Multi-fidelity data fusion for inelastic woven composites: Combining recurrent neural networks with transfer learning

Downloaded from: https://research.chalmers.se, 2026-10-17 21:01 UTC

Citation for the original published paper (version of record):
Ghane, E., Fagerström, M., Mirkhalaf, M. (2025). Multi-fidelity data fusion for inelastic woven composites: Combining recurrent neural networks
with transfer learning. Composites Science and Technology, 267.
http://dx.doi.org/10.1016/j.compscitech.2025.111163

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Composites Science and Technology 267 (2025) 111163

Contents lists available at ScienceDirect

Composites Science and Technology

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

Multi-fidelity data fusion for inelastic woven composites: Combining recurrent neural networks with transfer learning

Ehsan Ghane [a,*], Martin Fagerström [b], Mohsen 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: https://github.com/EsanGhaneh/Multi-fidelity-Data-Fusion-for-In-elastic-Woven-Composites.git

Keywords:
Woven composites
Meso-scale
Plasticity
Fast-Fourier transform
Gated recurrent networks
Transfer-learning

ABSTRACT

Surrogate deep learning models provide an efficient solution for reducing the computational demands of homogenizing complex meso-scale woven composites to study their elasto-plastic mechanical behaviors. This research introduces a comprehensive framework using transfer learning that combines data from a mean-field homogenization approach with high-fidelity full-field simulations. In a design space characterized by diverse loading conditions and micro-scale constitutive material properties, the goal is to address the challenges of generating sufficient datasets for training a data-hungry gated recurrent neural network (GRU). Multiple datasets of varying precision are generated and used, containing multi-axial stress–strain responses under two load types: random walking and proportional cyclic loading. Moreover, this study emphasizes the importance of temporal correlations in the dataset, which align with the physically path-dependent behavior of most non-linear materials, and demonstrates that temporal correlations are crucial for training time-series models. These correlations also provide the foundation for data augmentation using a linear interpolation technique within time-series stress analyses, enabling accurate predictions of homogenized meso-scale stresses based on strain trajectories and microstructural properties. Results demonstrate that integrating transfer learning with neural networks successfully incorporates a limited number of high-fidelity data with more accessible but low-fidelity data. With this framework, surrogate models for predicting the complex behavior of woven composites will be accurate and efficient, marking an important advancement in material modeling.

1. Introduction

Woven composites [1] incorporate yarns oriented in two or three [2] directions, typically resulting in a dense weave along the primary load-bearing direction. A woven reinforcement textile composite can conform to complex curvatures while maintaining desirable and balanced mechanical properties. Textile composites come in various weaving patterns, which use over-and-under interlacing to improve drapeability and resin wettability [3]. The cost of manufacturing necessitates the creation of comprehensive modeling strategies. However, woven composites present computational challenges for predicting mechanical response because of their complex geometry and interlacing. Due to their complicated behavior under mechanical loads, developing accurate and computationally efficient modeling approaches is an important field to study.

Prediction methods for non-linear mechanical properties of woven composites can be broadly classified into analytical, semi-analytical, and numerical approaches. Analytical techniques, such as the rule of mixtures, asymptotic homogenization [4], and mean-field homogenization (MFH) [5,6], offer efficient solutions for predicting elastic and non-linear behaviors but are often limited in capturing complex nonlinear mechanisms. Semi-analytical methods, including the method of cells [7], transformation field analysis [8], and clustering analysis [9,10], derive macroscopic constitutive relationships from local-scale behaviors, offering a balance between accuracy and computational cost.

Numerical methods, particularly computational homogenization approaches such as finite element (FE) analysis [11–13] and mesh-free methods [14], are more powerful for simulating path-dependent behaviors, including damage and plasticity in textile composite architectures. Although computationally demanding, these methods provide greater accuracy, especially when modeling complex failure modes. At the meso-scale, models such as [15–18] focus on damage and deformation within the woven architecture, capturing detailed interactions at the yarn and fiber level. Furthermore, stochastic analysis techniques

* Corresponding author.
E-mail address: ehsan.ghane@physics.gu.se (E. Ghane).

https://doi.org/10.1016/j.compscitech.2025.111163
Received 13 September 2024; Received in revised form 7 February 2025; Accepted 22 March 2025
Available online 5 April 2025
0266-3538/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

[19–22] offer insights into the reliability and performance of woven composites under uncertainty, addressing variations in material properties and microstructures [23].

Recently, neural network surrogate modeling has been developed for rapid and comprehensive solutions. Neural networks are advanced algorithms that build relationships between inputs and outputs, not exclusively with a physical explanation. While this lack of explanation can be a disadvantage, neural networks use universal functions to map input parameters to output parameters by establishing relationships. Using neural networks, one can generate a fast, accurate surrogate model based on experimental data, simulated responses, or a combination of both [24]. Research on artificial neural network (ANN) models for composites involves using geometric parameters, material properties, and loading conditions as inputs to predict axial and shear stiffness [25,26], fatigue [27], and stress [6,28,29] as outputs. In a supervised learning approach, given ample training data, a well-trained neural network can make predictions as accurate as high-fidelity computational homogenization techniques [29–31]. Kim et al. [32] used an ANN to predict the stress–strain behavior of woven composites by incorporating time segmentation as an input feature.

To have an overview of recent developments, challenges, and potential solutions in the data-driven modeling of composite materials, an interested reader is referred to [33]. It should be noted that the highest level of accuracy that an ANN can achieve is at the level of its training data. Thus, the closer the data are to the physics of the problem, the more accurate the ANN will be. However, providing enough data to train a neural network is not always straightforward.

In response to address the data-hungry issue of neural networks, potential solutions have been presented. Several attempts have been made to create synthetic data resembling the characteristics of the training set, employing techniques such as variational autoencoders and generative adversarial networks [34]. The intention is to use generative networks to produce representative synthetic samples for training surrogate models. Yet, high-quality training data is still necessary to train these networks. Synthetic samples from generative networks that lack physical constraints cannot help the surrogate network learn with less data. Instead, they can mislead the learning process, which contradicts information theory's data processing inequality [35].

An alternative approach to creating surrogate models based on neural networks is to incorporate physics directly into the network structure. Physics-informed neural networks (PiNNs) incorporate knowledge of physical laws such as conservation principles, boundary conditions, and kinematic relations by adding new terms related to these constraints in the neural network's loss function. PiNNs have been used for the elasticity of porous materials [36,37] and to predict internal state variables for woven composites. It enables extrapolation beyond the training dataset [38]. However, their effectiveness is highly sensitive to the relative weighting of loss terms [39].

The physically recurrent neural network [40] is another physics-based framework that integrates constitutive material models directly into feed-forward neural networks by embedding microscale constitutive relations as modified hidden layers in a multi-layer perceptron. These physically informed recurrent blocks capture path dependencies like elasto-plasticity through internal variables. Physically recurrent neural networks have shown promise in reducing training data needs, enabling extrapolation to non-monotonic behaviors. However, its application to complex hierarchical structures remains an open challenge.

Transfer learning is a powerful strategy for addressing data acquisition challenges. In this approach, a neural network trained on a well-resourced source task is adapted for a related target task by additional training using high-fidelity data. The goal is to utilize the model's pre-learned features and patterns to enhance learning on the target task. This method leverages optimized initial weights and biases from the source task, improving convergence speed and efficiency on the target task. It has been successfully applied to integrate data in the elastic regime and to study the elasto-plastic behavior of short-fiber composites [41]. Transfer learning, however, requires a thorough study of the available datasets and their relative dependencies.

Another popular approach involves using data augmentation techniques tailored to the training data and expanding the dataset size for training without extra simulations. In [42], RVE rotation is an effective method for enriching data in short-fiber composite homogenization. Yet, the rotation technique as a way of augmentation is material-specific and requires additional feature inputs, such as fiber rotation during the network's training.

Strain and stress multivariate responses can be considered time series because they are recorded or observed sequentially over running time during physical or simulation experiments. Each measurement depends on its previous states due to the material's mechanical properties and load history, leading to time-correlated data points. The temporal dependency and the continuous recording of stress variables make stress responses potential candidates for time-series analysis as they evolve under varying strain-loading conditions. The present study combined transfer learning with a universally applicable augmentation technique utilizing linear interpolation of time-series data to enrich the information extracted from each data sample. Additionally, analyzing strain–stress data as a time series yields valuable statistics. Before training, these statistics enable us to determine if initiating from a specific source domain will benefit a target domain.

This study addresses the challenge of the data-intensive requirements for training a gated recurrent neural network (GRU) model tailored for the elasto-plastic behavior of woven composites. Unlike our previous work [25], which focused solely on low-fidelity data, the current study addresses incorporating high-fidelity data as efficiently as possible. The term *transfer* here encompasses both loading scenarios transfer learning and multi-fidelity data fusion approaches. The introduced model leverages multiple data sources. The input datasets include random walk and cyclic loading simulations across various constitutive material properties. The output datasets are generated through the Mori-Tanak MFH approach [6], chosen for its accessibility. MFH dataset is complemented by an efficient full-field *fast Fourier transform* (FFT) strategy [43,44]. The FFT homogenization method is crucial for incorporating the geometry and mesoscale structural details into the data generation process [26]. Additionally, a transfer learning approach is employed to fine-tune [45] the models across different scenarios, highlighting successful and unsuccessful fine-tuning strategies.

The structure of this paper is as follows: Section 2 reviews prior research on the MFH model, detailing the data generation process and the training of the low-fidelity path-dependent neural network model. This section covers aspects such as the sub-scale MFH modeling approach, the materials involved, and the design of computational experiments. In Section 3, we show the originality of the present study compared to earlier work and data augmentation strategies that utilize time-series statistics. The results obtained, along with comparisons between initial mean field and re-trained full-field GRU models, are discussed in Section 5. This section also includes a broader discussion and comparative analysis of the developed GRU models. Finally, Section 6 provides concluding remarks.

2. An overview of the low-fidelity RNN model

In a previous study [6], we introduced a machine-learning approach using recurrent neural networks to model the elasto-plastic behavior of woven composites. Data samples with elastic fibers and elasto-plastic matrix consisted of input features that included eight constitutive material properties and the fiber volume fraction (as described in Section 2.1), along with six independent components of the strain tensor. The output comprised a 6D pseudo-time history of homogenized stress components. The history-dependent RNN model was developed as a surrogate and utilized training data generated through a two-step mean-field homogenization approach implemented in Digimat-MF.

The homogenization approach was initially developed for modeling elasto-plastic composites by Doghri and Tinel [46] and implemented numerically in Digimat-MF. The Digimat-MF adopts a two-step procedure. In the first step, the RVE is divided into two pseudo-grains (PGs), where each PG represents a two-phase system consisting of the matrix and fibers with the same shape, aspect ratio, orientation, rate-independent constitutive model, and material parameters in either the warp or weft direction. The Hill-type incremental formulation is used for the homogenization of each pseudo-grain. The tangent stiffness matrix evolves based on the elasto-plastic deformation of the matrix, allowing for history-dependent stress–strain behavior. The overall RVE response is obtained by orientation averaging over both PGs in the second step. This approach contrasts with the one-step (direct) Mori–Tanaka method, which applies homogenization at the global level and may introduce asymmetries in stiffness when multiple orientations are present. Using the pseudo-grain approach and weighted averaging of the orientation distribution function [46,47], Digimat-MF provides a computationally efficient method of modeling woven composites [6].

The MFH simulations provided an extensive dataset critical for training neural networks to predict random and cyclic loads [6]. While a recent work [48] demonstrates that the Mori–Tanaka method can account for waviness and mesoscale geometry in woven fabrics, our current approach does not explicitly consider these features when homogenizing woven composite behavior. Accordingly, neither the specific weave pattern nor crimp (or undulations) is considered when using the MFH method to homogenize woven composite behavior. On the other hand, the MFH method efficiently averages stress in response to imposed strain across sub-scale constituents, reducing computational costs while still effectively approximating the composite behavior.

Building on the surrogate modeling foundation in [6], the current study aims to decrease the amount of high-fidelity full-field data needed, where the geometry of the meso-scale weave pattern is explicitly considered in the homogenization process. Thus, critical aspects of the previous framework are reiterated with modifications suited to the current objectives.

2.1. Constitutive material model

The matrix is modeled with an elasto-plastic behavior (J_2-plasticity with linear-exponential hardening [49]), while the reinforcement is considered elastic. As a simplification, the fiber properties are considered isotropic in this study. While this assumption is valid for many types of fibers, including glass fibers, it does not capture the transversely isotropic behavior of, e.g., carbon fibers. However, we emphasize that the current study is on method development whereby we argue that the simplifying assumption of fiber isotropy is reasonable and that the effect of transversely isotropy can be addressed in future studies.

The matrix yield function is given by

$$\Phi(\sigma,\kappa) = \sigma_{\mathrm{vM}} - (\sigma_{\mathrm{y}} + H\bar{\varepsilon^{\mathrm{p}}} + H_{\infty}\left(1 - e^{-m\bar{\varepsilon^{p}}}\right)) \leq 0, \tag{1}$$

where σ_{y} is the yield stress, and σ_{vM} is the von Mises stress. H_{∞} is referred to as the hardening modulus, H is the linear hardening modulus, $m > 0$ is the hardening exponent, and $\bar{\varepsilon}^{\mathrm{p}} \geq 0$ is the accumulated plastic strain. A detailed formulation of the material models can be found in [6]. The study utilizes Sobol sequence sampling to generate a diverse and uniform distribution of micro-structural constitutive material properties. This quasi-random sampling technique [50] facilitates the creation of a comprehensive design space for training the ANN models. The range of properties considered for the static features is similar to the previous study and is repeated in Table 1. Each data sample contains a set of microscale constitutive material properties and specific six-dimensional random walk or cyclic strain loading paths, as described below.

Table 1
Ranges of material parameters used to generate the simulation dataset and specific constitutive material properties for a sample carbon fiber/epoxy resin, used for comparing FFT homogenization with FE and MFH methods.

	Parameter	Range	Specific sample
Fiber	Young's modulus E_{F}	69–700 GPa	385.5 GPa
	Poisson's ratio ν_{F}	0.25–0.49	0.37
	Fiber volume fraction V_f	0.10–0.48	0.29
Matrix	Young's modulus E_m	2–10 GPa	6 GPa
	Poisson's ratio V_m	0.2–0.49	0.34
	Yield stress σ_{y}	31–66 MPa	48.5 MPa
	Linear hardening modulus H	1–200 MPa	100.5 MPa
	Hardening modulus H_{∞}	10–30 MPa	20 MPa
	Hardening exponent m	1–500	250.5

2.2. Load generator algorithm

A random walk algorithm is used to generate six-dimensional arbitrary input strain loading paths [51]. The algorithm adds the previous step with two white noises[1] generated from Gaussian noise with different scales. The first white noise serves as the main component on a larger scale (drifts); the other serves as the distractor on a tiny scale (noises). The algorithm enables the simulation of multi-axial stress–strain histories under random-walk loading conditions for both mean- and full-field simulations. To fine-tune the trained networks to cases with high sparsity in input space,[2] such as proportional pure shear cyclic loading, a second strategy for data generation is adopted. This strategy focused solely on cyclic loads in shear load cases where plasticity is significant in woven composites. Key factors in the cyclic loads include the peak strain value, the load ratio (fraction of the maximum positive strain to the minimum negative strain value), and the number of cycles.

Previously in [6], The initial dataset comprised random strain paths and was used for initial training and evaluation. Subsequently, transfer learning was employed to fine-tune the trained network on a second dataset featuring cyclic loading paths.

2.3. RNN model training

Among various neural network models available for time-series.[3] prediction, such as foundation models [53–55] and convolutional neural networks [56], GRUs and LSTMs are particularly prominent in material modeling [28,29,33,42]. Both GRUs [57] and LSTMs [58], which are specialized forms of RNNs, excel in handling sequential data, specifically time series. This makes them well-suited as surrogate models for predicting the path-dependent responses of materials.

Six strain tensor components (dynamic data) are combined in the current study with the eight intrinsic material properties and the fiber volume fraction (as listed in Table 1), repeating these static inputs across all time steps before feeding them to the network. This approach results in a total of 15 features being processed by the GRU layers at each time step. In prior research [6], various configurations of these networks were explored to optimize network architecture and hyperparameters. The present study focuses exclusively on GRU networks, given their faster training times compared to LSTMs and the previous findings that the accuracy differences between them are negligible.

[1] A process X_t is said to be *white noise* [52] if the following conditions hold: (1) Each X_t has zero mean and finite variance. (2) X_{t_1} and X_{t_2} are uncorrelated if $t_1 \neq t_2$.

[2] The input space includes six independent components of strain tensors and eight constitutive material properties and fiber volume fraction. By sparsity, we mean some components of input strain tensor are zero, like the pure in-plane shear case where there is only one component of strain, ϵ_{12}.

[3] Time-series refers to a collection of observations made sequentially, indexed in (abstract) time [52].

Refer to Section S.1 of supplementary materials for a comprehensive explanation of GRU units.

The output signals are stress tensor components. In the regression layer, the mean squared error loss function is calculated as follows

$$L = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} L_i, \quad \text{with} \quad L_i = \frac{1}{2N_T} \sum_{t=1}^{N_T} \sum_{c=1}^{6} (\hat{\sigma}_c^{(t)} - \sigma_c^{(t)})^2, \tag{2}$$

where $\hat{\sigma}_c^{(t)}$ and $\sigma_c^{(t)}$ are the predicted and desired normalized stress component[4] at time step t in a training sample, respectively. F is the number of output features, N_T is the data sequence length, and N_{batch} is the batch size.

Overfitting and convergence of neural networks are checked with the loss function during training. However, since the loss function is applied to normalized data, it is not the best metric for evaluating trained neural networks. This makes comparisons across different samples under varying load conditions quite difficult and non-intuitive. To provide a more intuitive measure of the quality of the network, error metrics based on the von Mises stress have been employed.

3. Enhanced high-fidelity model with transfer learning

The aim is to develop a GRU model as a surrogate for high-fidelity full-field simulations of woven composites. The composite is assumed to have a balanced weave with fixed geometrical features, and failure mechanisms are excluded from consideration. While the geometry of the meso-scale structure remains unchanged, the micro-structural constituent's properties change in each simulation. To harness the advantages of previously available datasets and the new high-fidelity datasets generated in this study, we initially trained our model on a dataset from MFH simulations of the previous study, then employed transfer learning to fine-tune the model with data from FFT simulations before ultimately validating the model against FFT simulations. The following sections describe the computational framework for high-fidelity FFT data generation, data preparation, augmentation, and transfer learning.

3.1. Computational homogenization of woven composite

Building on the methodology from Section 2, this research utilizes the FFT technique to model meso-scale woven composites. FFT is a highly efficient algorithm for computing the discrete Fourier transform and its inverse. FFT homogenization [43] is particularly advantageous for analyzing woven structures since it can efficiently handle their periodicity and complex geometric properties. By converting the spatial topology of a structure into the frequency domain, FFT facilitates the effective computation of stress distributions under various strain load conditions. Moreover, FFT methods can achieve results comparable to those obtained from FE analysis but usually require significantly less computational time and avoid the difficulties associated with FE mesh generation. A more detailed description of FFT homogenization for a boundary value problem is provided in Section S.3 of supplementary materials or can be found in [59,60].

In the data generation process, the RVE model utilized in the FE and FFT homogenization processes includes a balanced weave with 15 yarns per centimeter in the warp and weft directions, featuring a yarn spacing ratio of 0.1 and a 0.5 crimp factor. A fixed ellipsoidal yarn cross-section with a width of 0.05 mm and height of 0.5 mm is considered. The FFT simulations are conducted using 64 by 64 by 64 voxel (A voxel represents a discrete computational element in each dimension of the simulation grid, with each voxel having dimensions of 21 × 21 × 1.7 μm). Higher grid resolutions have been tested, but the improvement in accuracy is marginal compared to the increased computational cost, confirming the appropriateness of the chosen resolution. The FE simulations are conducted using 64 by 64 by 32 three-dimensional fully integrated with 8 nodes voxel elements after a convergence study that showed further refinement yielded negligible changes in the stress–strain response (see Fig. 1).

To illustrate a comparative analysis, Fig. 2 showcases the shear stress–strain responses of the RVE with a specific set of constitutive material properties (specific values in Table 1) subjected to a random and a three-cycle in-plain shear loading. Additionally, MFH results of the same material system are shown. The volume of reinforcing fibers in a composite RVE relative to the total volume of the RVE, the fiber volume fraction (V_f^{RVE}), is restricted to be constant and equal in MFH and FFT simulations by controlling the filament count and yarn density.

FFT homogenization closely aligns with voxel-based mesh FE homogenization outcomes. However, in terms of computational efficiency, during cyclic loading, as an example, the FE simulation required 7040 s, and FFT homogenization took 440 s, approximately. In comparison, MFH was completed in merely 30 s. The simulations are conducted on Digimat-FE software package [61] on a 16-core Processor computer with a *NVIDIA® GeForce RTX™ 4090* [62] system. MFH serves primarily as an approximation technique and fails to describe woven composite behavior accurately. This discrepancy is severe under random loading conditions. In Fig. 2, it is also evident that under cyclic loading, MFH deviations from FE and FFT results become more pronounced after the first cycle. Although FFT simulations require more time than MFH simulations, limiting the quantity of data generated, they provide a level of accuracy closer to computationally demanding FE simulations. Consequently, FFT simulations have been selected for the development of a high-fidelity dataset.

3.2. Data preparation

Thorough examination and cleaning of the dataset are essential in the data preparation phase for modeling strain and stress dynamics. Fig. 3 contains four plots of two stress components σ_{11} and σ_{12} from random and cyclic datasets, and illustrates the distribution of output labels. Notably, an outlier in the random loading dataset is detected since a stress component maximum value (here σ_{11}) exceeds the other values by more than 50%. As part of the data cleaning process, the decision is made to remove the sample containing this outlier.

Eliminating the outlier dramatically affects the statistical properties of the dataset. The removed simulation from the training dataset significantly shifts the mean distribution of the σ_{11} components, according to Fig. 3(b), resulting in a more balanced distribution centered around zero through time. This adjustment in the data distribution is crucial as it enhances the robustness of the training process by reducing the potential bias that outliers can introduce. The impact of this adjustment is discussed extensively in Section 5.2, where it is shown to facilitate more accurate and reliable learning outcomes.

In a woven composite, fibers are primarily aligned in specific directions, typically referred to as the warp and weft directions. It means the applied load is carried directly by the strong fibers, causing significantly higher stress values in normal directions, like σ_{11} compared to shear directions, like σ_{12} in Fig. 3. To further enhance data quality and model performance, the *robust scaler* [63] method is employed for data normalization. This scaling technique is particularly known more effective in datasets with outliers, as it removes the median and scales the data according to the Interquartile Range (IQR), which is the range between the 1st quartile (25th percentile) and the 3rd quartile (75th percentile). Centering and scaling are performed independently on each feature by only computing the median and IQR from the training set samples.

[4] The symmetric stress tensor in Voigt notation is represented as a six-component column vector as follows: $\tilde{\sigma} = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}, \sigma_{23})^T \equiv (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6)^T$.

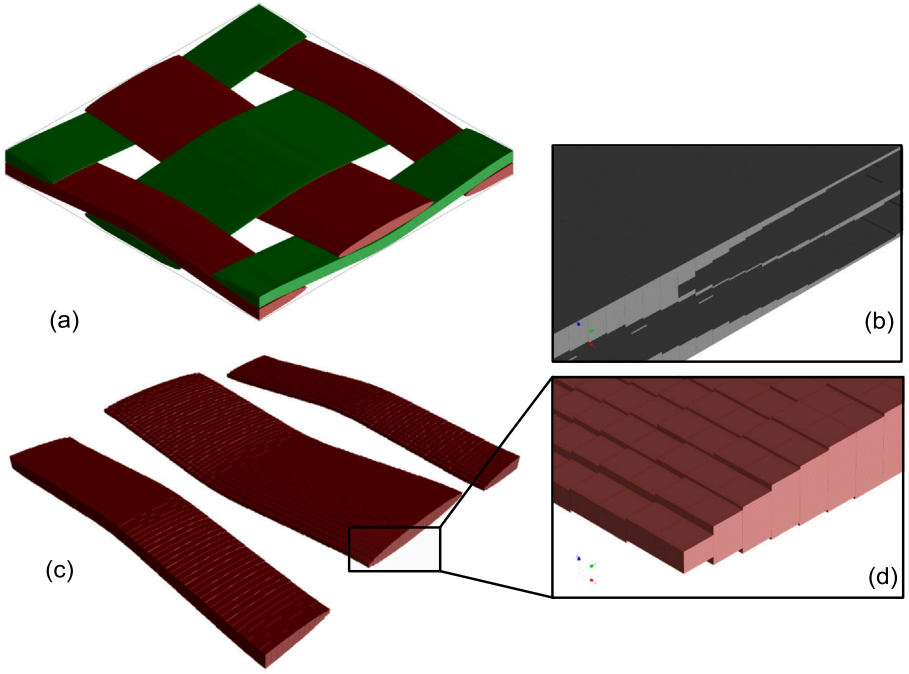


Fig. 1. (a) Mesoscale woven RVE generated in Digimat-FE and used in FE and FFT simulations, (b) detailed voxel-based FE mesh of the matrix, (c) the warp yarn, and (d) zoomed-in view of the yarn mesh.

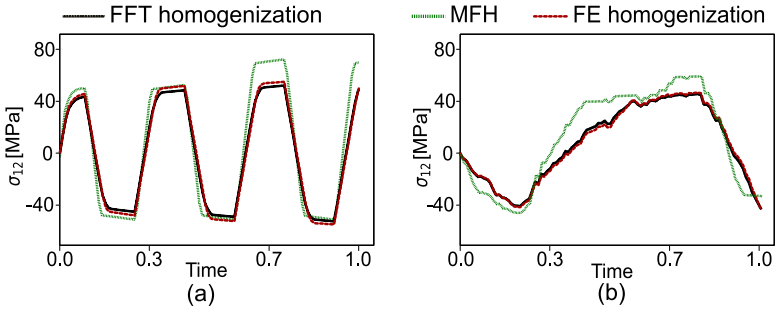


Fig. 2. FFT homogenization (black dashed lines) comparison with FE homogenization (solid red lines) and mean-field homogenization [6,46] (dashed green lines) for a sample (a) cyclic in-plain shear and (b) random walk load path.

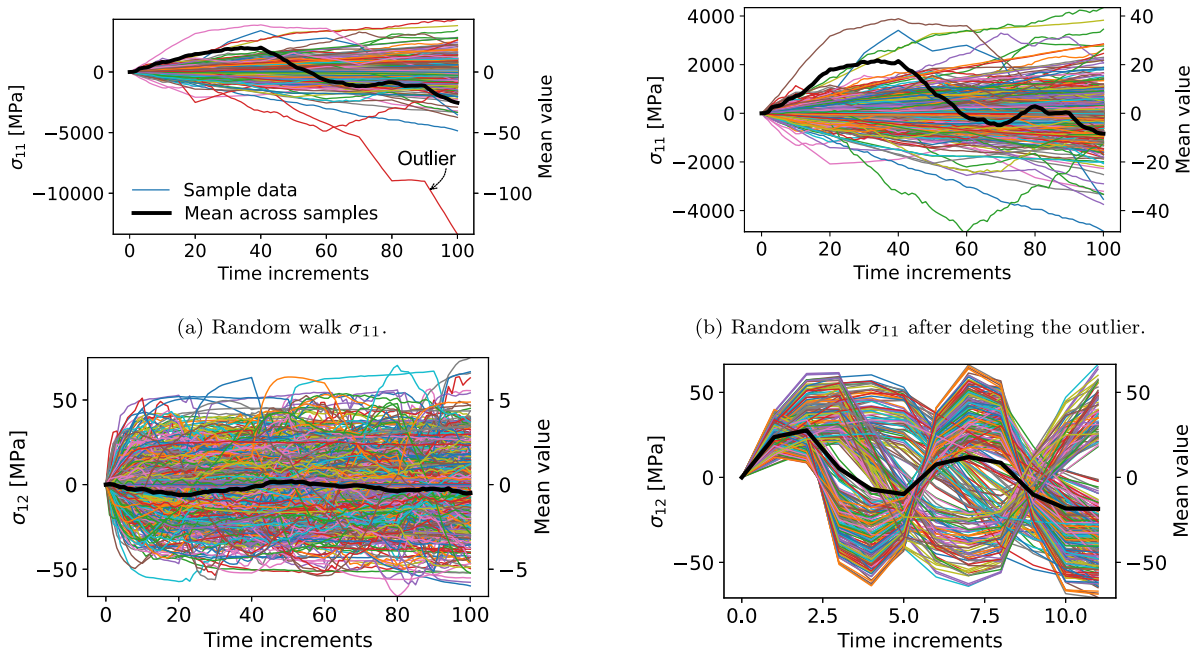
3.3. Transfer learning

Transfer learning emerges as a highly effective strategy for addressing the challenges of limited data availability in the field of material behavior prediction [6,41,64–66]. The source task here is a GRU network trained on a large dataset generated from mean-field simulations [6] characterized by random walk behavior. Having learned the underlying patterns and dynamics of MFH simulations, this network contains a wealth of knowledge about woven composite behavior under random walk loading and different microscale constitutive material properties. Transfer learning enables us to utilize the pre-trained GRU network effectively when shifting focus to a target task with limited data generated through high-fidelity full-field FFT simulations. By transferring the learned weights and features from the low-fidelity MFH model to the FFT model, it is hypothesized that the network will require less data to adjust and optimize its parameters for the new task.

Two common strategies in transfer learning are *freezing* and *fine-tuning* [67]. Freezing involves keeping the weights of specific layers of a pre-trained model fixed, while only updating the weights of the remaining layers during training on a new task. First layers of a network learn general features that are not specific to the source task, and following layers learn specific target tasks. The other approach, fine-tuning involves adjusting the weights of all the layers of a pre-trained model during the training process on the new task. This allows the model to adapt more thoroughly to the new data, often leading to better performance as it tweaks high-level and low-level features to suit the new task. In a similar case, [68], after experimenting with both freezing and fine-tuning strategies, it is found that fine-tuning the entire network's parameters is most suitable. Fig. 4 illustrates the transfer learning steps employed in this study.

The initial training set comprises an extensive collection of 28,000 samples from random non-proportional loading scenarios. Then, the network is trained on 1,000 high-fidelity samples from full-field FFT simulations under random loading conditions. The final training phase fine-tunes the model further by focusing on specific and challenging load cases, including pure in-plane and out-of-plane shear cyclic loading, each containing 300 high-fidelity FFT samples. This crucial step exposes the model to extreme conditions which often results in significant plastic deformation, significantly enhancing its predictive accuracy. Ultimately, the trained GRU model is evaluated against FFT random and cyclic simulations, which serve as the benchmark for this study.

Sparse datasets are characterized by numerous missing, zero, non-informative values. When a model encounters sparse inputs during



(a) Random walk σ_{11}.

(b) Random walk σ_{11} after deleting the outlier.

(c) Random walk σ_{12}.

(d) Cyclic σ_{12}.

Fig. 3. Output results of training data high-fidelity FFT simulations: tensile σ_{11} and pure in-plain shear σ_{12} stress components. (a) Random walk data for σ_{11}; (b) Random walk data for σ_{11} after outlier removal; (c) Random walk data for σ_{12}; (d) Cyclic pure in-plane shear data for σ_{12}. The cyclic dataset in (d) was generated with a fixed strain amplitude, number of cycles, and load ratio to explore cyclic deformation responses. The mean stress for each dataset is indicated by the thick black line.

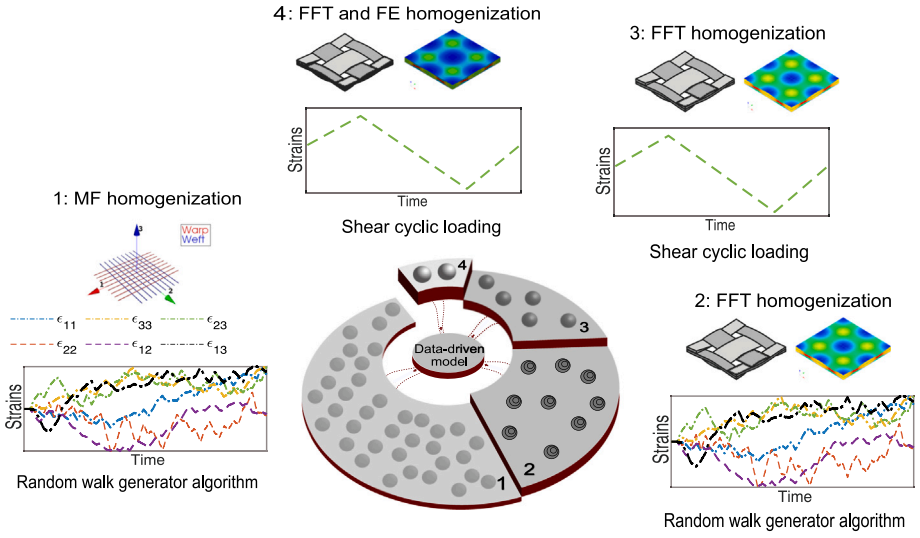
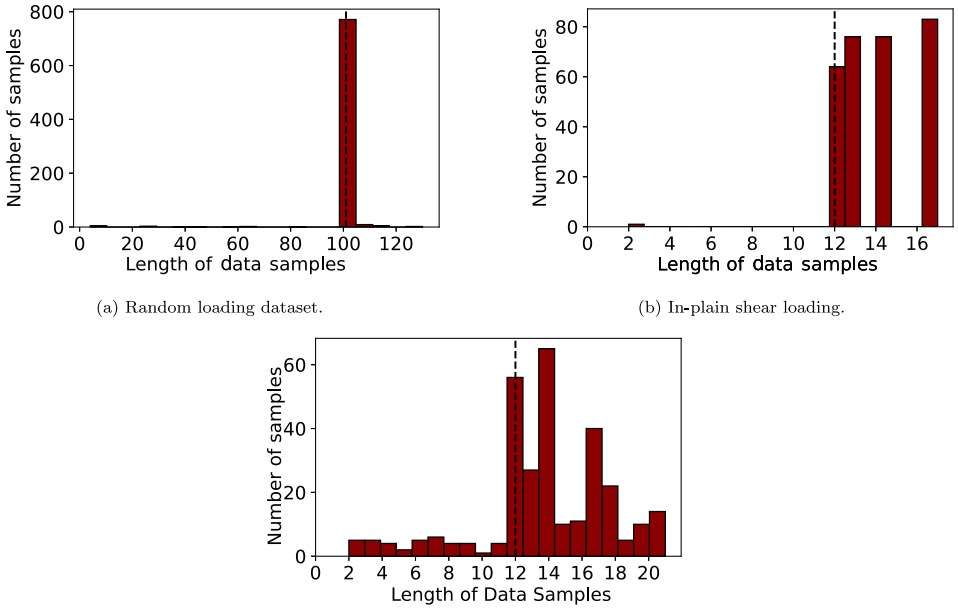


Fig. 4. An illustration of the step-wise transfer learning process utilized for data fusion in woven composite behavior modeling, with data fidelity represented through color contours. Step (1) involves training the GRU model on a random-walk MFH dataset. In Step (2), the model, initialized from the previous training, is further trained using FFT simulation data. Step (3) focuses on fine-tuning the model using a pure shear FFT cyclic dataset. Finally, Step (4) evaluates the model's performance against FFT and FE simulations.

testing, it experiences a distribution shift [69], as the input patterns differ from those seen during training. During the training phase, the model is typically exposed to dense input patterns, where most or all features are active, allowing it to learn from a rich and complete dataset. However, in the testing phase, the model may encounter sparse inputs with only a few active features. This results in a mismatch in the input distribution, potentially degrading the model's generalization ability, as it has not been trained to handle such sparse patterns effectively [70]. Consequently, the model may struggle to make accurate predictions, leading to increased error and reduced reliability in real-world applications. This problem arises in uniaxial stress loadings like pure-shear loading scenarios, where only one component is non-zero, resulting in sparse data. This sparsity complicates the learning process, hindering the network's ability to find meaningful patterns.

3.4. Data augmentation

Data augmentation aims to create synthetic datasets processed according to desired invariances that cover unexplored input space while the output targets are the same as those in the original datasets [56]. Although many methods exist for data augmentation in computer vision, more attention is needed for comprehensive and systematic data augmentation of time-series data in common tasks, including time-series forecasting, anomaly detection, and classification. Data augmentation



(a) Random loading dataset.

(b) In-plain shear loading.

(c) Out-of-plain shear loading.

Fig. 5. Frequency of happening (distribution) variable-length time-series data resulted from FFT simulations, (a) for random loading dataset, (b) in-plain shear loading dataset, and (c) out-of-plain shear loading dataset.

becomes increasingly complex in the context of multivariate time series, which require careful consideration of the intricate dynamics among variables over time.

The woven composite homogenization results in stress multivariate responses. Each predicted stress value depends on its previous states due to the material's mechanical properties and load history, leading to time-correlated data points. Window cropping or slicing, window wrapping, flipping, and noise injection are among the usual data augmentation in time series [71]. However, the characteristics of multivariate stresses limit the application of conventional data augmentation methods. Composite materials' behaviors differ under tension and compression in the presence of plasticity. Loading begins from a zero strain value, and the subsequent behavior of the material is strongly linked to the accumulation of internal variables from the onset of loading. Thus, window-based methods for the augmentation of data are complicated by such dependencies.

In this study, the term *length* is used when referring to the sequential nature of time-series data, consistent with common terminology in time-series modeling [52]. Meanwhile, *number of time increments* refers to the computational steps required to converge the simulation models during the data generation. This distinction helps clarify the relationship between the representation of sequential data and the underlying computational process.

Data obtained from FFT simulations vary in length scales due to the variations in the number of required time increments. Fig. 5 displays the number of simulations verses the time increments taken for convergence. Notably, only a few samples have fewer than 101 time increments (see Fig. 5(a)) for the random case dataset; these can be omitted from the training process. Samples with more than 101 time increments are included in the training process after shortening. The result is a uniform, consistent dataset for training the GRU network. Fig. 5(b) shows that pure in-plane shear cyclic simulations usually converge after 12 or more time increments. Pure out-of-plane shear cycles with a wide range of time-series lengths are even more challenging data as FFT simulations converge at different number of time increments. Training a GRU network to predict path-dependent behavior is particularly challenging for short series as each data sample contains limited information.

Linear interpolation offers a practical solution for data augmentation and uniformity. The suggestion is to equalize the length for each sample by linear interpolation to increase the number of time increments to a consistent value across all datasets. It is important to analyze the impact of interpolation on the underlying data dynamics before applying the model. The following statistical analysis investigates how the structure of the time-series data samples can affect the learning process. Statistical analysis of data can be used to evaluate whether or not two time-series datasets from different domains will help before applying transfer learning or if they will result in negative transfers (discussed in Section 5.1). This potentially reduces the computational effort for re-training a network for a particular target task.

4. Data analysis

An autocorrelation function (ACF) is a standard tool for assessing the relationships between variables in time-series data. ACF can evaluate the impact of linear interpolation on time-series stress data predictions. Detailed statistical analyses and theoretical discussions on this topic are available in Section S.2 of supplementary materials, which focuses on time-series analysis.

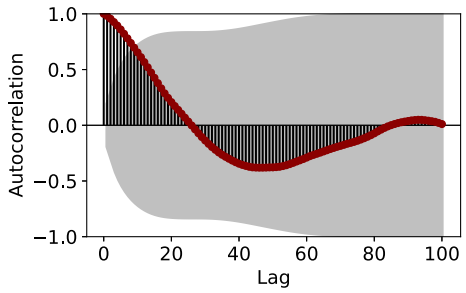
The ACF for stress data is briefly reformulated here. The sample autocovariance for stress data based on two different time increments t_1 and t_2 is defined as

$$\hat{\gamma}_{\sigma^{(t_1)}\sigma^{(t_2)}} = \frac{1}{m}\sum_{i=1}^{m}(\sigma_{(i)}^{(t_1)} - \bar{\sigma}^{(t_1)})(\sigma_{(i)}^{(t_2)} - \bar{\sigma}^{(t_2)}), \tag{3}$$

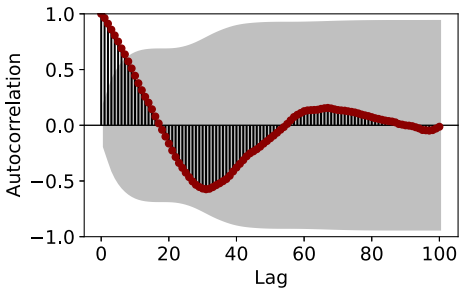
where m is the number of simulations in the training data and $\bar{\sigma}$ is the sample mean of a stress component for each time increment among all the training data. The autocorrelation for stress data is defined as

$$r_{\sigma^{(t_1)}\sigma^{(t_2)}} = \frac{\hat{\gamma}_{\sigma^{(t_1)}\sigma^{(t_2)}}}{(m-1)s_{\sigma^{(t_1)}}s_{\sigma^{(t_2)}}}, \tag{4}$$

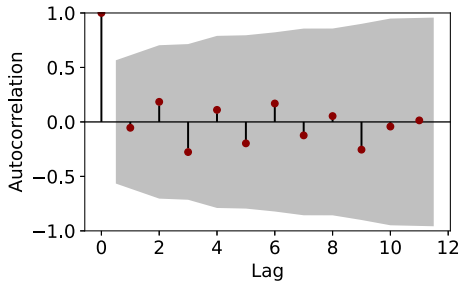
where $\hat{\gamma}_{\sigma^{(t_1)}\sigma^{(t_2)}}$ is the stress autocovariance according to Eq. (3), and $s_{\sigma^{(t)}}$ is the standard deviation of the stress samples. Data is collected from FFT simulations with equidistant stress values up to the convergence time as the maximum applied strain is reached. Consequently, we can define the lag parameter k as the unit temporal distance between stress (or strain) values at time t and $t-k$, denoted as $\sigma^{(t)}$ and $\sigma^{(t-k)}$ respectively. Section 4.1 analyzes the temporal stress responses and explores the impact of data augmentation.



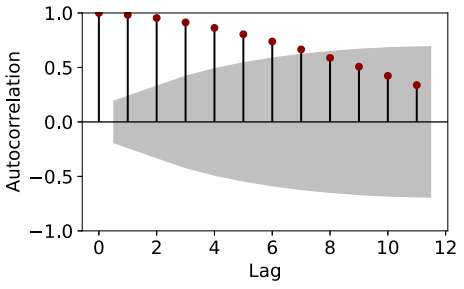
(a) ACF for σ_{11} random loads.



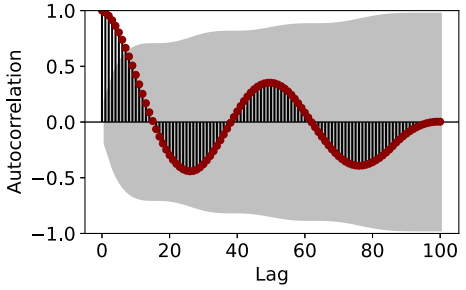
(b) ACF for σ_{12} FFT random loads.



(c) ACF for σ_{12} FFT cyclic loads without interpolation (for k up to the original time-series length).



(d) ACF for σ_{12} FFT cyclic loads with interpolation (showing only up to k=12).



(e) ACF for σ_{12} FFT cyclic loads with interpolation (for k up to the new time-series length).

Fig. 6. Auto-correlation functions for different stress components and interpolation methods. The shaded areas are the standard deviation of ACF among all the samples in the training datasets.

4.1. Analyzing temporal correlations in stress time-series

To systematically analyze correlation, graphical representations of the ACF versus lags display varying behaviors in different datasets of loading conditions. As an example, for σ_{11} under random loading in Fig. 6(a) and σ_{12} in Fig. 6(b), the ACF graphs show distinct correlations between time increments, highlighting substantial temporal dependencies. Since σ_t and σ_{t+1} values are happening after each other on the time axis, we see a noticeable strong positive $r_{\sigma_t\sigma_{t+1}}$. This strong positive autocorrelation is dampened as lags increase. For example, there is a very slight autocorrelation between σ_t and σ_{t+20}. This indicates that when the network attempts to capture evolving stress in a sample at each time interval, the stress value at the next step strongly correlates with the current state. This is similar to plasticity theory, where the current state is calculated based on the proceeding state and the material history. However, the correlation analysis at different sample lags suggests that the stress data by itself has additional information from more distant time increments, extending up to ten increments from the current state. This demonstrates that the network must have the capability to integrate longer-term dependencies within the data.

The AFC plot according to FFT samples subjected to pure in-plain shear loading (Fig. 3(d)) is plotted in Fig. 6(c). Also, ACF tends to be near zero at lag $k \neq 0$. By applying linear interpolation as a way of augmentation by increasing the number of time increments from 12 to 101 (equal to the number of time increments in the random dataset), a significant change is observed in the ACF plots; what was previously a non-correlated time series in shear loading conditions now displays correlations between interpolated time steps at Fig. 6(d). This emergence of meaningful correlations through interpolation suggests that the augmentation method does more than just align data lengths — it can alter the inherent statistical properties of the time series, as it can be seen in Fig. 6(e). The effect of interpolation on GRU training as a mechanism for augmentation will be discussed in Section 5.

The sample mean stress for each component is defined as the mean stress among all samples at each increment. The thicker black lines in Fig. 3 for two example components show the sample means which are almost constant and close to zero, and stress datasets do not follow a particular trend. Nevertheless, the stress responses cannot be considered *stationary* and analyzed by conventional stationary time-series methods, like ARMA [52]. Because the autocorrelation does not essentially depend on the distance $t_2 - t_1$, it can be infer from the large standard deviation of ACF for different lags in Fig. 6.

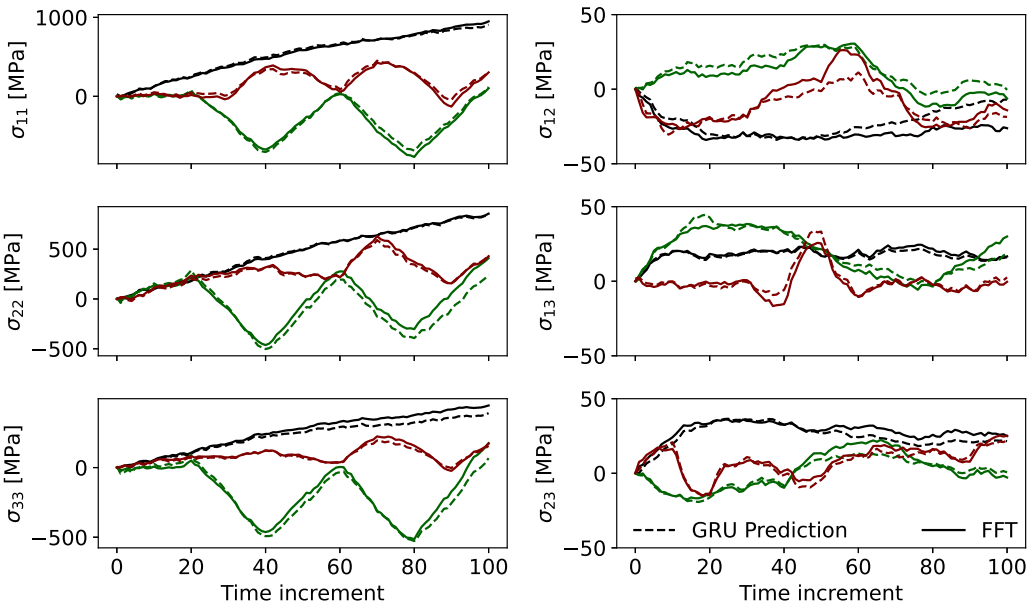


Fig. 7. Fine-tuned high-fidelity model predictions on three randomly selected full-field random walk path samples from the unseen dataset. The solid lines are the target stress results from the FFT simulations, while the dashed lines are the network predictions.

5. Results and discussion

The GRU network is implemented using Python and the deep learning library PyTorch. *NVIDIA® GeForce RTX™ 4090* [62] GPU is utilized to train various networks. Generated datasets are divided randomly into 80% training, 10% validation, and 10% testing sets. During training, the GRU network processes the training set through multiple epochs, with data shuffled in each epoch. The validation set is crucial for adjusting hyperparameters like batch size, learning rate, and regularization parameters to enhance model performance. The test set, comprising unseen data, is used to evaluate the network's final performance.

Section 5.1 discusses the process of training a network from scratch using random high-fidelity data and optimizing hyperparameters. Section 5.2 evaluates the application of transfer learning from previously developed networks using mean-field low-fidelity data to initialize and train a network with high-fidelity FFT random loading data.

5.1. Training of the surrogate GRU model

The training process uses ADAM optimizer [72]. Among the hyperparameters, the learning rate is tuned for each network training step automatically using Spotpython [73]. For simplicity, a fixed 50% dropout after the first GRU layer and the minibatch size 32 is used. A piece-wise learning rate decay strategy is employed, reducing the learning rate by 10% every ten epochs. Following our previous study [6], the GRU network architecture has three layers, each with 512 GRU units and one fully connected layer.

In order to prevent over-fitting, early stopping is used instead of a fixed number of epochs. Training stops when the models' performance on the validation set plateaus (a patient equal to 10 is used) while the loss on the training set continues to decrease. The plateau region on the validation set indicates convergence in the training. Plots of the loss functions are presented in Figure S.2 of supplementary materials.

5.2. Prediction on random walk and cyclic paths

Targeting the random walk FFT test set, the most promising fine-tuning strategy is obtained according to following. First, train the model on the MFH random-walk load path. Second, fine-tuning it with FFT random walk. With the transferred network, stress values can be predicted from sparse feature samples throughout the loading increments.

The results on the test sets show effective learning occurred using the transfer learning technique and utilizing the knowledge gained from MFH. Note that the performance after training exclusively on MFH data provides a reasonable baseline. However, this approach results in an error on the FFT random test set that is approximately two times larger than that of a model trained from scratch on FFT samples. Therefore, fine-tuning with high-fidelity data is necessary to further improve accuracy. Additionally, different optimizer settings are tested, but the results reported in the paper reflect the best outcomes to keep the study concise. Fig. 7 shows successful predictions of the GRU network for all six stress components on three randomly selected samples from the test set.

The transition to conventional cyclic paths presents a significant challenge, particularly when applying the trained network to cyclic load tasks. The input features for cyclic loading cases are notably sparse, involving only one (in pure shear conditions) active sequential feature alongside static features representing the microstructure and plasticity is more pronounced compared to random loads. Therefore, ANN models often struggle to predict stresses from these sparse inputs accurately, or predict nonzero fictitious values for the inactive components. Although one can train a network specifically for the target stress component, the network will lack generality. To address this challenge, four scenarios are examined: (1) training a GRU network from scratch using the original cyclic FFT samples (presented in Fig. 3(d)), with the corresponding ACF shown in Fig. 6(c); (2) training a GRU network from scratch on the augmented cyclic FFT dataset (with the ACF depicted in Fig. 6(e)); (3) employing transfer learning to acquire prior knowledge from a dataset comprising random MFH and FFT data, followed by fine-tuning with the original cyclic FFT; and (4) utilizing transfer learning with the augmented cyclic FFT dataset.

Firstly, the impact of augmentation is examined through the first two scenarios. As demonstrated in Fig. 8, the FFT cyclic homogenization is accurately predicted when using the linear interpolation augmentation framework. However, this accuracy diminishes when training is conducted with data of the original length.

Secondly, the GRU network, initially trained on the random MFH dataset and fine-tuned on random FFT data, is further fine-tuned using the cyclic FFT samples. The bars on the right side in Fig. 9(a) illustrate that utilizing augmented cyclic FFT significantly reduced the RMSE

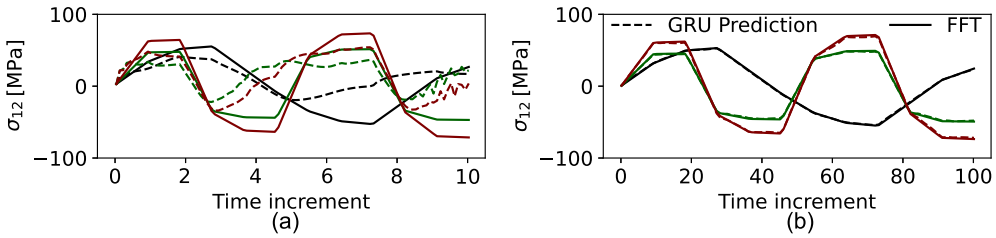


Fig. 8. GRU model prediction on three σ_{12} unseen full-field sample. (a) The original data length was used for training from scratch, and (b) an augmented dataset was used for training.

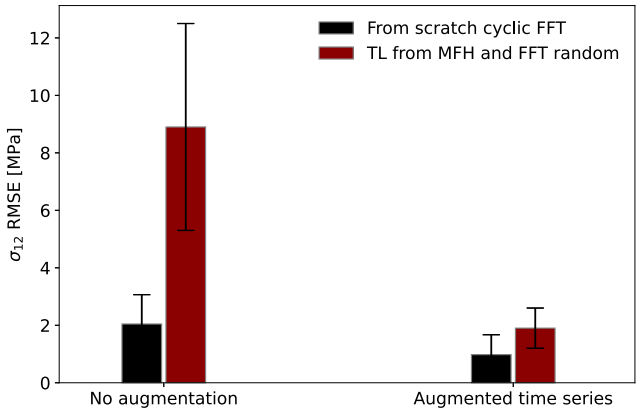
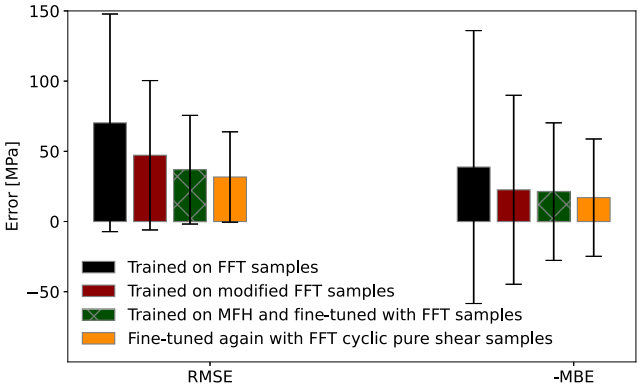
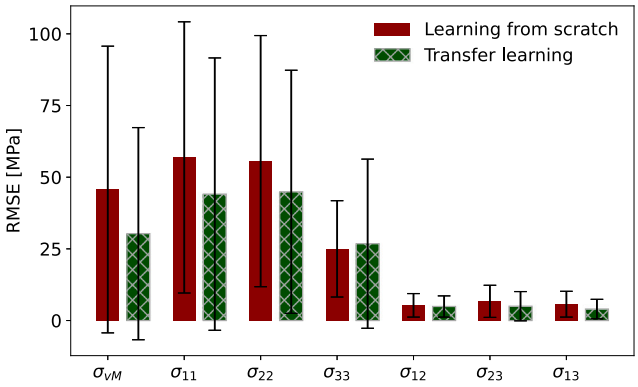


Fig. 9. GRU model prediction on σ_{12} unseen cyclic full-field dataset. Augmentation reduces the RMSE, but a negative transfer effect is observed. TL stands for transfer learning.



(a) Error measures based on von Mises stress



(b) RMSE based on different FFT samples stress components

Fig. 10. Network performance against unseen high-fidelity random walk FFT samples with and without transfer learning. (a) RMSE and MBE are based on von Mises stress, (b) RMSE is based on individual stress components, and the von Mises stress.

error [6] on the test set of cyclic loads. However, the error increased in original and augmented cases compared to training from scratch. Despite initializing the model from scratch with different random seeds, the transfer learning model consistently performed worse on unseen cyclic samples compared to the model trained from scratch. In such cases where the negative transfer [74] is pronounced, a conventional network, trained on the limited labeled data specific to the target application, often outperforms a model that incorporates both the limited labeled target data and the source data. This situation tends to manifest when the source domain (six components of stress result from random paths) has little resemblance to the target domain (two components of shear stress result from cyclic paths).

The network performance is evaluated against random FFT samples despite the negative transfer occurrence when testing against cyclic FFT samples. The error bar chart depicted in Fig. 10(a) illustrates the progressive improvement of the developed GRU model when tested against 100 FFT random walk samples. As mentioned in Section 3.2, the initial modification of the training set, which involved removing a single outlier, resulted in significant reductions in error measures: 30% in mean and 32% in standard deviation. Additionally, employing the MFH dataset for pre-training the GRU network, instead of initializing training from scratch, led to a decrease of 21% in the mean error and 40% in the standard deviation. Introducing the network to high-fidelity specific cyclic loads towards the end of the training process did not lead to negative transfer; instead, it enhanced performance by 14% on random walk paths. The MBE measure [6] is also progressively reduced through transfer learning, ensuring a more reliable and unbiased model.

The poor transfer observed in Fig. 9 can be attributed to differences in the datasets, which arise from variations in loading scenarios. These differences pose challenges in achieving an optimal solution during fine-tuning. Specifically, when the network is initially trained on cyclic loads with a single active loading feature and tested on the same data type, it effectively learns the relationship between the active feature and the output. However, when trained on random loads, the network must capture dependencies across all six loading features. Testing this model on cyclic loads introduces a rare scenario characterized by extreme plasticity and sparsity, resulting in a distribution shift in the input space. Consequently, transfer learning on cyclic loads leads to negative transfer, where the model performs worse than one trained on cyclic loads from scratch. In contrast, when the network is first trained on random loads and subsequently fine-tuned on cyclic loads, it demonstrates improved performance on random loads (Fig. 10). This suggests that prior exposure to random loads enables the model to develop a more robust understanding of plasticity mechanics, which enhances its adaptability.

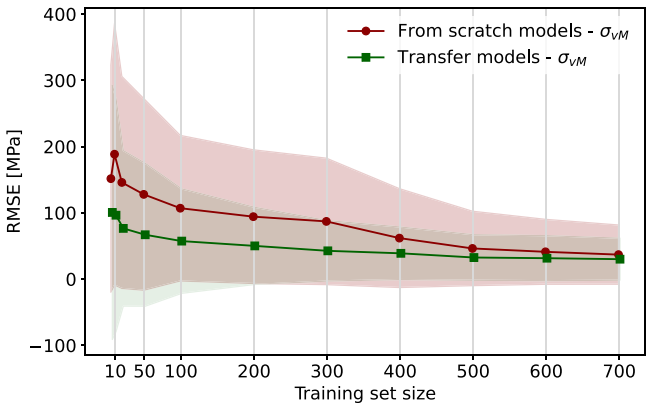


Fig. 11. Mean RMSE for von Mises stress (σ_{vM}) across different training set sizes. The plot compares the performance of "from scratch" and "transfer" models, with shadows indicating standard deviation. The transfer learning model was pre-trained on random loads of MFH data and fine-tuned with random FFT data. Details on each component can be found in Figure S.3.

For a more detailed evaluation, the performance of the GRU model trained on modified FFT samples is compared with that of the fine-tuned network, as shown in Fig. 10(b). This comparison highlights a general reduction in prediction errors across various stress components, demonstrating improvements beyond just the von Mises stress. However, an exception is noted in the case of the σ_{33} component, where the model experienced a small negative transfer, resulting in decreased accuracy for this specific stress component. It underscores how transfer learning has complex dynamics, in which improvements in some areas might mislead others.

Fig. 11 highlights the performance of the models in predicting the von Mises stress across different training set sizes when evaluated against the random load high-fidelity test set. The transfer learning model, which is pre-trained on the random loads MFH dataset and fine-tuned with the FFT dataset, demonstrates superior accuracy compared to the models trained from scratch. For larger training sets (e.g., 600 or 700 samples), the difference between the two learning methods is minimal in both mean RMSE and standard deviation. Conversely, with smaller training sets (300 and below in Fig. 11), which are common in practice, both the mean error and the standard deviation across test samples are significantly reduced, indicating robustness against individual sample variations.

6. Conclusions

To tackle the intricacies of time-consuming woven composite homogenization analysis, GRU surrogate models are proposed. However, these models are extremely data-intensive, necessitating large datasets for effective training. This study explores the application of transfer learning in the supervised learning of nonlinear behaviors in woven composites, particularly under conditions where matrix plasticity exists. We have utilized the predictive capabilities of GRUs trained on FFT homogenization data to serve as an accurate surrogate model for mesoscale homogenization. To overcome data limitations, we initialized the GRU network using an existing dataset from MFH, which contains three times the data points compared to full-field datasets.

We considered four datasets. The first used a random walk loading path with fast MFH simulation for diverse strain and stress trajectories. The second also used a random walk but with FFT homogenization. The third and fourth involved conventional cyclic strain paths, presenting challenges with fewer data samples. After training on random loads and fine-tuning with cyclic loads, the network can accurately predict stress under standard cyclic conditions.

This research also highlights the importance of temporal correlations in datasets, which represent the path-dependent physical characteristics of materials. These correlations are essential for training time-series models such as GRUs, enabling them to learn path-dependent responses accurately. Autocorrelation analysis is also a method for investigating the impact of linear interpolation on data augmentation. It examines whether the augmentation influences the dataset's correlations, thereby affecting how it is learned during the training process.

In conclusion, this study demonstrates that transfer learning can effectively integrate readily available, albeit less accurate, data with high-fidelity but sparse and scarce data in modeling history-dependent material behavior. However, it also highlights the risk of negative transfer when the source and target data domains lack sufficient relatedness. This finding motivates future research to explore alternative methods to transfer learning for such applications. This proposed approach facilitates the use of limited high-fidelity data to train representative surrogate models.

Further work is needed to enhance model reliability, particularly in handling outliers and extreme cases where prediction errors are highest. Future research will focus on identifying composite configurations where the model performs poorly and investigating their underlying causes. Additionally, to improve the practical applicability of this approach, we propose implementing uncertainty quantification techniques, such as Monte Carlo dropout and Bayesian deep learning, to provide confidence intervals for predictions. These enhancements will equip engineers with clear reliability metrics, allowing them to determine whether the model's outputs are trustworthy for specific applications.

CRediT authorship contribution statement

Ehsan Ghane: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Martin Fagerström:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization. **Mohsen Mirkhalaf:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, 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.

Acknowledgments

Ehsan Ghane and Mohsen Mirkhalaf gratefully acknowledge financial support from the Swedish Research Council (VR grant: 2019-04715) and the University of Gothenburg, Sweden. Martin Fagerström is thankful for the support through Vinnova's strategic innovation programme LIGHTer, in particular via the project LIGHTer Academy Phase 4 (grant no. 2023-01937).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compscitech.2025.111163.

Data availability

The neural network codes and data generated for this research are available on a GitHub repository (https://github.com/EsanGhaneh/Multi-fidelity-Data-Fusion-for-In-elastic-Woven-Composites.git).

References

[1] Philippe Boisse, Nahiene Hamila, Stepan V Lomov, Naim Naouar, Emmanuelle Vidal-Sallé, Mechanics of Woven Fabrics, Elsevier, 2023.

[2] Carolyn Oddy, Magnus Ekh, Martin Fagerström, Macroscale modelling of 3D-woven composites: Elasto-plasticity and progressive damage, Int. J. Solids Struct. 250 (2022) 111696.

[3] Omar Bacarreza, Pihua Wen, M.H. Aliabadi, Micromechanical modelling of textile composites, in: Woven Composites, World Scientific, 2015, pp. 1–74.

[4] Weipeng Liu, Changjun Yin, Multiscale structural analysis for plain weave composite plates using asymptotic homogenization method, Mech. Adv. Mater. Struct. (2022) 1–16.

[5] Ling Wu, Laurent Adam, Ludovic Noels, Micro-mechanics and data-driven based reduced order models for multi-scale analyses of woven composites, Compos. Struct. 270 (2021) 114058.

[6] Ehsan Ghane, Martin Fagerström, Mohsen Mirkhalaf, Recurrent neural networks and transfer learning for predicting elasto-plasticity in woven composites, Eur. J. Mech. A Solids (2024) 105378.

[7] M.M. Shokrieh, R. Ghasemi, R. Mosalmani, A general micromechanical model to predict elastic and strength properties of balanced plain weave fabric composites, J. Compos. Mater. 51 (20) (2017) 2863–2878.

[8] Un-Il Ri, Jun-Hyok Ri, Hyon-Sik Hong, Multiscale computational homogenization for woven composites using the cluster-based nonuniform transformation field analysis, Arch. Appl. Mech. 93 (5) (2023) 1797–1808.

[9] Xinxing Han, Jiaying Gao, Mark Fleming, Chenghai Xu, Weihua Xie, Songhe Meng, Wing Kam Liu, Efficient multiscale modeling for woven composites based on self-consistent clustering analysis, Comput. Methods Appl. Mech. Engrg. 364 (2020) 112929.

[10] Tong-Rui Liu, Yang Yang, Omar R Bacarreza, Shaoqiang Tang, MH Aliabadi, An extended full field self-consistent cluster analysis framework for woven composite, Int. J. Solids Struct. 281 (2023) 112407.

[11] Tomo Takeda, Satoru Takano, Yasuhide Shindo, Fumio Narita, Deformation and progressive failure behavior of woven-fabric-reinforced glass/epoxy composite laminates under tensile loading at cryogenic temperatures, Compos. Sci. Technol. 65 (11–12) (2005) 1691–1702.

[12] Hao Huang, Zhongde Shan, Jianhua Liu, Zitong Guo, Zheng Sun, Dong Wang, Chenchen Tan, Juncheng Luo, Damage failure analysis of flexible-oriented three dimensional woven composites with void defects: An experimental and numerical study, Compos. Part A: Appl. Sci. Manuf. (2024) 108106.

[13] Dmitry S. Ivanov, Stepan V. Lomov, Modeling of 2D and 3D woven composites, in: Polymer Composites in the Aerospace Industry, Elsevier, 2020, pp. 23–57.

[14] Y.H. Chen, M.H. Aliabadi, Meshfree-based micromechanical modelling of twill woven composites, Compos. Part B: Eng. 197 (2020) 108190.

[15] M. Kharratzadeh, M.M. Shokrieh, M. Salamat-Talab, Effect of interface fiber angle on the mode I delamination growth of plain woven glass fiber-reinforced composites, Theor. Appl. Fract. Mech. 98 (2018) 1–12.

[16] A.E. Bogdanovich, Multi-scale progressive failure modeling: From nano-structured carbon fibers to textile composites, in: The Structural Integrity of Carbon Fiber Composites: Fifty Years of Progress and Achievement of the Science, Development, and Applications, Springer, 2017, pp. 767–793.

[17] Ehsan Ghane, Bijan Mohammadi, Entropy-damage mechanics for the failure investigation of plain weave fabric composites, Compos. Struct. 250 (2020) 112493.

[18] Jie Sun, Ang Wang, Hongneng Cai, Xuecheng Han, Zhiyuan Wei, Yaping Huang, Multiscale progressive damage model for plain woven composites, Int. J. Mech. Sci. 259 (2023) 108604.

[19] Xiao-Yi Zhou, Sheng-Yu Qian, Neng-Wei Wang, Wen Xiong, Wen-Qing Wu, A review on stochastic multiscale analysis for FRP composite structures, Compos. Struct. 284 (2022) 115132.

[20] Bing Wang, Guodong Fang, Hongyue Wang, Jun Liang, Fuhong Dai, Songhe Meng, Uncertainty modelling and multiscale simulation of woven composite twisted structure, Compos. Sci. Technol. 217 (2022) 109118.

[21] Andy Vanaerschot, Brian N. Cox, Stepan V. Lomov, Dirk Vandepitte, Stochastic multi-scale modelling of textile composites based on internal geometry variability, Comput. Struct. 122 (2013) 55–64.

[22] Ludovic Noels, Toward stochastic multiscale methods in continuum solid mechanics, Adv. Appl. Mech. 55 (2022) 1–254.

[23] Ziheng Tang, Licheng Guo, Tao Zheng, Zhixing Li, Ruijian Sun, Kai Huang, A combined machine learning and numerical approach for evaluating the uncertainty of 3D angle-interlock woven composites, Compos. Struct. 294 (2022) 115726.

[24] Jan Niklas Fuhg, Govinda Anantha Padmanabha, Nikolaos Bouklas, Bahador Bahmani, WaiChing Sun, Nikolaos N. Vlassis, Moritz Flaschel, Pietro Carrara, Laura De Lorenzis, A review on data-driven constitutive laws for solids, 2024, arXiv preprint arXiv:2405.03658.

[25] E. Ghane, M. Fagerström, S.M. Mirkhalaf, A multiscale deep learning model for elastic properties of woven composites, Int. J. Solids Struct. 282 (2023) 112452.

[26] Menglei Li, Bing Wang, Jiqiang Hu, Gao Li, Peng Ding, Chunming Ji, Artificial neural network-based homogenization model for predicting multiscale thermo-mechanical properties of woven composites, Int. J. Solids Struct. 301 (2024) 112965.

[27] Jie Chen, Yongming Liu, Fatigue modeling using neural networks: A comprehensive review, Fatigue Fract. Eng. Mater. Struct. 45 (4) (2022) 945–979.

[28] Ling Wu, Nanda Gopala Kilingar, Ludovic Noels, et al., A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths, Comput. Methods Appl. Mech. Engrg. 369 (2020) 113234.

[29] Shiguang Deng, Shirin Hosseinmardi, Libo Wang, Diran Apelian, Ramin Bostanabad, Data-driven physics-constrained recurrent neural networks for multiscale damage modeling of metallic alloys with process-induced porosity, Comput. Mech. (2024) 1–31.

[30] Ling Wu, Ludovic Noels, Self-consistency reinforced minimal gated recurrent unit for surrogate modeling of history-dependent non-linear problems: Application to history-dependent homogenized response of heterogeneous materials, Comput. Methods Appl. Mech. Engrg. 424 (2024) 116881.

[31] Jan Niklas Fuhg, Craig M. Hamel, Kyle Johnson, Reese Jones, Nikolaos Bouklas, Modular machine learning-based elastoplasticity: Generalization in the context of limited data, Comput. Methods Appl. Mech. Engrg. 407 (2023) 115930.

[32] Dug-Joong Kim, Gyu-Won Kim, Jeong-hyeon Baek, Byeunggun Nam, Hak-Sung Kim, Prediction of stress-strain behavior of carbon fabric woven composites by deep neural network, Compos. Struct. 318 (2023) 117073.

[33] Mohsen Mirkhalaf, Iuri Rocha, Micromechanics-based deep-learning for composites: Challenges and future perspectives, Eur. J. Mech. A Solids 105 (2024) 105242.

[34] Yingzhou Lu, Minjie Shen, Huazheng Wang, Xiao Wang, Capucine van Rechem, Wenqi Wei, Machine learning for synthetic data generation: a review, 2023, arXiv preprint arXiv:2302.04062.

[35] Normand J. Beaudry, Renato Renner, An intuitive proof of the data processing inequality, Quantum Info. Comput. 12 (5–6) (2012) 432–441.

[36] Jiajun Wu, Jindong Jiang, Qiang Chen, George Chatzigeorgiou, Fodil Meraghni, Deep homogenization networks for elastic heterogeneous materials with two-and three-dimensional periodicity, Int. J. Solids Struct. 284 (2023) 112521.

[37] Jiajun Wu, Qiang Chen, Jindong Jiang, George Chatzigeorgiou, Fodil Meraghni, Adaptive deep homogenization theory for periodic heterogeneous materials, Compos. Struct. 340 (2024) 118171.

[38] M. El Fallaki Idrissi, Francis Praud, Fodil Meraghni, Francisco Chinesta, George Chatzigeorgiou, Multiscale Thermodynamics-Informed Neural Networks (MuTINN) towards fast and frugal inelastic computation of woven composite structures, J. Mech. Phys. Solids 186 (2024) 105604.

[39] Salah A. Faroughi, Nikhil M. Pawar, Celio Fernandes, Maziar Raissi, Subasish Das, Nima K. Kalantari, Seyed Kourosh Mahjour, Physics-guided, physics-informed, and physics-encoded neural networks and operators in scientific computing: Fluid and solid mechanics, J. Comput. Inf. Sci. Eng. 24 (4) (2024) 040802.

[40] M.A. Maia, I.B.C.M. Rocha, P. Kerfriden, F.P. van der Meer, Physically recurrent neural networks for path-dependent heterogeneous materials: Embedding constitutive models in a data-driven surrogate, Comput. Methods Appl. Mech. Engrg. 407 (2023) 115934.

[41] Hon Lam Cheung, Mohsen Mirkhalaf, A multi-fidelity data-driven model for highly accurate and computationally efficient modeling of short fiber composites, Compos. Sci. Technol. 246 (2024) 110359.

[42] Hon Lam Cheung, Petter Uvdal, Mohsen Mirkhalaf, Augmentation of scarce data—A new approach for deep-learning modeling of composites, Compos. Sci. Technol. (2024) 110491.

[43] Jan Zeman, Tom W.J. de Geus, Jaroslav Vondřejc, Ron H.J. Peerlings, Marc G.D. Geers, A finite element perspective on nonlinear FFT-based micromechanical simulations, Internat. J. Numer. Methods Engrg. 111 (10) (2017) 903–926.

[44] Hervé Moulinec, Pierre Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Comput. Methods Appl. Mech. Engrg. 157 (1–2) (1998) 69–94.

[45] Julian N. Heidenreich, Colin Bonatti, Dirk Mohr, Transfer learning of recurrent neural network-based plasticity models, Internat. J. Numer. Methods Engrg. (2023).

[46] Issam Doghri, L. Tinel, Micromechanical modeling and computation of elasto-plastic materials reinforced with distributed-orientation fibers, Int. J. Plast. 21 (10) (2005) 1919–1940.

[47] Ling Wu, Ludovic Noels, L. Adam, Issam Doghri, A multiscale mean-field homogenization method for fiber-reinforced composites with gradient-enhanced damage models, Comput. Methods Appl. Mech. Engrg. 233 (2012) 164–179.

[48] F. Lamon, P.A. Carraro, L. Maragoni, M. Quaresimin, A multiscale strategy for assessing the micro-scale stress distribution in the matrix of textile composites, Compos. Part A: Appl. Sci. Manuf. 159 (2022) 107026.

[49] Juan C. Simo, Thomas J.R. Hughes, Computational Inelasticity, vol. 7, Springer Science & Business Media, 2006.

[50] Phil Gregory, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support, Cambridge University Press, 2005.

[51] J. Friemann, B. Dashtbozorg, M. Fagerström, S.M. Mirkhalaf, A micromechanics-based recurrent neural networks model for path-dependent cyclic deformation of short fiber composites, Internat. J. Numer. Methods Engrg. 124 (10) (2023) 2292–2314.

[52] Wayne A. Woodward, Bivin Philip Sadler, Stephen Robertson, Time Series for Data Science: Analysis and Forecasting, Chapman and Hall/CRC, 2022.
[53] Mingtian Tan, Mike A. Merrill, Vinayak Gupta, Tim Althoff, Thomas Hartvigsen, Are language models actually useful for time series forecasting?, 2024, arXiv preprint arXiv:2406.16964.
[54] Mononito Goswami, Konrad Szafer, Arjun Choudhry, Yifu Cai, Shuo Li, Artur Dubrawski, Moment: A family of open time-series foundation models, 2024, arXiv preprint arXiv:2402.03885.
[55] Markus J. Buehler, Mechgpt, a language-based strategy for mechanics and materials modeling that connects knowledge across scales, disciplines, and modalities, Appl. Mech. Rev. 76 (2) (2024) 021001.
[56] Christopher M. Bishop, Hugh Bishop, Deep Learning: Foundations and Concepts, Springer Nature, 2023.
[57] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, 2014, arXiv preprint arXiv:1406.1078.
[58] Ilya Sutskever, Oriol Vinyals, Quoc V. Le, Sequence to sequence learning with neural networks, Adv. Neural Inf. Process. Syst. 27 (2014).
[59] Sébastien Brisard, Overview of FFT-based homogenization techniques from the Galerkin point of view, in: Conférence Internationale de Géotechnique, des Ouvrages et Structures, CIGOS 2015, 2015.
[60] Franz Roters, Martin Diehl, Pratheek Shanthraj, Philip Eisenlohr, C. Reuber, Su Leen Wong, Tias Maiti, Alireza Ebrahimi, Thomas Hochrainer, H.-O. Fabritius, et al., DAMASK–the Düsseldorf Advanced Material Simulation Kit for modeling multi-physics crystal plasticity, thermal, and damage phenomena from the single crystal up to the component scale, Comput. Mater. Sci. 158 (2019) 420–478.
[61] Digimat-FE, 2016, https://www.e-xstream.com/products/digimat/tools?fe=1. (Accessed 24 October 2016).
[62] NVIDIA® GeForce RTX™ 4090, 2023, https://www.nvidia.com/en-us/geforce/graphics-cards/40-series/rtx-4090/. (Accessed 04 July 2023).
[63] scikit-learn developers, RobustScaler, 2024, https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html. (Accessed 03 September 2024).
[64] Aleksandr Dekhovich, O. Taylan Turan, Jiaxiang Yi, Miguel A. Bessa, Cooperative data-driven modeling, Comput. Methods Appl. Mech. Engrg. 417 (2023) 116432.
[65] Jiyoung Jung, Yongtae Kim, Jinkyoo Park, Seunghwa Ryu, Transfer learning for enhancing the homogenization-theory-based prediction of elasto-plastic response of particle/short fiber-reinforced composites, Compos. Struct. 285 (2022) 115210.
[66] Carlos Mora, Jonathan Tammer Eweis-Labolle, Tyler Johnson, Likith Gadde, Ramin Bostanabad, Probabilistic neural data fusion for learning from an arbitrary number of multi-fidelity data sets, Comput. Methods Appl. Mech. Engrg. 415 (2023) 116207.
[67] Qiang Yang, Yu Zhang, Wenyuan Dai, Sinno Jialin Pan, Transfer Learning, Cambridge University Press, 2020.
[68] Dengpeng Huang, Jan Niklas Fuhg, Christian Weißenfels, Peter Wriggers, A machine learning based plasticity model using proper orthogonal decomposition, Comput. Methods Appl. Mech. Engrg. 365 (2020) 113008.
[69] Weixing Wang, Haojin Yang, Christoph Meinel, Hasan Yagiz Özkan, Cristian Bermudez Serna, Carmen Mas-Machuca, Feature distribution shift mitigation with contrastive pretraining for intrusion detection, 2024, arXiv preprint arXiv:2404.15382.
[70] Anbo Meng, Haitao Zhang, Zhongfu Dai, Zikang Xian, Liexi Xiao, Jiayu Rong, Chen Li, Jianbin Zhu, Hanhong Li, Yiding Yin, et al., An adaptive distribution-matched recurrent network for wind power prediction using time-series distribution period division, Energy 299 (2024) 131383.
[71] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, Huan Xu, Time series data augmentation for deep learning: A survey, 2020, arXiv preprint arXiv:2002.12478.
[72] Diederik P. Kingma, Jimmy Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
[73] Thomas Bartz-Beielstein, PyTorch hyperparameter Tuning-A tutorial for spotpython, 2023, arXiv preprint arXiv:2305.11930.
[74] Zahra Taghiyarrenani, Sławomir Nowaczyk, Sepideh Pashami, Mohamed-Rafik Bouguelia, Multi-domain adaptation for regression under conditional distribution shift, Expert Syst. Appl. 224 (2023) 119907.