PAPER • OPEN ACCESS

Prediction of the evolution of the stress field of polycrystals undergoing elastic-plastic deformation with a hybrid neural network model

View the article online for updates and enhancements.

MACHINE LEARNING
Science and Technology

PAPER

Prediction of the evolution of the stress field of polycrystals undergoing elastic-plastic deformation with a hybrid neural network model

Ari Frankel[1], Kousuke Tachida[1] and Reese Jones[1]

Sandia National Laboratories, Livermore, CA 94551 United States of America

E-mail: rjones@sandia.gov

Abstract

Crystal plasticity theory is often employed to predict the mesoscopic states of polycrystalline metals, and is well-known to be costly to simulate. Using a neural network with convolutional layers encoding correlations in time and space, we were able to predict the evolution of the dominant component of the stress field given only the initial microstructure and external loading. In comparison to our recent work, we were able to predict not only the spatial average of the stress response but the evolution of the field itself. We show that the stress fields and their rates are in good agreement with the two dimensional crystal plasticity data and have no visible artifacts. Furthermore the distribution of stress throughout the elastic to fully plastic transition match the truth provided by held out crystal plasticity data. Lastly we demonstrate the efficacy of the trained model in material characterization and optimization tasks.

1. Introduction

The modeling of the mechanical behavior of polycrystalline metals has wide spread technological relevance. The prediction of the stress states of polycrystals as they deform is difficult, especially during transition to plasticity. Homogenization theory can provide bounds and approximations to the macroscopic response, but is, by construction, insensitive to the details of the microscopic states influencing the response of a particular polycrystal. On the other hand, crystal plasticity theory [1–7] provides a constitutive model that enables detailed meso-scale simulations of metals; however, it is complex and costly to simulate. Many higher level models of the phenomenology of plastic deformation have been developed over decades [8, 9] and are often calibrated to represent polycrystalline metals. While this approach is sufficient in the small grain/large sample limit (representative volumes [10, 11] and larger), there is a need for accurate, efficient models of the response of aggregates with smaller sample sizes (stochastic volumes [12]). These models can be used for stand-ins for material design/optimization, and also for uncertainty quantification and other statistical analysis.

The high cost for traditional finite element formulations of crystal plasticity engendered innovative approaches beyond straight-forward finite element implementations, such as Green's function based spectral methods, see [13] for a review and [14–19] for details. Included in these developments is the use of databases [14, 20] and generic functional bases [21, 22] which can be seen as steps toward data-driven modeling. Kalidindi and co-workers have been especially active in the field and a summary of their current developments can be found in [22].

In a wider context, data-driven modelling has experienced recent rapid growth. In fact, across a range of commercial and scientific fields, data-driven modeling and, in particular, machine learning [23, 24] has risen as a tool to form representative models of complex underlying physics without resorting to gross approximations. Given a large enough dataset and a trainable model of sufficient and appropriate complexity, many machine learning techniques can provide a high fidelity representation of the underlying physics.

Although nascent, the field of applying machine learning to materials science has garnered sufficient interest to elicit topical reviews, such as [25–27]. There are also a number of notable individual contributions employing a variety of approaches to the tasks of classification and reconstruction/synthesizing microstructures, and predicting their physical response. Key ingredients in these tasks are selection of (a) the inputs by feature extraction from image, either manually or by the network, and (b) the output, such as classes of the expected microstructures or their stress response. Convolutional neural networks (CNNs) [28] typically play a role in feature extraction and latent space discovery since they were developed to efficiently process image data and have the distinct advantage of being able to discern the features salient to the selected outputs when trained successfully.

Image-based learning encompasses a variety of approaches. Image-based transfer learning is a scheme that leverages developments in computer science, such as Gatys *et al* [29]. In transfer learning, models trained in similar contexts are reused with minimal additional training. Lubbers *et al* [30] adapted a convolutional neural network trained for image recognition to form low-dimensional representations of microstructures in order to generate statistical reconstructions. In the related field of modeling atomic phenomena, Noh *et al* [31] used Gaussian smearing of atomic positions together with CNN to circumvent the more traditional approach of feature selection [32]. Li *et al* [33] also used transfer learning for microstructure reconstruction for a wide range of classes of microstructure. Clustering is another of the variety of methods employed for classification. Papanikolaou *et al* [34] employed dislocation dynamics data and a clustering algorithm to classify materials based on prior plastic deformation. Liu *et al* [20] employed a database of microstructures and a clustering algorithm to predict plastic localization. Using a variety of techniques, such as visual bag of words, texture and shape statistics, and pre-trained convolutional neural networks, support vector machines, clustering, and random forests, Chowdhury *et al* [35] performed feature extraction and classification of dendritic morphologies.

The task of microstructure reconstruction/generation has a strong image processing component and is particularly germane to a variety of materials studies. Chen and coworkers [36–39] have produced a considerable body of work in this emerging field, primarily focussed on microstructure reconstruction [38]. Xu *et al* [36] used pre-selected and winnowed image features in a supervised learning process to represent microstructures. Bostanabad *et al* [38, 39] used a Gibbs sampler and a classification tree to construct binary microstructures. Generative models such as variational autoencoders (VAEs) [40] and generative adversarial networks (GANs) [41] have been employed for microstructure reconstruction and optimization. Cang *et al* [42] used a VAE to represent granular microstructures and compared its performance to the more traditional approach of using a Markov random field model. Yang *et al* [43] trained a GAN to provide latent variables that may be used to optimize the properties of a material structure. Kalidindi and coworkers [44–48] have also generated numerous papers in the field of microstructure analysis and reconstruction. Particularly relevant is the work of Cecen *et al* [46] and Yang *et al* [47], which compared more traditional models, based on statistical and physical descriptors, with data-driven methods like principal component analysis (PCA) and compared to a model where a CNN was used to encode the spatial correlations in the elastic response of a two phase material. They showed showed that the CNN model was superior to the reduced order models using the physical descriptors and regression, as well as classical models based on homogenization theory.

Other work has also focused on predicting the physical response of microstructures. Niezgoda [26, 49, 50] has been active in the field of applying machine learning to materials science, in part with Kalidindi [50] and also with commercial groups [26, 49]. As mentioned Kalidindi and coworkers have be actively pursuing data-driven modeling for microstructural problems with emphasis on crystal plasticity. In one of their latest works Montes de Oca Zapiain and Kalidindi [21, 22] used a physics-based features with spherical harmonic kernels to encode crystal orientation and coefficients tuned with regression to model plastic stretch rates and simulate multiscale 3D stress evolution. Also related to the present effort, Yuan *et al* [49] used a random forest model with PCA of texture and the underlying model parameters as inputs to predict the history for multiple loading conditions. Mangal and Holm [51] used a random forest decision tree and pre-selected features to predict local environments that lead to high stress in uniaxial tension. Bessa *et al* [37] developed a methodology for applying machine learning to modelling physical response with particular focus on sampling of the training data. Jones *et al* [52] designed neural network models of elastic-plastic behavior based on the traditional stress and flow rule framework and key aspects of classical representation theory. Yang *et al* [48] compared a CNN based model for making point estimates for a two phase elastic material to more traditional feature based techniques, and used the trained model in multiscale predictions. Recently, Frankel *et al* [53] demonstrated the use of convolutional neural networks to predict the elastic modulus of polycrystalline aggregates and combined the CNN architecture with a recurrent neural network to predict the system average stress evolution through yield and plastic flow. Although the results were, in part, a proof-of-concept of the use of data-driven modeling for crystal plasticity, the architecture was limited to predicting the homogenized stress states and had reduced accuracy as time progressed.
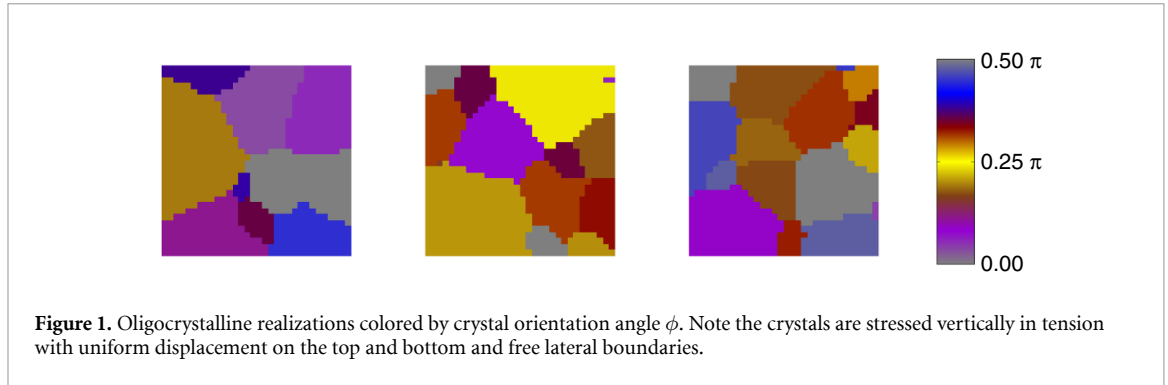
In this work we create a data-driven, full-field predictor of the dominant component of the stress of plastically-deforming polycrystals with the use of a novel hybrid neural network architecture based on a convolutional-long short term memory neural network (ConvLSTM) [54]. The ConvLSTM *simultaneously* (versus sequentially) processes the spatial and temporal aspects of the data with trainable causal filters in time that have underlying, embedded spatial convolutions. As a convolutional neural network, the ConvLSTM is efficient in terms of number of parameters needed to represent processes that have temporal and spatial locality of action such as most physical initial value problems. Our architecture employs a ConvLSTM at its core and separate convolutional unit to take a complex initial image of one type (microstructure) to predict the evolution of physically related images of another type (stress field) and makes no assumptions of the form of the functions describing the evolution. The convolutional aspects of the network allow for the material features relevant to the output stress to be discovered through correlation with the output instead of guessed/pre-supposed. Given the high-fidelity of the predictions we can confidently assert that subsequent use of the model in structure-property explorations is based on the relevant underlying features, as opposed to reduced models based on pre-selected features. As with our previous work [53] (which predicted the evolving average stress of three dimensional (3D) microstructures), we predict the *evolution* of the stress response through elastic-plastic transition to full plastic flow which is considerably more difficult for a unified network that merely predicting the initial elastic response and involves previous states to predict future ones. Unlike our previous work, we predict the entire stress field instead of its spatial average by the use of the ConvLSTM which combines the application of the spatial filters of a CNN with the time kernels of a long short term memory (LSTM) recurrent neural network. In our previous work the output of the spatial convolution was reduced to a small number of features before feeding into a separate LSTM to capture the time dependence. In this work the initial field is not reduced through an encoder but is fed forward and, together with the loading, is used to predict the stress field evolution through time. To the best of our knowledge this is the first neural network model to predict the elastic-plastic *evolution* of the stress field from an initial microstructure. Another improvement over our previous work is we ameliorate the growth in error with time using an additional convolutional layer that processes the spatial and temporal information simultaneously. The extension of this network to three dimensional (3D) data and other types of microstructure is straight-forward and will be discussed in the concluding section.

There are significant advantages of the approach we demonstrate over recent developments in the field. The first is the embedding of the temporal evolution of the stress states in the data-driven model through a recurrent component of the architecture which is combined with the spatial convolution. Another is that neural networks have a universal approximation property that ensures that given a large enough network and sufficient data to train it, the neural network can learn to represent any function of interest [55, 56]. This means that as opposed to traditional physics-based models or simpler regression models, our architecture has the capacity to converge to an unbiased representation of the underlying mechanics. As we will demonstrate, it has deep learning capabilities to embed features of the initial microstructure that are not pre-selected instead and are instead best correlated with the selected output. Furthermore, given that the method is image based and agnostic to the type of microstructure and the particular physical process that drives the evolution of the output field field, it should be directly extensible to a wide class of physics, materials science, and engineering problems.

In section 2, we describe the polycrystal dataset and then, in section 3, we discuss the particular ConvLSTM-based model in the wider context of neural networks and compare variants of the proposed model. In section 4, we show results for the prediction of the stress state evolution, including error analysis, statistical comparisons, and a material exploration demonstration. Lastly, we discuss extensions and additional applications in section 5.

2. Crystal plasticity data

Crystal plasticity (CP) simulations typically consist of a finite element representation of the grain structure, including a mesh defining the grain boundaries and material models informed by the crystal orientation of each grain and the allowed slip systems. So as not to be data limited in training for our neural network (described in section 3.2), we employed over 10^4 two dimensional (2D) simulations of polycrystals. This is particularly crucial to make the potential weaknesses in the neural network architecture apparent and unambiguous. Each sample consisted of the tensile stress response of 1 μm^2 squares of polycrystalline steel with a single face-centered cubic (FCC) phase. The cost for a 3D sample was on the order of 100 cpu-hours in previous studies [52, 53] where we generated approximately 10^3 tension samples. In this study, they are on the order of 1 cpu-hour. Also, given the relative magnitude of the components of the stress tensor in tension as well as the need to expand the network with the number of components, we chose to only model the dominant normal component of the stress in the tension direction. Each grain was assigned a viscoplastic

Figure 1. Oligocrystalline realizations colored by crystal orientation angle ϕ. Note the crystals are stressed vertically in tension with uniform displacement on the top and bottom and free lateral boundaries.

constitutive model with power-law hardening and the crystal interactions in aggregate were governed by compatibility.

2.1. Realizations

To produce the 2D realizations, we utilized the well-tested algorithms in the Dream.3d software package [57] for creating 3D polycrystals. In particular, we created 1 μm^3 cubes with a realistic polycrystalline morphology using Dream.3d and then took slices of these three dimensional (3D) representations. For ease of slicing and subsequent simulation efficiency, the cubes were 32^3 element structured meshes with voxelated grain boundaries. To minimize correlation between 2D samples we sliced each cube on every 7th plane in the out-of-plane dimension. For the 2D slices only the in-plane texture was preserved so that the crystal orientation can be described by the in-plane rotation by angle ϕ of a $\langle 100 \rangle$ canonically oriented FCC crystal. This procedure has the benefit of being connected to, if not a full representation of, realistic 3D polycrystals.

Figure 1 shows representative realizations from this process colored by the texture angle $\phi \in [0, \pi/2]$. The pixelated grain boundaries are apparent, as are the fairly equiaxed grains. Figure 2 shows descriptive statistics of the ensemble of 2D realizations. The orientation angle distribution is approximately uniform and is correlated over the spatial domain of the realizations. Also, roughly a quarter (24.7%) of the elements (across realizations) have a grain boundary on their border. Clearly the 2D (and 3D) samples are stochastic volume elements (SVEs) since the correlations are long compared to the size of the domain. This spatial correlation was inherited from the process that created the 3D cubes. Since the samples are not at the representative volume element (RVE) limit, we will refer to the realizations as oligocrystals (as opposed to polycrystals) in the subsequent sections. The number of grains per realization ranges from 5 to 25, with 14 being the most prevalent (corresponding to an average grain size of 0.15 μm^2 and number density of 0.071 μm^{-2}). The distribution in grain size is approximately log normal with mode at approximately 0.1 μm^2. Again, this distribution is inherited from the process used to create the 3D realizations. The distribution of misorientation between grains has slight preference for low angle boundaries but there are also a significant number of high angle boundaries. Lastly, the grain boundary length distribution is fairly compact with a mode at approximately 6 μm.

2.2. Mechanical response

Given the crystal orientation, the response of each grain in the oligocrystal followed a widely employed elastic-viscoplastic constitutive response [1–7] representative of austenitic (FCC) steel. For the crystal elasticity, we employed a finite deformation Saint Venant model, which is a linear relation between the second Piola-Kirchhoff stress **S** in the intermediate configuration and the elastic Lagrangian strain $\mathbf{E}^e$,

$$\mathbf{S} = \mathbb{C} : \mathbf{E}^e . \tag{1}$$

The Cauchy stress $\boldsymbol{\sigma}$ is obtained through the usual relation involving the deformation gradient. Since elastic modulus tensor $\mathbb{C}$ respects FCC crystal symmetries, there are only three relevant non-zero components, C_{11}, C_{12}, C_{44}, and these were set to 204.6, 137.7, 126.2 GPa [58], respectively. In 3D, plastic flow can occur on any of the 12 FCC slip planes (with Schmid tensors $\{\mathbf{S} \otimes \mathbf{n}\} = \langle 111 \rangle \otimes \{110\}$) in each crystal subject to compatibility between grains. For this 2D study, the allowed in-plane rotation by ϕ leaves the third component of the slip **S** and normal **n** vectors unchanged; and, a plane strain constraint is enforced via projection of the motion into the plane. Note that in this quasi-2D formulation, multiple slip systems have the same projection into the plane. Furthermore, the effective yield strength varies across samples in part due to the underlying variation in single crystal yield strength. A single crystal oriented at $\phi = \pi/4$ has the maximum yield strength, and the orientations $\phi = 0, \pi/2$ have near maximum values (approximately 3% lower), while $\phi = \pi/8, 3\pi/8$ have minimum effective yield strengths at $\approx 15\%$ lower. As in a previous study

Figure 2. Descriptive statistics of the oligocrystal ensemble.
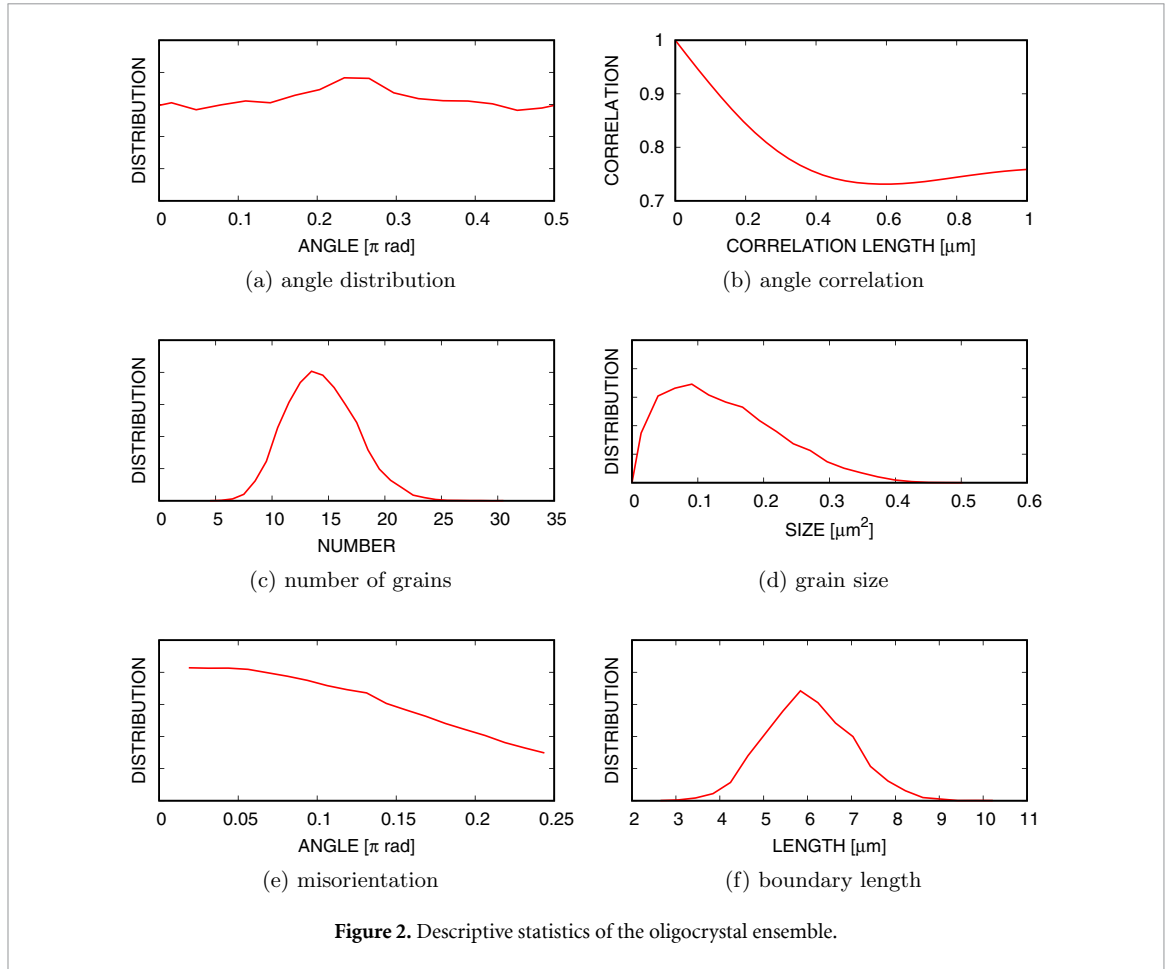
[53], we employed a power-law form for the slip rate relation

$$\dot{\gamma}_\alpha = \dot{\gamma}_0 \left| \frac{\tau_\alpha}{g_\alpha} \right|^{m-1} \tau_\alpha \,, \tag{2}$$

driven by the shear stress τ_α resolved on slip system α. We chose the reference slip rate $\dot{\gamma}_0 = 1.0 \ \mathrm{s}^{-1}$, the rate exponent $m = 20$, and the initial slip resistances $g_\alpha(t=0) = 122.0$ MPa. The slip resistance evolves according to [59, 60]

$$\dot{g}_\alpha = (H - R g_\alpha) \sum_\alpha |\dot{\gamma}_\alpha| \tag{3}$$

where the hardening modulus is $H = 355.0$ MPa and the recovery constant is $R = 2.9$. See [52, 53] for more details.

Under the plane strain constraint, each oligocrystal realization was subjected to quasi-static uniaxial tension at a constant engineering strain-rate of $\dot{\epsilon} = 1 \ s^{-1}$ up to 0.3% strain with free lateral boundaries and uniform displacement on the top and bottom. These boundary conditions create a nominal homogeneous stress state and the loading range was chosen to cover the elastic-plastic transition and the subsequent plastic flow, as can be seen in the stress-strain data shown in figure 3 (representative stress fields are shown in figure 9 and will be discussed in section 4). We used the Albany finite element code [61] for these simulations. Note that a richer set of loading conditions could be explored, at the cost of increasing the size and cost of the training data manifold [52]. Here we focus on tension since this is the most common loading condition in high-throughput experiments, see Salzbrenner *et al* [62] for example.

3. Neural network model

Given the relatively recent application of machine learning to materials science, we first give a brief overview of the relevant techniques, specifically neural networks that can be applied to predicting the evolution of the stress field in an oligocrystal given an initial grain structure and the loading. Then we describe the particular architecture and training of the proposed network.
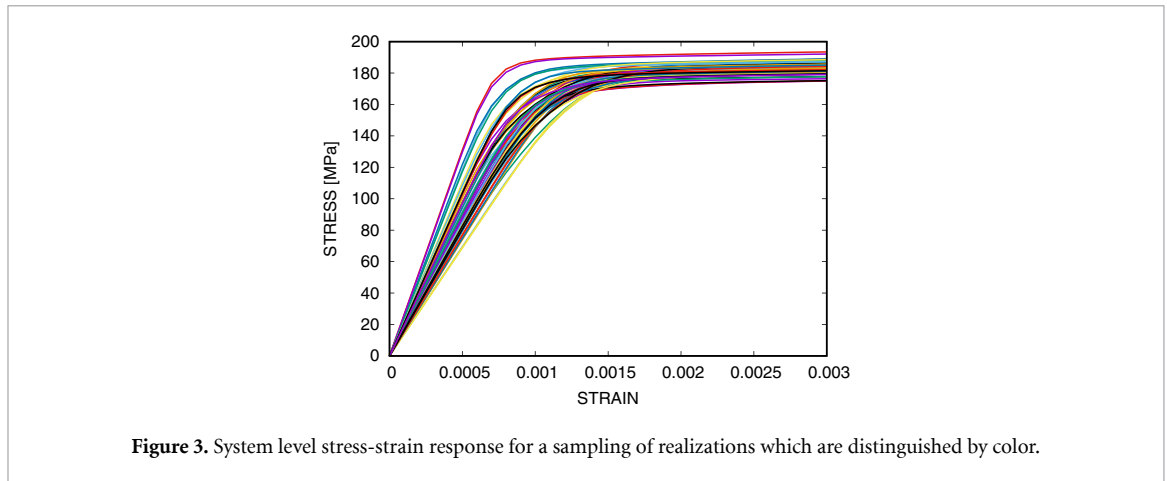
Figure 3. System level stress-strain response for a sampling of realizations which are distinguished by color.

The basic type of artificial neural network (NN) is the *multilayer perceptron* (MLP) [63]. This feed-forward model is a directed graph that feeds the selected inputs through layers of nodes that are densely connected between neighboring layers. Each node transforms its inputs via a non-linear activation function with trainable weights and biases. The result of all the transformations of the inputs is given by the output layer that is compared to data in training/calibrating the model using a selected error/loss metric. A MLP neural network with L layers and N nodes per layer requires the optimization of $O(LN^2)$ parameters.
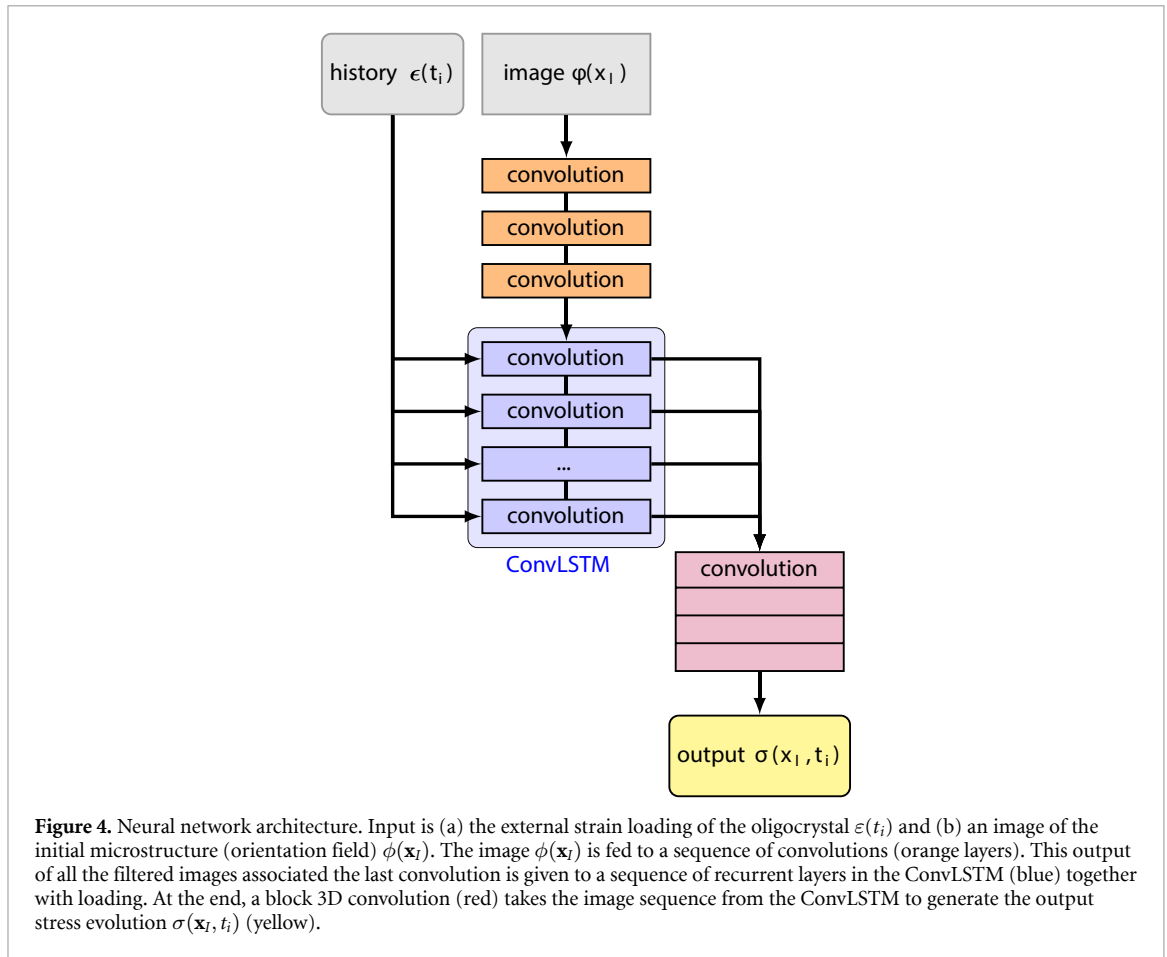
For image or field data, the number of parameters needed to model a function of the image is too large to be trained practically, and would require a prohibitive amount of data in the current materials science context. However, for field data, it is physically plausible that spatial correlations exist that decay with distance, see figure 2(b) for example. The *convolutional neural network* (CNN) [24] was developed to take advantage of this locality. Rather than prescribing a unique set of weights to every single pixel, the convolutional neural network uses a convolutional filter with compact support and single set of weights that is moved across the entire image to produce output of the same dimension. The result of this operation is a new image and multiple such sets of filters can be applied in parallel. In general, further non-linear operations, reductions of the image, or convolutions may be performed to enhance the richness of the feature space of the image or reduce the dimensionality before finally yielding the desired output.

A similar heuristic is commonly used for time-series data, where a long sequence of data may require an inordinate number of parameters to fit to an output. Based on causality, we expect that a future state will depend most strongly on the current state along with some additional latent variables controlling non-stationarity of the signal being modeled. Thus *recurrent neural network* (RNN) architectures were created to take as input to a single layer MLP the features from the current state as well as the output from the previous state in order to make predictions over a time sequence. In this sense, they have some features in common with traditional time integrators. A number of advanced variants of the recurrent neural network design, including the *long short term memory* (LSTM) [64], include additional hidden variables that take into account a longer running history/memory of the signal and thus improve the accuracy and stability of training these neural networks.

The *convolutional LSTM* (ConvLSTM) [54] was designed to combine the heuristics of the CNN and LSTM architectures for making predictions of a sequence of images. Where an LSTM uses a small MLP as a non-linearity for feature processing and prediction, a convolutional LSTM uses a set of convolution operations on each image to create the next image. Specifically, the input image from the current state, as well as the output image from the previous state and hidden variables, are used as inputs to make subsequent predictions. This architecture has particular appeal for microstructural mechanics since it exploits both spatial and temporal correlations in the data to be modelled. For our application, the physical topology of the grains influences the stress state of the microstructure in a way that simple reduced order statistics like average grain size cannot capture reliably, and so we need a model that can analyze the full-field physical state of the system at once. At the same time, the evolution of the stress as a function of strain is expected to depend heavily on the previous stress states as well as the evolution of the latent variables associated with plastic dissipation, which indicates a sequential model with a longer-term memory.

3.1. Architecture

Figure 4 shows a schematic of the proposed architecture which has a ConvLSTM at its core. The inputs (gray) to the network are: (a) a time sequence $\varepsilon(t_i)$, and (b) an image $\phi(\mathbf{x}_I)$ (2D array) corresponding to the first time, t_0. The outputs (yellow) are images, $\sigma(\mathbf{x}_I, t_i)$, for each time t_i in the sequence (a 3D array). Note that we
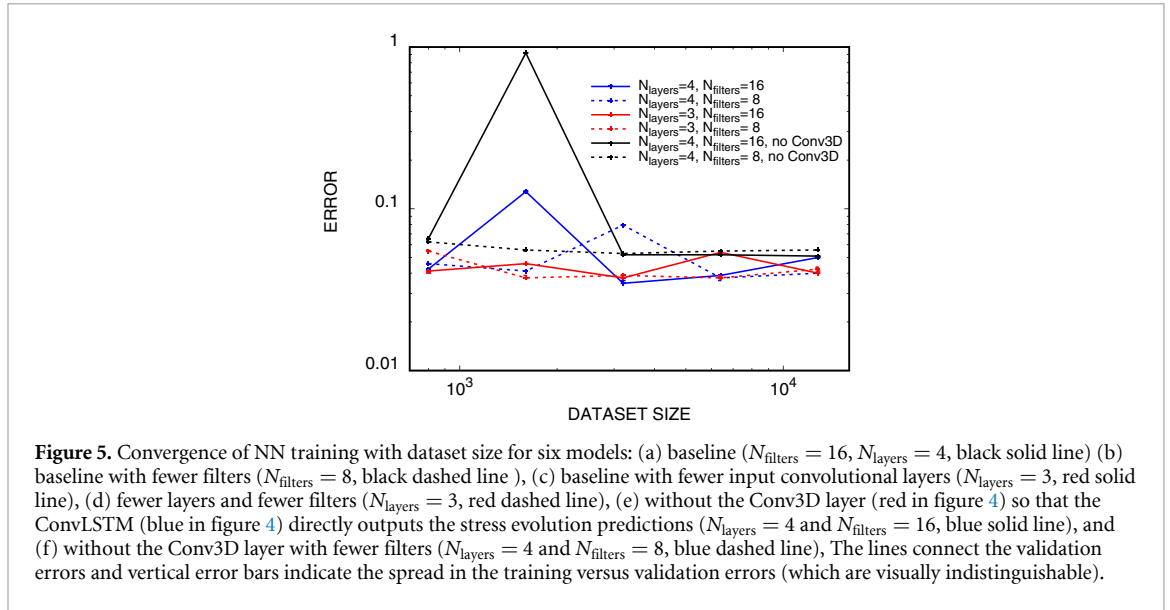
Figure 4. Neural network architecture. Input is (a) the external strain loading of the oligocrystal $\varepsilon(t_i)$ and (b) an image of the initial microstructure (orientation field) $\phi(\mathbf{x}_I)$. The image $\phi(\mathbf{x}_I)$ is fed to a sequence of convolutions (orange layers). This output of all the filtered images associated the last convolution is given to a sequence of recurrent layers in the ConvLSTM (blue) together with loading. At the end, a block 3D convolution (red) takes the image sequence from the ConvLSTM to generate the output stress evolution $\sigma(\mathbf{x}_I, t_i)$ (yellow).

train only to the normal component of stress in the (horizontal) tension direction σ since it dominates the response and modeling all components would require a larger network and more training samples. Unlike in traditional applications where the input image and output images are of the same quantity, in our network the initial image $\phi(\mathbf{x}_I)$ is the crystal orientation field characterizing texture, while the output images are the stress field $\sigma(\mathbf{x}_I, t_i)$ as a function of time/external loading strain $\varepsilon(t_i)$. Here, $\mathbf{x}_I$ are element locations which are identified with image pixels I and $\phi(\mathbf{x}_I)$ with the color of the pixel I. Image size $N \times N = 32 \times 32$ is fixed by the data described in section 2. The image of the initial microstructure is fed into a sequence of N_{layers} 2D convolutional layers (Conv2D, orange) with rectified linear unit (ReLU) non-linear activations [65]. Each layer has N_{filters} filters applied in parallel. Each of the filters has a kernel size of 4×4, with zero-padding to maintain the image size pre- to post-filter.

The output of the initial convolutional layers and the loading history is combined in the ConvLSTM (blue); its recursive structure is shown unrolled in figure 4 for illustration. To embed the time-dependence of the problem, each processed image is repeated for each of the time steps, and an additional filter is added to each whose value is constant across the image and equal to the value of the time step. Thus the input to the convolutional LSTM is a 4-dimensional tensor of size $N_{\text{times}} \times N \times N \times (N_{\text{filters}} + 1)$, indexed by time step, image row, image column, and filter value, respectively. The ConvLSTM used the same kernel size, number of filters, padding, and activation as the initial 2D convolutional layers. Its output was sequence of images, one for each time step. Finally, a 3D convolutional layer (Conv3D, red) with 1 filter was applied across the two space and one time dimensions of the input data, as processed by the ConvLSTM. Its output (yellow) is the time sequence of images of the stress field $\sigma(\mathbf{x}_I, t_i)$. To keep the cost of training manageable, the stress field data was limited to the normal component of stress in the tension direction i.e. the most significant component.

The Keras-Tensorflow framework [66] was used to construct and train the proposed neural network architecture.

3.2. Training
To train the proposed network and some variants, a dataset of 16,000 input-output pairs $\{[\epsilon(t_i), \phi(\mathbf{x}_I)], \sigma(\mathbf{x}_I, t_i)\}$ from the CP simulations described in section 2 was split into 80% for training and

Figure 5. Convergence of NN training with dataset size for six models: (a) baseline ($N_{filters} = 16$, $N_{layers} = 4$, black solid line) (b) baseline with fewer filters ($N_{filters} = 8$, black dashed line), (c) baseline with fewer input convolutional layers ($N_{layers} = 3$, red solid line), (d) fewer layers and fewer filters ($N_{layers} = 3$, red dashed line), (e) without the Conv3D layer (red in figure 4) so that the ConvLSTM (blue in figure 4) directly outputs the stress evolution predictions ($N_{layers} = 4$ and $N_{filters} = 16$, blue solid line), and (f) without the Conv3D layer with fewer filters ($N_{layers} = 4$ and $N_{filters} = 8$, blue dashed line), The lines connect the validation errors and vertical error bars indicate the spread in the training versus validation errors (which are visually indistinguishable).

20% for testing. One tenth of the training set was used for network validation to monitor for convergence during training and prevent over-fitting. This amounts to 11,520 training samples, 1,280 validation samples, and 3,200 testing/prediction samples. The networks were trained using an Adam optimizer [67] guided by a mean-squared-error loss function summed over all pixels I and times t_i. The Adam optimizer [67] used an initial learning rate of 0.001 and a batch-size of 128. GPUs were used to overcome the computational expense of training the network so that each training epoch took approximately 100 seconds. Training proceeded for 1500 epochs or until the validation error stopped decreasing (approximately 100 gpu-hours). To aid in the training, the stress field data was also normalized such that the average value of the stress at the final time step was equal to 1. In preliminary studies datasets of 1000 samples with resolutions 16×16 and 64×64 were used in training with comparable results to the 32×32 resolution we used for the final studies and was a compromise between microstructure resolution and computational cost.

The network shown in figure 4 has a large space of hyperparameters, especially if we were to assign kernel width, etc per layer. We did a non-exhaustive exploration of the hyperparameter space to determine the relative performance of variants of the architecture. Figure 5 compares the root mean square (pixel-wise) validation error versus size of training dataset for a few architecture variants. For each of the trainings the training and validation errors were comparable (to within 6%), which indicated that the networks were not over-fit to the data. The flattening of the error curves in figure 5 is evidence of limits to the learning and accuracy of the models, and may indicate that the networks were not complex enough or the increase in the size of the dataset provides no additional information. The main finding of the architecture exploration is that the Conv3D appears to be needed to reduce the training error. Regarding the other modifications, namely changing the number of Conv2D layers and the number of filters, the results are more equivocal. Since we were only able to train each network variant once due to the training costs, we conjecture that the scatter in the losses of the best networks is likely due to the vagaries of the stochastic optimization method (and the early stopping criterion). Given that the architecture with $N_{filters} = 16$ and $N_{layers} = 4$ achieved marginally the lowest error, we focussed on it in the subsequent results. Further optimization of the network architecture should be possible.

4. Results

After training the model we evaluated the quality of its predictions for the response of different microstructures to the same loading conditions as the training data, and demonstrated its use in material evaluation/optimization.

4.1. Quality of the predictions

We used the trained model to predict the response of the 3,200 testing samples not employed in training (described in section 3.2) and compared the stress history $\sigma(\mathbf{x}_I, t)$ predictions to the corresponding held out CP data. Note that, given the strain rate ($\dot{\varepsilon} = 1/s$), the time t and the loading strain $\varepsilon(t)$ have the same values.

In this section we focus on the errors in the (post-training) predictions of held out data for both the full-field stress $\sigma(\mathbf{x}_I, t)$ and its spatial average $\bar{\sigma}(t)$ over time. The pixel/element-wise stress error field

error($\mathbf{x}_I, t$) is simply the difference between the predicted $\sigma(\mathbf{x}_I, t)$ and true $\sigma(\mathbf{x}_I, t)$ stress fields:

$$\text{error}(\mathbf{x}_I, t) = \sigma(\mathbf{x}_I, t) - \breve{\sigma}(\mathbf{x}_I, t) \tag{4}$$

for a particular realization. The root mean square error (RMSE) of the pixel/element-wise stress is given by the ensemble average $\langle \bullet \rangle$ of these errors across all the 3,200 realizations used in testing

$$\text{RMSE}(t) = \sqrt{\left\langle \frac{1}{N} \sum_I \text{error}(\mathbf{x}_I, t)^2 \right\rangle}, \tag{5}$$

where here $N = 32^2$ is the number of pixels/elements. The root mean square error of the spatially averaged stress $\bar{\sigma}(t) = \frac{1}{N} \sum_I \sigma(\mathbf{x}_I, t)$ is

$$\overline{\text{RMSE}}(t) = \sqrt{\left\langle (\bar{\sigma}(t) - \breve{\bar{\sigma}}(t))^2 \right\rangle} \tag{6}$$

Note the total root mean square pixel-wise error summed across pixels I and discrete times t_i

$$\text{TMSE}(t) = \sqrt{\left\langle \frac{1}{NN_t} \sum_i \sum_I \text{error}(\mathbf{x}_I, t_i)^2 \right\rangle}, \tag{7}$$

was used to drive the stochastic descent algorithm training the network, with N_i being the length of the time sequence. (As stated in section 3.2 the final training errors were comparable to the validation errors.)

Figure 6 shows the correlation of the NN predictions with the true/CP stresses and the RMSEs of the normalized stresses over the entire held out testing ensemble of 3,200 samples. The stresses are normalized by their maximum value of over the strain history. Both the correlation and error of the spatial average ($\overline{\text{RMSE}}(t)$ and the full-field (RMSE(t)) stress evolution are shown. The spatial average displays a cancellation of errors and better correlation with data than the full-field but both are good. The ensemble-average stress-strain response with ensemble variance is shown in the top panel to aid comparison with trends in the correlations and errors. The absolute error in the normalized stress and its spread for both the average and the field are worse preceding the elastic-plastic transition, where there is the highest variance in the ensemble response and the most complexity in the local states. Note that the spread of pixel-wise errors is significant; however, the spread for the predictions of the average derived from the full-field predictions appears to be negligible. Interestingly the correlation coefficient was highest around the elastic-plastic transition, which implies model bias in this regime since is where the relative errors are the highest (and that variance in the output aids the correlation). Unlike our previous work [53] which displayed monotonic increase in error with time (with a MLP-based LSTM architecture used to predict the spatial average stress history), the Conv3D structure in the present NN appears to level and reduce the error over the entire history (as well as effectively predicting the evolution of the entire field). However, some artifacts of the zero padding in time can be seen in the error curves in figure 6, for example the kink in correlation and error in at the last time-step, and the predicted (zero load) first stress states are not identically zero (not shown).

We also compared the distributions of per pixel stresses over the ensemble at fixed loading levels. Figure 7 shows the stress distributions in the elastic, elastic-plastic and plastic regimes. Clearly, they are nearly identical. The discrepancy for the elastic-plastic regime is the highest, as expected, since this is the most complex regime of the simulations. The distribution of errors in the normalized stress fields are show in figure 8. The distributions of error across the testing ensemble for fixed times in figure 8(a) exhibit peaked, bell-shaped distributions with slight but noticeable bias to under-prediction of the stress and greatest spread in near the elastic-plastic transition. Both observations are corroborated by the information in figure 6. As indicated in figure 6 the errors are most pronounced near $\varepsilon = 0.07\%$ where figure 8(a) shows tails of the error distribution extending to roughly $\pm\, 0.15$. Figure 8(b) shows the spatial distribution of error at fixed times for one realization, with the true stress fields shown for comparison. It is evident that most of the stress error field is white or near white indicating that the error near zero especially in full plastic flow, and is largest at some but not all of the extreme stress locations near grain boundaries. It appears that the trained filters are more sensitive to the jumps in field values at grain boundaries than the nearly constant field values in the interior of grains.

Figure 9 shows the predictions for the test realizations with the minimum, median, and maximum TMSE (0.016 3, 0.025 8, 0.040, respectively). It demonstrates that the predictions of the normal stress field evolution are high fidelity, with very few visible artifacts in the sequence of images (upper rows: true/CP data, lower rows: predicted/NN). The correspondence is corroborated by the numerical correlation shown in figure 6. As
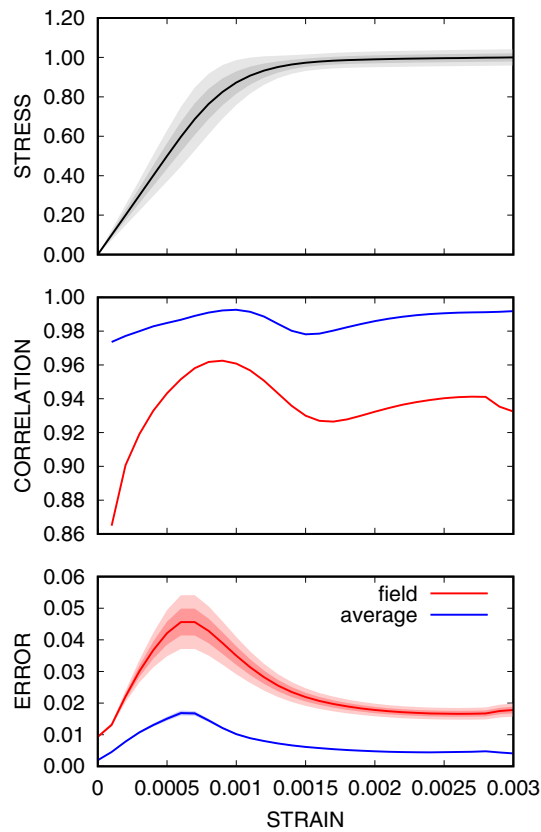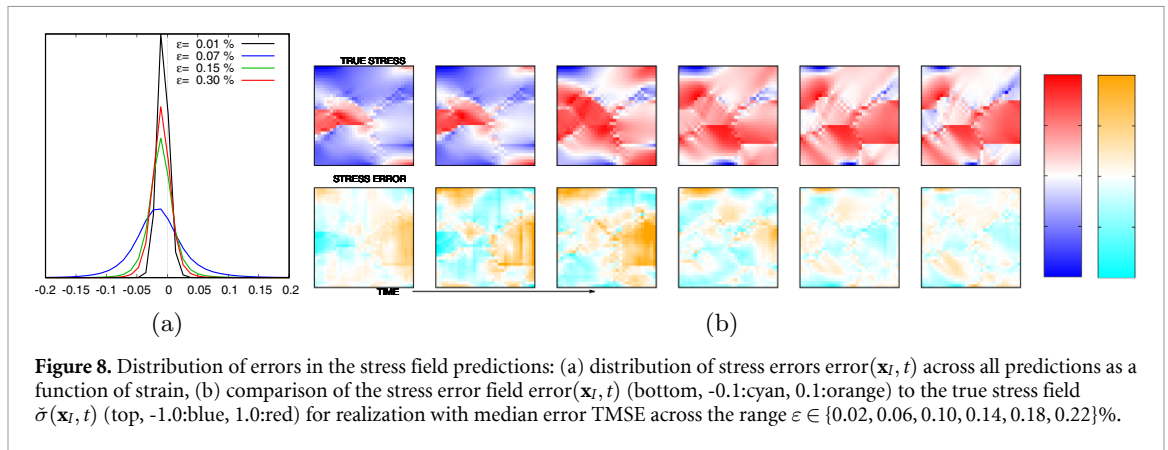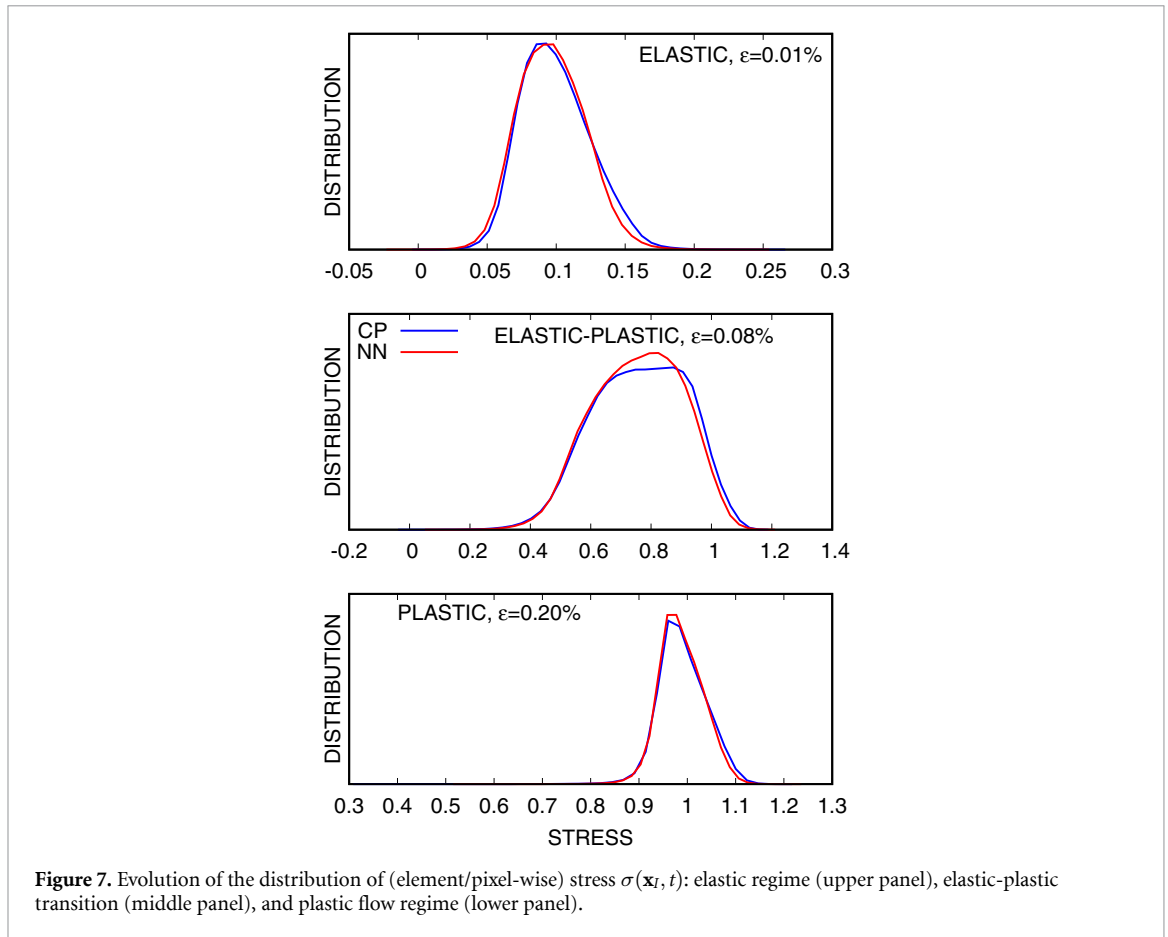
Figure 6. Normalized average stress response $\breve{\bar{\sigma}}(t)$ (upper panel), correlations (middle panel), and root mean squared errors $\overline{\mathrm{RMSE}}(t)$ and $\mathrm{RMSE}(t)$ (lower panel) versus loading strain $\varepsilon(t)$. Upper panel shows the mean response (black line), and one (dark gray) and two (light gray) standard deviations with strain. Middle panel shows the correlation of predictions for full-field (red line) and average (blue line). Lower panel shows the mean squared error of predictions for full-field (red line) and average (blue line) together with one and two standard deviation bands. Note that bands around the error in the average $\overline{\mathrm{RMSE}}(t)$ are barely visible since they are effectively coincident with the blue trend line.

mentioned, the initial prediction at $t = 0$ (not shown for brevity), before any deformation has occurred, is not identically zero; likewise some relative errors are visible even in the best case, figure 9(a). This is a result of the loss function, equation (7), used to train the network being based on absolute errors and hence not as sensitive to errors in small stresses. This behavior could be corrected with an alternative structure for the NN to enforce zero stress at zero strain, or padding the time history with more appropriate values, or minimized with particular weighting or penalty in the loss function to emphasize the early stress states. These errors, as are the errors throughout the history, are arguably negligible since the mean response (predicted/NN (blue), true/CP (red)) lie on top of each other. Both the predicted and true stress fields (blue-white-red image sequences) show stress concentrations in correspondence with the boundaries of the grain structure (color).
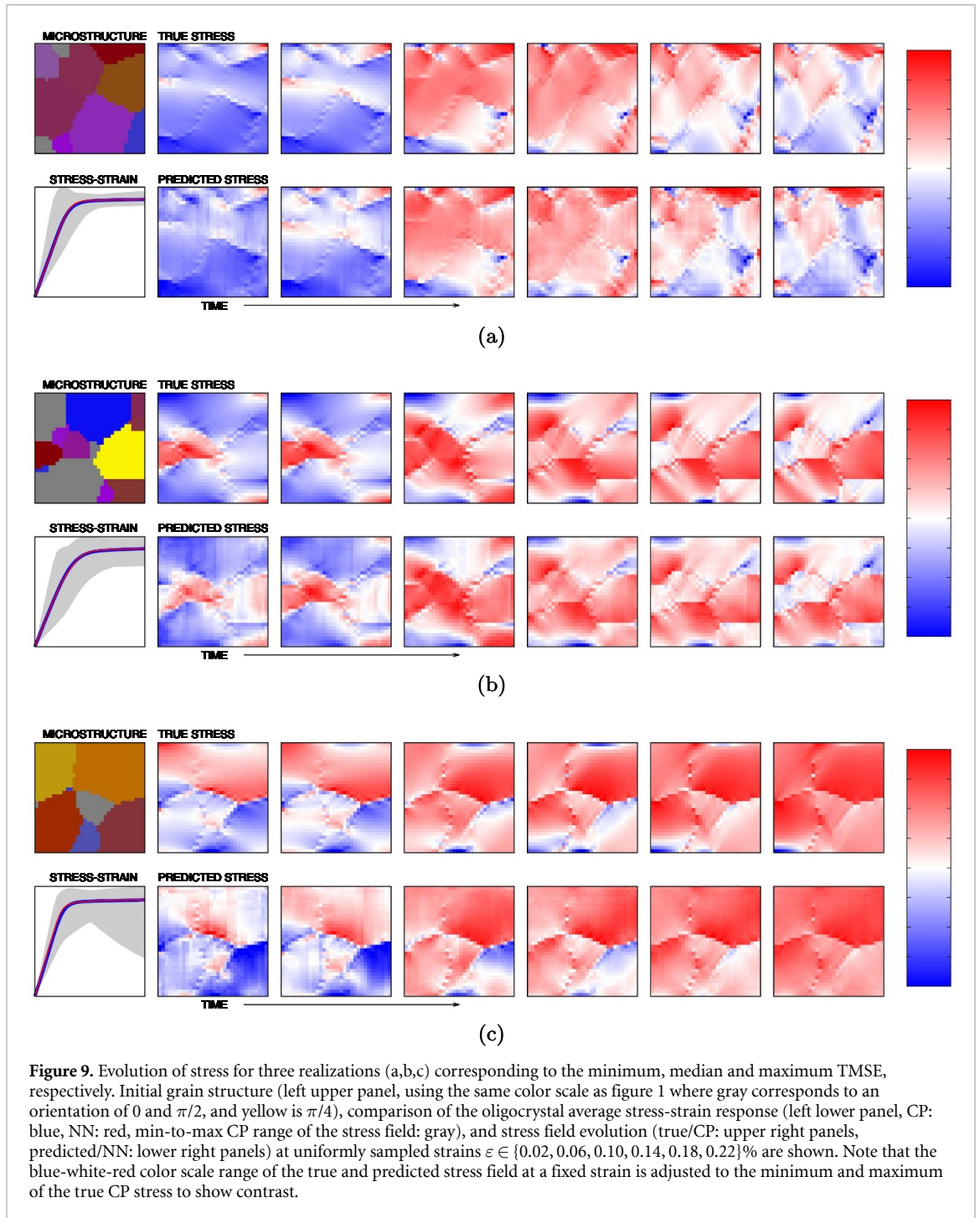
Predicting the stress rates is arguably a more difficult task since they all tend to zero as the plastic flow ensues. Figure 10 shows that the network predictions of the full-field stress rates are in good correlation with the CP data and, given their positivity, satisfy the basic dissipation requirement implied by the second law of thermodynamics. Nevertheless, there are noticeable errors in the average rates early on in the maximum error realization, figure 10(c), and more minor fluctuations in the median and minimum error realizations. This is possibly due to zero strain stress errors previously mentioned and the emphasis on higher stress values in the error norm, TMSE. Spurious undulations in the rate fields are also apparent and to be expected since differences such as time rates tend to have higher errors than the primary field. The errors tend to manifest as striations normal to the loading direction early in the deformation which then become more grid-like as fully plastic flow develops and the stress field becomes nearly constant. This is could be related to by the limited width of the convolution kernels and the fact that the C0 activation functions do not produce entirely smooth outputs so that differences show exacerbated fluctuations. We could ameliorate these errors by adding error in the rates to the loss function used in training, equation (7), and possibly with wider convolution kernels.

These results show that the predictions of mean, distribution, and pixel/element-wise stresses are in good correspondence with the held out true values.

Figure 7. Evolution of the distribution of (element/pixel-wise) stress $\sigma(\mathbf{x}_I, t)$: elastic regime (upper panel), elastic-plastic transition (middle panel), and plastic flow regime (lower panel).

Figure 8. Distribution of errors in the stress field predictions: (a) distribution of stress errors error$(\mathbf{x}_I, t)$ across all predictions as a function of strain, (b) comparison of the stress error field error$(\mathbf{x}_I, t)$ (bottom, -0.1:cyan, 0.1:orange) to the true stress field $\breve{\sigma}(\mathbf{x}_I, t)$ (top, -1.0:blue, 1.0:red) for realization with median error TMSE across the range $\varepsilon \in \{0.02, 0.06, 0.10, 0.14, 0.18, 0.22\}\%$.
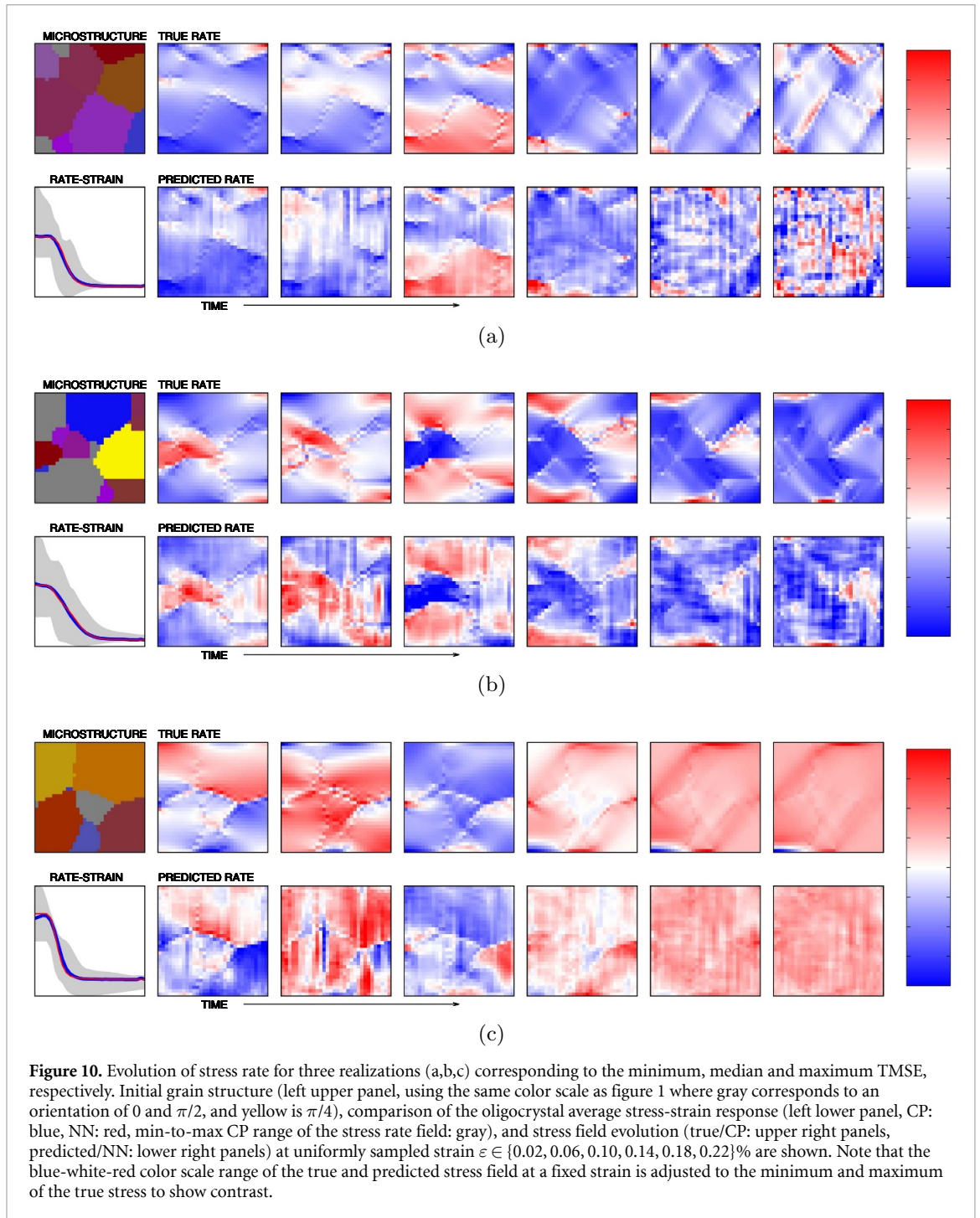
4.2. Structure-property exploration

As a demonstration of the efficacy of using the proposed network model in exploring the effects of microstructure on mechanical performance, we used the NN model to screen 16,000 additional oligocrystals (generated with the same process as for the calibration, refer to section 2) based on their yield strength and the homogeneity of their stress fields. Using thresholds shown in figure 11(a) and (b) for the yield strength and the maximum (over time) of the standard deviation the stress field, we were able to partition the ensemble into high yield structures (11.1%), low variance structures (7.6%) and structures that have both properties (2.4%) in a few minutes. To determine mechanistic sources for these properties, we compared the distributions of the subpopulations to that of the full population for a variety of grain statistics. Figure 11(c) show interesting differences in the crystal textures in the various populations. The high yield structures have a distinct prevalence of grains with $\phi = \pi/4$ orientations such that the slip planes are oblique to the loading direction. The low variance structures have a bimodal distribution of textures with preference for near but off $\phi = \pi/4$ orientations. The distribution of structures with both properties is unimodal but with broader peak than the high yield population. Figure 11(d) shows slight shift to lower grain boundary lengths (per

Figure 9. Evolution of stress for three realizations (a,b,c) corresponding to the minimum, median and maximum TMSE, respectively. Initial grain structure (left upper panel, using the same color scale as figure 1 where gray corresponds to an orientation of 0 and $\pi/2$, and yellow is $\pi/4$), comparison of the oligocrystal average stress-strain response (left lower panel, CP: blue, NN: red, min-to-max CP range of the stress field: gray), and stress field evolution (true/CP: upper right panels, predicted/NN: lower right panels) at uniformly sampled strains $\varepsilon \in \{0.02, 0.06, 0.10, 0.14, 0.18, 0.22\}\%$ are shown. Note that the blue-white-red color scale range of the true and predicted stress field at a fixed strain is adjusted to the minimum and maximum of the true CP stress to show contrast.
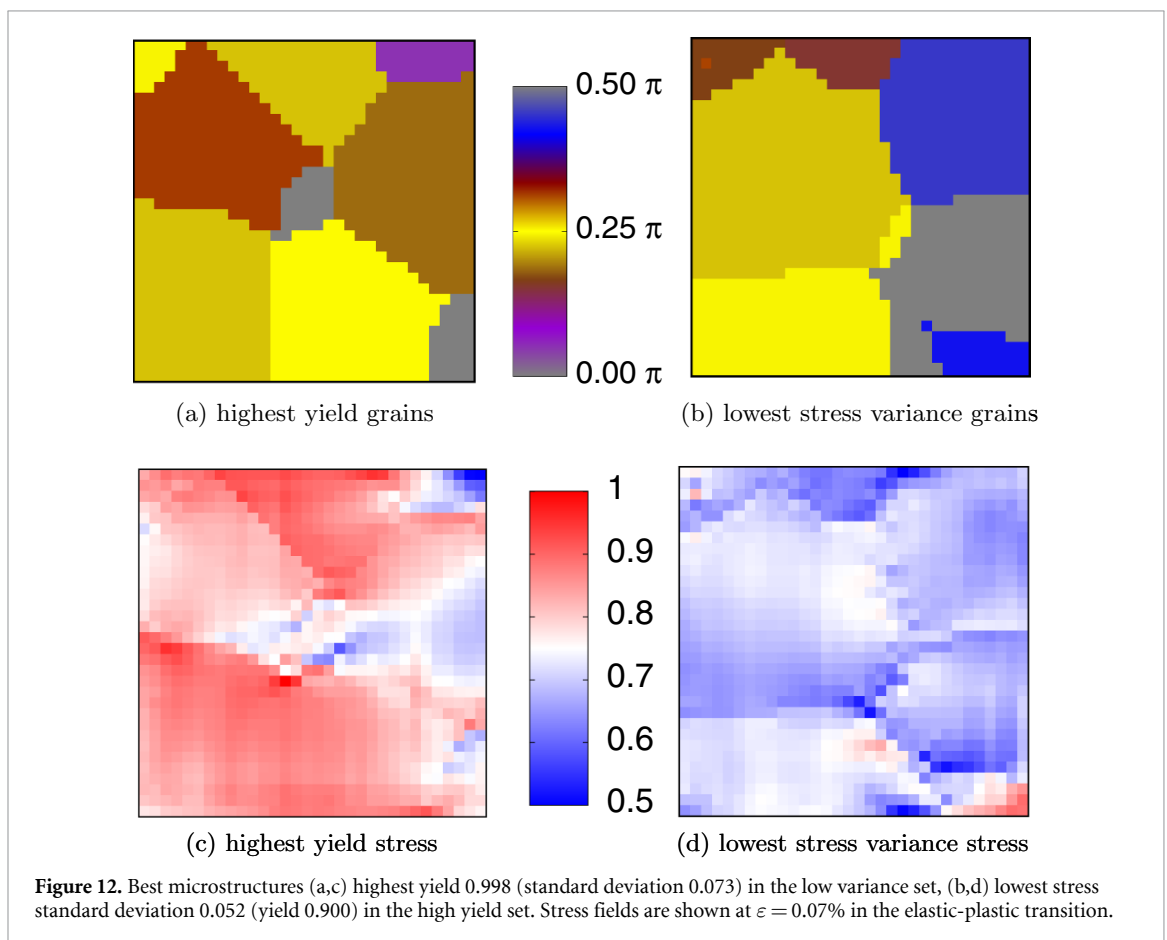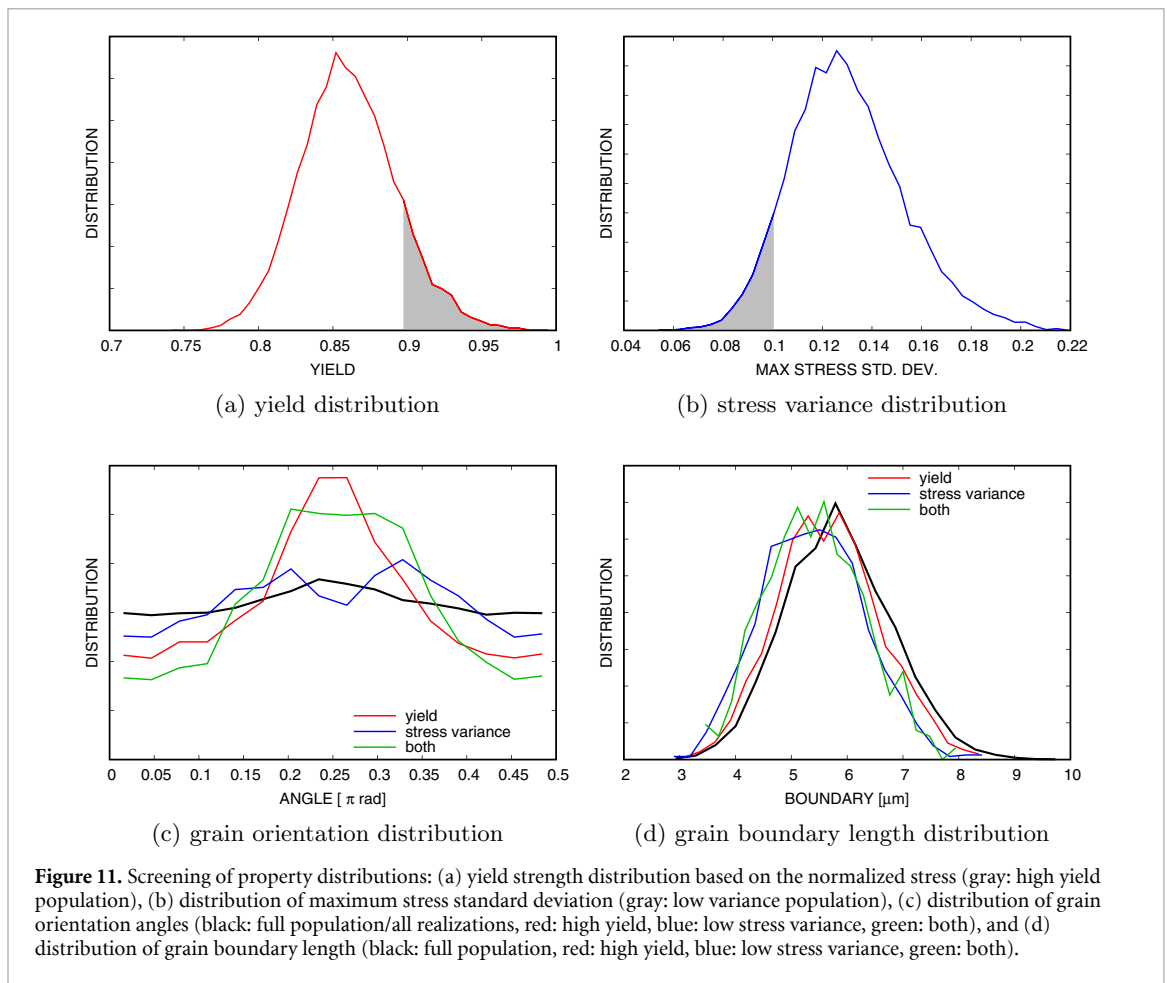
simulation cell) for the superior (high strength and low variance) microstructures. Since the misorientation was calculated such that it scales with boundary length, the distributions for the high yield and low variance populations also shifted to slightly lower values compared to distribution for the full population. We also examined the distributions of grain density and grain size for the high performing structures. They were were essentially equivalent with the distributions for the full population.

Figure 12(a) shows the highest yield microstructure and figure 12(b) the lowest stress variance microstructure; and, figure 12(c) and (d) show the corresponding stress fields in the elastic-plastic transition regime. Both the highest yield microstructure, figure 12(a), and the lowest variance, high yield microstructure, figure 12(b), have high grain-to-grain misorientation. As mentioned in section 2.2, the effective yield strength varies with orientation in this quasi-2D formulation, which makes interpretation of the results in terms of maximum resolved shear stress difficult; however, both oligocrystals have a high percentage of $\phi \approx 0, \pi/2$ grains and these abut $\phi \approx \pi/4$ grains. It is apparent that the highest yield microstructure has a rough central symmetry and the lowest variance, high yield microstructure has loose

Figure 10. Evolution of stress rate for three realizations (a,b,c) corresponding to the minimum, median and maximum TMSE, respectively. Initial grain structure (left upper panel, using the same color as figure 1 where gray corresponds to an orientation of 0 and $\pi/2$, and yellow is $\pi/4$), comparison of the oligocrystal average stress-strain response (left lower panel, CP: blue, NN: red, min-to-max CP range of the stress rate field: gray), and stress field evolution (true/CP: upper right panels, predicted/NN: lower right panels) at uniformly sampled strain $\varepsilon \in \{0.02, 0.06, 0.10, 0.14, 0.18, 0.22\}\%$ are shown. Note that the blue-white-red color scale range of the true and predicted stress field at a fixed strain is adjusted to the minimum and maximum of the true stress to show contrast.

bilateral symmetry. Note that the responses of these microstructures are only marginally different, although it is notable that the microstructure shown in figure 12(a) has the highest yield 0.99, where as the low variance microstructure figure 12(b) had a yield of 0.90 which was on the lower boundary of the selected range. It is also notable that both the microstructures have grain boundary lengths well below the average of the distribution, 4.7 and 4.4 μm, respectively. It is likely that these results are strongly influenced by boundary effects given the correlation length (refer to figure 2(b)) versus the size of the samples [68]. Nevertheless, it is a strength of the machine learning model that it can efficiently find these high scoring oligocrystals in a fixed ensemble of 16,000 microstructures. Also, it should be noted that the subtleties in the results would be unlikely to be uncovered by traditional means, such as exploration based on assumed sensitivities. This demonstrates that connecting this trained network to a generative network capable of creating realistic oligocrystals with the characteristics we created with Dream.3d could find near optimal microstructures in a material discovery effort.

(a) yield distribution

(b) stress variance distribution

(c) grain orientation distribution

(d) grain boundary length distribution

Figure 11. Screening of property distributions: (a) yield strength distribution based on the normalized stress (gray: high yield population), (b) distribution of maximum stress standard deviation (gray: low variance population), (c) distribution of grain orientation angles (black: full population/all realizations, red: high yield, blue: low stress variance, green: both), and (d) distribution of grain boundary length (black: full population, red: high yield, blue: low stress variance, green: both).

(a) highest yield grains

(b) lowest stress variance grains

(c) highest yield stress

(d) lowest stress variance stress

Figure 12. Best microstructures (a,c) highest yield 0.998 (standard deviation 0.073) in the low variance set, (b,d) lowest stress standard deviation 0.052 (yield 0.900) in the high yield set. Stress fields are shown at $\varepsilon = 0.07\%$ in the elastic-plastic transition.

5. Conclusion

Using a neural network architecture based on a ConvLSTM that efficiently encoded the spatial and temporal correlations in the data, we were able to predict the evolution the stress field of oligocrystals with remarkable fidelity using just the initial microstructure and the external loading as inputs. Admittedly there are limitations to the dataset, mostly necessitated by cost, such as the grain boundaries being stair-stepped on a structured grid and the strict geometric compatibility of interactions across grain boundaries; however, taking these as ground truth the ability of the proposed network to represent elastic-plastic process based on these initial structures is remarkable. We also demonstrated that the network model can be used to facilitate material screening, optimization and design tasks. A more sophisticated method would use the generator component of a GAN trained to generate microstructures [43] for material optimization over latent space/manifold embedded in the GAN. In addition to structure-property discovery, the model can also be used as a high efficiency surrogate model for uncertainty quantification and global sensitivity studies. It is worth noting that computing the crystal plasticity training data took approximately 500 processor-days on a cluster, the NN training took 4 processor-days on a GPU, after which tens of thousands of evaluations of the NN model took minutes on a laptop.

Although not as mature recent developments to accelerate crystal plasticity calculations [16, 18, 19], machine learning techniques are promising, probably faster to evaluate once trained, and arguable more widely applicable to mesoscale/microscale modeling. For the particular problem at hand, application of the proposed *evolving, full-field* NN to other loading modes/non-monotonic loading should be straight-forward, albeit with the significant sampling burden quantified in [52] and the path-dependence of plasticity. Likewise, extension to 3D data should be straightforward in that only a ConvLSTM capable of handling 4D data (3 spatial dimensions plus time) is necessary. In fact, Yang *et al* [48] demonstrated that training a CNN based model over a wide variety of loading conditions is possible, at least for the average response of an elastic material. A specific scheme would use architectures like that shown in figure 4 parallel to model the coefficient functions of the natural tensor basis for the stress response, refer to [52, 69]. Also the proposed approach should be effective in predicting the complex evolution other types of microstructures, such as voids, multiple phases, fractures, given appropriate training data and means of encoding them. Given the subtleties in the structure-property demonstration, we will investigate the interpretability [70] of the proposed network which should aid in this endeavor and reveal what hidden physical aspects learned by the network are informing the predictions. This should lead to simpler, more humanly digestible understanding of the complex behavior. We would also like to extend the model to use multi-fidelity training augmented with limited experimental images and mechanical response [71–73]. Another topic of interest is the sensitivity of these types of network to the resolution of the grain structures and their spatial correlation. In addition to the topic of interpretability, the imposition of physical constraints on the models [52, 69] and merging human understanding with machine learning is another topic of active research with significant open questions.

This work demonstrates that the proposed data driven, image based neural network model is effective means of modeling evolving spatio-temporal processes. Unlike previous approaches which have a fixed capacity for accuracy due to pre-selection of specific features and do not intrinsically handle evolving processes, deep learning can reach arbitrary levels of non-linearity and has the potential for learning the key informative features and input-output correlations implicitly for microstructure evolution. We have shown that our neural network model of spatio-temporal evolution can converge to an sufficiently accurate representation of the underlying data-generating process given a large enough dataset and network. Furthermore, the architecture, figure 4, is ambivalent to the data represented in the images of initial microstructure other than this field needs to be correlated with the evolving field output of interest, and, hence, the proposed network is adaptable to a wide class of physical evolution problems where an initial field strongly influences the evolution of another field of interest. In addition the architecture is directly extensible by adding input channels and filters to handling more than one type of image simultaneously so that the effects of porosity on crystal plastic behavior [74], for instance, can be modelled. To our knowledge no other broadly applicable, data-driven method has demonstrated the capability to predict complex evolving microstructural fields over space and time with comparable accuracy and the model trained quickly on a relatively small dataset (considering other deep learning models and the complexity of the task at hand). This method can enable high-speed model querying and uncertainty quantification as well as provide a model of simulation-experiment discrepancy.

Data availability

The data that support the findings of this study will be openly available [75] following a delay.

Acknowledgments

ORCID iD

Reese Jones ⬤ https://orcid.org/0000-0002-2332-6279

References

[1] Taylor G I 1934 The mechanism of plastic deformation of crystals Part I Theoretical *Proc. R. Soc. London* A **145** 362–87

[2] Kroner E 1961 On the plastic deformation of polycrystals *Acta. Metall.* **9** 155–61

[3] Bishop J F W and Hill R 1951 XLVI. A theory of the plastic distortion of a polycrystalline aggregate under combined stresses *London Edinburgh Dublin Phil. Mag. J. Sci.* **42** 414–27

[4] Bishop J F W and Hill R 1951 CXXVIII. A theoretical derivation of the plastic properties of a polycrystalline face-centred metal *London Edinburgh Dublin Phil. Mag. J. Sci.* **42** 1298–307

[5] Mandel J 1965 Généralisation de la théorie de plasticité de WT Koiter *Int. J. Solids Struct.* **1** 273–95

[6] Dawson P R 2000 Computational crystal plasticity *Int. J. Solids Struct.* **37** 115–30

[7] Roters F, Eisenlohr P, Hantcherli L, Tjahjanto D D, Bieler T R and Raabe D 2010 Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications *Acta. Mater.* **58** 1152–1211

[8] Hill R 1998 *The Mathematical Theory of Plasticity* (Oxford: Oxford University Press)

[9] Lubliner J 2008 *Plasticity Theory* (Mineola, NY: Dover Publications)

[10] Hill R 1963 Elastic properties of reinforced solids: some theoretical principles *J. Mech. Phys. Solids* **11** 357–72

[11] Drugan W J and Willis J R 1996 A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites *J. Mech. Phys. Solids* **44** 497–524

[12] Ostoja-Starzewski M 2006 Material spatial randomness: From statistical to representative volume element *Probab. Eng. Mech.* **21** 112–32

[13] Lebensohn R A, Rollett A D and Suquet P 2011 Fast Fourier transform-based modeling for the determination of micromechanical fields in polycrystals *JOM* **63** 13–18

[14] Knezevic M, Kalidindi S R and Fullwood D 2008 Computationally efficient database and spectral interpolation for fully plastic Taylor-type crystal plasticity calculations of face-centered cubic polycrystals *Int. J. Plast.* **24** 1264–76

[15] Knezevic M, Al-Harbi H F and Kalidindi S R 2009 Crystal plasticity simulations using discrete Fourier transforms *Acta. Mater.* **57** 1777–84

[16] Zecevic M, McCabe R J and Knezevic M 2015 A new implementation of the spectral crystal plasticity framework in implicit finite elements *Mech. Mater.* **84** 114–26

[17] Azdoud Y, Cheng J and Ghosh S 2017 Wavelet-enriched adaptive crystal plasticity finite element model for polycrystalline microstructures *Comput. Methods Appl. Mech. Eng.* **327** 36–57

[18] Eghtesad A, Zecevic M, Lebensohn R A, McCabe R J and Knezevic M 2018 Spectral database constitutive representation within a spectral micromechanical solver for computationally efficient polycrystal plasticity modelling *Computat. Mech.* **61** 89–104

[19] Gupta A, Bettaieb M B, Abed-Meraim F and Kalidindi S R 2018 Computationally efficient predictions of crystal plasticity based forming limit diagrams using a spectral database *Int. J. Plast.* **103** 168–87

[20] Liu Z, Fleming M and Liu W K 2018 Microstructural material database for self-consistent clustering analysis of elastoplastic strain softening materials *Comput. Methods Appl. Mech. Eng.* **330** 547–77

[21] Montes de Oca Zapiain D, Popova E and Kalidindi S R 2017 Prediction of microscale plastic strain rate fields in two-phase composites subjected to an arbitrary macroscale strain rate using the materials knowledge system framework *Acta. Mater.* **141** 230–40

[22] Montes de Oca Zapiain D and Kalidindi S R 2019 Localization models for the plastic response of polycrystalline materials using the material knowledge systems framework *Modelling Simul. Mater. Sci. Eng.* **27** 074008

[23] Hastie T, Tibshirani R, Friedman J and Franklin J 2005 The elements of statistical learning: data mining, inference and prediction *Math. Intelligencer* **27** 83–5

[24] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (Cambridge, MA: MIT press)

[25] Liu Y, Zhao T, Wangwei J and Shi S 2017 Materials discovery and design using machine learning *Journal of Materiomics* **3** 159–77

[26] Dimiduk D M, Holm E A and Niezgoda S R 2018 Perspectives on the impact of machine learning, deep learning and artificial intelligence on materials, processes and structures engineering *Integrating Materials and Manufacturing Innovation* **7** 157–72

[27] Bostanabad R, Zhang Y, Xiaolin Li, Kearney T, Catherine Brinson L, Apley D W, Liu W K and Chen W 2018 Computational microstructure characterization and reconstruction: Review of the state-of-the-art techniques *Prog. Mater. Sci.* **95** 1–41

[28] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in Neural Information Processing Systems* pp 1097–105

[29] Gatys L, Ecker A S and Bethge M 2015 Texture synthesis using convolutional neural networks *Advances in Neural Information Processing Systems 28 (NIPS 2015)* pp 262–70

[30] Lubbers N, Lookman T and Barros K 2017 Inferring low-dimensional microstructure representations using convolutional neural networks *Phys. Rev.* E **96** 052111

[31] Noh J, Kim J, Stein H S, Sanchez-Lengeling B, Gregoire J M, Aspuru-Guzik A and Jung Y 2019 Inverse design of solid-state materials via a continuous representation *Matter* **1** 1370–84

[32] Faber F, Lindmaa A, Anatole von Lilienfeld O and Armiento R 2015 Crystal structure representations for machine learning models of formation energies *Int. J. Quantum Chem.* **115** 1094–101

[33] Xiaolin Li, Zhang Y, Zhao H, Burkhart C, Catherine Brinson L and Chen W 2018 A transfer learning approach for microstructure reconstruction and structure-property predictions *Sci. Rep.* **8** 1–13

[34] Papanikolaou S, Tzimas M, Song H, Reid A C E and Langer S A 2017 Learning crystal plasticity using digital image correlation: Examples from discrete dislocation dynamics Arxiv:1709.08225

[35] Chowdhury A, Kautz E, Yener Bulent and Lewis D 2016 Image driven machine learning methods for microstructure recognition *Comput. Mater. Sci.* **123** 176–87

[36] Hongyi X, Liu R, Choudhary A and Chen W 2015 A machine learning-based design representation method for designing heterogeneous microstructures *J. Mech. Design* **137** 051403

[37] Bessa M A, Bostanabad R, Liu Z, Hu A, Apley D W, Brinson C, Chen W and Liu W K 2017 A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality *Comput. Methods Appl. Mech. Eng.* **320** 633–67

[38] Bostanabad R, Bui A T, Xie W, Apley D W and Chen W 2016 Stochastic microstructure characterization and reconstruction via supervised learning *Acta. Mater.* **103** 89–102

[39] Bostanabad R, Chen W and Apley D W 2016 Characterization and reconstruction of 3D stochastic microstructures via supervised learning *J. Microsc.* **264** 282–97

[40] Kingma D P and Welling M 2013 Auto-encoding variational Bayes Arxiv:1312.6114

[41] Goodfellow I, Pouget-Abadie J, Mirza M, Bing X, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets *Advances in Neural Information Processing Systems* pp 2672–80

[42] Cang R, Hechao Li, Yao H, Jiao Y and Ren Y 2018 Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model *Comput. Mater. Sci.* **150** 212–21

[43] Yang Z, Xiaolin Li, Catherine Brinson L, Choudhary A N, Chen W and Agrawal A 2018 Microstructural materials design via deep adversarial learning methodology *J. Mech. Design* **140** 111416

[44] Deshpande P D, Gautham B P, Cecen A, Kalidindi S, Agrawal A and Choudhary A 2013 Application of statistical and machine learning techniques for correlating properties to composition and manufacturing processes of steels *Proc. of the 2nd World Congress on Integrated Computational Materials Engineering (ICME)* (Berlin: Springer) pp 155–60

[45] Yabansu Y C, Steinmetz P, Kalidindi S R and Nestler B 2017 Extraction of reduced-order process-structure linkages from phase-field simulations *Acta. Mater.* **124** 182–94

[46] Cecen A, Dai H, Yabansu Y C, Kalidindi S R and Song L 2018 Material structure-property linkages using three-dimensional convolutional neural networks *Acta. Mater.* **146** 76–84

[47] Yang Z, Yabansu Y C, Al-Bahrani R, Liao W-keng, Choudhary A N, Kalidindi S R and Agrawal A 2018 Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets *Comput. Mater. Sci.* **151** 278–87

[48] Yang Z, Yabansu Y C, Jha D, Liao W-keng, Choudhary A N, Kalidindi S R and Agrawal A 2019 Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches *Acta. Mater.* **166** 335–45

[49] Yuan M, Paradiso S, Meredig B and Niezgoda S R 2018 Machine learning–based reduce order crystal plasticity modeling for ICME applications *Integrating Mater. Manufacturing Innovation* **7** 214–30

[50] Niezgoda S R, Kanjarla A K and Kalidindi S R 2013 Novel microstructure quantification framework for databasing, visualization and analysis of microstructure data *Integrating Mater. Manufacturing Innovation* **2** 54–80

[51] Mangal A and Holm E A 2018 Applied machine learning to predict stress hotspots I: Face centered cubic materials *Int. J. Plast.* **111** 122–34

[52] Jones R E, Templeton J A, Sanders C M and Ostien J T 2018 Machine learning models of plastic flow based on representation theory *Comput. Model. Eng. Sci.* **117** 309–42

[53] Frankel A L, Jones R E, Alleman C and Templeton J A 2019 Predicting the mechanical response of oligocrystals with deep learning *Comput. Mater. Sci.* **169** 109099

[54] Xingjian S H I, Chen Z, Wang H, Yeung D-Y, Wong W-K and Woo W-chun 2015 Convolutional LSTM network: A machine learning approach for precipitation nowcasting *Advances in Neural Information Processing Systems* pp 802–10

[55] Cybenko G 1989 Approximation by superpositions of a sigmoidal function *Math. Control Signals Syst.* **2** 303–14

[56] Hornik K 1991 Approximation capabilities of multilayer feedforward networks *Neural Netw.* **4** 251–7

[57] Groeber M A and Jackson M A 2014 Dream.3D: A digital representation environment for the analysis of microstructure in 3D *Integrat. Mater. Manuf. Innov.* **3** 5

[58] Ledbetter H M 1984 Monocrystal-polycrystal elastic constants of a stainless steel *Phys. Status Solidi (a)* **85** 89–96

[59] Kocks U F 1976 Laws for work-hardening and low-temperature creep *J. Eng. Mater. Technol.* **98** 76–85

[60] Mecking H, Kocks U F and Fischer H 1976 Hardening, recovery and creep in fcc mono-and polycrystals *Presented at the 4th Intern. Conf. on Strength of Metals and Alloys, Nancy, 30 Aug.-3 Sep. 1976*

[61] Salinger A G *et al* 2016 Albany: using component-based design to develop a flexible, generic multiphysics analysis code *Int. J. Multiscale Computational Engineering* **14** 415–38

[62] Salzbrenner B C, Rodelas J M, Madison J D, Jared B H, Swiler L P, Shen Y-L and Boyce B L 2017 High-throughput stochastic tensile performance of additively manufactured stainless steel *J. Mater. Process. Technol.* **241** 1–12

[63] Rosenblatt F 1961 Principles of neurodynamics. perceptrons and the theory of brain mechanisms *Technical Report* Cornell Aeronautical Lab Inc., Buffalo NY

[64] Hochreiter S and Schmidhuber Jurgen 1997 Long short-term memory *Neural Comput.* **9** 1735–80

[65] Resteghini L, Lanzi P L, Nebuloni R, Riva C, Capsoni C and Gabellini P 2013 Single-objective genetic algorithm for dynamic optimization of reconfigurable antenna systems *Antennas and Propagation (EuCAP), 2013 7th Conf. on* IEEE 1333–5

[66] Abadi M *et al* 2015 TensorFlow: Large-scale machine learning on heterogeneous systems Software available from tensorflow.org

[67] Kingma D P and Jimmy B 2014 Adam: A method for stochastic optimization Arxiv:1412.6980

[68] Uchic M D, Dimiduk D M, Florando J N and Nix W D 2004 Sample dimensions influence strength and crystal plasticity *Science* **305** 986–9

[69] Ling J, Jones R and Templeton J 2016 Machine learning strategies for systems with invariance properties *J. Comput. Phys.* **318** 22–35

[70] Mengnan D, Liu N and Xia H 2018 Techniques for interpretable machine learning Arxiv:1808.00033
[71] Lebensohn R A, Brenner R, Castelnau O and Rollett A D 2008 Orientation image-based micromechanical modelling of subgrain texture evolution in polycrystalline copper *Acta. Mater.* **56** 3914–26
[72] Lim H, Carroll J D, Battaile C C, Buchheit T E, Boyce B L and Weinberger C R 2014 Grain-scale experimental validation of crystal plasticity finite element simulations of tantalum oligocrystals *Int. J. Plast.* **60** 1–18
[73] Hansong J, Song Q, Gupta M K and Liu Z 2020 A pseudorandom based crystal plasticity finite element method for grain scale polycrystalline material modeling *Mech. Mater.* **144** 103347
[74] Srivastava A and Needleman A 2013 Void growth versus void collapse in a creeping single crystal *J. Mech. Phys. Solids* **61** 1169–84
[75] Talirz L *et al* 2020 Materials Cloud. www.materialscloud.org