

Multi-objective surrogate-assisted calibration of CPFEM models using macroscopic response and *in situ* EBSD measurements of grain reorientation trajectories



J. Choi ^{a,b}, O. Muránsky ^{a,b,*}, M.C. Messner ^c, T. Wei ^a, T. Hu ^c, J.J. Kruzic ^b,
M.D. McMurtrey ^d

^a Australian Nuclear Science and Technology Organisation (ANSTO), Lucas Heights, NSW, Australia

^b School of Mechanical and Manufacturing Engineering, UNSW Sydney, Sydney, NSW 2052, Australia

^c Applied Materials Division, Argonne National Laboratory, Lemont IL 60439, USA

^d Idaho National Laboratory, Idaho Falls, ID 83415, USA

ARTICLE INFO

Keywords:

Crystal plasticity finite element method
Multi-objective optimisation
Surrogate modelling
In situ electron back-scatter diffraction (EBSD)
Grain reorientation trajectories

ABSTRACT

Crystal plasticity finite element method (CPFEM) models are widely used to simulate the deformation behaviour of polycrystalline materials, but their calibration is often limited by their high computational cost and the non-convexity of the optimisation landscape. This study develops a multi-objective surrogate-assisted calibration workflow that couples a multi-objective genetic algorithm (MOGA) with an adaptively trained deep neural network (DNN) surrogate model to efficiently identify CPFEM parameters from experimental data. The workflow is demonstrated on three crystal plasticity (CP) formulations of increasing complexity — Voce hardening (VH), two-coefficient latent hardening (LH2), and six-coefficient latent hardening (LH6) — using *in situ* electron backscatter diffraction (EBSD) measurements of Alloy 617 under uniaxial tensile loading. The CPFEM models are calibrated against the experimentally observed stress-strain response and reorientation trajectories of eight grains, then validated against eight additional trajectories and overall texture evolution. Across the CP formulations, the macroscopic response was reproduced reliably, while differences emerged in the robustness and accuracy of the grain-scale predictions. Including grain reorientation trajectories in the multi-objective calibration improved texture evolution predictions and filtered out physically inconsistent parameter sets that can arise from calibrating against only the stress-strain data. The workflow also demonstrates good transferability of calibrated parameters from a low- to a high-fidelity microstructural model. These results provide practical guidance for integrating *in situ* microstructural data into CPFEM through efficient, repeatable, and physically meaningful multi-objective calibration.

1. Introduction

Material models are of technological importance in materials science and engineering, enabling the prediction of material behaviour under various conditions. Empirical and semi-empirical elasto-plastic models are widely utilised in engineering applications due to their straightforward implementation and computational efficiency [1–4], relying primarily on experimental data to establish relationships between stress and strain. However, these models do not directly consider the alloy's microstructural influences on their elasto-plastic behaviour. To address this limitation, microstructure-informed models have been developed

and are widely adopted by the research community [5–9]. In particular, crystal plasticity finite element method (CPFEM) models simulate the elasto-plastic deformation of polycrystalline materials at the crystal scale by considering the alloy's microstructure, deformation mechanisms, and grain-to-grain interactions [10,11]. While CPFEM models offer enhanced accuracy by integrating detailed microstructural information, they are significantly more computationally expensive, posing challenges for their use in practical engineering applications [12,13].

One of the first steps in the deployment of CPFEM models in engineering finite element analyses is to calibrate these models against experimental data to ensure that the simulated responses accurately

* Corresponding author.

E-mail address: ondrej.muransky@ansto.gov.au (O. Muránsky).

reflect observed material behaviour under simple experimental loading conditions. The calibration process of CPFEM models involves the numerical identification of the constitutive material parameters that govern the material's behaviour at the microstructural level, such as the slip system strengths, hardening coefficients, and interaction terms. The existing literature frequently relies on manual trial-and-error methods to calibrate CPFEM models against a single objective, such as a uniaxial stress-strain curve. However, this approach depends heavily on prior knowledge of the constitutive parameters and can be extremely time consuming, given the extensive parameter space of CPFEM models necessitating numerous evaluations [12–14]. In addition, calibrating against a single objective often leads to an under-constrained problem where multiple sets of parameters yield similar results, leading to ambiguity and reduced reliability in the calibrated models. Hence, many authors use additional data as objectives to constrain the large parameter space, such as the internal lattice strains of grain families [15–17], reorientation trajectories of individual grains [18], or overall texture evolution [19–21]. However, the calibration of models against multiple objectives introduces complexities that make the process even more computationally expensive. As such, various optimisation methodologies have been employed for the multi-objective calibration of CPFEM models [10,14,18,22].

These optimisation methods are broadly categorised into gradient-based and direct-search optimisation methods, each with their own advantages and limitations. Gradient-based optimisation methods, such as gradient descent, Adam, and Newton's methods, are efficient and well-suited for high-dimensional problems [23–25]. However, these methods rely heavily on the initial parameter values and are highly susceptible to converging towards local minima, particularly in the complex and non-linear parameter spaces characteristic of CPFEM simulations [12,23]. Gradient-based methods also require differentiability in their objective functions, in which noise in the experimental data could deteriorate convergence behaviour [12]. In contrast, direct-search optimisation methods do not rely on gradient information, making them generally more robust and better suited for exploring diverse regions of large parameter spaces [3,26]. These methods are also more effective at locating global minima and can handle multiple objective functions simultaneously. A popular example is the genetic algorithm (GA), which has been used extensively to identify the constitutive parameters in various crystal plasticity (CP) constitutive laws [12,27,28]. Despite these advantages, direct-search methods typically require a substantial number of evaluations to achieve convergence, which can be prohibitive given the high computational cost of each CPFEM simulation.

To address the computational challenges associated with calibrating CPFEM models using direct-search optimisation, surrogate modelling techniques have emerged as a viable solution to approximate the input-output relationships of computationally expensive models with significantly reduced computational costs. A notable example of such techniques is symbolic regression, which has recently been used in Ref. [29] to derive interpretable mathematical relationships from CPFEM simulations of Ti-6Al-4 V. Other well-established techniques include the response surface methodology (RSM) [30], Kriging method [31], and deep neural networks (DNNs) [32]. By developing a surrogate model, the calibration process can leverage these faster approximations to explore the parameter space more efficiently. This enables the application of direct-search optimisation methods to the surrogate model, thereby minimising the number of required CPFEM simulations [33,34]. Once an optimal set of constitutive parameters is identified through the surrogate-based optimisation, these parameters can be used to calibrate the original CPFEM model with much lower computational overhead. Consequently, surrogate modelling not only alleviates the computational burden but also facilitates the practical application of robust optimisation techniques in the calibration of complex microstructure-informed material models [10,33,34].

In the present work, *in situ* electron backscatter diffraction (EBSD)

was performed on an Alloy 617 specimen under uniaxial tensile loading at room temperature to obtain both the macroscopic stress-strain response and grain-scale reorientation trajectories. These experimental datasets were used to calibrate CPFEM models with three CP formulations: (i) a Voce hardening (VH) formulation, (ii) a two-coefficient latent hardening (LH2) formulation, and (iii) a six-coefficient latent hardening (LH6) formulation — implemented within low- and high-fidelity finite element (FE) microstructural representations of the specimen gauge. The calibration employed a multi-objective surrogate-assisted workflow in which a multi-objective genetic algorithm (MOGA) was coupled with an adaptively trained deep neural network (DNN) surrogate model to efficiently identify CP parameters that reproduce both the macroscopic stress-strain curve and the reorientation trajectories of eight selected grains in the low-fidelity model. The performance of the CPFEM models was then validated against the trajectories of eight additional grains and the overall texture evolution. Finally, the optimised parameter sets from the low-fidelity calibrations were applied to the corresponding high-fidelity CPFEM models to assess their transferability across different microstructural fidelities within the FE framework.

2. Experiment

The following section describes the experimental procedures and datasets used in this study. The section includes details on the Alloy 617 material (Section 2.1), the data acquisition during the *in situ* EBSD-based tensile testing (Section 2.2), the analysis of grain reorientation across EBSD maps (Section 2.3), and the generation of finite element (FE) microstructural models from the initial EBSD map (Section 2.4).

2.1. Material

Alloy 617 is a solid solution and precipitation strengthened Ni-Cr-Co alloy, offering an exceptional combination of thermal, mechanical, physical, and corrosion-resistance [35]. In particular, the alloy maintains high strength and excellent creep resistance at elevated temperatures. In addition, its corrosion resistance supports applications in chemical processing environments [36,37], while its high ductility and toughness at lower temperatures allow for use in certain cryogenic systems [38]. These properties allow Alloy 617 to perform reliably across a broad temperature range, expanding its versatility in critical engineering applications. In particular, the nuclear industry is pursuing the alloy as structural material for intermediate heat exchangers in the next generation (Gen IV) of nuclear reactor systems [39], which are expected to operate at significantly higher temperatures (up to 950°C) [35,40,41] than the current fleet of nuclear reactor systems (300°C). The aerospace industry also employs Alloy 617 in the combustion chambers and exhaust systems of jet engines and gas turbines due to the alloy's ability to withstand thermal and mechanical stresses [42–44]. The alloy studied in the present work has been provided by Idaho National Laboratory and was a part of the material used for the development of the ASME Section III, Division 5 code case permitting the use of Alloy 617 in nuclear high-temperature applications [45].

2.2. Data acquisition during tensile loading

The *in situ* tensile testing was performed using a Zeiss UltraPlus scanning electron microscope (SEM) equipped with a Kammerath & Weiss GmbH tensile loading stage. The test was conducted at room temperature at a post-yield displacement rate of $0.224\mu\text{m s}^{-1}$ and strain rate of 10^{-4}s^{-1} , following a stepwise loading approach in crosshead control mode. At each step, the miniaturised tensile specimen (see Fig. 1a) was first loaded to a specified displacement level and then unloaded to 15MPa to facilitate the acquisition of the electron back-scatter diffraction (EBSD) data. The EBSD measurements were performed using an electron beam energy of 20kV, with each map taking

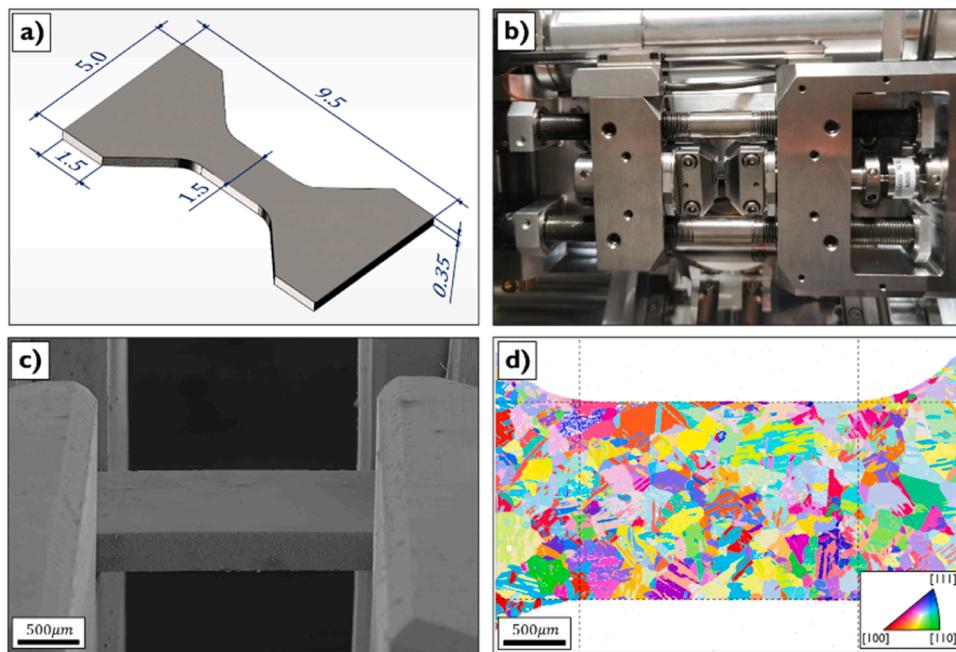


Fig. 1. Images of the experimental data acquisition — a) dimensioned drawing of the specimen, b) photograph of the specimen in the loading stage, c) SEM micrograph of the specimen in the loading stage, and d) initial EBSD map shown as IPF-x coloured map.

approximately 30 minutes to acquire. After the data was collected, the specimen was reloaded to a higher displacement level in the subsequent step and unloaded again for further EBSD mapping. In total, 24 EBSD orientation maps were acquired, each corresponding to a different level of imparted plastic strain. These maps are presented in [Appendix A](#).

The plot in [Fig. 2](#) presents the true stress-strain response for the miniaturised tensile specimen with the unloading/reloading portions removed. In this plot, the dashed blue lines mark the points at which the EBSD datasets were collected. The plot also compares this response with that of a standard-sized specimen tested at Idaho National Laboratory under identical strain rate and temperature conditions following ASTM

E8-11 [\[46,47\]](#). The close agreement of the two responses suggest that the stepwise loading procedure and the specimen-size differences have minimal influence on the measured mechanical behaviour. Additionally, the EBSD maps were acquired at a magnification of $30 \times$ with a $120\mu\text{m}$ aperture and a spatial resolution of $5\mu\text{m}$. This configuration enabled relatively fast data acquisition of around 10 to 30 min per map, depending on the level of imparted plastic strain. The map size of $2865\mu\text{m} \times 1920\mu\text{m}$ was sufficient to capture nearly the entire initial gauge section of the specimen, which has approximate dimensions of $2200\mu\text{m} \times 1560\mu\text{m}$. However, as shown in [Fig. 1c](#), partial shadowing from the loading stage obstructed the EBSD detector, resulting in missing data along the upper edges of the specimen.

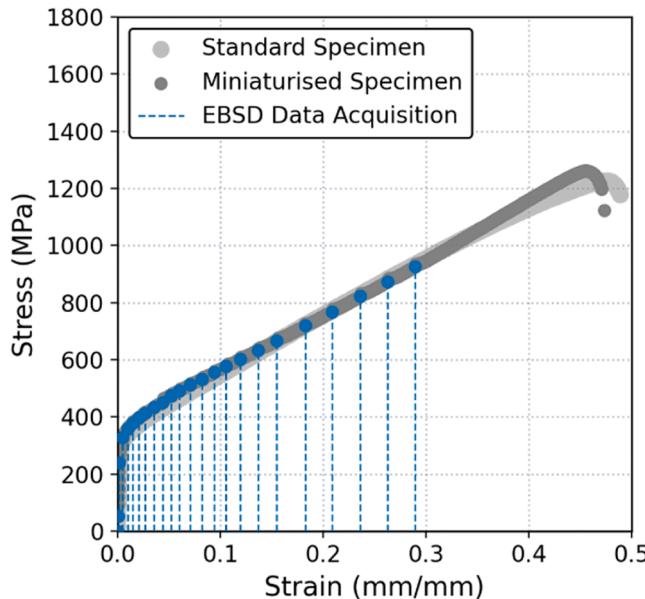


Fig. 2. Comparison of the recorded true stress-strain curves obtained during the in-situ EBSD experiment on the miniaturised tensile specimen ([Fig. 1](#)) and on a standard tensile specimen [\[47\]](#). Dashed blue lines indicate the stress-strain positions at which the EBSD datasets were collected ([Fig. 19](#)).

2.3. Grain reorientation across EBSD orientation maps

Grain tracking is essential for understanding the evolution of crystallographic texture as a function of plastic deformation in the specimen gauge during tensile loading. In this work, an automated grain-tracking algorithm was developed to monitor the orientation of individual grains across the sequence of 24 EBSD maps, as described in [Appendix B](#). The algorithm operates by matching grains between consecutive maps based on the normalised distance between their centroids and the similarity of their crystallographic orientations. A grain in a subsequent EBSD map is considered a match if its centroid lies within a predefined distance threshold and its orientation difference is within an acceptable tolerance relative to the grain in the preceding map. This iterative matching procedure enables continuous tracking of individual grains from the initial undeformed state to high levels of plastic strain. Once tracked, the reorientation trajectories of individual grains were constructed by stitching their crystallographic orientations throughout their deformation sequence. The constructed reorientation trajectories exhibited notable noise due to measurement variability, which was mitigated through smoothing using spherical linear interpolation (SLERP). Once smoothed, the reorientation trajectories were manually inspected to ensure that they did not significantly deviate from the raw data.

In this study, the reorientation trajectories of 16 grains ($g_{1..16}$) were selected from the tracked EBSD dataset for use in the calibration and validation procedures. These grains were chosen to span a broad range

of initial orientations relative to the loading axis, to provide a diverse spatial distribution across the specimen, and to ensure well-tracked reorientation trajectories throughout the deformation process. The location of the selected grains on the initial ($\epsilon = 0.0\%$) and final ($\epsilon = 29.0\%$) EBSD maps are shown in Fig. 3, while their deformation paths are summarised in Appendix B. In addition, the inverse pole figures (IPFs) in Fig. 4 compare the raw and SLERP-smoothed reorientation trajectories of $g_{1..16}$ in the loading direction (x).

2.4. Finite element (FE) model generation

Initial analysis of the collected EBSD orientation data and subsequent grain segmentation was performed using the MTEX MATLAB toolbox [48]. In the first EBSD map (Fig. 1d), any missing data was reconstructed by extrapolating crystal orientations from neighbouring regions, which enabled the generation of finite element (FE) microstructural models of the specimen gauge. This EBSD dataset contains a total of 571 grains and was acquired before tensile testing at a spatial resolution of $5\mu\text{m}$. The EBSD data was downsampled using the MTEX function, 'reduce', and used to generate microstructural FE models of the specimen gauge. The models were generated by creating a 2.5D volumetric representation from the initial 2D EBSD orientation map. This was accomplished by replicating the 2D data along the specimen thickness (z-axis), thereby forming a voxelated volume. The voxellation was then converted into a hexahedral (hex) mesh using the 'Sculpt' module in Cubit Coreform 2022.4 [49] and scaled to $300\mu\text{m}$ to reflect the actual thickness of the specimen. This process produces a 2.5D representation of the specimen [50], since the mesh extends the 2D surface microstructure into a third dimension but does not account for variations in grain geometry through the depth of the specimen. In this conversion process, voxels corresponding to individual grains were mapped onto FE blocks, where each block is formed from a subset of hex elements used to represent that grain. It should be noted that some grains were lost during the conversion from the voxellation into the 2.5D mesh due to limitations inherent in the voxel merging and meshing process.

Using the meshing processes, a low- and high-fidelity FE model was generated at a downsampled resolution of $40\mu\text{m}$ (containing 132 grains) and $10\mu\text{m}$ (containing 481 grains), respectively, as shown in Fig. 5. The low-fidelity model was generated with three layers of elements along the specimen's thickness, containing a total of 17,109 hex elements. In contrast, the high-fidelity mesh was generated with ten layers of elements, containing a total of 574,560 hex elements. While the high-fidelity model offered a more accurate representation of the microstructure, its higher element count substantially amplified the computational costs of each CPFEM evaluation.

3. Microstructure-informed models

In this study, the crystal plasticity finite element method (CPFEM) was utilised to simulate the multiscale behaviour of the Alloy 617 specimen under uniaxial tensile loading. The crystal plasticity (CP)

formulations were implemented in the nuclear engineering material library (NEML) [51] to describe the mechanical and microstructural response of the tested alloy. The CP formulations were integrated with the finite element (FE) models through Deer [52], an application of the open-source multiphysics object-oriented simulation environment (MOOSE) [53] that links to NEML. The formulations include a Voce hardening (VH) formulation [54], a two-coefficient latent hardening (LH2) formulation [55], and a six-coefficient latent hardening (LH6) formulation [55]. These phenomenological formulations were chosen to represent progressively increasing levels of model complexity and parameterisation, while remaining widely used in CPFEM literature. While physics-based hardening formulations (e.g., Kocks-Mecking, dislocation-density-based, or GND-based formulations) could provide additional mechanistic insights [56–58], they were not considered in this work as they are not yet available within the current MOOSE-NEML framework. Accordingly, the following section will provide an overview of CP (Section 3.1), as well as descriptions of the VH (Section 3.2), LH2 (Section 3.3), and LH6 (Section 3.4) formulations. In addition, a summary of the constitutive parameters for all three formulation can be found in Appendix C.

3.1. Crystal plasticity formulation

Crystal plasticity (CP) formulations were defined to simulate the multiscale behaviour of Alloy 617. In these formulations, the plastic velocity gradient (L^P) of a single crystal is defined as a sum of the shear deformation of each slip system [γ_α] is defined using the standard power law model, as shown in Eq. (1), where $\dot{\gamma}_0$ represents the reference slip rate, n represents the strain rate sensitivity parameter, τ_α represents the resolved shear stress (RSS), and τ_α^* represents the critical resolved shear stress (CRSS).

$$\dot{\gamma}_\alpha = \dot{\gamma}_0 \left| \frac{\tau_\alpha}{\tau_\alpha^*} \right|^{n-1} \left(\frac{\tau_\alpha}{\tau_\alpha^*} \right) \quad (1)$$

The reference slip rate ($\dot{\gamma}_0$) is defined based on the uniaxial strain rate ($\dot{\epsilon}$) and Taylor factor (M) [60], as shown in Eq. (2). The Taylor factor for face-centred cubic (FCC) polycrystalline metals with random orientations under uniaxial tensile deformation has been previously calculated to be 3.07 using a full constraint Taylor model [60–63]. As such, with a post-yield strain rate of 10^{-4}s^{-1} (Section 2.2), the reference slip rate was set to $3.25 \times 10^{-5}\text{s}^{-1}$ to reduce the rate sensitivity in the model response.

$$\dot{\gamma}_0 = \dot{\epsilon}/M \quad (2)$$

The CRSS of a slip system (τ_α^*) describes the threshold shear stress required to initiate slip. When the RSS exceeds the CRSS, dislocations will begin moving along the slip system, and the crystal will begin plastically deforming [64]. The CP formulations define the CRSS of a slip system as the sum of its static strength (τ_0) and its dynamic strength (τ'_α), as shown in Eq. (3). In this equation, the static strength represents the

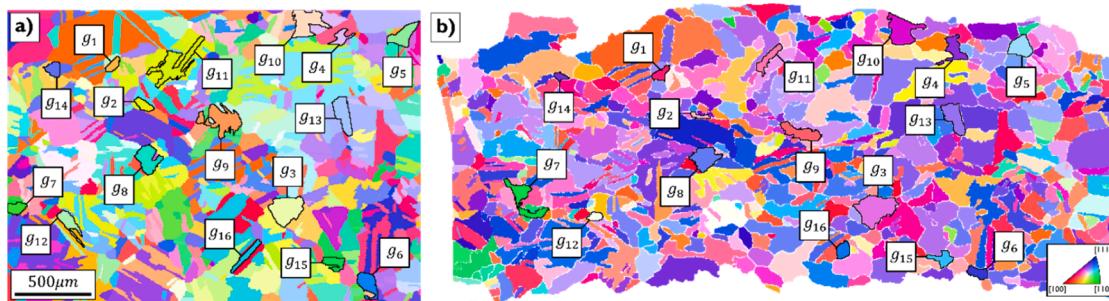


Fig. 3. Selected grains ($g_{1..16}$) on the a) initial EBSD map ($\epsilon = 0.0\%$) and b) final EBSD map ($\epsilon = 29.0\%$) shown as IPF-x coloured maps.

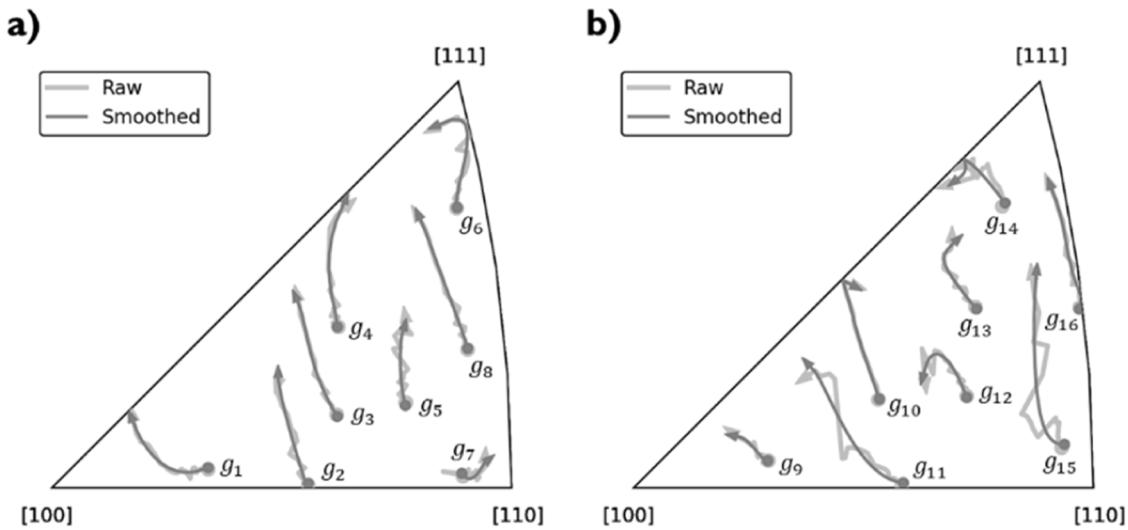


Fig. 4. Inverse pole figures (IPFs) of the raw and SLERP-smoothened reorientation trajectories of grains a) g₁₋₈ and b) g₉₋₁₆ in the x-direction.

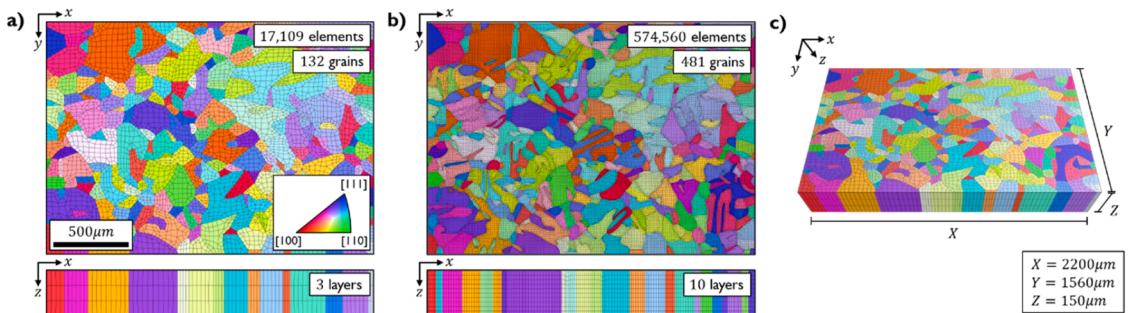


Fig. 5. IPF-x coloured maps of a) the low-fidelity mesh (top and side view), b) the high-fidelity mesh (top and side view), and c) the low-fidelity mesh (angled view).

inherent resistance of all the slip systems, while the dynamic strength represents the additional resistance that evolves based on time-dependent factors.

$$\tau_a^* = \tau_0 + \tau'_a \quad (3)$$

3.2. Voce hardening (VH) formulation

The Voce hardening (VH) formulation [54] assumes that all slip systems share the same dynamic strength (τ'), accounting for all three stages of single crystal plasticity behaviour. The VH formulation is defined in Eq. (4) [54], where b represents the saturation rate parameter and τ_s represents the saturation strength. Hence, the VH formulation is governed by four adjustable constitutive material parameters — τ_0 , n , b , and τ_s .

$$\dot{\tau}' = b(\tau_s - \tau') \sum_{\alpha \in A} |\dot{\gamma}_\alpha| \quad (4)$$

3.3. Two-coefficient latent hardening (LH2) formulation

In contrast, the latent hardening formulations [55] assume that the dynamic strength of the slip systems (τ'_α) evolve independently, accounting for only the first two stages of single crystal plasticity behaviour (i.e., easy glide and linear hardening). The latent hardening formulations simulate hardening behaviour by considering the interaction strengths between slip systems, as shown in Eq. (5) [55]. In this equation, $H_{\alpha\beta}$ represents an element of the interaction matrix that

quantifies how much slip system β contributes to the hardening of slip system α .

$$\dot{\tau}'_\alpha = \sum_{\beta \in A} H_{\alpha\beta} |\dot{\gamma}_\beta| \quad (5)$$

For FCC crystal structures with 12 independent {111}⟨110⟩ slip systems, the interaction matrix is a 12×12 matrix designed to capture the influence of dislocation interactions between the slip systems. The 144 elements of the interaction matrix are typically reduced to two independent coefficients [65–67]. In this two-coefficient interaction matrix (H_2), the diagonal elements (i.e., $\alpha = \beta$) correspond to the self-hardening coefficient ($h_{\alpha\alpha}$) and the off-diagonal elements (i.e., $\alpha \neq \beta$) correspond to the latent hardening coefficient ($h_{\alpha\beta}$). The two-coefficient latent hardening (LH2) formulation employs the H_2 interaction matrix and is thus governed by four adjustable constitutive material parameters — τ_0 , n , $h_{\alpha\alpha}$, and $h_{\alpha\beta}$.

3.4. Six-coefficient latent hardening (LH6) formulation

According to Franciosi *et al.* [68], the elements of the interaction matrix can also be reduced to six independent coefficients due to the diagonal symmetry of the matrix and the ternary symmetry of the four ⟨111⟩ axes [69]. The six-coefficient interaction matrix (H_6) is shown in Eq. (6), in which the diagonal elements correspond to the self-hardening coefficient ($h_{\alpha\alpha}$) and the off-diagonal elements correspond to five distinct latent hardening coefficients (h_1 , h_2 , h_3 , h_4 , and h_5). The latent hardening coefficients represent the strengths of the coplanar, Hirth

lock, collinear, glissile junction, and sessile junction interactions [69, 70], respectively.

$$H_6 = \begin{bmatrix} h_{aa} & h_1 & h_1 & h_3 & h_4 & h_4 & h_2 & h_4 & h_5 & h_2 & h_5 & h_4 \\ h_1 & h_{aa} & h_1 & h_4 & h_2 & h_5 & h_4 & h_3 & h_4 & h_5 & h_2 & h_4 \\ h_1 & h_1 & h_{aa} & h_4 & h_5 & h_2 & h_5 & h_4 & h_2 & h_4 h_4 & h_3 & h_3 \\ h_3 & h_4 & h_4 & h_{aa} & h_1 & h_1 & h_2 & h_5 & h_4 & h_2 & h_4 & h_5 \\ h_4 & h_2 & h_5 & h_1 & h_{aa} & h_1 & h_5 & h_2 & h_4 & h_4 & h_3 & h_4 \\ h_4 & h_5 & h_2 & h_1 & h_1 & h_{aa} & h_4 & h_4 & h_3 & h_5 & h_4 & h_2 \\ h_2 & h_4 & h_5 & h_2 & h_5 & h_4 & h_{aa} & h_1 & h_1 & h_3 & h_4 & h_4 \\ h_4 & h_3 & h_4 & h_5 & h_2 & h_4 & h_1 & h_{aa} & h_1 & h_4 & h_2 & h_5 \\ h_5 & h_4 & h_2 & h_4 & h_3 & h_1 & h_1 & h_1 & h_{aa} & h_4 & h_5 & h_2 \\ h_2 & h_5 & h_4 & h_2 & h_4 & h_5 & h_3 & h_4 & h_4 & h_{aa} & h_1 & h_1 \\ h_5 & h_2 & h_4 & h_4 & h_3 & h_4 & h_4 & h_2 & h_5 & h_1 & h_{aa} & h_1 \\ h_4 & h_4 & h_3 & h_5 & h_4 & h_2 & h_4 & h_5 & h_2 & h_1 & h_1 & h_{aa} \end{bmatrix} \quad (6)$$

The six-coefficient latent hardening (LH6) formulation employs the H_6 interaction matrix and is thus governed by eight adjustable constitutive material parameters — τ_0 , n , h_{aa} , h_1 , h_2 , h_3 , h_4 , and h_5 .

4. Calibration workflow

In this study, the multi-objective genetic algorithm (MOGA) was combined with a deep neural network (DNN) surrogate model to calibrate the CPFEM models against the experimental *in situ* EBSD data. A multi-objective surrogate-assisted calibration workflow was developed to iteratively reduce the discrepancies between the responses produced by the CPFEM and surrogate models. The following section will provide overviews of the MOGA optimisation (Section 4.1) and DNN surrogate model (Section 4.2), as well as describe how the MOGA and surrogate model were integrated into the calibration workflow (Section 4.3).

4.1. Multi-objective genetic algorithm (MOGA) optimisation

The genetic algorithm (GA) is a population-based, direct search optimisation algorithm inspired by the principles of natural selection. In the context of material model calibration, the objective of the GA is to search for a suitable set of constitutive parameters that enables the material model's simulated response to align with the experimentally observed material behaviour. The GA achieves this by iteratively evolving a population of candidate solutions (i.e., sets of constitutive parameters) over multiple generations. Every generation, the GA assesses the optimality of each solution by evaluating the material model and calculating the discrepancies between the experimental and simulated responses through the use of an objective function [3, 26, 71]. The GA then combines the most optimal solutions to produce the next generation of potential solutions, with this process repeated until a pre-defined termination criterion is met.

The multi-objective genetic algorithm (MOGA) combines the concepts of the GA with multi-objective optimisation (MOO) to simultaneously minimise multiple objective functions. The MOGA assesses the solutions based on how well they balance conflicting objectives, with a solution deemed Pareto-efficient if no objective can be improved without compromising another. Over successive generations, the MOGA collects these Pareto-efficient solutions to gradually develop the Pareto front. This allows the MOGA to simultaneously consider multiple objectives without the need for normalisation, weighting factors, or prioritisation schemes. Once the MOGA terminates, the best solution (i.e., the knee point) is identified from the Pareto front based on the minimal squared sum of the normalised objective values, as detailed in Appendix D. The knee-point parameters represent the optimal trade-off between the objectives and are used to calibrate the material model, yielding simulated responses that best align with the experimentally observed material behaviour.

In this study, the non-dominated sorting genetic algorithm-II (NSGA-II) implementation of the MOGA [72] is used to minimise two objective

functions. The first objective function (E_σ) calculates the discrepancies between the experimental and simulated stress-strain curves using the normalised root mean square error (NRMSE), expressed as a dimensionless decimal value. This is shown in Eq. (7), where σ_{exp} represents a stress value in the experimental curve, σ_{sim} represents a stress value in the simulated curve, $\bar{\sigma}_{exp}$ represents the average of stresses in the experimental curve, and N represents the number of stress values to evaluate. This objective function is used to assess the alignment of the experimental and simulated stress-strain responses at $N = 32$ uniformly spaced strain intervals along the curves. Since the calibration focuses on plastic deformation rather than the elastic regime, this resolution is sufficient for capturing the relevant behaviour.

$$E_\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{(\sigma_{exp})_i - (\sigma_{sim})_i}{\bar{\sigma}_{exp}} \right)^2} \quad (7)$$

The second objective function (E_ϕ) calculates discrepancies between the experimental and simulated grain reorientation trajectories using the mean geodesic distance, expressed in radians. This is shown in Eq. (8), where q_{exp} represents a quaternion from an experimental trajectory, q_{sim} represents a quaternion from a simulated trajectory, and G represents the number of grains being evaluated. This objective function is used to assess the alignment of the experimental and simulated reorientation trajectories similarly at $N = 32$ uniformly spaced strain intervals along the trajectories. Note that the geodesic distance was chosen for the objective function over the misorientation to reduce the computational cost of the calibration process. As long as the orientations being compared are fairly close together in the 3D rotation space (i.e., SO(3)), this measure is identical to a misorientation calculation accounting for crystal symmetry.

$$E_\phi = \frac{1}{GN} \sum_{i=1}^G \sum_{j=1}^N 2\cos^{-1} \left((q_{exp})_{ij} \cdot (q_{sim})_{ij} \right) \quad (8)$$

The behaviour of the MOGA is governed by user-defined hyperparameters, such as crossover probability, mutation probability, initial population, number of offspring, and number of generations. Following Ref. [3, 12, 71] and optimisations with similar material models, the values of these hyperparameters were set to 80%, 1%, 100, 100, and 500, respectively. In addition, bounds were also defined for each constitutive parameter to control the shape and size of the MOGA's search space. The bounds for the VH, LH2, and LH6 formulations were initially defined based on literature and adjusted using a Taylor mean field model [51] to ensure the optimisations remained confined to feasible and physically meaningful regions of the parameter space. The parameter bounds and the referenced literature are summarised in Appendix C.

4.2. Deep neural network (DNN) surrogate modelling

The MOGA assesses the optimality of a set of constitutive parameters by evaluating multiple objective functions. When calibrating a material model, these objective functions require the model's simulated response to compare with the experimentally observed material behaviour. In a single calibration, the MOGA will assess numerous parameter sets across multiple generations, which necessitates running the model a substantial number of times. However, the high computational cost of running the CPFEM models makes their direct calibration using the MOGA impractical [3, 12, 73]. The computational burden of the calibration process can be reduced by developing a surrogate model to approximate the behaviour of the CPFEM model more efficiently. Given a set of constitutive parameters, the surrogate model can approximate the stress-strain response and reorientation trajectories simulated by the CPFEM model at a fraction of the original computational cost. By employing the surrogate model to evaluate the objective functions, the MOGA is able to obtain an optimal set of constitutive parameters to

calibrate the CPFEM model much faster by using significantly fewer resources.

In this study, the surrogate model was developed using a deep neural network (DNN) implemented in PyTorch [74]. The input layer of the DNN contains $P + 1$ features, which include the P adjustable constitutive parameters of the CPFEM model and a strain value. The input layer feeds into three hidden layers with 128, 256, and 512 neurons, respectively, each followed by a rectified unit linear activation function. The output layer contains $4G + 1$ features, consisting of four quaternion components for each of the G grains selected as well as a stress value. By predicting the macroscale stresses and grain orientations at various strains, the DNN can approximate the entire stress-strain response and G reorientation trajectories for a given set of constitutive parameters. The training dataset for the DNN was composed of the results of multiple CPFEM simulations. For each simulation, the macroscale stress and G grain orientations were extracted from the simulated material response at 32 evenly spaced strain intervals. Consequently, each CPFEM simulation provided 32 input-output pairs, with each pair corresponding to the $P + 1$ input features and the $4G + 1$ output features. The extracted input-output pairs from all the CPFEM simulations were then compiled to form the training dataset.

Once the training dataset was generated, the input and output features were normalised through linear and logarithmic scaling to lie within a $[0, 1]$ range. The dataset was then partitioned using K-fold cross validation with five splits, ensuring that the entire dataset would be used for training and validating the DNN across five iterations. The DNN was trained using the Adam optimiser and a L_1 regularisation coefficient of 1×10^{-7} . The training process was guided by the mean square error (MSE) loss function to minimise the discrepancies between the DNN-predicted and CPFEM-simulated material responses. The training was performed with an initial learning rate of 1×10^{-3} , which was adaptively reduced by a factor of 10 with a patience of 100 epochs. The DNN was trained for up to 2000 epochs, with early termination triggered when the learning rate dropped below 1×10^{-7} to prevent overfitting.

4.3. Adaptive calibration workflow

In initial tests, surrogate models were trained against a dataset of 16 CPFEM evaluations and calibrated using the MOGA. While the surrogate models reliably reproduced the calibration dataset, the optimised parameters consistently showed poor transferability to the CPFEM models, resulting in substantial discrepancies between the CPFEM responses and the experimental data. While the accuracy of the CPFEM models improved when using larger training datasets, they required additional CPFEM evaluations that significantly increased the computational expense. As such, a multi-objective surrogate-assisted calibration workflow was developed to effectively and efficiently calibrate the CPFEM models, as summarised in Fig. 6.

The workflow begins by sampling the parameter space using the Latin hypercube sampling (LHS) scheme within the bounds defined in Appendix C. To ensure a fair comparison across the CPFEM models, a sampling size of $2P$ was adopted following Refs. [75–77] and a sensitivity study presented in Appendix E, where P represents the number of adjustable constitutive parameters. For each sampled set of parameters, the CPFEM model is evaluated, and the simulated material responses are compiled to generate an initial training dataset. The surrogate model is then trained using the training dataset and calibrated using the MOGA. These MOGA optimisations use an initial population that is seeded with the knee-point parameters (Appendix D) from the previous CPFEM evaluations, which minimised both objective functions (E_σ and E_ϕ) most optimally. The CPFEM model is then evaluated using the optimised parameters, and the discrepancies between the surrogate model's approximations and the CPFEM responses are calculated by repurposing the objective functions in Eqs. (7) and (8).

If $E_\sigma < 1\%$ and $E_\phi < 0.1^\circ$, the calibration workflow terminates, and

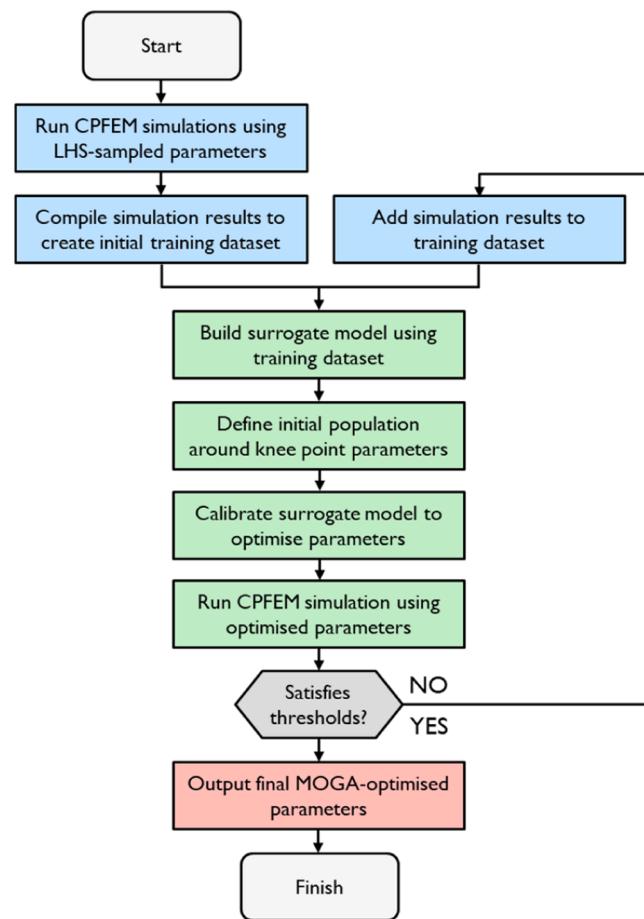


Fig. 6. Flow chart of the multi-objective surrogate-assisted calibration workflow.

the final knee-point parameters are extracted. If the criteria are not met, the CPFEM response at the MOGA-optimised parameters is incorporated into the training dataset, after which the surrogate model is retrained, and the optimisation and evaluation steps are repeated. The iterative refinement of the surrogate model at the MOGA-optimised parameters ensures that the surrogate model is most accurate in the regions of the parameter space that are actively explored by the MOGA. This refinement process continues, with each cycle improving the surrogate model's accuracy, until the accuracy thresholds ($E_\sigma < 1\%$ and $E_\phi < 0.1^\circ$) are satisfied. Although these thresholds were not selected through formal optimisation, preliminary testing indicated that they provided a practical balance between surrogate fidelity and computational cost. Importantly, the termination criterion is governed by the accuracy of the surrogate model in approximating the CPFEM responses, rather than the accuracy of the CPFEM model in reproducing the experimentally observed behaviour. This choice was made to improve the generalisability of the calibration workflow by accommodating for biases in the calibration dataset, such as reorientation trajectories that are inherently less reproducible due to inaccuracies in the FE models.

5. Results and discussion

In this study, a calibration workflow was developed that couples a multi-objective genetic algorithm (MOGA) with an adaptively trained deep neural network (DNN) surrogate model. While the MOGA offers strong global search capabilities for multi-objective optimisation, it typically requires a large number of model evaluations — something that becomes computationally prohibitive when each evaluation involves a computationally expensive CPFEM simulation. To address this,

the direct MOGA–CPFEM interaction is replaced with a DNN surrogate model that was adaptively trained to emulate the CPFEM predictions at a fraction of the computational cost. Each CP formulation (VH, LH2, and LH6) was coupled with a low-fidelity FE model and calibrated five independent times, with a new surrogate model developed in each run to avoid bias from the initial dataset or network realisation. The best-performing parameter set obtained using the low-fidelity CPFEM model was then applied to the high-fidelity microstructural model to assess both the performance and robustness of the developed calibration workflow as well as the transferability of the optimised parameters across microstructural fidelities.

The following section will first discuss the development and accuracy of the DNN surrogate models (Section 5.1) before evaluating the predictive performance of the calibrated CPFEM models, comparing the stress–strain response, grain reorientation trajectories, and overall texture evolution across the three CP formulations (Section 5.2). The section will then assess the variability of the optimised parameters across independent calibration runs (Section 5.3) and discuss the impact of mesoscale constraints on the calibration results (Section 5.4). Finally, the section will evaluate the transferability of the optimised parameters from the low- to high-fidelity microstructural FE models (Section 5.5). In addition, an assessment of the calibration costs for each CPFEM model can be found in Appendix F.

5.1. Performance of trained DNN surrogate models

Deep neural network (DNN) surrogate models were adaptively trained to replicate the predictions of the low-fidelity CPFEM models, reproducing both the macroscopic stress–strain response (Fig. 2) and the mesoscale reorientation trajectories of eight grains ($g_{1..8}$; Fig. 4). The surrogate models were trained using an initial dataset of 8 simulations for the VH and LH2 formulations, and 16 simulations for the LH6 formulation to accommodate its larger number of material parameters. The surrogate model was iteratively refined using additional low-fidelity CPFEM simulations evaluated at the MOGA-optimised parameter values, until the surrogate model satisfied the accuracy thresholds, $E_\sigma < 1\%$ and $E_\phi < 0.1^\circ$. This adaptive training strategy allowed the surrogate model to become progressively more accurate in regions of the parameter space that are most relevant to the calibration, where the CPFEM responses aligned closest with the experimental observations. However, this

strategy limits the generalisability of the surrogate models beyond the explored regions of the parameter space, preventing its direct application to other experimental datasets without further training.

A new surrogate model was developed from scratch for each independent calibration run to avoid bias from any particular initial dataset or DNN realisation. Fig. 7 presents the evolution of the surrogate model accuracy as more CPFEM simulations are progressively added into the DNN training set across five independent calibration runs of the VH formulation. The surrogate model accuracy is measured using the objective functions, E_σ and E_ϕ (Eqs. (7) and (8)), which in this case quantify the deviation (error) between the surrogate model predictions and the CPFEM predictions across the full range of strain values used in the calibration process. As shown in Fig. 7, the VH formulation achieved the accuracy thresholds after an average of 13 re-training cycles which corresponds to 21 CPFEM evaluations in the training set. The LH2 formulation required an average of 16 cycles (24 evaluations) to reach the same accuracy, while the more complex LH6 formulation required an average of 22 cycles (38 evaluations), reflecting the higher dimensionality of its parameter space. The variability in the number of re-training cycles required to converge reflects the stochastic nature of the employed LHS scheme and the MOGA optimisation process.

Once the DNN surrogate models achieved the predefined accuracy thresholds ($E_\sigma < 1\%$ and $E_\phi < 0.1^\circ$), they consistently reproduce the CPFEM predictions, regardless of how many CPFEM evaluations were used during their training process. A representative example of the surrogate model's accuracy is presented in Fig. 8, showing its predictions (black) with the corresponding low-fidelity CPFEM model's predictions (red), and the experimental data (grey). The close agreement between both models for the stress–strain responses (Fig. 8a) and grain $g_{1..8}$ reorientation trajectories (Fig. 8b) confirms that the adaptively trained DNN surrogate model can provide an accurate representation of the CPFEM model. Reaching this level of surrogate accuracy is a critical prerequisite for the proposed workflow, enabling the MOGA-based CPFEM model calibration to be achieved at a fraction of the computation cost without compromising accuracy.

5.2. Performance of calibrated CPFEM models

The trained surrogate DNN models described in Section 5.1 were next deployed within the MOGA-based calibration workflow (Fig. 6) to

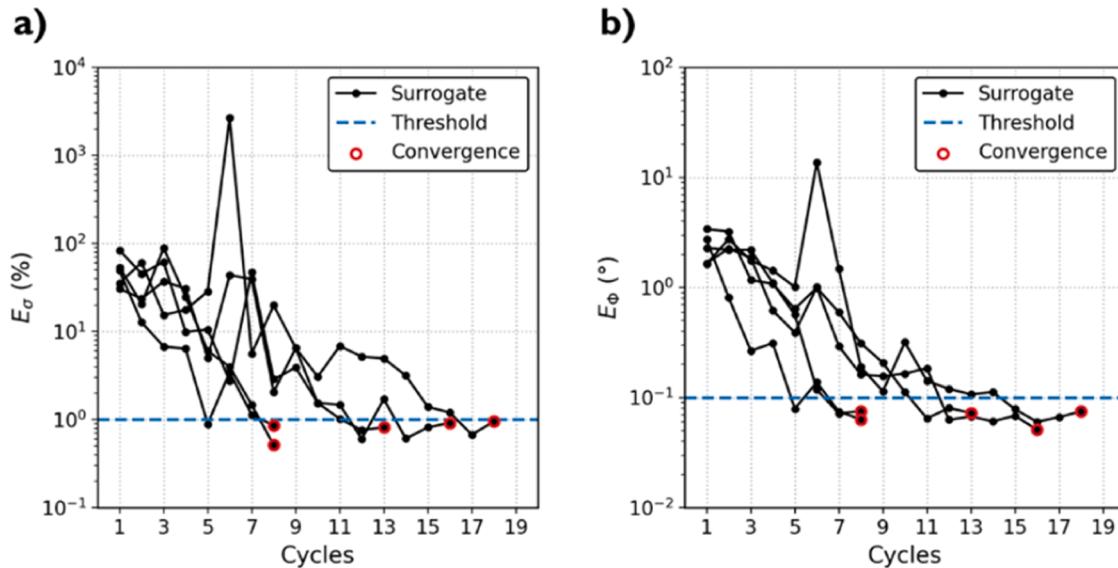


Fig. 7. Discrepancies between the surrogate model approximations and the low-fidelity VH simulations across five calibration runs using an initial dataset size of 8. Plot a) shows the discrepancies in the stress–strain response (%) while plot b) shows the discrepancies in the reorientation trajectories (°). Each plot shows the discrepancies (black line), accuracy threshold (blue dashed line), and the termination cycle (red circle) for each calibration run.

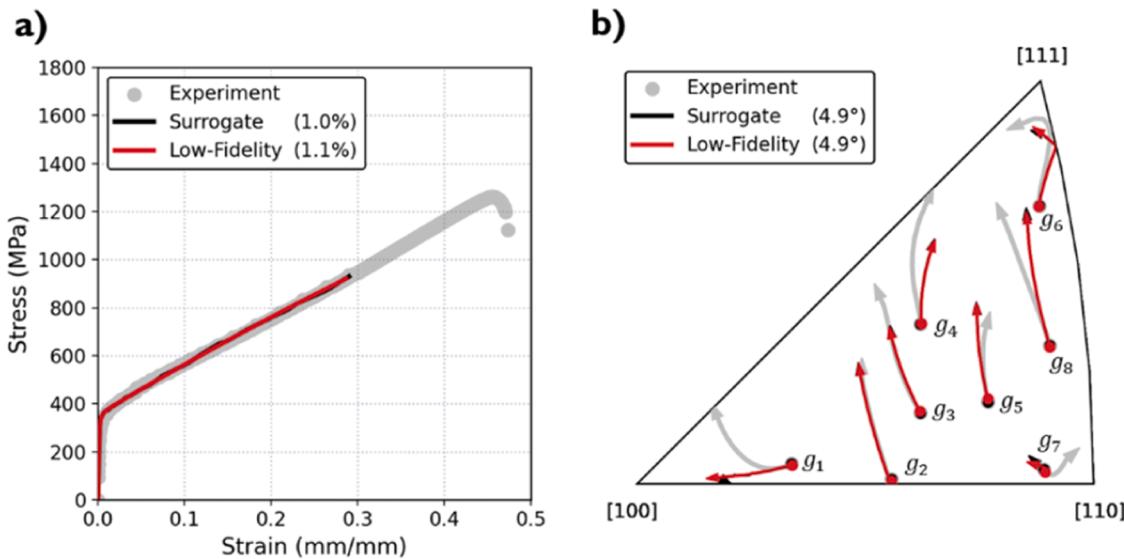


Fig. 8. Comparison of the surrogate model approximations and the low-fidelity VH simulations, with the experimental data in grey, the approximations in black, and the VH responses in red. These plots compare the a) stress–strain responses and b) reorientation trajectories of $g_{1\ldots 8}$, including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)).

replace the computationally expensive CPFEM evaluations. To evaluate the robustness and repeatability of the complete workflow — encompassing both adaptive surrogate model development and MOGA-based optimisation — each of the three CP formulations (VH, LH2, and LH6) was integrated with the low-fidelity FE model and calibrated five times. As outlined above, the multi-objective calibration simultaneously targeted the experimental stress–strain response and the reorientation

trajectories of eight selected grains ($g_{1\ldots 8}$), while validation was performed against the trajectories of eight additional grains ($g_{9\ldots 16}$) and the overall texture evolution. The calibration and validation results for the VH, LH2, and LH6 models are presented in Fig. 9, Fig. 10, and Fig. 11, respectively, comparing the simulated stress–strain responses, reorientation trajectories, and texture evolution with the experimental measurements. In these figures, the best calibration run is identified

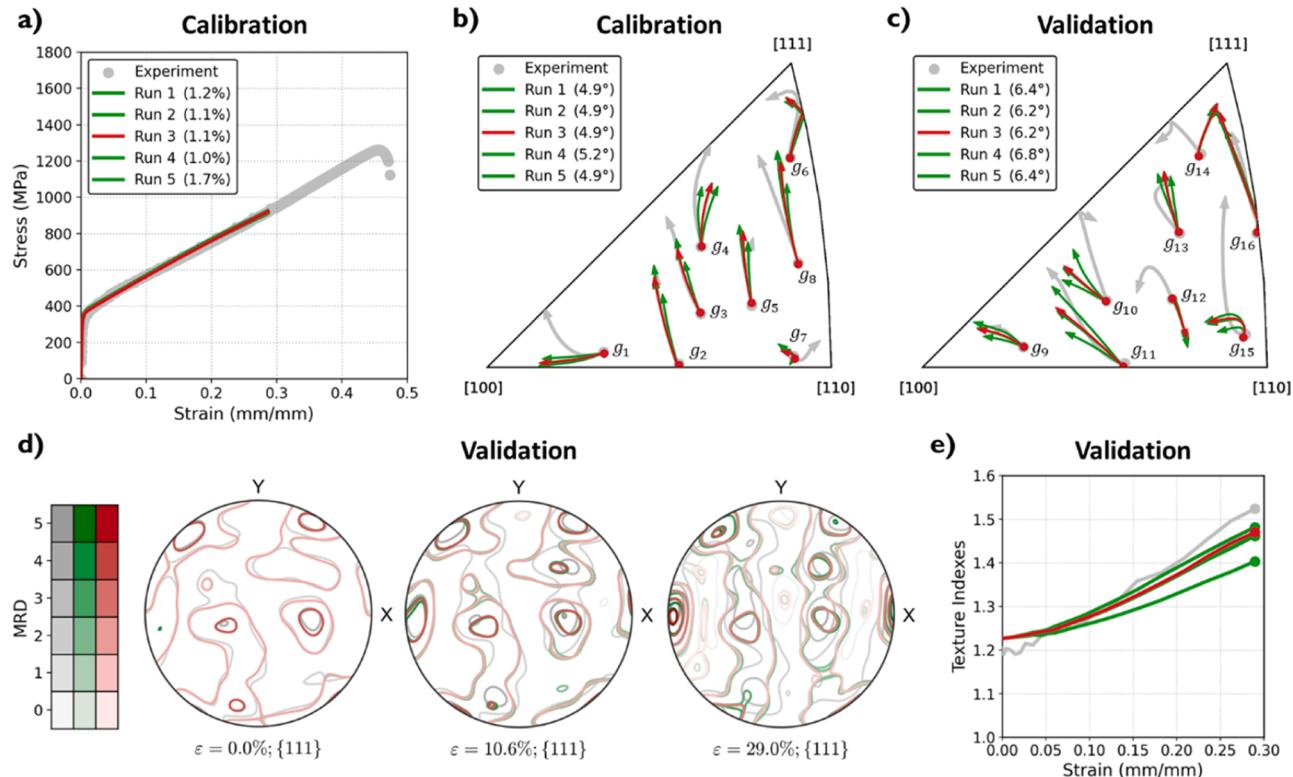


Fig. 9. Calibration and validation results for the VH formulation after five runs, comparing the experimental data in grey with the simulated responses in green (all runs) and red (best run). The plots in the top row compare the a) stress–strain responses (calibration), b) reorientation trajectories of $g_{1\ldots 8}$ (calibration), and c) reorientation trajectories of $g_{9\ldots 16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured $\{111\}$ pole figures at $\epsilon = 0.0\%$, 10.6% , and 29.0% , and e) texture indexes over 0.0% to 29.0% .

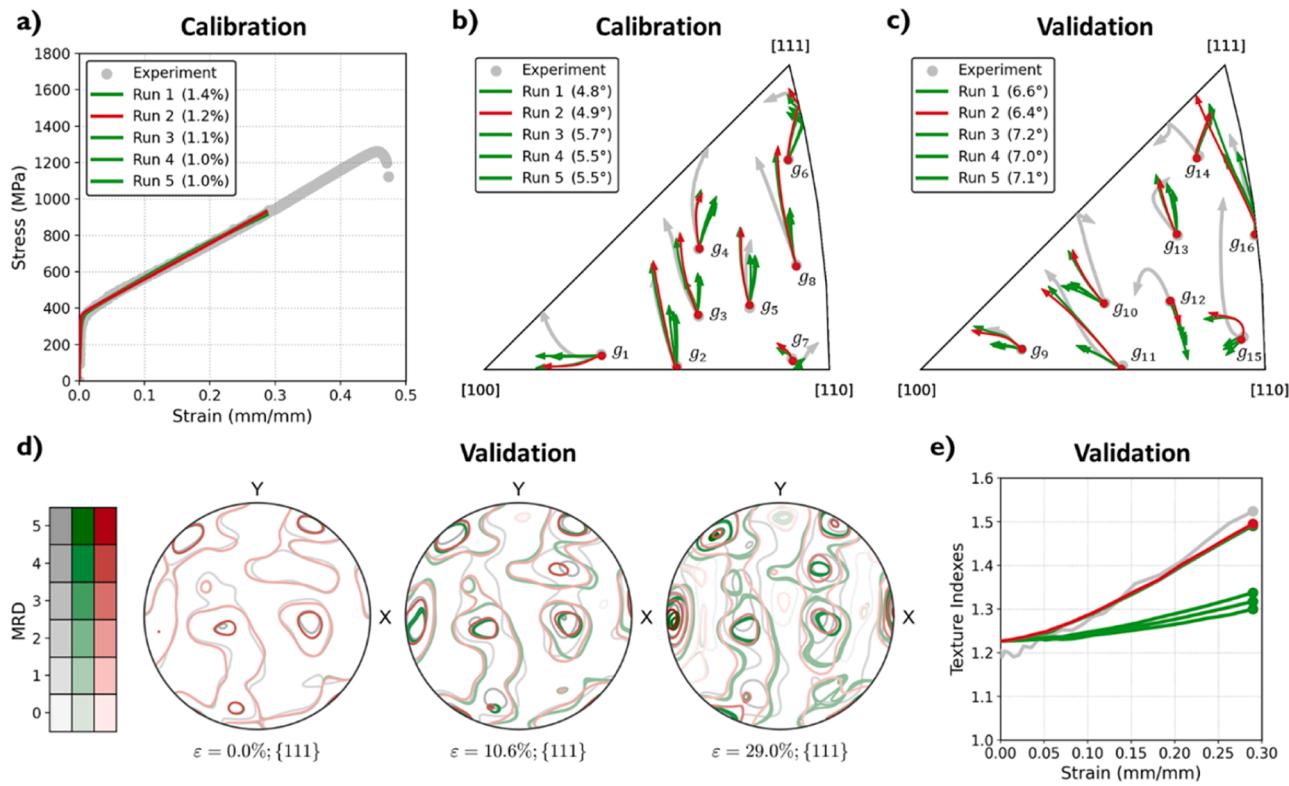


Fig. 10. Calibration and validation results for the LH2 formulation after five runs, comparing the experimental data in grey with the simulated responses in green (all runs) and red (best run). The plots in the top row compare the a) stress–strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured $\{111\}$ pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% , and e) texture indexes over 0.0% to 29.0% .

based on the minimal squared sum of the normalised objective values (Appendix D) and is shown in red, while the other runs are shown in green. The overall texture is characterised using contoured $\{111\}$ pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% , complemented by texture indexes calculated over the same strain range. Note that the influence of the selected calibration grain reorientation trajectories is examined in Appendix G, where the calibration and validation grains are swapped to demonstrate that the workflow is largely insensitive to the choice of calibration grains.

Across all three formulations, the simulated stress–strain responses (Figs. 9a, 10a, and 11a) showed excellent agreement with the experimental data. The VH and LH2 formulations exhibited minimal variability across the five calibration runs, with a mean standard deviation (SD) of 3.76 MPa and 5.05 MPa across the strain range, respectively. This consistency highlights the robustness of the calibration workflow and demonstrates that even the simplest VH formulation can reliably reproduce the macroscopic mechanical response of Alloy 617 under uniaxial loading. In contrast, the LH6 formulation exhibited noticeably greater run-to-run variability with a SD of 10.6 MPa , suggesting less stable convergence of the optimisation process due to the formulation's larger parameter space.

The reorientation trajectories of the eight calibration grains ($g_{1..8}$; Figs. 9b, 10b, and 11b) were reproduced with reasonable accuracy for all three CPFEM models. Grains exhibiting complex, non-monotonic rotation paths in orientation space showed poorer agreement than those following simpler, monotonic trajectories. This behaviour reflects the local nature of grain interactions as well as the limitations associated with constructing a 3D volumetric mesh from a single 2D EBSD orientation map. In this 2.5D representation (Section 2.4), the through-thickness microstructural variations — such as grain morphology, local neighbourhood topology, or orientation gradients — are not captured, which can influence the predicted slip activity and rotation

behaviour of individual grains. These simplifications likely contributed to the discrepancies observed for certain grain trajectories. Introducing stochastic orientation perturbations or randomised microstructural layers along the specimen thickness could, in future, help quantify the sensitivity of the calibration to such assumptions and provide a measure of robustness. Notably, the run-to-run variability in the simulated $g_{1..8}$ trajectories was most pronounced for the LH2 and LH6 formulations, with standard deviations of 0.908° and 0.859° , respectively, compared with 0.256° for the VH formulation.

The reorientation trajectories of the eight validation grains ($g_{9..16}$; Figs. 9c, 10c, and 11c) exhibited greater run-to-run variability to the calibration grains, with SDs of 0.53° , 1.68° , and 1.15° for the VH, LH2, and LH6 formulations, respectively. The simulated trajectories also showed a clear decline in accuracy, as indicated by their higher E_ϕ values. The reduced consistency and predictive accuracy of the $g_{9..16}$ trajectories is expected due to their exclusion from the calibration dataset, further exacerbated by complex rotation paths that the FE model was unable to reliably reproduce. In many cases, these grains (i.e., g_{11} , g_{12} , g_{14} , g_{15} , and g_{16}) were located near mesh boundaries or affected by segmentation artefacts — such as merging or splitting during grain tracking — which likely degraded their correspondence between the experimental microstructure and the low-fidelity model (see Appendix B for details of the grain-tracking algorithm applied to the *in situ* EBSD maps). While the LH2 and LH6 formulations offered greater flexibility in simulating non-monotonic reorientation behaviour, they also showed higher sensitivity to noise in the calibration data. In contrast, the VH formulation delivered more consistent predictions across runs, suggesting a more favourable balance between model complexity and the level of experimental constraint provided under uniaxial loading.

Despite the variability in the mesoscale reorientation trajectory predictions, all three CPFEM models showed good performance in pre-

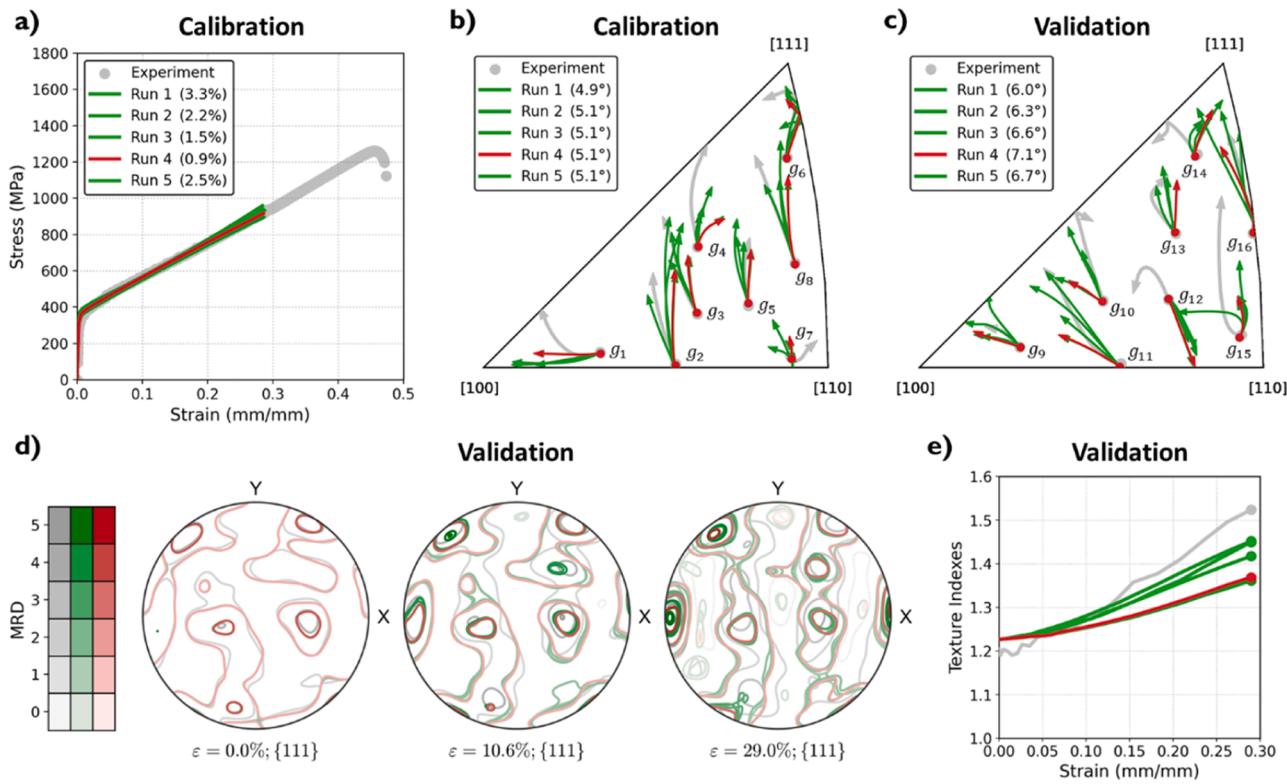


Fig. 11. Calibration and validation results for the LH6 formulation after five runs, comparing the experimental data in grey with the simulated responses in green (all runs) and red (best run). The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured {111} pole figures at $\varepsilon = 0.0\%$, 10.6%, and 29.0%, and e) texture indexes over 0.0% to 29.0%.

dicting the overall texture evolution (Fig. 9d, e, 10d, e, and 11d, e). The simulated {111} pole figures and texture indexes were generally in good agreement with the experimental data, particularly at lower and intermediate strain levels. At higher strains, the VH and LH2 formulations continued to capture the dominant texture components well, while the LH6 formulation displayed some misalignment in both intensity and orientation. Notably, for the LH6 formulation, some runs produced texture indexes well below the experimental value despite similar accuracy in the calibration grain trajectories. These results highlight a limitation of relying on a restricted set of grain reorientation trajectories in the calibration process. In other words, accurately reproducing the mesoscale grain trajectories may not necessarily guarantee accurate texture predictions.

The simulated results from the best calibration run of each CPFEM formulation are presented in Fig. 12, comparing the VH (cyan), LH2 (orange), and LH6 (purple) responses — these are the best of five calibration runs discussed above and presented in Figs. 9, 10, and 11. It becomes clear from Fig. 12 that all three models reproduced the stress-strain response with excellent accuracy (Fig. 12a), confirming that the calibration workflow can reliably capture the macroscopic mechanical behaviour of Alloy 617 under uniaxial loading. The grain reorientation trajectories of $g_{1..16}$ (Fig. 12b, c) were also predicted with broadly similar accuracy, although the LH6 responses consistently exhibited larger deviations and greater run-to-run scatter than the VH and LH2 responses, despite its greater flexibility and larger parameter space. A comparable trend is observed for the overall texture evolution (Fig. 12d, e), where the VH and LH2 formulations more accurately tracked the intensity and orientation of the dominant texture components, while the LH6 formulation displayed larger deviations in the pole figure contours and final texture index, along with greater run-to-run variability.

Overall, the present results demonstrate that all three tested CP formulations can capture the key aspects of Alloy 617's deformation behaviour under uniaxial tensile loading, including the macroscopic stress-strain response, individual grain rotation behaviour, and texture evolution as a function of imparted plastic strain. Of the three, the VH formulation delivered the most consistent and robust performance, providing the best balance between accuracy and reliability. The LH2 formulation achieved comparable accuracy in many cases and offered greater flexibility for modelling complex grain rotations, but at the cost of increased sensitivity to noise and reduced parameter identifiability. In contrast, the LH6 formulation — despite its broader representational capacity — did not outperform the simpler CP formulations and showed substantially greater variability between calibration runs in all evaluated metrics. Additionally, the VH formulation was substantially cheaper to calibrate, requiring an average of 4.83 CPU-hours on 192 cores (dual AMD EPYC 96-core processors), compared with 35.04 h for LH2 and 55.48 h for LH6 (see Appendix F). These findings show that VH is the most practical formulation for this application, underscoring the need to match model complexity to the available experimental data.

5.3. Robustness of MOGA-based calibration workflow

As discussed in Sections 5.1 and 5.2, the developed calibration workflow — combining adaptively trained DNN surrogate models with a MOGA — enables efficient calibration of computationally expensive CPFEM models. To assess its robustness, five independent calibration runs were performed for each CP formulation. These runs revealed some variability in the simulated responses, particularly in the mesoscale reorientation trajectories of individual grains. Such variability reflects the inherent sensitivity of CPFEM simulations to the chosen material parameters and raises the question of how consistently these parameters

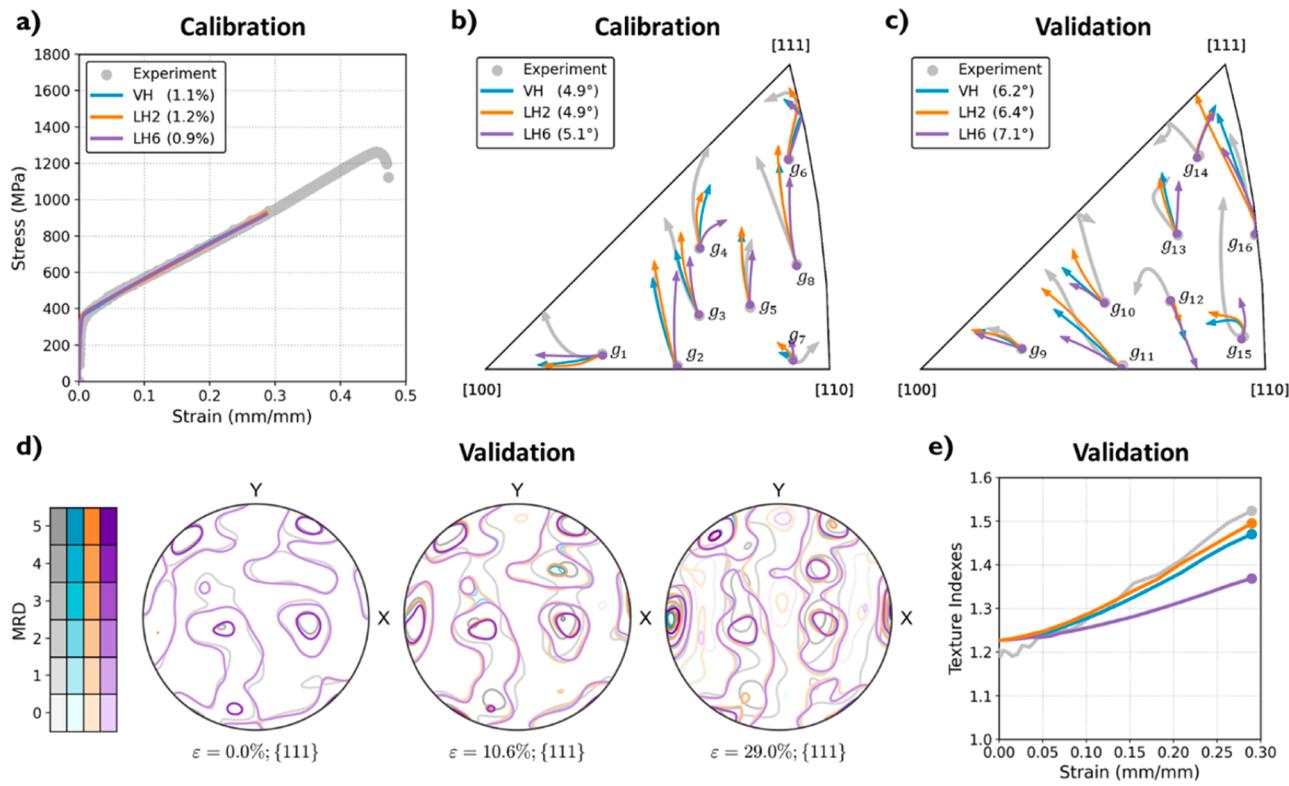


Fig. 12. Comparison of the low-fidelity responses of the three CPFEM models, with the experimental data in grey, the VH responses in cyan, the LH2 responses in orange, and the LH6 responses in purple. The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_o and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured {111} pole figures at $\varepsilon = 0.0\%$, 10.6%, and 29.0%, and e) texture indexes over 0.0% to 29.0%.

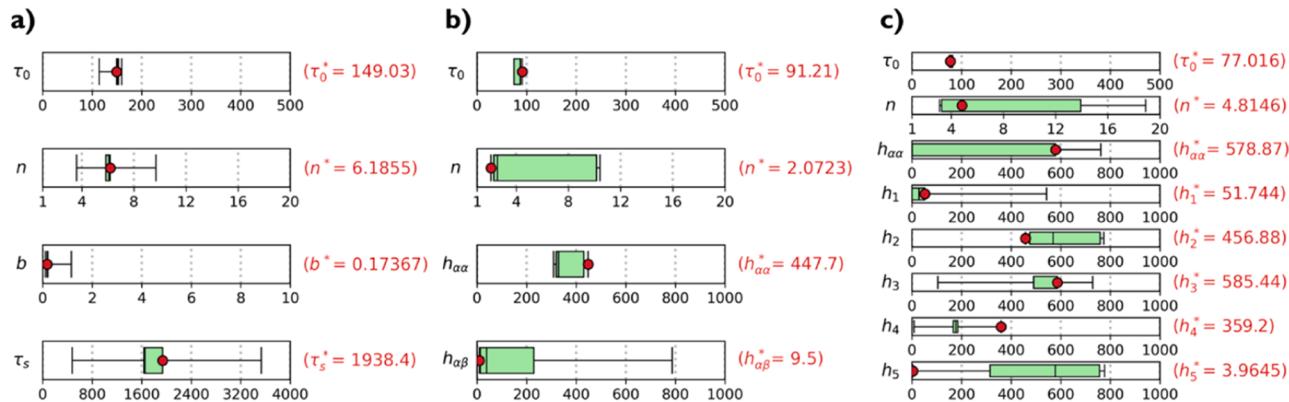


Fig. 13. Boxplots of the optimised parameters for the a) VH formulation, b) LH2 formulation, and c) LH6 formulation across five calibration runs, showing the mean (black horizontal line), standard deviation (box), range (whiskers), and bounds (limits). The knee-point values are overlaid as red circles, with their exact values annotated on the right.

can be identified. To explore this, Fig. 13 presents the distributions of the optimised material parameters for the VH, LH2, and LH6 formulations across their five calibration runs. These distributions include $\{\tau_0, n, b, \tau_s\}$ for the VH formulation, $\{\tau_0, n, h_{\alpha\alpha}, h_{\alpha\beta}\}$ for the LH2 formulation, and $\{\tau_0, n, h_{\alpha\alpha}, h_1, h_2, h_3, h_4, h_5\}$ for the LH6 formulation.

It is clear from Fig. 13 that the variability of the identified material parameters differs substantially between CPFEM models, suggesting that the robustness of the calibration workflow varies with the employed CP formulation. For the VH formulation (Fig. 13a), τ_0 (initial slip system strength) and b (hardening rate) exhibit narrow distributions across calibration runs, indicating that these parameters are well constrained by the available experimental datasets. In contrast, n (strain-rate

sensitivity) and τ_s (saturation strength) display broader spreads, reflecting the absence of calibration data at multiple strain rates and into the late-stage saturation regime (Stage III), limiting the constraint of these two material parameters. The inclusion of grain reorientation trajectories in the objective function appears to reinforce the stability of the VH formulation, as τ_0 and b strongly influence the primary slip system activation and are already tightly constrained by the stress-strain curve. The LH2 formulation (Fig. 13b) retains relatively low variability in τ_0 , but shows significantly wider spreads for n , $h_{\alpha\alpha}$ (self-hardening strength), and $h_{\alpha\beta}$ (latent hardening strength). Under a single monotonic tensile path, the additional reorientation data only moderately reduces the parameter non-uniqueness, as many slip system interactions remain

weakly activated and therefore do not leave a distinct signature in the orientation trajectories. The LH6 formulation (Fig. 13c) shows a similar pattern for τ_0 , which remains relatively consistent between runs, but displays pronounced variability in n , h_{aa} (self-hardening strength), and $h_{1..5}$ (latent hardening strengths). The variability of these parameters reflects their poor constraint due to the limited activation of secondary slip systems under uniaxial tensile loading [75]. Without more diverse deformation paths — such as multiaxial or non-monotonic loading — to stimulate these interactions, the optimiser struggles to distinguish among multiple mathematically viable but physically inconsistent solutions [75]. This further highlights a need to match the model complexity to the type of available experimental datasets.

The optimised parameter values from the best calibration runs (Appendix D) of all three formulations are presented in Fig. 13 as red symbols. It should be noted upfront that the strain-rate sensitivity parameter, n , should be interpreted with caution across all formulations. Its variation between the VH, LH2, and LH6 formulations primarily reflects the lack of experimental data at multiple strain rates to constrain it — if such data were available, n , would be expected to converge to a consistent value independent of the chosen CP formulation. With this limitation in mind, the remaining parameters reveal distinct trends across the three models. The optimised material parameters for the VH formulation ($\tau_0^* = 149.03\text{MPa}$, $n^* = 6.1855$, $b^* = 0.1737$, and $\tau_s^* = 1938.4\text{MPa}$) indicate a high initial slip system strength and moderate hardening rate. In the LH2 formulation, the optimised parameters ($\tau_0^* = 91.21\text{MPa}$, $n^* = 2.0723$, $h_{aa}^* = 447.7\text{MPa}$, and $h_{ab}^* = 9.5\text{MPa}$) indicate a decrease in initial strength and suggest that most of the additional hardening is concentrated in self-hardening rather than latent interactions. Finally, the LH6 formulation exhibits an even lower initial strength ($\tau_0^* = 77.016\text{MPa}$, $n^* = 4.8146$, $h_{aa}^* = 578.87\text{MPa}$) and highly scattered latent hardening values ($h_1^* = 51.744\text{MPa}$, $h_2^* = 456.88\text{MPa}$, $h_3^* = 585.44\text{MPa}$, $h_4^* = 359.2\text{MPa}$, and $h_5^* = 3.9645\text{MPa}$), underscoring the poor constraint of these parameters under the present uniaxial loading condition.

The contrast between the VH and LH2 formulations further highlights how parameter identifiability depends on model structure. Although both formulations involve fitting only four material parameters, the inclusion of latent hardening terms in the LH2 formulation changes the allocation of hardening between τ_0 and the interaction coefficients. This reallocation is a mathematical consequence of parameter coupling, in which the absence of data that explicitly constrains latent hardening allows the optimiser to reduce τ_0 and adjust h_{aa} and h_{ab} to reproduce the same macroscopic stress-strain curve. As a result, τ_0 becomes model-dependent rather than a unique estimate of the critical resolved shear stress (CRSS) for the active slip system. Determining the true value of τ_0 would thus require independent experimental constraints, such as multiple strain-rate curves to decouple τ_0 from n , or non-proportional and multiaxial loading to activate and constrain latent hardening. This issue becomes even more pronounced in the LH6 formulation, where the expanded parameter space further amplifies coupling effects. Without sufficiently diverse experimental loading conditions to activate secondary slip systems, the optimiser distributes hardening across many poorly constrained interaction terms, yielding parameter sets that lack physical uniqueness.

This non-uniqueness is a well-recognised characteristic of inverse CPFEM calibration problems, where several parameter sets can yield comparable macroscopic and mesoscale responses. In the present study, the inclusion of grain reorientation trajectories helps to filter out non-physical solutions but cannot fully remove this degeneracy because the available experimental data provide only limited independent constraints. The issue is analogous to other under-constrained inverse problems, such as nano-indentation, where distinct constitutive parameter combinations can reproduce similar load-displacement curves. As demonstrated in indentation studies, introducing additional physically measurable quantities — such as the plastic-zone radius [78]

— can improve parameter identifiability. By analogy, CPFEM calibrations could incorporate physics-informed observations (e.g., local lattice strains, slip activity maps, or orientation-gradient fields) to further constrain the feasible parameter space and enhance uniqueness [16,79].

Taken together, these results show that, given the single monotonic tensile curve at one strain rate available for calibration, the VH formulation provides the most reliable and interpretable parameter set. The parameters governing the initial slip system strength and hardening rate are well constrained, whereas the strain-rate sensitivity parameter remains weakly constrained across all formulations. This limitation reflects the lack of experimental data at multiple strain rates rather than any deficiency of the CP formulations themselves. The LH2 formulation can reproduce the same stress-strain response, but the additional latent hardening terms are only loosely identifiable under simple uniaxial loading. The LH6 formulation is clearly over-parameterised for the available dataset and produces parameter sets that lack physical uniqueness, consistent with the large spreads and run-to-run variability observed in stress-strain, grain reorientation, and texture predictions (Section 5.2). Based on this variability analysis alone, the inclusion of grain reorientation trajectories provides only a moderate reduction in parameter non-uniqueness. Overall, the VH formulation yields the most robust and interpretable parameters, the LH2 formulation introduces moderate ambiguity, and the LH6 formulation is over-parameterised under the present calibration conditions.

5.4. Impact of mesoscale constraints on model calibration

As shown in Section 5.3, the inclusion of mesoscale grain reorientation trajectories in the multi-objective calibration framework has only a limited effect on reducing the statistical variability of material parameters under the present loading conditions. However, it plays an important role in filtering out physically unrealistic solutions that would otherwise pass undetected in a single-objective framework. To clarify this distinction, the VH formulation was calibrated without mesoscale constraints, using only the macroscopic stress-strain curve while excluding the reorientation trajectories from the calibration process. This single-objective approach mirrors the most common practice in CPFEM parameter identification, where model fitting is performed exclusively against the stress-strain response without incorporating additional experimental datasets capable of constraining the parameter space. The predictions obtained from the VH formulation calibrated using single-objective approach are presented in Fig. 14, while the corresponding distributions of the optimised parameters are presented in Fig. 15.

The results in Fig. 14 can be directly compared with those in Fig. 9, which shows the VH responses from the developed multi-objective calibration workflow that also incorporated the grain reorientation trajectories in its objective function. Several observations emerge from this comparison. Firstly, the stress-strain response is predicted with excellent accuracy in both single- (Fig. 14a) and multi-objective approaches (Fig. 9a). This demonstrates the strong capacity of the VH formulation in matching the global mechanical response even when unconstrained by mesoscale reorientation behaviour. Secondly, the predictions for the $g_{1..16}$ trajectories remain reasonably accurate despite their exclusion from the calibration, with the average geodesic distance error (E_ϕ) increasing only slightly from 5.68° (Fig. 9b, c) to 6.24° (Fig. 14b, c). Thirdly, and most critically, the predictions for texture evolution show greater discrepancies in the single-objective approach. This can be seen in the best single-objective calibration run, which yields a final texture index of approximately 1.30 (Fig. 14). In contrast, the best multi-objective run yields a final index of 1.47 (Fig. 9e), which is much closer to the experimental value of 1.52.

At first glance, these findings may appear to contradict the results of Section 5.3, which indicates that the inclusion of the grain reorientation trajectories in the calibration process has only a moderate effect on constraining the statistical spread of parameters. In reality, the two sets

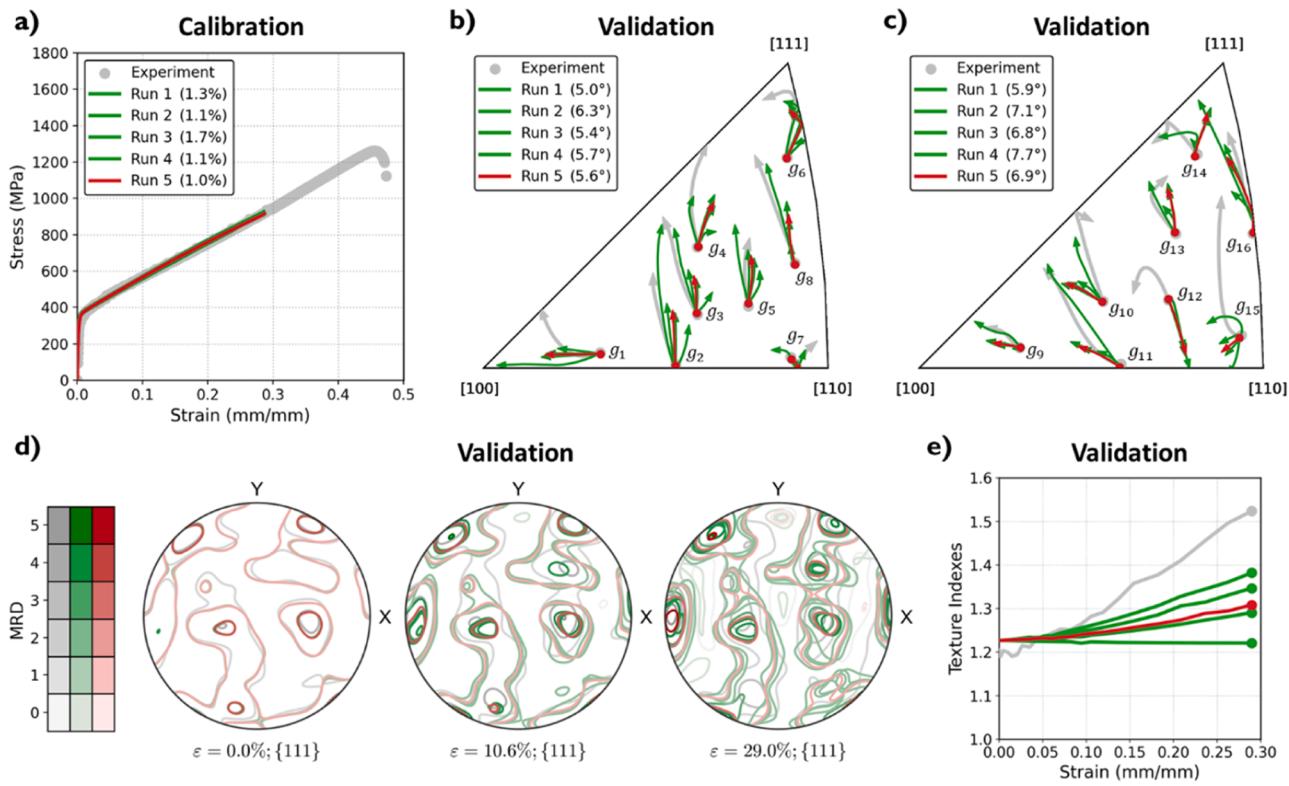


Fig. 14. Supplementary calibration and validation results for the VH formulation after five single-objective runs, comparing the experimental data in grey with the simulated responses in green (all runs) and red (best run). The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1,8}$ (validation), and c) reorientation trajectories of $g_{9,16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured $\{111\}$ pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% .

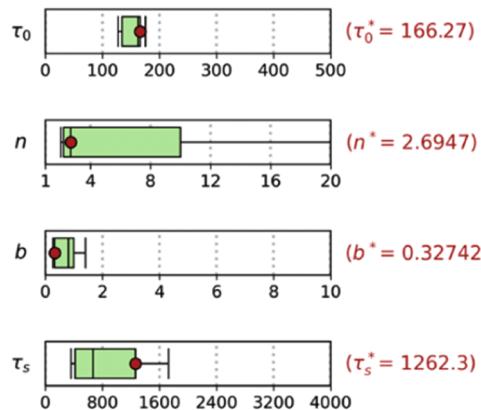


Fig. 15. Supplementary boxplots of the optimised parameters for the VH formulation across five single-objective calibration runs, showing the mean (black horizontal line), standard deviation (box), range (whiskers), and bounds (limits). The knee-point values are overlaid as red circles, with their exact values annotated on the right.

of results are consistent once the distinction between parameter identifiability and physical plausibility is recognised. Under the present loading conditions — single-path monotonic tension at a single strain rate — grain reorientation trajectories and the stress-strain curve are not strongly independent constraints for material parameters. Consequently, their inclusion as an objective function does not sharply reduce the run-to-run variability of these parameters, as illustrated by the box plots in Figs. 13 and 15. Instead, these mesoscale constraints act as a decisive physical filter. Many parameter combinations can match the macroscopic stress-strain response, but only a subset produce

microstructural evolution consistent with the measured reorientation behaviour and the resulting texture. Without these mesoscale constraints, the optimisation can converge to mathematically valid yet physically inconsistent material parameter sets, leading to degraded predictions of texture evolution despite an equally accurate macroscopic fit.

Thus, in the present study, the value of including grain trajectories lies less in tightening the statistical identifiability of weakly constrained parameters and more in ensuring that the selected material parameter set is physically consistent across scales. This conclusion is particularly relevant because single-objective, stress-strain-only calibration remains the dominant approach in CPFEM modelling. The present results therefore provide strong evidence that developed MOGA-based model calibration incorporating mesoscale constraints should be adopted more widely to obtain parameter sets that are both mathematically optimal and physically meaningful.

5.5. Transferability across microstructural fidelities

All calibration runs in previous sections were performed using a 2.5D low-fidelity FE microstructural representation of the Alloy 617 specimen gauge, comprising 132 grains (from the experimentally identified 571 grains) and 17,109 hexahedral elements. While this model is referred to as ‘low fidelity’, it still captures nearly one quarter of the measured grains, preserving a representative subset of the microstructure. To assess the transferability of the found material parameters to a more realistic microstructural representation, the knee-point parameters from the best calibration run of each CP formulation (VH, LH2, and LH6; Figs. 9, 10, and 11) were applied without modification to a 2.5D high-fidelity FE model of the same specimen, comprising 481 grains and 574,560 hexahedral elements. This high-fidelity FE model of the

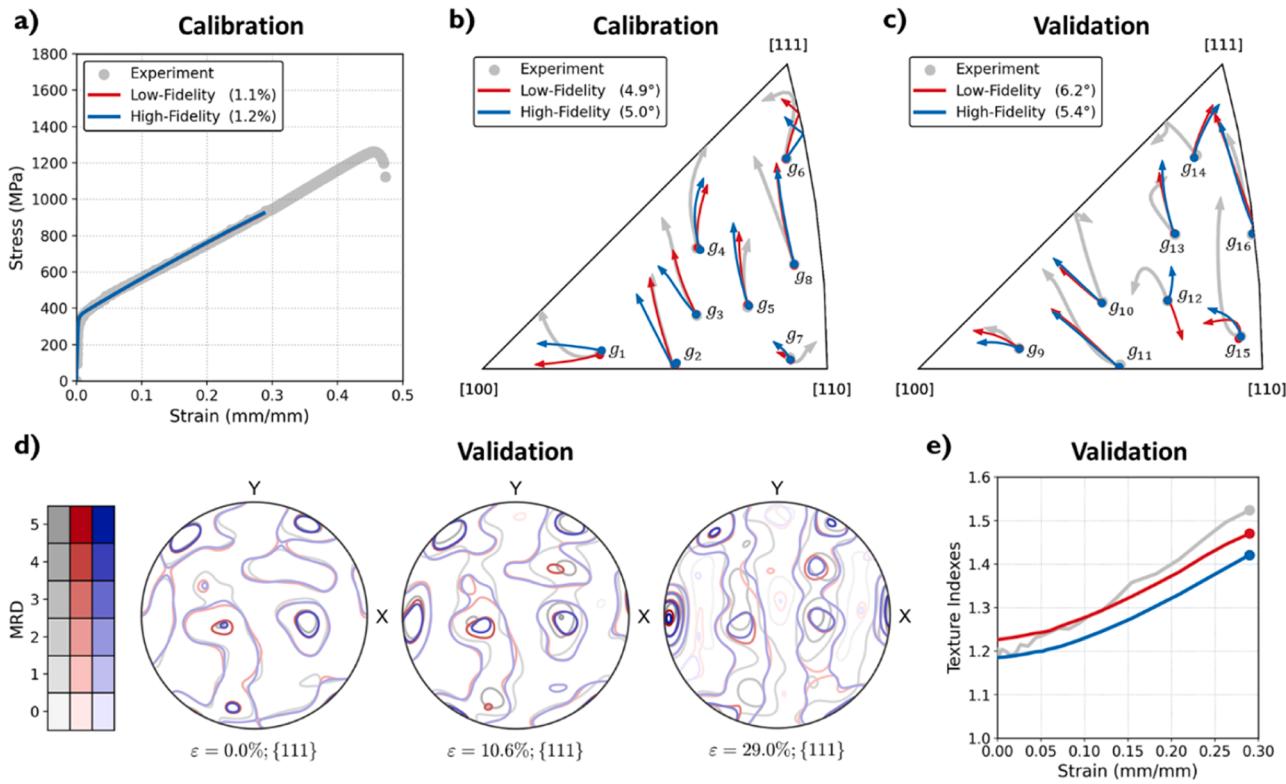


Fig. 16. Comparison of the low- and high-fidelity responses from the VH formulation, with the experimental data in grey, the low-fidelity responses in red, and the high-fidelity responses in blue. The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured {111} pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% , and e) texture indexes over 0.0% to 29.0% .

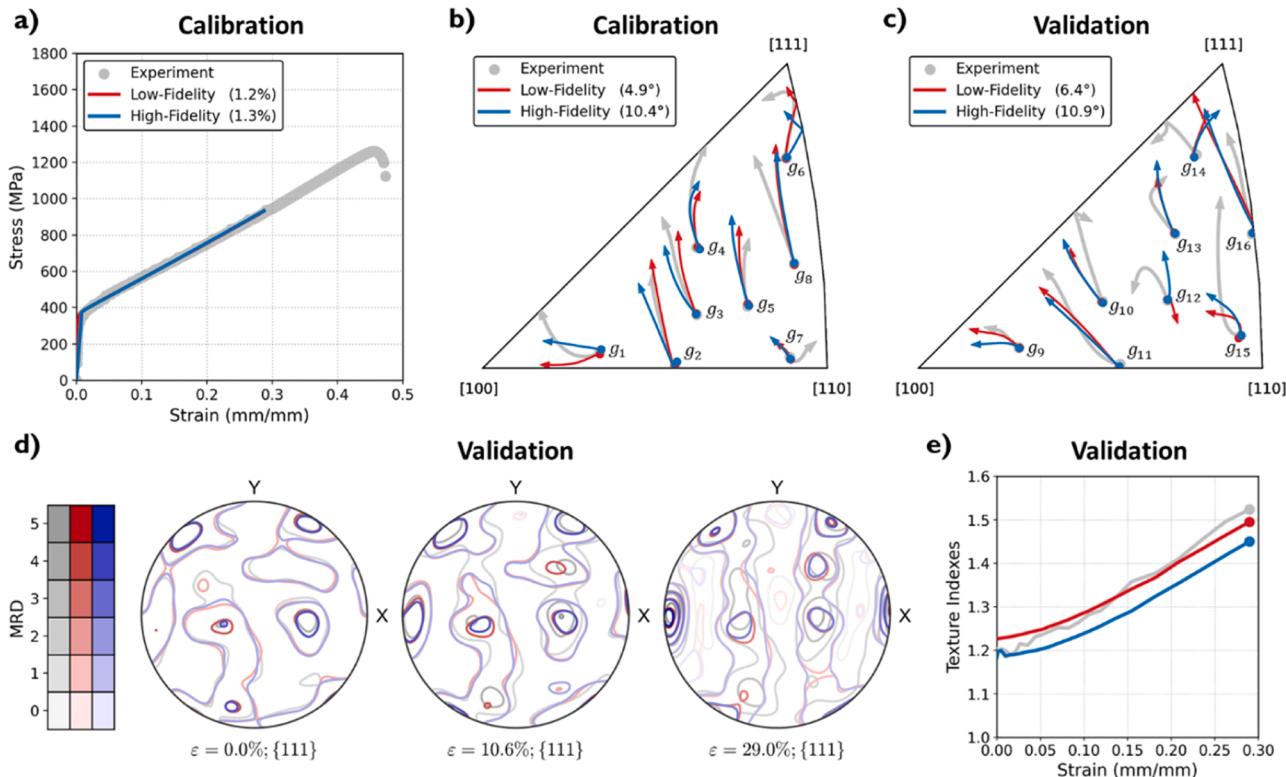


Fig. 17. Comparison of the low- and high-fidelity responses from the LH2 formulation, with the experimental data in grey, the low-fidelity responses in red, and the high-fidelity responses in blue. The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured {111} pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% , and e) texture indexes over 0.0% to 29.0% .

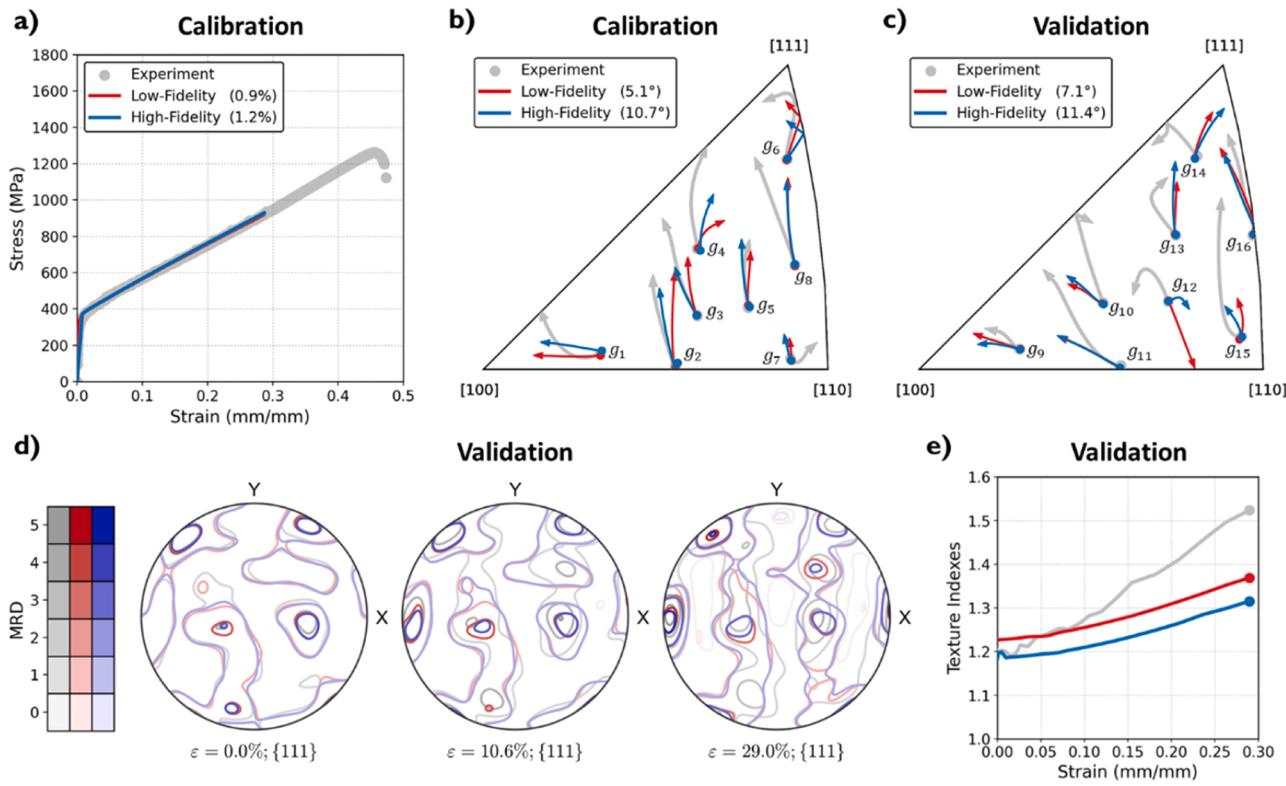


Fig. 18. Comparison of the low- and high-fidelity responses from the LH6 formulation, with the experimental data in grey, the low-fidelity responses in red, and the high-fidelity responses in blue. The plots in the top row compare the a) stress–strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (calibration), and c) reorientation trajectories of $g_{9..16}$ (validation), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured $\{111\}$ pole figures at $\varepsilon = 0.0\%$, 10.6% , and 29.0% .

microstructure preserves a much larger fraction of the experimentally measured grain structure and neighbourhood topology, thereby modifying the local interaction fields experienced by individual grains. The plots in Figs. 16, 17, and 18 compare the low- and high-fidelity predictions for the VH, LH2, and LH6 formulations, respectively, showing the stress–strain curves, reorientation trajectories of $g_{1..16}$, and overall texture evolution.

Across all three formulations, the stress–strain curves (Figs. 16a, 17a, and 18a) remain in excellent agreement with the experimental data and are virtually indistinguishable between low- and high-fidelity CPFEM models. This demonstrates that the macroscopic mechanical response is largely insensitive to microstructural resolution, at least within the range examined here, and that the calibrated parameters capture the bulk deformation behaviour in a transferable manner. From an engineering perspective, this suggests that moderate-fidelity microstructures, such as the 132-grain model used for calibration, are sufficient for identifying parameters that govern global stress–strain behaviour, offering substantial computational savings without compromising accuracy.

For the VH formulation (Fig. 16b, c), the average geodesic distance error for $g_{1..16}$ remains essentially unchanged when moving from the low- to high-fidelity model, with certain grains (e.g., g_{12}) even exhibiting improvements in accuracy. In contrast, the LH2 and LH6 formulations show reduced accuracy in the reorientation trajectory predictions, with average geodesic distance errors (E_ϕ) increasing from $\sim 6^\circ$ in the low-fidelity calibration to $\sim 11^\circ$ in the high-fidelity predictions (Figs. 17b, c and 18b, c). This reduction in accuracy reflects the greater sensitivity of the latent-hardening-based formulations to local

neighbourhood changes, which are significantly altered by the more complete grain topology in the high-fidelity representation. These results build on the findings of Section 5.2, confirming that the VH formulation delivers the most stable and repeatable (robust) predictions when it comes to grain reorientation trajectories, both within the low-fidelity calibration model and when transferred to a higher-fidelity representation.

For the overall texture evolution, both low- and high-fidelity models reproduce the main features of the measured $\{111\}$ pole figures (Figs. 16d, e, 17d, e, and 18d, e). Consistent with the mesoscale predictions of grain reorientation trajectories discussed above, the VH formulation provides the most stable predictions, whereas the LH2 and LH6 formulations show greater sensitivity to the higher-resolution microstructural topology. Across all three formulations, however, the high-fidelity models yield texture indices that are about 0.45 units lower than their low-fidelity counterparts, corresponding to an under-prediction of texture sharpness. This reduction can be attributed to the greater microstructural detail in the high-fidelity FE models, which disperses dominant orientations that may be artificially concentrated in the lower-fidelity representation.

Overall, these results show that transferring calibrated parameters from low- to high-fidelity microstructural models is entirely feasible for capturing the global stress–strain response and predicting the texture evolution. While the VH formulation yields reliable grain rotation predictions, the LH2 and LH6 formulations exhibit degraded accuracy due to their increased sensitivity to local interaction effects at higher microstructural resolutions. While direct calibration on a high-fidelity model would likely improve these mesoscale predictions, the

associated computational cost is substantial. As such, for applications focused on reproducing the macroscopic response or achieving stable mesoscale predictions, it is effective to calibrate FE models that use moderately reduced fidelities and VH-type formulations.

6. Conclusion

This study presented a multi-objective surrogate-assisted calibration workflow for crystal plasticity finite element method (CPFEM) models, combining a multi-objective genetic algorithm (MOGA) with an adaptively trained deep neural network (DNN) surrogate model. The calibration workflow was demonstrated on three crystal plasticity (CP) formulations — a Voce hardening (VH) model, a two-coefficient latent hardening (LH2) model, and a six-coefficient latent hardening (LH6) model — coupled with a 2.5D low-fidelity microstructural finite element (FE) model of the specimen gauge. The CPFEM models were calibrated in five independent runs to confirm the repeatability of the workflow and enable direct analysis of variability in the optimised material parameters. Across all formulations, the macroscopic stress-strain response was reproduced reliably. The VH formulation delivered the most consistent results across calibration runs, with stable predictions of grain reorientation trajectories and texture evolution. The LH2 formulation matched the macroscopic accuracy of VH but showed higher sensitivity to noise in the calibration data and to changes in the fidelity of microstructural representation. The LH6 formulation exhibited substantial run-to-run variability across all assessed metrics, indicating that its larger parameter space is under-constrained under the present single-rate, monotonic loading conditions.

Analysis of the calibrated material parameters explains these trends. In the VH formulation, the initial slip system strength and hardening rate were tightly constrained, whereas the strain-rate sensitivity remained weakly constrained across all formulations due to the absence of multi-rate data. Adding latent hardening terms in the LH2 and LH6 formulations increased modelling flexibility but reduced parameter identifiability. Different parameter combinations could reproduce the same macroscopic response, reflecting parameter compensation effects and weak constraints on latent interactions under the given loading path. Including grain reorientation trajectories in the multi-objective calibration only moderately reduced statistical variability but acted as an important mesoscale filter, removing physically inconsistent solutions that could otherwise degrade texture evolution predictions. The transferability of the calibrated material parameters was also examined by applying the optimised parameter sets from the low-fidelity calibrations to a 2.5D high-fidelity microstructural FE model of the same specimen. Across all formulations, the prediction of the macroscopic stress-strain response was clearly preserved. At the mesoscale, the VH formulation retained good prediction of grain rotations and overall texture evolution, whereas the CPFEM predictions implementing the LH2 and LH6 formulations showed larger discrepancies, highlighting their sensitivity to local neighbourhood changes introduced by the higher fidelity of the FE microstructural model.

Overall, under the present single-rate, monotonic uniaxial loading, the VH formulation provides the most reliable and interpretable calibration outcome and is recommended for parameter reporting and predictive use in similar conditions. The LH2 and LH6 formulations could be potentially more useful when additional constraints exist, such as multiple strain rates, multiaxial or non-proportional loading, or physics-informed parameter bounds to reduce non-uniqueness and parameter compensation. Regardless of formulation, including grain

reorientation trajectories in the objective is recommended. Even when they do not benefit the model stress-strain accuracy, they improve physical consistency and enhance predictions of texture evolution. Calibration on microstructures with moderately reduced fidelities offers an efficient route to reliable macroscopic predictions, but grain-scale predictions should be validated and, if necessary, refined on higher-fidelity representations of microstructure.

Finally, the adopted 2.5D microstructural representation reduces through-thickness complexity, enabling direct integration of *in situ* EBSD data while keeping the computational cost tractable. The agreement achieved under uniaxial loading indicates that the key deformation mechanisms were captured. It is recognised that the parameters identified in this framework may not be fully transferable to a fully three-dimensional microstructure, where additional morphological constraints would influence the local deformation behaviour. Extending the same multi-objective surrogate-assisted calibration workflow to fully three-dimensional microstructures, richer loading paths, and additional physics such as creep and damage should further improve parameter identifiability and broaden its applicability to structural materials in demanding environments.

Data availability

The data that supports the findings of this research are available upon reasonable request. We are committed to sharing our data to enable further research and replication of our findings. Please contact the corresponding author to request the data, and we will provide it as soon as possible.

CRediT authorship contribution statement

J. Choi: Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **O. Muránsky:** Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **M.C. Messner:** Writing – review & editing, Supervision, Software, Resources, Methodology, Investigation. **T. Wei:** Methodology. **T. Hu:** Writing – review & editing, Software. **J.J. Kružic:** Writing – review & editing, Supervision, Resources. **M.D. McMurtrey:** Writing – review & editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors sincerely thank Zhiyang Wang (ANSTO) for his outstanding efforts in setting up and performing the *in situ* EBSD experiment, which was critical to the success of this study. The authors also thank Annastasia Bedford (ANSTO) for preparing the EBSD specimens used in the *in situ* experiment. The project received funding from the NSW Government and ANSTO through the Future Now scholarship programme. The contributions of these two organisations have been invaluable in the successful completion of this work. Additionally, the work of MCM and TH was sponsored by the U.S. Department of Energy under Contract No. DE-AC02-06CH11357 with Argonne National Laboratory, managed and operated by UChicago Argonne LLC.

Appendix A

Figs. A1 and A2 visualise the experimentally observed and simulated deformation of the Alloy 617 specimen along 24 irregular strain intervals. Specifically, Fig. A1 shows electron backscatter diffraction (EBSD) maps of the specimen's microstructure, obtained through *in situ* EBSD. In contrast, Fig. A2 shows the deformation of the high-fidelity finite element mesh of the specimen gauge, simulated using the Voce hardening (VH) model.

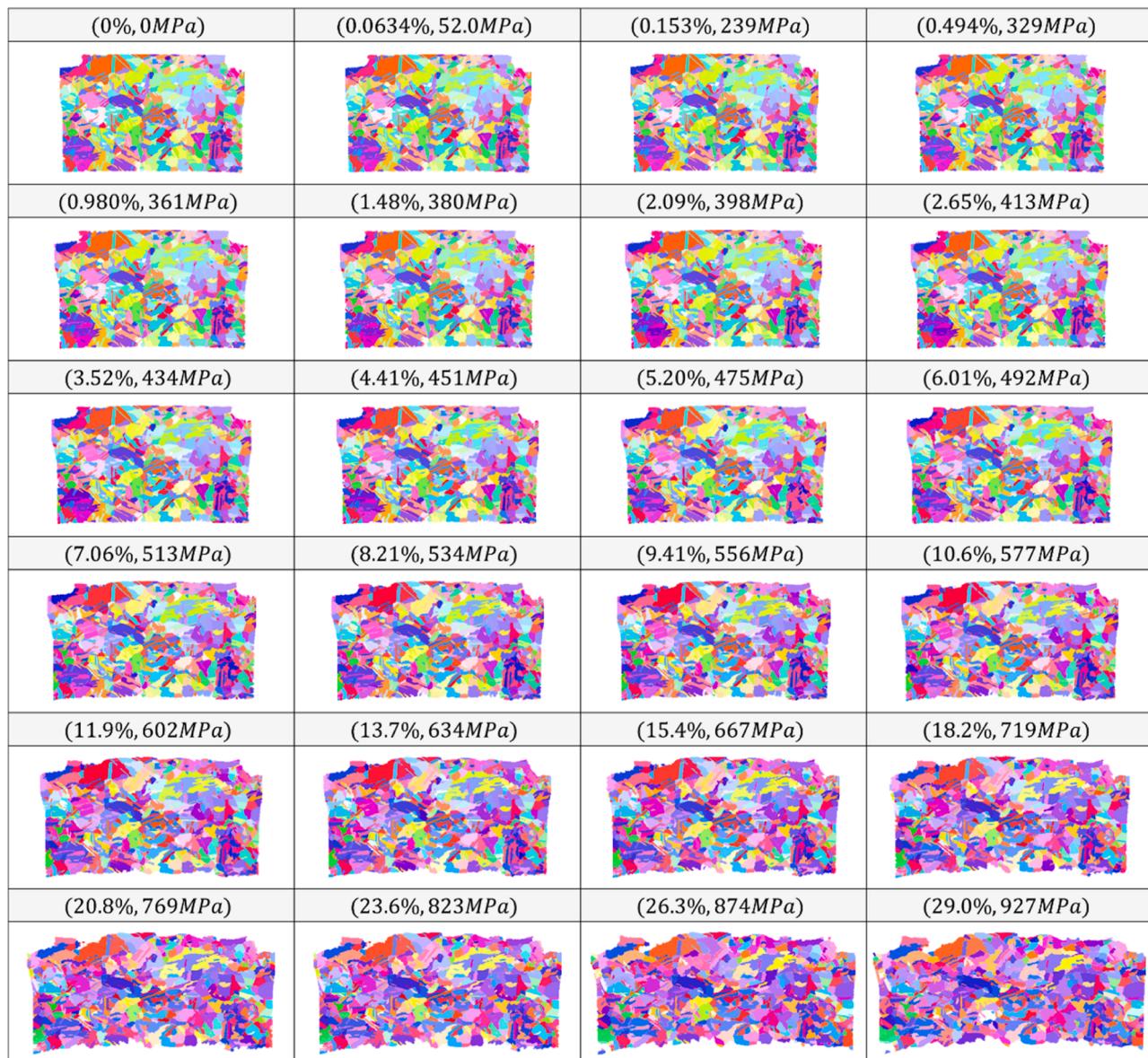


Fig. A1. Electron backscatter diffraction (EBSD) maps of the Alloy 617 specimen shown as IPF-x coloured maps.



Fig. A2. Simulated deformation of the high-fidelity finite element mesh using the Voce hardening (VH) model shown as IPF-x coloured maps.

Appendix B

An algorithm was developed to track the evolution of individual grains in the specimen as it deformed, using a sequence of electron backscatter diffraction (EBSD) maps. This was achieved by repeatedly matching grains between consecutive EBSD maps based on their centroids and crystallographic orientations. A match between two grains from different EBSD maps would thus indicate that those two grains represent the same physical grain in the specimen. The algorithm represented the EBSD maps as separate graphs, where G_i represents the graph corresponding to the i -th EBSD map in the sequence. The nodes in G_i represented the grains in the i -th EBSD map, with each node containing information about the centroid and orientation of the corresponding grain. Edges between two graphs represented potential grain matches between two EBSD maps.

The algorithm began by matching grains between the first and second EBSD map, which involved creating a list of potential edges between the nodes in G_1 and G_2 . The algorithm created a potential edge between a node in G_1 ($n_{1,j}$) to a node in G_2 ($n_{2,k}$) if the distance between the normalised centroids of the corresponding grains was less than 0.2. The normalised centroid distance (D_C) was calculated using Eq. (9), where x and y are the centroid coordinates, and X and Y are the EBSD map dimensions along the x- and y-axes, respectively. This approach eliminated unnecessary comparisons between grains that were too far apart to be potential matches.

$$D_C = \sqrt{\left(\frac{x_{ij}}{X_i} - \frac{x_{i+1,k}}{X_{i+1}}\right)^2 + \left(\frac{y_{ij}}{Y_i} - \frac{y_{i+1,k}}{Y_{i+1}}\right)^2} \quad (9)$$

Once all the potential edges were created, each edge was allocated a weight ($w_{j,k}$) based on the sum of the normalised centroid distance (D_C) and

normalised geodesic distance (D_G) of the two grains (i.e., $w_{j,k} = D_C + D_G$). The normalised geodesic distance was calculated using Eq. (10), where q represents the orientation of a grain in its quaternion form.

$$D_G = \frac{2}{\pi} \cos^{-1} (q_{ij} \cdot q_{i+1,k}) \quad (10)$$

After assigning weights to all the potential edges, the algorithm arranged the edges in ascending order of their weights. The algorithm then iteratively selected edges with the lowest weights to establish bijective matching between the nodes in G_1 and G_2 , such that each node in G_1 was matched to at most one node in G_2 , and vice versa. Nodes in G_2 that were not attached to any edges (i.e., unmatched grains) were removed, such that only grains that had been previously matched were retained. This approach excluded grains that had not been part of the initial EBSD map, such as those grains formed due to fragmentation or nucleation. In addition, nodes in G_1 that were not attached to any edges were moved to G_2 . This enabled the algorithm to maintain continuity in the tracking of a grain's evolution, even when the grain temporarily lacked matches. The algorithm repeats the process for subsequent consecutive EBSD maps (i.e., G_i and G_{i+1}), until the grains from the initial EBSD map have been tracked across the entire sequence of maps.

This algorithm was applied to the 24 EBSD maps of the Alloy 617 specimen described in Section 2 to track the evolution of individual grains and construct their reorientation trajectories. From the tracked grains, 16 grains were selected to construct reorientation trajectories for use in the calibration ($g_{1..8}$) and validation ($g_{9..16}$) of the CPFEM models. The evolution of the selected grains is shown in Fig. A3. In these figures, the selected grains are shown at five different strains corresponding to five EBSD maps, specifically at 0.0, 5.2, 10.6, 20.8, and 29.0 %. Additionally, the relative size of each grain is preserved within its evolutionary sequence, but not across different grains.

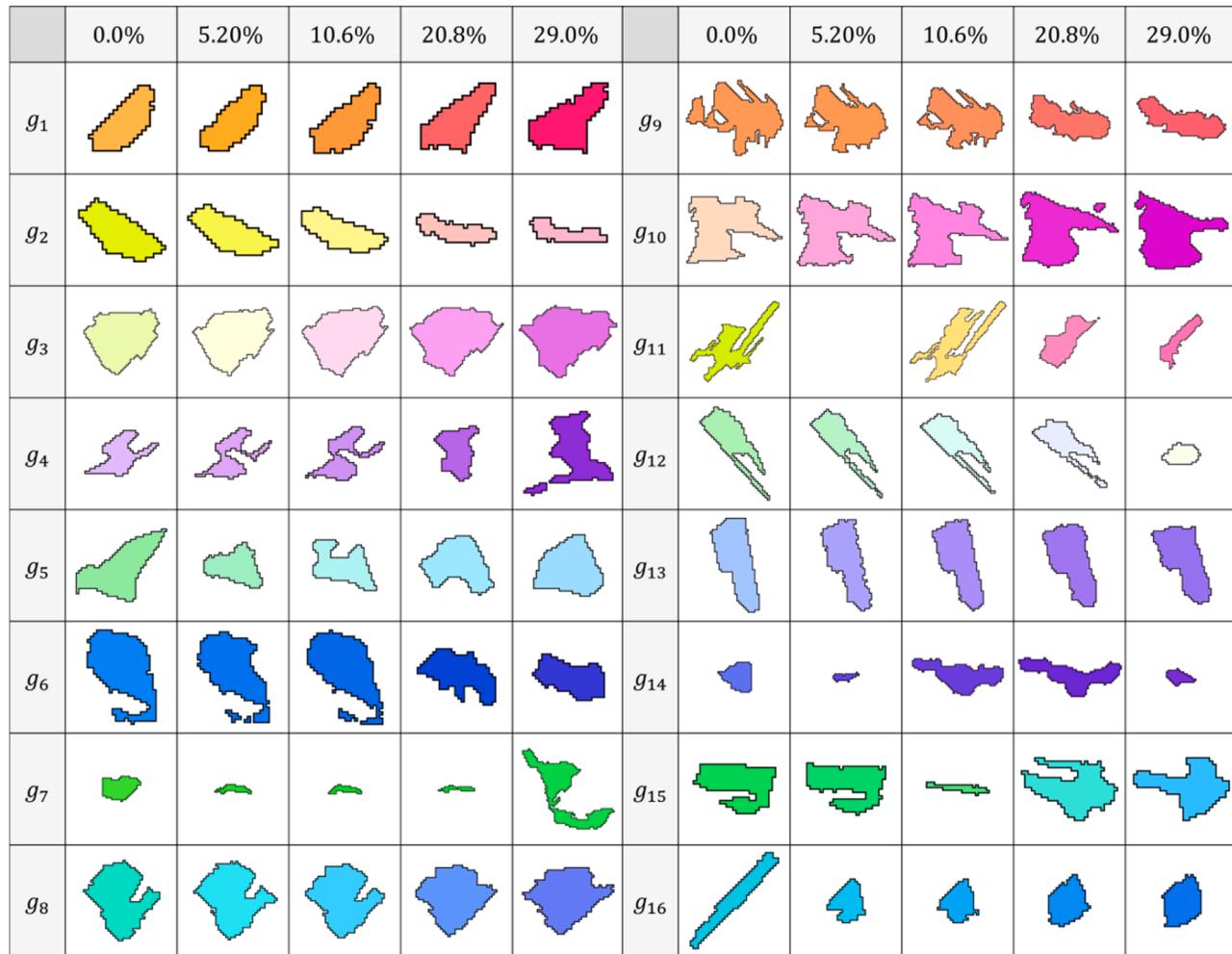


Fig. A3. Evolution of the grains selected for calibration ($g_{1..8}$) and validation ($g_{9..16}$) at $\varepsilon = 0.0\%, 5.2\%, 10.6\%, 20.8\%$, and 29.0% .

The effectiveness of the grain tracking algorithm depends on the accuracy of the grain segmentation procedure. The employed segmentation method could not recognise separated grain segments as being part of the same grain. As a result, the algorithm would temporarily track different segments of the same grain. This can be seen in grains that exhibit merging behaviour (e.g., g_7 and g_{14}) and splitting behaviour (e.g., g_{11} and g_{12}), as well as grains that contain fully-sectioning twins (e.g., g_{16}). However, as long as the tracked segments retained similar orientations to the actual grains, the reorientation trajectories would remain representative of the grain's overall behaviour. Another issue with the grain tracking algorithm was with unmatched grains, where the algorithm temporarily failed to track the grain across one or more EBSD maps, as seen in g_{11} . This issue typically occurred when the grain underwent significant deformation, temporary fragmentation, or subtle shifts in centroid position and orientation, preventing the grain from being tracked. To address this issue, the grain tracking algorithm reintegrated previously tracked grains into the tracking sequence, thereby maintaining continuity. Since the EBSD maps were captured at relatively small strain intervals, the algorithm was still able to construct a continuous reorientation trajectory for g_{11} .

Appendix C

Table C1 contains a summary of the constitutive parameters for the Voce hardening (VH) model, the two-coefficient latent hardening (LH2) model, and the six-coefficient latent hardening (LH6) model. The table also contains the parameters bounds used in the calibration of those hardening models as well as the referenced literature used to determine those bounds. Note that the parameter bounds were defined to provide a sufficiently large search space to demonstrate the capabilities of the adaptive calibration workflow with the hardening models.

Table C1

Summary of constitutive parameters for the Voce hardening (VH), two-coefficient latent hardening (LH2), and six-coefficient latent hardening (LH6) formulations.

Model	Parameter	Symbol	Units	Bounds	References
VH, LH2, LH6	Static slip system strength	τ_0	MPa	(0, 500)	[80,81]
VH, LH2, LH6	Strain rate sensitivity parameter	n	—	(1, 20)	[81–83]
VH	Saturation slip system strength	τ_s	MPa	(0, 4000)	[80–82]
VH	Saturation rate parameter	b	—	(0, 10)	[81,83]
LH2, LH6	Self-hardening coefficient	h_{aa}	MPa	(0, 1000)	[55,68,69]
LH2	Latent hardening coefficient	h_{ap}	MPa		
LH6	Coplanar interaction coefficient	h_1	MPa		
LH6	Hirth interaction coefficient	h_2	MPa		
LH6	Collinear interaction coefficient	h_3	MPa		
LH6	Glissile interaction coefficient	h_4	MPa		
LH6	Lomer interaction coefficient	h_5	MPa		

Appendix D

The alignment between the simulated and experimental responses is evaluated using two objective functions — E_σ from Eq. (7) which quantifies the discrepancies between the experimental and simulated stress-strain responses, and E_ϕ from Eq. (8) which quantifies the discrepancies between the experimental and simulated reorientation trajectories. Among multiple simulated responses, the response that achieves the best agreement with the experimental data is selected as the knee point on the Pareto front. In this study, the knee point is identified by normalising the objective values (E_σ and E_ϕ) to the range [0,1], and then selecting the solution that minimises the sum of their squares (i.e., $\hat{E}_\sigma^2 + \hat{E}_\phi^2$). In other words, the knee point corresponds to the minimal Euclidean distance from the origin (0, 0) in the normalised objective value space.

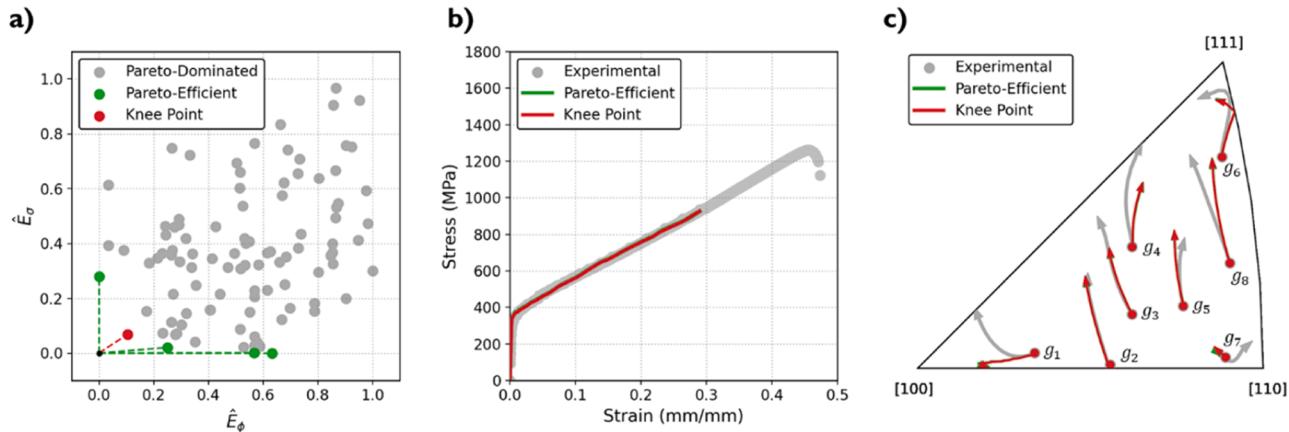


Fig. A4. Results from the Pareto front of the final population of a MOGA optimisation performed on the surrogate model for the VH formulation. Panel a) presents the normalised objective values (\hat{E}_σ and \hat{E}_ϕ), showing the Pareto-dominated solutions (grey), Pareto-efficient solutions (green), and knee point solution (red). Panels b) and c) present the stress-strain response and reorientation trajectories of $g_{1..8}$ predicted by the surrogate model, comparing the experimental data (grey), Pareto-efficient solutions (green), and knee point solution (red).

The knee point is used to determine the best solution in the final population of the MOGA optimisations and to select the best calibration run among multiple independent runs. An example of the knee-point identification is presented in Fig. A4, obtained from a MOGA optimisation of the surrogate model for the VH formulation. The normalised objective values (\hat{E}_σ and \hat{E}_ϕ) of the MOGA's final population are shown in Fig. A4a, highlighting the Pareto-efficient solutions in green and the knee-point solution in red. The corresponding stress-strain responses and reorientation trajectories of $g_{1..8}$ predicted by the surrogate model are also presented in Fig. A4b and c, respectively, which show the surrogate model's responses across the Pareto front achieving comparably close agreement with the experimental data.

Appendix E

The adaptive calibration workflow (Section 4.3) was designed to efficiently calibrate the CPFEM models. In the first iteration of the workflow, the simulation results from several CPFEM evaluations were used to generate the initial training dataset. In subsequent iterations, the results from an additional CPFEM evaluation were used to update the training dataset. Since CPFEM evaluations account for a substantial portion of the calibration workflow's computational expense, reducing the number of CPFEM evaluations is crucial.

As such, a sensitivity study was conducted to determine the initial dataset set that minimises the total number of CPFEM evaluations while sufficiently calibrating the CPFEM model. The sensitivity study involved calibrating the Voce Hardening (VH) model across varying initial dataset sizes, ranging from 2 to 16 evaluations in increments of 2. The calibration runs were repeated until termination, five times for each initial dataset size, with the total number of iterations averaged across the runs. The averaged results were then plotted to visualise the influence of the initial dataset size on the number of additional evaluations required, as shown in Fig. A5.

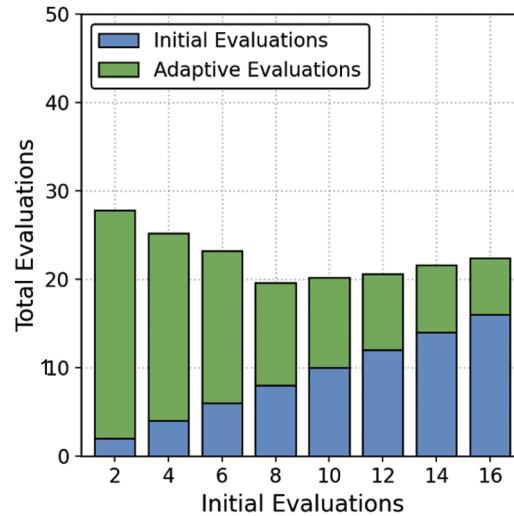


Fig. A5. Plot of the total number of evaluations to calibrate the Voce hardening (VH) model with the adaptive surrogate-modelling-based calibration workflow using various initial dataset sizes.

When the calibration workflow was performed with an initial dataset size ranging from 2 to 8, the number of additional evaluations decreased as the initial dataset size increased. This was expected, since a larger initial dataset would improve the accuracy of the surrogate model in approximating the VH model's responses, reducing the need for additional evaluations. Importantly, the total number of evaluations also decreased. However, beyond an initial dataset size of 8, no further reduction in additional evaluations was observed, which led to an overall increase in the total number of evaluations. These results suggest that an initial dataset size of 8 minimises the total number of evaluations, and thus the computational cost of the calibration process.

Appendix F

In addition to predictive performance and transferability, the practical applicability of CPFEM models also depends heavily on their computational cost. To quantify this aspect, each CPFEM model was evaluated using 192 cores across dual AMD EPYC 96-core processors. The average runtime for a single CPFEM evaluation varied significantly across the three hardening models. Specifically, each evaluation of the low-fidelity VH formulation required approximately 0.23 h of CPU time, while the LH2 and LH6 formulations required roughly 1.46 h. This increased cost for the latent hardening models reflects the additional internal state variables and nonlinear coupling introduced by the latent hardening framework, which increases both the number and stiffness of the equations being solved during each simulation.

As described in Section 5.1, the surrogate-assisted calibration workflow required an initial set of CPFEM simulations followed by several additional evaluations as part of the adaptive refinement process. Specifically, the VH and LH2 formulations required 8 initial evaluations, while the LH6 formulation required 16 due to its larger parameter space. On average, an additional 13, 16, and 22 evaluations were needed for the VH, LH2, and LH6 formulations, respectively, to reach convergence in the multi-objective optimisation. The expected computational cost of calibrating each CPFEM model is summarised in Fig. A6. The plot clearly illustrates that the computational demand increases significantly with model complexity. The total CPU time required to calibrate the LH2 and LH6 formulations is approximately 7–12 times greater than that required for the VH formulation. The steep increase in runtime for the LH6 formulation reflects not only the cost per evaluation but also the greater number of evaluations required to sufficiently explore and constrain its higher-dimensional parameter space.

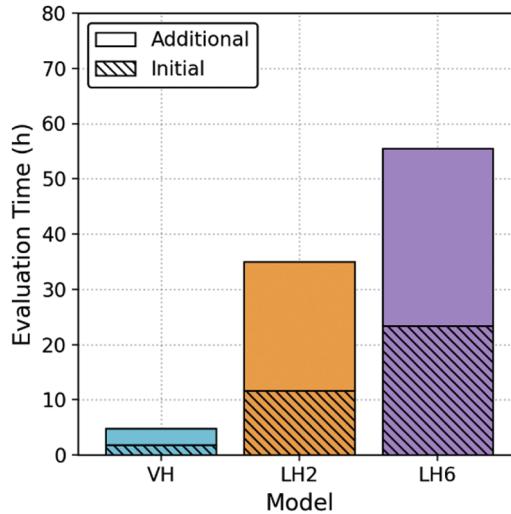


Fig. A6. Comparison of the computational requirements for the calibration of the VH formulation (in cyan), LH2 formulation (in orange), and LH6 formulation (in purple). The average costs for the initial evaluations are represented by single-hashed bars while the costs for the additional evaluations are represented by plain bars.

Taken together with the results presented in previous sections, these findings suggest that the VH formulation provides the most favourable balance between computational efficiency and predictive accuracy. It consistently reproduces the stress-strain response, grain reorientation behaviour, and texture evolution of Alloy 617 under uniaxial tensile loading, while requiring only a fraction of the computational resources consumed by the more complex latent hardening formulations. While the LH2 and LH6 formulations offer greater flexibility and may be better suited for applications involving more complex loading paths or multiaxial stress states, the VH formulation emerges as the most practical choice for routine applications focused on uniaxial deformation with *in situ* validation data.

Notably, the computational cost of the calibration workflow is expected to increase with both microstructural size and loading complexity. Larger microstructures containing more grains will require finer mesh resolutions to accurately represent microstructural heterogeneity, where the cost of each CPFEM evaluation will increase roughly in proportion to the total number of finite elements. Likewise, non-monotonic or multiaxial loading conditions often require smaller time increments and additional solver iterations per increment, further increasing evaluation costs [84]. These factors may also introduce greater variability in the model responses, necessitating more CPFEM evaluations to adequately sample the parameter space and train the surrogate model. That said, an assessment of the workflow's scalability under such conditions is beyond the scope of the present study.

Appendix G

To further assess the robustness of the calibration workflow, an additional test was performed on the VH formulation by swapping the calibration and validation grain datasets. Using this alternate approach, the VH formulation underwent five additional calibration runs against the experimental stress-strain curve and the reorientation trajectories of $g_{9..16}$. After each run, the calibrated model was validated against the $g_{1..8}$ trajectories and the overall texture evolution. The calibration and validation results are presented in Fig. A7, where the grains $g_{1..8}$ were used for calibration and $g_{9..16}$ for validation.

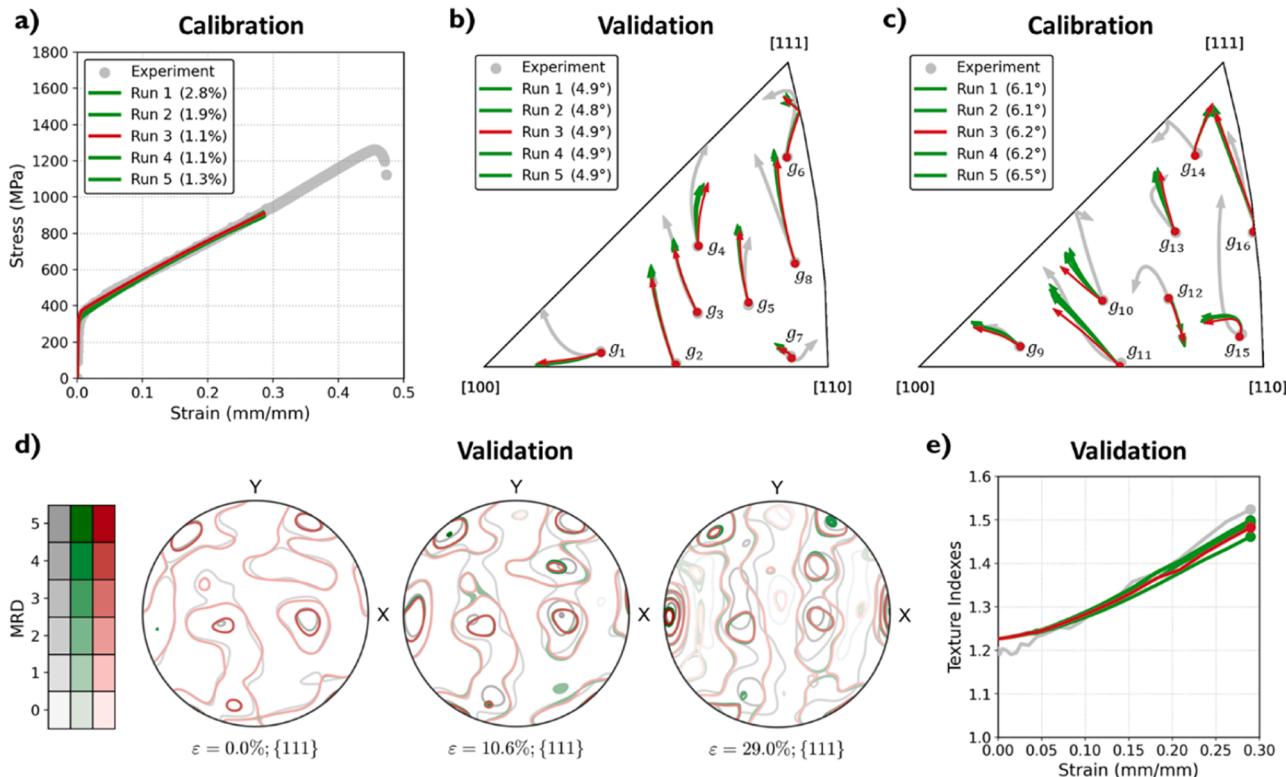


Fig. A7. Supplementary calibration and validation results for the VH formulation after five runs, comparing the experimental data in grey with the simulated responses in green (all runs) and red (best run). The plots in the top row compare the a) stress-strain responses (calibration), b) reorientation trajectories of $g_{1..8}$ (validation), and c) reorientation trajectories of $g_{9..16}$ (calibration), including the corresponding objective values (E_σ and E_ϕ from Eqs. (7) and (8)). The plots in the bottom row compare the overall texture evolution (validation) using d) contoured {111} pole figures at $\varepsilon = 0.0\%$, 10.6%, and 29.0%, and e) texture indexes over 0.0% to 29.0%.

The macroscopic stress-strain predictions maintained excellent agreement with the experimental data in both calibration approaches. At the mesoscale, both sets of reorientation trajectories (i.e., $g_{1..8}$ and $g_{9..16}$) showed slight improvements when using this alternate approach, with the average geodesic distance error (E_ϕ) decreasing from 4.96° to 4.88° for $g_{1..8}$, and 6.4° to 6.22° for $g_{9..16}$. The overall texture evolution also improved marginally, with the average texture index at $\varepsilon = 29\%$ increasing from 1.46 to 1.48, which is closer to the experimental value of 1.52. While these improvements could suggest that the complex reorientation paths of $g_{9..16}$ offer better constraints on grain rotation and texture predictions, they could also be coincidental. Ultimately, the workflow is largely insensitive to the choice of calibration grains, provided that they exhibit representative and trackable rotation behaviour.

References

- [1] P. Zerbe, B. Schneider, E. Moosbrugger, M. Kaliske, A viscoelastic-viscoplastic-damage model for creep and recovery of a semicrystalline thermoplastic, *Int. J. Solids Struct.* 110–111 (2017) 340–350.
- [2] C.M. Stewart, A.P. Gordon, Strain and damage-based analytical methods to determine the Kachanov-Rabotnov tertiary creep-damage constants, *Int. J. Damage Mech.* 21 (8) (2012) 1186–1201.
- [3] J. Choi, L. Bortolan Neto, R.N. Wright, J.J. Kružic, O. Muránsky, On the prediction of creep behaviour of alloy 617 using Kachanov-Rabotnov model coupled with multi-objective genetic algorithm optimisation, *Int. J. Press. Vessels Pip.* 199 (2022) 104721.
- [4] J.J. Skrzypek, Material Damage Models For Creep Failure Analysis and Design of Structures, Springer Vienna, Vienna, 1999, pp. 97–166.
- [5] C. Gierden, J. Kochmann, J. Waimann, B. Svendsen, S. Reese, A review of FE-FFT-based two-scale methods for computational modeling of microstructure evolution and macroscopic material behavior, *Arch. Comput. Methods Eng.* 29 (6) (2022) 4115–4135.
- [6] L. Minfei, G. Yidong, C. Ze, W. Zhi, S. Erik, Š. Branko, Microstructure-informed deep convolutional neural network for predicting short-term creep modulus of cement paste, *Cem. Concr. Res.* 152 (2022) 106681.
- [7] G. Venkataramani, K. Kirane, S. Ghosh, Microstructural parameters affecting creep induced load shedding in Ti-6242 by a size dependent crystal plasticity FE model, *Int. J. Plast.* 24 (3) (2008) 428–454.
- [8] G. Weber, M. Pinz, S. Ghosh, Machine learning-enabled self-consistent parametrically-upscaled crystal plasticity model for Ni-based superalloys, *Comput. Methods Appl. Mech. Eng.* 402 (2022) 115384.
- [9] T. Hu, H.H.J. Choi, M.C. Messner, A Mechanistic Model for Creep and Thermal Aging in Alloy 709, United States, 2023, p. 56, p. Medium: ED; Sizep.
- [10] O. Sedaghat, H. Abdolvand, A non-local crystal plasticity constitutive model for hexagonal close-packed polycrystals, *Int. J. Plast.* 136 (2021) 102883.
- [11] H. Abdolvand, Development of microstructure-sensitive damage models for zirconium polycrystals, *Int. J. Plast.* 149 (2022) 103156.
- [12] K. Sedighiani, M. Diehl, K. Traka, F. Roters, J. Sietsma, D. Raabe, An efficient and robust approach to determine material parameters of crystal plasticity constitutive laws from macro-scale stress-strain curves, *Int. J. Plast.* 134 (2020) 102779.
- [13] A. Venkataraman, M. Messner, An Initial Framework for the Rapid Qualification of Long-Term Creep Rupture Strength Via Microstructural Modeling, Office of Scientific and Technical Information (OSTI), 2021.

- [14] B. Clausen, D. Brown, C. Tomé, L. Balogh, S. Vogel, Engineering related neutron diffraction measurements probing strains, texture and microstructure, in: Proceedings of the 31st Riso International Symposium on Materials Science, Technical University of Denmark, Riso National Laboratory for Sustainable Energy, 2010.
- [15] A. Aburakhia, A. Bonakdar, M. Molavi-Zarandi, J. Kelleher, H. Abdolvand, Deformation mechanisms of additively manufactured Hastelloy-X: a neutron diffraction experiment and crystal plasticity finite element modeling, *Mater. Des.* 222 (2022) 111030.
- [16] H. Abdolvand, M.R. Daymond, Internal strain and texture development during twinning: comparing neutron diffraction measurements with crystal plasticity finite-element approaches, *Acta Mater.* 60 (5) (2012) 2240–2248.
- [17] S.R. Agnew, D.W. Brown, C.N. Tomé, Validating a polycrystal model for the elastoplastic response of magnesium alloy AZ31 using in situ neutron diffraction, *Acta Mater.* 54 (18) (2006) 4841–4852.
- [18] E. Mengiste, D. Piedmont, M.C. Messner, M. Li, J. Stubbins, J.-S. Park, X. Zhang, M. Kasemer, Effect of irradiation-induced strength anisotropy on the reorientation trajectories and fragmentation behavior of grains in BCC polycrystals under tensile loading, *Acta Mater.* 263 (2024) 119503.
- [19] Y.-c. Shao, T. Tang, D. Li, W. Tang, Y. Peng, Crystal plasticity finite element modelling of the extrusion texture of a magnesium alloy, *Model. Simul. Mat. Sci. Eng.* 23 (2015).
- [20] F. Lin, M. Giannetta, M. Jugle, S. Couper, B. Dunleavy, L. Miyagi, Texture development and stress-strain partitioning in periclase + halite aggregates, *Minerals* (2019).
- [21] O. Ibragimova, A. Brahma, W. Muhammad, J. Lévesque, K. Inal, A new ANN based crystal plasticity model for FCC materials and its application to non-monotonic strain paths, *Int. J. Plast.* 144 (2021) 103059.
- [22] P. Acar, Crystal plasticity model calibration for Ti-7Al alloy with a multi-fidelity computational scheme, *Integr. Mater. Manuf. Innov.* 7 (4) (2018) 186–194.
- [23] T. Chen, M.C. Messner, Training material models using gradient descent algorithms, *Int. J. Plast.* 165 (2023) 103605.
- [24] Z. Yang, A.-W. Elgamal, Application of unconstrained optimization and sensitivity analysis to calibration of a soil constitutive model, *Int. J. Numer. Anal. Methods Geomech.* 27 (2003) 1277–1297.
- [25] A. Andrade-Campos, S. Thuijller, P. Philippe, F. Teixeira-Dias, On the determination of material parameters for internal variable thermoelastic-viscoplastic constitutive models, *Int. J. Plast.* 23 (2007) 1349.
- [26] J. Fernández, J. López-Campos, A. Segade Robleda, J.A. Vilán, A genetic algorithm for the characterization of hyperelastic materials, *Appl. Math. Comput.* 329 (2018) 239–250.
- [27] J. Wang, W. Jiang, Numerical assessment on fatigue damage evolution of materials at crack tip of CT specimen based on CPFE, *Theor. Appl. Fract. Mech.* 109 (2020) 102687.
- [28] P.T.N. Nguyen, F. Abbès, J.S. Lecomte, C. Schuman, B. Abbès, Inverse identification of single-crystal plasticity parameters of HCP zinc from nanoindentation curves and residual topographies, *Nanomaterials (Basel)* 12 (3) (2022).
- [29] B. Dong, T. Gu, Y. Zhang, H. Proudhon, Y.F. Jia, X.J. Pei, X. Long, F.Z. Xuan, Interpretable prediction of sample size-dependent fatigue crack formation lifetime using deep symbolic regression and polycrystalline plasticity models, *Int. J. Fatigue* 199 (2025) 109057.
- [30] G.E.P. Box, N.R. Draper, Empirical Model-Building and Response Surface, John Wiley & Sons, Inc, 1986.
- [31] D.G. Krige, A statistical approach to some basic mine valuation problems on the Witwatersrand, by D.G. Krige, published in the Journal, December 1951: introduction by the author, *J. South Afr. Inst. Min. Metall.* 52 (9) (1952) 201–203.
- [32] S. Hwang, S.-K. Choi, Deep learning-based surrogate modeling via physics-informed artificial image (PiAI) for strongly coupled multidisciplinary engineering systems, *Knowl. Based Syst.* 232 (2021) 107446.
- [33] C. Bonatti, B. Berisha, D. Mohr, From CP-FFT to CP-RNN: recurrent neural network surrogate model of crystal plasticity, *Int. J. Plast.* 158 (2022) 103430.
- [34] H. Dorward, D.M. Knowles, E. Demir, M. Mostafavi, M.J. Peel, Calibration and surrogate model-based sensitivity analysis of crystal plasticity finite element models, *Mater. Des.* 247 (2024) 113409.
- [35] W. Ren, R. Swindeman, A review on current status of alloys 617 and 230 for gen IV nuclear reactor internals and heat exchangers, *Journal of Pressure Vessel Technology-transactions of The ASME -*, *J. Press. Vessel Technol.* 131 (2009) 044002–044015.
- [36] K. Maile, Qualification of Ni-based alloys for advanced ultra supercritical plants, *Proc Eng.* 55 (2013) 214–220.
- [37] X. Xie, C.-y. Chi, S. Zhao, J. Dong, F. Lin, Superalloys and the development of advanced ultra-supercritical power plants, *Mater. Sci. Forum* 747 (2013) 594–603.
- [38] P. Mandal, A. Lal, A.M. Sajil, M. Jacob, Cryogenic behavior of TiG welded Ni-based superalloy (IN617), 2022, pp. 207–219.
- [39] W. Ren, Considerations of Alloy 617 Application in the Gen IV Nuclear Reactor Systems - Part II, Metallurgical Property Challenges, 2009.
- [40] J.K. Wright, Next Generation Nuclear Plant Steam Generator and Intermediate Heat Exchanger Materials Research and Development Plan, Office of Scientific and Technical Information (OSTI), 2010.
- [41] R.E. Mizia, Next Generation Nuclear Plant Intermediate Heat Exchanger Acquisition Strategy, Office of Scientific and Technical Information (OSTI), 2008.
- [42] Q. Meng, Z. Wang, Creep damage models and their applications for crack growth analysis in pipes: a review, *Eng. Fract. Mech.* 205 (2019) 547–576.
- [43] L.J. Carroll, C. Cabet, M.C. Carroll, R.N. Wright, The development of microstructural damage during high temperature creep-fatigue of a nickel alloy, *Int. J. Fatigue* 47 (2013) 115–125.
- [44] H.E. McCoy, J.F. King, Mechanical properties of inconel 617 and 618, United States, 1985.
- [45] R.N. Wright, Draft ASME boiler and pressure vessel code cases and technical bases for use of alloy 617 for constructions of nuclear component under section III, division 5, United States, 2021, p. Medium: ED; Size: 384 p.
- [46] ASTM E 8-11, Standard Test Methods for Tension Testing of Metallic Materials, ASTM International, 2011.
- [47] J. Wright, N. Lybeck, Tensile Properties of Alloy 617 Bar Stock, Idaho National Lab. (INL), Idaho Falls, ID (United States), 2013.
- [48] F. Niessen, T. Nyssönen, A.A. Gazder, R. Hielscher, Parent grain reconstruction from partially or fully transformed microstructures in MTEX, *J. Appl. Crystallogr.* 55 (1) (2022) 180–194.
- [49] Coreform Cubit, Coreform LLC, Orem, UT, 2022.
- [50] S. François, M. Schenven, P. Galván, G. Lombaert, G. Degrande, A 2.5D coupled FE-BE methodology for the dynamic interaction between longitudinally invariant structures and a layered halfspace, *Comput. Methods Appl. Mech. Eng.* 199 (23) (2010) 1536–1548.
- [51] N.E.M.L.: The Nuclear Engineering material Model Library, Argonne National Laboratory, 2024.
- [52] A.N. Laboratory, DEER, 2022.
- [53] C. Permann, D. Gaston, D. Andrs, R. Carlsen, F. Kong, A. Lindsay, J. Miller, J. Peterson, A. Slaughter, R. Stogner, R. Martineau, MOOSE: enabling massively parallel multiphysics simulation, 2019.
- [54] E. Voce, The relationship between stress and strain for homogeneous deformation, *J. Inst. Met.* 74 (1948) 537–562.
- [55] J. Mandel, Generalisation de la theorie de plasticite de W. T. Koiter, *Int. J. Solids Struct.* 1 (3) (1965) 273–295.
- [56] A. Ma, F. Roters, D. Raabe, A dislocation density based constitutive model for crystal plasticity FEM including geometrically necessary dislocations, *Acta Mater.* 54 (8) (2006) 2169–2179.
- [57] H. Mecking, U. Kocks, Kinetics of flow and strain-hardening, *Acta metall.* 29 (11) (1981) 1865–1875.
- [58] A. Alankar, I.N. Mastorakos, D.P. Field, A dislocation-density-based 3D crystal plasticity model for pure aluminum, *Acta Mater.* 57 (19) (2009) 5936–5946.
- [59] R.J. Asaro, Micromechanics of crystals and polycrystals, in: J.W. Hutchinson, T. Y. Wu (Eds.), Advances in Applied Mechanics, Elsevier, 1983, pp. 1–115.
- [60] G.I. Taylor, The spectrum of turbulence, *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* 164 (919) (1938) 476–490.
- [61] K. Zhang, B. Holmedal, T. Máni, A. Saai, Assessment of advanced Taylor models, the Taylor factor and yield-surface exponent for FCC metals, *Int. J. Plast.* 114 (2019) 144–160.
- [62] H.J. Bunge, Some applications of the Taylor theory of polycrystal plasticity, *Krist. Tech.* 5 (1) (1970) 145–175.
- [63] U.F. Kocks, The relation between polycrystal deformation and single-crystal deformation, *Metall. Mater. Trans.* 1 (5) (1970) 1121–1143.
- [64] Chapter 6 - experimental studies of Peierls-Nabarro-type friction forces in metals and alloys, in: D. Caillard, J.L. Martin (Eds.), Pergamon Materials Series, Pergamon, 2003, pp. 159–224.
- [65] S. Alkan, A. Ojha, H. Sehitoglu, Determination of latent hardening response for FeNiCoCrMn for twin-twin interactions, *Acta Mater.* 147 (2018) 149–164.
- [66] M. Zecevic, M. Knezevic, Latent hardening within the elasto-plastic self-consistent polycrystal homogenization to enable the prediction of anisotropy of AA6022-T4 sheets, *Int. J. Plast.* 105 (2018) 141–163.
- [67] Y. Wu, M. Bönisch, S. Alkan, W. Abuzaid, H. Sehitoglu, Experimental determination of latent hardening coefficients in FeMnNiCoCr, *Int. J. Plast.* 105 (2018) 239–260.
- [68] P. Franciosi, M. Berveiller, A. Zaoui, Latent hardening in copper and aluminium single crystals, *Acta Metall.* 28 (3) (1980) 273–283.
- [69] C. Gérard, G. Cailletaud, B. Bacroix, Modeling of latent hardening produced by complex loading paths in FCC alloys, *Int. J. Plast.* 42 (2013) 194–212.
- [70] C. Gérard, B. Bacroix, M. Bornert, G. Cailletaud, J. Crépin, S. Leclercq, Hardening description for FCC materials under complex loading paths, *Comput. Mater. Sci.* 45 (3) (2009) 751–755.
- [71] M. Franulović, R. Basan, I. Prebil, Genetic algorithm in material model parameters' identification for low-cycle fatigue, *Comput. Mater. Sci.* 45 (2) (2009) 505–510.
- [72] J. Blank, K. Deb, Pymoo: multi-objective optimization in Python, *IEEE Access* 8 (2020) 89497–89509.
- [73] Y. Lin, W. Yang, Application of multi-objective genetic algorithm based simulation for cost-effective building energy efficiency design and thermal comfort improvement, *Front. Energy Res.* 6 (2018).
- [74] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in PyTorch, 2017.
- [75] G. Manache, C. Melching, Sensitivity of Latin hypercube sampling to sample size and distributional assumptions, 2007.
- [76] S. Tripoppoom, W. Yu, K. Sepehrnoori, J. Miao, Chapter 2 - methodology, in: S. Tripoppoom, W. Yu, K. Sepehrnoori, J. Miao (Eds.), Assisted History Matching For Unconventional Reservoirs, Gulf Professional Publishing, 2021, pp. 17–35.
- [77] M. Vorechovsky, Hierarchical refinement of Latin hypercube samples, *Comput.-Aided Civ. Infrastruct. Eng.* (2014) accepted.
- [78] X. Long, Z. Shen, Q. Jia, J. Li, R. Dong, Y. Su, X. Yang, K. Zhou, Determine the unique constitutive properties of elastoplastic materials from their plastic zone evolution under nanoindentation, *Mech. Mater.* 175 (2022) 104485.

- [79] H. Abdolvand, M.R. Daymond, C. Mareau, Incorporation of twinning into a crystal plasticity finite element model: evolution of lattice strains and texture in Zircaloy-2, *Int. J. Plast.* 27 (11) (2011) 1721–1738.
- [80] M. Shahmardani, N. Vajragupta, A. Hartmaier, Influence of crystal plasticity parameters on the strain hardening behavior of polycrystals, *Crystals* 11 (12) (2021) 1473.
- [81] M.C. Messner, T.-L. Sham, Reference constitutive model for alloy 617 and 316H stainless steel for use with the ASME division 5 design by inelastic analysis rules, United States, 2021, p. Medium: ED; Size: 68 p.
- [82] T. Diehl, Modeling of elastic-viscoplastic behavior and its finite element implementation, 1988.
- [83] V.-T. Phan, M. Messner, S. Sham, A unified engineering inelastic model for 316H stainless steel, 2019.
- [84] X. Wang, F. Cazes, J. Li, A. Hocini, K. Ameyama, G. Dirras, A 3D crystal plasticity model of monotonic and cyclic simple shear deformation for commercial-purity polycrystalline Ti with a harmonic structure, *Mech. Mater.* 128 (2019) 117–128.