



Contents lists available at ScienceDirect

International Journal of Plasticity

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

A machine learning model to predict yield surfaces from crystal plasticity simulations

Anderson Nascimento^{a,*}, Sharan Roongta^b, Martin Diehl^{c,d},
Irene J. Beyerlein^{a,e}

^a Department of Mechanical Engineering, University of California, Santa Barbara, CA, 93106-5070, USA

^b Max-Planck-Institut für Eisenforschung, Max-Planck-Str. 1, Düsseldorf 40237, Germany

^c Department of Materials Engineering, KU Leuven, Kasteelpark Arenberg 44, Leuven 3001, Belgium

^d Department of Computer Science, KU Leuven, Celestijnenlaan 200 A, Leuven 3001, Belgium

^e Materials Department, University of California, Santa Barbara, CA, 93106-5050, USA

ARTICLE INFO

Keywords:

A. Yield condition
B. Crystal plasticity
Machine learning
Neural network

ABSTRACT

We introduce a microstructurally informed machine learning model for predicting the anisotropic yield surfaces of polycrystalline materials. A full-field, spatially resolved crystal plasticity model is employed to generate a data set describing the yield response of an aluminum alloy, enabling the training of a neural network yield function and the calibration of 3D yield criteria of plastically anisotropic polycrystals. This novel formulation explores the flexibility of neural networks to describe complex-shaped yield loci and avoids common problems associated with conventional 3D yield functions, such as the non-trivial parameter identification and non-uniqueness of the anisotropy coefficients. Here, Bayesian optimization is applied to obtain an optimal neural network architecture and allows for an automated model design. The neural network yield function is able to learn intrinsic properties such as the convexity of the yield hull and tension–compression symmetry from a relatively small number of data points. The fully data-driven yield criterion can accurately reproduce multiaxial flow response and planar anisotropy despite of its material blind initial state. Stress gradients can also be computed from the neural network through automatic differentiation as derived quantities with good fidelity. This allows the calculation of *r*-values and provides a pathway for implementing the neural network yield model into finite element codes.

1. Introduction

Polycrystalline alloys often inherit crystallographic texture from thermo-mechanical treatments during both processing and manufacturing (Raabe et al., 2004). Such texture renders the plastic flow behavior of these materials intrinsically anisotropic. In order to accurately describe the directional dependence of plastic flow, crystal plasticity (CP) models that track lattice rotation of individual grains have been developed. While these models can predict both polycrystal anisotropy and texture evolution with high fidelity, their feasibility for engineering scale simulations of metal forming processes is undermined by their high computational cost (Roters et al., 2010). In fact, the plastic anisotropy of a wide range of metals and alloys at the continuum level is often described by analytical yield functions due to their efficiency and accurate yield prediction (Shutov and Ihlemann, 2013).

* Corresponding author.

E-mail address: andersonw@ucsb.edu (A. Nascimento).

<https://doi.org/10.1016/j.ijplas.2022.103507>

Received 2 November 2022; Received in revised form 11 December 2022

Available online 28 December 2022

0749-6419/© 2022 Elsevier Ltd. All rights reserved.

The first yield criterion was proposed by von Mises (1928). This quadratic model asserts that plastic flow starts when the J_2 invariant of the CAUCHY stress reaches a critical value and constitutes the basis on which more accurate and specialized criteria have been developed. A likewise quadratic model, featuring pressure insensitivity and orthotropic symmetry was formulated by Hill (1948) and is notably precise for the yield prediction of body-centered cubic (bcc) metals (Han et al., 2020). While ensuring convexity, the quadratic model limits the predictive capacity for materials with higher equi-biaxial flow stress, which often leads to complex-shaped yield loci. This quadratic requirement has been relaxed in generalized models such as the ones proposed by Hosford (1979) and Hill (1979) furnishing additional fitting capacity to these functions.

Despite the certain level of flexibility provided by the newly introduced variable exponent and anisotropy coefficients in both Hosford's and Hill's yield criteria, these functions are yet not able to fully capture the flow behavior of appreciably anisotropic materials. In this context, models based on linear transformation of the stress tensor (Barlat et al., 2007) have demonstrated accuracy in the description of the yield locus of materials featuring different types of symmetry. For plane stress and orthotropic symmetry, Banabic et al. (2000) proposed a model with seven anisotropy coefficients which has shown considerable flexibility and Barlat et al. (1991) used a linear transformation based yield function, Yld91, to describe orthotropic materials under general stress conditions. In fact, Yld91 is part of a family of phenomenological yield surfaces which includes, but is not limited to, Yld91 (Barlat et al., 1991), Yld2000-2D (Barlat et al., 2003), Yld2004-13p, Yld2004-18p (Barlat et al., 2005) and Yld2004-27p (Aretz et al., 2010). These functions feature respectively 6, 8, 13, 18 and 27 anisotropy parameters that allow for a better description of anisotropic properties. Van Houtte et al. (2009) presented a convex yield function able to describe the response of materials displaying stress differential effects, in particular, HCP metals and pre-strained cubic metals. Lou et al. (2022) developed a general anisotropic yield function also capable of modeling strength differential effects and differential hardening response under a variety of stress states.

While the advanced 3D yield functions have proven accurate in modeling plastic anisotropy, a relevant difficulty in their usage remains in the non-trivial parameter identification process. As an example, the initially proposed approach for calibrating Yld2004-18p consists of nine mechanical and two virtual tests. The mechanical ones are comprised of seven uniaxial, balanced biaxial and the disk compression tests. These tests are experimentally practical and provide meaningful in-plane data points. For an out-of-plane probing, however, the lack of simple experiments for attaining the yz and zx stress components shifted the data retrieval from experimental tests to polycrystal simulations. Although the chosen set of tests provides valuable description of the yield surface shape, most of the stress space is left unexplored (Zhang et al., 2016), which motivated the investigation of further calibration methods.

Experimental calibration approaches are based purely on data obtained via mechanical tests. Over the years, different calibration methods have been analyzed. For instance, Malo et al. (1998) used tensile and bending tests to fit Hill's yield criterion without the use of r -values, introducing a purely stress-based calibration method. Abedini et al. (2018) employed a genetic algorithm to obtain the optimal anisotropy coefficients of the Yld2000-2D yield function based on the mean-squared error between uniaxial yield stresses and r -values at different orientations. Cazacu (2020) used experimental data at different directions to determine the optimal parameters associated with the orthotropic yield criterion proposed by Karafillis and Boyce (1993). Du et al. (2022) used 19 cruciform biaxial tensile tests under different loading paths to describe the yield response of an FCC AA6016-T4 and a BCC DP490. Despite various calibration techniques, experimental methods are usually bound to stress states that can be achieved via mechanical tests, which can limit their capacity to describe complex yield loci.

Hybrid approaches that incorporate experimental and simulation data into the calibration have also been developed to overcome experimental limitations. Grytten et al. (2008) studied different calibration methods for the Yld2004-18p yield function, employing uniaxial tension tests at different directions from the RD, compression tests along the ND and stress data points obtained from 690 simulations adopting the full-constraint Taylor model (Taylor, 1938) at evenly distributed strain paths. Zhang et al. (2015) studied the anisotropy of an AA1050 aluminum alloy with the Yld2004-18p yield function using three different methods: uniaxial tension data at different loading directions, 201 virtual stress data points obtained by different CP models and a mixed approach combining both the experimental tests and the CP data points. Iftikhar et al. (2021) experimentally probed the initial yield surface of an AA6061 aluminum alloy in the $\sigma_{11} - \sqrt{3}\sigma_{12}$ stress-space and used a Taylor-type CPFE model to simulate the yield surface evolution. While experimental data aid in the validation of hybrid methods, it is also worth noting the greater size of the simulation based datasets mentioned above and the general trend towards thorough virtual sampling of the stress space.

Finally, fully simulation-based calibration frameworks have gained notoriety due to their high accuracy and capacity to probe stress states not easily reached experimentally. The FACET method developed by Van Houtte et al. (2009) proposed a calibration scheme that used predictions from multilevel models, such as Taylor and Alamel (Van Houtte et al., 1999, 2005) homogeneously distributed throughout the strain rate space. While this approach overcomes the probing limitations stemming from mechanical tests, the high number of evaluations of the multilevel models imposes a high computational cost. Zhang et al. (2016) developed an approach in which the planar anisotropic properties of an AA3104 aluminum alloy were gathered from randomly generated virtual yield stress data points. The random sampling method they proposed aimed for a more comprehensive data retrieval and consequently a more accurate yield description. Despite this success, an optimal stress probing and calibration framework has yet to be developed.

An additional difficulty associated with these advanced yield functions is the non-uniqueness of the anisotropy coefficients. van den Boogaard et al. (2016) observed, through a sensitivity analysis, dependencies between specific parameters of Yld2004-18p and suggested a reduction to 14 coefficients instead. While analyzing the parameter identifiability of anisotropic yield functions, Zhang et al. (2022) reiterated the correlation between different coefficients, outlined the role of the respective optimization algorithms for parameter determination and further emphasized the importance of carefully chosen information-rich

stress data points in the calibration process. Hence, while the sound mathematical formalism of the linear transformation based yield functions, along with their accurate yield description for a wide range of metals and alloys should be restated, the complex heuristics associated with their parameter determination cannot be overlooked.

Alternative avenues for plastic flow prediction have been studied, and machine learning (ML) based approaches have gained notoriety. The fundamental concept of most ML frameworks consists of inferring from the training data, a surrogate model with predictive capacity that satisfies the underlying physics of the described phenomenon. In this context, [Vlassis and Sun \(2020\)](#) introduced a deep learning formulation to derive the stored elastic energy, yield surface and plastic flow. This modeling approach explored customized loss functions that included first and second order derivatives to ensure that the final neural network (NN) would achieve thermodynamical consistency and interpretability. [Zhang and Mohr \(2020\)](#) successfully described the stress–strain response of a material featuring J_2 plasticity with isotropic hardening using a deep NN, bypassing an explicit representation of the yield surface, flow rule and hardening law. [Bonatti et al. \(2022\)](#) showed that Recurrent Neural Networks can describe the mechanical response of an AA6016-T4 aluminum alloy, acting as CP surrogate models. [Ibragimova et al. \(2021\)](#) demonstrated that NNs trained with data obtained from CP simulations under unique monotonic loading conditions could predict the stress–strain curves and the texture evolution of FCC metals. [Ali et al. \(2019\)](#) successfully predicted the stress–strain response and texture evolution of an AA6063-T6 aluminum alloy under uniaxial tension and simple shear. [Bonatti and Mohr \(2021\)](#) developed a NN model capable of predicting the forming limits of a DP780 steel sheet from the previous loading history. [Hartmaier \(2020\)](#) used a Support Vector Classifier to formulate a decision rule defining whether a stress tensor input would be in the elastic or plastic regime, *i.e.*, effectively acting as a yield surface surrogate. [Shoghi and Hartmaier \(2022\)](#) extended this model with an optimal data generation approach and were able to predict anisotropy in the principal stress space with high accuracy. [Mianroodi et al. \(2021\)](#) used a Deep NN to calculate local stresses in complex microstructures based on the spacial distribution of the material properties achieving substantial computational speed-up. [Fuhg et al. \(2022\)](#) presented a ML framework for predicting bulk yield response under plane stress conditions from CP simulations of textured polycrystals. Their approach effectively incorporates texture parameters into the prediction of macroscopic yield loci from convex NNs.

While the studies described above show the notable flexibility, accuracy and potential of NNs in the scope of yield prediction, important simplifications have been made. Such assumptions include dimensional reduction to the principal stress space, J_2 plasticity, or the prediction of very specific stress states. Additionally, often very large training datasets, on the scale of tens of thousands data points, are required when trying to predict more complex behavior such as entire stress–strain curves and texture evolution. Here, we develop a deep NN based surrogate model, with performance comparable to advanced phenomenological yield functions. We keep the training process similar to the simulation-based calibration of conventional 3D yield functions, providing the pathway for a seamless integration into conventional FE codes. Further, due to optimal design of the NN we can remain on the small data regime and use only 220 yield data points in the training of the NN yield function. The problem is approached in its most general formulation, *i.e.*, in the six dimensional stress space. Full-field CP modeling is employed in conjunction with a spectral solver based on fast FOURIER transform (FFT) ([Lebensohn et al., 2012](#)) to simulate the mechanical response of a highly resolved microstructure ([Roters et al., 2019](#)). The CP simulations probe the anisotropic properties of an aluminum alloy and provide the required flow stress data to be used in the calibration of analytical yield functions and training of the NN. The anisotropy parameters of Yld2004-18p and Yld2004-13p are obtained and their respective yield loci are used for comparison and benchmarking against the NN. Despite the reduced size of our training data set, the obtained NN model is able to capture fundamental properties such as convexity of the yield hull and tension–compression symmetry. It also shows a strong agreement with the advanced yield functions in the prediction of multiaxial stress states as well as planar yield properties, such as uniaxial flow stresses and r -values.

The present work is structured as follows: Initially, an overview of the CP formulation is provided in Section 2, after which an outline of the mathematical formulation of the phenomenological yield functions is given in Section 3. Further, Section 4 describes the essential aspects of neural network regression, while their application into the yield prediction and relevant implementation features is provided in Section 5. The results are presented in Section 6 and discussed in Section 7. Conclusions and suggestions for future work are given in Section 8.

2. Crystal plasticity constitutive description

The mathematical formulation of the constitutive model is based on the multiplicative decomposition of the deformation gradient $\mathbf{F}$ in its elastic ($\mathbf{F}_e$) and plastic ($\mathbf{F}_p$) parts:

$$\mathbf{F} = \mathbf{F}_e \mathbf{F}_p \quad (1)$$

with $\mathbf{F}_e$ accounting for the reversible elastic distortion and lattice rotation, while $\mathbf{F}_p$ is associated with plastic slip. The elastic deformation gradient is used to calculate the second Piola–Kirchhoff stress tensor $\mathbf{S}$ by means of Hooke's law, precisely:

$$\mathbf{S} = \mathbb{C} : \mathbf{E}_e \quad (2)$$

where $\mathbf{E}_e$ represents the elastic Green–Lagrange strain tensor, defined as:

$$\mathbf{E}_e = (\mathbf{F}_e^T \mathbf{F}_e - \mathbf{I}) / 2 \quad (3)$$

and $\mathbb{C}$ denotes the fourth-rank elastic stiffness tensor, fully defined by three independent elastic parameters in the case of cubic crystals. The plastic response, on the other hand, is driven by the plastic velocity gradient $\mathbf{L}_p$ which is defined as a function of the respective deformation gradient $\mathbf{F}_p$, i.e.:

$$\mathbf{L}_p = \dot{\mathbf{F}}_p \mathbf{F}_p^{-1} = \sum_{\alpha} \dot{\gamma}^{\alpha} \mathbf{S}_0^{\alpha} \quad (4)$$

with $\dot{\gamma}^{\alpha}$ indicating the slip rate on the slip system α and $\mathbf{S}_0^{\alpha} = \mathbf{m}^{\alpha} \otimes \mathbf{n}^{\alpha}$ denoting the Schmid tensor obtained by the tensor product between the unit vectors describing the slip direction and slip plane normal, respectively. The evolution of the slip rate is regarded as a function of the resolved shear stress τ^{α} , and the slip resistance g^{α} in each slip system (Hutchinson, 1976):

$$\dot{\gamma}^{\alpha} = \dot{\gamma}_0 \left| \frac{\tau^{\alpha}}{g^{\alpha}} \right|^n \text{sgn}(\tau^{\alpha}) \quad (5)$$

with $\dot{\gamma}_0$ describing the reference slip rate and n the stress exponent. Finally, the interaction between an arbitrary slip system β and the hardening response of the slip system α is modeled within the evolution of the slip resistance as:

$$g^{\alpha} = \sum_{\beta} h_{\alpha\beta} |\dot{\gamma}^{\beta}| \quad (6)$$

where $h_{\alpha\beta}$ is the strain hardening matrix, capturing the anisotropy of the hardening response.

3. Phenomenological yield surfaces

The advanced phenomenological yield surfaces Yld2004-13p and Yld2004-18p (Barlat et al., 2005) provide a reliable benchmark for comparison with the newly proposed NN yield criterion. The predictive accuracy of these functions, with regards to multiaxial yield stress, uniaxial flow, r-values and earing profiles has been thoroughly studied in virtual, experimental and hybrid frameworks. Additionally, they have been used in the analysis of different materials, including but not limited to the aluminum alloys AA7003-T6, AA6063-T6 (Aghani et al., 2009), AA5083-H116 (Grytten et al., 2008), AA7075 (Rong et al., 2021), AA1050 (Zhang et al., 2015), AA3014 (Zhang et al., 2016) and advanced high strength steel (Ma et al., 2022) with different calibration methods and distinct degrees of anisotropy in each of these studies. Here, a concise description of the mathematical formulation of Yld2004-13p and Yld2004-18p is presented.

3.1. Yld2004-13p yield criterion

The Yld2004-13p yield function introduced by Barlat et al. (2005) is defined in terms of the principal values of the linearly transformed stress deviator. It includes 13 anisotropy coefficients and its equivalent stress ϕ has the following general form:

$$\phi = \phi(\tilde{\mathbf{S}}', \tilde{\mathbf{S}}'') \quad (7)$$

$$= \left| \tilde{S}'_1 - \tilde{S}'_2 \right|^m + \left| \tilde{S}'_2 - \tilde{S}'_3 \right|^m + \left| \tilde{S}'_3 - \tilde{S}'_1 \right|^m - \left\{ \left| \tilde{S}'_1 \right|^m + \left| \tilde{S}'_2 \right|^m + \left| \tilde{S}'_3 \right|^m \right\} + \left| \tilde{S}''_1 \right|^m + \left| \tilde{S}''_2 \right|^m + \left| \tilde{S}''_3 \right|^m$$

where the components of $\tilde{\mathbf{S}}'$ and $\tilde{\mathbf{S}}''$ are the principal values of the transformed stress deviator, $\mathbf{s}'$ and $\mathbf{s}''$, via the following contractions:

$$\tilde{\mathbf{s}}' = \mathbf{C}' \mathbf{s} \quad \text{and} \quad \tilde{\mathbf{s}}'' = \mathbf{C}'' \mathbf{s} \quad (8)$$

with $\mathbf{s}$ representing the deviatoric stress while $\mathbf{C}'$ and $\mathbf{C}''$ denote the fourth rank tensors describing the material anisotropy and embedding the calibration coefficients. Such tensors are expressed in Voigt notation as the linear transformation matrices:

$$\mathbf{C}' = \begin{bmatrix} 0 & -1 & -c'_{13} & 0 & 0 & 0 \\ -c'_{21} & 0 & -c'_{23} & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c'_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c'_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c'_{66} \end{bmatrix}, \quad \mathbf{C}'' = \begin{bmatrix} 0 & -c''_{12} & -c''_{13} & 0 & 0 & 0 \\ -c''_{21} & 0 & -c''_{23} & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c''_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c''_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c''_{66} \end{bmatrix} \quad (9)$$

The number of coefficients of Yld2004-13p can be reduced from thirteen to nine if only plane stress states are considered. A further reduction to the number of parameters, however, could lead to locking effects (Barlat et al., 2003).

3.2. Yld2004-18p yield criterion

The Yld2004-18p is a particularly flexible yield function, and due to its additional parameters it can better capture out-of-plane anisotropy (Barlat et al., 2005). It is also based on the linear transformation of the deviatoric stress and is formulated as:

$$\phi = \phi(\tilde{\mathbf{S}}', \tilde{\mathbf{S}}'') \quad (10)$$

$$= \left| \tilde{S}'_1 - \tilde{S}''_1 \right|^m + \left| \tilde{S}'_1 - \tilde{S}''_2 \right|^m + \left| \tilde{S}'_1 - \tilde{S}''_3 \right|^m + \left| \tilde{S}'_2 - \tilde{S}''_1 \right|^m + \left| \tilde{S}'_2 - \tilde{S}''_2 \right|^m + \left| \tilde{S}'_2 - \tilde{S}''_3 \right|^m + \left| \tilde{S}'_3 - \tilde{S}''_1 \right|^m + \left| \tilde{S}'_3 - \tilde{S}''_2 \right|^m + \left| \tilde{S}'_3 - \tilde{S}''_3 \right|^m$$

Eq. (10) can also be succinctly represented by:

$$\phi(\tilde{\mathbf{S}}', \tilde{\mathbf{S}}'') = \sum_{i=1}^3 \sum_{j=1}^3 |\tilde{S}'_i - \tilde{S}''_j|^m \quad (11)$$

where, similarly to Yld2004-13p, the components of $\tilde{\mathbf{S}}'$ and $\tilde{\mathbf{S}}''$ are the principal values of the transformed stress deviator, $\tilde{s}'$ and $\tilde{s}''$, obtained respectively as $\tilde{s}' = \mathbf{C}'\mathbf{s}$ and $\tilde{s}'' = \mathbf{C}''\mathbf{s}$, with $\mathbf{C}'$ and $\mathbf{C}''$ as:

$$\mathbf{C}' = \begin{bmatrix} 0 & -c'_{12} & -c'_{13} & 0 & 0 & 0 \\ -c'_{21} & 0 & -c'_{23} & 0 & 0 & 0 \\ -c'_{31} & -c'_{32} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c'_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c'_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c'_{66} \end{bmatrix}, \quad \mathbf{C}'' = \begin{bmatrix} 0 & -c''_{12} & -c''_{13} & 0 & 0 & 0 \\ -c''_{21} & 0 & -c''_{23} & 0 & 0 & 0 \\ -c''_{31} & -c''_{32} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c''_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c''_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c''_{66} \end{bmatrix} \quad (12)$$

Given the pronounced relevance of in-plane anisotropic properties for sheet metal forming, 14 of the above coefficients describe in-plane yield response, while the parameters c'_{55} , c'_{66} , c''_{55} and c''_{66} provide the description of out-of-plane anisotropic properties (Yoon et al., 2006).

4. Neural network overview

Regression problems handle the prediction of continuous variables based on models constructed from a given pool of data. The problem addressed in this work falls into that category since it consists, fundamentally, in predicting a continuous and smooth function of the stress state at yield onset. In this context, NNs provide a robust tool in the study of yield anisotropy and will be used herein due to their demonstrated accuracy in a variety of flow prediction problems, which include but are not limited to, J_2 plasticity (Jang et al., 2021), flow stress variation during hot rolling (Aghasafari et al., 2014), flow behavior at elevated temperatures of low alloy steel (Lin et al., 2008) and magnesium alloys (Sabokpa et al., 2012). A concise overview of the mathematical formulation of NN regression is presented in what follows.

4.1. Artificial neural networks

The neuron represents the elementary unit of a NN, and performs a basic linear operation defined by:

$$a(\mathbf{w}, \mathbf{x}) = \sum_{i=0}^N w_i x_i \quad (13)$$

where $\mathbf{w}$ denotes the vector of weights (including the bias), which are learned in the training step, $\mathbf{x}$ represents the N -dimensional input vector and a indicates the activation of each neuron. The assemblage of neurons establishes the next level in the NN hierarchical architecture. Specifically, a layer is defined by a combination of multiple neurons and implements the above calculation on all its units, precisely, the output of the neuron j of the first layer is given by:

$$z_j^{(1)} = f \left(\sum_{i=0}^N w_{ji}^{(1)} x_i \right) \quad (14)$$

where f defines a nonlinear differentiable activation function. Here, it should be noted that the NN's ability to model non-linearity stems from nonlinear character of the activation function. Commonly used activation functions include hyperbolic tangent, logistic sigmoid and rectified linear unit (ReLU). The last will be used in this work due to its computational efficiency and better gradient propagation in comparison with the other activation functions (Sharma et al., 2020). The better gradient propagation partially mitigates vanishing gradients for deeper NNs. The expression for the ReLU activation function is given by:

$$f(x) = \max(0, x) \quad (15)$$

The output of the first layer, activated by Eq. (15), is fed forward to the next layer, which will perform a similar operation. In general, the neuron j of the layer n , computes:

$$z_j^{(n)} = f \left(\sum_{k=0}^{N_s} w_{jk}^{(n)} z_k^{(n-1)} \right), \quad n \geq 2 \quad (16)$$

where N_s indicates the size of the preceding layer. It can be observed from Eq. (14) and Eq. (16), that only the first layer operates directly on the input data, whereas subsequent layers operate on activated output from the previous layer. The sequential stacking of multiple layers completes the architecture of a deep feed-forward NN and the final output is calculated as a linear combination over the last layer as a function of the weights and input data, i.e.:

$$y_j(\mathbf{w}, \mathbf{x}) = \sum_{k=0}^{N_s} w_{jk}^{(N_l+1)} z_k^{(N_l)} \quad (17)$$

where N_l denotes the number of hidden layers.

The training of a NN is done by adjusting the weights based on the error calculated with a chosen objective function. In fact, the prediction accuracy of the output variable is strongly associated with the optimal weights obtained during the training process. While a probabilistic interpretation of the output provides a more general analysis of the training step, the classical approach will be used here for simplicity (Bishop, 2006). According with the latter, the NN is regarded as a parametric nonlinear function of the input vector, and the optimal set of weights is obtained by minimizing a predefined loss function. Precisely, the error associated with the weights $\mathbf{w}$, given a set of input vectors $\{\mathbf{x}_n\}$ and target vectors $\{\mathbf{t}_n\}$, is commonly defined by:

$$E(\mathbf{w}|\mathbf{x}_n, \mathbf{t}_n) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{y}_n(\mathbf{w}, \mathbf{x}_n) - \mathbf{t}_n\|^2 \quad (18)$$

where N denotes the number of training data points. The minimization of the mean square error (MSE) of Eq. (18), is done with the backpropagation algorithm (Rumelhart et al., 1988), which consists of an adapted gradient descent optimization scheme and makes use of the partial derivatives of the loss function with respect to any arbitrary weight w to find an approximate minimum. Such an optimal set of weights in combination with the layered feed-forward architecture described above, fully define a deterministic function of the inputs (Bishop, 2006) that can be deployed for prediction of the target variable.

5. The machine learning virtual laboratory (MLVL)

The framework adopted in this study defines a machine learning virtual laboratory (MLVL), integrating the design, training and verification of NN for yield prediction into the virtual laboratory arrangement. The concept of virtual material testing was demonstrated by Kraska et al. (2009) and further elaborated in the works of Zhang et al. (2016) and Ma et al. (2022) by using a simulation based virtual lab for initial yield surface prediction, and Han et al. (2020) while studying the yield surface evolution during forming simulations. Here, we further expand the virtual lab framework by integrating microstructurally informed CP models (Zhang et al., 2015) with an optimally designed data-driven yield criterion.

The overarching configuration of the MLVL consists of a few major steps, as shown in Fig. 1. The initial stage provides the material description and enables the input of microstructural features into the MLVL. A representative volume element (RVE) incorporates the chosen material into a discretized computational model and is further probed under a variety of periodic boundary conditions with DAMASK (Roters et al., 2019), deploying a FFT-based spectral solver in the solution of elastoplastic boundary value problems. This allows the generation of a flow stress data set which will be later used in the training of the NN yield function surrogate. The next step involves the automated design of an optimal architecture for the NN. Here, SigOpt (Clark and Hayes, 2019), an experimentation platform for hyperparameter tuning, is used to obtain the optimal number of layers and the size of each layer via Bayesian optimization. The final stage concerns the training and verification of the model and renders a calibrated NN yield criterion with high fidelity in the prediction of multiaxial flow response, strength and deformation anisotropy as well as plastic flow vectors. These steps are further described in the following sections.

5.1. Material characterization & virtual testing

The framework outlined above is designed to be general and suitable for modeling different degrees of anisotropy. Aiming to demonstrate the capabilities of our proposed NN yield function, we employ an aluminum alloy with the CP constitutive parameters presented in Table 1, while the original grain orientations were obtained from an AA3104-H19 alloy (Sun et al., 2021; Wu et al., 2003). A statistically representative texture consisting of 1000 orientations is sampled from the experimental data obtained by X-ray diffraction using the hybrid integer approximation (hybridIA) developed by Eisenlohr and Roters (2008). The orientation density maps of the reconstructed microstructure are shown in Fig. 2. The model is discretized with $32 \times 32 \times 32$ Fourier points providing a resolution of approximately 33 points per grain, which is suitable for macroscopic yield prediction at low computational cost. A relatively high number of grains was adopted here to allow for a statistically meaningful microstructural representation as well as to reduce statistical discrepancy between the mechanical response of single grains and that of the homogenized RVE, which has been associated with local non-convexity of the yield hull (Hu et al., 2015; Zhang et al., 2016). A cubic RVE embedding the grain distribution sampled from the reconstructed microstructure is used.

The yield data points are obtained by virtual experiments with DAMASK (Roters et al., 2019) under the application of arbitrary periodic boundary conditions (BCs). The application of BCs for virtual testing and the corresponding deformed shape of the RVEs are exemplified in the second step of Fig. 1 with the model under uniaxial tension along RD, TD, and pure shear. The stress-strain curves are presented in Fig. 3 with the corresponding yield stress identified. Finally, both Figs. 3 and 2 initially suggest an overall mildly anisotropic response, which stems mostly from the weak textured used here and is incorporated into the CP model by means of the individual grain orientations.

Table 1
Parameters for CP model of the aluminum alloy under analysis (Eisenlohr et al., 2013).

Property	C_{11}	C_{12}	C_{44}	$\dot{\gamma}_0$	g_0	g_∞	h_0	n	a
Value	106.75	60.41	28.34	0.001	31	63	75	20	2.25
Unit	GPa	GPa	GPa	s^{-1}	MPa	MPa	MPa	–	–

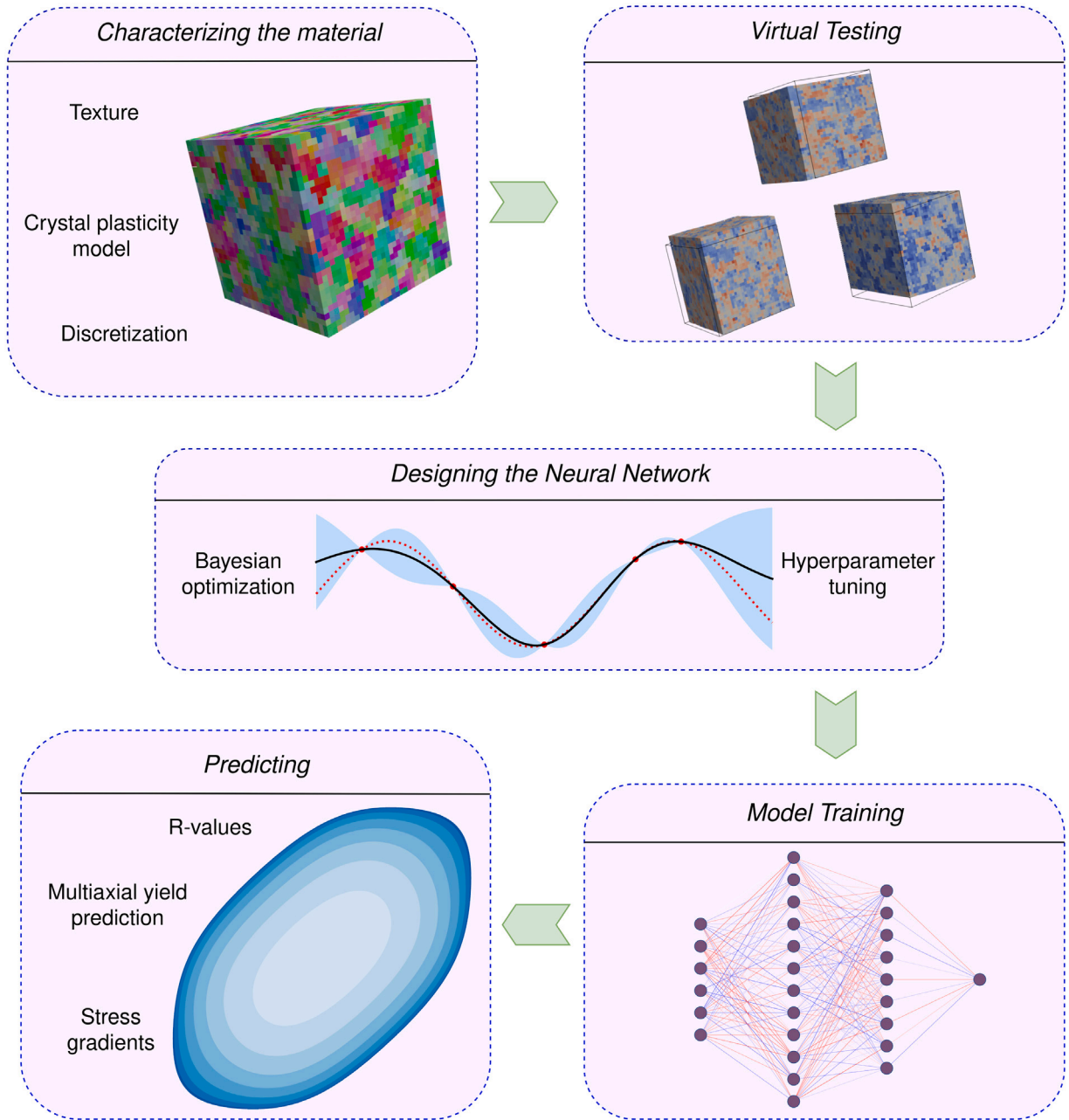


Fig. 1. Outline of the architecture of the machine learning virtual laboratory (MLVL). The workflow consists of material description, virtual testing with a FFT-based spectral solver, optimizing the NN design, training the NN yield function and finally deploying the model for prediction of desired quantities.

5.2. Neural network design

The design of the NN architecture is an important step in the MLVL workflow. However, in view of the black-box nature of NNs, choosing an adequate configuration of hyperparameters is not a trivial task and often relies on experience and complex heuristics. Additionally, training with sub-optimal parameters severely hinders the performance of the trained models and usually leads to inferior predictive capabilities (Smith, 2018). In this context, different optimization schemes have been employed for hyperparameter identification and viable techniques have emerged, which include genetic algorithms (Tsai et al., 2006), racing algorithms (Birattari et al., 2010) and Bayesian optimization (Eggenberger et al., 2013). The last has gained attention due to its efficient functional evaluation (Claesen and De Moor, 2015) and computationally inexpensive implementation, and will therefore, be used in this work to determine the optimal NN design.

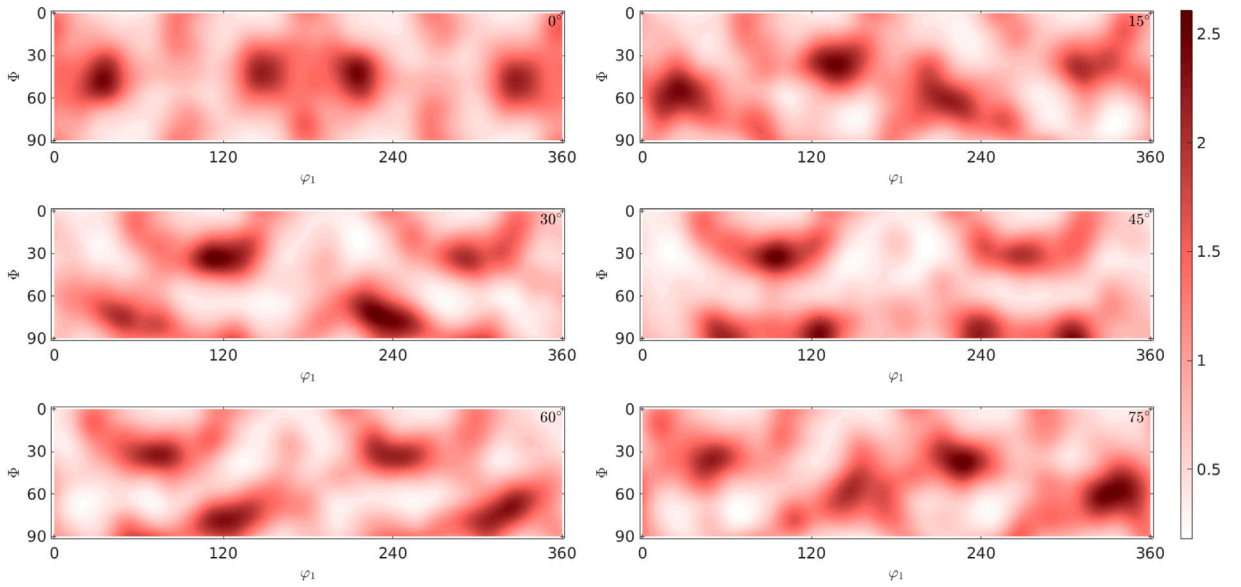


Fig. 2. Orientation density maps obtained from the ODF of the AA3104-H19 aluminum alloy reconstructed microstructure featuring 1000 grains sampled with HYBRIDIA (Eisenlohr and Roters, 2008).

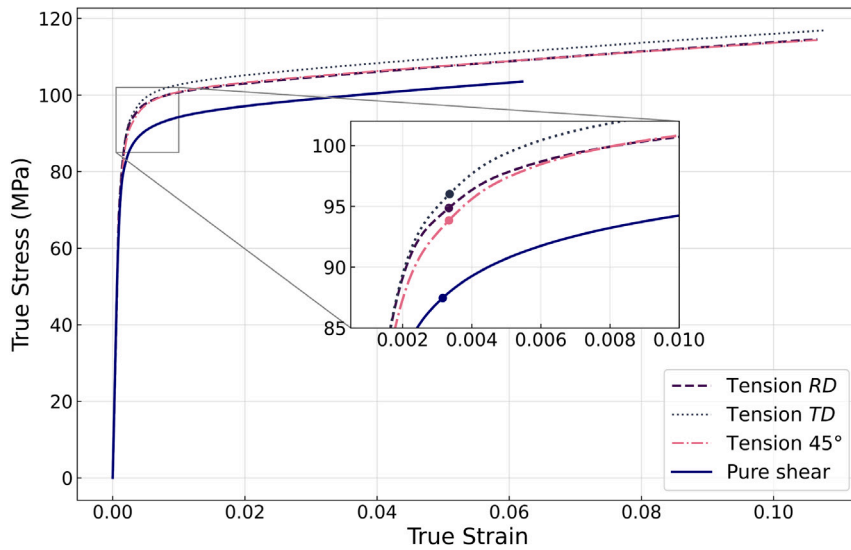


Fig. 3. Virtual stress-strain curves obtained from the CP model with identified macroscopic yield stress at 0.2% offset of the linear-elastic regime.

Fig. 4 presents the framework used herein. For the sake of simplicity, only the fundamental architectural parameters are optimized, namely the number of layers and the size of each layer. However, at the cost of having a high dimensional optimization space, the approach described here can effortlessly be extended to include additional hyperparameters such as learning rate, type of layer, class of activation function and number of epochs. The initial set of parameters used to initialize the NN model is defined by two fully connected layers with 250 neurons each. As illustrated in Fig. 4, each combination of layers with a certain size leads to a corresponding NN configuration, which in turn has its efficiency measured by the total loss on the test set. The architecture associated with the minimal loss is regarded as the optimal arrangement and used henceforth. The Bayesian optimization consists of three sub-steps: Initially a Gaussian Process surrogate of the loss function is built, following, an estimate of the best architecture to be tried next is calculated through the Expected Improvement acquisition function (Moćkus, 1975). This function measures the potential improvement upon the incumbent optimum of every point in the domain. Finally, the hyperparameter space is sampled once again at the point indicated by the acquisition function. It should be noted that this approach not only allows for an efficient

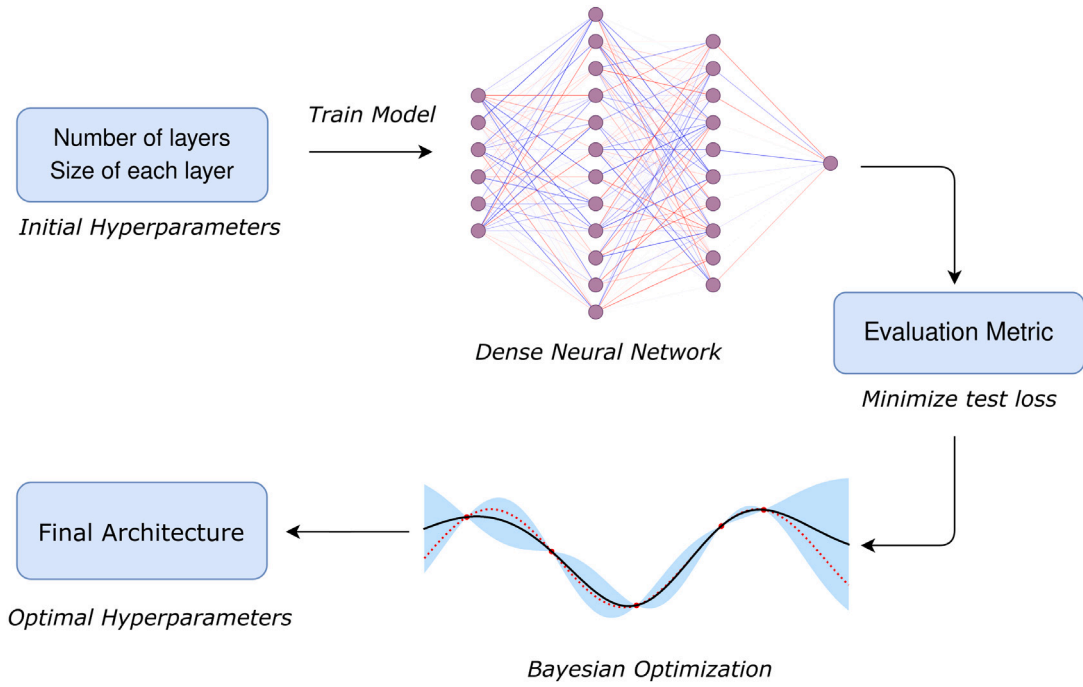


Fig. 4. Schematic workflow used in the determination of the optimal NN design for yield surface prediction.

NN design, but also substantially increases the level of automation of the MLVL. The final architecture is presented in Fig. 5 and discussed in Section 6.

5.3. Model training

The training process was intended to be similar to simulation based yield surface calibration methods, facilitating the incorporation of current and future approaches into the MLVL. In this context, random stress states at 0.2% offset of the linear-elastic regime obtained from the data set generated via virtual testing, as described in Section 5.1, were used to train the NN yield function and determine the anisotropy parameters of Yld2004-13p and Yld2004-18p.

The set of optimal coefficients for the advanced yield functions is obtained by the minimization of a quadratic loss function. Precisely, given a set of stress tensors σ_n at yield onset, the error associated with the anisotropy coefficients c'_{ij} and c''_{ij} , as defined in Eq. (8), Eq. (9) and Eq. (12), is expressed by:

$$E(c'_{ij}, c''_{ij} | \sigma_n) = \frac{1}{n} \sum_{k=1}^n \left(\frac{\phi(c'_{ij}, c''_{ij}, \sigma_k)}{\bar{\sigma}} - 1 \right)^2 \quad (19)$$

where $\phi(c'_{ij}, c''_{ij}, \sigma_k)$ denotes the equivalent stress estimated by the yield criterion. The error definition above explores the property that all stress-states belonging to the same yield locus lead to equal equivalent stresses (Aretz and Barlat, 2013), and uses the flow stress at uniaxial tension along RD ($\bar{\sigma}$) as the reference value. The minimization of Eq. (19) is performed by the Nelder–Mead or simplex search algorithm (Nelder and Mead, 1965), which has been widely used for parameter estimation and does not require derivative information providing, therefore, an efficient optimization scheme.

A custom loss function analogous to Eq. (19) is defined for the training of the NN yield function, i.e., for the determination of an optimal set of weights $\mathbf{w}$. Specifically, the error associated with the weight matrix $\mathbf{w}$, given the training set σ_n , is denoted by:

$$E(\mathbf{w} | \sigma_n) = \frac{1}{n} \sum_{k=1}^n \left(\frac{\phi(\mathbf{w}, \sigma_k)}{\bar{\sigma}} - 1 \right)^2 \quad (20)$$

with $\phi(\mathbf{w}, \sigma_k)$ indicating the equivalent stress obtained from the NN defined by the feed-forward formulation outlined in Section 4. The optimization of the weights applies the backpropagation algorithm (Rumelhart et al., 1988) over multiple epochs, or training cycles sweeping through the entire data set, allowing each data point to update the model parameters. The minimization of both Eq. (19) and Eq. (20) employed a total of 220 yield stress data points, which falls within the range of data set sizes employed in virtual yield surface fitting (Zhang et al., 2015, 2016).

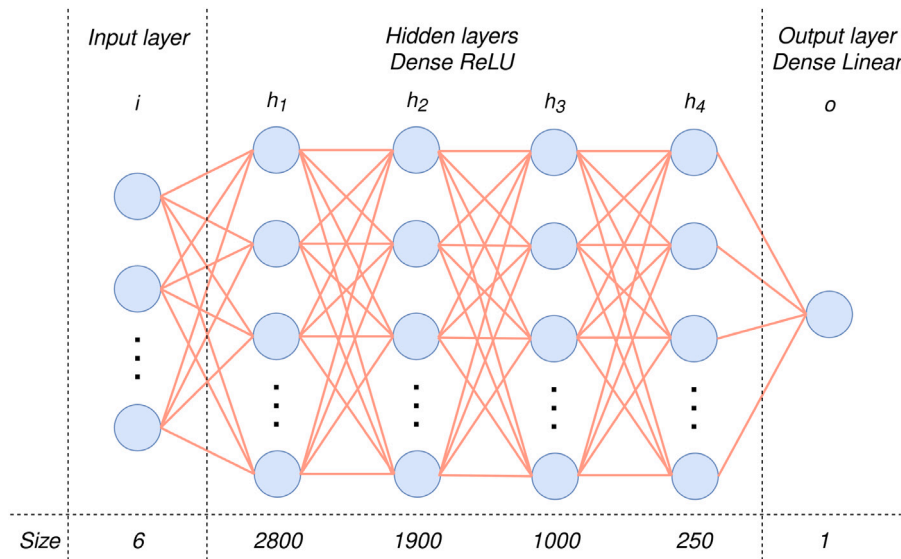


Fig. 5. Optimal architecture of the NN yield function obtained via Bayesian optimization featuring the six dimensional stress input, scalar equivalent stress output and four fully connected dense hidden layers.

6. Results

6.1. Design of the NN architecture

Following the approach described in Section 5, the architecture of the NN yield function was initially obtained via Bayesian optimization. Precisely, the optimal configuration was defined by the number of layers and the size of each layer such that the test loss was minimal. A total of 150 different configurations were used in the optimization process. The optimization was performed on a machine powered by an 11th Gen Intel Core i9-11900H clocked at 2.50 GHz and 16 processors with computational overhead limited as much as possible. Under these conditions the average computational time for each evaluation was 28.58s while the total wall-clock time of the optimization step was 1 h 11 min 27 s. Each evaluation includes all the steps defined in Sections 5.2 and 5.3. This is a marginal computational cost when compared to the cost required to obtain representative datasets from CP simulations (Eisenlohr et al., 2013). Fully connected hidden dense layers were used in conjunction with ReLU activation functions in order to furnish advanced fitting capability and computational efficiency, as highlighted in Section 4. The architecture is shown schematically in Fig. 5, outlining that four hidden layers with a relatively large number of neurons provided the best NN arrangement.

6.2. Yield surface models

Equipped with a well defined NN architecture, the initial calibration of Yld2003-13p, Yld2004-18p and the training of the NN model was performed with data obtained via random sampling of the stress space. Specifically, 220 yield data points were used, from which 110 describe plane stress states and the remainder characterize out-of-plane stress states. Such emphasis on plane stresses is justified due to their relevance for sheet metal forming (Barlat et al., 2003). An additional set of 220 random data points was set apart as testing batch allowing the evaluation of the NN on unseen data. It should be emphasized that this last step is not a prerequisite for accurate calibration, but rather provides an assessment framework for the current study.

The accuracy of the trained yield models has been associated with the representativeness of the training data (Zhang et al., 2022); thus, in order to estimate the average performance of the proposed NN yield function, multiple calibration rounds were observed. Specifically, 100 training cycles, each employing 1000 epochs with different randomized training and testing data. The history of the NN loss at each cycle along with the average loss curve, is presented in Fig. 6, where it can be seen that both the train and test loss indicate an accurate NN yield function with high representational capacity. Additionally, it should be noted that despite the lower training loss, there are no signs of overfitting, which is often expressed by a clear uptick on the test loss over the number of epochs. Further evidence of learning and generalization is provided by the adequate description of the yield locus, under multiaxial and uniaxial stress states, as will be highlighted in Section 6.3 and Section 6.5, respectively. Finally, it is also noteworthy the onset of a plateau on the test error approximately at 600 epochs, which indicates a margin for reducing the number of epochs and consequently increase the computational efficiency of the NN model training step.

The final test loss of the NN model in each cycle is compared against the corresponding test error of Yld2004-13p and Yld2004-18p. The results are presented in Fig. 7, where it can be observed that in average the NN yield function has higher accuracy, albeit by a small margin, when contrasted with both Yld2004-13p and Yld2004-18p. The NN, however, shows a wider scatter

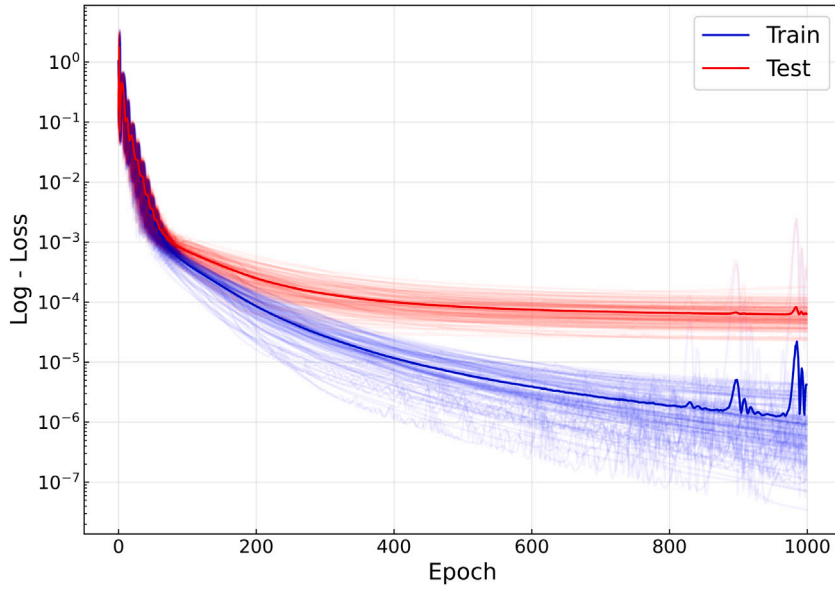


Fig. 6. History of the NN yield function training and testing losses on each calibration cycle and average loss history over 100 cycles.

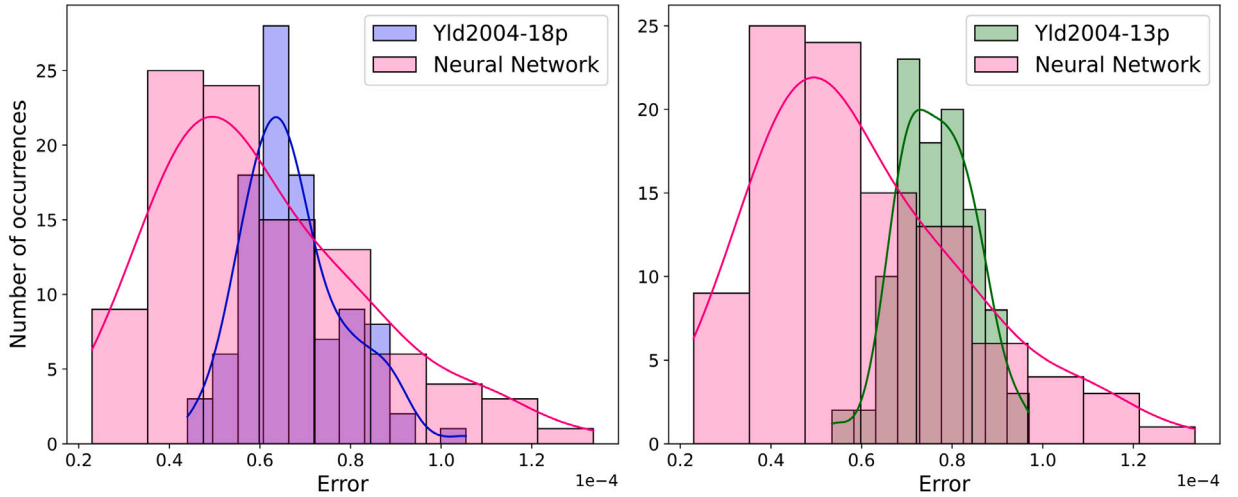


Fig. 7. Comparison between the test loss of the 3D yield functions Yld2004-18p and Yld2004-13p (Barlat et al., 2005) against the NN yield model over different randomized calibration cycles.

error range which could potentially be associated with a stronger dependency on the representativeness of the training data. Tangentially, it is worth noting the narrow difference between the precision of Yld2004-13p and Yld2004-18p, despite the latter having five additional anisotropy parameters initially meant to provide more flexibility and precision. This could be associated with the observed correlation between different parameters (van den Boogaard et al., 2016), outlining that adding further coefficients does not necessarily lead to substantial accuracy improvement. The most notable observation from Fig. 7, however, is the comparable performance of the NN flow criterion to 3D yield functions, despite the former being trained with the same number of data points and its material blind initial state.

In order to further analyze the predictive capacity of the NN yield function, an arbitrary training cycle was chosen as reference and compared against the calibrated Yld2004-13p and Yld2004-18p. The identified parameters for the phenomenological yield functions are presented in Tables 2 and 3. Here, the initial guess for the parameters was set to their respective isotropic values, i.e., $c'_{ij} = c''_{ij} = 1$ and $m = 2$ (or 4) (Zhang et al., 2015), reducing both Yld2004-13p and Yld2004-18p to Hershey's isotropic yield

Table 2

Yld2004-13p (Barlat et al., 2005) anisotropy parameters identified for the reference aluminum alloy.

Yld2014-13p	c'_{12}	c'_{13}	c'_{21}	c'_{23}	c'_{44}	c'_{55}	c'_{66}	
	–	1.0077	0.9614	0.9853	1.0095	0.6098	1.1894	
	c''_{12}	c''_{13}	c''_{21}	c''_{23}	c''_{44}	c''_{55}	c''_{66}	m
	0.9344	1.0874	1.2967	1.0700	1.2917	0.4643	1.0219	5.8630

Table 3

Yld2004-18p (Barlat et al., 2005) anisotropy parameters identified for the reference aluminum alloy.

Yld2014-18p	c'_{12}	c'_{13}	c'_{21}	c'_{23}	c'_{31}	c'_{32}	c'_{44}	c'_{55}	c'_{66}	
	1.1687	1.0202	0.9501	1.0697	0.8373	0.7931	0.6482	1.1994	0.8183	
	c''_{12}	c''_{13}	c''_{21}	c''_{23}	c''_{31}	c''_{32}	c''_{44}	c''_{55}	c''_{66}	m
	0.9102	1.1648	0.8921	1.1305	0.9878	0.9531	0.5625	0.5625	1.2065	5.9255

function (Hershey, 1954). Additionally, in order to allow for greater flexibility, the exponent m is regarded as an unknown parameter instead of set to its usual values, namely $m = 8$ and $m = 6$ for FCC and BCC metals respectively (Logan and Hosford, 1980). The results and discussion that follows are based on these set of parameters and the corresponding NN function obtained in the same calibration cycle.

6.3. Multiaxial yield surfaces

Figs. 8 and 9 show the multiaxial yield contours of Yld2004-18p and Yld2004-13p contrasted with the respective yield contours of the NN yield function in isolines of normalized shear in increments of approximately 0.08 from 0 in the RD-TD plane. The yield onset is assumed at 0.2% of strain offset from the linear-elastic regime and $\bar{\sigma}$ denotes the uniaxial tensile flow stress along RD. The good agreement between the NN yield function and both benchmark yield criterion can be observed. For the material under consideration, exhibiting mild anisotropy, as suggested by Fig. 2, the radius around the equibiaxial tension axis is relatively large, leading to a smooth curvature, which was accurately predicted by the three yield functions. This is consistent with the relatively low values for the exponent m obtained in Tables 2 and 3, usually associated with random or weak textures. In contrast, higher m values often describe pronounced anisotropy and lead to angular shaped yield loci, commonly observed for strong cube textures with narrow scatter (Wu et al., 2004; Zhang et al., 2019).

An additional observation from Figs. 8 and 9, is the slight decrease of agreement between the compared yield functions as the normalized shear $\sigma_{12}/\bar{\sigma}$ increases. A larger error at higher shear contours was also reported by Zhang et al. (2022), while analyzing

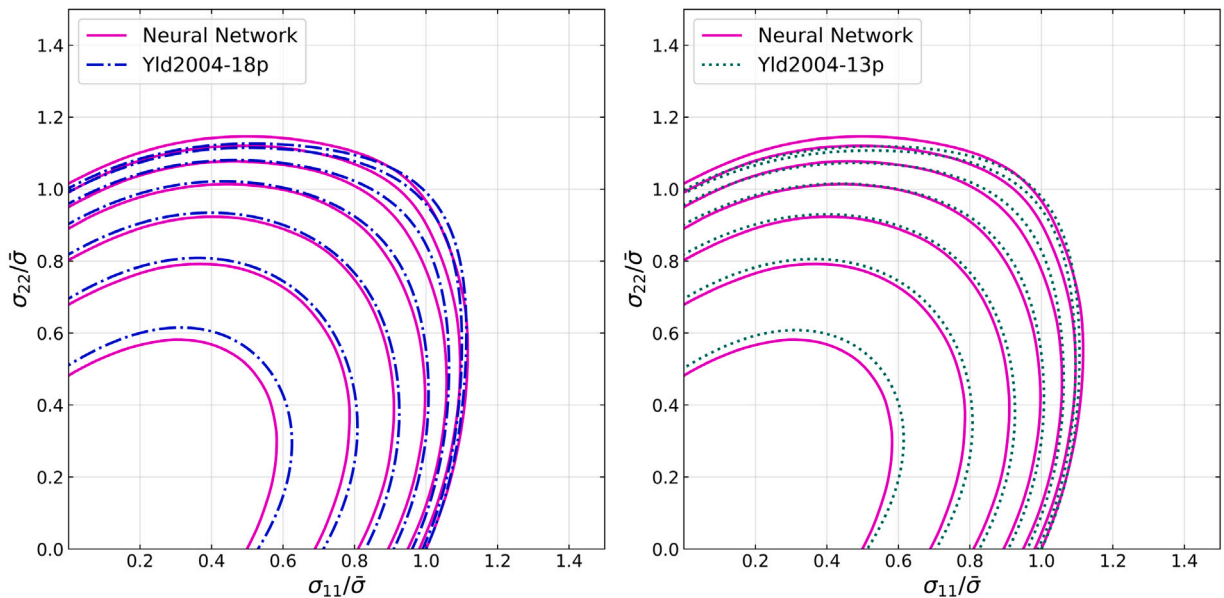


Fig. 8. Multiaxial yield surface of Yld2004-18p vs NN (left) and Yld2004-13p vs NN (right) in the tension tricomponent stress subspace $\{\sigma_{11} - \sigma_{22} - \sigma_{12}\}$ at isolines of normalized σ_{12} in increments of approximately 0.08 from 0 for the aluminum alloy described in Section 5.1.

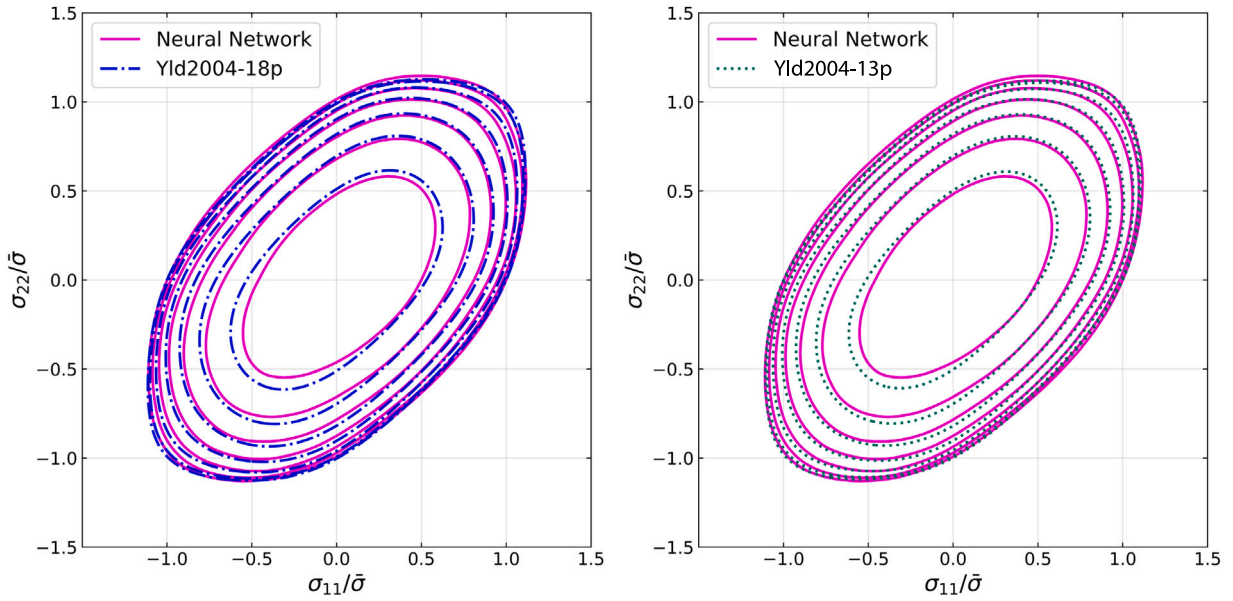


Fig. 9. Multiaxial yield surface of Yld2004-18p vs NN (left) and Yld2004-13p vs NN (right) in the tricomponent stress subspace $\{\sigma_{11} - \sigma_{22} - \sigma_{12}\}$ at isolines of normalized σ_{12} in increments of approximately 0.08 from 0 for the aluminum alloy described in Section 5.1.

the Yld2000-2D tricomponent plane stress yield function (Barlat et al., 2003). In this context, the sharper radius of curvature and therefore stronger gradients at elevated shear when compared to the biaxial plane, makes this region particularly difficult to be accurately represented since a greater flexibility is required from the yield function and might lead to higher error. This phenomenon could be partially mitigated by increasing the number of data points with a more pronounced shear component, effectively providing a more thorough sampling of this region of the stress-space, or alternatively assigning higher weight to these data points.

Fig. 10 displays the biaxial yield contour and the contour at a normalized shear height of 0.3 of the calibrated NN with marked deformation states and their respective symmetry lines. The clear alignment between the tension-compression pairs with the independently drawn symmetry lines in both cases indicates the NN's capacity to learn symmetry properties. Minor violations of the tension-compression symmetry, however, can be observed in Fig. 9 at the highest shear contour. This is consistent with the lower accuracy on the yield prediction at high shear levels, and further underscores the need of representative data in this region of the stress space.

6.4. Yield surface convexity

Convexity is an additional requirement for a stable yield function. In this context, an initial visual inspection on the yield contours of Figs. 8, 9 and 10 indicates that the convexity was captured by the NN yield function. Additionally, a further sign of a convex NN yield locus is provided by the flow vector, obtained as the derivative of the equivalent stress with respect to the CAUCHY stress, $\partial\phi/\partial\sigma$. Fig. 11 shows the flow vector of the three functions calculated at the biaxial yield contour, where its outwards normal character can be observed, and reiterates the convexity of the yield hull. Incidentally, it should be noted that despite of the black-box character of the NN, the stress gradient can be easily calculated via automatic differentiation at any given stress state lying on the yield locus.

A more disciplined assessment of the convexity of the NN yield locus, however, can be done applying Drucker's postulate (Drucker, 1950). This principle ensures the convexity of the yield surface and the local normality of the plastic strain increment (Lubarda et al., 1996). In its most simple form, the principle of maximum plastic work is mathematically stated as $(\sigma - \sigma^0) : d\epsilon^P \geq 0$. If an associated flow rule is assumed and the arbitrary stress state σ^0 set to $\mathbf{0}$, this criterion can be expressed as $\sigma : n \geq 0$, where n denotes the flow vector obtained from the yield surface gradient. In fact, this postulate provides a necessary and sufficient condition for positive plastic dissipation and a unique relationship between stress and strain increments (Barlat et al., 2005) with the latter being an important aspect for stable integration of the elastoplastic boundary value problems (Stoughton and Yoon, 2006). Here, Drucker's postulate will be used as an indicator of the convexity of the yield locus determined by the NN, following the approach presented by Vlassis and Sun (2020).

Fig. 12 shows the results obtained by applying the principle of maximum plastic work over 2000 stress data points lying on the yield locus. It demonstrates that all contractions lead to positive values, with the minimum of approximately 0.56. Therefore, Drucker's criterion was not violated in any of the instances observed. While this approach provides a sound indicator of convexity, it should be noted that it does not constitute a formal proof.

It can be observed from Fig. 11 not only the normal character of the flow vector, but also the misalignment between the corresponding flow directions predicted by the NN yield function and the benchmark yield criteria. Here, the L_2 norm of the

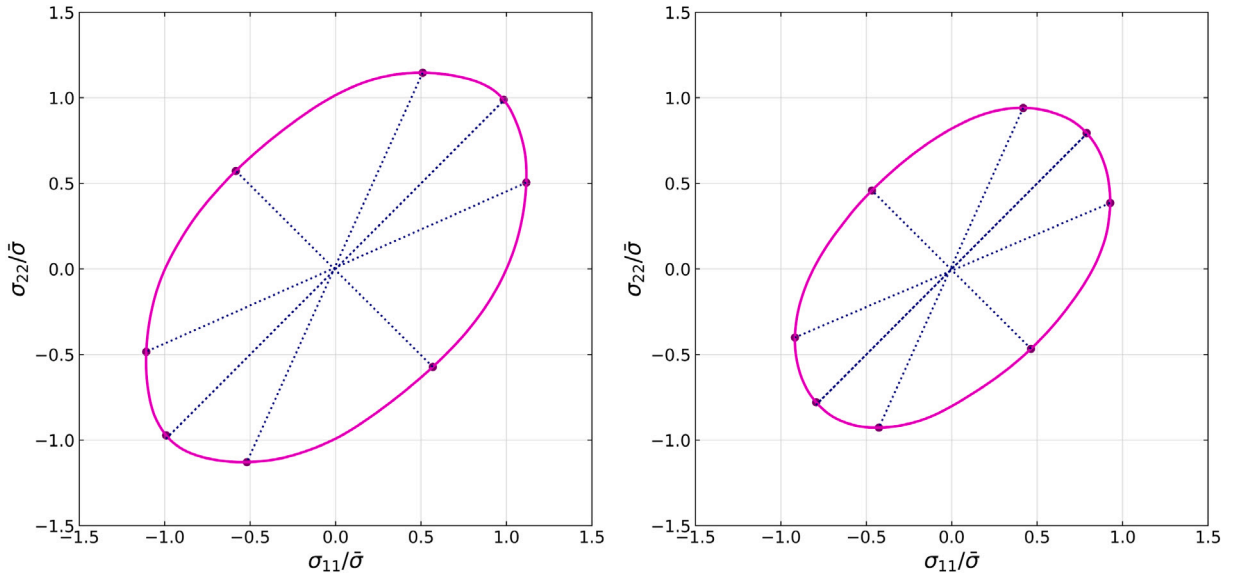


Fig. 10. Yield contours at the biaxial plane (left) and at 0.3 normalized σ_{12} (right) of the NN yield function with connecting symmetry lines highlighting the tension-compression symmetry captured by the NN.

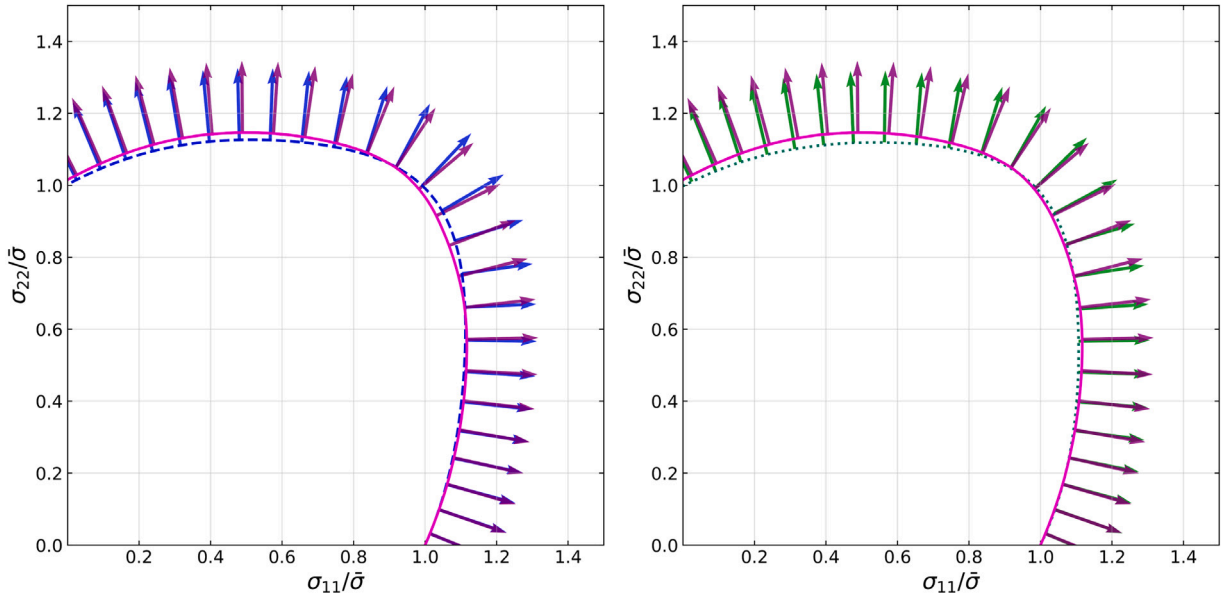


Fig. 11. Biaxial yield contour of the NN vs Yld2004-18p (left) and NN vs Yld2004-13p (right) with their respective flow vectors obtained through automatic differentiation.

difference between the normalized stress gradients calculated by the NN, Yld2004-13p and Yld2004-18p is used to estimate the discrepancy between the determined plastic strain rates. Fig. 13 shows the results obtained for over 2000 flow vectors calculated at arbitrary points on the yield surface. The mean of the difference norm lies around 0.2, which is equivalent to an misalignment angle of approximately 11° in the 6 dimensional symmetric stress space. This result is commensurate, albeit marginally higher, than the results reported by Vlassis and Sun (2020) while studying the plastic flow difference between distinct NNs estimates in the principal stress space. Additionally, the similar histograms in Fig. 13 indicate that while there is a meaningful contrast between the flow vector obtained by the NN and the advanced yield functions, the plastic flow predicted by both Yld2004-13p and Yld2004-18p is nearly identical. It is important to highlight that the stress gradient is a derived field, for which the NN's weights have not been optimized and yet a reasonably good prediction is achieved.

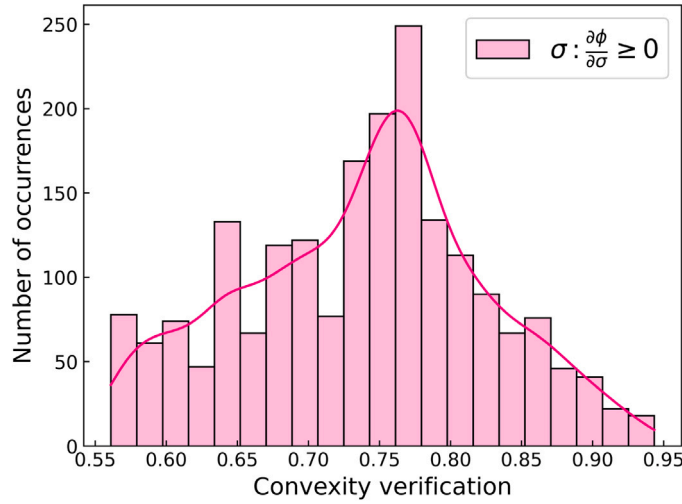


Fig. 12. Verification of Drucker's condition over 2000 stress data points lying on the yield surface obtained the NN flow criterion.

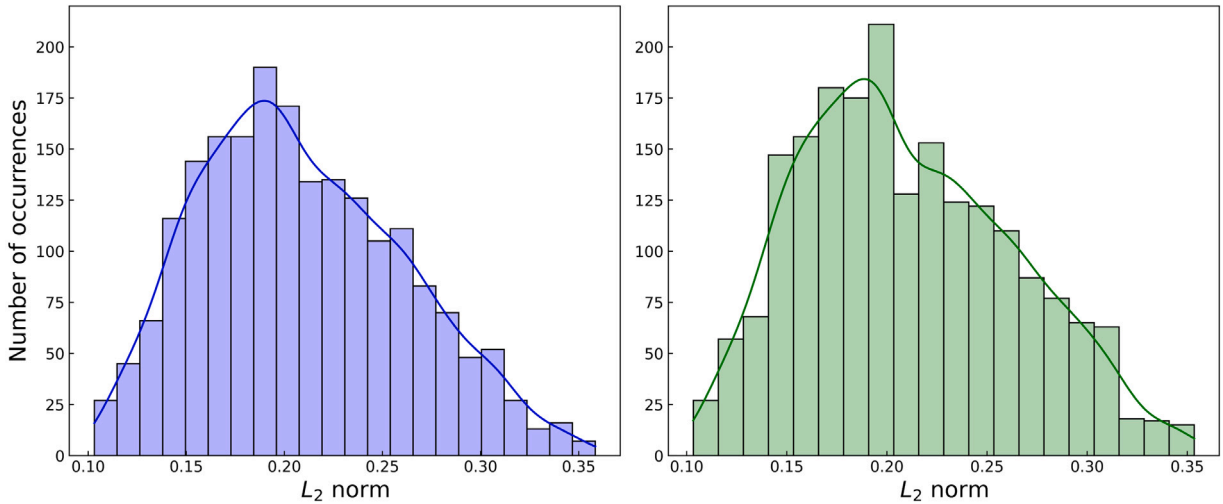


Fig. 13. L_2 norm of the difference between the normalized flow vectors predicted by the NN vs Yld2004-18p (left) and NN vs Yld2004-13p for 2000 stress data points.

6.5. Planar yield response

The strength variation is presented in Fig. 14. Here, the normalized yield stress is defined as the ratio between the directional flow stress and the tensile uniaxial yield stress along RD. The former was calculated at every 15° from RD to TD. The limited strength variation shown in Fig. 14 outlines a virtually planar isotropic yield response. This is consistent with the weak texture displayed in Fig. 2 and the results obtained from the CP simulations. Despite an offset of about 1% from the simulation data, the three yield criteria captured the overall trend, characterized by a diminishing strength up to 45° from RD, followed by an increase until 90° . It should be emphasized that in view of the narrow variation range observed in Fig. 14, the three yield functions show a good agreement when predicting uniaxial yield behavior. We can also notice from Fig. 14 that the NN accuracy is better or worse when compared to Yld2004-18p depending on the loading direction; however, its performance is overall better as indicated by Fig. 7. A precise estimate of elementary uniaxial flow properties is often regarded as a required feature if reasonable prediction of complex stress states is to be expected (Barlat et al., 2005). In this context, the capacity of the NN yield function to reproduce the strength variation described by advanced yield surfaces should be underscored. This particularly applies since the random sample technique adopted here does not allocate higher weight to uniaxial stress states, as it is commonly done in conventional calibration methods (Zhang et al., 2015).

The plastic anisotropy, measured by means of the r-values, is presented in Fig. 15. Here, the Lankford coefficients are determined at every 15° from RD following the formulation presented by Stewart and Cazacu (2011). The r-values vary in the approximate range

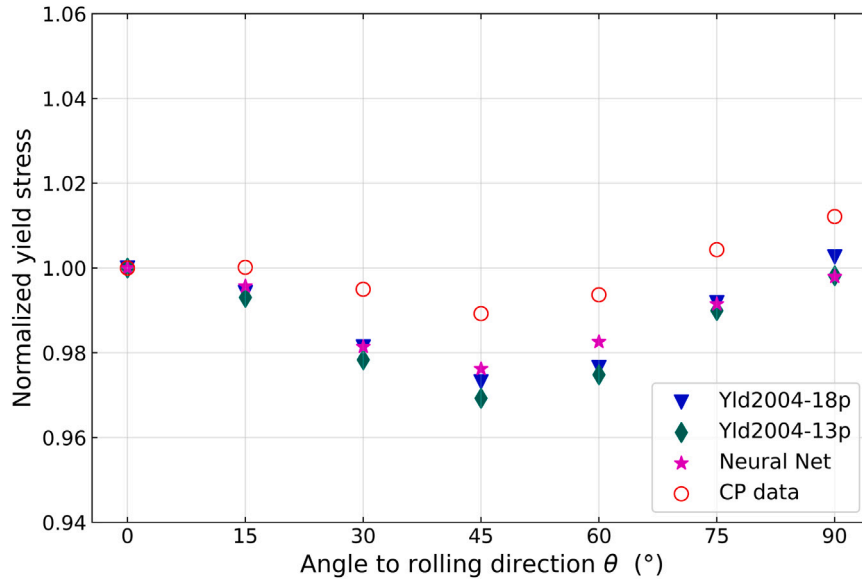


Fig. 14. Normalized yield stress versus θ from RD at 0.2% of strain offset of the linear-elastic regime predicted by the 3D yield surfaces Yld2004-13p, Yld2004-18p (Barlat et al., 2005), the NN yield function and CP simulations.

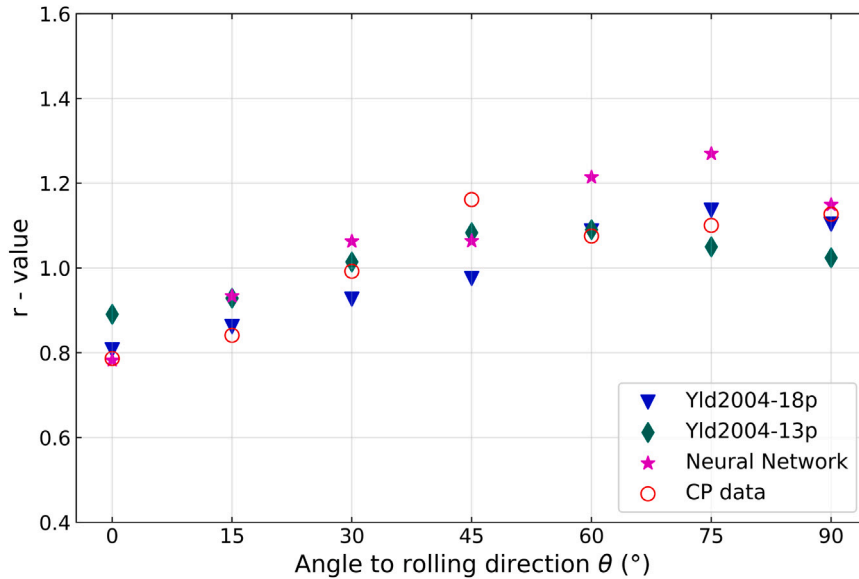


Fig. 15. R-value versus θ from RD at 0.2% of strain offset of the linear-elastic regime predicted by the 3D yield surfaces Yld2004-13p, Yld2004-18p (Barlat et al., 2005), the NN yield function and CP simulations.

of 0.8–1.2, *i.e.*, about 20% deviation from unity. Such broader variation indicates a more pronounced planar deformation anisotropy when compared with the planar strength response. While the initial rising trend on the r-values, up to 45°, was well captured by the three yield functions, a modest difference from the CP data can be observed. It should be highlighted that these results are despite the fact that r-values were neither used in the training of the NN nor in the calibration of Yld2004-18p and Yld2004-13p. We emphasize that a greater discrepancy between the calculated r-values can also be observed among the benchmark yield functions themselves, in particular when compared with the strength anisotropy shown in Fig. 14. Indeed, a relatively high sensitivity of the r-values with respect to the stress gradient was detected, which is aligned with a similar observation reported by Zhang et al. (2016). Finally, while this sensitivity could introduce spurious variations and partially reduce the prediction quality, the overall accuracy of the r-value calculations by the NN yield function are commensurate with the precision of advanced yield functions and is expected to improve once r-values are incorporated in the training in future studies.

7. Discussion

The machine learning virtual laboratory framework proposed here aims to be resilient and accommodate different degrees of anisotropy and material properties, mostly due to the optimal design or the NN yield function, which is sensitive to the microstructural information provided. In order to showcase its application, however, an aluminum alloy, as described in Section 5.1 was used, and the corresponding results were presented in Section 6.

An immediate observation from Fig. 5 is the relatively high number of neurons and the depth of the NN model. As outlined in Section 6, the optimal hyperparameters are automatically obtained via Bayesian optimization. Such large architecture could be associated with the greater complexity of the problem delegated to the NN when compared with the one assigned to Yld2004-18p and Yld2004-13p. In fact, the last two, albeit phenomenological, incorporate into their analytical model pronounced features of the class of material under analysis, such as convexity of the yield hull and tension–compression symmetry (Barlat et al., 2005). The NN, however, does not possess these inbuilt attributes at the outset and must learn such underlying physical properties in addition to provide an accurate fitting. Furthermore, in favor of practical usage of the model, the number of data points necessary in the training has been kept near the size of data sets used in simulation based calibration of the reference yield criteria. This is a specially demanding requirement given the above-mentioned material blind initial state of the NN, and additionally underscores the need of a robust architecture with high representational capacity, which in turn has been directly linked to the number and size of the layers comprising a NN model (Goodfellow et al., 2016).

The NN yield function obtained through the optimized architecture provides an accurate yield description as evidenced by Fig. 7. Here, the accuracy of the NN is not measured by a hard threshold, but rather, by comparison with the corresponding loss of the Yld2004-18p and Yld2004-13p yield functions calculated over the test data. Given the accuracy of the benchmark yield functions (Yoon et al., 2006) and considering that the mean loss of the NN on the test data was lower than the mean loss of both benchmark functions, we believe that MLVL framework and the NN model have been proven accurate. Additionally, the shape of the initial yield surface is adequately captured by the NN model, as indicated by Figs. 8 and 9, where we can observe a strong agreement between the contours at different shear levels. The capacity to learn the shape of the yield surface is directly linked to the accurate description of the initial anisotropy. Furthermore, while the evolution of the yield surface is not analyzed here, it should be emphasized that the same framework applies to subsequent yield surfaces with the only difference being the level of plastic strain at which the stress data points are obtained.

One relevant underlying property of the flow behavior under consideration concerns tension–compression symmetry of the yield locus. Specific classes of material symmetry can be incorporated into the yield functions of the Yld family by means of the linear transformation matrices presented in Eq. (8) (Barlat et al., 2003). Meanwhile, the NN was able to learn the symmetry of the yield surface directly from the training data. In fact, the ability to learn tension–compression symmetry, to an adequate degree and without any prior knowledge of the material under analysis further underscores the robustness of the optimally designed NN. It also suggests the potential to learn different symmetry properties, or the lack thereof. This could be particularly useful in the description of macroscopic plasticity models for HCP polycrystals, which are currently less developed than their FCC and BCC counterparts (Stewart and Cazacu, 2011; Li et al., 2016) due to the difficulty in modeling the pronounced strength differential effects as a result of the directionality of twinning (Tomé et al., 2001), leading to highly non-symmetrical yield loci.

The convexity of the NN yield function was tested here, but not by a strict method. Precisely, as presented in Section 6.4, the visual inspection of the yield contours, the normality of the stress gradients and the application of Drucker's postulate, furnish a strong indication that the obtained yield hull is convex. A recently proposed and more sophisticated algorithm for evaluating the convexity of black-box functions was outlined by Tamura and Gallagher (2019) and could be employed in case a formal proof is required. The convexity of Yld2004-13p and Yld2004-18p, on the other hand, is ensured by the anisotropy exponent m , in particular, both functions will provide a convex yield hull if m is greater than approximately 1.7 (Barlat et al., 2005). Furthermore, it is noteworthy that automatic differentiation (Bartholomew-Biggs et al., 2000) was used to calculate the derivative of the black-box NN yield function as well as the derivatives of the benchmark yield criteria (Güneş Baydin et al., 2015). Here, the capacity to compute the flow vector as a tensorial gradient of a NN should be emphasized, since in addition to assist in the convexity analysis, it also provides a viable pathway to integrate the current model into incremental associated flow rules with substantially reduced effort.

As presented in Section 6.5, the prediction of planar properties such as strength and deformation anisotropy with overall good accuracy despite uniaxial stress states and r -values not being explicitly used in the training is a further indication of the representative capacity of the NN yield function. In fact, incorporating r -values into the training could be a possible avenue to augment the accuracy of calculated flow vectors, since they are intrinsically associated with plastic strain rates. Indeed, multiple yield function calibration methods include uniaxial r -values in addition to stress data (Han et al., 2020). It is noteworthy that increasing the dataset size could also improve the NN accuracy. However, one of the goals and highlights of this work is to achieve an accurate yield description with minimum data set sizes. Finally, the NN architecture optimization presented here focused on two hyperparameters, namely the number of layers and the size for each layer to keep the optimization space small. There are, however, additional hyperparameters that could be included in the optimization framework, such as number of epochs, learning rate, and layer type, that could potentially improve the accuracy of the NN yield function.

8. Conclusions

We presented an approach to obtain a fully data driven yield model suited for sheet metal forming. This model was incorporated into a fully simulation based framework and integrates microstructurally informed crystal plasticity (CP) constitutive models into the testing procedure. The optimal architecture of the neural network (NN) yield function is determined via Bayesian optimization and both the NN model and the Yld2004-18p and Yld2004-13p yield criteria are trained with randomly sampled stress data points. The resulting yield functions are used in the study of the multiaxial yield response, planar strength and deformation anisotropy of a reference aluminum alloy, and the capacity of the NN flow criterion to assimilate fundamental material features such as convexity and tension-compression symmetry was also considered.

The performance of the proposed model is commensurate with that of advanced phenomenological yield functions in the prediction of complex multiaxial stress states and planar anisotropic properties, such as normalized yield stresses and r -values. Additionally, the NN yield function was able to reproduce, to an adequate degree, the convexity of the yield hull as well as the tension-compression symmetry of the material under consideration. Furthermore, despite of the black-box nature of the NN, stress gradients can be conveniently calculated through automatic differentiation, providing a pathway for seamless incorporation into associated flow rules and conventional finite element codes. Combined, these observations render a promising outlook in the model proposed herein. Possible extensions of the model include the analysis of the yield surface evolution as well as the study of materials exhibiting a more involved constitutive response, such as pronounced anisotropy and strength differential behavior.

CRediT authorship contribution statement

Anderson Nascimento: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Sharan Roongta:** Methodology, Software, Visualization, Writing – review & editing. **Martin Diehl:** Methodology, Software, Visualization, Writing – review & editing. **Irene J. Beyerlein:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

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.

Data availability

Data will be made available on request.

Acknowledgments

Anderson Nascimento and Irene J. Beyerlein are grateful for the support by the National Science Foundation, United States under Grant Number 2051390. Use was made of computational facilities purchased with funds from the National Science Foundation, United States (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute, United States and the Materials Research Science and Engineering Center, United States (MRSEC; NSF DMR 1720256) at UC Santa Barbara. Sharan Roongta would like to acknowledge the support of Materials innovation institute (M2i), Netherlands (www.m2i.nl) supported by the Dutch government under the project number 17019e. Martin Diehl acknowledges funding through the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - DI 1873/4-1 and through Internal Funds KU Leuven, Belgium.

References

- Abedini, A., Butcher, C., Rahmaan, T., Worswick, M.J., 2018. Evaluation and calibration of anisotropic yield criteria in shear Loading: Constraints to eliminate numerical artefacts. *Int. J. Solids Struct.* 151, 118–134. <http://dx.doi.org/10.1016/J.IJSOLSTR.2017.06.029>.
- Achani, D., Hopperstad, O.S., Lademo, O.G., 2009. Behaviour of extruded aluminium alloys under proportional and non-proportional strain paths. *J. Mater. Process. Technol.* 209 (10), 4750–4764. <http://dx.doi.org/10.1016/J.JMATPROTEC.2008.12.002>.
- Aghasafari, P., Abdi, H., Salimi, M., 2014. Artificial neural network modeling of flow stress in hot rolling. *ISIJ Int.* 54 (4), 872–879. <http://dx.doi.org/10.2355/ISIJINTERNATIONAL.54.872>.
- Ali, U., Muhammad, W., Brahme, A., Skiba, O., Inal, K., 2019. Application of artificial neural networks in micromechanics for polycrystalline metals. *Int. J. Plast.* 120, 205–219. <http://dx.doi.org/10.1016/J.IJPLAS.2019.05.001>.
- Aretz, H., Aegerter, J., Engler, O., 2010. Analysis of earing in deep drawn cups. In: *AIP Conference Proceedings*. 1252, (1), American Institute of Physics, pp. 417–424. <http://dx.doi.org/10.1063/1.3457585>.
- Aretz, H., Barlat, F., 2013. New convex yield functions for orthotropic metal plasticity. *Int. J. Non-Linear Mech.* 51, 97–111. <http://dx.doi.org/10.1016/J.IJNONLINEMEC.2012.12.007>.
- Banabic, D., Comsa, D.S., Balan, T., 2000. A new yield criterion for orthotropic sheet metals under plane stress conditions. In: *7th TPR 2000, Cold Metal Forming Conf.*. Printek Publishing House, pp. 217–224.
- Barlat, F., Aretz, H., Yoon, J.W., Karabin, M.E., Brem, J.C., Dick, R.E., 2005. Linear transformation-based anisotropic yield functions. *Int. J. Plast.* 21 (5), 1009–1039. <http://dx.doi.org/10.1016/j.ijplas.2004.06.004>.

- Barlat, F., Brem, J.C., Yoon, J.W., Chung, K., Dick, R.E., Lege, D.J., Pourboghrat, F., Choi, S.H., Chu, E., 2003. Plane stress yield function for aluminum alloy sheets—part 1: theory. *Int. J. Plast.* 19 (9), 1297–1319. [http://dx.doi.org/10.1016/S0749-6419\(02\)00019-0](http://dx.doi.org/10.1016/S0749-6419(02)00019-0).
- Barlat, F., Lege, D.J., Brem, J.C., 1991. A six-component yield function for anisotropic materials. *Int. J. Plast.* 7 (7), 693–712. [http://dx.doi.org/10.1016/0749-6419\(91\)90052-Z](http://dx.doi.org/10.1016/0749-6419(91)90052-Z).
- Barlat, F., Yoon, J.W., Cazacu, O., 2007. On linear transformations of stress tensors for the description of plastic anisotropy. *Int. J. Plast.* 23 (5), 876–896. <http://dx.doi.org/10.1016/J.IJPLAS.2006.10.001>.
- Bartholomew-Biggs, M., Brown, S., Christianson, B., Dixon, L., 2000. Automatic differentiation of algorithms. *J. Comput. Appl. Math.* 124 (1–2), 171–190. [http://dx.doi.org/10.1016/S0377-0427\(00\)00422-2](http://dx.doi.org/10.1016/S0377-0427(00)00422-2).
- Birattari, M., Yuan, Z., Balaprakash, P., Stützle, T., 2010. F-race and iterated F-race: An overview. *Experimental Methods for the Analysis of Optimization Algorithms* 311–336. http://dx.doi.org/10.1007/978-3-642-02538-9_13.
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Springer, New York.
- Bonatti, C., Berisha, B., Mohr, D., 2022. From CP-FFT to CP-RNN: Recurrent neural network surrogate model of crystal plasticity. *Int. J. Plast.* 158, 103430. <http://dx.doi.org/10.1016/J.IJPLAS.2022.103430>.
- Bonatti, C., Mohr, D., 2021. Neural network model predicting forming limits for Bi-linear strain paths. *Int. J. Plast.* 137, 102886. <http://dx.doi.org/10.1016/J.IJPLAS.2020.102886>.
- van den Boogaard, T., Havinga, J., Belin, A., Barlat, F., 2016. Parameter reduction for the Yld2004-18p yield criterion. *Int. J. Mater. Form.* 9 (2), 175–178. <http://dx.doi.org/10.1007/s12289-015-1221-3>.
- Cazacu, O., 2020. New expressions and calibration strategies for Karafillis and Boyce (1993) yield criterion. *Int. J. Solids Struct.* 185–186, 410–422. <http://dx.doi.org/10.1016/J.IJUSOLSTR.2019.09.004>.
- Claesen, M., De Moor, B., 2015. Hyperparameter Search in Machine Learning. In: *The XI Metaheuristics International Conference*. <http://dx.doi.org/10.48550/arxiv.1502.02127>.
- Clark, S., Hayes, P., 2019. SigOpt web page. URL <https://sigopt.com>.
- Drucker, D.C., 1950. Some implications of work hardening and ideal plasticity. *Quart. Appl. Math.* 7 (4), 411–418. <http://dx.doi.org/10.1090/QAM/34210>.
- Du, K., Huang, S., Li, X., Wang, H., Zheng, W., Yuan, X., 2022. Evolution of yield behavior for AA6016-T4 and DP490—Towards a systematic evaluation strategy for material models. *Int. J. Plast.* 154, 103302. <http://dx.doi.org/10.1016/J.IJPLAS.2022.103302>.
- Eggensperger, K., Feurer, M., Hutter, F., Bergstra, J., Snoek, J., Hoos, H.H., Leyton-Brown, K., 2013. Towards an empirical foundation for assessing Bayesian optimization of hyperparameters. In: *NIPS Workshop on Bayesian Optimization in Theory and Practice*.
- Eisenlohr, P., Diehl, M., Lebensohn, R.A., Roters, F., 2013. A spectral method solution to crystal elasto-viscoplasticity at finite strains. *Int. J. Plast.* 46, 37–53. <http://dx.doi.org/10.1016/J.IJPLAS.2012.09.012>.
- Eisenlohr, P., Roters, F., 2008. Selecting a set of discrete orientations for accurate texture reconstruction. *Comput. Mater. Sci.* 42 (4), 670–678. <http://dx.doi.org/10.1016/j.commatsci.2007.09.015>.
- Fuhg, J.N., van Wees, L., Obstalecki, M., Shade, P., Bouklas, N., Kasemer, M., 2022. Machine-learning convex and texture-dependent macroscopic yield from crystal plasticity simulations. *Materialia* 23, 101446. <http://dx.doi.org/10.1016/J.MTLA.2022.101446>.
- Goodfellow, I.J., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press, URL <http://www.deeplearningbook.org>.
- Grytten, F., Holmedal, B., Hopperstad, O.S., Børvik, T., 2008. Evaluation of identification methods for YLD2004-18p. *Int. J. Plast.* 24 (12), 2248–2277. <http://dx.doi.org/10.1016/J.IJPLAS.2007.11.005>.
- Güneş Baydin, A., Pearlmutter, B.A., Andreyevich Radul, A., Mark Siskind, J., 2015. Automatic differentiation in machine learning: a survey. *J. Mach. Learn. Res.* 18, 1–43. <http://dx.doi.org/10.48550/arxiv.1502.05767>.
- Han, F., Diehl, M., Roters, F., Raabe, D., 2020. Using spectral-based representative volume element crystal plasticity simulations to predict yield surface evolution during large scale forming simulations. *J. Mater. Process. Technol.* 277 (July 2019), 116449. <http://dx.doi.org/10.1016/j.jmatprotec.2019.116449>.
- Hartmaier, A., 2020. Data-oriented constitutive modeling of plasticity in metals. *Materials* 13 (7), 1600. <http://dx.doi.org/10.3390/MA13071600>.
- Hershey, A.V., 1954. The plasticity of an isotropic aggregate of anisotropic face-centered cubic crystals. *J. Appl. Mech.* 21 (3), 241–249. <http://dx.doi.org/10.1115/1.4010900>.
- Hill, R., 1948. A theory of the yielding and plastic flow of anisotropic metals. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* 193 (1033), 281–297. <http://dx.doi.org/10.1098/rspa.1948.0045>.
- Hill, R., 1979. Theoretical plasticity of textured aggregates. *Math. Proc. Camb. Phil. Soc.* 85 (1), 179–191. <http://dx.doi.org/10.1017/S0305004100055596>.
- Hosford, W.F., 1979. On yield loci of anisotropic cubic metals. In: *Proceedings of the Seventh North American Metalworking Research Conference*. Dearborn, pp. 191–197.
- Hu, G., Huang, S., Lu, D., Zhong, X., Li, Z., Brocks, W., Zhang, K., 2015. Subsequent yielding of polycrystalline aluminum after cyclic tension–compression analyzed by experiments and simulations. *Int. J. Solids Struct.* 56–57, 142–153. <http://dx.doi.org/10.1016/J.IJUSOLSTR.2014.11.022>.
- Hutchinson, J.W., 1976. Bounds and self-consistent estimates for creep of polycrystalline materials. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* 348, 101–127. <http://dx.doi.org/10.1098/rspa.1976.0027>.
- Ibragimova, O., Brahme, A., Muhammad, W., Lévesque, J., Inal, K., 2021. A new ANN based crystal plasticity model for FCC materials and its application to non-monotonic strain paths. *Int. J. Plast.* 144, 103059. <http://dx.doi.org/10.1016/J.IJPLAS.2021.103059>.
- Ifitkhar, C.M.A., Li, Y.L., Kohar, C.P., Inal, K., Khan, A.S., 2021. Evolution of subsequent yield surfaces with plastic deformation along proportional and non-proportional loading paths on annealed AA6061 alloy: Experiments and crystal plasticity finite element modeling. *Int. J. Plast.* 143, 102956. <http://dx.doi.org/10.1016/J.IJPLAS.2021.102956>.
- Jang, D.P., Fazily, P., Yoon, J.W., 2021. Machine learning-based constitutive model for J2- plasticity. *Int. J. Plast.* 138, 102919. <http://dx.doi.org/10.1016/j.ijplas.2020.102919>.
- Karafillis, A.P., Boyce, M.C., 1993. A general anisotropic yield criterion using bounds and a transformation weighting tensor. *J. Mech. Phys. Solids* 41 (12), 1859–1886. [http://dx.doi.org/10.1016/0022-5096\(93\)90073-O](http://dx.doi.org/10.1016/0022-5096(93)90073-O).
- Kraska, M., Doig, M., Tikhomirov, D., Raabe, D., Roters, F., 2009. Virtual material testing for stamping simulations based on polycrystal plasticity. *Comput. Mater. Sci.* 46 (2), 383–392. <http://dx.doi.org/10.1016/J.COMMATSCI.2009.03.025>.
- Lebensohn, R.A., Kanjarla, A.K., Eisenlohr, P., 2012. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. *Int. J. Plast.* 32–33, 59–69. <http://dx.doi.org/10.1016/J.IJPLAS.2011.12.005>.
- Li, H., Hu, X., Yang, H., Li, L., 2016. Anisotropic and asymmetrical yielding and its distorted evolution: Modeling and applications. *Int. J. Plast.* 82, 127–158. <http://dx.doi.org/10.1016/J.IJPLAS.2016.03.002>.
- Lin, Y.C., Zhang, J., Zhong, J., 2008. Application of neural networks to predict the elevated temperature flow behavior of a low alloy steel. *Comput. Mater. Sci.* 43 (4), 752–758. <http://dx.doi.org/10.1016/J.COMMATSCI.2008.01.039>.
- Logan, R.W., Hosford, W.F., 1980. Upper-bound anisotropic yield locus calculations assuming <111>- pencil glide. *Int. J. Mech. Sci.* 22 (7), 419–430. [http://dx.doi.org/10.1016/0020-7403\(80\)90011-9](http://dx.doi.org/10.1016/0020-7403(80)90011-9).
- Lou, Y., Zhang, C., Zhang, S., Yoon, J.W., 2022. A general yield function with differential and anisotropic hardening for strength modelling under various stress states with non-associated flow rule. *Int. J. Plast.* 158, 103414. <http://dx.doi.org/10.1016/J.IJPLAS.2022.103414>.
- Lubarda, V.A., Mastilovic, S., Knap, J., 1996. Some comments on plasticity postulates and non-associative flow rules. *Int. J. Mech. Sci.* 38 (3), 247–258.

- Ma, H., Li, Y., Zhang, H., Li, Q., Chen, F., Cui, Z., 2022. A virtual laboratory based on full-field crystal plasticity simulation to characterize the multiscale mechanical properties of AHSS. *Sci. Rep.* 12 (1), 1–16. <http://dx.doi.org/10.1038/s41598-022-09045-8>.
- Malo, K.A., Hopperstad, O.S., Lademo, O.G., 1998. Calibration of anisotropic yield criteria using uniaxial tension tests and bending tests. *J. Mater. Process. Technol.* 80–81, 538–544. [http://dx.doi.org/10.1016/S0924-0136\(98\)00202-7](http://dx.doi.org/10.1016/S0924-0136(98)00202-7).
- Mianroodi, J.R., H. Siboni, N., Raabe, D., 2021. Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials. *Npj Comput. Mater.* 7 (1), 1–10. <http://dx.doi.org/10.1038/s41524-021-00571-z>.
- von Mises, R., 1928. Mechanik der plastischen Formänderung von Kristallen. *ZAMM -J. Appl. Math. Mech.* 8 (3), 161–185. <http://dx.doi.org/10.1002/ZAMM.19280080302>.
- Močkus, J., 1975. On Bayesian methods for seeking the extremum. In: *Optimization Techniques IFIP Technical Conference*. Springer Berlin Heidelberg, pp. 400–404. http://dx.doi.org/10.1007/978-3-662-38527-2_55.
- Nelder, J.A., Mead, R., 1965. A Simplex Method for Function Minimization. *Comput. J.* 7 (4), 308–313. <http://dx.doi.org/10.1093/COMJNL/7.4.308>.
- Raabe, D., Roters, F., Barlat, F., Chen, L.-Q., 2004. Continuum Scale Simulation of Engineering Materials. Wiley, <http://dx.doi.org/10.1002/3527603786>.
- Rong, H., Ying, L., Hu, P., Hou, W., 2021. Characterization on the thermal anisotropic behaviors of high strength AA7075 alloy with the Yld2004-18p yield function. *J. Alloys Compd.* 877, 159955. <http://dx.doi.org/10.1016/J.JALLCOM.2021.159955>.
- Roters, F., Diehl, M., Shanthraj, P., Eisenlohr, P., Reuber, C., Wong, S.L., Maiti, T., Ebrahimi, A., Hochrainer, T., Fabritius, H.O., Nikolov, S., Friák, M., Fujita, N., Grilli, N., Janssens, K.G., Jia, N., Kok, P.J., Ma, D., Meier, F., Werner, E., Stricker, M., Weygand, D., Raabe, D., 2019. DAMASK – The Düsseldorf Advanced Material Simulation Kit for modeling multi-physics crystal plasticity, thermal, and damage phenomena from the single crystal up to the component scale. *Comput. Mater. Sci.* 158 (January 2018), 420–478. <http://dx.doi.org/10.1016/j.commatsci.2018.04.030>.
- Roters, F., Eisenlohr, P., Hantcherli, L., Tjahjanto, D.D., Bieler, T.R., Raabe, D., 2010. Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications. *Acta Mater.* 58 (4), 1152–1211. <http://dx.doi.org/10.1016/J.ACTAMAT.2009.10.058>.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1988. Learning representations by back-propagating errors. *Nature* 323 (6088), 533–536. <http://dx.doi.org/10.1038/323533a0>.
- Sabokpa, O., Zarei-Hanzaki, A., Abedi, H.R., Haghdadi, N., 2012. Artificial neural network modeling to predict the high temperature flow behavior of an AZ81 magnesium alloy. *Mater. Des.* 39, 390–396. <http://dx.doi.org/10.1016/J.MATDES.2012.03.002>.
- Sharma, S., Sharma, S., Athaiya, A., 2020. Activation functions in neural networks. *Int. J. Eng. Appl. Sci. Technol.* 4, 310–316.
- Shoghi, R., Hartmaier, A., 2022. Optimal data-generation strategy for machine learning yield functions in anisotropic plasticity. *Front. Mater.* 231. <http://dx.doi.org/10.3389/FMATS.2022.868248>.
- Shutov, A., Ihlemann, J., 2013. On the phenomenological modelling of yield surface distortion. In: *Proceedings of 10th World Congress on Computational Mechanics*. Editora Edgard Blücher, pp. 1662–1670. <http://dx.doi.org/10.48550/arXiv.1309.3091>.
- Smith, L.N., 2018. A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay. <http://dx.doi.org/10.48550/arxiv.1803.09820>.
- Stewart, J.B., Cazacu, O., 2011. Analytical yield criterion for an anisotropic material containing spherical voids and exhibiting tension–compression asymmetry. *Int. J. Solids Struct.* 48 (2), 357–373. <http://dx.doi.org/10.1016/J.IJSOLSTR.2010.10.009>.
- Stoughton, T.B., Yoon, J.W., 2006. Review of Drucker's postulate and the issue of plastic stability in metal forming. *Int. J. Plast.* 22 (3), 391–433. <http://dx.doi.org/10.1016/J.IJPLAS.2005.03.002>.
- Sun, F., Liu, P., Liu, W., 2021. Multi-level deep drawing simulations of AA3104 aluminium alloy using crystal plasticity finite element modelling and phenomenological yield function. *Res. Article Adv. Mech. Eng.* 13 (3), 1–19. <http://dx.doi.org/10.1177/16878140211001203>.
- Tamura, K., Gallagher, M., 2019. Quantitative measure of nonconvexity for black-box continuous functions. *Inform. Sci.* 476, 64–82. <http://dx.doi.org/10.1016/J.INS.2018.10.009>.
- Taylor, G.I., 1938. Plastic strain in metals. *J. Inst. Met.* 62, 307–324.
- Tomé, C.N., Maudlin, P.J., Lebensohn, R.A., Kaschner, G.C., 2001. Mechanical response of zirconium—I. Derivation of a polycrystal constitutive law and finite element analysis. *Acta Mater.* 49 (15), 3085–3096. [http://dx.doi.org/10.1016/S1359-6454\(01\)00190-2](http://dx.doi.org/10.1016/S1359-6454(01)00190-2).
- Tsai, J.T., Chou, J.H., Liu, T.K., 2006. Tuning the structure and parameters of a neural network by using hybrid Taguchi-genetic algorithm. *IEEE Trans. Neural Netw.* 17 (1), 69–80. <http://dx.doi.org/10.1109/TNN.2005.860885>.
- Van Houtte, P., Delannay, L., Samajdar, I., 1999. Quantitative prediction of cold rolling textures in low-carbon steel by means of the Lamel model. *Texture, Stress, and Microstructure* 31 (3), 109–149. <http://dx.doi.org/10.1155/TSM.31.109>.
- Van Houtte, P., Li, S., Seefeldt, M., Delannay, L., 2005. Deformation texture prediction: from the Taylor model to the advanced Lamel model. *Int. J. Plast.* 21 (3), 589–624. <http://dx.doi.org/10.1016/J.IJPLAS.2004.04.011>.
- Van Houtte, P., Yerra, S.K., Van Bael, A., 2009. The Facet method: A hierarchical multilevel modelling scheme for anisotropic convex plastic potentials. *Int. J. Plast.* 25 (2), 332–360. <http://dx.doi.org/10.1016/J.IJPLAS.2008.02.001>.
- Vlassis, N.N., Sun, W.C., 2020. Sobolev training of thermodynamic-informed neural networks for smoothed elasto-plasticity models with level set hardening. *Comput. Methods Appl. Mech. Engrg.* 377, <http://dx.doi.org/10.1016/j.cma.2021.113695>.
- Wu, P.D., Jain, M., Savoie, J., MacEwen, S.R., Tuğcu, P., Neale, K.W., 2003. Evaluation of anisotropic yield functions for aluminum sheets. *Int. J. Plast.* 19 (1), 121–138. [http://dx.doi.org/10.1016/S0749-6419\(01\)00033-X](http://dx.doi.org/10.1016/S0749-6419(01)00033-X).
- Wu, P.D., MacEwen, S.R., Lloyd, D.J., Neale, K.W., 2004. Effect of cube texture on sheet metal formability. *Mater. Sci. Eng. A* 364 (1–2), 182–187. <http://dx.doi.org/10.1016/J.MSEA.2003.08.020>.
- Yoon, J.W., Barlat, F., Dick, R.E., Karabin, M.E., 2006. Prediction of six or eight ears in a drawn cup based on a new anisotropic yield function. *Int. J. Plast.* 22 (1), 174–193. <http://dx.doi.org/10.1016/j.ijplas.2005.03.013>.
- Zhang, H., Diehl, M., Roters, F., Raabe, D., 2016. A virtual laboratory using high resolution crystal plasticity simulations to determine the initial yield surface for sheet metal forming operations. *Int. J. Plast.* 80, 111–138. <http://dx.doi.org/10.1016/j.ijplas.2016.01.002>.
- Zhang, K., Holmedal, B., Hopperstad, O.S., Dumoulin, S., Gawad, J., Van Bael, A., Van Houtte, P., 2015. Multi-level modelling of mechanical anisotropy of commercial pure aluminium plate: Crystal plasticity models, advanced yield functions and parameter identification. *Int. J. Plast.* 66, 3–30. <http://dx.doi.org/10.1016/j.ijplas.2014.02.003>.
- Zhang, K., Holmedal, B., Mánik, T., Saai, A., 2019. Assessment of advanced Taylor models, the Taylor factor and yield-surface exponent for FCC metals. *Int. J. Plast.* 114, 144–160. <http://dx.doi.org/10.1016/J.IJPLAS.2018.10.015>.
- Zhang, A., Mohr, D., 2020. Using neural networks to represent von Mises plasticity with isotropic hardening. *Int. J. Plast.* 132, 102732. <http://dx.doi.org/10.1016/J.IJPLAS.2020.102732>.
- Zhang, Y., Van Bael, A., Andrade-Campos, A., Coppieters, S., 2022. Parameter identifiability analysis: Mitigating the non-uniqueness issue in the inverse identification of an anisotropic yield function. *Int. J. Solids Struct.* 243, 111543. <http://dx.doi.org/10.1016/J.IJSOLSTR.2022.111543>.