Modelling and Simulation in
Materials Science and Engineering

PAPER • **OPEN ACCESS**

Combining simulation and experimental data via surrogate modelling of continuum dislocation dynamics simulations

To cite this article: Balduin Katzer *et al* 2024 *Modelling Simul. Mater. Sci. Eng.* **32** 055026

OPEN ACCESS
IOP Publishing
Modelling and Simulation in Materials Science and Engineering

Modelling Simul. Mater. Sci. Eng. **32** (2024) 055026 (31pp) https://doi.org/10.1088/1361-651X/ad4b4c

Combining simulation and experimental data via surrogate modelling of continuum dislocation dynamics simulations

Balduin Katzer[1,2], **Daniel Betsche**[3], **Felix von Hoegen**[1], **Benjamin Jochum**[3], **Klemens Böhm**[3] and **Katrin Schulz**[1,2,*]

[1] Karlsruhe Institute of Technology (KIT), Institute for Applied Materials (IAM), Kaiserstr. 12, 76131 Karlsruhe, Germany
[2] Karlsruhe University of Applied Sciences (HKA), Moltkestr. 30, 76133 Karlsruhe, Germany
[3] Karlsruhe Institute of Technology (KIT), Institute for Program Structures and Data Organization (IPD), Am Fasanengarten 5, 76131 Karlsruhe, Germany

E-mail: katrin.schulz@kit.edu

Received 20 March 2024; revised 30 April 2024
Accepted for publication 14 May 2024
Published 4 June 2024

Abstract

Several computational models have been introduced in recent years to yield comprehensive insights into microstructural evolution analyses. However, the identification of the correct input parameters to a simulation that corresponds to a certain experimental result is a major challenge on this length scale. To complement simulation results with experimental data (and vice versa) is not trivial since, e.g. simulation model parameters might lack a physical understanding or uncertainties in the experimental data are neglected. Computational costs are another challenge mesoscale models always have to face, so comprehensive parameter studies can be costly. In this paper, we introduce a surrogate model to circumvent continuum dislocation dynamics simulation by a data-driven linkage between well-defined input parameters and output data and vice versa. We present meaningful results for a forward surrogate formulation that predicts simulation output based on the input parameter space, as well as for the inverse approach that derives the input parameter space based on simulation as well as experimental output quantities. This enables, e.g. a direct derivation of the

---

[*] Author to whom any correspondence should be addressed.

input parameter space of a continuum dislocation dynamics simulation based on experimentally provided stress–strain data.

Keywords: plasticity, continuum dislocation dynamics, surrogate modelling, inverse modelling, microstructure prediction, machine learning, data science

1. Introduction

Plasticity in crystalline materials is a result of its underlying microstructure evolution. Plastic deformation arises from dislocation motion and dislocation interaction [1, 2]. Various computational models have been established to study dislocation-based plasticity at different length scales, such as molecular dynamics (MD), discrete dislocation dynamics (DDD), continuum dislocation dynamics (CDD) and crystal plasticity (CP) [3]. Besides comparing different length and time scales with each other or combining them in a multi-scale model, the comparability between simulation and experiment is of essential importance [4]. Experiments yield essential input parameters for simulations on the meso- and macroscale. They have to be taken into account to validate simulation results in order to apply simulations to predictive analyses or enable simulation-assisted interpretation of experiments.

Figure 1 shows a common process of combining simulation and experiment. One derives the input for a simulation from experiments, conducts the simulation and validates the output of a simulation again with experiments [5–7]. A challenge in this procedure is that often several input parameters for the simulation are unknown and have to be estimated since they cannot be derived from the experiment itself or from additional experimental methods [5]. This is, e.g. the case for microstructural parameters such as the initial dislocation microstructure or dislocation reaction coefficients. Thus, the identification of the correct and distinct input for a simulation of the microstructure evolution from experiments is a major challenge [6]. Although, advanced experimental methods have been developed such as transmission electron microscopy, electron backscatter diffraction (EBSD) or digital image correlation (DIC), the methods are often limited to a two dimensional representation of the microstructure [5, 8, 9]. Thus, simulations are necessary to capture the three dimensional microstructure evolution. Ultimately, complementing simulation and experimental data leads to insights into the fundamental material behavior.

In materials mechanics, a simple and self-evident way to set up a simulation is to compare the mechanical behavior based on the stress–strain curve between simulation and experiment. Since each elastic-plastic stress–strain curve arises from its underlying microstructure evolution, the correct stress–strain behavior indicates that a possibly correct microstructure representation has been processed in the simulation. However, there might be several microstructural representations that yield the same overall stress–strain behavior and it is not clear which one is the correct one with respect to the considered experiment.

The stress–strain behavior depends on the initial dislocation microstructure and the applied boundary conditions. In simulations, further numerical parameters are required, while boundary conditions can usually be set in good agreement with an experiment for simple loading scenarios, e.g. tension, compression or bending [10]. Initial conditions and simulation parameters for a specific specimen to be experimentally tested can not be extracted directly in most cases [11]. Statistical scattering of parameters has to be taken into account. Often, the choice of numerical parameters is not explicitly discussed in many publications, which limits the understanding of their impact. Thus, simulation domain knowledge and parameter studies are employed to approach the desired stress–strain curve. This leads to time-consuming trial and error approaches without guarantee of completeness and undetermined conclusions.

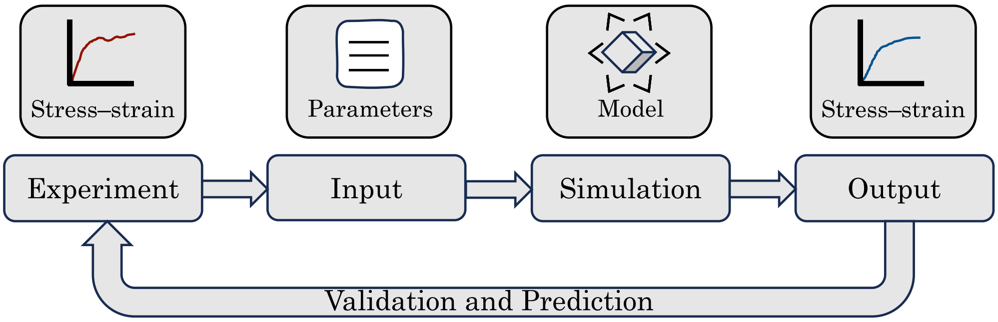


Figure 1. Process of validation of simulations and predictive analysis.

In recent years, statistical and machine learning approaches yielded strategies to systematically identify fundamental material parameters and properties. Various strategies for materials design from simulation or experimental databases are applicable [12, 13]. An overview of machine learning and data mining methods applied in material science is given by Bock *et al* [14]. An overview on more sophisticated deep learning methods is provided by Choudhary *et al* [15]. Frequently, optimization concepts are utilized to iteratively achieve a certain result. For dislocaiton dynamics, in 2D DDD simulations, e.g. a machine learning approach is used by Barros de Moraes *et al* [16] to estimate parameters of the probability density function of a transport model for dislocations. In 3D DDD simulations, Steinberger *et al* [17] evaluate the relevance of different dislocation densities for dislocation density evolution using a Bayesian algorithm. For multiscale crystal plasticity models, Kuhn *et al* [18] use a Bayesian optimisation based on a Gaussian process to inversely identify material parameters.

In parallel to optimization methods, surrogate modelling has been of growing interest to reduce the computational cost by deriving an approximation of the result space [19]. Alizadeh *et al* [20] provide a study of common approaches to build surrogate models for simulations in a variety of physical applications. In a review of more than 200 papers, they order them by methods, and compare the accuracy and computational costs of the surrogate models and their corresponding simulations. A surrogate model is a ‘stand-alone’ model that mimics a simulation and directly interconnects input and output on a data-driven manner. Such a surrogate model concept predicts a simulation output from a simulation input, which is called ‘forward surrogate model’, or the model can predict the simulation input from the simulation output, which is called ‘inverse surrogate model’ [21].

Forward surrogate models are applicable, for example, in multiscale materials modelling to inform models on a larger length scale by surrogate models based on a smaller length scale [22, 23]. Fernandez *et al* [24] apply an artificial neural network (ANN) surrogate to model the constitutive behavior of grain boundaries by using a MD database. Mudunuru *et al* [25] show that the estimation of failure in brittle materials is several orders of magnitude faster by a k-nearest neighbor (kNN) surrogate model compared to high-fidelity simulations. Liang *et al* [26] propose an encoder-decoder surrogate model to directly predict stress distributions on the material surface based on finite element analysis. A long short-term memory (LSTM) surrogate model is proposed, e.g. to predict crystal orientations in crystal plasticity simulations using fcc materials under uniaxial tension [27] or to predict phase-field simulations based microstructure evolution [28].

Inverse surrogate models can approximate the solution of inverse problems. Ktari *et al* [29] identify material parameters based on force-displacement curves by an inverse ANN surrogate model. The applicability to solve inverse problems has been recently demonstrated based on physics-informed neural networks that take into account underlying physical equations to the loss function of the considered neural network [30–33]. Parameter identification by connecting simulation and experiment results is subject of ongoing research [34–36]. A prominent example for the fusion of simulation and experiment is showcased for the plastic deformation of indentation by a neural network based parameter identification from finite element simulations [37] and experimental validation of the parameters [38]. Recently, frameworks for forward and inverse predictions between microstructure properties and mechanical behavior from simulative and experimental data are demonstrated, e.g. for indentation of metals based on a neural network trained on finite element simulations [39] or for compression of architectured materials based on a LSTM trained on MD simulations [40].

Taking into account the recent progress in forward and inverse surrogate modelling, in this paper, we want to address the fundamental challenge of coupling experimental and simulation approaches in dislocation based plasticity. We want to answer the question whether there is a meaningful approach to estimate the input parameter space for continuum dislocation dynamics simulations based on easily accessible experimental data? It is investigated to what extent a surrogate model can represent the microstructural evolution in continuum dislocation dynamics without oversimplifying the complex physical interactions.

The paper is structured as follows: In section 2, we introduce the methodology of the considered continuum dislocation dynamics formulation and propose a concept of a coupled forward and inverse surrogate model. We demonstrate the generation of the simulation database and describe the considered parameter space. We formulate potential surrogate model candidates and metrics for comparing their performance. In section 3, the results of the comparison of the surrogate model candidates are presented. We demonstrate the predictive capability of a forward surrogate model as well as of the inverse formulation by identifying input parameter spaces based on experimental stress–strain curves. Finally, we discuss the sensitivity of the input parameters on the mechanical behavior and its implications for the generation of a surrogate model in section 4. Here, we also discuss uncertainties and limitations of the presented surrogate concept. Section 5 gives a conclusion and outlook.

2. Methods

2.1. Methodology

In this work, we present a surrogate model concept for a dislocation based continuum (CDD) model of crystal plasticity simulation. Figure 2 depicts the concept of coupling a forward surrogate model and an inverse surrogate model approach for CDD simulation. Here, the forward surrogate model aims to map the simulation input to a mechanical behavior. The inverse surrogate model aims to map an arbitrary mechanical behavior to the corresponding simulation input parameter space.

The following three key components of the surrogate approach will be addressed: (i) generation of a simulation database of CDD simulations representing statistical parameter spaces; (ii) evaluation of the possibilities to circumvent the CDD simulation by a forward surrogate model; (iii) potential of an inverse surrogate model, that enables the prediction of input parameter spaces, e.g. based on stress–strain curves. These components are summarized in figure 3 and the respective procedures are outlined in more detail in the following.

(i) First, we generate a representative database based on CDD simulations. As simulation input, we systematically vary a set of input parameters within a chosen parameter space. The

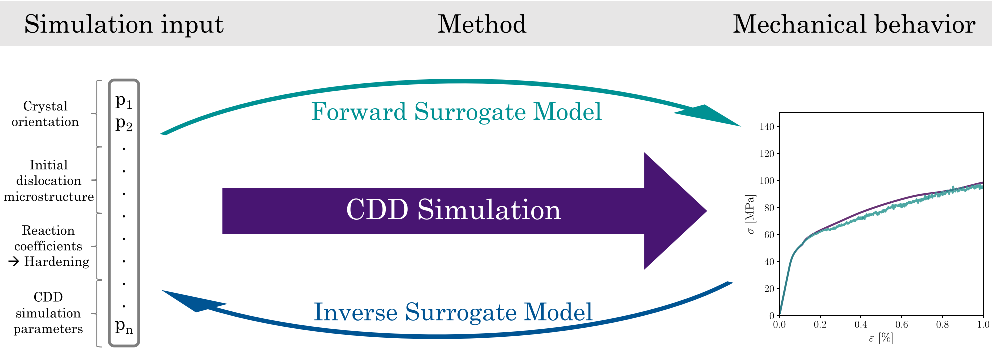


Figure 2. Concept of the surrogate model approach including forward and inverse surrogate for CDD modelling.

Simulation Database

Continuum Dislocation Dynamics (CDD) simulations

Eleven different parameters → Parameter space $P \in \mathbb{R}^{11}$

5832 tensile test simulations with varying initial parameters

Forward Surrogate Model

Prediction of mech. behavior (output) based on initial parameters (input)

Bypassing CDD simulation

Training and evaluation of various surrogate model candidates based on CDD simulation database

Inverse Surrogate Model

Prediction of init. parameters (output) based on mechanical behavior (input)

Estimation of initial parameter probability using cross-validation

Estimation of initial parameters from experimental mechanical behavior

Figure 3. Components of the coupled forward and inverse surrogate approach.

parameter space is described in detail in 2.3. In this work, we include various parameters such as crystal orientation, elastic properties, initial dislocation microstructure, reaction constants as well as numerical simulation parameters. For simplicity, we do not vary the loading scenario, the specimen size and the boundary conditions. As simulation output, we obtain the mechanical behavior in terms of a stress–strain curve as well as the dislocation density evolution.

(ii) Second, a forward surrogate model is trained based on the generated simulation database with the objective to investigate the potential of a surrogate to mimic CDD on a data-driven manner. The forward surrogate model circumvents CDD and predicts a mechanical behaviour (output) based on a well-defined set of simulation input parameters (input). Different surrogate model candidates are evaluated to discuss the suitability for this use case. Different aspects are taken into account such as the accuracy, the amount of data needed and the computational costs. Furthermore the interpolation capability within the parameter space as well as the feature importance is studied for the forward surrogate model.

(iii) Third, we introduce an inverse surrogate model for the prediction of the simulation input parameter space (output) from stress–strain curves (input). We identify most suitable model candidates and an appropriate evaluation metric. We evaluate the limits of the inverse prediction of input parameters based on cross-validation to provide not only the prediction of the best input parameters, but a probability estimation of the input parameter space. The identification of simulation input parameters from experimental stress–strain curves is chosen to discuss the approach.

2.2. Continuum dislocation dynamics

The CDD formulation for fcc crystalline materials employed in this study is based on the framework presented in [41]. In general, the CDD model is a direct homogenization of three-dimensional DDD for fcc crystals [42–44]. The formulation incorporates two interlinked problems: the elastic (external) problem that computes the elastic stress field for a given deformation state based on a discontinuous Galerkin method, and the internal problem that describes the microstructure evolution under a given stress field leading to plastic deformation. The two problems are coupled by the plastic shear strain. For the internal problem, the dislocation microstructure is characterized by dislocation densities on each slip system, specifying the dislocation line length per averaging volume as well as the dislocation curvature density. Following the approaches outlined by Sudmanns *et al* [43], we distinguish between mobile and network dislocation densities. The mobile dislocation density can be decomposed again into statistically stored dislocation density and geometrically necessary dislocation (GND) densities accounting for screw and edge character. The network dislocation density comprises the density of Lomer junctions and a stabilized dislocation density that captures the dislocation density attached to Lomer junctions. The model considers dislocation nucleation [44], dislocation multiplication processes such as glissile reaction and cross-slip [42], as well as dislocation stabilization and annihilation, such as Lomer and collinear reactions [43]. The governing equations including stress interaction and reaction terms are given in appendix A. Regarding the physical consistency, the considered CDD model has been evaluated by comparison to three-dimensional DDD simulations [42–44] as well as to experiments [44–46] showing physically meaningful microstructural evolution results.

2.3. Database

Using the three-dimensional CDD formulation, we conducted simulations with varying input parameters of an uniaxial tensile test of a cubic $(5\ \mu\mathrm{m})^3$ specimen mimicking a face centered cubic single crystal. To ensure the numerical reliability of the simulation results, we conducted temporal and spatial convergence studies. The simulation is carried out with a constant strain rate of $\dot{\varepsilon} = 5000\ \mathrm{s}^{-1}$ up to a total strain $\varepsilon = 1.0\%$. The database is set up consisting of simulation inputs and corresponding simulation outputs. The input is the set of simulation input parameters. The output is a time series, i.e. a data series at certain strain states. Here, we exemplarily choose two data series: stress over strain data and dislocation density over strain data.

2.3.1. Parameter space. For the generation of the simulation database, we focus on a selection of parameters that are varied with respect to the CDD model. We examine eleven input parameters considering the elastic material properties and crystal orientation, dislocation microstructure as well as reaction and interaction constants that are varied as shown and specified in table 1. The variation leads to an overall number of 5832 simulations of permutated input parameters. Values that are varied exclusively as tuples are described by the tuple specifications and are shown in brackets () in table 1. The parameter space is described in more detail in the following:

Material We apply the elastic properties of three materials (Al, Au, Cu) with isotropic material properties. Each material is defined by the Young's modulus E, the Poisson's ratio ν and the Burgers vector b.

Table 1. Parameter space of the simulation input parameters considered in CDD. The values that are exclusively varied as tuples, are indicated by brackets () in the specification.

Parameter	Specifications	Values (sets of values)
Mesh size	M [-]	125, 1000
Material	($material$, E, ν, b) ([-], [GPa], [-], (μm))	(Cu, 117, 0.35, 0.254) (Al, 72, 0.347, 0.254) (Au, 78, 0.440, 2.870)
Dislocation distribution	$(\rho^i_{\mathrm{tot},0}, \rho^i_{\mathrm{mob},0}, \rho^i_{\mathrm{sta},0}, \rho^i_{\mathrm{lom},0})$ $([\frac{1}{\mu\mathrm{m}^2}], [\frac{1}{\mu\mathrm{m}^2}], [\frac{1}{\mu\mathrm{m}^2}], [\frac{1}{\mu\mathrm{m}^2}])$	(0.1, 0.1, 0.0, 0.0) (0.1, 0.05, 0.025, 0.025) (0.1, 0.0, 0.05, 0.05) (1.0, 1.0, 0.0, 0.0) (1.0, 0.5, 0.25, 0.25) (1.0, 0.0, 0.5, 0.5) (10, 10, 0.0, 0.0) (10, 5.0, 2.5, 2.5) (10, 0.0, 5.0, 5.0)
Crystal orientation	CO [-]	$\langle 100\rangle$, $\langle 110\rangle$, $\langle 123\rangle$
Source constant	C_{fr} [-]	1.5, 2.0
Curvature constant	$C_{\mathrm{fr,multi}}$ [-]	0, 0.1, 0.5
Reaction constants	(C_{lom}, C_{col}) ([-], [-])	(0.032, 0.016) (0.064, 0.032) (0.128, 0.064)
Cross-slip constant	(material, τ_{III}, V_{act}) ([–], [GPa], [b^3])	(Cu, 0.028, 300) (Cu, 0.028, 42.0) (Al, 0.005, 300) (Al, 0.005, 1.7) (Au, 0.01, 300) (Au, 0.01, 17.2)

Crystal Orientation Based on the crystal orientation, different resolved shear stresses result on the slip systems having significant impact on the microstructural evolution. Therefore, we exemplarily consider the three different crystal orientations $\langle 100\rangle$, $\langle 111\rangle$ and $\langle 123\rangle$.

Initial dislocation density configuration The initial dislocation density configuration affects the microstructure evolution with a direct impact on the stress–strain behaviour. We vary the initial dislocation density and its distribution based on different dislocation density quantities. Here, the total initial dislocation density $\rho_{\mathrm{tot},0}$ is calculated as the sum of each total dislocation density per slip system $\rho^i_{\mathrm{tot},0}$. Each total dislocation density per slip system can consist of mobile dislocation density $\rho^i_{\mathrm{mob},0}$ and network dislocation density $\rho^i_{\mathrm{net},0}$. The network dislocation density consists of Lomer dislocation density $\rho^i_{\mathrm{lom},0}$ and stabilized dislocation density $\rho^i_{\mathrm{sta},0}$, which leads to

$$\rho_{\mathrm{tot},0} = \sum_{i=1}^{12} \left(\rho^i_{\mathrm{mob},0} + \rho^i_{\mathrm{lom},0} + \rho^i_{\mathrm{sta},0}\right). \tag{1}$$

We assume the initial Lomer dislocation density to be equal to the initial stabilized dislocation density such that $\rho^i_{\mathrm{lom},0} = \rho^i_{\mathrm{sta},0}$. This simplification is based on the assumption that each stabilised dislocation density is always part of connected Lomer dislocation density [47]. To describe the variation of the initial dislocation distribution, we introduce a parameter η_d that is defined as

$$\eta_d = \frac{\rho_{\mathrm{lom},0} + \rho_{\mathrm{sta},0}}{\rho_{\mathrm{tot},0}}. \tag{2}$$

The parameter describes the proportion of the dislocation density that is mobile or bound in the network, i.e. as a Lomer or stabilised dislocation density, in the total dislocation density. Based on the initial dislocation density distribution, we calculate the yield stress τ^i_y based on [48]. For the CDD formulation, we define the yield stress as

$$\tau^i_y = \mu b^i \sqrt{\sum_{j=1}^{12} a^{ji} \cdot \begin{cases} \left(\rho^i_{\mathrm{lom},0}\right), & a^{ji} = a_{\mathrm{lom}} \\ \left(\rho^i_{\mathrm{mob},0} + \rho^i_{\mathrm{sta},0}\right), & \text{else} \end{cases}} \tag{3}$$

with the interaction coefficients a^{ji} derived from DDD simulations [49].

Dislocation nucleation The dislocation nucleation is incorporated by a formulation of a homogenized dislocation source model [44]. The source model is activated when the resolved shear stress τ^i_{res} on a slip system exceeds a critical shear stress τ^i_{crit}, which is defined as

$$\tau^i_{\mathrm{crit}} = \max\left(\frac{1}{C_{\mathrm{fr}}}\mu b \sqrt{\rho^i_{\mathrm{tot}}}, \tau_{\mathrm{CRSS}}\right). \tag{4}$$

The constant C_{fr} is set to control the critical shear stress for activating the source model, μ the shear modulus and τ_{CRSS} is a reference value for the critical shear stress in case of vanishing dislocation densities.

Dislocation cross-slip Modelling the cross-slip of screw dislocation density is carried out by considering a probability term based on Verdier *et al* [50]. Incorporating the term in CDD according to [42], the cross-slip probability P_{cs} within a simulation time step enters as a parameter in the evolution equation of the cross-slip model and is given as

$$P_{\mathrm{cs}} = f(\exp(V_{\mathrm{act}}), \exp(\tau_{III})). \tag{5}$$

V_{act} denotes the activation volume and τ_{III} defines the shear stress at the transition to stage III hardening. Both parameters are material-dependent and are varied in accordance to the chosen material. Furthermore, the model includes the generation of the curvature density based on dislocation multiplication due to cross-slip. It occurs if the resolved shear stress τ^i_{res} on a slip system exceeds the critical shear stress τ^i_{crit} of the source model. The parameter $C_{\mathrm{fr,multi}}$ is a constant between 0 and 1, which captures the shape of the bow-out of the cross-slipped dislocation [44].

Reaction model We consider dislocation reactions of dislocation on different slip systems on a continuum level [43, 51]. Here, we consider the input parameters for collinear and Lomer reactions. Collinear as well as Lomer reactions affect the mobile as well as the network dislocation density. The dislocation reaction model is formulated according to [43] as

$$\partial_t\left(\rho^k_{\mathrm{mob},r} + \rho^k_{\mathrm{sta},r}\right) = -C_r\left(\frac{1}{b}\partial_t\gamma^i\sqrt{\rho^j_{\mathrm{mob}} + \rho^j_{\mathrm{sta}}} + \frac{1}{b}\partial_t\gamma^j\sqrt{\rho^i_{\mathrm{mob}} + \rho^i_{\mathrm{sta}}}\right), \tag{6}$$

where C_r corresponds to a collinear coefficient C_{col} or a Lomer coefficient C_{lom}, which leads to reactions of mobile $\rho^k_{\mathrm{mob},r}$ and stabilized $\rho^k_{\mathrm{sta},r}$ dislocation density of reaction r. The two interacting slip systems i and j yield a reaction on slip system k. It has to be remarked that k can be identical to i or j or represent a third slip system [51]. Furthermore, Lomer reactions lead to the generation of Lomer dislocation density ρ^k_{lom}, which is given as

$$\partial_t \rho^k_{\mathrm{lom}} = -\partial_t \left(\rho^k_{\mathrm{mob,lom}} + \rho^k_{\mathrm{sta,lom}} \right). \tag{7}$$

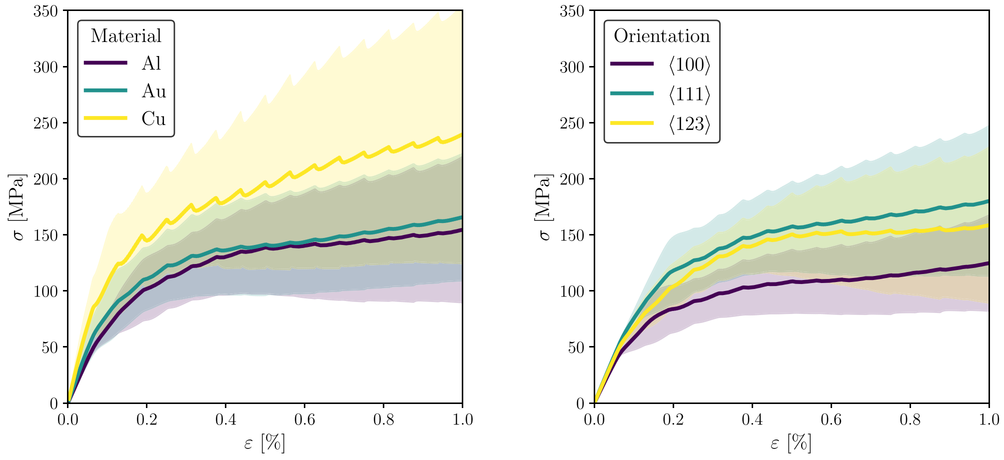
It accounts for the transfer of mobile and stabilized to Lomer dislocation density.

Mesh Size To control the impact of the spatial discretization of the simulations setup two different mesh sizes are chosen. The cubic sample has been analyzed for a total element number of 125 hexahedrons ($5 \times 5 \times 5$) with an edge length of 1.0 μm and a number of 1000 hexahedrons ($10 \times 10 \times 10$) with an edge length of 0.5 μm.

2.3.2. Exploration of database. To obtain an understanding of the simulation output based on the variation of input parameters, the generated simulation database is statistically evaluated. Figure 4(a) depicts the stress–strain behaviour of the three different elastic material properties considered. A higher mean yield stress and larger hardening are visible for Cu compared to those of Al and Au. The stress deviation is increasing with ongoing plastic deformation and is overall larger for Cu than for Al and Au.

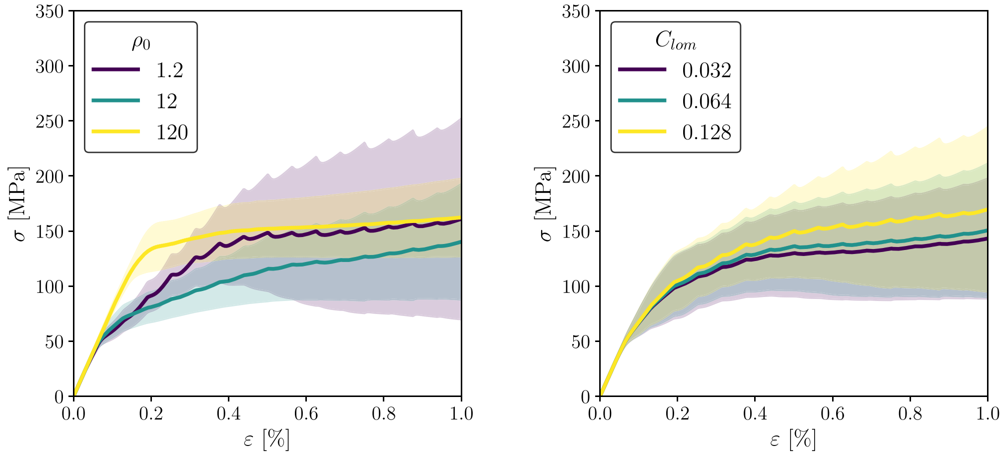
Focusing exemplarily on the subset of the database for Al, figure 4(b) depicts the stress–strain behaviour for the three considered crystal orientations. Here, the material parameters are equal for each crystal orientation. It can be observed that plastic yielding starts at lower stresses for $\langle 100 \rangle$ orientation. Thus, the transition from elastic to plastic deformation occurs at an earlier strain state. Higher mean stresses are obtained for $\langle 111 \rangle$ and $\langle 123 \rangle$ orientation. The deviation of stress during plastic deformation occurs to be comparable for each crystal orientation. Figure 4(c) shows the influence of the initial dislocation density. The mean yield stress is found to be largest for the highest initial dislocation density. A low dislocation density of 1.2 μm^{-2} exhibits a similar yield stress compared to an initial dislocation density of 12 μm^{-2}. Simulations with initial dislocation densities of 1.2 μm^{-2} exhibit slight fluctuations for the stress–strain behavior. Simulations of larger initial dislocation density do rarely show this effect. Figure 4(d) depicts the stress–strain behaviour of the three considered Lomer reaction coefficients. Lower coefficients lead to lower mean stress–strain curves. The stress deviation is similar for each Lomer reaction coefficient.

The CDD model distinguishes between mobile and network dislocation density. Since the initial dislocation density has a significant impact on the macroscopic response, as shown in figure 4(c), we analyse the influence of the distribution of initial dislocation density using η_d. We study three initial cases, i.e. only mobile dislocation density with $\eta_d = 0$, equally distributed mobile and network dislocation density with $\eta_d = 0.5$ and only network density with $\eta_d = 1$. To analyze this effect in the dataset, we use the subset of the database for Al with an initial dislocation density of 10 μm^{-2} per slip system. Figure 5(a) displays the stress–strain behaviour. The case $\eta_d = 1$ shows an increased mean stress during plastic deformation. The hardening behaviour and the deviation of stress is similar for each case. Figure 5(b) displays the mobile dislocation density evolution over strain. During elastic deformation, the mobile dislocation density stays constant. After the transition from elastic to plastic deformation, a significant dependency of the mobile dislocation density evolution on η_d can be observed. $\eta_d = 0$ leads to a reduction of mobile dislocation density. In contrast, the presence of initial network density leads to a generation of mobile dislocation density. An overall decrease



(a) Material

(b) Crystal orientation of Al



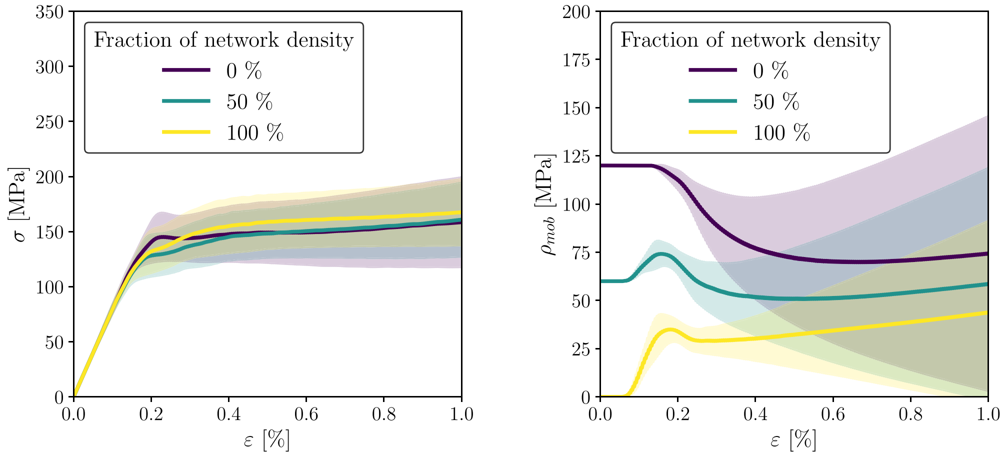
(c) Initial dislocation density of Al

(d) Lomer reaction coefficient of Al

Figure 4. Mean and standard deviation of the stress–strain behavior with respect to different parameters settings.

is observed for the mean mobile dislocation density evolution. However, the deviation of the mobile dislocation density evolution increases while straining, whereby the deviation increases larger for simulations with lower initial network dislocation density.

To better understand the correlations between input parameters and simulation output, the Pearson correlations between input and output are calculated. We analyze the following material characteristics, which are derived from the stress and the dislocation density evolution: resulting overall yield stress σ_y, calculated local yield stress τ_y, maximal stress $\sigma_{\max}$, stress at total strain $\sigma_{\varepsilon_{\max}}$, hardening $\theta_\sigma = (\sigma(\varepsilon_{\text{tot}}) - \sigma_y)$, maximum of dislocation density $\rho_{\max}$, dislocation density at total strain $\rho_{\varepsilon_{\max}}$, change of mobile dislocation density $\theta_\rho = (\rho_{\text{mob}}(\varepsilon_{\text{tot}}) - \rho_{\text{mob},0})$. Figure 6(left) depicts all Pearson correlations between input and material characteristics. It can be observed that the initial total dislocation density correlates strongly with the yield stress as well as with the maximal mobile dislocation density. The initial dislocation density, which is



(a) Stress (b) Mobile dislocation density

Figure 5. Dependence of the distribution of initial dislocation density based on factor η_d for simulations mimicking Al and having a initial dislocation density of 10 μm^{-2} per slip system based on the mean and standard deviation.

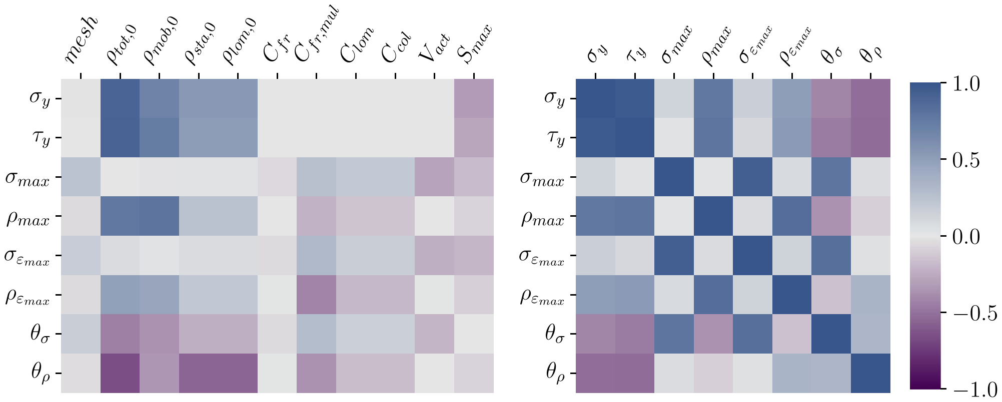


Figure 6. Pearson correlation of input vs output parameters (left) and of output parameters themselves (right).

mobile, Lomer or stabilized density, exhibits a similar correlation. For the initial densities, a negative correlation with the change of dislocation density θ_ρ is found. Other correlations are less pronounced. Complementary, figure 6(right) shows the Pearson correlation between all material characteristics. The diagonal is 1 as expected for self correlation. The overall yield stress and the calculated local yield stress show a strong correlation. $\sigma_{\max}$ and $\sigma_{\varepsilon_{\mathrm{tot}}}$ correlate strongly as well. σ_y exhibits a negative correlation with θ_σ or θ_ρ. With this analysis, the Pearson correlation provides a first insight into the simulation data. However, it has to be remarked that it only reveals the correlation between two parameters. This first evaluation of the database shows that the input parameter variation yields a reasonable database that can be applied in the following for surrogate modelling.

2.4. Forward and inverse surrogate model

The mathematical formulation and the numerical methods comprised in the CDD approach are based on the solution of partial differential equations. To solve these equations comes along with high computational costs. To overcome this computational effort, we introduce a surrogate model that replaces the solution of the partial differential equations by a data-driven approach that mimics a simulation by a data-based linkage between input and output. We distinguish between a forward surrogate model, which predicts output from input and an inverse surrogate model, which predicts the input parameter space based on simulative or experimental output quantities. Surrogate models can be constructed by a variety of machine learning approaches, including neural networks and decision trees. Here, the primary objective is to identify a surrogate model that facilitates rapid and accurate predictions for the linkage of input and output in the context of CDD simulations. We distinguish between two categories of models: one for deriving individual point-wise predictions and another for deriving sequential predictions. In the following the potential model candidates are introduced. Detailed information about each surrogate model candidate is provided in appendix B.

Point-wise prediction A point-wise approach predicts one data point for each point in time, which, e.g. can be the strain state for a stress–strain curve. In this work we investigate three point-wise prediction model candidates known to show good performance for regression problems. First, we use a kNN model, which identifies the k samples from the training set most similar to the input and then predicts the average of the selected series. Second, we use a response surface method (RSM) model, which is an optimization heuristic to find minima in response functions. Here, the response function is the latent function given by the simulation, i.e. the relation between our input features and target. Third, we include Gradient Boosting Trees (GBTs). Tree-based models like GBT are well known to deliver strong performance across various applications, especially when dealing with small data sets.

Sequence prediction A sequence approach predicts an output sequence from an input sequence. We choose methods based on Recurrent Neural Networks (RNNs) and include LSTM Network models. Since our input is not a sequence but static data, we employ different techniques to transform the static input to a sequence input. A straightforward approach is to couple a LSTM with Repeated Inputs (RepInLSTM). It integrates static data into the LSTM by feeding the static data as a LSTM input at each strain state together with the previous predictions. Another approach is to use a LSTM with Raw Hidden State (RawInLSTM), which passes the static data as a hidden state into the LSTM at each strain state. The previous prediction is fed again as a LSTM input. A third approach is the consideration of a hybrid model. It is a combination of a point-wise GBT and a sequence-wise LSTM (GBT+LSTM) and combines the advantages of both approaches. A step-wise prediction using the GBT model in addition to LSTM yields a sequence-to-sequence mapping of the predictions of the GBT.

To train each surrogate model candidate based on the simulation database, the database must be preprocessed. Non-floating point input parameters such as material, crystal orientation and mesh size are treated as categorical variables and transformed into floating point input parameters. Then, all input parameters are normalized between 0 and 1. For the evaluation of the surrogate model candidates, we normalize also the simulation output between 0 and 1. The simulation input can be described as a parameter vector $\mathbf{X} \in \mathbb{R}^{1\times P}$ consisting of all input parameters P. The simulation output is a data series $\mathbf{Y} \in \mathbb{R}^{1\times T}$ with T data points. The forward surrogate model predicts the output $\mathbf{Y}^{\mathbf{SM}}$ based on the input $\mathbf{X}$. The inverse surrogate model predicts the input $\mathbf{X}^{\mathbf{SM}}$ based on the output $\mathbf{Y}$. Formally, for the total number N of all simulations, we receive an input matrix $\mathbf{X} \in \mathbb{R}^{N \times P}$ and an output matrix $\mathbf{Y} \in \mathbb{R}^{N \times T}$.

To evaluate the model performance of the data-driven predictions, a variety of error measures are established in literature based on the comparison of ground truth and predicted data [52]. In this work, we apply the commonly used metrics to evaluate the performance of regression models: All are commonly used metrics to evaluate the performance of regression models.

The root mean square error (RMSE) measures the average magnitude of the residuals between predicted and actual values, defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}. \tag{8}$$

The mean absolute percentage error (MAPE) quantifies the percentage difference between predicted and actual values, calculated by

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{9}$$

The absolute percentage error (APE) is the percentage difference between each predicted and actual value given as

$$\mathrm{APE} = \left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{10}$$

The coefficient of determination (R^2) indicates the proportion of the variance in the dependent variable that is predictable from the independent variables. It can be written as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}. \tag{11}$$

For each error measure, y_i represents the actual values, $\hat{y}_i$ represents the predicted values and $\bar{y}$ is the mean of the actual values.

To evaluate the surrogate model candidates, we perform k cross-validations to increase the validity of our results. We apply a train-test split of $\frac{k-1}{k} / \frac{1}{k}$ for each cross-validation. We receive k error estimations based on the considered error measure and compare the model candidates based on these cross-validation scores. We compute the mean and the standard deviation of the considered error measure of all k cross-validation folds. The expected generalization capability corresponds to the mean of the cross-validation scores [53]. Since the model performance can strongly depend on the appropriate hyperparameter choices, we conducted the hyperparameter tuning as described in appendix C.

3. Results

3.1. Surrogate model candidate selection

To select the most suitable surrogate model candidate for the forward and inverse surrogate modelling approach, we evaluate the proposed model candidates based on their accuracy. For the forward surrogate model approach, each model candidate is trained and evaluated for the stress and dislocation density evolution defined as targets, based on a 10-fold cross-validated RMSE. Table 2 displays the performance of each forward model candidate with respect to each target and to each considered material. The results showing the lowest RMSE are marked

Table 2. Comparison of accuracy of forward surrogate model candidates based on mean and standard deviation of the RMSE measure.

Model	Target	Dislocation density			Stress		
	Material	Al	Au	Cu	Al	Au	Cu
kNN	mean std	0.109 ±0.011	0.104 ±0.010	0.117 ±0.019	0.099 ±0.011	0.076 ±0.006	0.090 ±0.008
RSM	mean std	0.035 ±0.002	0.036 ±0.004	0.040 ±0.003	0.038 ±0.003	0.031 ±0.002	0.042 ±0.003
GBT	mean std	**0.031** ±0.002	**0.022** ±0.002	**0.033** ±0.004	0.015 ±0.002	0.012 ±0.003	**0.027** ±0.003
RawInLSTM	mean std	0.041 ±0.057	0.067 ±0.043	0.045 ±0.052	0.015 ±0.005	0.011 ±0.001	0.040 ±0.029
RepInLSTM	mean std	0.047 ±0.048	0.052 ±0.042	0.036 ±0.053	**0.012** ±0.008	**0.009** ±0.002	0.034 ±0.021
GBT+LSTM	mean std	0.063 ±0.061	0.033 ±0.034	0.055 ±0.056	0.019 ±0.014	0.030 ±0.006	0.051 ±0.031

in bold. It is found that the point-wise baseline model candidates kNN and RSM give higher RMSE values than GBT across all data sets. Based on the comparison of the RMSE values, the GBT performs best for the dislocation density predictions of the considered materials. For the stress prediction, no model strictly outperforms the others. However, the sequence-wise RepInLSTM gives the best results for the stress prediction in Al and Au. For Cu, GBT performs slightly better than the other models. It can be observed that the RawInLSTM has overall a slightly worse performance compared to the RepInLSTM. The hybrid approach exhibits less accurate results compared to each individual point-wise and sequence-wise constituent. To summarize, the point-wise GBT demonstrates an overall low mean MSE and small standard deviations for the forward surrogate model for both targets. Thus, we identify the GBT to be most appropriate for the forward surrogate model for the subsequent evaluations.

Considering the inverse surrogate model approach, we follow a similar procedure. We evaluate the model candidates based on the prediction accuracy of a simulation input parameter set based on stress–strain data. The results of the RMSE values for the inverse investigation of our data are given in table 3. It can be observed that the GBT exhibits the highest overall accuracy compared to the other candidates. The results show that the hybrid approach has a smaller mean RMSE than the sequence-wise approaches. Within the two sequence-wise approaches, RepInLSTM occurs to be more accurate compared to RawInLSTM. The lowest accuracy is given by RSM and the kNN provides slightly better predictions than the sequence-wise approaches. Thus, the GBT model occurs best suited for the inverse surrogate model approach. It is remarked here, that the evaluation of the model candidates is presented here based on the RMSE measure. The same trend has been received by the application of the R^2 metric. For brevity, the exact numbers are not shown here. Taking into account the evaluation results, we select the GBT for the forward as well as the inverse surrogate model approach for the subsequent analyses.

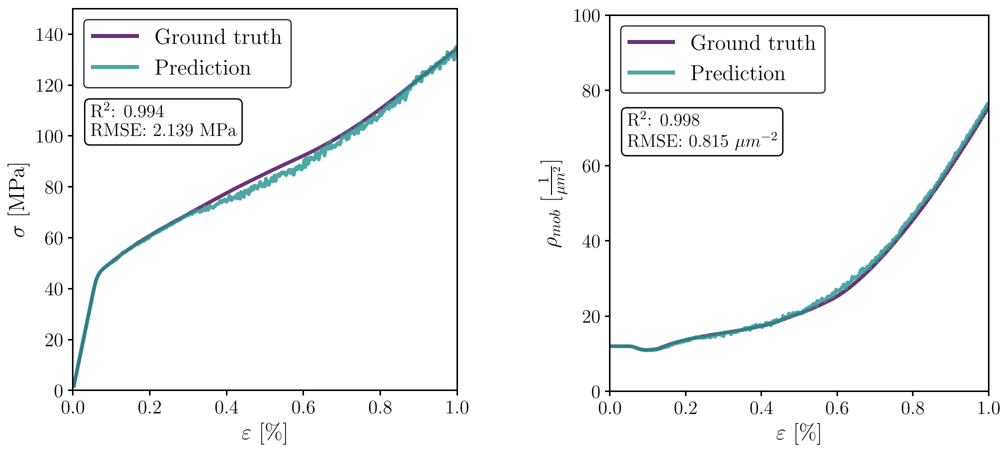
Table 3. Comparison of accuracy of inverse surrogate model candidates based on mean and standard deviation of the RMSE measure.

Model	Target	Simulation input parameter space		
	Material	Al	Au	Cu
kNN	mean std	0.319 ±0.004	0.316 ±0.005	0.315 ±0.004
RSM	mean std	1.249 ±0.163	2.742 ±0.635	5.841 ±1.17
GBT	mean std	**0.175** ±0.003	**0.179** ±0.002	**0.182** ±0.002
RawInLSTM	mean std	0.389 ±0.001	0.389 ±0.003	0.392 ±0.002
RepInLSTM	mean std	0.347 ±0.006	0.356 ±0.019	0.360 ±0.007
GBT+LSTM	mean std	0.295 ±0.002	0.301 ±0.014	0.321 ±0.026

3.2. Forward surrogate model

To qualitatively explore the GBT forward surrogate model in detail, we demonstrate the 10-fold cross-validation accuracy with a 90/10 train-test split based on a subset of the database for Al. An exemplary forward prediction is shown in figure 7, for (a) the stress over strain and (b) the mobile dislocation density over strain behavior. The RMSE for the stress prediction results to a value of $2.139\,\mathrm{MPa}$ and for the mobile dislocation density prediction to a value of $0.815\,\mu\mathrm{m}^{-2}$. In can be seen that during elastic deformation, the prediction of stress and dislocation density fits very well to the ground truth. With the onset of plastic deformation, the prediction starts to show slight deviations from the ground truth. However, the trend of the prediction of stress as well as of the dislocation density is resembling to the ground truth. The increment between each predicted data point exhibits a slightly jagged behaviour during plastic deformation.

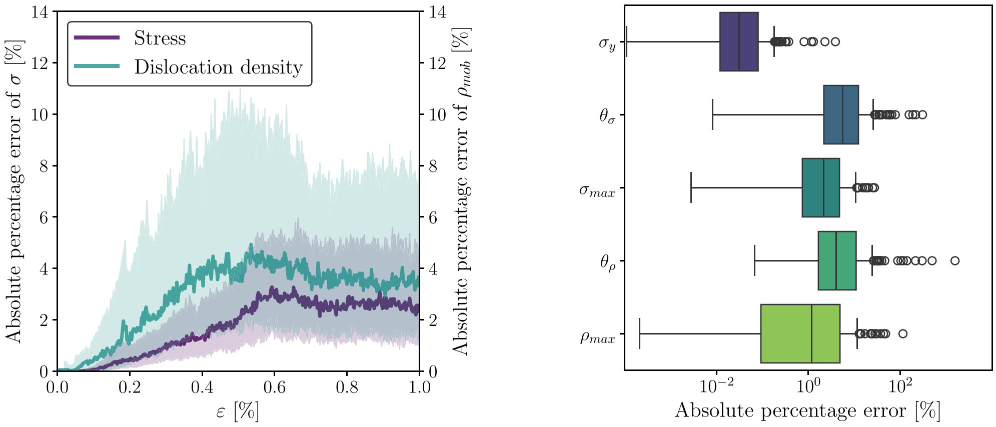
To quantify the accuracy of the GBT forward surrogate model, we display the APE of the stress and dislocation density evolution at each strain state in figure 8(a). The bold line indicates the median APE and the shaded background indicates the area between lower and upper quartiles of the APE, for the targets stress and dislocation density, respectively. During elastic deformation up to 0.1% strain the APE is close to zero. During plastic deformation, the median APEs as well as the interquartile range of the APEs increase with ongoing straining. The dislocation density prediction starts to deviate at a slightly earlier strain state than the stress prediction. The median predictions converge after a strain of 0.6% to a constant APE with a slightly larger error of about 4% median APE of the dislocation density prediction compared to about 2.5% median APE of the stress prediction. To quantify the GBT forward surrogate model from a macroscopic perspective, we evaluate the predictions based on macroscopic material characteristics shown in figure 8(b). The box plot for each material characteristic displays the median APE as a bold vertical line, the interquartile range of the APE as a box, 1.5 times the interquartile range of the APE as whiskers and outliers as markers. The prediction of the yield stress σ_y shows a good accuracy and almost all predictions of σ_y have an APE below 1%. The



(a) Stress

(b) Dislocation density

Figure 7. Example of GBT forward surrogate model prediction for (a) stress and (b) dislocation density.

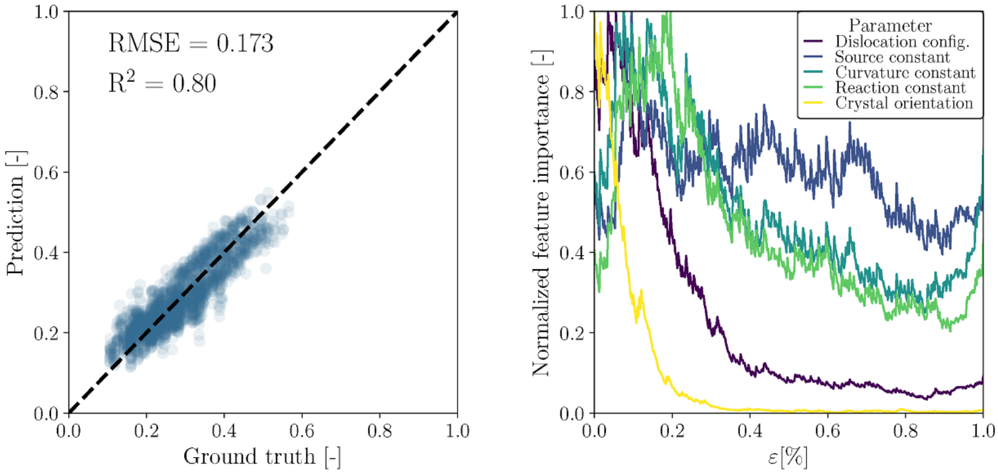


(a) Absolute percentage error of stress and dislocation density evolution

(b) Absolute percentage error of a selection of material characteristics

Figure 8. Accuracy of the GBT forward surrogate model based on the absolute percentage error (APE) measure of (a) stress and dislocation density evolution and (b) a selection of material characteristics.

behavior changes for material characteristics describing plastic deformation. The predicted hardening rate θ_σ has a median APE of around 10% and whiskers ranging from 0.01% to 20% APE. Outliers may also have an APE of more than 100%, i.e. the hardening rate is strongly overestimated in these cases. The APE of the change of dislocation density θ_ρ is resembling to the predictions of θ_σ. The predictions of maximum stress σ_{max} and maximum dislocation density ρ_{max} reveal a median APE of around 2% and 1%, respectively. The interquartile range for the predictions of ρ_{max} ranges between two orders of magnitude compared to a range of



(a) Prediction accuracy for each simulation input parameter set

(b) Feature importance during a stress-strain evolution for the inverse surrogate model

Figure 9. Prediction over ground truth evaluation and analysis of feature importance for the inverse surrogate model.

one order of magnitude for $\sigma_{\max}$. However, only few predictions for $\sigma_{\max}$ and $\rho_{\max}$ are outliers with APEs larger than 10%.

3.3. *Inverse surrogate model*

To explore the quantitative accuracy of the inverse GBT surrogate model, we choose the same procedure as for the forward model. We demonstrate the 10-fold cross-validation accuracy with a 90/10 train-test split based on a subset of the database for Al. Figure 9(a) shows the prediction over ground truth values and thus illustrates the prediction accuracy of each of the 1944 input parameter sets derived from the provided stress–strain data. Here, each depicted ground truth value corresponds to the mean of each normalized parameter set. The overall prediction shows a RMSE value of 0.173 and a R^2 value of 0.8 To investigate the results further, we analyse the importance of features during the stress–strain evolution for the inverse model. Figure 9(b) quantifies the normalized feature importance during loading for the simulation input parameters. The results show the variation of impact of several features on the input parameter prediction of the inverse GBT surrogate model. The prediction of the initial dislocation densities is mainly dependent on the onset of yielding until a strain state of 0.2%. The ongoing plastic deformation has smaller impact on the inverse prediction of initial dislocation densities. Similarly, the model predicts the crystal orientation most dominantly by the stress–strain data given during elastic deformation and the onset of plastic deformation. In contrast, the feature importance for the prediction of the source, curvature and reaction constant is broader distributed. These constants are less dependent on the range of elastic deformation. The results indicate that the onset of yielding impacts the prediction of these constants most. However, ongoing plastic deformation has still a significant influence on the estimation of the source, curvature and reaction constant.

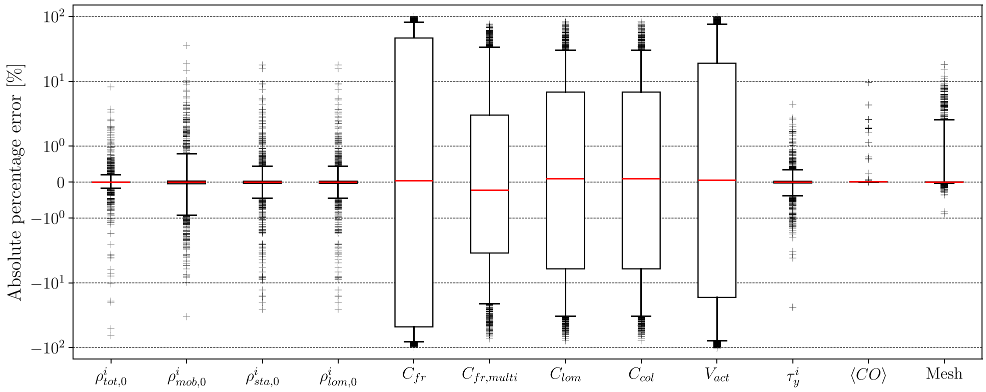
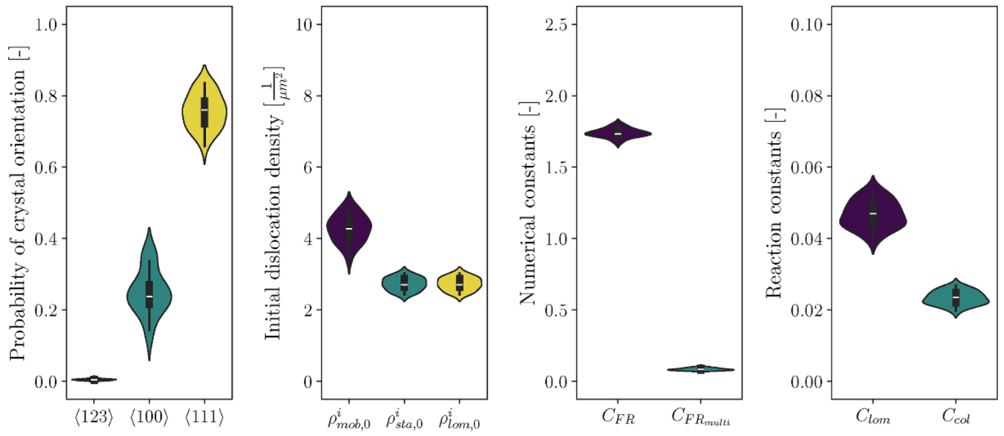


Figure 10. Evaluation of each simulation input parameter prediction performance based on the inverse GBT surrogate model.

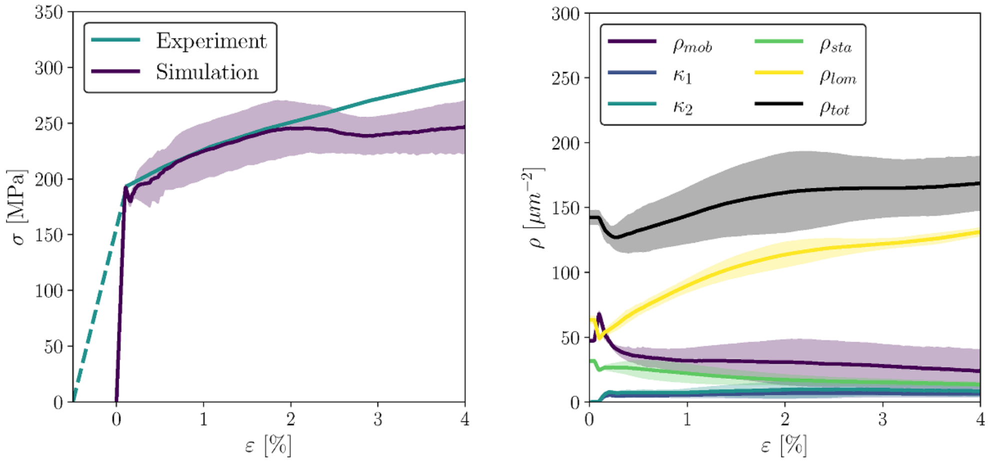
To evaluate the prediction of all considered simulation input parameters, we analyse the APEs, see figure 10. Each box plot displays the range between the lower and upper quartiles as a box, the median value as a bold bar, whiskers with 5%–95% confidence interval (CI), and outliers. The prediction of the initial dislocation densities ρ^i_{tot}, ρ^i_{mob}, ρ^i_{sta} and ρ^i_{lom} show a lower CI than $\pm 1\%$ and only few outliers. The prediction performance of the crystal orientation probability $\langle \mathrm{CO}\rangle$ of $\langle 123\rangle$, $\langle 100\rangle$ and $\langle 111\rangle$ is accurate with the lowest APE. The prediction of the yield stress τ^i_y yields accurate results with a CI smaller than $\pm 1\%$. The prediction of the source constant C_{fr} exhibits a high uncertainty with a CI of $\pm 80\%$ and an interquartile range of $\pm 40\%$. The prediction error of V_{act} exhibits a similar uncertainty compared to C_{fr}, however, with a smaller interquartile range of around $\pm 20\%$. The results of the curvature constant $C_{\mathrm{fr,multi}}$ is more accurate compared to C_{fr} with a CI of $\pm 30\%$ and an interquartile range of $\pm 4\%$. The APE of the reaction constants C_{lom} and C_{col} are comparable to the APE of $C_{\mathrm{FR,multi}}$. The estimation of the mesh size reveals a low overall APE, however, with a bias of the upper confidence bound of around 3% compared to the lower confidence bound, which indicates a slight overestimation. Overall, the APE exhibit a median close to zero for all considered parameters. No significant bias of underestimating or overestimating input parameter values is observed when considering the interquartile ranges. The parameter categories dislocation density, crystal orientation and mesh reveal a high accuracy. Larger uncertainties are exhibited for the reaction and interaction constants.

3.4. Input parameter prediction from experiment

Considering the surrogate approach, we address the challenge of identifying the simulation input parameter set that corresponds to a given experimental stress–strain curve. We apply the inverse surrogate model to stress–strain data from an experimental compression test provided by Kiener *et al* [54]. We consider the stress–strain curve of a squared Cu sample with an edge length of 4.82 μm in $\langle 111\rangle$ crystal orientation. For the parameter prediction, we apply a two step transformation on the experimental data to prepare it for the surrogate procedure. Since the elastic stiffness of the experimental stress–strain curve is affected by the system compliance such as e.g. stiffness of the testing machine as already reported by Kiener *et al* [54] for a comparison of DDD data and experiments, we first transform the experimental elastic stiffness to the theoretic elastic stiffness of Cu. Afterwards, we set the initial strain of the experimental



(a) Simulation input parameter prediction based on experimental stress-strain curve.



(b) Experimental and simulative stress-strain curve.

(c) Dislocation density distribution evolution.

Figure 11. Inverse parameter identification from experiment. (a) Simulation input parameter prediction from an experimental stress–strain curve. (b) Comparison of experimental and simulative stress–strain behavior resulting from predicted simulation input parameters. (c) Simulative results of dislocation density evolution from CDD simulations resulting from predicted input parameters.

stress–strain curve to zero. In a second step, we cut the experimental stress–strain curve at 1% strain, which corresponds to the limit inherent in the training database. Since the experiment represents a Cu sample, we restrict the inverse GBT surrogate model on the sub-dataset for Cu only to reduce the uncertainty of the inverse surrogate model. To estimate now the distribution of each parameter, we apply five repetitions of a 10-fold cross-validation. The predicted parameter distributions based on the experimental data are shown in figure 11(a). Here, each predicted simulation input parameter reveals a unimodal distribution. The results show that the prediction of the initial dislocation densities (ρ^i_{mob}, ρ^i_{sta}, ρ^i_{lom}), the reaction constants

(C_{lom}, C_{col}) as well as the multiplication constants (C_{FR}, $C_{\mathrm{FR,multi}}$) yield some probability distribution. The predicted initial dislocation density distributions indicate an initial dislocation network with more initial network density than initial mobile density. As a first simple evaluation of the inverse surrogate model, we predict the crystal orientation probability. Although already known from the experiment, we use the inverse model to predict the crystal orientation. The results of figure 11(a) show that the model yields a high probability of having a $\langle 111 \rangle$ crystal orientation that is the correct orientation given by the experiment.

To further evaluate the predicted simulation input parameters, we conduct CDD simulations applying the identified parameter probablity distributions. We choose five simulation input parameter sets of the predicted distributions by selecting the median, two overestimated and two underestimated parameter sets within the interquartile range of each predicted parameter, and conduct the CDD simulations. Figure 11(b) displays the mean and the standard deviation of the resulting simulative stress–strain curves up to a total strain ε of 4%. Since the elastic stiffness differs between experiment and simulation, which might be based on the impact of the stiffness of the testing machine, the experimental elastic stiffness is indicated as green dashed line. It is transformed as described above to the theoretic elastic stiffness of Cu that is now identical to the simulation stiffness. Thus, a comparison of the elastic stiffness is not part of the analysis. Comparing the experimental and the simulative stress–strain curve, the yield stress is resembling between experiment and simulation. With ongoing plastic deformation, the simulated stress–strain behavior shows good accordance to the experiment up to 2% strain. For larger strains, the simulative stress–strain curve starts to deviate from the experimental stress–strain curve and exhibits a lower stress evolution compared to the experiment, leading to a stress difference of 50 MPa at 4% strain. The standard deviation of the simulated stress–strain data is about ± 20 MPa during plastic deformation for the five simulation input parameter sets. To provide an insight into the dislocation microstructure evolution, the evolution of the dislocation densities in the CDD simulations is shown in figure 11(c). The diagram displays the mobile ρ_{mob}, stabilized ρ_{sta}, Lomer ρ_{lom}, screw GND κ_1, edge GND κ_2 and the total ρ_{tot} dislocation density evolution. Due to the five different input parameter sets, the initial dislocation density distribution slightly differs. During elastic deformation, the dislocation densities remain constant. A small region of dislocation density transfer from ρ_{lom} to ρ_{mob} and ρ_{sta} can be observed just after reaching the yield stress. In the plastic regime, ρ_{lom} as well as edge and screw GND density increase, whereas ρ_{mob} and ρ_{sta} slightly decrease during ongoing plastic deformation. The overall trend of the dislocation density evolution slightly changes at 2% strain to a less pronounced increase. The results show that the quantitative ratio of Lomer and mobile dislocation density reveals a large amount of persisting Lomer dislocation density and a reduction of mobile dislocation density during plastic deformation.

4. Discussion

This introduced surrogate model bypasses continuum dislocation dynamics simulations through a data-driven linkage of input parameter sets and output quantities. The work comprises the generation of a CDD simulation database representing simulations of a variety of different input parameter sets. Based on the selection and evaluation of model candidates, a surrogate model for forward and inverse prediction has been trained. Besides the speedup of CDD computations by the forward surrogate model, the demonstration of the simulation input parameter prediction based on experimental stress–strain data has been presented.

The evaluation of the surrogate model candidates has shown that a point-wise predictor, namely the GBT, performs overall better than a sequence-wise predictor such as the LSTM with respect to the lowest mean RMSE measure as shown in table 2 for the forward and in table 3 for the inverse surrogate model. We observe that the forward RepInLSTM surrogate model shows more accurate results for the stress prediction of Al and Au compared to the GBT. However, the GBT outperforms all other candidates for the stress prediction of Cu and for the dislocation density prediction of each considered material. Similarly to the forward problem, the GBT outperforms each model candidate for the inverse surrogate model except for Cu, where the kNN performs equally well based on the lowest RMSE measure. This is different to the forward surrogate model, where kNN shows less overall performance based on the RMSE measure. This leads to the conclusion that distinct features are present in the stress–strain data, that map the proximity of the stress–strain data to the proximity of the simulation input parameters, but not vice versa. Ultimately, we select the GBT for the forward and the inverse surrogate model. However, it has to be remarked that the GBT does not include the material history in its prediction, which is contradicting plasticity theory on a first glance [55]. In literature, LSTMs are frequently employed for data series predictions since it has been found that they are able to take into account the influence of preceding microstructure evolution [27, 28, 56]. This is not explicitly found in our study since the GBT shows better results compared to the LSTM. We assume this can be explained here by only considering one system set-up with the same loading scenario, strain rate and total strain. The overall microstructural evolution history of the variations might be too similar to have a significant impact on the model. Therefore, the consideration of the material history information does not necessarily provide an advantage for the prediction of the LSTM compared to the GBT as indicated by the ambiguous best model candidates with the lowest mean RMSE for stress in table 2. The prediction accuracy of the hybrid approach, i.e. the combination of GBT and LSTM, also does not leverage the GBT and the LSTM for higher accuracy. Incorporating different loading scenarios, specimen sizes or strain rates might be of particular interest for further development of forward and inverse surrogate modelling. However, a comparison of various model candidates is reasonable based on the present observation that more complex surrogate models such as the LSTMs are not necessarily superior to simpler surrogate models such as GBTs.

The proposed forward GBT surrogate model demonstrates accurate predictions for the stress and dislocation density evolution as shown exemplarily in figure 7. Based on the absolute prediction error for the stress σ and the dislocation densities ρ_{mob}, the model exhibits accurate results during elastic deformation until 0.1% strain. The prediction accuracy slightly decreases with ongoing plasticity leading to an APE of 2.5% for stress and 4% for ρ_{mob} at 1% strain as shown in figure 8(a). This APE can be considered small compared to the median stress of 120 MPa at 1% strain for the Al sub-dataset as shown in figure 4(a). As indicated by the absolute percentage error of the material characteristics in figure 8(b), the slightly decreasing prediction accuracy with ongoing plasticity affects especially the prediction of material characteristics describing plastic deformation such as the hardening rate. The prediction error may be affected by strongly non-linear correlations between evolutionary input parameters and mechanical behavior as indicated by the low linear correlation for evolutionary parameters in figure 6. We assume that an extension of the database to a larger regime of parameter values, could support the optimization of the surrogate model accuracy during plastic deformation.

In principal, such a forward surrogate model now applies to be incorporated into large scale plasticity models. This aims to latest research in data-integrated multi-scale approaches for faster and small scale informed modelling [24, 28, 57]. Considering the good accuracy of the results, the significant reduction of the computational costs by the surrogate model has to be highlighted. Whereas a CDD simulation requires a significant amount of CPU resources, the proposed forward surrogate model predicts the desired output several orders of magnitude faster. This facilitates, e.g. the implementation of CDD into an accelerated multi-scale framework in the future.

The inverse surrogate model demonstrates the ability to predict the simulation input parameter set from simulation as well as from experimental stress–strain data. The prediction based on simulation stress–strain data shows high accuracy as displayed by the prediction vs ground truth diagram in figure 9(a). Based on the normalized prediction error, the prediction accuracy has been evaluated for each considered simulation input parameter as shown in figure 10. We observe parameter predictions with a confidence interval of less than $\pm 1\%$ as e.g. for the initial dislocation densities, the yield stress or the crystal orientation that indicate a very good prediction accuracy. The crystal orientation prediction depends most dominantly in the regime of yielding as shown by the high feature importance until 0.1% strain in figure 9(b). The high accuracy can be explained by the small but distinct differences for the onset of yielding for each considered crystal orientation as shown in figure 4(b). The prediction of the initial dislocation density depends most dominantly on the region during the onset of yielding and slightly beyond as shown by the high feature importance between 0.1% and 0.3% strain in figure 9(b). Since the initial dislocation densities are linked to the yield stress formulation in CDD (cp. equation (3)), we assume that the initial dislocation density prediction is consistently accurate if the yield stress is accurately predicted. Furthermore, the distribution of the dislocation densities reveals a high prediction accuracy despite a less distinct plastic deformation behavior as displayed in figure 5(a). In contrast, simulation input parameters, which affect the evolution of the dislocation microstructure during plastic deformation such as the source constant C_{fr}, the curvature constant $C_{\mathrm{fr,multi}}$ or the reaction constants C_{lom} or C_{col}, are less accurately predicted with confidence intervals ranging from $\pm 40\%$ to $\pm 80\%$ as displayed in figure 10. The prediction of these parameters depends on the entire plastic deformation region as indicated by the high feature importance starting from 0.1% until 1.0% strain in figure 9(b). We assume that the fewer prediction accuracy is related to a less distinct plastic deformation behavior, e.g. for the proximity of the stress–strain behavior for different Lomer reaction coefficient shown in figure 4(d). Furthermore, as shown by the Pearson correlation in figure 6, these parameters are a less distinct correlated to the stress and dislocation density evolution as indicated by the Pearson coefficient around 0, which can be interpreted as a low linear or a potential nonlinear influence compared to input parameters with a high Pearson coefficient. For the source and curvature constant, we observe a decreased prediction accuracy which can be explained by the fact that these parameters are only varied between two values. However, the proposed inverse GBT surrogate model is able to translate the information from plastic deformation to an accurate evolutionary input parameter prediction that can be seen by the interquartile range of the prediction error for the curvature and reaction constant is below $\pm 10\%$. With respect to the overall prediction error, it can be concluded that the inverse GBT surrogate model exhibits accurate results for the prediction of the CDD simulation input parameter sets based on stress–strain data.

The applicability of the inverse surrogate model to the direct identification of simulation input parameters from experimental stress–strain data has been investigated. We applied the inverse surrogate model to experimental results for a compression test presented by Kiener *et al* [54]. The simulation input parameters derived by the inverse surrogate model from the experimental stress–strain data reveal a unimodal distribution for each considered parameter as displayed in figure 11(a). Since we do not observe a multimodal distribution, i.e. a stress–strain curve corresponding to two or more distinct parameter sets, we interpret the results that the surrogate model is able to identify distinct features for each parameter leading to a uniform distribution. However, it can not be ruled out that an expansion of the parameter space might yield to more than one global minimum of each considered parameter resulting in multimodal distributions. To validate the predicted input parameter distributions, we computed the corresponding CDD simulations including the parameters. The stress–strain results of the CDD simulations based on the predicted simulation input parameters are then compared again to the experimental stress–strain data. A very good accordance can be observed in the regime until 1% strain as shown by the stress–strain data in figure 11(b). Beyond 1% strain, the simulation results of the stress–strain data can be considered still accurate and only starts to deviate from the experimental data for a strain larger than 2%. The deviation of simulation and experiment beyond 2% strain indicates that the microstructural evolution processes are not accurately represented within the CDD simulation for these strain states as indicated by the evolution of total dislocation density in figure 11(c). However, it has to be remarked that the training data for the surrogate model comprised data for strains up to 1%. The good accordance of the results up to 2% yield the conclusion that certain extrapolation from the data is possible. But the limits has to be investigated further. Possible explanations for the extrapolation behavior of the results could be based on simplifying assumptions of further microstructural parameters. The incorporation of certain input parameters, that have been kept constant in this study as, e.g. the glissile and coplanar reaction constants, need to be taken into account for further investigations. Furthermore, an increase in Lomer dislocation density, i.e. a stable dislocation network has been observed as shown in figure 11(c). In the CDD model, we incorporate unzipping of Lomer junctions and therefore the dissolution of dislocation network density in a simplified way [45]. However, the incorporation of a more accurate dissolution process of dislocation network density in CDD can be discussed, since we observe a limitation of the content of Lomer dislocation density in DDD simulations [47, 51]. But, as an overall conclusion to minimize the observed deviations in the regime of higher strains, we assume that the incorporation of simulations with higher strains into the database would yield a more precise input parameter prediction and avoid the well-known issues of extrapolation in data-driven models.

In general, we observe a moderate total dislocation density evolution, which is similar to other investigations of micro-compression tests [45]. However, more pronounced total dislocation density evolutions for this set-up can be found based on DDD analyses [54]. Here, only a small fraction of the total dislocation density is GND density, see figure 11(c). This is resembling to DDD simulation results, which have been conducted in alignment to the same considered experiment of Kiener *et al* [54]. Additionally, we observe an increasing contribution of Lomer dislocation density to the total dislocation density with ongoing plastic deformation as shown in figure 11(c). This indicates an increasing amount of network dislocation density that has been observed as well in DDD simulations in conjunction with a decreasing strain rate [58]. We assume that the experimental stress–strain data, which is conducted at a lower strain rate compared to CDD simulation, incorporate information about the stability of the network dislocation density, which is determined by the inverse surrogate model through the magnitude of the reaction constants. However, further analysis on the derivation of physical-based constitutive laws for dislocation network stability and evolution from DDD simulations is required for a better understanding of these network processes [59–61]. Summarizing, the prediction of simulation input parameters based on quite limited consideration of strain states still exhibits meaningful results and enables the derivation of input parameter sets for CDD simulations representing the experimental mechanical behavior and the microstructural evolution beyond the strain states, the inverse surrogate model is trained on. In general, the presented inverse modelling approach can be applied to any experimental data series as long as the simulation is capable to output this data. For example, one could consider the surface deformation evolution from DIC measurements or the GND density evolution from HR-EBSD measurements.

5. Conclusion

In this work, we introduced a surrogate modelling approach to link input parameter sets and output quantities for continuum dislocation dynamics simulations on a data-driven manner. We coupled a forward surrogate model that predicts the stress and the dislocation density evolution from a simulation input parameter set and an inverse surrogate model that derives the simulation input parameters such as the initial microstructure and model parameters based on simulation or experimental stress–strain data. We explored the mutual influence of the simulation input parameters on the mechanical behaviour, and incorporated it into a surrogate model for single crystalline materials. The main outcomes of this work are:

- A forward and an inverse surrogate model trained on a CDD database comprising a variety of elastic properties, crystal orientations, initial dislocation microstructures as well as numerical and evolutionary parameters.
- A parameter space exploration revealing insights into linear and non-linear correlations of the simulation input parameters and the resulting microstructural evolution and its mechanical behavior.
- An evaluation of surrogate model candidates revealing strengths and weaknesses of various machine learning approaches for capturing this work's scenario of mapping simulation input parameters to mechanical behavior.

The proposed forward surrogate model facilitates the derivation of CDD output data by several orders of magnitudes lower computational cost compared to the classical CDD simulation. The proposed inverse surrogate model provides a method to unravel the simulation input parameters from experimental data. The results indicate that the presented surrogate concept can be applied for a faster and more reliable combination of simulation and experimental data in the field of engineering and mechanics.

Data availability statement

Ongoing project the data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The financial support for this work in the context of the DFG research Projects SCHU 3074/4-1 and BO 2129/16-1 is gratefully acknowledged. This work was performed on the computational resource HoreKa funded by the Ministry of Science, Research and the Arts Baden-Wuerttemberg and DFG.

Conflict of 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.

Appendix A. Mathematical formulation of the continuum dislocation dynamics model

The mathematical formulation of the continuum dislocation dynamics method is based on a higher order formulation, which incorporates information about dislocation orientation and dislocation curvature [62, 63]. The CDD implementation incorporates a dislocation network evolution model [42] and dislocation source model [45] into a dicontinuous Galerkin framework [41]. In general, the model interlinks an external elastic problem with an internal microstructure problem. The external problem is solved by linear elasticity by

$$\boldsymbol{\sigma} = \mathbb{C}\left(\varepsilon - \varepsilon_{\text{pl}}\right), \tag{A.1}$$

with the elasticity tensor $\mathbb{C}$, the Cauchy stress tensor $\boldsymbol{\sigma}$, the infinitesimal strain tensor ε and the infinitesimal plastic strain tensor ε_{pl}. The equation takes into consideration the elastic properties of each desired material, which is simplified in our case to an isotropic material behavior with the Young's modulus E and the Poisson's ratio ν. The infinitesimal plastic strain tensor $\boldsymbol{\varepsilon}_{\text{pl}}$ equals to symmetric part of the plastic distortion tensor $sym(\boldsymbol{\beta}_{\text{pl}})$ under the assumption of small deformations, which can be calculated by

$$\boldsymbol{\beta}_{\text{pl}} = \sum_{i}^{12} \gamma^i \boldsymbol{M}^i, \tag{A.2}$$

with the Schmid tensor $\boldsymbol{M}^i$ and the plastic shear strain γ^i. The Schmid tensor equals to $\boldsymbol{m}^i \otimes \boldsymbol{d}^i$ with the normal vector $\boldsymbol{m}^i$ and the slip direction $\boldsymbol{d}^i = \frac{1}{b^i}\boldsymbol{b}^i$ of Burgers vector $\boldsymbol{b}^i$. Each considered crystal orientation rotates the slip normal and the slip direction, which affects the Schmid tensor and thus the overall plastic strain. The internal problem is coupled to the external problem by the Orowan equation

$$\partial_t \gamma^i = \upsilon^i b \rho^i_{\text{tot}}, \tag{A.3}$$

where υ^i is the dislocation velocity and ρ^i_{tot} is the total dislocation density of each slip system. The dislocation velocity is calculated by

$$\upsilon^i = \frac{b^i}{B} \operatorname{sgn}\left(\tau^i_{\text{eff}}\right) \max\left(0, |\tau^i_{\text{eff}} - \tau^i_y|\right) \tag{A.4}$$

with the drag coefficient B, the yield stress τ^i_y and the effective shear stress τ^i_{eff}. For the yield stress, we use the formulation $\tau^i_y = \mu b^i \sqrt{a_{ji}\rho^i}$ described by Franciosi [48], where the interaction matrix a_{ji} quantifies the strength of different reaction mechanisms. The effective shear stress τ^i_{eff} equals to $\tau^i - \tau^i_b$, which considers the back stress τ^i_b described in more detail in Groma *et al* [64]. The total dislocation density ρ^i_{tot} is the sum of mobile ρ^i_{mob} and network ρ^i_{net} dislocation density. The network dislocation density ρ^i_{net} incorporates the sessile dislocation network described by the Lomer dislocation density ρ^i_{lom} and the mobile dislocation network described by the stabilized dislocation density ρ^i_{sta}. The dislocation microstructure evolution of the CDD formulation consists of a set of evolution equations incorporating the dislocation densities ρ^i_{mob}, ρ^i_{lom}, ρ^i_{sta} as well as the GND density $\boldsymbol{\kappa}^i$ of edge and screw character and the curvature dislocation density q^i on each slip system i. The evolution equations are described as

$$\begin{aligned} \partial_t \rho^i_{\text{mob}} = {} & -\nabla \cdot \left(\upsilon^i\left(\boldsymbol{\kappa}^i \times n^i\right)\right) + \upsilon^i q^i + \left(\partial_t \bar{\rho}^i_{\text{mob,gli}} + \partial_t \bar{\rho}^i_{\text{mob,cs}} + \partial_t \hat{\rho}^i_{\text{mob}}\right) \\ & - \left(\partial_t \bar{\rho}^i_{\text{mob,lom}} + \partial_t \rho^i_{\text{mob,react}} + \partial_t \rho^i_{\text{mob,cs}}\right), \end{aligned} \tag{A.5}$$

$$\partial_t \boldsymbol{\kappa}^i = \nabla \times \left(\rho^i_{\text{mob}} \upsilon^i m^i\right) + \partial_t \bar{\boldsymbol{\kappa}}^i_{\text{cs}} - \partial_t \boldsymbol{\kappa}^i_{\text{cs}}, \tag{A.6}$$

$$\partial_t \rho^i_{\mathrm{lom}} = \partial_t \bar{\rho}^i_{\mathrm{lom,lom}} - \partial_t \hat{\rho}^i_{\mathrm{lom}}, \tag{A.7}$$

$$\partial_t \rho^i_{\mathrm{sta}} = \partial_t \bar{\rho}^i_{\mathrm{sta,lom}} - \left(\partial_t \hat{\rho}^i_{\mathrm{sta}} + \rho^i_{\mathrm{sta,react}}\right), \tag{A.8}$$

$$\partial_t q^i = -v^i \nabla \cdot \left(\frac{q^i}{\rho^i} \boldsymbol{\kappa}^i\right) - \boldsymbol{A}^i \nabla^2 v^i + \left(\partial_t \bar{q}^i_{gli} + \partial_t \bar{q}^i_{cs} + \partial_t \bar{q}^i_{\mathrm{prod}}\right) - \left(\partial_t q^i_{\mathrm{react}} + \partial_t q^i_{\mathrm{cs}}\right), \tag{A.9}$$

with the alignment tensor

$$\boldsymbol{A}^i = \frac{1}{2}\left(\left(\rho^i_{\mathrm{mob}} + ||\boldsymbol{\kappa}^i||\right) \frac{\boldsymbol{\kappa}^i}{||\boldsymbol{\kappa}^i||} \otimes \frac{\boldsymbol{\kappa}^i}{||\boldsymbol{\kappa}^i||} + \left(\rho^i_{\mathrm{mob}} - ||\boldsymbol{\kappa}^i||\right) \frac{\boldsymbol{\kappa}^i}{||\boldsymbol{\kappa}^i||} \otimes \frac{\boldsymbol{\kappa}^i}{||\boldsymbol{\kappa}^i||}\right). \tag{A.10}$$

Hereby, notations with $\bar{()}$ indicate the evolution based on reaction mechanisms and $\hat{()}$ indicate the evolution based on Lomer dissolution. The considered Lomer, glissile and collinear dislocation reactions are indicated with $()_{\mathrm{react}}$. The cross slip mechanism is notated by $()_{\mathrm{cs}}$ and is formulated by a probability term

$$P = \min\left(\beta \frac{L^i_{\mathrm{screw}}}{L_0} \frac{\delta t}{t_0} \exp\left(\frac{V_{\mathrm{act}}}{k_b T}\left(\tau^i - \tau_{III}\right)\right), 1\right). \tag{A.11}$$

Here, β is a normalizing constant, L^i_{screw} is the average length of a screw dislocation, L_0 and t_0 are reference values for the length and the time, respectively, δt is the simulation time step, k_{B} is the Boltzmann constant, T is the temperature, and V_{act} and τ_{III} are the material dependent activation volume and stage 3 transition shear stress, respectively. For estimating the bow-out of cross-slip, a factor $C_{\mathrm{bow-out}}$ is introduced that is multiplied to the $\boldsymbol{\kappa}^i_{\mathrm{cs}}$ and q^i_{cs}. It describes if a generated screw dislocation is line like or half-circle like by

$$C_{\mathrm{bow-out}} = \begin{cases} 1, & \tau^i \leqslant C_{\mathrm{fr,multi}} \tau^i_{\mathrm{crit}} \\ \frac{\pi}{2}, & \tau^i < C_{\mathrm{fr,multi}} \tau^i_{\mathrm{crit}} \end{cases} \tag{A.12}$$

where $C_{\mathrm{fr,multi}}$ is a constant between 0 and 1. τ^i_{crit} is a critical shear stress defined as

$$\tau^i_{\mathrm{crit}} = \max\left(\frac{1}{C_{\mathrm{fr}}} \mu b \sqrt{\rho^i_{\mathrm{tot}}}, \tau_{\mathrm{CRSS}}\right). \tag{A.13}$$

τ_{CRSS} is a reference value for the critical shear stress in case of vanishing dislocation densities and C_{fr} is a constant to set adjust the average dislocation nucleation length described by $l^i = C_{\mathrm{fr}}(\rho^i_{\mathrm{tot}})^{-0.5}$. In addition to the varied parameters explained in section 2.3.1, table A1 displays the parameters, which are not varied in this work.

Appendix B. Formulation of the surrogate model candidates

Here, we provide a detailed description of our surrogate model candidates. We define the surrogate model inputs and outputs using the following scheme. The estimator function $\hat{f}$ takes $x \in X$ as inputs and $y \in Y$ as outputs at time steps $t = 0, .., T$. Each input x consists of a set of P parameters. We divide into point-wise predictors and sequence-wise predictors in the following.

Point-wise prediction candidates

k-nearest neighbor Nearest neighbor models are regression models that are based on the assumption that samples in close proximity also share a high similarity. To get an estimate for

Table A1. CDD parameters with constant values.

Parameters	Values
Drag coefficient	5×10^{-5} sPa
Interaction matrix a_{ji}	$a_{\text{self}} = 0.300$ $a_{\text{coplanar}} = 0.152$ $a_{\text{Hirth}} = 0.083$ $a_{\text{Lomer}} = 0.326$ $a_{\text{glissile}} = 0.661$ $a_{\text{collinear}} = 0.578$
Cross-slip	$\beta = 10^5$ $L_0 = 1\ \mu$m $t_0 = 1$
Reaction constant	$C_{\text{gli}} = 0.128$ $C_{\text{copl}} = 0.0$

a new sample $x' \in X$, we use the k examples found in the training set that are most similar to x'. We call them the neighborhood. The euclidean distance is a common measure for similarity. Given a training dataset X with targets Y, we can compute the k-Nearest Neighbour (k-NN) estimate as

$$\hat{y} = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i \tag{B.1}$$

using $x \in X$, $y \in Y$ and $N_k(x)$ as the neighbourhood of x that minimize the distance metric. If we find two neighbors that are equally distant, we select them randomly.

Response surface method The RSM is a heuristic to find a response function, which is a latent function given by the simulation, i.e. the relation between our input parameters and outputs. This leads to a response surface, which is model by an nth-order polynomial regression to approximate the response surface. The target estimate of a second-order model using least square regression has the form

$$\hat{y} = \beta_0 + \sum_{j}^{m} (\beta_j x_j) + \sum_{j}^{m} \sum_{j'}^{m} (\beta_{j,j'} x_j x_{j'}) + \epsilon, \tag{B.2}$$

where β_i is the ith coefficient and ϵ the residual. Second-order RSM models can capture non-linear dependencies.

Gradient boosted tree Gradient boosted trees (GBTs) are tree-based models built from multiple decision trees. Boosting is an ensemble method to reduce the variance of a single estimator. An ensemble consists of multiple estimators and is an estimator itself. The aggregated estimate of the GBT is

$$\hat{y} = \hat{f}(X) = \hat{f}_0 + \sum_{j}^{b} \hat{f}_j(X), \tag{B.3}$$

where λ is the learning rate of the boosted tree, b is the number of boosting trees and $\hat{f}_j$ is the j_{th} boosting tree, which will be minimized in the direction of the squared-loss function $Y-\hat{f}_{j-1}(X)$.

Sequence-wise prediction candidates

Recurrent neural networks Recurrent neural networks (RNNs) are ANNs, which has forward and backward connections between neurons. A neuron is a unit that takes a number of input values and computes a corresponding output. It processes the input by weighting the input and applying an activation function on the result of each neuron. By giving the ANN feedback of the prediction based the incorporation of information past inputs in processing of future inputs enables forward and backward transfer of information. An optimization algorithm computes the loss function, which is the difference between predicted outputs and actual outputs, and iteratively reduces the desired loss function.

Long short-term memory A common implementation of RNNs for data sequences are LSTM Networks, originally proposed by Hochreiter and Schmidhuber [65] and extended by Gers *et al* [66] that is commonly used today. In short, LSTMs retain a memory and are fed by the hidden state h_t, the cell state c_t and the input vector x. Each time an LSTM processes an input, it updates its cell state, makes a prediction, and then uses its feedback connections to pass on its cell state to the next iteration. It processes inputs, includes memorized information, and memorizes parts of the information. A so-called forget gate resets the cell state. For more detailed information, we refer to the aforementioned original literature. The formal notation of the LSTM estimate is

$$\hat{y} = \hat{f}(x,t) = \begin{cases} \text{LSTM}\left([y_0, x], h_t, c_t\right) & \text{if}\, t = 0 \\ \text{LSTM}\left(\left[\hat{f}(x,t-1), x\right], h_t, c_t\right) & \text{if}\, t > 0 \end{cases}. \tag{B.4}$$

The square brackets represent concatenation. In this work, we subdivide the LSTM into two different models: First, we feed the parameter space together with the previous prediction as the input vector into the LSTM, which we call a LSTM with repeated inputs (*RepInLSTM*). Second, we feed the previous prediction as the input vector and the parameter space as hidden state into the LSTM, which we call LSTM with raw hidden state (*RawInLSTM*). Because for the first time step, no previous prediction is available, the initial hidden state h_0, cell state c_0 and target y_0 need to be integrated into the LSTM. We use a dummy for the initial target, which is $y_0 = 0$, which assumes, e.g. zero initial stress. For the *RepInLSTM*, we choose $h_0 = 0$ and $c_0 = 0$, and for the *RawInLSTM*, we choose $h_0 = x$ and $c_0 = 0$. We chose to develop two LSTM models, since we wanted to investigate to what extent it differs whether the input parameters are initially specified as hidden states or whether they are specified anew at each point in time.

Appendix C. Hyperparameter tuning of the surrogate model candidates

The surrogate model performance can strongly depend on the considered hyperparameter choices. We conducted a grid search using a 10-fold cross-validation for each model candidate. The performance of k-NN models is affected by k. Since k-NN is distance-based, we added a constraint, which distinguishes between the closest k neighbors across all dimensions and across each dimension individually. For RSM, the polynomial degree p mainly affects the

Table C1. Surrogate model hyperparameter grid.

Model	Hyperparameters
k-NN	$k \in \{1,3,5,10\}$ Distance across all dimensions $\in \{\text{True},\text{False}\}$
RSM	$p \in \{2,3,4,5\}$
GBT	Maximal tree depth $\in \{2,10,-1\}$ Number of trees $\in \{100,500\}$ Learning rate $\in \{0.01,0.1,1\}$ Early stopping patience $\in \{-1,1,10\}$
LSTM	Batch size $\in \{64,128\}$ Adaptive learning rate estimator $\in \{\text{True},\text{False}\}$

model performance. The model performance of the tree-based GBT is impacted by the maximal depth of the trees, the number of trees and the learning rate. For LSTM, several hyperparamters have impact on the model such as the number of hidden layers, learning rate or batch size. Here, we choose 30 hidden layers for all LSTMs and use ADAM as an optimizer [67]. We apply an adaptive learning rate estimator to find the most appropriate learning rate [68]. We define an early stopping by a minimal improvement parameter of $\delta_{\text{min}} = 1 \times 10^{-6}$ and we vary the batch size. We use the same setup of hyperparameters for the considered LSTM candidates. The hyperparameter grid of each model candidate is shown in table C1.

ORCID iDs

Balduin Katzer https://orcid.org/0000-0002-2236-0256
Katrin Schulz https://orcid.org/0000-0001-6427-0948

References

[1] Lomer W 1951 *London, Edinburgh Dublin Phil. Mag. J. Sci.* **42** 1327–31
[2] Hirth J P 1961 *J. Appl. Phys.* **32** 700–6
[3] Bertin N, Sills R B and Cai W 2020 *Annu. Rev. Mater. Res.* **50** 437–64
[4] van der Giessen E *et al* 2020 *Modelling Simul. Mater. Sci. Eng.* **28** 043001
[5] Roters F, Eisenlohr P, Hantcherli L, Tjahjanto D, Bieler T and Raabe D 2010 *Acta Mater.* **58** 1152–211
[6] Pollock T M and LeSar R 2013 *Curr. Opin. Solid State Mater. Sci.* **17** 10–18
[7] Dingreville R, Karnesky R A, Puel G and Schmitt J H 2015 *J. Mater. Sci.* **51** 1178–203
[8] Kiener D, Motz C, Dehm G and Pippan R 2009 *Int. J. Mater. Res.* **100** 1074–87
[9] Dehm G, Jaya B, Raghavan R and Kirchlechner C 2018 *Acta Mater.* **142** 248–82
[10] Coman C D 2019 *Linear Elasticity: General Considerations and Boundary-Value Problems* (Springer) pp 243–80
[11] Pilania G, Balachandran P V, Gubernatis J E and Lookman T 2020 *Data-Based Methods for Materials Design and Discovery: Basic Ideas and General Methods* vol 1 (Morgan & Claypool Publishers LLC) pp 1–188
[12] Gubernatis J E and Lookman T 2018 *Phys. Rev. Mater.* **2** 120301
[13] Guo K, Yang Z, Yu C H and Buehler M J 2021 *Mater. Horiz.* **8** 1153–72

[14] Bock F E, Aydin R C, Cyron C J, Huber N, Kalidindi S R and Klusemann B 2019 *Front. Mater.* **6** 110
[15] Choudhary K *et al* 2022 *npj Comput. Mater.* **8** 59
[16] de Moraes E A B, D'Elia M and Zayernouri M 2023 *Comput. Methods Appl. Mech. Eng.* **403** 115743
[17] Steinberger D, Song H and Sandfeld S 2019 *Front. Mater.* **6** 141
[18] Kuhn J, Spitz J, Sonnweber-Ribic P, Schneider M and Böhlke T 2021 *Optim. Eng.* **23** 1489–523
[19] Forrester A I J, Sóbester A and Keane A J 2008 *Engineering Design via Surrogate Modelling* (Wiley)
[20] Alizadeh R, Allen J K and Mistree F 2020 *Res. Eng. Des.* **31** 275–98
[21] Bilicz S, Lambert M and Gyimóthy S 2010 *Inverse Problems* **26** 074012
[22] Reimann D, Nidadavolu K, ul Hassan H, Vajragupta N, Glasmachers T, Junker P and Hartmaier A 2019 *Front. Mater.* **6** 181
[23] Anand G, Ghosh S, Zhang L, Anupam A, Freeman C L, Ortner C, Eisenbach M and Kermode J R 2022 *J. Inst. Eng.* D **104** 867–77
[24] Fernández M, Rezaei S, Rezaei Mianroodi J, Fritzen F and Reese S 2020 *Adv. Model. Simul. Eng. Sci.* **7** 1
[25] Mudunuru M K, Panda N, Karra S, Srinivasan G, Chau V T, Rougier E, Hunter A and Viswanathan H S 2019 *Appl. Sci.* **9** 2706
[26] Liang L, Liu M, Martin C and Sun W 2018 *J. R. Soc. Interface* **15** 20170844
[27] Pandey A and Pokharel R 2021 *Scr. Mater.* **193** 1–5
[28] Montes de Oca Zapiain D, Stewart J A and Dingreville R 2021 *npj Comput. Mater.* **7** 3
[29] Ktari Z, Leitão C, Prates P A and Khalfallah A 2021 *Mech. Mater.* **153** 103673
[30] Raissi M, Perdikaris P and Karniadakis G 2019 *J. Comput. Phys.* **378** 686–707
[31] Karniadakis G E, Kevrekidis I G, Lu L, Perdikaris P, Wang S and Yang L 2021 *Nat. Rev. Phys.* **3** 422–40
[32] Haghighat E, Raissi M, Moure A, Gomez H and Juanes R 2021 *Comput. Methods Appl. Mech. Eng.* **379** 113741
[33] Xu C, Cao B T, Yuan Y and Meschke G 2023 *Comput. Methods Appl. Mech. Eng.* **405** 115852
[34] Ferdousi S, Chen Q, Soltani M, Zhu J, Cao P, Choi W, Advincula R and Jiang Y 2021 *Sci. Rep.* **11** 14330
[35] Andrade-Campos A, Bastos N, Conde M, Gonçalves M, Henriques J, Lourenço R, Martins J M P, Oliveira M G, Prates P and Rumor L 2022 *IOP Conf. Ser.: Mater. Sci. Eng.* **1238** 012059
[36] Schulte R, Karca C, Ostwald R and Menzel A 2023 *Eur. J. Mech.* A **98** 104854
[37] Tyulyukovskiy E and Huber N 2006 *J. Mater. Res.* **21** 664–76
[38] Klötzer D, Ullner C, Tyulyukovskiy E and Huber N 2006 *J. Mater. Res.* **21** 677–84
[39] Lu L, Dao M, Kumar P, Ramamurty U, Karniadakis G E and Suresh S 2020 *Proc. Natl Acad. Sci.* **117** 7052–62
[40] Lew A J, Jin K and Buehler M J 2023 *npj Comput. Mater.* **9** 80
[41] Schulz K, Wagner L and Wieners C 2019 *Int. J. Plast.* **120** 248–61
[42] Sudmanns M, Stricker M, Weygand D, Hochrainer T and Schulz K 2019 *J. Mech. Phys. Solids* **132** 103695
[43] Sudmanns M, Bach J, Weygand D and Schulz K 2020 *Modelling Simul. Mater. Sci. Eng.* **28** 065001
[44] Zoller K and Schulz K 2020 *Acta Mater.* **191** 198–210
[45] Zoller K, Kalácska S, Ispánovity P D and Schulz K 2021 *C. R. Physique* **22** 267–93
[46] Zoller K, Gruber P, Ziemann M, Görtz A, Gumbsch P and Schulz K 2023 *Comput. Mater. Sci.* **216** 111839
[47] Katzer B, Zoller K, Bermuth J, Weygand D and Schulz K 2023 *Scr. Mater.* **226** 115232
[48] Franciosi P 1985 *Acta Metall.* **33** 1601–12
[49] Akhondzadeh S, Sills R B, Bertin N and Cai W 2020 *J. Mech. Phys. Solids* **145** 104152
[50] Verdier M, Fivel M and Groma I 1998 *Modelling Simul. Mater. Sci. Eng.* **6** 755–70
[51] Katzer B, Zoller K, Weygand D and Schulz K 2022 *J. Mech. Phys. Solids* **168** 105042
[52] Naser M Z and Alavi A H 2021 *Architecture, Structures and Construction* (Springer)
[53] Hastie T, Tibshirani R and Friedman J 2009 *The Elements of Statistical Learning* (Springer)
[54] Kiener D, Guruprasad P, Keralavarma S, Dehm G and Benzerga A 2011 *Acta Mater.* **59** 3825–40
[55] Gorji M B, Mozaffar M, Heidenreich J N, Cao J and Mohr D 2020 *J. Mech. Phys. Solids* **143** 103972
[56] Farizhandi A A K and Mamivand M 2023 *Comput. Mater. Sci.* **223** 112110
[57] Bonatti C, Berisha B and Mohr D 2022 *Int. J. Plast.* **158** 103430

[58] Fan H, Wang Q, El-Awady J A, Raabe D and Zaiser M 2021 *Nat. Commun.* **12** 1845
[59] Akhondzadeh S, Bertin N, Sills R B and Cai W 2021 *Mater. Theory* **5** 2
[60] Starkey K, Hochrainer T and El-Azab A 2022 *J. Mech. Phys. Solids* **158** 104685
[61] Katzer B, Betsche D, Böhm K, Weygand D and Schulz K 2024 *Scr. Mater.* **240** 115841
[62] Hochrainer T, Zaiser M and Gumbsch P 2007 *Phil. Mag.* **87** 1261–82
[63] Hochrainer T, Sandfeld S, Zaiser M and Gumbsch P 2014 *J. Mech. Phys. Solids* **63** 167–78
[64] Groma I, Csikor F and Zaiser M 2003 *Acta Mater.* **51** 1271–81
[65] Hochreiter S and Schmidhuber J 1997 *Neural Comput.* **9** 1735–80
[66] Gers F A, Schmidhuber J and Cummins F 2000 *Neural Comput.* **12** 2451–71
[67] Kingma D P and Ba J 2014 Adam: a method for stochastic optimization (arXiv:1412.6980)
[68] Smith L N 2015 Cyclical learning rates for training neural networks (arXiv:1506.01186)