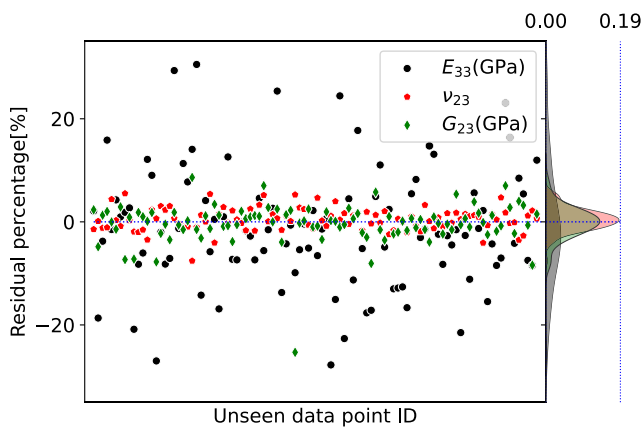
The micro-ANN trained to predict the micromechanical behavior of unidirectional composites has been used then to predict local material properties of yarns in a woven composite. Following the micromechanical study, a total of 400 mesoscale data points are considered. A second ANN is trained on 80% of the training data based on the micromechanical features and the engineering elastic properties of woven RVE computed from finite element analysis. Training results for the most optimized meso-ANN model, with three hidden layers including 64 hidden neurons each, is summarized in Fig. 10. Results in Fig. 11 show a good agreement between the predicted values by meso-ANN and the output of FE simulations of the mesoscale woven RVE. Scatter plot of residuals with respect to true values are presented in Fig. 12. There, it can be seen that the predictions have a maximum error of less than 10% when compared to true values with a confidence interval of 95%. Furthermore, it can be concluded that the residual distribution is close to normal with a mean of less than 0.1 and standard deviation presented in Table 2. The error distributions indicates that the model captured the predictive information, and what left behind (residuals) is stochastic (i.e., unpredictable).

6. Conclusions

In this study, artificial neural networks are employed for efficient and accurate prediction of elastic properties of woven composites as a function of varying fiber and matrix properties and micro- and mesoscale characteristics, such as local fiber volume fraction and weaving pattern (the latter is kept as fixed in the current paper). An



(a) In-plane properties



(b) Out-of-plane properties

Fig. 12. Meso-ANN residuals of predictions vs. expected values of mesoscale properties for the test (unseen data) set. Marginal graph also indicates that the distribution of residuals is close to normal.

Table 2
Standard deviation of residuals for the meso-ANN.

Mesoscale properties	STD [%]
E_{11}	6.70
ν_{12}	6.94
G_{12}	4.70
E_{33}	11.48
ν_{23}	2.19
G_{23}	3.88

automatic process has been developed to create the database for training and validating the micromechanical and mesoscale neural network models based on finite element analysis. More specifically, both the microscale and mesoscale data are generated via computational homogenization on two sets of 3D Representative Volume Elements representing each scale. Based on the data generated, two feed-forward neural networks concerning micromechanical UD composites and mesoscale woven composites were then trained and validated. 4 hidden layers for micro-ANN and 3 for meso-ANN, both with 64 nodes in each layer equipped with ReLu activation functions found proper.

The ANN model predictions compare well with microscale and mesoscopic data, which demonstrates the predictive capability of the models. Based on the current setup, the CPU time required to evaluate all elastic parameters using the finite element method is approximately 40 (for samples with one element through the thickness) and 2390 s per sample for the micro- and the mesoscale RVEs, respectively. The developed multiscale neural network model is capable of predicting the parameters in less than a second. Thus, the developed ANN models can be used both for fast and accurate calculation of the elastic properties of a wide range of UD composites with varying constituent and fiber volume fractions, as well as for the prediction of similar properties for woven fabric composites. Although we made the restriction to a single weave type for demonstration purposes, a natural extension of the work is to also consider different ranges of geometrical features of woven RVEs.

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

Data will be made available on request.

MultiscaleElasticWovenNeuralNet (Original data) (GitHub)

Acknowledgments

S.M. Mirkhalaf and E. Ghane gratefully acknowledge financial support from the Swedish Research Council (VR grant: 2019-04715) and the University of Gothenburg. M. Fagerström is thankful for the support through Vinnova's strategic innovation programme LIGHTer, in particular via the project LIGHTer Academy Phase 3 (grant no. 2020-04526).

References

Aboudi, Jacob, Arnold, Steven M., Bednarcyk, Brett A., 2013. Micromechanics of Composite Materials: A Generalized Multiscale Analysis Approach. jKJ2VqSwu08C, Butterworth-Heinemann, Google-Books-ID.

Adumitroaie, Adi, Barbero, Ever J., 2011. Beyond plain weave fabrics – II. Mechanical properties. Compos. Struct. 93 (5), 1449–1462. http://dx.doi.org/10.1016/j.compstruct.2010.11.016, URL https://www.sciencedirect.com/science/article/pii/S0263822310003909.

Bai, Xiaoming, Bessa, Miguel A., Melro, António R., Camanho, Pedro P., Guo, Licheng, Liu, Wing K., 2015. High-fidelity micro-scale modeling of the thermo-visco-plastic behavior of carbon fiber polymer matrix composites. Compos. Struct. 134, 132–141. http://dx.doi.org/10.1016/j.compstruct.2015.08.047, URL https://linkinghub.elsevier.com/retrieve/pii/S026382231500728X.

Barbero, Ever J., 2011. Introduction To Composite Materials Design, second ed. CRC Press, http://dx.doi.org/10.1201/9781439894132.

Bessa, M.A., Bostanabad, R., Liu, Z., Hu, A., Apley, Daniel W., Brinson, C., Chen, W., Liu, Wing Kam, 2017. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. Comput. Methods Appl. Mech. Engrg. 320, 633–667. http://dx.doi.org/10.1016/j.cma.2017.03.037, URL https://www.sciencedirect.com/science/article/pii/S0045782516314803.

Bischof, Rafael, Kraus, Michael, 2021. Multi-objective loss balancing for physics-informed deep learning. http://dx.doi.org/10.13140/RG.2.2.20057.24169, URL http://arxiv.org/abs/2110.09813, arXiv:2110.09813,

Bonatti, Colin, Berisha, Bekim, Mohr, Dirk, 2022. From CP-FFT to CP-RNN: Recurrent neural network surrogate model of crystal plasticity. Int. J. Plast. 158, 103430. http://dx.doi.org/10.1016/j.ijplas.2022.103430, URL https://www.sciencedirect.com/science/article/pii/S074964192200208X.

Brown, L.P., Long, A.C., 2021. 8 - modeling the geometry of textile reinforcements for composites: TexGen. In: Boisse, Philippe (Ed.), Composite Reinforcements for Optimum Performance (Second Edition). In: Woodhead Publishing Series in Composites Science and Engineering, Woodhead Publishing, pp. 237–265. http://dx.doi.org/10.1016/B978-0-12-819005-0.00008-3, URL https://www.sciencedirect.com/science/article/pii/B9780128190050000083.

Buitinck, Lars, Louppe, Gilles, Blondel, Mathieu, Pedregosa, Fabian, Mueller, Andreas, Grisel, Olivier, Niculae, Vlad, Prettenhofer, Peter, Gramfort, Alexandre, Grobler, Jaques, Layton, Robert, Vanderplas, Jake, Joly, Arnaud, Holt, Brian, Varoquaux, Gaël, 2013. API design for machine learning software: experiences from the scikit-learn project. http://dx.doi.org/10.48550/arXiv.1309.0238, URL https://arxiv.org/abs/1309.0238v1.

Burhenne, Sebastian, Jacob, Dirk, Henze, Gregor P., 2011. Sampling based on sobol' sequences for Monte Carlo techniques applied to building simul ations. In: Proceedings of Building Simulation. Sydney. pp. 1816–1823, URL http://www.ibpsa.org/proceedings/BS2011/P_1590.pdf.

Chollet, Francois, 2021. Deep Learning with Python, Second Edition. mjVKEAAAQBAJ, Simon and Schuster, Google-Books-ID.

Dige, Nishant, Diwekar, Urmila, 2018. Efficient sampling algorithm for large-scale optimization under uncertainty problems. Comput. Chem. Eng. 115, 431–454. http://dx.doi.org/10.1016/j.compchemeng.2018.05.007, URL https://www.sciencedirect.com/science/article/pii/S009813541830437X.

Digimat-FE, 2020. URL https://www.e-xstream.com/products/digimat/tools?fe=1.

Friemann, J., Dashtbozorg, B., Fagerström, M., Mirkhalaf, S.M., 2023. A micromechanics-based recurrent neural networks model for path-dependent cyclic deformation of short fiber composites. Internat. J. Numer. Methods Engrg. 124 (10), 2292–2314. http://dx.doi.org/10.1002/nme.7211, URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85147506285&doi=10.1002%2fnme.7211&partnerID=40&md5=b3ca3ff1a882fdd35d8e834695932a2c.

Furtado, C., Pereira, L.F., Tavares, R.P., Salgado, M., Otero, F., Catalanotti, G., Arteiro, A., Bessa, M.A., Camanho, P.P., 2021. A methodology to generate design allowables of composite laminates using machine learning. Int. J. Solids Struct. 233, 111095. http://dx.doi.org/10.1016/j.ijsolstr.2021.111095, URL https://linkinghub.elsevier.com/retrieve/pii/S0020768321001852.

Geers, M.G.D., Kouznetsova, V.G., Brekelmans, W.A.M., 2010. Computational homogenization. In: Pippan, Reinhard, Gumbsch, Peter (Eds.), Multiscale Modelling of Plasticity and Fracture By Means of Dislocation Mechanics. In: CISM International Centre for Mechanical Sciences, Springer, pp. 327–394. http://dx.doi.org/10.1007/978-3-7091-0283-1_7.

Géron, Aurélien, 2019. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques To Build Intelligent Systems. "O'Reilly Media, Inc.", Google-Books-ID.

Haghighat, Ehsan, Raissi, Maziar, Moure, Adrian, Gomez, Hector, Juanes, Ruben, 2021. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Comput. Methods Appl. Mech. Engrg. 379, 113741. http://dx.doi.org/10.1016/j.cma.2021.113741, URL https://linkinghub.elsevier.com/retrieve/pii/S0045782521000773.

Heidari-Rarani, M, Bashandeh-Khodaei-Naeini, K, Mirkhalaf, SM, 2018. Micromechanical modeling of the mechanical behavior of unidirectional composites – A comparative study. J. Reinf. Plast. Compos. 37 (16), 1051–1071. http://dx.doi.org/10.1177/0731684418779441.

Herakovich, Carl T., 1998. Mechanics of fibrous composites. New York: John Wiley & Sons, Inc, 1998..

Ivanov, Dmitry S., Lomov, Stepan V., 2020. 2 - modeling of 2D and 3D woven composites. In: Irving, Philip, Soutis, Constantinos (Eds.), Polymer Composites in the Aerospace Industry (Second Edition). In: Woodhead Publishing Series in Composites Science and Engineering, Woodhead Publishing, pp. 23–57. http://dx.doi.org/10.1016/B978-0-08-102679-3.00002-2, URL https://www.sciencedirect.com/science/article/pii/B9780081026793000022.

Karapiperis, K., Stainier, L., Ortiz, M., Andrade, J.E., 2021. Data-driven multiscale modeling in mechanics. J. Mech. Phys. Solids 147, 104239. http://dx.doi.org/10.1016/j.jmps.2020.104239, URL https://linkinghub.elsevier.com/retrieve/pii/S0022509620304531.

Kingma, Diederik P., Lei, Jimmy, 2015. Adam: A method for stochastic optimization. arXiv: 1412.6980 [cs.LG].

Linka, Kevin, Hillgärtner, Markus, Abdolazizi, Kian P., Aydin, Roland C., Itskov, Mikhail, Cyron, Christian J., 2021. Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. J. Comput. Phys. 429, 110010. http://dx.doi.org/10.1016/j.jcp.2020.110010, URL https://www.sciencedirect.com/science/article/pii/S0021999120307841.

Liu, Xin, Tian, Su, Tao, Fei, Yu, Wenbin, 2021. A review of artificial neural networks in the constitutive modeling of composite materials. Composites B 224, 109152. http://dx.doi.org/10.1016/j.compositesb.2021.109152, URL https://linkinghub.elsevier.com/retrieve/pii/S1359836821005321.

Mehlig, Bernhard, 2021. Machine Learning with Neural Networks: An Introduction for Scientists and Engineers, first ed. Cambridge University Press, http://dx.doi.org/10.1017/9781108860604, URL https://www.cambridge.org/core/product/identifier/9781108860604/type/book.

Melro, A.R., Camanho, P.P., Pinho, S.T., 2012. Influence of geometrical parameters on the elastic response of unidirectional composite materials. Compos. Struct. 94 (11), 3223–3231. http://dx.doi.org/10.1016/j.compstruct.2012.05.004, URL https://www.sciencedirect.com/science/article/pii/S0263822312002139.

Mentges, N., Dashtbozorg, B., Mirkhalaf, S.M., 2021. A micromechanics-based artificial neural networks model for elastic properties of short fiber composites. Composites B 213, 108736. http://dx.doi.org/10.1016/j.compositesb.2021.108736, URL https://linkinghub.elsevier.com/retrieve/pii/S1359836821001281.

Mozaffar, Mojtaba, Paul, Arindam, Al-Bahrani, Reda, Wolff, Sarah, Choudhary, Alok, Agrawal, Ankit, Ehmann, Kornel, Cao, Jian, 2018. Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks. Manuf. Lett. 18, 35–39. http://dx.doi.org/10.1016/j.mfglet.2018.10.002, URL https://linkinghub.elsevier.com/retrieve/pii/S2213846318300804.

Peng, Grace C.Y., Alber, Mark, Tepole, Adrian Buganza, Cannon, William, De, Suvranu, Dura-Bernal, Salvador, Garikipati, Krishna, Karniadakis, George, Lytton, William W., Perdikaris, Paris, Petzold, Linda, Kuhl, Ellen, 2020. Multiscale modeling meets machine learning: What can we learn? arXiv:1911.11958 [physics].

Renardy, Marissa, Joslyn, Louis R., Millar, Jess A., Kirschner, Denise E., 2021. To sobol or not to sobol? The effects of sampling schemes in systems biology applications. Math. Biosci. 337, 108593. http://dx.doi.org/10.1016/j.mbs.2021.108593, URL https://linkinghub.elsevier.com/retrieve/pii/S0025556421000419.

Saltelli, Andrea, Annoni, Paola, Azzini, Ivano, Campolongo, Francesca, Ratto, Marco, Tarantola, Stefano, 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Comm. 181 (2), 259–270. http://dx.doi.org/10.1016/j.cpc.2009.09.018, URL https://www.sciencedirect.com/science/article/pii/S0010465509003087.

Shokrieh, MM, Ghasemi, R., Mosalmani, R., 2017. A general micromechanical model to predict elastic and strength properties of balanced plain weave fabric composites. J. Compos. Mater. 51 (20), 2863–2878. http://dx.doi.org/10.1177/0021998317716530.

Svenning, Erik, Fagerström, Martin, Larsson, Fredrik, 2016. On computational homogenization of microscale crack propagation. Internat. J. Numer. Methods Engrg. 108 (1), 76–90.

Thomas, Akshay J., Barocio, Eduardo, Bilionis, Ilias, Pipes, R. Byron, 2022. Bayesian inference of fiber orientation and polymer properties in short fiber-reinforced polymer composites. Compos. Sci. Technol. 228, 109630. http://dx.doi.org/10.1016/j.compscitech.2022.109630, URL https://linkinghub.elsevier.com/retrieve/pii/S0266353822003724.

Varandas, Luís F., Catalanotti, Giuseppe, Melro, António R., Falzon, Brian G., 2020. On the importance of nesting considerations for accurate computational damage modelling in 2D woven composite materials. Comput. Mater. Sci. 172, 109323. http://dx.doi.org/10.1016/j.commatsci.2019.109323, URL https://linkinghub.elsevier.com/retrieve/pii/S0927025619306226.

Vlassis, Nikolaos N., Sun, WaiChing, 2021. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Comput. Methods Appl. Mech. Engrg. 377, 113695. http://dx.doi.org/10.1016/j.cma.2021.113695, URL https://www.sciencedirect.com/science/article/pii/S0045782521000311.

Wen, P.H., Aliabadi, M.H., 2009. Mesh-free micromechanical model for woven fabric composite elastic moduli. J. Multiscale Model. 01 (2), 303–319. http://dx.doi.org/10.1142/S175697370900013X, URL https://www.worldscientific.com/doi/10.1142/S175697370900013X.

Wu, Ling, Adam, Laurent, Noels, Ludovic, 2021. Micro-mechanics and data-driven based reduced order models for multi-scale analyses of woven composites. Compos. Struct. 270, 114058. http://dx.doi.org/10.1016/j.compstruct.2021.114058, URL https://www.sciencedirect.com/science/article/pii/S0263822321005183.

Yan, Shibo, Zou, Xi, Ilkhani, Mohammad, Jones, Arthur, 2020. An efficient multiscale surrogate modelling framework for composite materials considering progressive damage based on artificial neural networks. Composites B 194, 108014. http://dx.doi.org/10.1016/j.compositesb.2020.108014, URL https://linkinghub.elsevier.com/retrieve/pii/S1359836820303279.

Zobeiry, Navid, Reiner, Johannes, Vaziri, Reza, 2020. Theory-guided machine learning for damage characterization of composites. Compos. Struct. 246, 112407. http://dx.doi.org/10.1016/j.compstruct.2020.112407, URL https://www.sciencedirect.com/science/article/pii/S0263822320301756.