A Microstructure-Informed GRU-Based Autoregressive Framework for Constitutive Modelling

Ho Yin Byan SIU[1], Pratheek Shanthraj[2], David A.W. Barton[3], Burcu Tasdemir[1]
[1]School of Electrical, Electronic and Mechanical Engineering, University of Bristol, Bristol, UK
[2]UK Atomic Energy Authority, Culham Campus, Abingdon, UK
[3]School of Engineering Mathematics and Technology, University of Bristol, Bristol, UK

oo24075@bristol.ac.uk

Abstract

Rapid and accurate assessment of the mechanical response of structural materials under extreme conditions is essential for next-generation scientific applications. Machine learning approaches have shown significant promise in providing efficient and accurate material models. In this study, a Gated Recurrent Unit (GRU) based machine learning (ML) model is developed to predict the stress response of fixed three-dimensional grain structures under arbitrary monotonic deformation. Moreover, this model operates with an incremental mechanism; therefore, it can be directly integrated into finite element analysis (FEA) frameworks. A major challenge associated with ML models operating in an incremental basis is the accumulation of errors over timesteps. Minor errors in early timesteps can propagate, leading to significant path discrepancies in later timesteps. To address this issue, an autoregressive (AR) training method was applied during model training, which resulted in significant performance improvement. Initially, a single microstructure was used for model development, with the aim of extending the approach to multiple microstructures in subsequent stages. Unseen testing results indicate that the current model achieves an average error of approximately 0.5656 MPa^2 over a 200-step sequence. The method was further validated on additional microstructures, where comparable prediction accuracy was observed, demonstrating that the model is capable of generalising across different microstructural configurations.

Data

The data used in this study were provided by UKAEA from a previous study [1]. This dataset is fully synthetic and generated using full-field DAMASK CPFFT simulations. The simulation was built on volume elements (VEs), each containing 20 randomly located and oriented grains, with material properties of FCC aluminium. Each unique microstructure VE was simulated to obtain the mechanical response under various rate-dependant monotonic deformation paths. Further details on the data generation process can be found in [1].

Each microstructure contains about 3000 time-series sequences of deformation-stress response. These were split in the ratio of 0.75:0.05:0.2 for training, validation, and testing, respectively, for surrogate model development.

Model and Training

A GRU-based machine learning model is developed in this study. Recurrent neural network (RNN) models have been widely applied to history-dependent solid mechanics problems, where the material response depends on prior loading paths. For example, RNNs have been used as surrogate models for multiscale finite element simulations of history-dependent materials [2] and to learn constitutive behaviour in viscoelastic systems [3], demonstrating their strong capability in capturing path-dependent responses.

The ML model developed in this study operates in an incremental manner. At each timestep $\boldsymbol{t}$, the model takes as input $\mathbf{x}_t$: [current deformation ($\boldsymbol{F}^T{}_t$), current stress ($\boldsymbol{\sigma}^T{}_t$), deformation increment ($\Delta\boldsymbol{F}_t$)], along with the hidden state vector ($\boldsymbol{h}_{t-1}$) of the VE and predicts the output $\boldsymbol{y}_t$: [increment of stress ($\Delta\boldsymbol{\sigma}_t$)], as well as the updated hidden state vector ($\boldsymbol{h}_t$). The hidden state vector is a variable used by the GRU unit to represent the internal state of the CP model. At each timestep, the hidden state is used by the model along with the input ($\mathbf{x}_t$) and its value is updated at the end of each timestep.

An important aspect of this work is the use of an autoregressive (AR) training strategy. In contrast to many data-driven constitutive modelling approaches that employ direct mappings from strain to stress, in this study, the predicted stress increment ($\Delta\boldsymbol{\sigma}_t$) is fed back to update the current stress state, forming part of the input for the next timestep ($\boldsymbol{\sigma}^T{}_{t+1}$).

However, this mechanism is highly sensitive to error accumulation, as small errors at early timesteps can lead to significant deviations in the predicted trajectory at later stages. While fully connected neural network (FCNN) -based approaches [4],[5] have demonstrated good performance for incremental constitutive modelling with relatively compact input representations, they are not well-suited to this problem as the model works on all the components of the stress tensor. Therefore, an autoregressive training approach is employed. At each training step, instead of relying solely on ground-truth sequences, the model performs autoregressive rollout of the output sequence, which is then used to construct subsequent inputs. This rollout is used not only in prediction but also during training. As the model is trained under the same mechanism as the intended application, this could effectively mitigate the error accumulation issue mentioned above. Tests on the unseen data have shown the AR training was an effective solution. Figure 1 illustrates the AR workflow.

Regarding the model architecture, the input and hidden state are first passed through two fully connected neural network (NN) layers, followed by GRU layers.

The selected model hyperparameters include 4 GRU layers, a hidden state size of 90, and 564 neurons in each fully connected layer. The ReLU activation function and the mean squared error (MSE) loss function were used in training. The model was trained for 15000 epochs with learning rate 1e-4.

Results

Figure 2 demonstrates the model predictions for one sequence from the unseen test dataset. The plot shows the total stress trajectories of all 6 components of the stress tensor over 200 timesteps. These trajectories were obtained using the AR rollout and follow a sequential state-update mechanism analogous to the incremental mechanism used in FEA.

It is observed (Figure 2) that two stress components exhibit relatively larger errors; however, overall, the predicted trajectories closely follow the ground truth. Across all 589 unseen testing sequences, each with 200 steps, the average error per step is 0.5656 MPa^2.

It is worth noting that the AR training approach incurs a higher computational cost compared to conventional training, due to the need for sequence rollout during training. The computational cost increases with sequence length; however, this can be partially mitigated by increasing the batch size. Another potential issue is gradient explosion, which is common in RNN-based models due to sequential error propagation. However, no significant gradient instability was observed in this study.

Further validation was conducted by training separate models on datasets corresponding to different microstructures. Consistent prediction accuracy was achieved across these cases, demonstrating that the proposed model architecture provides a robust and generalisable solution for this class of problems.

As this methodology has been shown to perform well for a single microstructure, the next step is to develop a model capable of handling multiple and randomly selected microstructures. Subsequently, the trained model will be implemented in FEA software, benefiting from its incremental formulation. Previous work has demonstrated the feasibility of integrating machine learning models into FEA frameworks, for example through user-defined hardening (UHARD) subroutines under uniaxial cyclic loading conditions [6]. However, such implementations are typically limited to single stress components. In contrast, the present approach is designed to capture the full stress tensor.

References

[1] M. D. Atkinson, M. D. White, A. J. Plowman, P. Shanthraj, Microstructure sensitive recurrent neural network surrogate model of crystal plasticity, arXiv preprint, (2025) arXiv:2510.06904.

[2] F. Ghavamian, A. Simone, Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network, Computer Methods in Applied Mechanics and Engineering, 357 (2019) 112594.

[3] G. Chen, Recurrent neural networks (RNNs) learn the constitutive law of viscoelasticity, Computational Mechanics, 67(3) (2021) 1009–1019.

[4] B. Tasdemir, A. Pellegrino, X. Su, V. L. Tagarielli, Learning the non-proportional multiaxial elastic–plastic response of an aluminium alloy with neural networks, Materials & Design, 253 (2025) 113956.

[5] B. Tasdemir, A. Pellegrino, V. Tagarielli, A strategy to formulate data-driven constitutive models from random multiaxial experiments, Scientific Reports, 12(1) (2022) 22248.

[6] J. Lee, B. Tasdemir, M. Martin, D. Knowles, M. Mostafavi, Development of a data-driven user-defined hardening model for cyclic behaviour via transformer, Materials & Design, (2026) 115417.

Figures

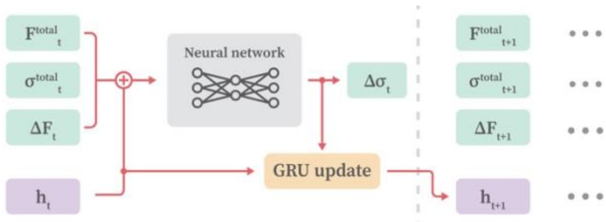

Figure 1. AR training workflow demonstrating how this approach uses the model with the most updated parameter to rollout both input and output time sequences.

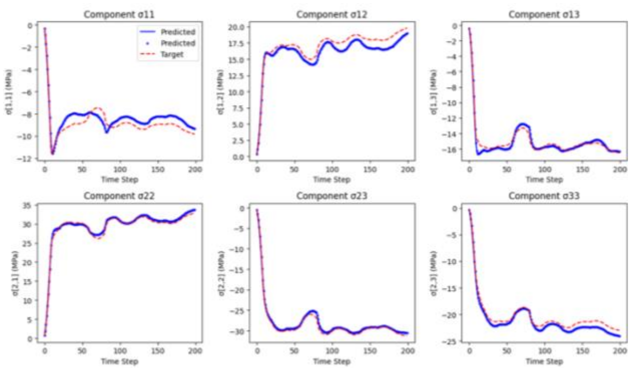

Figure 2. Model prediction for unseen testing sequence demonstrating the stress trajectory of all 6 components in the stress tensor along 200 timesteps.