Report Title:	High Throughput Computational Framework of Materials Properties for Extreme Environments
Type of Report:	Final Scientific/Technical Report
Reporting Period Start Date:	12/15/2017
Reporting Period End Date:	09/30/2022
Principal Author(s):	Zi-Kui Liu (Dorothy Pate Enright Professor, PI) Shun-Li Shang (Research Professor, Co-PI) Allison M. Beese (Associate Professor, Co-PI) Yi Wang (Research Professor) Adnan Eghtesad (Postdoc) John Shimanek (PhD student) Shipin Qin (PhD student) Shuang Lin (PhD student) Hui Sun (PhD student) Brandon Bocklund (PhD student)
Date Report was Issued:	December 2022
DOE Award Number:	DE-FE0031553
Name and Address of Submitting Organization:	Pennsylvania State University, University Park, PA 16802

DISTRIBUTION LIMITATION NOTICES

No distribution limitation, all contents in this report are open to the public.

ABSTRACT

This project aims to establish a framework capable of efficiently predicting the properties of structural materials for service in harsh environments over a wide range of temperatures and over long periods of time. The approach is to develop and integrate high throughput first-principles calculations in combination with machine learning (ML) methods, perform high throughput CALPHAD (calculations of phase diagrams) modeling, and carry out finite element method (FEM) simulations. Relevant to high temperature service in fossil power system, nickel-based superalloys such as Inconel 740 and Haynes 282 as well as the associated (Ni-Cr-Co)-Al-C-Fe-Mn-Mo-Nb-Si-Ti system, were investigated.

The present framework was built on the concept of phase-based property data, in which properties of individual phases are modeled as a function of internal and external independent variables. This project established an open-source infrastructure with the following capabilities:

- High throughput implementation of first-principles calculations at finite temperatures and variable compositions using both accurate phonon calculations and the efficient Debye model for thermodynamic properties, elastic constants, diffusion coefficients, vacancy formation, stacking and twin faults, and dislocation mobility; i.e., using the developed code DFTTK [1];
- Machine learning capabilities to predict the above properties so that the number of first-principles calculations can be significantly reduced; e.g., using the developed code SIPFENN [2];
- High throughput CALPHAD modeling of the above properties as a function of temperature and composition using our unique capability based on ESPEI and PyCalphad [3];
- New capabilities to predict the stress-strain behavior of individual phases; and
- New models for tensile strength prediction in common FEM software with the crystal plasticity finite element simulations (CPFEM).

1 Table of Contents

1 EXECUTIVE SUMMARY

The present project aims to establish a framework capable of the efficient prediction of properties of structural materials for service in harsh environments over a wide range of temperatures and over long periods of time, through developing and integrating high throughput first-principles calculations based on density functional theory (DFT) in combination with machine learning methods, high throughput CALPHAD modeling, and finite element method (FEM) simulations.

This project was built on the expertise of the PIs and the state-of-the-art capabilities such as open-source high throughput DFT-based predictions, CALPHAD modeling using various software (e.g.: atomate, PyCalphad, and ESPEI) and the machine learning methods in the community. Our ESPEI is capable of testing and developing new models with uncertainty quantifications. The multi-scale integration from electronic structures, phonon properties, phase stability, and phase transformation to macroscopic tensile property extends our early initiatives in the field, which enabled the PI Liu to coin and trademark the term Materials Genome®. Therefore, our technology can be expected to contribute to a developmental breakthrough for high throughput computational design of multicomponent alloy compositions and tensile properties at elevated temperatures for long-term service. Furthermore, the open-source strategy of our approaches and tools enable scientists and engineers in the community to adapt the framework and continue their own development of specific approaches and tools. Additionally, the incorporation of the state-of-the-art machine learning methods developed in the community into the proposed framework can significantly reduce the computing expense in DFT-based first-principles calculations.

This project have **six** major achievaments in this project; as detailed in the following sections.

First. Development of thermodynamic databases. The CALPHAD-based thermodynamic database for Ni-based alloys includes the key elmenets of Ni-Cr-Fe-Mo-Nb, and the DFT-based database for Ni-, Fe-, and Co-based alloys includes the key elements of Al, B, C, Cr, Cu, Hf, La, Mn, Mo, N, Nb, O, P, Re, Ru, S, Si, Ta, Ti, V, W, Y, and Zr.

Second. Development of simulation tools for high throughput DFT-based first-principles calculations at finite temperatures via DFTTK; for machine learning predictions of enthalpy of formation via SIPFENN; and for high throughput CALPHAD modeling via PyCalphad and ESPEI.

Third. Development of ideal shear strength database for FEM simulations of Ni-based alloys based on first-principles calculations, including Ni_3Al and the dilute Ni-X and Ni-X-Z alloys.

Fourth. Development of DFT-informed FEM simulations to predict strain-stress curves, as applied for single and polycrystal Ni and Ni-X alloys.

Fifth. Exploration of machine learning methodology to predict stacking fault energy and microscale features responsible for the formation of local stress and strain hotspots within the grains and near the grain boundaries.

Sixth. Application of thermodynamic and kinetic database to explore critical Al concentration to form external Al_2O_3 scale on Ni-Al alloys.

The above six achievaments make it possible to establish a framework capable of the efficient prediction of properties of structural materials for service in harsh environments over a wide range of temperatures and over long periods of time through an integration of high throughput DFT-based simulations, high throughput CALPHAD modeling, machine learning, and FEM simulations.

2 ACTUAL ACCOMPLISHMENTS COMPARED WITH PROJECT OBJECTIVES

Objectives set in the proposed project:

"The proposed computational framework will establish an open-source infrastructure in the Python environment including the following specific capabilities:

- *High throughput implementation of first-principles calculations at finite temperatures and variable compositions using both the accurate phonon calculations and the efficient Debye model for thermodynamic properties, elastic constants, diffusion coefficients, vacancy formation, stacking and twin faults, grain boundary and interfacial energy, and dislocation mobility;*
- *Exploration of machine learning methods to predict the above properties so that the amount of first-principles calculations can be significantly reduced;*
- *High throughput CALPHAD modeling of the above properties as a function of temperature and composition using our unique capability based on pycalphad and ESPEI;*
- *New capabilities to be developed in the proposed project to predict the stress-strain behavior of individual phases with dislocations and cohesive strength of grain boundary and interfaces; and*
- *New models for tensile strength prediction in common finite element method (FEM) analysis software."*

Acutal accomplishments with respect to the check points set in the proposal: ALL DONE; see below the decision points, and more details in Section 3: Report Details.

Decision point 1: High throughput implementation of first-principles calculations
Criteria for success: Development of Python-based software for high throughput DFT-based first-principles calculations

Acutal accomplishment w.r.t. decision point 1: ***DONE.***
Software DFT Tool Kit (DFTTK) has been developed [1], see also https://www.dfttk.org

Decision point 2: Exploration of machine learning (ML) methods to predict properties
Criteria for success: Development of Python-based software and exploration of ML methods.

Acutal accomplishment w.r.t. decision point 2: ***DONE.***
Software SIPFENN [2] has been developed to predict enthalpy of formation, and ML methods have been explored to predict stacking fault energy [4] and to reveal the microscale features responsible for the formation of local stress and strain hotspots within the grains and near the grain boundaries [5].

Decision point 3: High throughput CALPHAD modeling and improvement of PyCalphad/ESPEI
Criteria for success: PyCalphad and ESPEI are further improved and Ni-based thermodynamic database are generated.

Acutal accomplishment w.r.t. decision point 3: ***DONE.***
ESPEI/PyCalphad were developed, see details in https://espei.org and https:/pycalphad.org. In addition, the Ni-Cr-Fe-Mo-Nb database was generated.

Decision points 4 and 5: Development of finite element method (FEM) for strain-stress predictions
Criteria for success: Novel CPFEM is developed to predict strain-stress curves.

Acutal accomplishment w.r.t. decision points 4 and 5: ***DONE.***
DFT-informed FEM simulations were developed to predict strain-stress curves as applied for single and polycrystal Ni and Ni-X alloys [6–8]

The milestone status is listed in Table 1**.**

Table 1: Milestone status report.

Milestone title /Description	Planned completion date	Actual completion date	Verification method	Comments
Milestone 1: Completion of literature survey of thermodynamic and other properties in the (Ni-Cr-Co)-Al-C-Fe-Mn-Mo-Nb-Si-Ti system relevant to the prediction of tensile properties.	3/31/2018	4/30/2018	Literature report. Found practice methods to predict CRSS.	Done.
Milestone 2: DFT calculations and CALPHAD modeling of thermodynamic and other properties for all binary phases in (Ni-Cr-Co)-Al-C-Fe-Mn-Mo-Nb-Si-Ti.	12/31/2018	9/30/2021	Database and publications	Done
Milestone 3: DFT calculations of tensile properties for γ-phase and TC-PRISMA simulations of phase distribution and implementation of tensile strength model for γ-phase	7/31/2019	9/30/2021	Database and publications	Done
Milestone 4: Development of properties database including all phases by CALPHAD modeling	5/31/2019	9/30/2021	Properties database	Done
Milestone 5: DFT calculations and modeling of tensile properties for γ'-phase; TC-PRISMA simulations of phase distributions; open-source code to prepare input for FEM simulations; and results from experimental validations	6/30/2020	6/30/2022	Database, publications. and open-source code (DFTTK)	Done
Milestone 6: Open-source code for high-throughput DFT calculations	8/31/2020	9/30/2021	Open-source code	Done. DFTTK released.
Milestone 7: Open-source code to high-throughput CALPHAD modeling for the properties of interest in the proposed project, i.e., the PyCalphad code	9/30/2020	9/30/2021	Open-source code	Done. ESPEI and PyCalphad. Released and improved.
Milestone 8: A complete FEM model to predict tensile properties of real alloys together with validation from experiments	10/31/2020	9/30/2022	Publications	Done.
Milestone 9: Final report together with the complete open-source codes for high throughput DFT, high throughput CALPHAD, and FEM simulations	12/31/2020	12/31/2022	Final report, open-source codes	Done. This report.

* TCP represents “topologically closed packed”

3 REPORT DETAILS

Aiming to establish a high throughput computational framework capable of the efficient prediction of properties of structural materials for service in harsh environments over a wide range of temperatures and over long periods of time, we developed high throughput DFT-based first-principles calculations in combination with machine learning (ML) methods, high throughput CALPHAD modeling, and finite element method (FEM) simulations. The present project has the following key accomplishments within five years (including no cost extension of two years). For each of the accomplishments, we show the key details, results and discussion, and conclusions. Also we refer to our publications and our websites for each accomplishment when available:

- Development of open-source software for DFT, ML, and CALPHAD calculations
- Development of thermodynamic database and its applications
- DFT-based pure shear deformation to predict ideal shear strength and stacking fault energy
- DFT-informed FEM simulations

3.1 Development of open-source software packages

3.1.1 Machine learning for thermodynamic properties

We developed a user-friendly open-source tool (SIPFENN) for predicting the formation enthalpy of any atomic structure on the millisecond timescale based on a structure file (POSCAR/CIF/etc.) or on the microsecond timescale if the descriptor has been calculated.

Unlike the vast majority of other reported models, SIPFENN does not require any intensive model training from the user but is delivered ready to use on a desktop PC or a laptop. SIPFENN does not require any input other than structure files, what allows effortless integration into any CALPHAD study within minutes. On the computational end, pymatgen (www.pymatgen.org) converts structure data into a unified form, and if needed, performs an additional project-specific analysis. Descriptors are generated with Magpie and then passed through one of the trained neural networks to predict the values of formation energy. Our code is built to allow easy switch of the neural network, which is encoded in a popular open-source format of MXNet. We leverage that by making it easy for others to re-train the network to include new data into the model at a fraction of computational cost (transfer learning). The model itself is built using a structure-informed descriptor, open DFT databases, and new neural network architectures designed by us. On a random 5% subset of the OQMD (www.oqmd.org), we achieve a mean absolute error (MAE) of 28 meV/atom. In addition to trivial cross-validation, we also tested our model on a few special quasirandom structures (SQS's) and 243 endmembers of 5-sublattice model of Fe-Cr-Ni σ-phase, getting very similar results in both cases.

This work was published in Computational Materials Science [2]:

- *A.M. Krajewski, J.W. Siegel, J. Xu, Z.-K. Liu, Extensible Structure-Informed Prediction of Formation Energy with improved accuracy and usability employing neural networks, Comput. Mater. Sci. 208 (2022) 111254.*

Figure 1 shows the SIPFENN schematic description of operation [2]; see also our website for more details: https://www.phaseslab.com/sipfenn.

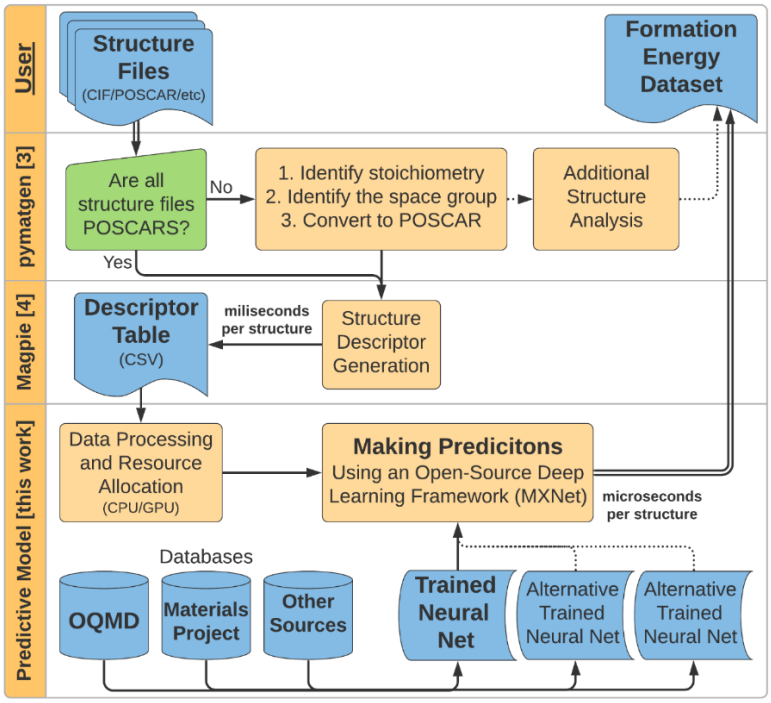


Figure 1. SIPFENN schematic description of operation [2].

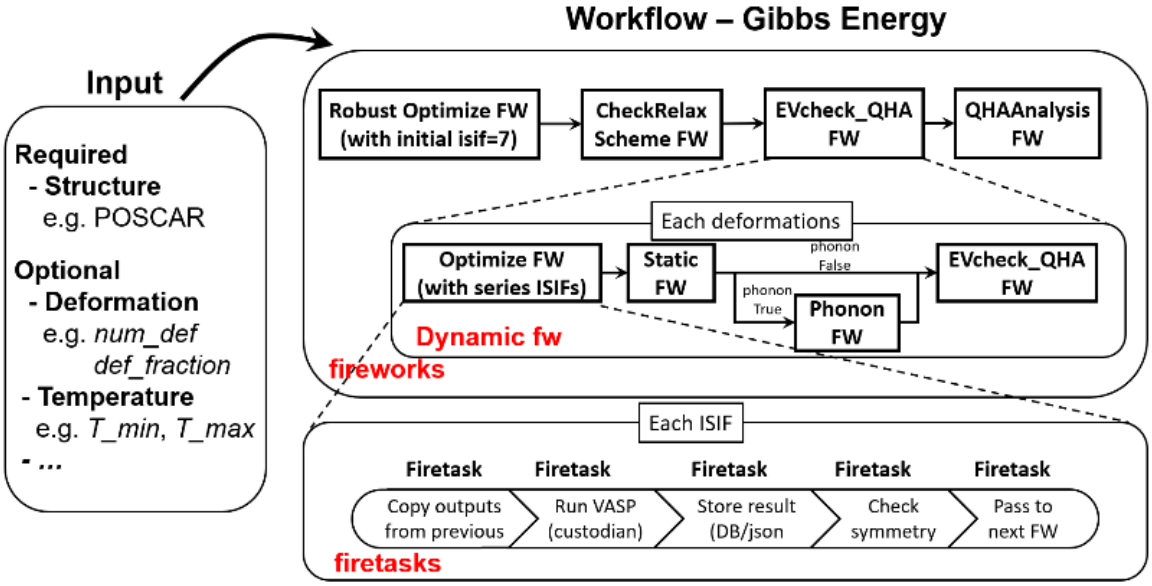


Figure 2. Schematic DFTTK workflow with FW (firework) for job control and ISIF for structure relaxations in VASP (www.vasp.at); see details in [1].

3.1.2 DFTTK for high throughput first-principles calculations

We developed the DFT Tool Kit (www.dfttk.org) and published a paper about DFTTK in the journal of CALPHAD :

- Y. Wang, M. Liao, B.J. Bocklund, P. Gao, S.-L. Shang, H. Kim, A.M. Beese, L.-Q. Chen, Z.-K. Liu, DFTTK: Density Functional Theory ToolKit for high-throughput lattice dynamics calculations, Calphad. 75 (2021) 102355.

In this paper, we present a software package in Python for high-throughput first-principles calculations of thermodynamic properties at finite temperatures, which we refer to as DFTTK (Density Functional Theory Tool Kit).

DFTTK is based on the atomate package and integrates our experiences in the last decades on the development of theoretical methods and computational software; see its workflow in Figure 2. It includes task submissions on all major operating systems and task execution on high-performance computing environments. The distribution of the DFTTK package comes with examples of calculations of phonon density of states, heat capacity, entropy, enthalpy, and free energy under the quasiharmonic phonon scheme for the stoichiometric phases of Al, Ni, Al_3Ni, AlNi, $AlNi_3$, Al_3Ni_4, Al_3Ni_5, and the fcc solution phases treated using the special quasirandom structures at the compositions of Al_3Ni, AlNi, and $AlNi_3$.

3.1.3 ESPEI/PyCalphad for high throughput CALPHAD modeling

ESPEI (www.espei.org) or Extensible Self-optimizing Phase Equilibria Infrastructure, is a tool for thermodynamic database development within the CALPHAD method. It uses PyCalphad (www.pycalphad.org) for calculating Gibbs energies of thermodynamic models. PyCalphad is a Python library for computational thermodynamics using the CALPHAD method.

In this project, ESPEI was updated to use site fraction to model phase diagrams. In the previous version of ESEPI, site fraction data cannot be considered as a direct input for CALPHAD modeling due to the complexity of site fraction calculations. However, with Bayesian parameter estimation in ESPEI, site fraction data together with other thermochemical and phase equilibrium data can be directly used to be fitted with Gibbs energy parameters by Markov Chain Monte Carlo (MCMC) optimization in ESPEI. For example, the experimental data of site fraction by Joubert [9] were used to modify the mu phase (another TCP phase) in the Nb-Ni system [10]. The comparison between site fraction of Chen's paper (Figure 3a) and the site fraction of the present work (Figure 3b) shows that after taking site fraction data as direct input for CALPHAD modeling, the discrepancy between simulation results and experimental data become much smaller.

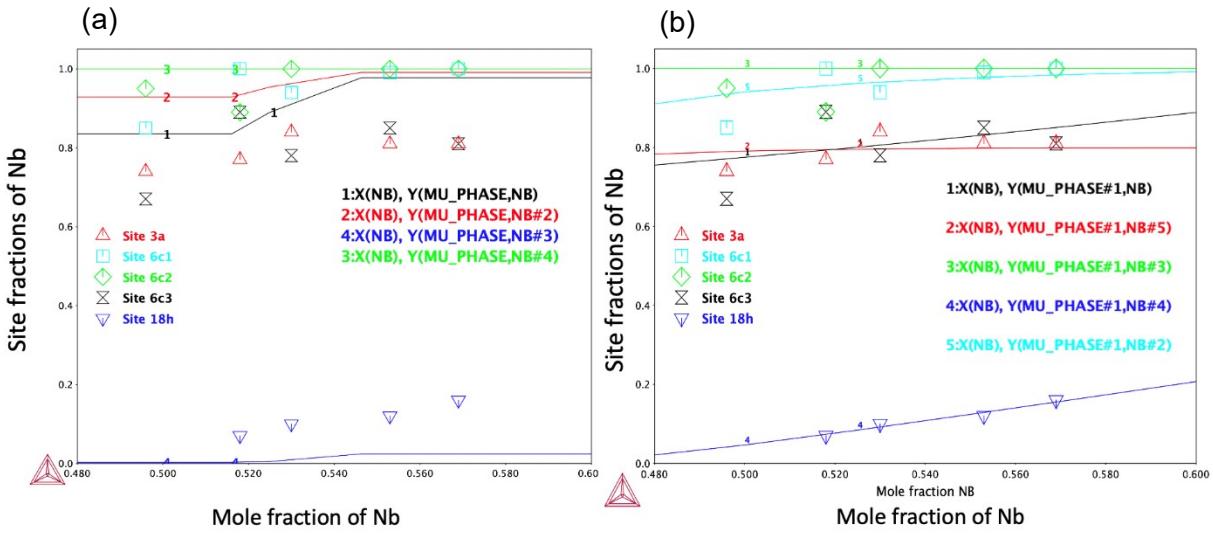


Figure 3. Site fraction of Nb in μ phase in Nb-Ni: (a) This work and (b) Chen's work [11].

In addition to implementation of site fraction, more updates of PyCalphad and ESPEI were listed in their websites:

- www.espei.org
- www.pycalphad.org

3.2 Development of thermodynamic database and its applications

3.2.1 DFT-based thermodynamic database and its applications

Based on mainly the developed DFTTK code [12] and the quasiharmonic phonon calculations using our code YPHON [13], we performed high-throughout DFT calculations at finite temperatures for over 3000 structures for the systems made of 26 elements, namely, Ni, Fe, and Co alloyed with Al, B, C, Cr, Cu, Hf, La, Mn, Mo, N, Nb, O, P, Re, Ru, S, Si, Ta, Ti, V, W, Y, and Zr.

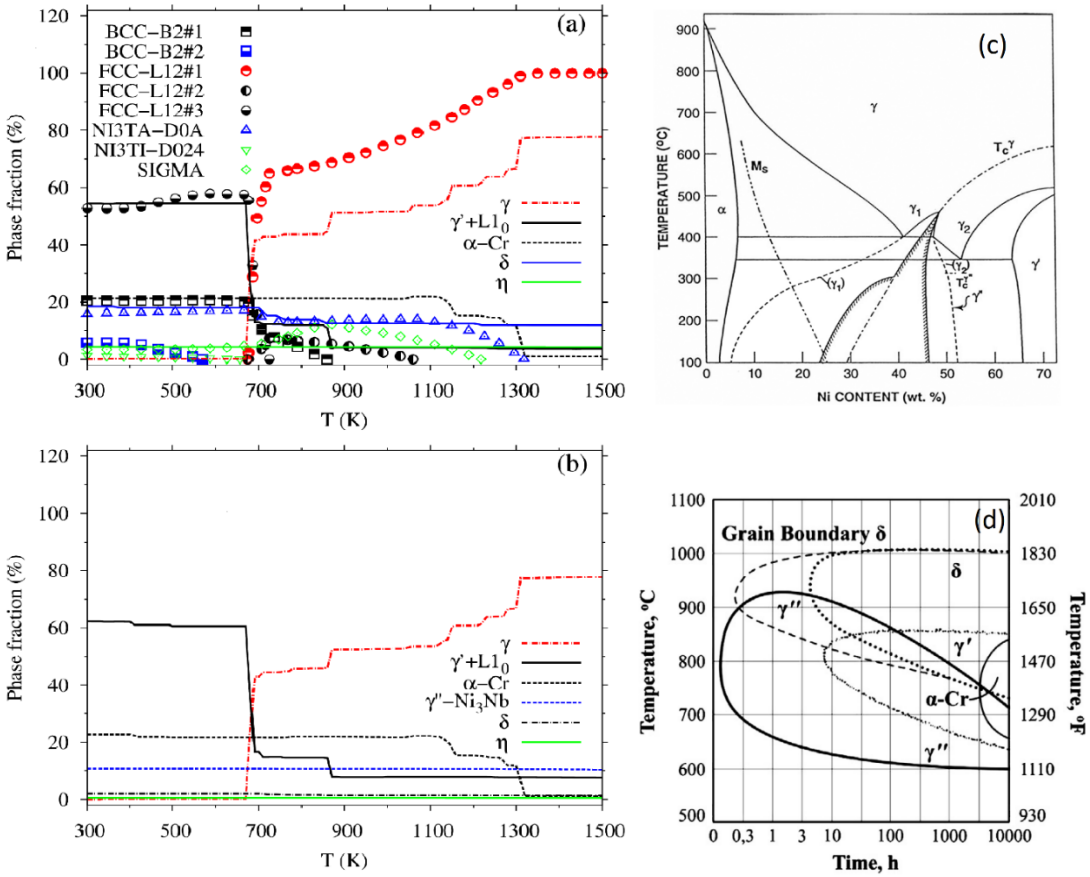


Figure 4. (a) Phase fractions of IN718 calculated by DFT (lines) without considering grain boundary constraint compared those calculated based on commercial database TCNI9 (symbols, www.thermocalc.com). (b) Phase fractions for IN718 calculated by DFT (lines) with considering grain boundary constraint. (c) Fe-Ni binary phase diagram (Courtesy of Yang et al. [14,15]). (d) T-T-T diagram for IN718 (Courtesy of Oradei-Basile and Radavich [16]).

This work aims to provide reliable thermochemical data that are mostly unavailable from experiments. Accordingly, a computational procedure was developed for predicting morphology evolutions as functions of temperature and composition. The applicability of this database was demonstrated using the AM Ni-based superalloy IN718 by monitoring the formations of various phases and their evolution during the heating-cooling cycles. Furthermore, to support the theoretical approach, a series of transmission electron microscopy (TEM) and energy dispersive X-ray spectroscopy (EDS)

measurements of AM IN718 samples were performed to obtain spatial elemental distribution data. The procedure was verified by the successful predictions of the formations of various phases at low temperatures [$L1_0$-FeNi, γ'-Ni_3(Fe,Al), α-Cr, δ-Ni_3(Nb,Mo), γ"-Ni_3Nb , and η-Ni_3Ti], intermediate temperatures [γ'-Ni_3Al, δ-Ni_3Nb, γ"-Ni_3Nb, α-Cr, and γ-Ni(Fe,Cr,Mo)], and high temperatures [δ-Ni_3Nb and γ-Ni(Fe,Cr,Mo)]. For example, Figure 4 shows the presently predicted phases in Ni-based Inconel 718 (IN718) in comparison with experimental Fe-Ni phase diagram [14,15] and the TTT (time-temperature-transformation) diagram for IN718 [16]. Figure 4 shows the presently predicted enthalpies of IN718 as a function of temperature in comparison with the results from commercial database TCNI9 (c.f., www.thermocalc.com) and experimental data [17,18]. It shows that the present results are in good agreement with experiments.

Currently this work was published:

- Y. Wang, F. Lia, K. Wang, K. McNamara, Y. Ji, X. Chong, S.-L. Shang, Z.-K. Liu, R.P. Martukanitz, L.-Q. Chen, A thermochemical database from high-throughput first-principles calculations and its application to analyzing phase evolution in AM-fabricated IN718, Acta Mater. 240 (2022) 118331.

3.2.2 CALPHAD-based thermodynamic database and its applications

As a model system to examine Ni-based Inconel alloys (e.g., 625 and 718), a thermodynamic database of the Ni-Cr-Fe-Mo-Nb system was established based on our DFT-based first-principles results for the topologically close pack (TCP) phases. We used the high throughput CALPHAD modeling software packages ESPEI and PyCalphad [19] to adjust phase boundaries with respect to experimental data. The adopted sublattice models for the key TCP phases are as follows:

- σ (5-sublattice to represent 5 Wyckoff positions)
- μ (5-sublattice to represent 5 Wyckoff positions)
- C14-laves (3-sublattice to represent 3 Wyckoff positions)
- δ (3-sublattice to represent 3 Wyckoff positions)
- γ' (2-sublattice to represent 2 Wyckoff positions), and
- γ'' (3-sublattice to represent 3 Wyckoff positions).

One of the examples is the Ni-Nb system shown in Figure 5 [10]. Figure 5 (a) is the phase diagram from the TCNI9 commercial database developed by Thermo-Calc [20], Figure 5(b) is the phase diagram calculated by the present database. It can be seen from Figure 5(b), a better fitting with experimental data is achieved for δ phase around compositions of 23.5 – 26.5 at. % Nb, and also a better agreement for μ phase around 49.8 - 58.3 at. % Nb. At the same time, Figure 5 (c-d) shows the Gibbs energy surface of δ phase calculated by PyCalphad [19]. Note that the three-sublattice model based on the Wyckoff positions are applied for the δ phase (Figure 5 (d)) instead of the two-sublattice model used by the TCNI9 (Figure 5 (c)). With the presently updated sublattice, it allows the CALPHAD model to actually describe the Gibbs energy by incorporating more endmembers (the edges in the Gibbs energy surface) as show in Figure 5 (d), which give a better description of thermochemical properties like site fractions shown in Figure 3 and reduce the limitation to model multi-component system.

The ternary systems (such as Ni-Fe-Mo, Ni-Fe-Nb, Ni-Mo-Nb, and Fe-Mo-Nb) were also optimized based on thermochemical and phase boundary data after all the binary systems were finished. For example, the isothermal section of Fe-Nb-Ni at 1373 K from the TCNI 9 database (Figure 6 (a)) and the present database (Figure 6 (b)) shows that the calculations from the present database has a better match with experimental data of the C14-laves phase. For the δ phase region, two databases both shows a good agreement with experimental data, while the present database indicates that the δ phase region is wider with increasing Fe. The isothermal section of Fe-Mo-Nb at 1373 K from TCNI 9 database

(Figure 6 (c)) and the present database (Figure 6 (d)) shows that the present database has a better agreement with experimental data for the C14-laves region and the μ phase region.

Currently we published one CALPHAD modeling work of Ni-Hf [21] and are preparing more manuscripts to report the above modeling results.

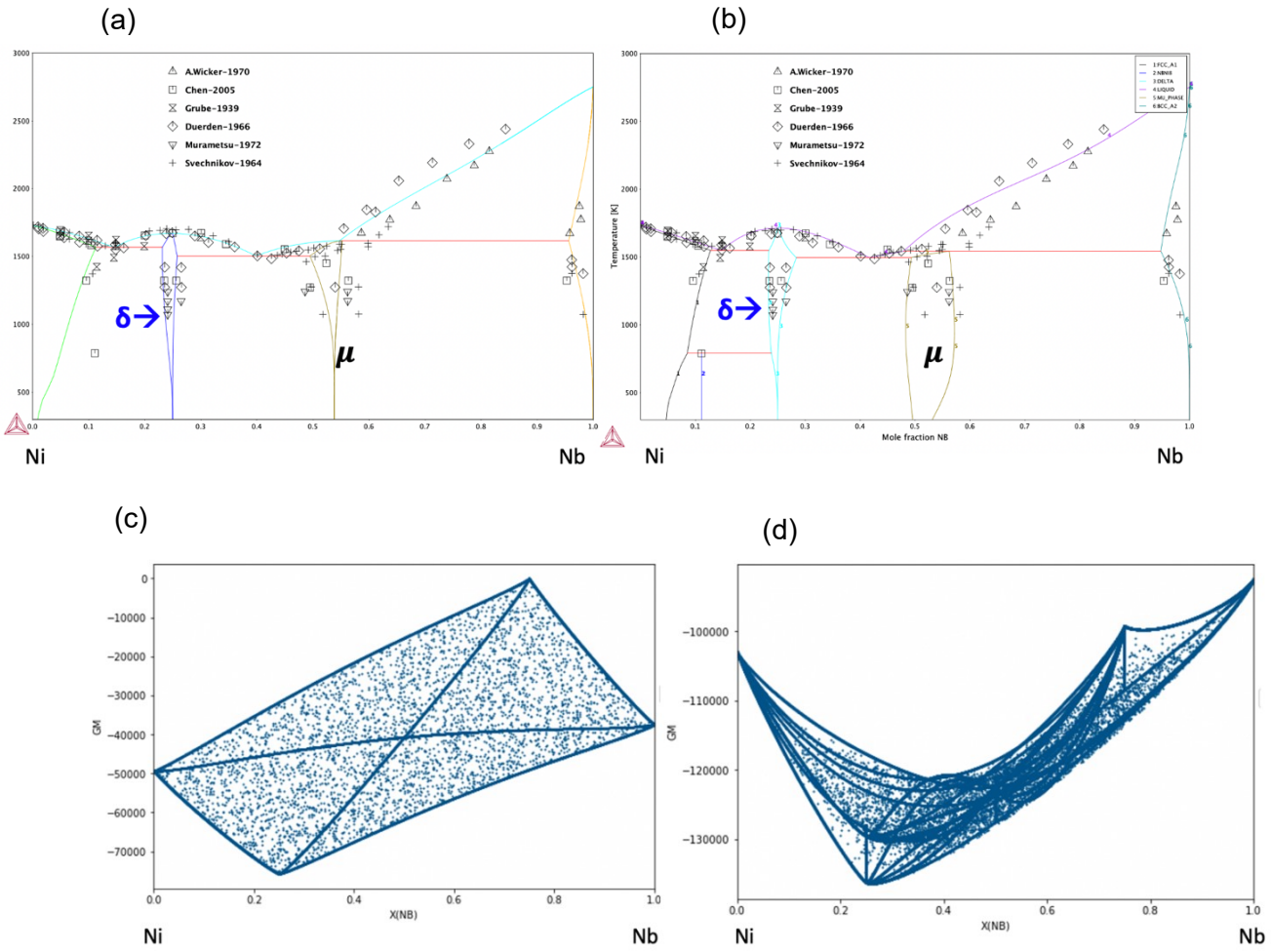


Figure 5. Nb-Ni phase diagram with experimental data and Gibbs energy surface of δ phase. (a) Nb-Ni phase diagram of TCNI 9 database. (b) Nb-Ni phase diagram of the present database. (c) Gibbs surface energy of δ from TCNI 9 database. (d) Gibbs surface energy of δ from the present database.

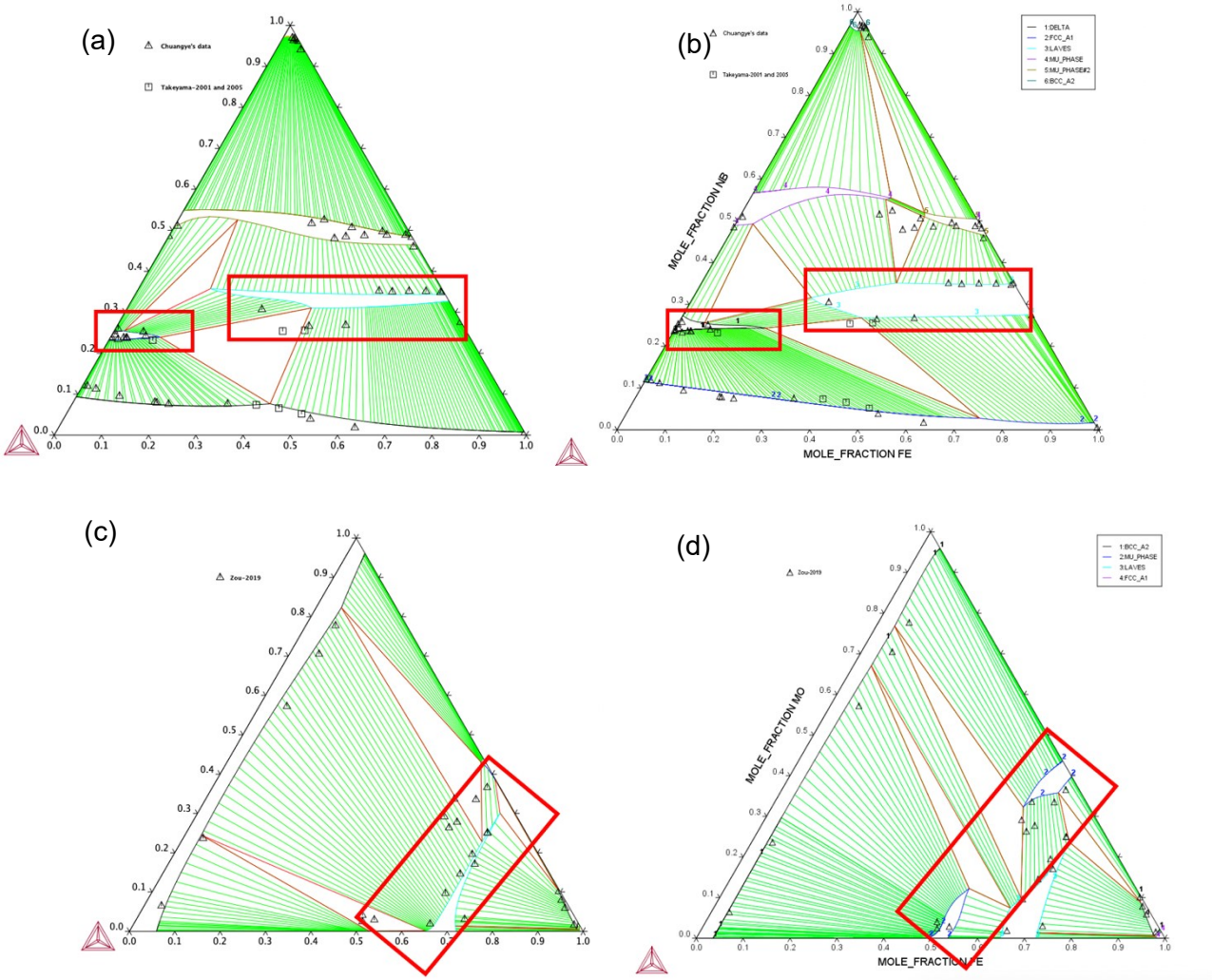


Figure 6. Isothermal section of Fe-Nb-Ni at 1373K from TCNI 9 (a) and the present database (b) and isothermal section of Fe-Mo-Nb at 1373 K from TCNI 9 (c) and the present database (d).

3.3 DFT-based pure shear deformation and its applications

Shear deformation and associated properties such as stacking fault energy, γ, ideal (or theoretical) shear strength, τ_{IS}, and critical resolved shear stress (CRSS), τ_{CRSS}, are fundamental for understanding and modeling a vast number of materials properties and phenomena related to dislocations, plastic deformation, crystal growth, and phase transitions. Notably, τ_{IS} is a key parameter to estimate the Peierls stress – the force required to move an individual dislocation. The Peierls stress, τ_P, is approximately equal to τ_{CRSS} at 0 K. For the case of a wide dislocation and according to Joós and Duesbery [22],

$$\tau_P = \frac{Kb}{a}\exp\left(-2\pi\zeta/a\right) \qquad \text{Eq. 2}$$

where *b* is the Burgers vector, *a* the row spacing of atoms within the slip plane, *K* an elastic factor, and ζ the half-width of the dislocation with $\zeta = \frac{Kb}{4\pi\tau_{IS}}$. For an isotropic crystal, $K_{iso} = \mu(\frac{sin^2\theta}{1-\nu} + cos^2\theta)$ with μ being shear modulus, ν Poisson's ratio, and θ the angle between the dislocation line and its Burgers vector. To predict the τ_{IS} values, we adopted the DFT-based alias shear deformation scheme. For

example of Ni and Ni_3Al, estimates of the Peierls stresses using τ_{IS} and elastic properties suggest the prevalence of edge dislocations in Ni and screw dislocations in Ni_3Al, agreeing with experimental observations regarding the dominance of edge dislocations in the first stage of crystal deformation in fcc metals and the yield-strength anomaly related to screw dislocations in Ni_3Al; see Figure 7.

In addition to Ni_3Al, we also predicted the effect of alloying elements (denoted X) on the ideal shear strength for 26 dilute Ni-based alloys, $Ni_{11}X$, as determined by first-principles calculations of pure alias shear deformations; see Figure 8. We found that the variations in ideal shear strength are quantitatively explored with correlational analysis techniques, showing the importance of atomic properties such as size and electronegativity. The shear moduli of the alloys are affirmed to show a strong linear relationship with their ideal shear strengths, while the shear moduli of the individual alloying elements were not indicative of alloy shear strength.

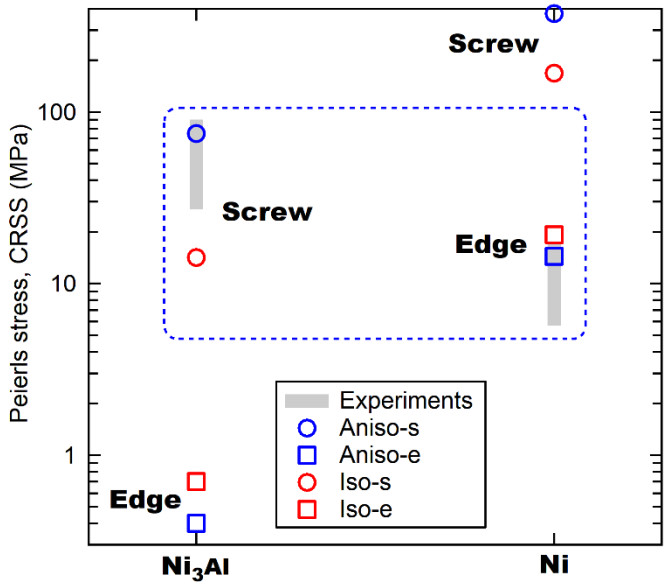


Figure 7. Calculated Peierls stresses at 0 K for four cases of edge (e) and screw (s) dislocations using elastic factors for isotropic (iso) and anisotropic (aniso) crystals in comparison with experimental CRSS values at room temperature for Ni_3Al and Ni.

											Al	Si
											4.58	4.17
Sc	Ti	V	Cr	Mn	Fe	Co	Ni, pure	Cu	Zn			
3.62	4.24	4.62	4.90	5.12	5.20	5.46	5.13	4.51	4.42			
Y_sv	Zr_sv	Nb_pv	Mo_pv	Tc_pv	Ru	Rh	Pd					
2.06	3.03	3.91	4.5	4.96	5.21	5.27	4.49					
	Hf_pv	Ta_pv	W_pv	Re	Os	Ir	Pt					
	3.23	3.93	4.57	5.02	5.24	5.36	4.91					

Figure 8. Calculated ideal shear strengths for 26 alloying elements and pure Ni (i.e., $Ni_{11}X$) in units of GPa. The “pv” and “sv” listed after an atomic symbol indicates that the DFT calculations were performed while treating p or s states as valence states. Warmer colors indicate a higher value of ideal shear strength (blue = low, yellow = middle, red = high).

Currently we published two papers and one will be published soon.

- S.L. Shang, J. D. Shimanek, S. Qin, Y. Wang, A.M. Beese, Z.K. Liu, Unveiling dislocation characteristics in Ni3Al from stacking fault energy and ideal strength: A first-principles study via pure alias shear deformation, Phys. Rev. B. 101 (2020) 024102.
- J. D. Shimanek, S.-L. Shang, A.M. Beese, Z.-K. Liu, Insight into ideal shear strength of Ni-based dilute alloys using first-principles calculations and correlational analysis, Comput. Mater. Sci. 212 (2022) 111564.
- S. Lin, S. Shang, J.D. Shimanek, Y. Wang, A.M. Beese, Z. Liu, Predicting ideal shear strength of dilute multicomponent Ni-based alloys by an integrated first-principles calculations, CALPAHD modeling approach and correlation analysis, (2023) To be published and attached in this final report.

3.4 DFT-informed FEM simulations

To reduce reliance on experimental fitting data within the crystal plasticity finite element method (CPFEM), an approach is proposed that integrates first-principles calculations based on density functional theory (DFT) to predict the strainhardening behavior of pure Ni single crystals; see Figure 9. Flow resistance was evaluated through the Peierls–Nabarro equation using the ideal shear strength and elastic properties calculated by DFT-based methods, with hardening behavior modeled by imposing strains on supercells in first-principles calculations. Considered alone, elastic interactions of pure edge dislocations capture hardening behavior for small strains on single-slip systems. For larger strains, hardening is captured through a strain-weighted linear combination of edge and screw flow resistance components. The rate of combination is not predicted in the present framework, but agreement with experiments through large strains (~0.4) for multiple loading orientations demonstrates a possible route for more predictive crystal plasticity modeling through incorporation of analytical models of mesoscale physics.

In addition to pure Ni, we also predicted the strain-stress curves of dilute $Ni_{11}X$ alloys using the same method, where the alloying elements X = Fe, Co, Cr, Al, V, Ti, and Nb [7].

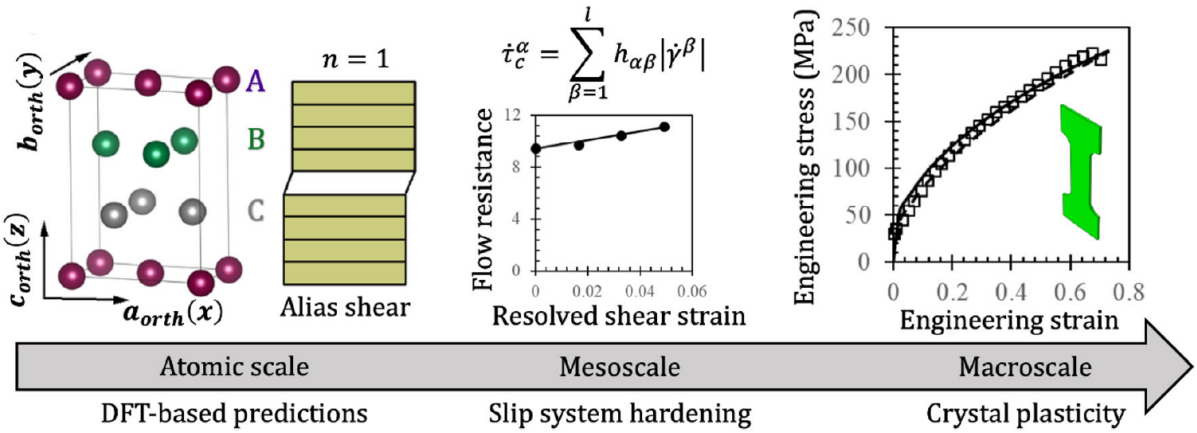


Figure 9. A schematic of the overall approach proposed in the current work, showing the transfer of information from the atomic scale ideal shear process to a mesoscale description of hardening on a slip system level to, finally, a description of macroscale deformation of single crystal samples.

We also used a full-field crystal plasticity model with a DFT-informed dislocation density (DD) hardening law to identify the key microstructural features correlated with micromechanical fields localization, or hotspots, in polycrystalline Ni. An ensemble learning approach to machine learning

interpreted with Shapley additive explanation was implemented to predict nonlinear correlations between microstructural features and micromechanical stress and strain hotspots; see Figure 10.

Results reveal that regions within the microstructure in the vicinity of the grain boundaries, higher Taylor and Schmid factors, and high intergranular misorientations, are more prone to being micromechanical hotspots. Additionally, under combined loading, intergranular misorientations are more responsible than Schmid factor in formation of stress hotspots while Schmid factors take precedence under high plastic strain localizations. The present work demonstrates a successful integration of physics-based crystal plasticity with DD-based hardening into machine learning models to reveal the microscale features responsible for the formation of local stress and strain hotspots within the grains and near the grain boundaries, as function of applied deformation states, grain morphology/size distribution, and microstructural texture, providing insights into micromechanical damage initiation zones in polycrystalline metals (Figure 10).

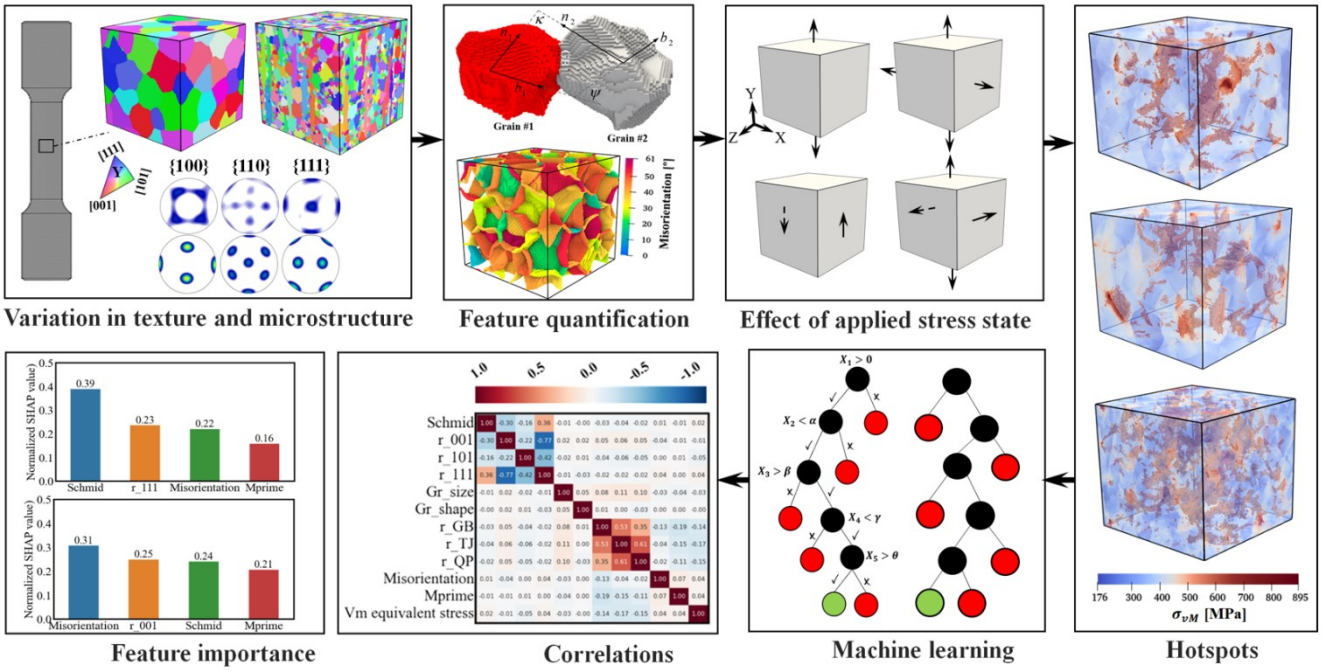


Figure 10. A schematic of the overall approach of physics-based crystal plasticity with DD-based hardening into machine learning models to reveal the microscale features responsible for the formation of local stress and strain hotspots within the grains and near the grain boundaries.

Currently we published three papers and one will be published soon to demonstrate the DFT-informed FEM simulations.

- J.D. Shimanek, S.-L. Shang, A.M. Beese, Z.-K. Liu, Insight into ideal shear strength of Ni-based dilute alloys using first-principles calculations and correlational analysis, Comput. Mater. Sci. 212 (2022) 111564.
- J.D. Shimanek, S. Qin, S.-L. Shang, Z.-K. Liu, A.M. Beese, Predictive Crystal Plasticity Modeling of Single Crystal Nickel Based on First-Principles Calculations, JOM. 74 (2022) 1423–1434.
- A. Eghtesad, J.D. Shimanek, S.-L. Shang, R. Lebensohn, M. Knezevic, Z.-K. Liu, A.M. Beese, Density functional theory-informed dislocation density hardening within crystal plasticity: Application to modeling deformation of Ni polycrystals, Comput. Mater. Sci. 215 (2022) 111803.
- A. Eghtesad, Q. Luo, S. Shang, R. Lebensohn, M. Knezevic, Z. Liu, A.M. Beese, Machine learning-enabled identification of micromechanical stress and strain hotspots predicted via dislocation density-based crystal plasticity simulations, (2023) To be published.

2. GRAPHICAL MATERIALS LIST(S)

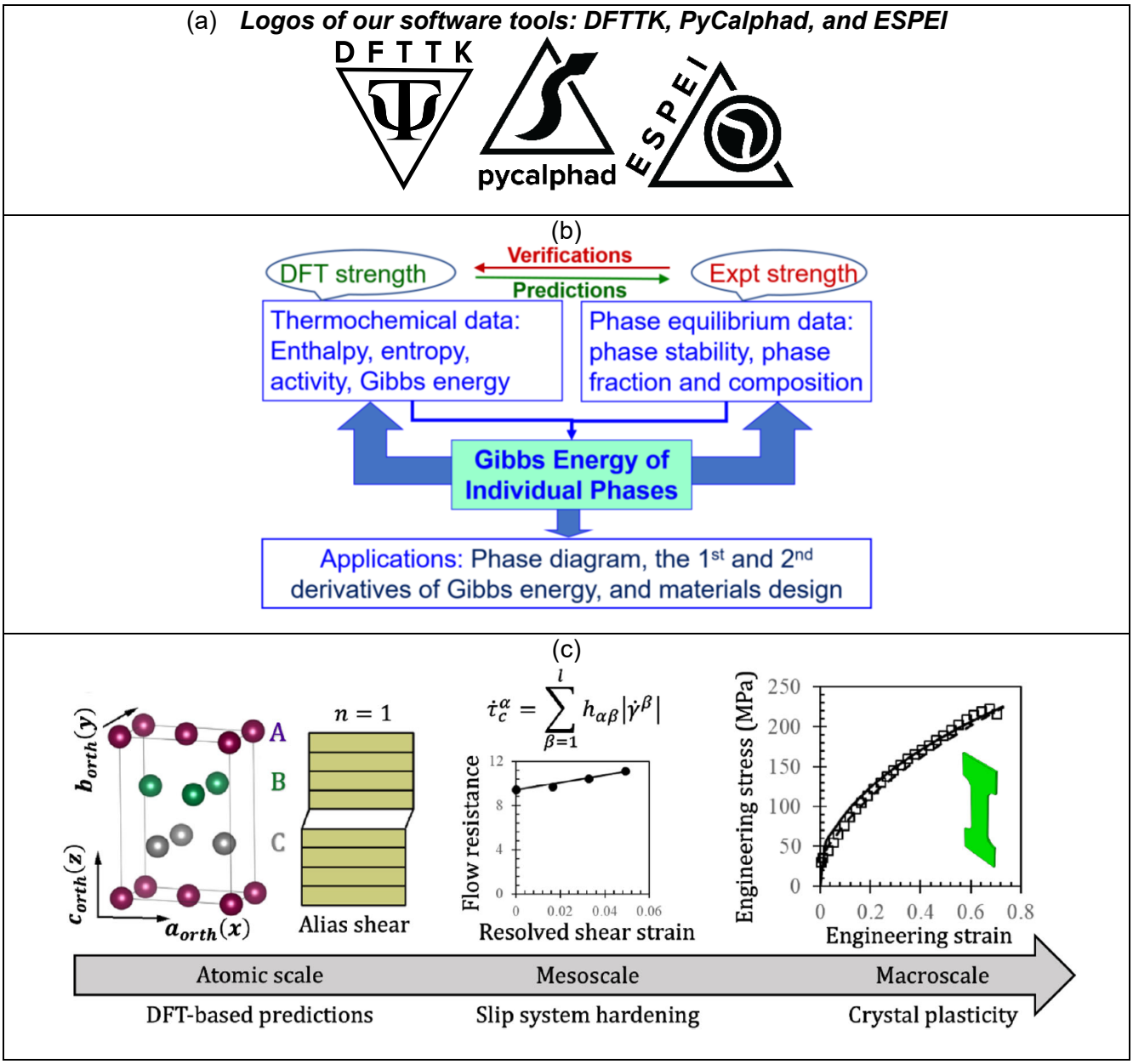


Figure 11. Key graphical materials for this project (a) the logos of our software tools: DFTTK, Pycalphad, and ESPEI, (b) the idea of CALPHAD modeling, and (c) the DFT-informed FEM simulations.

3. LIST OF KEY ACRONYMS AND ABBREVIATIONS

ABBREVIATIONS	EXPLANATIONS
CALPHAD	Calculations of phase diagram
DFT	Density functional theory
DFTTK	Density functional theory based tool kit

ESPEI	Extensible, self-optimizing phase equilibrium computer program
PyCalphad	Python based CALPHAD modeling code
SIPFENN	Extensible structure-informed prediction of formation energy by maching learning
FEM	Finite element method
CPFEM	Crystal plasticity finite element simulations

4. PRODUCTS

4.1. Two Ph.D. students graduated with two theses published

- Shipin Qin, "Influence of microstructure on the multiaxial plasticity and fracture of dual phase steels: experiments and multiscale computational modeling", Pennsylvania State University, Ph.D. thesis, 2020. https://etda.libraries.psu.edu/files/final_submissions/21517

- Brandon Bocklund, "Computational design of additively manufactured functionally graded materials by thermodynamic modeling with uncertainty quantification", Pennsylvania State University, Ph.D. thesis, 2021. https://etda.libraries.psu.edu/catalog/21192bjb54

4.2. Three Ph.D. students and one postdoc partially supported

- Shuang Lin (Ph.D. student, to be graduated within 2 years)
- Hui Sun (Ph.D. student, to be graduated within 1 year)
- John Shimanek (Ph.D. student, to be graduated within 1 year)
- Adnan Eghtesad (Postdoc)

4.3. Twenty publications (17 papers published and 3 to be published and attached)

- Fundamentals of thermodynamics (3 papers published)
  - Review of thermodynamics and its applications in 2020 by Liu [23]
  - Entropy and critical phenomena in 2019 by Liu et al. [24]
  - Theory of cross phenomena in 2022 by Liu (cited this project but indirectly) [25]
- Software tools development (3 papers published or to be published)
  - DFTTK in 2019 by Wang et al. [1]
  - SIPFENN in 2022 by Krajewski et al. [2]
  - ESPEI in 2022 by Hui et al. (to be published, attached) [10]
- Thermodynamic databases development (3 papers published or to be published)
  - DFT-based database in 2022 by Wang et al. [26]
  - CALPHAD-based thermodynamic modelings of the Ni-Hf [21] and Nb-Ni [10] systems
- Applications of thermodynamics and kinetics (1 paper published)
  - Predict critical Al concentration to form external Al_2O_3 scale in Ni-Al alloy in 2022 by Ross et al. [27]
- Ideal shear deformation and associated properties (e.g., input for DFT-informed FEM, 3 papers published or to be published)
  - Ni_3Al in 2020 by Shang et al. [28]
  - Dilute Ni-X alloys in 2022 by Shimanek et al. [7]
  - Dilute Ni-X-Z by Lin et al. (to be published, attached) [29]
- DFT-informed FEM (or ML-informed FEM) (4 papers published or to be published)

    - Single crystal Ni [6] and dilute Ni-X alloys [7]
    - Ni polycrystal [8]
    - Hotspots by ML (to be published, attached) [5]
- Machine learning with input data from DFT-based calculations (2 papers published or to be published)
    - Stacking fault energy [4]
    - Hotspots by ML (to be published, attached) [5]
- Experimental work of AM (1 paper published)
    - Inconel 625 manufactured by laser powder bed fusion [30]
- Other works indirectely related to this project (3 papers published)
    - Partially supported by this project to study strain-controlled antiferromagnetic memory [31]; to study synergetic effects of solute and strain in biocompatible Zn-based and Mg-based alloys [32]; and to study zentropy theory for positive and negative thermal expansion [33]

4.4. Websites related to the results of this project

- http://espei.org for ESPEI (Extensible Self-optimizing Phase Equilibria Infrastructure): A tool for automated thermodynamic database development within the CALPHAD method.
- https://pycalphad.org for *pycalphad* code: A Python library for computational thermodynamics using the CALPHAD method.
- https://www.dfttk.org for DFTTK: Density functional theory (DFT) workflows for finite temperature thermodynamics based on the *atomate* (atomate.org) workflows.
- https://www.phaseslab.com/sipfenn for SIPFENN: Structure-Informed Prediction of Formation Energy using Neural Networks.
- https://www.phaseslab.com/MPDD for the Material-Property-Descriptor Database.

5. REFERENCES

[1] Y. Wang, M. Liao, B.J. Bocklund, P. Gao, S.-L. Shang, H. Kim, A.M. Beese, L.-Q. Chen, Z.-K. Liu, DFTTK: Density Functional Theory ToolKit for high-throughput lattice dynamics calculations, Calphad. 75 (2021) 102355. https://doi.org/10.1016/j.calphad.2021.102355.

[2] A.M. Krajewski, J.W. Siegel, J. Xu, Z.-K. Liu, Extensible Structure-Informed Prediction of Formation Energy with improved accuracy and usability employing neural networks, Comput. Mater. Sci. 208 (2022) 111254. https://doi.org/10.1016/j.commatsci.2022.111254.

[3] R.A. Otis, Z.-K. Liu, pycalphad: CALPHAD-based Computational Thermodynamics in Python, J. Open Res. Softw. 5 (2017) 1. https://doi.org/10.5334/jors.140.

[4] X. Chong, S.-L. Shang, A.M. Krajewski, J.D. Shimanek, W. Du, Y. Wang, J. Feng, D. Shin, A.M. Beese, Z.-K. Liu, Correlation analysis of materials properties by machine learning: illustrated with stacking fault energy from first-principles calculations in dilute fcc-based alloys, J. Phys. Condens. Matter. 33 (2021) 295702. https://doi.org/10.1088/1361-648X/ac0195.

[5] A. Eghtesad, Q. Luo, S. Shang, R. Lebensohn, M. Knezevic, Z. Liu, A.M. Beese, Machine learning-enabled identification of micromechanical stress and strain hotspots predicted via dislocation density-based crystal plasticity simulations, (2023) To be published.

[6] J.D. Shimanek, S. Qin, S.-L. Shang, Z.-K. Liu, A.M. Beese, Predictive Crystal Plasticity Modeling of Single Crystal Nickel Based on First-Principles Calculations, JOM. 74 (2022) 1423–1434. https://doi.org/10.1007/s11837-022-05175-6.
[7] J.D. Shimanek, S.-L. Shang, A.M. Beese, Z.-K. Liu, Insight into ideal shear strength of Ni-based dilute alloys using first-principles calculations and correlational analysis, Comput. Mater. Sci. 212 (2022) 111564. https://doi.org/10.1016/j.commatsci.2022.111564.
[8] A. Eghtesad, J.D. Shimanek, S.-L. Shang, R. Lebensohn, M. Knezevic, Z.-K. Liu, A.M. Beese, Density functional theory-informed dislocation density hardening within crystal plasticity: Application to modeling deformation of Ni polycrystals, Comput. Mater. Sci. 215 (2022) 111803. https://doi.org/10.1016/j.commatsci.2022.111803.
[9] J.-M. Joubert, Y. Feutelais, Contribution of the Rietveld method to non-stoichiometric phase modeling. Part II: α-Tl5Te3 and μ Nb-Ni as experimental examples, Calphad. 26 (2002) 427–438.
[10] H. Sun, S.-L. Shang, R. Gong, B.J. Bocklund, A.M. Beese, Z.-K. Liu, Thermodynamic modeling with uncertainty quantification in the Nb-Ni system using the upgraded PyCalphad and ESPEI, ArXiv. (2022). https://doi.org/10.48550/arxiv.2204.11813.
[11] H. Chen, Y. Du, Refinement of the thermodynamic modeling of the Nb – Ni system, 30 (2006) 308–315. https://doi.org/10.1016/j.calphad.2006.02.005.
[12] Python-based open source software developed at Pennsylvania State University, ESPEI (http://espei.org), PyCalphad (http://pycalphad.org), SIPFENN (https://www.phaseslab.com/sipfenn), and DFTTK (http://github.com/PhasesResearchLab) based on atomate (https://github.com/hackingmaterials/atomate), (n.d.).
[13] Y. Wang, L.-Q. Chen, Z.-K. Liu, YPHON: A package for calculating phonons of polar materials, Comput. Phys. Commun. 185 (2014) 2950–2968. https://doi.org/10.1016/j.cpc.2014.06.023.
[14] C.-W. Yang, D.B. Williams, J.I. Goldstein, A revision of the Fe-Ni phase diagram at low temperatures (<400 °C), J. Phase Equilibria. 17 (1996) 522–531. https://doi.org/10.1007/BF02665999.
[15] J. Yang, J.I. Goldstein, The formation of the Widmanstätten structure in meteorites, Meteorit. Planet. Sci. 40 (2005) 239–253. https://doi.org/10.1111/j.1945-5100.2005.tb00378.x.
[16] A. Oradei-Basile, J.F. Radavich, A Current T-T-T Diagram for Wrought Alloy 718, in: E. A Loria (Ed.), The Minerals, Metals & Materials Society, Warrendale, PA, 1991: pp. 325–335.
[17] K.C. Mills, Recommended values of thermophysical properties for selected commercial alloys, Woodhead, Cambridge, England, 2002.
[18] H. Hosaeus, A. Seifter, E. Kaschnitz, G. Pottlacher, Thermophysical properties of solid and liquid Inconel 718 alloy, High Temp. High Press. 33 (2001) 405–410.
[19] S. Shang, ESPEI : Extensible , Self-optimizing Phase Equilibrium Infrastructure for Magnesium Alloys ESPEI : Extensible , Self-optimizing Phase Equilibrium Infrastructure for Magnesium Alloys, (2010).
[20] J.-O. Andersson, T. Helander, L. Höglund, P. Shi, B. Sundman, Thermo-Calc & DICTRA: Computational tools for materials science, Calphad. 26 (2002) 273–312. https://doi.org/10.1016/S0364-5916(02)00037-8.
[21] A.J. Ross, T. Gheno, P.K. Ray, M.J. Kramer, X.L. Liu, G. Lindwall, B. Zhou, S.L. Shang, B. Gleeson, Z.-K. Liu, A first-principles based description of the Hf-Ni system supported by high-temperature synchrotron experiments, Thermochim. Acta. 668 (2018) 142–151. https://doi.org/10.1016/J.TCA.2018.08.011.
[22] B. Joós, M.S. Duesbery, The Peierls Stress of Dislocations: An Analytic Formula, Phys. Rev. Lett. 78 (1997) 266–269. https://doi.org/10.1103/PhysRevLett.78.266.
[23] Z.-K. Liu, Computational thermodynamics and its applications, Acta Mater. 200 (2020) 745–792. https://doi.org/10.1016/j.actamat.2020.08.008.
[24] Z.-K. Liu, B. Li, H. Lin, Multiscale Entropy and Its Implications to Critical Phenomena, Emergent Behaviors, and Information, J. Phase Equilibria Diffus. 40 (2019) 508–521.

https://doi.org/10.1007/s11669-019-00736-w.
[25] Z.-K. Liu, Theory of cross phenomena and their coefficients beyond Onsager theorem, Mater. Res. Lett. 10 (2022) 393–439. https://doi.org/10.1080/21663831.2022.2054668.
[26] Y. Wang, F. Lia, K. Wang, K. McNamara, Y. Ji, X. Chong, S.-L. Shang, Z.-K. Liu, R.P. Martukanitz, L.-Q. Chen, A thermochemical database from high-throughput first-principles calculations and its application to analyzing phase evolution in AM-fabricated IN718, Acta Mater. 240 (2022) 118331. https://doi.org/10.1016/j.actamat.2022.118331.
[27] A. Ross, S. Shang, H. Fang, G. Lindwall, X.L. Liu, W. Zhao, B. Gleeson, M.C. Gao, Z. Liu, Tailoring critical Al concentration to form external Al2O3 scale on Ni–Al alloys by computational approach, J. Am. Ceram. Soc. 105 (2022) 7770–7777. https://doi.org/10.1111/jace.18707.
[28] S.L. Shang, J. Shimanek, S. Qin, Y. Wang, A.M. Beese, Z.K. Liu, Unveiling dislocation characteristics in Ni3Al from stacking fault energy and ideal strength: A first-principles study via pure alias shear deformation, Phys. Rev. B. 101 (2020) 024102. https://doi.org/10.1103/PhysRevB.101.024102.
[29] S. Lin, S. Shang, J.D. Shimanek, Y. Wang, A.M. Beese, Z. Liu, Predicting ideal shear strength of dilute multicomponent Ni-based alloys by an integrated first-principles calculations, CALPAHD modeling approach and correlation analysis, (2023) To be published.
[30] S. Qin, T.C. Novak, M.K. Vailhe, Z.-K. Liu, A.M. Beese, Plasticity and fracture behavior of Inconel 625 manufactured by laser powder bed fusion: Comparison between as-built and stress relieved conditions, Mater. Sci. Eng. A. 806 (2021) 140808. https://doi.org/10.1016/j.msea.2021.140808.
[31] H. Yan, Z. Feng, S. Shang, X. Wang, Z. Hu, J. Wang, Z. Zhu, H. Wang, Z. Chen, H. Hua, W. Lu, J. Wang, P. Qin, H. Guo, X. Zhou, Z. Leng, Z. Liu, C. Jiang, M. Coey, Z. Liu, A piezoelectric, strain-controlled antiferromagnetic memory insensitive to magnetic fields, Nat. Nanotechnol. 14 (2019). https://doi.org/10.1038/s41565-018-0339-0.
[32] Y.Q. Guo, S.H. Zhang, I.J. Beyerlein, D. Legut, S.L. Shang, Z.K. Liu, R.F. Zhang, Synergetic effects of solute and strain in biocompatible Zn-based and Mg-based alloys, Acta Mater. 181 (2019). https://doi.org/10.1016/j.actamat.2019.09.059.
[33] Z.-K. Liu, Y. Wang, S.-L. Shang, Zentropy Theory for Positive and Negative Thermal Expansion, J. Phase Equilibria Diffus. (2022) In press. https://doi.org/10.1007/s11669-022-00942-z.

6. APPENDICES

Three to be published papers are attached in the following pages, including:

- H. Sun, S.-L. Shang, R. Gong, B.J. Bocklund, A.M. Beese, Z.-K. Liu, “Thermodynamic modeling with uncertainty quantification in the Nb-Ni system using the upgraded PyCalphad and ESPEI”. (2023) To be published.

- S. Lin, S. Shang, J.D. Shimanek, Y. Wang, A.M. Beese, Z. Liu, Predicting ideal shear strength of dilute multicomponent Ni-based alloys by an integrated first-principles calculations, CALPAHD modeling approach and correlation analysis, (2023) To be published.

- Eghtesad, Q. Luo, S. Shang, R. Lebensohn, M. Knezevic, Z. Liu, A.M. Beese, “Machine learning-enabled identification of micromechanical stress and strain hotspots predicted via dislocation density-based crystal plasticity simulations”, (2023) To be published.

Thermodynamic modeling with uncertainty quantification in the Nb-Ni system using the upgraded PyCalphad and ESPEI

Hui Sun[a], Shun-Li Shang[a], Rushi Gong[a], Brandon J. Bocklund[a,b], Allison M. Beese[a], and Zi-Kui Liu[a]

a Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA, 16802, USA

b Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Abstract

The Nb-Ni system is remodeled with uncertainty quantification (UQ) using the upgraded software tools of PyCalphad and ESPEI with new capability by modeling site occupancy of Wyckoff position for the phases of interest. Specifically, the five- and three-sublattice models are used to model the topologically close pack (TCP) phases of μ-Nb_7Ni_6 and δ-$NbNi_3$ according exactly to their Wyckoff positions. The inputs for CALPHAD-based thermodynamic modeling include the predicted thermochemical data as a function of temperature by density functional theory (DFT) based first-principles and phonon calculations together with phase equilibrium and site occupancy data in the literature. In addition to phase diagram and thermodynamic properties, the CALPHAD predictions of site occupancies agree well with experimental data, namely the measured site occupancies of Nb in μ-Nb_7Ni_6. In addition, the UQ values estimated using the Markov Chain Monte Carlo (MCMC) method as implemented in ESPEI make it possible to quantify uncertainties in the Nb-Ni system, for example, site occupancy in μ-Nb_7Ni_6 and enthalpy of mixing in liquid.

Highlights

- New capability implemented into PyCalphad and ESPEI to model site occupancy
- TCP phases (μ-Nb_7Ni_6 and δ-$NbNi_3$) modeled well using sublattice models according exactly to their Wyckoff positions
- The Nb-Ni system remodeled with uncertainty quantification
- Finite-temperature thermochemical properties predicted by DFT-based first-principles and phonon calculations

Keywords: CALPHAD modeling; Nb-Ni; PyCalphad and ESPEI; First-principles and phonon calculations; Site occupancy; TCP phases; Uncertainty quantification.

1 Introduction

The topologically close pack (TCP) phases, also known as the Frank-Kasper phases [1], are intermetallic compounds with complex crystalline structures, which are frequently observed in Ni-based superalloys, for example, the $\sigma, \chi, P, R, \delta, \mu, M$, A15, and Laves phases [2]. TCP phases are usually brittle and detrimental, and hence, their type, amount, and distribution are of great importance for the performance of Ni-based superalloys [3]. For example, when Ni-based superalloys are highly alloyed with refractory elements (e.g., Cr, Mo, Nb, Ta, W, and Re) to achieve better strengths at high temperatures, the TCP phases will be formed and produce detrimental effects on the properties of superalloys by draining refractory elements from the matrix to reduce the solid solution strengthening like the γ phase with the FCC lattice in Ni-based superalloys [4] or the strengthening phase like the γ' phase with the $L1_2$ lattice in Co-based superalloys [5].

Relevant to the present work, the TCP phases in the Nb-Ni system are δ-$NbNi_3$ and μ-Nb_7Ni_6, which are deleterious to the performance of Ni-based superalloys since their formations will drain alloying element Nb from the matrix [6]. For example, the formation of a 10% volume fraction of δ-$NbNi_3$ can lower 40% of elongation in Ni-based superalloy of Inconel 718 [7]. The formation of around 10 vol.% of δ-$NbNi_3$ after stress relief heat treatment resulted in a 9% increase of the ultimate tensile strength but a 45% decrease of the fracture strain in Inconel 625 [8]. Similarly, μ-Nb_7Ni_6 also shows an undesirable influence on mechanical properties, for example, the precipitation of μ-Nb_7Ni_6 in the Ni-Mo-Cr alloy decreased its room temperature impact roughness by 100 joules in the Charpy test [9]. Therefore, an accurate thermodynamic description of the TCP phases in the Nb-Ni system is critical for a better understanding of their formations through thermodynamic calculations, so that the TCP phases can be avoided through designing chemistry and/or manufacturing process of superalloys.

In general, the TCP compounds are complex intermetallic phases with different elements in one or more of their Wyckoff sites. For example, δ-$NbNi_3$ includes three Wyckoff sites (2a, 2b, and 4e) with space group *Pmmn* (No. 59) [10] as shown in Table 1, while μ-Nb_7Ni_6 has five Wyckoff sites (3a, 6c(1), 6c(2), 6c(3), and 18h) with space group $R\bar{3}m$ (No. 166) [11] as shown in Table 1.

The crystallographic information of Wyckoff sites indicates that a three-sublattice model is expected for a complete description of δ-$NbNi_3$ and a five-sublattice model for μ-Nb_7Ni_6. However, the previous CALPHAD modeling works of Nb-Ni [12–15] cannot capture well the change of site occupancies in TCP phases, i.e., the occurrence of a given kind of atoms (or vacancy) in a given Wyckoff position; for example, Nb in Wyckoff sites of μ-Nb_7Ni_6 measured by Joubert et al. [16], since μ-Nb_7Ni_6 was described by the simplified sublattice models like $(Ni)_{0.47}(Nb)_{0.53}$[17], $(Nb, Ni)_7(Nb)_6$ [13], and $(Nb, Ni)_1Ni_4(Nb, Ni)_2Nb_6$ [14]; see details in Table 2 which summarizes the previous thermodynamic models for μ-Nb_7Ni_6 and δ-$NbNi_3$. It shows that Kaufman and Nesor [17] modeled μ-Nb_7Ni_6 and δ-$NbNi_3$ as stoichiometric compounds using $(Ni)_{0.47}$ $(Nb)_{0.53}$ for μ-Nb_7Ni_6, $(Ni)_{0.75}(Nb)_{0.25}$ for δ-$NbNi_3$, while Kejun et al. [14] used a four-sublattice model $(Nb, Ni)_1Ni_4(Nb, Ni)_2Nb_6$ for μ-Nb_7Ni_6 and a two-sublattice model $(Nb, Ni)_3(Nb, Ni)_1$ for δ-$NbNi_3$. Bolcavage and Kattner [13] used the model $(Nb, Ni)_7(Nb)_6$ for μ-Nb_7Ni_6. Note that Joubert et al. [18] adopted a five-sublattice model for μ-Nb_7Ni_6, i.e., $(Nb, Ni)_1Nb_2Nb_2(Nb, Ni)_2(Nb, Ni)_6$, which gives a better description of solubility and site occupancy in μ-Nb_7Ni_6. However, two of the sublattices in Joubert et al.'s model [18] include only one element Nb instead of two elements of Nb and Ni, in which the absence of all possible elements limits its model for further application in higher order systems. Most recently, Chen et al. [15] remodeled the Nb-Ni system with the $(Nb, Ni)_1Nb_4(Nb, Ni)_2(Nb, Ni)_6$ model for μ-Nb_7Ni_6 and the $(Nb, Ni)_3(Nb, Ni)_1$ model for δ-$NbNi_3$. Table 2 summarizes that none of the previous thermodynamic models of μ-Nb_7Ni_6 were based exactly on Wyckoff positions, resulting in a discrepancy regarding site occupancy between CALPHAD modeled results and experiments. The less accurate descriptions of TCP phases, especially their site occupancies, in the Nb-Ni system motivate the present CALPHAD remodeling. The sublattice model with all possible elements in each sublattice is needed to accurately describe the distribution of elements in different sublattices, which will be used to compare with other models and experimental data, and further benefit the model of multi-component systems.

It is worth mentioning that an arbitrary value of 5000 J/mol-atom was assigned as the enthalpy of formation for nonstable endmember compounds of TCP phases in Nb-Ni by Chen et al. [19], which is less reliable. To address this issue, in the present work, density functional theory (DFT) based

first-principles calculations (and phonon calculations for stable endmembers) are used to determine thermodynamic properties at 0 K (or as a function of temperature for stable endmembers) for all endmembers of μ-Nb_7Ni_6 and δ-$NbNi_3$.

In summary, the present work aims to remodel the Nb-Ni system in terms of the CALPHAD approach with UQ using the open-source tools of ESPEI (the Extensible, Self-optimizing Phase Equilibria Infrastructure) for database development [20], and PyCalphad for equilibrium thermodynamic calculations [21]. Here, the input for CALPHAD modeling includes the present first-principles and phonon calculations, and experimental data in the literature with an emphasis on site occupancy data. Note that ESPEI has been upgraded in the present work, making it possible to model site occupancies with UQ for Wyckoff sites.

2 Literature Review of Thermodynamic properties in Nb-Ni

2.1 The previous CALPHAD modeling

The Nb-Ni system has six phases including three solution phases (BCC, FCC, and liquid) and three intermetallic compounds (μ-Nb_7Ni_6, δ-$NbNi_3$, and $NbNi_8$) based on the summary by Chen et al. [19]. The Nb-Ni system has been modeled several times in terms of the CALPHAD approach [12–15]. Kaufman and Nesor [17] provided a modeled Nb-Ni phase diagram by considering μ-Nb_7Ni_6 and δ-$NbNi_3$ as stoichiometric compounds as shown in Table 2, which didn't match with experimental observations of Nb solubilities in these two phases measured by Murametsu et al. [22], Duerden et al. [23], and Chen et al. [24]. Kejun et al.'s modeling work [14] adopted the model of $(Nb, Ni)_1Ni_4(Nb, Ni)_2Nb_6$ for μ-Nb_7Ni_6, which cannot describe well the solubilities of Nb in the composition range of 50 – 54 at. % Nb around 1100℃ measured by Duerden et al. [23]. In addition, the enthalpies of formation of μ-Nb_7Ni_6 from Kejun et al.'s modeling show a large discrepancy around 13 kJ/mol-atom compared with experiments by Argent et al. [25]. In the modeling work by Bolcavage and Kattner [13], they did not consider the $NbNi_8$ compound because of the lacking of experimental data at that time. The predicted liquidus in the Nb-rich region is higher around 200 – 300 K than experimental data from Wicker et al. [26]. Joubert et al. [18] considered $NbNi_8$ in their modeling work. However, an arbitrary value of 5000 J/mol-atom was used to describe the enthalpy of formation for the pure element endmembers in the TCP phases.

This value is too positive to describe correctly the partially occupied Wyckoff sites in μ-Nb_7Ni_6 in comparison with experimental data by Joubert et al. [16].

The most recent modeling work was done by Chen et al. [19], which was widely used in thermodynamic modeling of ternary systems such as Fe-Nb-Ni [27], Nb-Ni-Zr [28], and Nb-Ni-Ti [29]. Chen et al.'s work [19] contains the $NbNi_8$ phase but fails to depict accurate phase boundaries between μ-Nb_7Ni_6 and δ-$NbNi_3$ and between μ-Nb_7Ni_6 and BCC measured by Murametsu et al. [22]. In addition, Chen et al.'s modeling didn't consider experimental liquidus temperatures with respect to BCC by Wicker et al. [26] and the enthalpies of mixing for the liquid at 1823 K by Chistyakov et al. [30].

Due to the limitation of the aforementioned sublattice models of TCP phases to describe well phase boundaries and thermochemical properties in the Nb-Ni system, an appropriate sublattice model derived from crystallographic information is needed to precisely describe the TCP phases. To this end, the sublattice model of $(Nb, Ni)_1(Nb, Ni)_1(Nb, Ni)_2$ is adopted to model δ-$NbNi_3$ and $(Nb, Ni)_1(Nb, Ni)_2(Nb, Ni)_2(Nb, Ni)_2(Nb, Ni)_6$ is used to describe μ-Nb_7Ni_6 in the present work, corresponding exactly to their Wyckoff positions as shown in Table 1.

2.1 Thermodynamic properties

2.1.1 Phase diagram data

Phase boundaries between FCC and liquid (0 – 15 at. % Nb) were measured using thermal analysis via heating curves by Duerden et al. [23], Pogodin et al. [28], and Grube et al. [32]; using differential thermal analysis (DTA) by Chen et al. [24] and Kajikawa [33]; and using the solid-liquid diffusion couple method (DCM) by Kajikawa [33]. All these measurements show good agreement, e.g., the temperature variation at each fixed composition is less than 40 K. All these data are hence used in the present CALPHAD modeling of the Nb-Ni system.

Phase boundaries between FCC and δ-$NbNi_3$ (0 – 15 at. % Nb) were measured by Pogodin et al. [31] and Grube et al. [32] by thermal analysis from heating curves. Chen et al. [24] used DFT results to analyze the FCC to δ-$NbNi_3$ transition at 1322 K, while Guseva et al. [34] detected the

FCC to δ-$NbNi_3$ transition using X-ray powder diffraction (XRD) at 1073 – 1473 K. Joubert et al. [18] assessed the homogenous regions of δ-$NbNi_3$ using electron probe micro-analysis (EPMA). All these measurements show good agreement with each other, e.g., the composition change is only 5 at. % Nb from 1000 K to 1500 K, and we hence use all these data in the present CALPHAD modeling.

The $NbNi_8$ phase was first observed using transmission electron microscopy (TEM) by Quist et al. in 1969 [35]. It was confirmed by Joubert et al. [18] that $NbNi_8$ is a stable phase by examining samples annealed at 723 K for 76 days using XRD. Wekken et al. [36] used the changes in electrical resistivity to detect the existence of $NbNi_8$, showing that $NbNi_8$ forms at 10.3 at. % Nb at 853 K. Chen et al. [24] also detected $NbNi_8$ with differential scanning calorimetry (DSC). Therefore, $NbNi_8$ is considered as a stable phase in the present work.

The Nb_2Ni phase was observed by Zhao et al. [37] by using the TEM method for sample annealing at 1523 K for 5h. However, this phase was not confirmed further using samples with a longer annealing time. The Nb_2Ni phase is hence excluded in the present modeling work.

Regarding solubility of Nb in δ-$NbNi_3$, Murametsu et al. [22] observed 24.0 – 26.6 at. % Nb in δ-$NbNi_3$ in the temperature range of 1023 K – 1303 K by EPMA. Chen et al. [24] reported the phase boundary around 23.4 – 25.7 at. % Nb by DTA at 1323 K for 336h. Duerden et al. [23] estimated the phase boundary around 23.5 – 26.5 at. % Nb using XRD at 1273 K. The phase boundaries of δ-$NbNi_3$ between δ-$NbNi_3$ and liquid were measured by Grube et al. [32], Duerden et al. [20], and Svechnikov et al. [38] using heating curves of thermal analysis, and by Chen et al. [24] using DTA. All these data are included in the present CALPHAD modeling.

Regarding solubility of Nb in μ-Nb_7Ni_6, Duerden et al. [23] estimated 50 – 54 at. % Nb in μ-Nb_7Ni_6 around 1373 K using the optical microscopy method. Svechnikov et al. [38] reported 49.8 - 58.3 at. % Nb in μ-Nb_7Ni_6 by using the heating curves of thermal analysis. Murametsu et al. [22] reported 48.6 - 56.2 at. % Nb in μ-Nb_7Ni_6 around 1023 K – 1303 K by EPMA. Joubert et al. [16] reported 49.6 - 56.9 at. % Nb in μ-Nb_7Ni_6 at 1273 K by EPMA. Chen et al. [24] estimated 49.5 – 56.3 at. % Nb in μ-Nb_7Ni_6 around 1273 K – 1303 K by EPMA. The phase boundaries of μ-Nb_7Ni_6

between μ-Nb_7Ni_6 and liquid were measured by Duerden et al. [23] and Svechnikov et al. [38] using the heating curves of thermal analysis and by Chen et al. [24] using DTA. All these data are included in the present CALPHAD modeling.

The measured temperatures of invariant reactions between liquid, δ-$NbNi_3$, and FCC phase agree well with each other from 1170 K to 1175 K by Chen et al. [24] using DTA, and by Duerden et al. [28] and Svechnikov et al. [39] using the heating curves of thermal analysis. However, the invariant temperatures between liquid, δ-$NbNi_3$, and μ-Nb_7Ni_6 have large uncertainties (from 1290 to 1320 K) measured by Duerden et al. [28] and Wicker et al. [26]. Nevertheless, all these data are included in the present CALPHAD modeling.

For phase boundaries between liquid and BCC, Svechnikov et al. [39] and Wicker et al. [26] measured them by quenching the samples, and Duerden et al. [23] measured them by heating, cooling, and quenching. Albeit these measurements exhibit noticeable discrepancies, as much as 300 K, the present modeling work considers all these experimental data.

2.1.2 Thermochemical data

The enthalpies of formation for the Nb-Ni system were measured by Argent et al. [25] through the calorimetry method with an error around 4 kJ/mol-atom from 12.5 – 75.0 at. % Nb. On the other hand, electromotive force (emf) measurements were adopted by Sokolvskaya et al. [40], Alekseev et al. [34], and Lyakishev et al. [42] to determine the enthalpies of formation for intermediate phases at 25.0 at. % Nb and 50.0 at. % Nb. However, the results from Alekseev et al. [41] and Sokolvskaya et al. [40] show great discrepancies around 8 kJ/mole-atom at 25.0 at. % Nb with respect to those from Argent et al. [25] as shown in Fig. 1. Compared with the DFT results from the Materials Project [43] and the Open Quantum Materials Database (OQMD) [44] as shown in Fig. 1, the DFT calculations using the generalized gradient approximation (GGA) are higher than experiments data by around 10%; agreeing with the general trends of higher enthalpies of formation from DFT-based calculations in comparison with those from experimental measurements [45]. The enthalpies of formation data from Argent et al. [25] are higher by around 3.4 kJ/mol-atom for compositions of 25.0 at. % Nb, but still compatible with the results from DFT-

based predictions [45]. Therefore, the data from Argent et al. [25] are more reliable than those from Sokolvskaya et al. [40] and hence adopted in the present CALPHAD modeling.

Two sets of measurements about enthalpy of mixing were available for the Ni-rich liquid phase. As shown in Fig. 2, the data from Schaefers et al. [46] at 1927 K and 2000 K show a larger difference around 15 kJ/mol-atom at 30.0 at. % Nb compared with those from Chistyakov et al. [30] at 2096 K and Sudavtsova et al. [40] at 2148 K. In general, the enthalpies of mixing for liquids and solids should be compatible with each other in the same alloy system. For example, in the Al-Cu system [47] the difference between the enthalpy of mixing in liquid and the enthalpy of formation in solid is around 4 kJ/mole-atom at 40.0 – 60.0 at. % Cu, and in the Fe-Ni system [48] the difference is around 5 kJ/mole-atom at 50.0 – 75.0 at. % Ni. The enthalpies of formation of solids in the Nb-Ni system are around 30 kJ/mole-atom at 25.0 at. % Nb, which is closer to the data of liquid (around 25 kJ/mole-atom) measured by Chistyakov et al. [30] and Sudavtsova et al. [40]. Since all the experiments were measured using the calorimetry method, the results from Chistyakov et al. [30] and Sudavtsova et al. [40] are adopted in the present work, while the experimental data of liquid from Schaefers et al. [46] is excluded because the difference is too large (up to 15 kJ/mole-atom at 25.0 at. % Nb) comparing with the enthalpies of formation of solids in the Nb-Ni system.

Experimental data of site occupancy are only available for Nb in μ-Nb_7Ni_6 by Joubert et al. [16] measured by EPMA at 1273 K. With the present upgrade of the ESPEI code (see details in Sec.3.2.2), these data are included in the present CALPHAD modeling.

3 Methodology

3.1 First-principles thermodynamics

DFT-based first-principles calculations can be used to predict Gibbs (and Helmholtz) energy of solid phase as a function of temperature. The expression for Gibbs energy under zero external pressure ($P = 0$ GPa, i.e., the Helmholtz energy) within the quasiharmonic approach is [49],

$$G = F(V,T)|_{P=0} + PV|_{P=0} = E_0(V) + F_{vib}(V,T) + F_{el}(V,T) \quad \text{Eq. 1}$$

where F is the Helmholtz energy, V is the volume, T is the absolute temperature, and P is the pressure. $E_0(V)$ is the static energy at 0 K without vibrational contribution. F_{vib} is the contribution of lattice vibrations and F_{el} the contribution by thermal electrons, and both are functions of V and T.

The energy versus volume (E-V) curves for each phase (or endmember) at 0 K were predicted by DFT-based first-principles calculations, usually employing 7 data points. The $E_0(V)$ curves were fitted by the following 4-parameter Birch-Murnaghan (BM4) equation of state (EOS) [49],

$$E_0(V) = k_1 + k_2 V^{-2/3} + k_3 V^{-4/3} + k_4 V^{-2} \quad \text{Eq. 2}$$

where k_1, k_2, k_3, and k_4 are fitting parameters. This EOS will result in four equilibrium properties at $P = 0$ GPa, including equilibrium energy E_0, volume V_0, bulk modulus B_0, and the pressure derivative of bulk modulus B'. The vibrational contribution F_{vib} can be predicted by the frequency-dependent phonon density of states (DOS) [50],

$$F_{\text{vib}}(T,V) = k_B T \int_0^{\infty} \ln\left[2 \sinh \frac{\hbar\omega}{2k_B T}\right] g(\omega)\, d\omega \quad \text{Eq. 3}$$

where $g(\omega)$ is the phonon DOS as a function of V and frequency ω. The thermal electronic contribution F_{el} can be predicted by Mermin statistics through $F_{el} = E_{el} - TS_{el}$, where E_{el} is the internal energy at V and T, and S_{el} the bare electronic entropy [51].

3.2 Details of first-principles calculations

DFT-based first-principles calculations were performed for BCC Nb and FCC Ni as reference states and the two TCP phases of δ-$NbNi_3$ and μ-Nb_7Ni_6 in the Nb-Ni system. δ-$NbNi_3$ was modeled by a three-sublattice model with a total of 8 endmembers; and μ-Nb_7Ni_6 was modeled by a five-sublattice model with 32 endmembers. Phonon calculations were performed for BCC Nb and FCC Ni and the stable endmembers of δ-$NbNi_3$ and μ-Nb_7Ni_6 including $Nb_2Ni_2Ni_4$ and $Nb_2Ni_2Ni_4$ of δ-$NbNi_3$, and $Nb_6Ni_6Nb_6Ni_{18}Nb_3$, $Nb_6Nb_6Nb_6Ni_{18}Nb_3$, $Nb_6Nb_6Nb_6Ni_{18}Ni_3$, and $Nb_6Nb_6Ni_6Ni_{18}Ni_3$ of μ-Nb_7Ni_6. Note that the crystal structures of μ-Nb_7Ni_6 and δ-$NbNi_3$ can be found in such as the Materials Project [43].

The Vienna *ab initio* Simulation Package (VASP) [52] was adopted for DFT-based first-principles and phonon calculations in the present work. The projector augmented wave (PAW) method was used to describe the ion-electron interaction [53], while the generalized gradient approximation (GGA) by Perdew, Burke, and Ernzerhof (PBE) was used to describe the exchange-correlation functional [54]. The plane-wave cutoff energy was set to be 367.945 eV for structural relaxations and phonon calculations, and 520 eV for the final static calculations to get accurate *E-V* data points and electron DOS's. The convergence criterion of electronic self-consistency was set as 6×10^{-5} eV/atom for relaxations, static calculations, and phonon calculations, while *k*-points meshes of (8×8×7) were used for relaxations and static calculations of δ-$NbNi_3$ and (5×5×1) for relaxations and static calculations of μ-Nb_7Ni_6. For phonon calculations, *k*-points meshes of (2×2×2) were adopted for δ-$NbNi_3$ and (1×1×1) for μ-Nb_7Ni_6. The selected electronic configurations include 5 charge valences for Nb and 10 for Ni, which are the same as those used by the Materials Project [43].

3.2 CALPHAD modeling

3.2.1 Thermodynamic models

There are three types of phases in the Nb-Ni system, i.e., the solution phases of BCC, FCC, and liquid, the stoichiometric compound of $NbNi_8$, and the non-stoichiometric TCP phases of δ-$NbNi_3$ and μ-Nb_7Ni_6. For the solution phases, the Redlich-Kister polynomial [55] was adopted to describe Gibbs energy,

$$G_m^\alpha = x_{Nb}G_{Nb}^\alpha + x_{Ni}G_{Ni}^\alpha + RT(x_{Nb}lnx_{Nb} + x_{Ni}lnx_{Ni}) + x_{Nb}x_{Ni}\sum_{k=0} {}^kL_{Nb,Ni}(x_{Nb} - x_{Ni})^k \quad \text{Eq. 3}$$

where x_{Nb} and x_{Nb} are the mole fractions of Nb and Ni in phase α. G_{Ni}^α and G_{Nb}^α are the Gibbs energies of pure elements, Ni and Nb, with respect to their standard element reference (SER) states at P = 1 bar and T = 298.15 K, obtained from the Scientific Group Thermodata Europe (SGTE) database [56]. R is the gas constant, T the temperature, and ${}^kL_{Nb,Ni}$ the kth interaction parameter between Nb and Ni,

$${}^kL_{Nb,Ni} = a + bT \quad \text{Eq. 4}$$

where a and b are model parameters.

$NbNi_8$ is treated as a stoichiometric compound with its Gibbs energy described by,

$$G_{Nb:Ni}^{NbNi_8} = {}^0G_{Nb}^{BCC} + 8\,{}^0G_{Ni}^{FCC} + A + BT \quad \text{Eq. 4}$$

where A and B are model parameters. ${}^0G_{Nb}^{BCC}$ and ${}^0G_{Ni}^{FCC}$ are the Gibbs energies of pure Nb and Ni in their stable structures, i.e., BCC and FCC, respectively. The values of ${}^0G_{Nb}^{BCC}$ and ${}^0G_{Ni}^{FCC}$ are taken from the SGTE database [56].

For non-stoichiometric compounds, the compound energy formalism (CEF) is used to describe the phase with its sublattices corresponding to its Wyckoff sites, see Table 1 [57]. In the CEF, the Gibbs energy for the phase of interest is described as follows,

$$G_{mf} = {}^0G_{mf} + RT\sum_t a^t \sum_i y_i^t ln y_i^t + {}^EG_{mf} \quad \text{Eq. 5}$$

where ${}^0G_{mf}$ is the Gibbs energy contribution of each endmember, R the gas constant, and T the temperature. $\sum_t a^t \sum_i y_i^t ln y_i^t$ is sublattice ratio a^t in the sublattice t times the ideal mixing in this sublattice, which is calculated by the site fraction y_i^t and site fraction y_i^t. ${}^EG_{mf}$ is the excess Gibbs energy which contains the contributions from the mixing in one sublattice where all other sublattices only contain one component; and from the mixing in more than one sublattice where more than one sublattices contain two or more components.

3.2.2 CALPHAD modeling by the upgraded ESPEI and PyCalphad

The open-source software tools, PyCalphad [21] and ESPEI [58], were employed in the present work to remodel the Nb-Ni system. PyCalphad is a Python-based code for computational thermodynamics using the CALPHAD method, focusing on calculating phase diagrams, investigating thermodynamic properties, and designing new materials [58]. ESPEI is a tool for database development using the CALPHAD approach [58], which uses PyCalphad as computational engine to perform thermodynamic calculations. ESPEI has two major features. First, ESPEI uses thermochemical data to choose and evaluate parameters to model Gibbs energy of individual phases. Second, ESPEI optimizes and quantifies uncertainties of model parameters using both thermochemical and phase equilibrium data through Bayesian parameter estimation via an ensemble Markov Chain Monte Carlo (MCMC) [59–61]. However, in the previous versions of ESPEI, the site occupancy data cannot be used as input for CALPHAD modeling.

In the present work, a new function that uses site occupancy data as input to optimize model parameters and quantify their uncertainties is implemented in ESPEI. ESPEI uses Bayesian parameter estimation to optimize model parameters, in which the acceptance of parameters is based on the posterior probability $p(\theta/D)$ of the model parameters θ under the data D. The posterior probability $p(\theta/D)$ can be calculated by the likelihood $p(D/\theta)$, the prior $p(\theta)$, and the evidence $p(D)$, i.e., $p(\theta/D) = p(D/\theta) * p(\theta)/p(D)$. The likelihood $p(D/\theta)$ is related to how experimental data are described by the proposed parameters, and the prior $p(\theta)$ is the probability distribution of each parameter. The flowchart of the present implementation is illustrated in Fig. 3, showing that ESPEI can take site occupancy as input using the JavaScript Object Notation (JSON) data format, propose new parameter values from the MCMC method, and calculate the log-type posterior probabilities from the prior and likelihood of site occupancy and other experimental data from the fixed temperature, pressure, and the number of moles. Then, the acceptance of the new parameters is decided by the Metropolis-Hastings criteria [61], comparing the posterior probabilities calculated from the new parameters with those from the current parameters. The JSON files that contain data about site occupancy are proposed and the codes to calculate the posterior probability of site occupancy are implemented into the error functions of ESPEI. To be consistent with the weighting of error from different types of data like activity, phase boundary, and thermochemical data, the likelihood for site occupancy data is normalized by the standard deviation of the error. The default value of the standard deviation of the site occupancy is set to be 0.01. After this implementation, the site occupancy data can be considered together with thermochemical and phase equilibrium data to fit all model parameters simultaneously, enabling the uncertainty propagation of site occupancies. In the present work, experimental data of site occupancy by Joubert et al. [16] were used to remodel μ-Nb_7Ni_6 in the Nb-Ni system as shown in Sec. 4.2.

Uncertainty quantification (UQ) in ESPEI can quantify uncertainties of model parameters using the possible values during the MCMC sampling process [61]. The UQ in ESPEI adopts the samples from different Markov chains in the MCMC optimization and leverages them to estimate the uncertainties for thermodynamic properties of interest. In the present work, the UQ of site occupancy is implemented into ESPEI, which allows the analysis of site occupancy from modeling parameters. For example, the uncertainties of site occupancy of μ-Nb_7Ni_6 and enthalpy of mixing

of liquid were used to show the possible range from the MCMC optimization. Note that four chains for each parameter were used during the MCMC optimization for 1000 steps which is enough to gain convergent results based on our tests. The standard deviation was set to be 0.01 in the initialization of chain values using the Gaussian distribution.

4 Results and discussion

4.1 Thermodynamic properties by first-principles calculations

Table 4 summarizes the space group and the predicted properties of V_0, B_0 , and B'_0 for the phases of BCC-Nb, FCC-Ni, δ-$NbNi_3$, and μ-Nb_7Ni_6 at 0 K using Eq. 2, in comparison with available experimental data [10,11,62,63].The δ-$NbNi_3$ has the highest bulk modulus (207.7 GPa), followed by μ-Nb_7Ni_6 (200.0 GPa), FCC-Ni (195.9 GPa), and BCC-Nb (173.5 GPa), indicating that the bonding in δ-$NbNi_3$ is strongest. The B' values increase from BCC-Nb (3.86), μ-Nb_7Ni_6 (4.48), δ-$NbNi_3$ (4.65), to FCC-Ni (4.81). For B_0 values from DFT-based predictions, both BCC-Nb and FCC-Ni show good agreement with experimental data [62,63]. BCC-Nb shows only a 1.0% difference and FCC-Ni has a 5.0% difference when compared with experiments. Table 4 shows that V_0 increases from BCC-Nb, μ-Nb_7Ni_6, FCC-Ni, to δ-$NbNi_3$. The difference of V_0 between DFT-based calculations and experiments is about 1.72%. Note that these differences are due mainly to the exclusion of vibrational contribution to DFT calculations and the uncertainty of exchange-correlation functional used in the calculations [45].

Fig. 4 shows the predicted values of entropy and enthalpy of BCC Nb as a function of temperature from the present DFT calculations using Eq. 1, which are in good agreement with the SGTE data [56] with an average difference of 4.83% as well as the standard deviation of 0.49 for entropy; and an average difference of 5.79% as well as the standard deviation of 0.97 for enthalpy. Similarly,

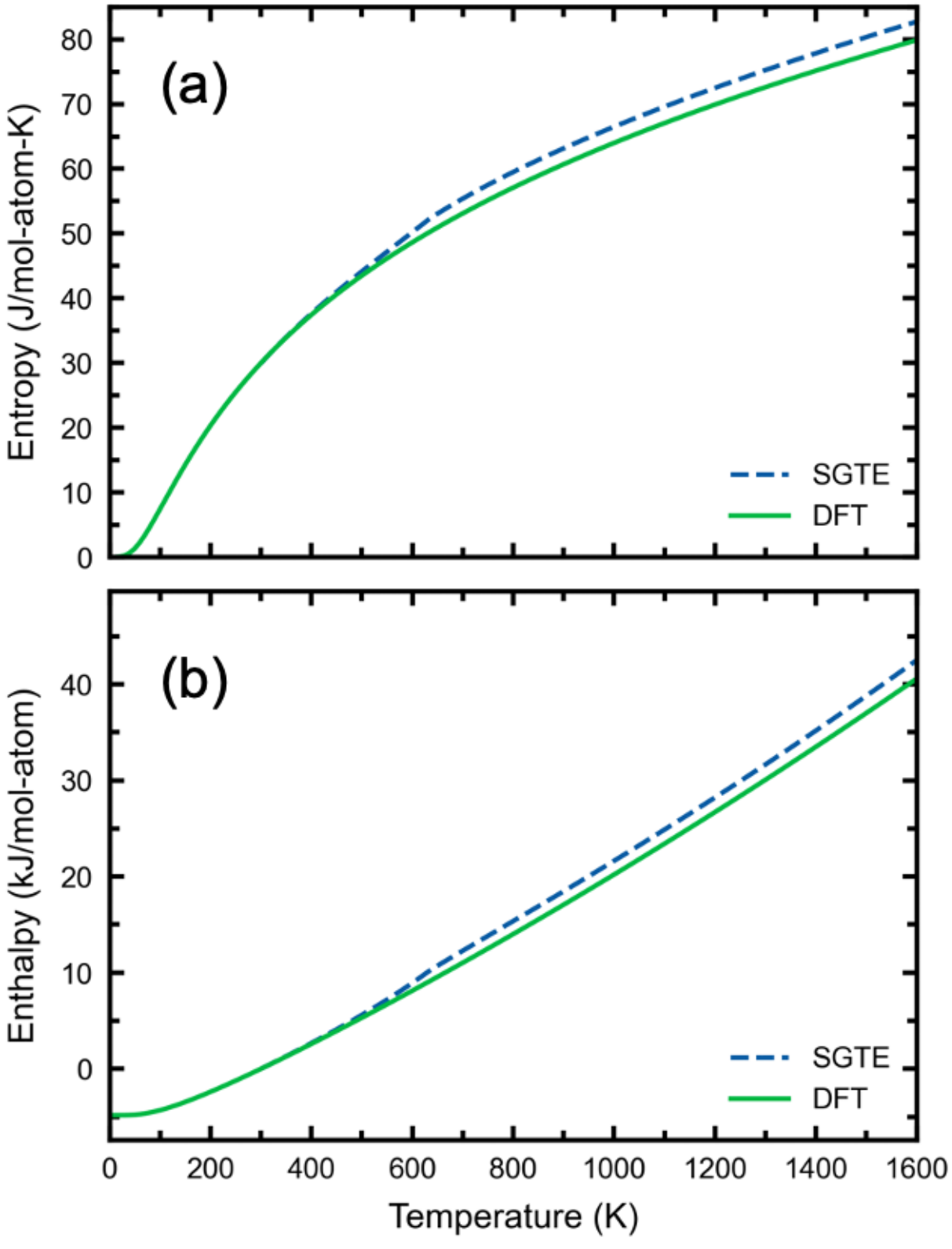


Fig. 5 presents the predicted entropy and enthalpy of FCC Ni as a function of temperature from present DFT calculations using Eq. 1, showing a good agreement with the SGTE database [56]

(with magnetism) with the average difference of 3.24% and the standard deviation of 1.21 for entropy; and the average difference of 6.44% and the standard deviation 2.42 for enthalpy.

a

b

Table 5 exhibits the predicted enthalpy of formation (ΔH_{form}) values of δ-$NbNi_3$ and μ-Nb_7Ni_6 from DFT-based first-principles calculations at both 0 K and room temperature, which are in good agreement with experiments [25], with a difference less than 3.3 kJ/mol-atom. The configurations on the convex hull around the compositions of 25 at. % Nb and 50 at. % Nb were chosen to represent the ΔH_{form} for δ-$NbNi_3$ ($Ni_1Nb_1Ni_2$) and μ-Nb_7Ni_6 ($Nb_2Nb_2Nb_2Ni_6Ni_1$). The predicted ΔH_{form} value of δ-$NbNi_3$ ($Ni_1Nb_1Ni_2$) is -28.4 kJ/mol-atom at 0 K (and -28.5 kJ/mol-atom at room temperature). The difference between the predicted ΔH_{form} value and experimental result (-31.8 kJ/mol-atom) is 3.2 kJ/mol-atom at 0 K (and 3.3 kJ/mol-atom at room temperature), which is within the reported error (4.73 kJ/mol-atom) of measurements by Argent et al. [25] using the calorimetry method for three samples at 25 at. % Nb. While the predicted ΔH_{form} value of μ-Nb_7Ni_6 ($Nb_2Nb_2Nb_2Ni_6Ni_1$) is -20.6 kJ/mol-atom at 0 K (and -20.4 kJ/mol-atom at room temperature). The difference between the predicted ΔH_{form} and experimental value (-22.6 kJ/mol-atom) is 2 kJ/mol-atom at 0 K (and 2.2 kJ/mol-atom at room temperature), which is also within the reported error (6.82 kJ/mol-atom) of measurements by Argent et al. [25] using the calorimetry method for six tests at 50 at. % Nb.

Fig. 6 shows the predicted phonon DOS curves at the equilibrium volumes of BCC Nb, FCC Ni, δ-$NbNi_3$, and μ-Nb_7Ni_6. The phonon DOS's of BCC Nb and FCC Ni show a good agreement with experimental data [64] as shown in Supplemental Fig. 12. Fig. 6 indicates that BCC Nb exhibits a higher phonon DOS at the lowest frequency region (e.g., < 5 THz) compared with those from FCC-Ni, μ-Nb_7Ni_6, and δ-$NbNi_3$; indicating phonon of BCC Nb has a larger contribution to Helmholtz energy due to vibrational entropy (S_{vib}) since $S_{vib} \propto -\int g(\omega) \ln(\omega) \mathrm{d}\omega$ [65]. The phonon DOS of μ-Nb_7Ni_6 is higher at the highest frequency region (e.g., > 8 THz), causing a smaller contribution to vibrational entropy (as well as Helmholtz energy). This trend is in accordance with bulk moduli predicted from DFT and observed from experiments [62,63]: in general the higher the

bulk modulus, the smaller the contribution to entropy will be, for example, B_0 = 173.5 GPa of BCC Nb and B_0 = 207.7 GPa of δ-$NbNi_3$; see Table 4.

4.2 Thermodynamic modeling by ESPEI/PyCalphad

The present model parameters are summarized in the thermodynamic database (TDB) file in the Supplemental Material. Fig. 1 illustrates the predicted ΔH_{form} values for the stable compounds (endmembers) in the Nb-Ni system at 298 K from the present CALPHAD modeling, agreeing reasonably well with the predictions by the present first-principles calculations and those in the literature (e.g., Materials Project [43] and OQMD [44]) and experimental data by Argent et al. [25]. As shown in a
b

Table 5, the predicted ΔH_{form} values at 25.0 at. % Nb from both 0 K and room temperature show a good agreement with experimental data [25] for δ-$NbNi_3$ (at 25 at. % Nb) with the difference of less than 3.4 kJ/mol-atom, which is in the range of experimental error (4.73 kJ/mol-atom). For μ-Nb_7Ni_6, the predicted ΔH_{form} values at 46.0 at. % Nb from both 0 K and room temperature show a good agreement with experimental data [25] for μ-Nb_7Ni_6 (at 50 at. % Nb) with a difference less than 2.2 kJ/mol-atom (experimental error is 6.82 kJ/mol-atom). As shown in Fig. 1, the predicted ΔH_{form} values from the present CALPHAD modeling show less than 6.5 kJ/mol-atom difference with those measured by Argent et al. [25] from 12.2 – 75.0 at. % Nb, which is slightly higher than the error bar (around 4 kJ/mol-atom) given by Argent et al. [25]. However, Argent et al. [25]'s ΔH_{form} values at 50 at. % Nb (-22.4 kJ/mol-atom) have a difference of 6.8 kJ/mol-atom using six experiments. The predicted ΔH_{form} value from CALPHAD modeling is -28.9 kJ/mol-atom at 50 at. % Nb, causing a 6.5 kJ/mol-atom difference in ΔH_{form} which is comparable with experimental error (6.8 kJ/mol-atom [25]). Note that Sokolvskaya et al. [40]'s data were not considered in the present CALPHAD modeling, since their ΔH_{form} value (-41.6 kJ/mol-atom at 50 at. % Nb) shows a 19.2 kJ/mol-atom lower than the one (-22.4 kJ/mol-atom) by Argent et al. [25]. Therefore, we

conclude that the present ΔH_{form} value at 50 at. % Nb (-28.9 kJ/mol-atom) agree reasonably well the experiments (-22.4 kJ/mol-atom) by Argent et al. [25].

Fig. 2 plots the presently predicted enthalpy of mixing (ΔH_{mix}) curve of liquid at 1823 K in comparison with experiments data measured by Schaefers et al. [46] at 1927 K and 2000 K, by Chistyakov et al. [30] at 2096 K, and by Sudavtsova et al. [40] at 2148 K. Fig. 2 shows that the ΔH_{mix} decreases with increasing the composition of Nb until 30.0 at. % Nb, then increases with increasing the composition of Nb. It is seen that the presently calculated results show a good agreement those by Chistyakov et al. [30] and Sudavtsova et al. [40] with an average difference of 2 kJ/mol-atom. As mentioned in Sec. 2, the present CALPHAD modeling excluded Schaefers et al. [46]'s data.

Table 6 summarizes the invariant reactions from the present modeling. It shows a good agreement with experiments [23,66] with the difference in compositions less than 2.3 at. % Nb, and the variance of the reaction temperature is less than 46 K. For example, the eutectic reaction from liquid to FCC and δ-$NbNi_3$, from experimental observation [23] the reaction composition is 16.0 at. % Nb for liquid phase, 12.7 at. % Nb for FCC, and 22.6 at. % Nb for δ-$NbNi_3$; and the reaction temperature is 1555 K. From the present predictions, these values are 14.7 at. % Nb, 12.8 at. % Nb, 23.3 at. % Nb, and 1547 K, respectively.

Fig. 7 shows the calculated phase diagram based on the present CALPHAD modeling, agreeing reasonably well with experimental data [16,18,19,22–24,26,38]. The present phase boundaries of δ-$NbNi_3$ between FCC and δ-$NbNi_3$ are predicted from 23.8 to 23.3 at. % Nb at 790 K – 1547 K, matching well with experimental data from Chen et al. [24] (around 23.5 at. % Nb at 1323 K), Joubert et al. [18] (23.6 at. % Nb at 1280 K), Duerden et al. [23] (23.6 – 23.6 at. % Nb at 1273 K - 1420K), and Murametsu et al. [22] (24.1 – 24.1 at. % Nb from 1070 - 1240 K). The phase boundaries of δ-$NbNi_3$ between μ-Nb_7Ni_6 and δ-$NbNi_3$ are predicted from 25.7 to 28.4 at. % Nb at 1000 K – 1494 K, agreeing well with experimental data by Chen et al. [24] (25.6 at. % Nb at 1323 K), Duerden et al. [23] (26.5 to 26.6 at. % Nb at 1273 K - 1420K), and Murametsu et al. [22] (26.5 at. % Nb at 1170 K). Phase boundaries of μ-Nb_7Ni_6 between μ-Nb_7Ni_6 and BCC were predicted as 57.2 – 56.6 at. % Nb at 1000 K -1400 K, matching well with measurements by

Duerden et al. [23] (54.0 at. % Nb at 1273K), Chen et al. [24] (56.3 at. % Nb at 1323K), Murametsu et al. [22] (56.2 – 56.2 at. % Nb at 1070 - 1240 K), Svechnikov et al. [38] (58.2 – 58.2 at. % Nb at 1074 -1276 K), and Joubert et al. [16] (56.9 at. % Nb at 1273K).

The comparison between the present modeling and the previous modeling work by Chen et al. [19] is shown in Fig. S2, showing that the present model has a better fit with experimental data on the phase boundaries of μ-Nb_7Ni_6 between μ-Nb_7Ni_6 and δ-$NbNi_3$, and the phase boundaries between liquid and BCC. Phase boundaries between liquid and BCC measured by Wicker et al. [26] were considered in the present work besides the heating, cooling, and quenching measurements by Duerden et al. [23]. From the present work, the phase boundaries between liquid and BCC show around 200 K lower than those from Chen et al.'s modeling work [19] at 50.0 – 85.0 at. % Nb, which gave a better match (with an average difference around 100 K) with the measured data by Wicker et al. [26]. The phase boundaries of μ-Nb_7Ni_6 between μ-Nb_7Ni_6 and δ-$NbNi_3$ are predicted around 48.3 – 49.3 at. % Nb at 1000 K – 1494 K in present work while those in Chen et al. [19] are 51.6 – 51.7 at. % Nb at 1000 K – 1458 K. The phase boundaries of μ-Nb_7Ni_6 between μ-Nb_7Ni_6 and δ-$NbNi_3$ from the present work match better with experimental data including from Murametsu et al. [22] (48.6 at. % Nb at 1240 K), Svechnikov et al. [38] (49.7 at. % Nb at 1276 K), and Chen et al. [24] (49.5 at. % Nb at 1273K), while Chen et al. [19] modeling work has a good agreement with experiments by Svechnikov et al. [38] (51.8 at. % Nb at 1074 K).

Fig. 9 shows the predicted site occupancy curves of Nb in μ-Nb_7Ni_6 from the present modeling compared with measured data by Joubert et al. [16]. It can be seen that the present CALPHAD predictions of site occupancy regarding Nb in μ-Nb_7Ni_6 agree well with experiments with an absolute error less than 0.062, especially when compared with site occupancy of Nb in μ-Nb_7Ni_6 predicted by Chen et al. [19] using a four-sublattice model, *cf.*, Table 7. Therefore, the sites $6c_2$ and $6c_3$ have the same site occupancy values from Chen et al. [19]'s model, causing the mean absolute error (MAE) with experiments [16] up to 0.044 in site $6c_2$ which is much higher than that from the present model (0.008 in site $6c_2$). At the same time, the MAE values compared with experiments are 0.128 at site $6c_1$, 0.032 at site $6c_3$, 0.088 at site 18h, and 0.086 at site 3a from Chen et al. [19] modeling, while the MAE values from the present model are 0.006 at site $6c_1$, 0.032 at site $6c_3$, 0.024 at site 18h, and 0.002 at site 3a. The standard deviations from the present work are

also smaller than those from Chen et al. [19]'s modeling, with the difference of standard deviation being 0.151 at site $6c_1$, 0.043 at site $6c_2$, 0.0 at site $6c_3$, 0.094 at site 18h, and 0.112 at site 3a. Therefore, with the present sublattice models based on Wyckoff sites of μ-Nb_7Ni_6 and δ-$NbNi_3$ and the new function in ESPEI to consider site occupancy data into modeling, these two TCP phases can be modeled well.

Fig. 10 shows the 95 % credible interval uncertainty propagation regions of site occupancy of each Wyckoff site concerning Nb in μ-Nb_7Ni_6 predicted by the parameters in the MCMC method as implemented in ESPEI during CALPHAD modeling, showing a good agreement with the uncertainty of experimental data [16]. It is seen that the uncertainty range of Nb in the first sublattice is around 0.2, corresponding to Wyckoff position $6c_1$. The shadow region in Fig. 10 covers all experimental data [16] except for the one at 49.6 at. % Nb that has a 0.2 difference with respect to the data point at 51.8 at. % Nb, which means the uncertainty of site occupancy at site $6c_1$ includes most of the uncertainty of experiments at site $6c_1$. Similarly, at the second, fourth, and fifth sublattices (corresponding to Wyckoff positions $6c_2$, 18h, and 3a, respectively), the uncertainty ranges of Nb are around 0.2, which also cover most experimental data except for the one at 49.6 at. % Nb at $6c_3$ site that has 0.06 (the average difference is 0.03) difference with respect to the data point at 51.8 at. % Nb. For the third sublattice corresponding to Wyckoff position $6c_3$, the uncertainty ranges of Nb do not appear due to the third sublattice of stable endmembers occupied by Nb around 49.6 -56.9 at. % Nb. The uncertainty propagation regions of site occupancy of Nb in μ-Nb_7Ni_6 cover 80% of experiments, which shows a good match with experimental data by considering that the standard deviation of experimental data is around 0.35. The good agreement between uncertainty propagation regions and the experimental data shows that the uncertainty during CALPHAD modeling reflects the uncertainty of experiments.

As another example, Fig. 11 shows the uncertainty propagation regions of ΔH_{mix} in liquid with a 95 % credible interval, indicating that the uncertainty increases from 0.0 – 40.0 at. % Nb, increases slightly from 40.0 – 80.0 at. % Nb, and then decreases from 80.0 – 100.0 at. % Nb, with the largest uncertainty around 3 kJ/mol-atom at 80.0 at. % Nb. It shows that a large increase of uncertainty in ΔH_{mix} of liquid appears at the Nb-rich region due to the parameters of the liquid phase being more sensitive at the Nb-rich region in the present CALPHAD modeling.

4 Conclusions

The present work combines thermodynamic data from DFT-based first-principles calculations and experiments, and the uncertainty quantification (UQ) method in CALPHAD modeling to model Gibbs energy expressions of the Nb-Ni system. The key conclusions are summarized as follows.

- First-principles calculations are used to predict thermochemical properties as a function of temperature for the TCP phases of δ-$NbNi_3$ and μ-Nb_7Ni_6, which provide the predicted ΔH_{form} as input for CALPHAD modeling.
- The present thermodynamic models of the δ-$NbNi_3$ and μ-Nb_7Ni_6 phases are built on their Wyckoff positions combined with site occupancy as input for CALPHAD modeling by ESPEI, making precise descriptions for both phase diagram and the distribution of site occupancies compared with available experimental data. For the phase diagram, the invariant reactions from the present modeling agree well with data from experiments [23,66].
- The uncertainty propagation regions have been employed to show uncertainties of enthalpy of mixing, and site fraction of Nb in the μ-Nb_7Ni_6 phase. The UQ regions of site occupancy of Nb in μ-Nb_7Ni_6 include 80% of the experimental data [16], indicating that the uncertainty during the modeling reflects the uncertainty of the experiments.

Acknowledgments:

The authors acknowledge the financial support by the Office of Naval Research (ONR) under Contract No. N00014-21-1-2608 and National Science Foundation (NSF) via Award CMMI-2050069. First-principles calculations were performed partially on the Roar supercomputer at the Pennsylvania State University's Institute for Computational and Data Sciences (ICDS), partially on the resources of the National Energy Research Scientific Computing Center (NERSC) supported by the U.S. DOE Office of Science User Facility operated under Contract No. DE-AC02-05CH11231, and partially on the resources of the Extreme Science and Engineering Discovery Environment (XSEDE) supported by NSF with Grant No. ACI-1548562. Part of this

work was performed under the auspices of the U.S. DOE at Lawrence Livermore National Laboratory under Contract No. DE-AC52–07NA27344.

Table 1: Wyckoff positions of the TCP phases of μ-Nb_7Ni_6 and δ-$NbNi_3$ in the Nb-Ni system.

Wyckoff position of δ phase [a]	x	y	z
2a	0	0	0.318
2b	0	0.5	0.651
4f	0.75	0	0.841
Wyckoff position of μ phase [b]			
3a	0	0	0
6c (1)	0	0	0.167
6c (2)	0	0	0.346
6c (3)	0	0	0.448
18h	0.5	0.5	0.590

[a] δ phase with space group Pmmn (no. 59), Pearson symbol oP8, strukturbericht designation $D0_a$, and prototype of β-Cu_3Ti [10].
[b] μ phase with space group $R\bar{3}m$ (no. 166), Pearson symbol hR13, strukturbericht designation $D8_5$, and prototype of Fe_7W_6 [11].

Table 2: Sublattice models used in the previous CALPHAD modeling.

References	Model for δ-$NbNi_3$	Model for μ-Nb_7Ni_6
Kaufman and Nesor [17]	$(Ni)_{0.75}$ $(Nb)_{0.25}$	$(Ni)_{0.47}$ $(Nb)_{0.53}$
Kejun et al. [14]	$(Nb, Ni)_3$ $(Nb, Ni)_1$	$(Nb, Ni)_1Ni_4(Nb, Ni)_2Nb_6$
Bolcavage and Kattner [13]	$(Nb, Ni)_3$ $(Nb, Ni)_1$	$(Nb, Ni)_7$ $(Nb)_6$
Joubert et al. [18]	$(Nb, Ni)_3$ $(Nb, Ni)_1$	$(Nb, Ni)_1Nb_2Nb_2(Nb, Ni)_2(Nb, Ni)_6$
Chen et al. [15]	$(Nb, Ni)_3$ $(Nb, Ni)_1$	$(Nb, Ni)_1Nb_4(Nb, Ni)_2(Nb, Ni)_6$

Table 3: Crystallographic information for the phases in Nb - Ni and their sublattice models used in the present CALPHAD modeling.

Phase name	Strukturbericht	Space group	Pearson symbol	Model
Liquid(L)				(Nb, Ni)
FCC	A1	$Fm\bar{3}m$	cF4	$(Nb, Ni)_1$ $(Va)_1$
HCP	A3	$P6_3/mmc$	hP2	$(Nb, Ni)_1$ $(Va)_1$
BCC_A2	A2	Im3m	cl2	$(Nb, Ni)_1$ (Va)3
δ-$NbNi_3$	$D0_a$	Pmmn	oP8	$(Nb, Ni)_1(Nb, Ni)_1(Nb, Ni)_2$
μ-Nb_7Ni_6	$D8_5$	$R\bar{3}m$	hR13	$(Nb, Ni)_1(Nb, Ni)_2(Nb, Ni)_2(Nb, Ni)_2(Nb, Ni)_6$
$NbNi_8$				$(Nb)_1(Ni)_8$

Table 4: Predicted equilibrium volume (V_0, Å^3/atom), bulk modulus B_0 (GPa), and the derivative of bulk modulus B'_0 from the present EOS fitting at 0 K in comparison with experimental data [10,11,62,63].

Phase	V_0 (Å^3/atom)	% Diff [a]	B_0 (GPa)	% Diff [b]	B'_0	Source
BCC-Nb	18.338	0.221	173.5	1.00	3.86	This work
	18.297		171.8			Expt. [62]
FCC-Ni	21.860	0.217	195.9	5.03	4.81	This work
	21.807		186.0			Expt. [63]
δ-$NbNi_3$	24.176	1.714	207.7		4.65	This work
	24.591					Expt. [10]
μ-Nb_7Ni_6	21.135	0.964	200.0		4.48	This work
	21.338					Expt. [11]

a

b

Table 5: The enthalpy of formation of the δ-$NbNi_3$ phase and μ-Nb_7Ni_6 phase from present DFT-based calculations at both 0 K and room temperature (RT) compared with experimental data at RT [25].

Phase	x_{Nb}	$\Delta_f H$ (kJ/mol-atom) at 0K	Difference (kJ/mol-atom)	$\Delta_f H$ (kJ/mol-atom) at RT	Difference (kJ/mol-atom)	Source
δ-$NbNi_3$	0.25	-28.4	3.4	-28.5	3.3	This work
	0.25	-31.8		-31.8		Expt. [25]
μ-Nb_7Ni_6	0.46	-20.6	2.0	-20.4	2.2	This work
	0.50	-22.6		-22.6		Expt. [25]

Table 6: Predicted invariant reactions in the Nb-Ni system from the present CALPHAD modeling in comparison with available experiments [23,66].

Type	Reaction compositions (at. % Nb)					Temperature (K)	Source
Eutectic	Liquid	↔	fcc	+	δ-$NbNi_3$		
	14.7		12.8		23.3	1547	This work
	16		12.7		22.6	1555	Expt. [23]
Congruent	Liquid	↔	δ-$NbNi_3$				
	25.0		25.0			1712	This work
	25.0		25.0			1675	Expt. [23]
Eutectic	Liquid	↔	δ-$NbNi_3$	+	μ-Nb_7Ni_6		
	41.7		28.4		49.3	1494	This work
	40.5		27.5		50	1448	Expt. [23]
Peritectic	Liquid	+	bcc	↔	μ-Nb_7Ni_6		
	47.9		94.8		56.3	1542	This work
	50.0		95.5		54	1568	Expt. [23]

Peritectic	fcc	+	δ-$NbNi_3$	↔	$NbNi_8$		
	8.5		23.8		11.1	790	This work
						808	Expt. [66]

Table 7. Site occupancies of Nb in μ-Nb_7Ni_6 from the present CALPHAD modeling compared with the modeling work by Chen et al. [19] work and experimental values [16]. Here, MAE indicates the mean absolute error and STD the standard deviation.

Composition	Type of results	$6c_1$	$6c_2$	$6c_3$	18h	3a
x_{Nb} = 0.496	Calc., this work	0.77	0.94	1.00	0.04	0.79
	Calc., Chen et al.	0.91	1.00	1.00	0.00	0.86
	Expt.	0.67	0.85	0.95	0.13	0.74
x_{Nb} = 0.518	Calc., this work	0.79	0.96	1.00	0.07	0.79
	Calc., Chen et al.	0.91	1.00	1.00	0.01	0.87
	Expt.	0.89	1.00	0.89	0.07	0.77
x_{Nb} = 0.530	Calc., this work	0.81	0.97	1.00	0.09	0.80
	Calc., Chen et al.	0.93	1.00	1.00	0.02	0.88
	Expt.	0.78	0.94	1.00	0.10	0.84
x_{Nb} = 0.533	Calc., this work	0.81	0.97	1.00	0.10	0.80
	Calc., Chen et al.	0.94	1.00	1.00	0.03	0.89
	Expt.	0.85	0.99	1.00	0.12	0.81
x_{Nb} = 0.569	Calc., this work	0.85	0.98	1.00	0.16	0.80
	Calc., Chen et al.	0.95	1.00	1.00	0.08	0.90
	Expt.	0.81	1.00	1.00	0.16	0.81
	MAE, this work	0.006	0.008	0.032	-0.024	0.002
	MAE, Chen et al.	0.128	0.044	0.032	-0.088	0.086
	STD, this work	0.070	0.049	0.068	0.052	0.031
	STD, Chen et al.	0.221	0.092	0.068	0.146	0.143

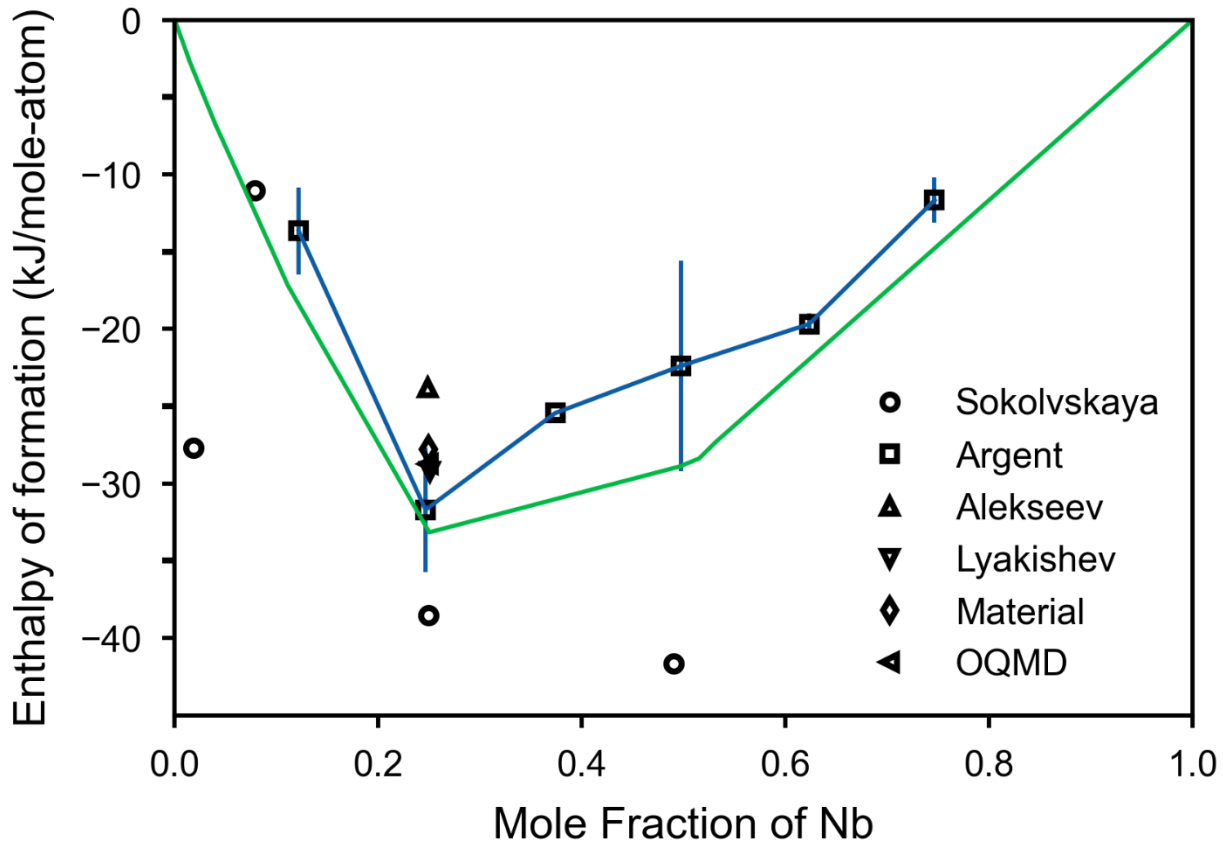


Fig. 1. The presently calculated enthalpies of formation at 298 K of the intermetallic compounds in the Nb-Ni system, in comparison with experimental data by Argent et al. [25], Sokolvskaya et al. [40], Alekseev et al. [34], and Lyakishev et al. [42], and the DFT results from the Materials Project (MP) [43] and the Open Quantum Materials Database (OQMD) [44].

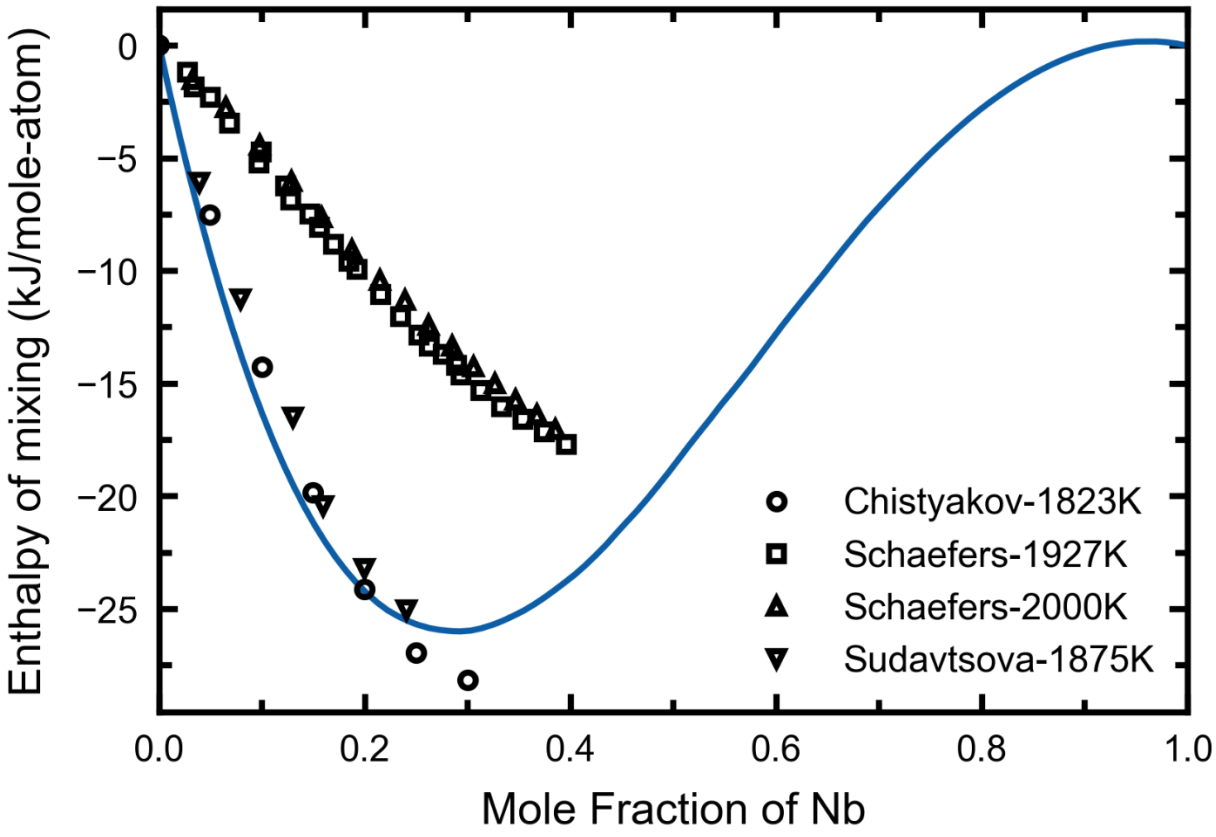


Fig. 2. Calculated enthalpy of mixing in the liquid phase in comparison with available experimental data by Schaefers et al. [46] at 1927 K and 2000 K, by Chistyakov et al. [30] at 2096 K, and by Sudavtsova et al. [40] at 2148 K.

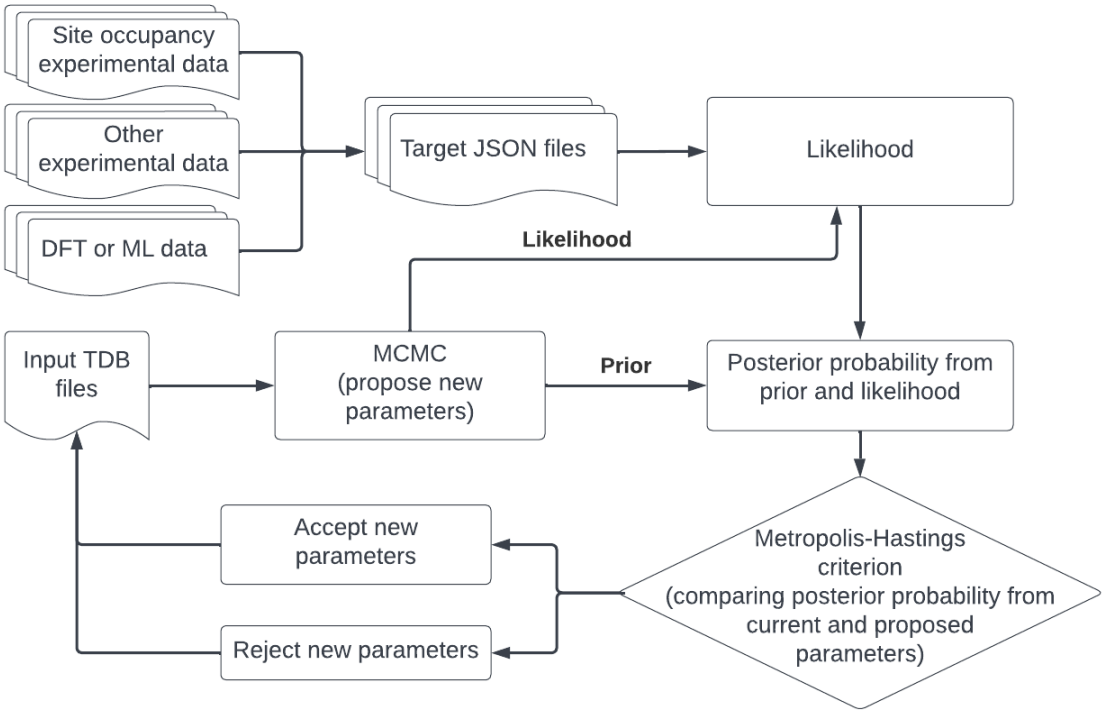


Fig. 3. Workflow of ESPEI implementation to consider site occupancy experimental data.

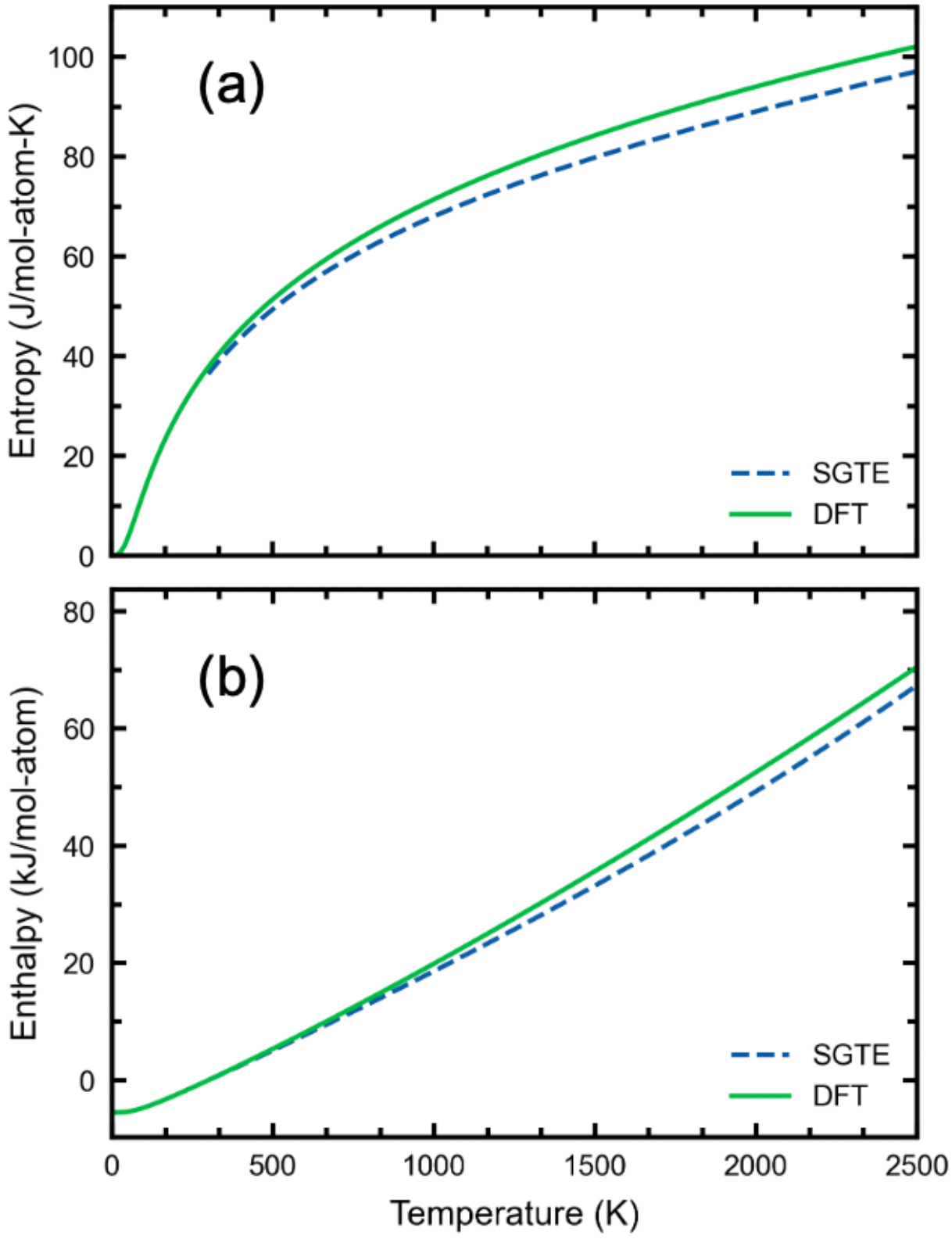

(a)
Entropy (J/mol-atom-K)
0
20
40
60
80
100
SGTE
DFT
(b)
Enthalpy (kJ/mol-atom)
0
20
40
60
80
SGTE
DFT
0
500
1000
1500
2000
2500
Temperature (K)

Fig. 4. (a) Entropy and (b) enthalpy of BCC-Nb as a function of temperature calculated from present work using DFT-based phonon calculations and Eq. xx, compared with the SGTE database [56].

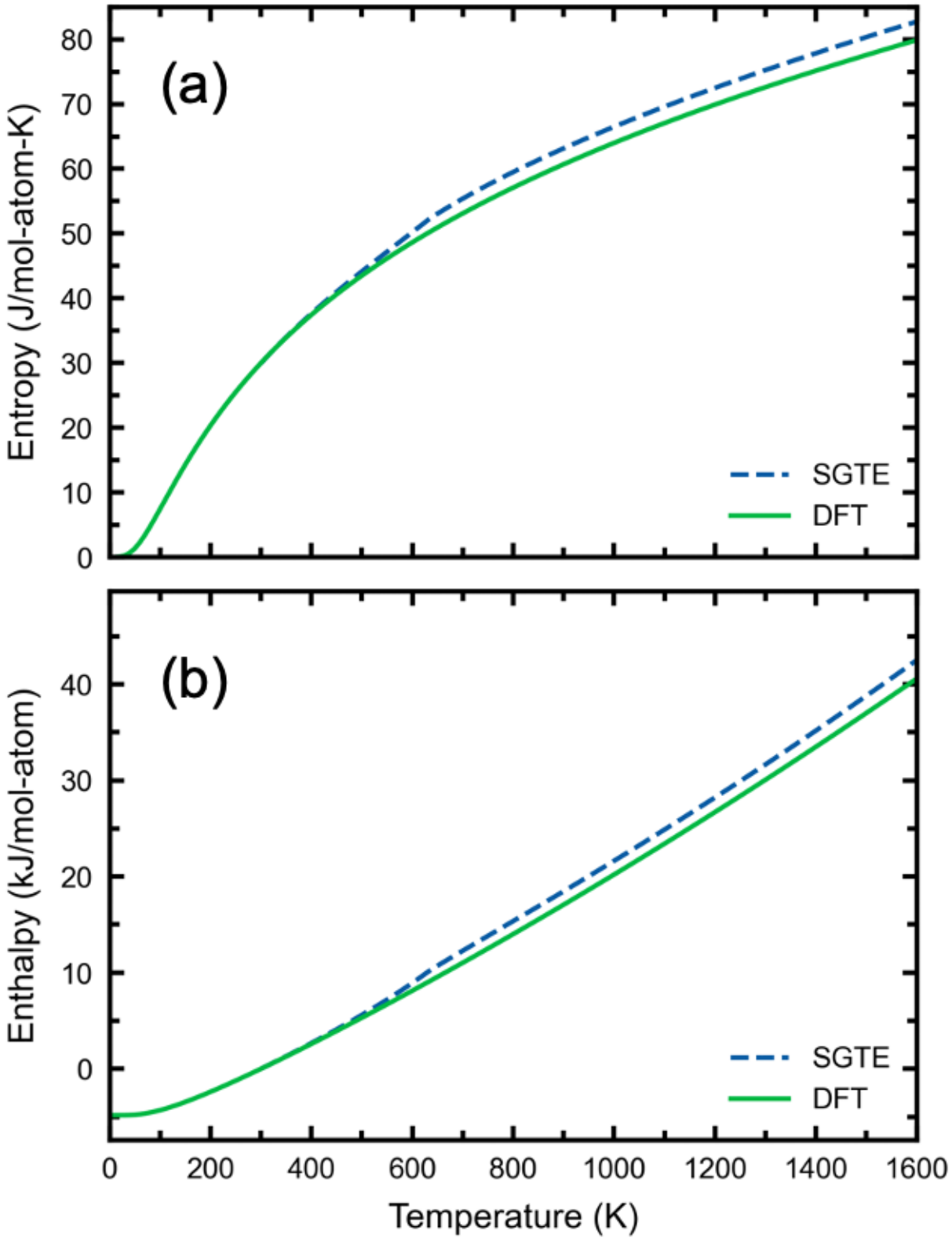


Fig. 5. (a) Entropy, (b) Enthalpy of FCC-Ni as a function of temperature calculated from present work using phonon calculations, compared with the SGTE database[56].

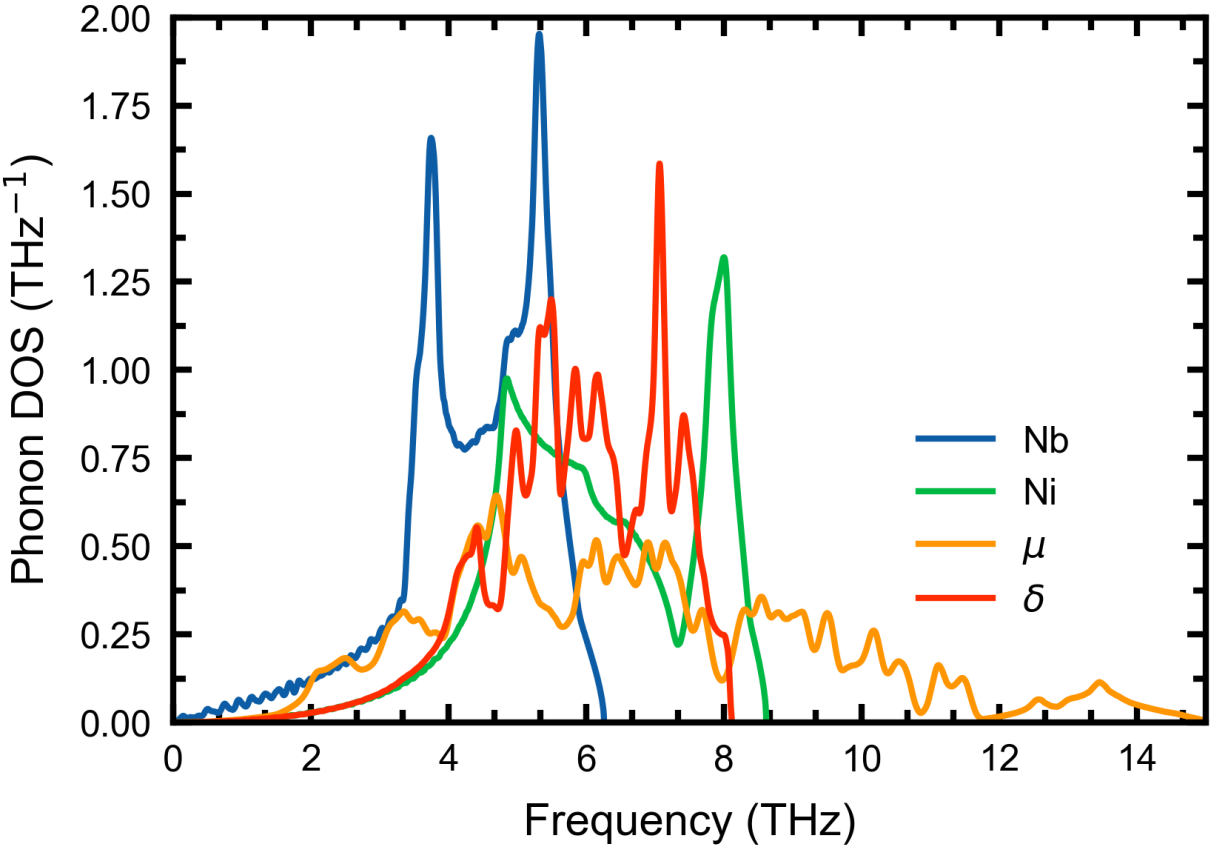


Fig. 6. Predicted phonon density of states (DOS) at the equilibrium volumes of the BCC-Nb, FCC-Ni, δ-$NbNi_3$, and μ-Nb_7Ni_6 using DFT-based phonon calculations.

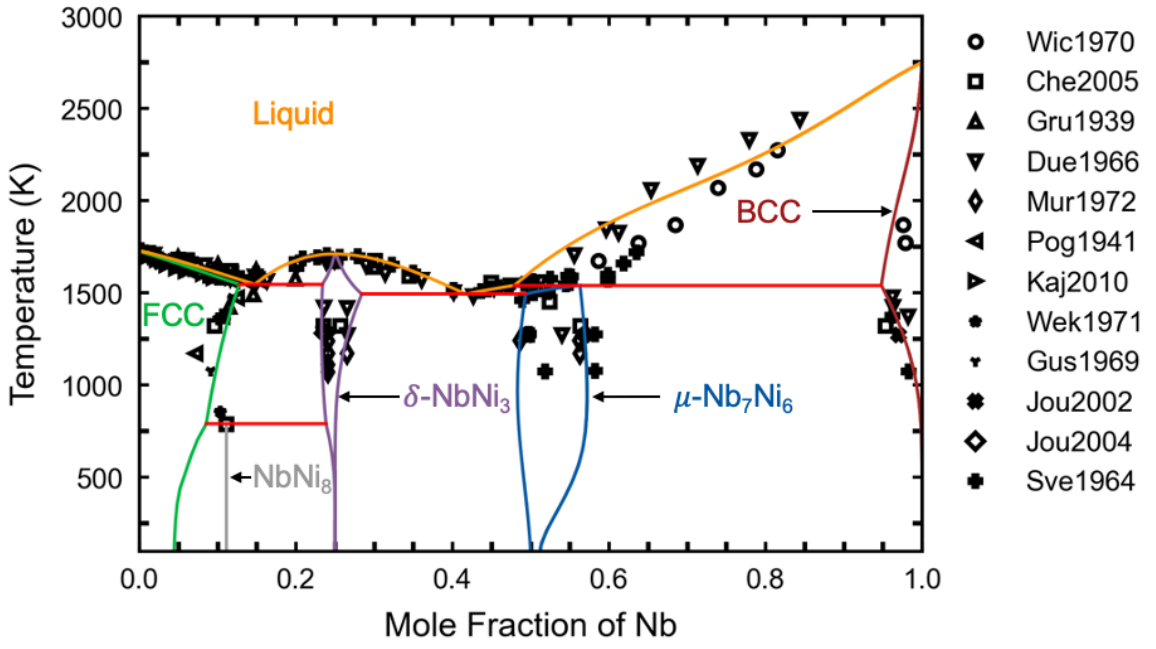


Fig. 7. Calculated phase diagram of the Nb-Ni system from the present CALPHAD modeling compared with available experimental data [16,18,19,22–24,26,38].

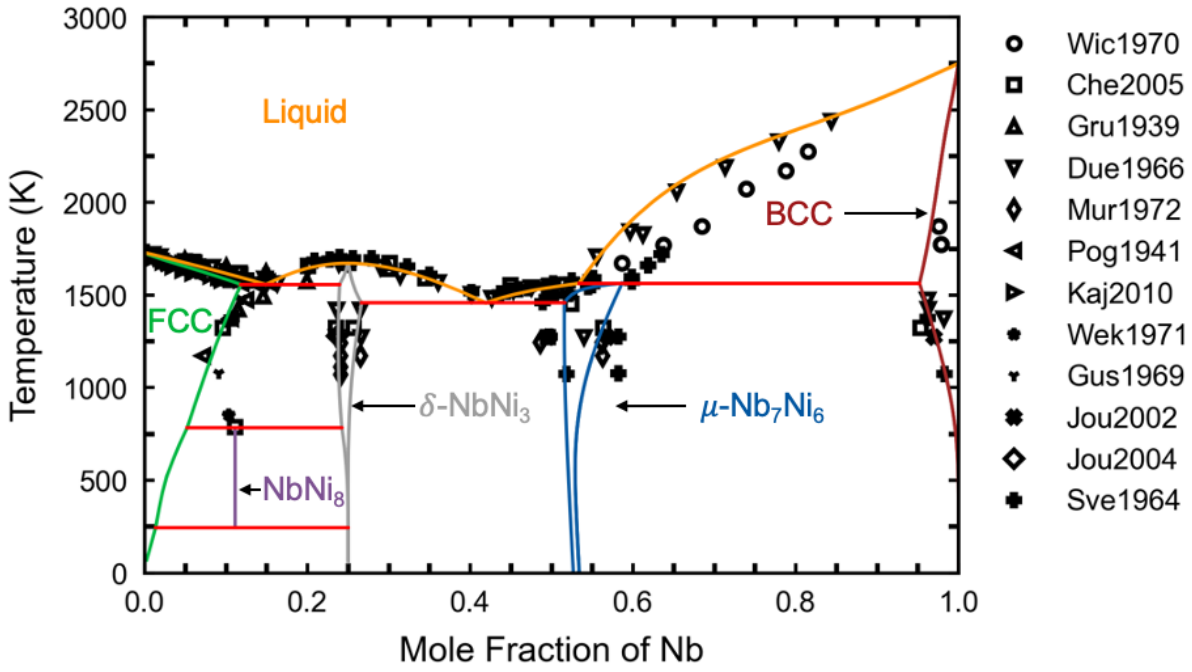


Fig. 8. Calculated phase diagram of the Nb-Ni system from Chen et al.'s CALPHAD modeling [19] in comparison with available experimental data [16,18,19,22–24,26,38]

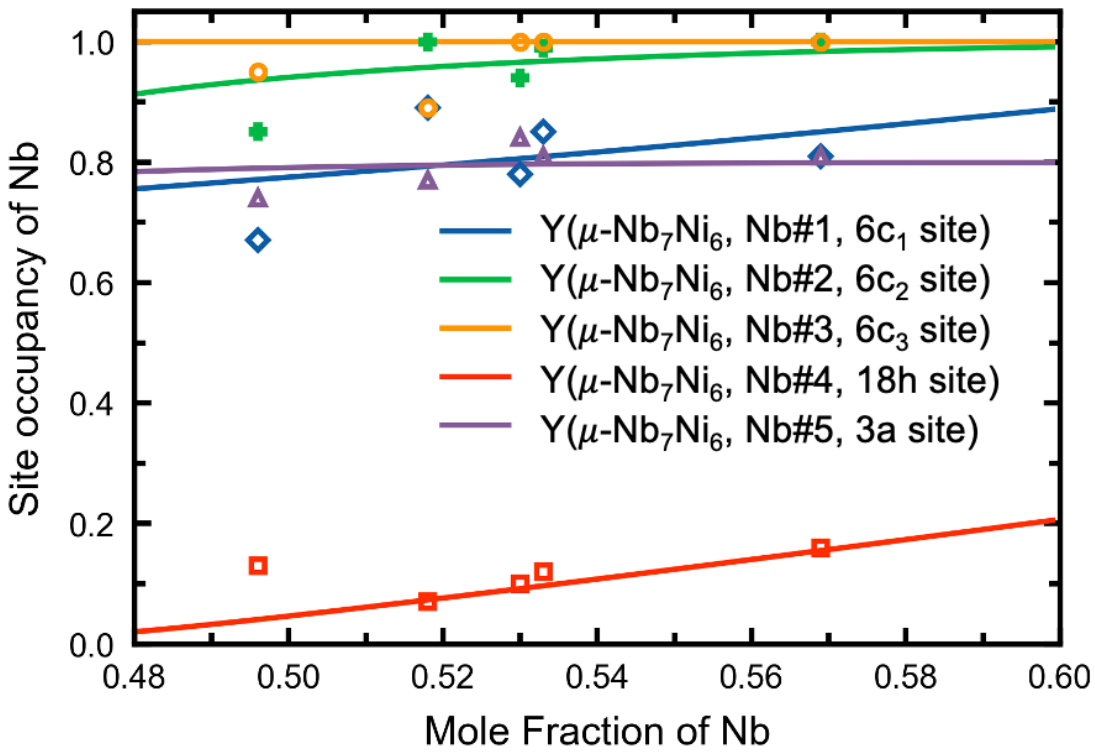


Fig. 9. Predicted site occupancies of Nb in μ-Nb_7Ni_6 at 1273 K (the lines) in comparison with experimental data (symbols) from Joubert et al. [16].

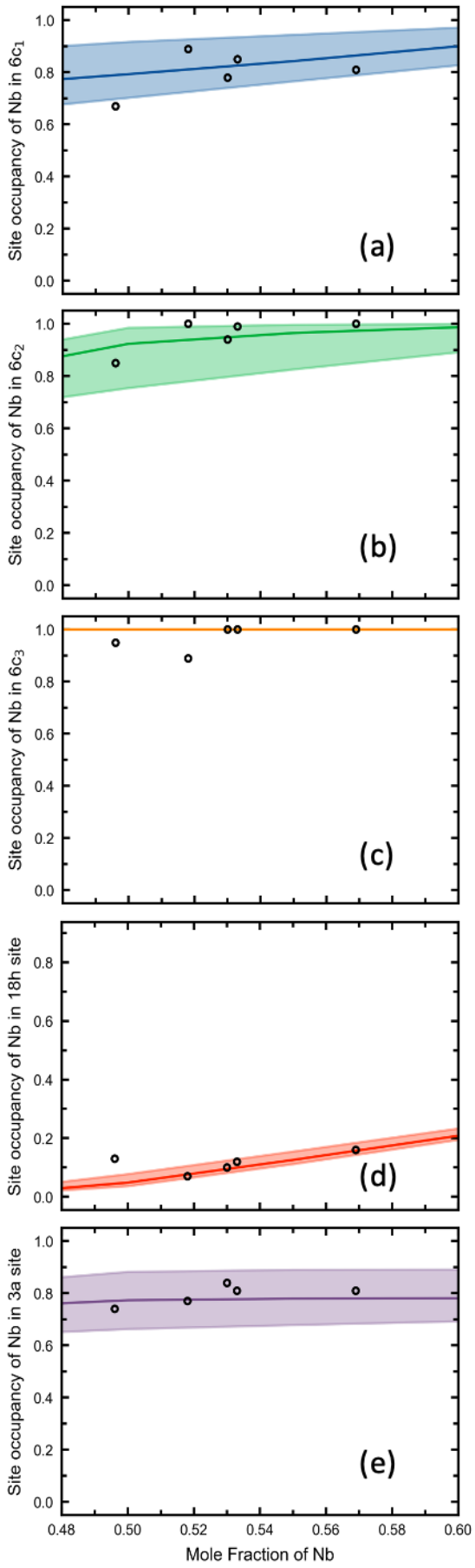

Site occupancy of Nb in 6c1
(a)
Site occupancy of Nb in 6c2
(b)
Site occupancy of Nb in 6c3
(c)
Site occupancy of Nb in 18h site
(d)
Site occupancy of Nb in 3a site
(e)
0.0
0.2
0.4
0.6
0.8
1.0
0.48
0.50
0.52
0.54
0.56
0.58
0.60
Mole Fraction of Nb

Fig. 10. Uncertainty quantification of site occupancies of Nb in the μ-Nb_7Ni_6 phase marked by the shaded region for each Wyckoff site: (a) , (b) …. [see also my comment in Fig 9]

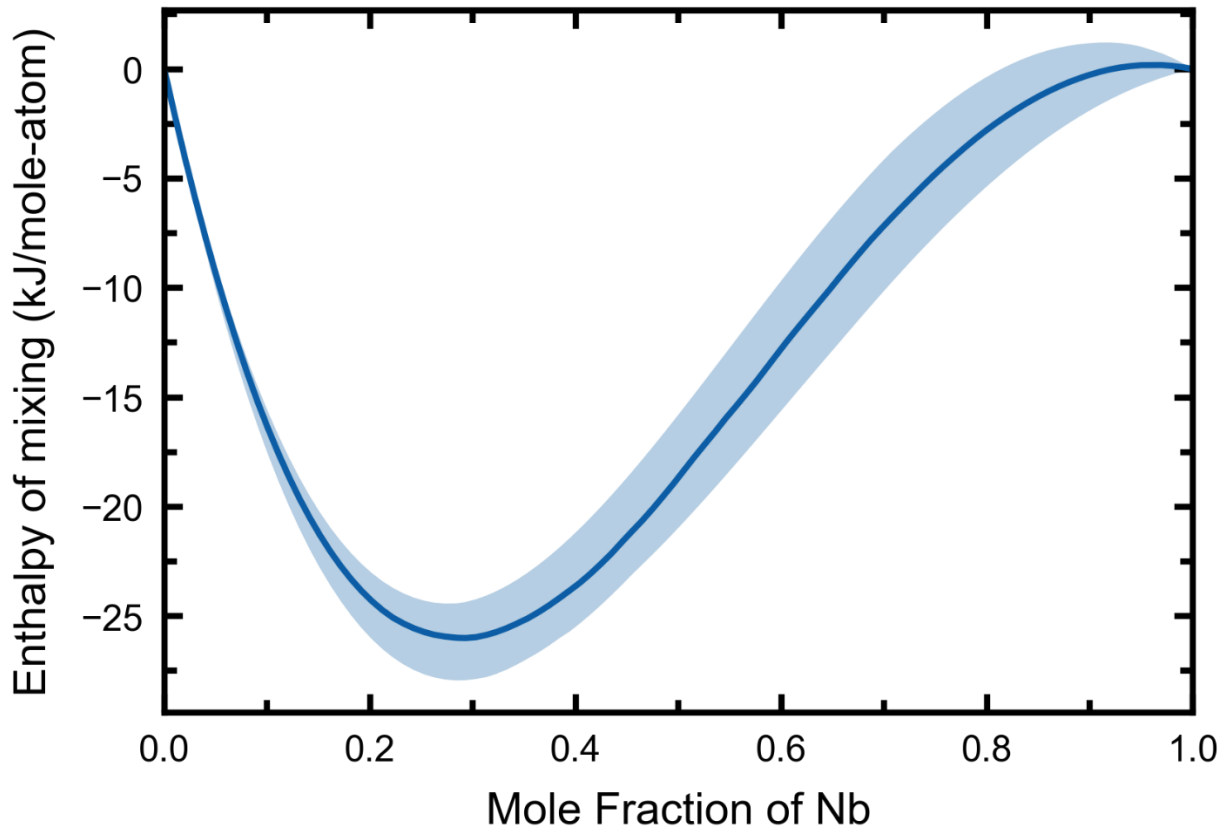


Fig. 11. Uncertainty quantification of the enthalpy of mixing in the liquid phase at 1823 K marked by shaded region.

References:

[1] F.C. Frank, J.S. Kasper, Complex alloy structures regarded as sphere packings. I. Definitions and basic principles, Acta Crystallogr. 11 (1958) 184–190. https://doi.org/10.1107/S0365110X58000487.

[2] B. Seiser, R. Drautz, D.G. Pettifor, TCP phase predictions in Ni-based superalloys: Structure maps revisited, Acta Mater. 59 (2011) 749–763. https://doi.org/10.1016/j.actamat.2010.10.013.

[3] T. Sugui, W. Minggang, L. Tang, Q. Benjiang, X. Jun, Influence of TCP phase and its morphology on creep properties of single crystal nickel-based superalloys, Mater. Sci. Eng. A. 527 (2010) 5444–5451. https://doi.org/10.1016/j.msea.2010.05.027.

[4] J.H. Alano, R.L. Siqueira, A.D. de Oliveira, G. dos Santos Vacchi, C.A. Della Rovere, S.E. Kuri, Effect of TCP phase formation on the electrochemical corrosion behavior of the nickel-based superalloy UNS N26455, Corros. Sci. 177 (2020). https://doi.org/10.1016/j.corsci.2020.108965.

[5] P. Pandey, S.K. Makineni, A. Samanta, A. Sharma, S.M. Das, B. Nithin, C. Srivastava, A.K. Singh, D. Raabe, B. Gault, K. Chattopadhyay, Elemental site occupancy in the L12 A3B ordered intermetallic phase in Co-based superalloys and its influence on the microstructure, Acta Mater. 163 (2019) 140–153. https://doi.org/10.1016/j.actamat.2018.09.049.

[6] P. Petrzak, K. Kowalski, M. Blicharski, Analysis of phase transformations in Inconel 625 alloy during annealing, Acta Phys. Pol. A. 130 (2016) 1041–1044. https://doi.org/10.12693/APhysPolA.130.1041.

[7] Y.Q. Mu, C.S. Wang, W.L. Zhou, L.Z. Zhou, Tensile Properties of Cast Alloy IN625 in Relation to δ Phase Precipitation, Acta Metall. Sin. (English Lett. 32 (2019) 535–540. https://doi.org/10.1007/s40195-018-0772-y.

[8] S. Qin, T.C. Novak, M.K. Vailhe, Z.K. Liu, A.M. Beese, Plasticity and fracture behavior of Inconel 625 manufactured by laser powder bed fusion: Comparison between as-built and stress relieved conditions, Mater. Sci. Eng. A. 806 (2021) 140808. https://doi.org/10.1016/j.msea.2021.140808.

[9] H.M. Tawancy, On the Precipitation of Intermetallic Compounds in Selected Solid-Solution-Strengthened Ni-Base Alloys and Their Effects on Mechanical Properties, Metallogr. Microstruct. Anal. 6 (2017) 200–215. https://doi.org/10.1007/s13632-017-0352-y.

[10] T. Fang, S.J. Kennedy, L. Quan, T.J. Hicks, The structure and paramagnetism of Ni3Nb, J. Phys. Condens. Matter. 4 (1992) 2405–2414. https://doi.org/10.1088/0953-8984/4/10/007.

[11] P.I. Kripyakevich, E.I. Gladyshevskii, E.N. Pylaeva, Compounds of the W6Fe7 type in the Ta-Ni and Nb-Ni systems, Sov. Phys. Crystallogr. 7 (1962) 165–168.

[12] L. Kaufman, H. Nesor, Ah = -267782 2092, 2 (1978) 81–108.

[13] A. Bolcavage, A Reassessment of the Calculated Ni-Nb P h a s e D i a g r a m, 17 (1996) 92–100.

[14] R. July, A thermodynamic calculation of the Ni-Nb phase diagram, 179 (1992) 177–185.

[15] H. Chen, Y. Du, Refinement of the thermodynamic modeling of the Nb–Ni system, Calphad. 30 (2006) 308–315.

[16] J.-M. Joubert, Y. Feutelais, Contribution of the Rietveld method to non-stoichiometric phase modeling. Part II: α-Tl5Te3 and μ Nb-Ni as experimental examples, Calphad. 26 (2002) 427–438.

[17] L. Kaufman, H. Nesor, Coupled phase diagrams and thermochemical data for transition metal binary systems—II, Calphad. 2 (1978) 81–108.

[18] J.-M. Joubert, B. Sundman, N. Dupin, Assessment of the niobium–nickel system, Calphad. 28 (2004) 299–306.

[19] H. Chen, Y. Du, Refinement of the thermodynamic modeling of the Nb – Ni system, 30 (2006) 308–315. https://doi.org/10.1016/j.calphad.2006.02.005.

[20] B. Bocklund, R. Otis, A. Egorov, A. Obaied, I. Roslyakova, Z.-K. Liu, ESPEI for efficient thermodynamic database development, modification, and uncertainty quantification: application to Cu–Mg—CORRIGENDUM, MRS Commun. 10 (2020) 1. https://doi.org/10.1557/mrc.2020.13.

[21] R. Otis, Z.-K. Liu, pycalphad: CALPHAD-based Computational Thermodynamics in Python, J. Open Res. Softw. 5 (2017) 1–11. https://doi.org/10.5334/jors.140.

[22] Y. Muramatsu, F. Roux, A. Vignes, Influence de la pression sur la croissance des couches des composes intermetalliques dans les couples de diffusion niobium/nickel, Trans. Japan Inst. Met. 16 (1975) 61–72.

[23] I.J. Duerden, W. Hume-Rothery, The equilibrium diagram of the system niobium-nickel, J. Less Common Met. 11 (1966) 381–387.

[24] J.C. Schuster, Experimental investigation of the Nb-Ni phase diagram, 0 (2005) 6019–6022. https://doi.org/10.1007/s10853-005-4553-4.

[25] and B.J.P. B.B. Argent, Heats of Formation of Nb−Ni Alloys, (n.d.).

[26] A. Wicker, C. Allibert, J. Oriole, E. Bonnier, Phase Equilibrium in Ni− Nb− Mo, Ni− Nb, and Ni− Mo Systems, CR Acad. Sci., Ser. C. 271 (1970) 273–275.

[27] M. Mathon, D. Connétable, B. Sundman, J. Lacaze, Calphad-type assessment of the Fe–Nb–Ni ternary system, Calphad. 33 (2009) 136–161.

[28] L.P. Déo, M.F. de Oliveira, Accuracy of a selection criterion for glass forming ability in the Ni–Nb–Zr system, J. Alloys Compd. 615 (2014) S23–S28.

[29] K. Santhy, K.C.H. Kumar, Thermodynamic reassessment of Nb-Ni-Ti system with order–disorder model, J. Alloys Compd. 619 (2015) 733–747.

[30] nb-ni-pub1.pdf, (n.d.).

[31] S.A. Pogodin, A.N. Selikmann, On the Phase Diagram of the System Nickel-Niobium, in: CR Acad. Sci. URSS, 1941: pp. 895–897.

[32] G. Grube, Uber die gewinnung von festen niob-nickel-legierungen durch reduktion von niobpentosyd be1 gegenwart von nickel., 367 (n.d.) 1–4.

[33] K. Kajikawa, K. Oikawa, F. Takahashi, H. Yamada, K. Anzai, Reassessment of liquid/solid equilibrium in Ni-rich side of Ni-Nb and Ni-Ti systems, Mater. Trans. 51 (2010) 781–786. https://doi.org/10.2320/matertrans.M2009351.

[34] L.N. Guseva, R.S. Mints, Y.S. Malkov, PHASE EQUILIBRIA IN THE NI-NI 3 AL-NI 3 NB SYSTEM AT 800-1200 C, IZVEST AKAD Nauk SSSR Met. SEPT-OCT. 1969,--5--, 186-190. (1969).

[35] W.E. Quist, C.J. van der Wekken, R. Taggart, D.H. Polonis, INTERMEDIATE COMPOUNDS Ni $ sub 8$ Nb (Cb) IN NICKEL-RICH NICKEL--NIOBIUM (COLUMBIUM) ALLOYS., Univ. of Washington, Seattle, 1969.

[36] C.J. van der Wekken, R. Taggart, D.H. Polonis, Short-Range Order and the Nucleation of Long-Range Order in Ni-Rich Nickel-Niobium Alloys, Met. Sci. J. 5 (1971) 219–223.

[37] J.F. Zhao, H.P. Wang, B. Wei, A new thermodynamically stable Nb2Ni intermetallic compound phase revealed by peritectoid transition within binary Nb-Ni alloy system, J. Mater. Sci. Technol. 100 (2022) 246–253. https://doi.org/10.1016/j.jmst.2021.07.001.

[38] V.N. Svechnikov, V.M. Pan, V.G. Korobeinikov, CONSTITUTION DIAGRAM OF NIOBIUM NICKEL ALLOY, Sb. Nauchnykh Tr. Instituta Metallofiz. Akad. Nauk Ukr. SSR (Ukrainian SSR) SR). (1964).

[39] V.S. Sudavtsova, Thermodynamic properties of binary Ni-V (Nb, Ta) melts, Metally. (1998) 44.

[40] E.M. Sokolvskaya, L.L. Mechkov, G.A. Tikhankin, No Title, Dokl. Akad. Nauk SSSR. 229 (1976) 914–916.

[41] G.D. Alekseev, D.M. Khazins, N.A. Kalinina, V. V Kruglov, On the accuracy of spatial position and tension of wire electrodes in proportional chambers, Prib. i Tekhnika Ehksperimenta. (1978) 47–50.

[42] Nb-NI-lyc.pdf, (n.d.).

[43] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. 1 (2013) 11002.

[44] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C.

Wolverton, The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies, Npj Comput. Mater. 1 (2015) 1–15.

[45] S.L. Shang, Y. Wang, T.J. Anderson, Z.K. Liu, Achieving accurate energetics beyond (semi-)local density functional theory: Illustrated with transition metal disulfides, Cu2ZnSnS4, and Na3PS4 related semiconductors, Phys. Rev. Mater. 3 (2019) 1–11. https://doi.org/10.1103/PhysRevMaterials.3.015401.

[46] K. Schaefers, J. Qin, M. Rösner-Kuhn, M.G. Frohberg, Mixing enthalpies of liquid Ni-V, Ni-Nb and Ni-Ta alloys measured by levitation alloying calorimetry, Can. Metall. Q. 35 (1996) 47–51.

[47] A. Kroupa, O. Zobač, K.W. Richter, The thermodynamic reassessment of the binary Al–Cu system, J. Mater. Sci. 56 (2021) 3430–3443. https://doi.org/10.1007/s10853-020-05423-7.

[48] G. Cacciamani, A. Dinsdale, M. Palumbo, A. Pasturel, The Fe-Ni system: Thermodynamic modelling assisted by atomistic calculations, Intermetallics. 18 (2010) 1148–1162. https://doi.org/10.1016/j.intermet.2010.02.026.

[49] S.-L. Shang, Y. Wang, D. Kim, Z.-K. Liu, First-principles thermodynamics from phonon and Debye model: Application to Ni and Ni3Al, Comput. Mater. Sci. 47 (2010) 1040–1048.

[50] Y. Wang, Z.-K. Liu, L.-Q. Chen, Thermodynamic properties of Al, Ni, NiAl, and Ni 3 Al from first-principles calculations, Acta Mater. (2004). https://doi.org/10.1016/j.actamat.2004.02.014.

[51] Y. Wang, Z.K. Liu, L.Q. Chen, Thermodynamic properties of Al, Ni, NiAl, and Ni3Al from first-principles calculations, Acta Mater. 52 (2004) 2665–2671. https://doi.org/10.1016/j.actamat.2004.02.014.

[52] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B. 54 (1996) 11169.

[53] P.E. Blöchl, Projector augmented-wave method, Phys. Rev. B. 50 (1994) 17953–17979. https://doi.org/10.1103/PhysRevB.50.17953.

[54] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865.

[55] O. Redlich, A.T. Kister, Algebraic representation of thermodynamic properties and the classification of solutions, Ind. Eng. Chem. 40 (1948) 345–348.

[56] A.T. Dinsdale, SGTE data for pure elements, Calphad. 15 (1991) 317–425. https://doi.org/10.1016/0364-5916(91)90030-N.

[57] M. Hillert, The compound energy formalism, J. Alloys Compd. 320 (2001) 161–176.

[58] B. Bocklund, R. Otis, A. Egorov, A. Obaied, I. Roslyakova, Z.K. Liu, ESPEI for efficient

thermodynamic database development, modification, and uncertainty quantification: Application to Cu-Mg, MRS Commun. 9 (2019) 618–627. https://doi.org/10.1557/mrc.2019.59.

[59] D. Foreman-Mackey, D.W. Hogg, D. Lang, J. Goodman, emcee : The MCMC Hammer , Publ. Astron. Soc. Pacific. 125 (2013) 306–312. https://doi.org/10.1086/670067.

[60] J. Goodman, J. Weare, Ensemble Samplers With Affine, Commun. Appl. Math. Comput. Sci. 5 (2010) 65–80.

[61] B.B.P. Carlin, T.A. Louis, Book Reviews, J. Am. Stat. Assoc. 109 (2014) 1325–1337. https://doi.org/10.1080/01621459.2014.963405.

[62] J. Trivisonno, S. Vatanayon, M. Wilt, J. Washick, R. Reifenberger, Temperature dependence of the elastic constants of niobium and lead in the normal and superconducting states, J. Low Temp. Phys. 12 (1973) 153–169. https://doi.org/10.1007/BF00654733.

[63] G. Simmons, H. Wang, A Handbook of Single Crystal Elastic Constants and Calculated Aggregate Properties, (1971).

[64] K.H. Hellwege, J.L. Olsen, Metals: Phonon states, electron states and Fermi surfaces. Subvolume a. Phonon states of elements, electron states and Fermi surfaces of alloys. Metalle: Phononenzustaende, Elektronenzustaende und Fermiflaechen. Teilband a. Phononenzustaende von Elementen, (1981).

[65] S.L. Shang, Y. Wang, Y. Du, Z.K. Liu, Entropy favored ordering: Phase stability of Ni3Pt revisited by first-principles, Intermetallics. 18 (2010) 961–964. https://doi.org/10.1016/j.intermet.2010.01.011.

[66] C.J. Van Der Wekken, R. Taggart, D.H. Polonis, Short-range order and the nucléation of long-range order in ni-rich nickel-niobium alloys, Met. Sci. J. 5 (1971) 219–223. https://doi.org/10.1179/030634571790439487.

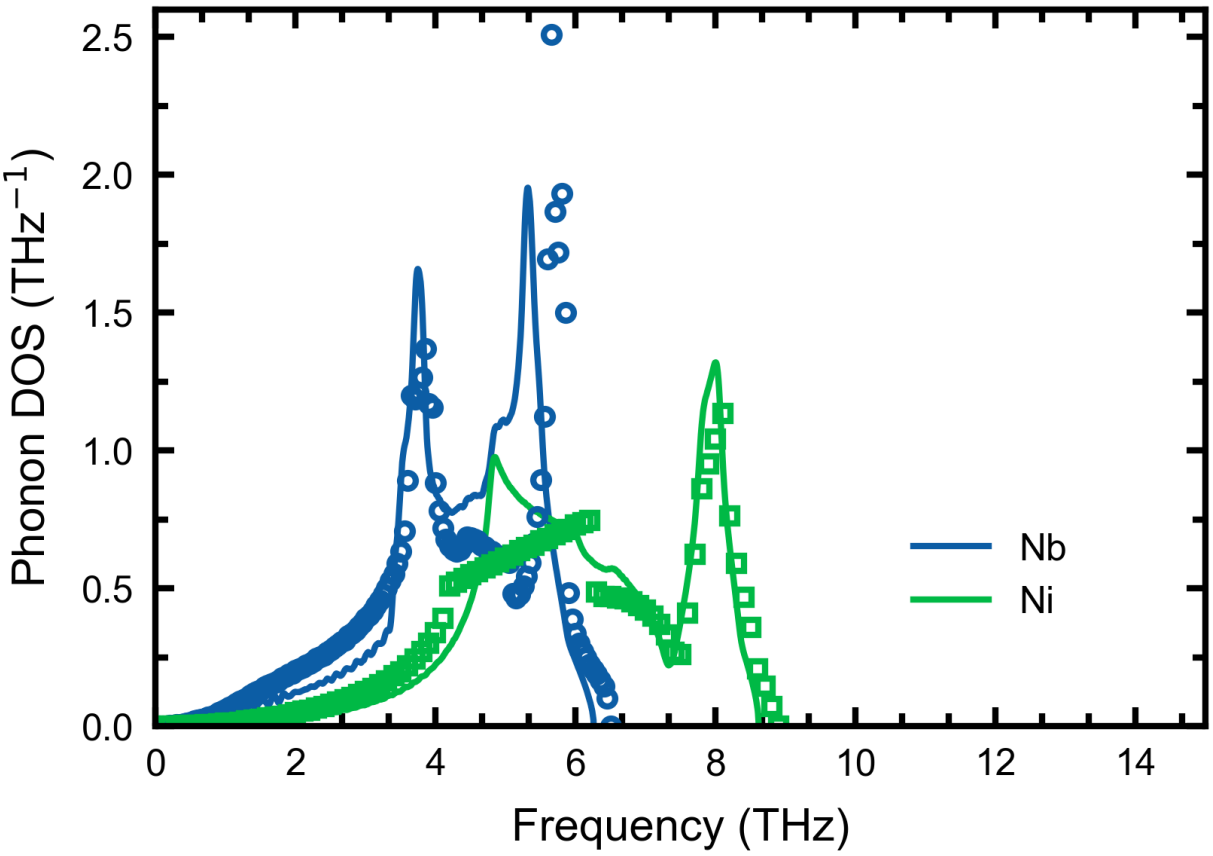


Fig. 12. Predicted phonon density of states of the BCC-Nb phase (blue line), the FCC-Ni phase (green line), using DFT-based phonon calculations in comparison with experimental data [64].

Predicting ideal shear strength of dilute multicomponent Ni-based alloys by an integrated first-principles calculations, CALPAHD modeling approach and correlation analysis

Shuang Lin[1,*], Shun-Li Shang[1,], John D. Shimanek[1], Yi Wang[1],

Allison M. Beese[1,2], and Zi-Kui Liu[1]

[1] Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA, 16802, USA

[2] Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA 16802, USA

* Corresponding author: spl5745@psu.edu

Abstract

The present work predicts the ideal shear strength of dilute $Ni_{34}XZ$ ternary alloys determined by first-principles calculations of pure alias shear deformations. The major elements of Inconel alloys are examined to understand composition and concentration effects on ideal shear strength. The alloy ideal shear strength decreased in the following order of alloying elements: Co > Mn~Fe > (pure) Ni > Cr > Al > Ti > Mo > Si > Nb. The ideal shear strength shows a roughly linear correlation with alloying concentration at the dilute side and concentration effect splits depending on the relative atomic volume of the alloying element compared to that of pure Ni. The alloying elements with larger atomic volume decrease the ideal shear strength of the alloy, and with smaller atomic volumes increase the ideal shear strength of the alloy. Binary and ternary interaction parameters are quantified by CALculation of PHAse Diagrams (CALPHAD) method, indicating

that atomic volume differences and the standard state crystal structure difference determine the degree of the interaction on the ideal shear strength. The ideal shear strength of the Ni-based dilute ternary system was estimated using a linear combination of binary systems. The variations of ideal shear strength are quantitatively correlated with features of the pure elements X and Z, showing that atomic size of alloying element is the most important factor strongly associated in calculating ideal shear strength.

1 Introduction

The ideal shear strength is the shear stress required to plastically deform a crystal lattice without defects [1]. This upper bound material property can be predicted by first-principles calculations based on density functional theory (DFT) [2]. Recently, motivated by reducing reliance on experimental fitting data in crystal plasticity finite element method (CPFEM), a DFT-based first-principles calculations approach was proposed to predict the strain-hardening behavior of pure Ni [3]. The flow resistance was evaluated by the Peierls–Nabarro model [4], with the input of the ideal shear strength and elastic properties calculated by DFT-based calculation. The elastic hardening behavior of the pure edge dislocations is captured at small strains. For large strains, the linear model was proposed to combine both edge and screw dislocation contributions to flow resistance, giving good agreement between the CPFEM results and experimental results of pure Ni single crystal [3]. A continuous work quantitively examined the effects of 26 alloying elements on the ideal shear strength in dilute $Ni_{11}X$ alloys via first-principles techniques [5]. Through the combination of the Mg alloys data, the trends of the ideal shear strength are explored by the feature selection method, showing that the elemental properties such as elemental volume and electronegativity are the most two important factors in ideal shear strength variation. The elastic modulus C_{55} of the Ni based alloys show a strong linear correlation with the ideal shear strength. The ideal shear strength data of $Ni_{11}X$ were severed to predict the macroscopic single crystal deformation behavior for small and large strain [5].

Although the ideal shear strength of the pure Ni and binary multicomponent systems $Ni_{11}X$ are well investigated in the previous study, the higher order composition effect on the ideal shear strength is still missing. One of the most successful approaches to model materials properties from

binary to ternary systems is the CALculation of PHAse Diagrams (CALPHAD) approach, which was originally developed for modeling thermodynamic properties by integrating experimental phase equilibrium and thermochemical data [6,7]. The CALPHAD approach relies on interaction parameters from unary, binary, and ternary systems to model more complex multi-component systems. Thus, with all interaction parameter data, the CALPHAD approach can cover the whole composition of the system, which includes uninvestigated compositions. In addition, the most challenging thing after adding the third element in the dilute alloys system is that the possible structures will exponentially increase when calculating the ideal shear strength. Therefore, the present work examines all the possible structures of $Ni_{34}XZ$ and chooses the lowest energy structure to perform the ideal shear strength calculation.

The present work aims to predict the ideal shear strength of the dilute ternary alloys $Ni_{34}XZ$ based on first-principles calculations. All possible structures after adding the third elements are examined with the system $Ni_{34}FeX$. The composition and concentration effects on ideal shear strength are explored combining with $Ni_{11}X$ data. The CALPHAD method evaluates all the binary parameters ${}^{0}\phi_{ij}$ and the ternary parameters ϕ_{ijk} with input from first-principles calculations. The feature selection method is also used in the present work to investigate correlation between the calculated ideal shear strength and features of pure elements.

2 Methods

2.1 Pure alias shear deformation

Pure alias shear deformation as shown in *Figure 1* (a) with one sliding plane involved was adopted to predict ideal shear strength. Here 'pure' means full relaxations of atomic positions, cell shape, and cell volume, except for the fixed shear angle. This type of shear deformation is more realistic because the displacement of one atomic layer (e.g., the top layer) influences only the next atomic layer at first, then the relaxation will propagate from top down through the entire cell with fixed shear displacement [8]. Previous investigations show that the $\{111\}\langle 11\bar{2}\rangle$ shear deformation is the primary slip system of face-centered-cubic (FCC) metals [9]. In the present work, DFT-based first-principles calculations were performed to determine ideal shear strength of dilute Ni-X-Z ternary alloys with alloying elements X and Z representing Al, Co, Cr, Fe, Mn, Mo, Nb, Si, and Ti. These elements were chosen based on the major compositions of Inconel alloys, for example Inconel 625 and Inconel 718 [10,11]. A supercell of the conventional FCC lattice, i.e., the 36-atom orthorhombic cell (α=β=γ=90°), with 3 {111} layers and 12 atoms on each layer is employed, was used to predict ideal shear strength based on the $\{111\}\ \langle 11\bar{2}\rangle$ shear deformation. *Figure* 1 (b) shows this 36-atom orthorhombic cell with the lattice vectors **a** along the $[11\bar{2}]$ direction, **b** the $[\bar{1}10]$, and **c** the [111], respectively. The initial lattice parameters along these three directions are 8.625, 7.470, and 6.099 Å, respectively. The deformed lattice vector $\bar{\mathrm{R}}$ can be obtained by [12],

$$\bar{\mathbf{R}} = \mathbf{RD} \qquad \textit{Eq. 1}$$

where **R** is the original lattice vector before alias shear, and **D** is the deformation matrix [8]:

$$\mathbf{D}_{[11\bar{2}]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \varepsilon & 0 & 1 \end{bmatrix} \qquad \textit{Eq. 2}$$

Here ε is the magnitude of engineering strain, defined as the ratio between displacement applied for first-principles calculations and the height of the supercell. The relaxation of pure alias shear deformation was performed by the external optimizer GADGET [13] to fix the angle during pure shear deformation of $\{111\}\langle 11\bar{2}\rangle$.

2.2 First-principles calculations

All DFT-based first-principles calculations in the present work were performed by the Vienna Ab initio Simulation Package (VASP) [14] using the projector augmented wave (PAW) method [15]. The exchange-correlation functional is described by the generalized gradient approximation (GGA) [16]. A 3 × 4 × 5 k-point meshes together with a 350 eV plane wave cutoff energy were employed based on convergent tests. The energy convergence criterion of electronic self-consistency was chosen as 2×10^{-5} eV per supercell in all the calculations. The reciprocal space energy integration was performed by the Methfessel-Paxton technique [17] for structural relaxations with a 0.2 eV smearing width. Because of the magnetic nature of Ni, the spin-polarized approximation was used in all first-principles calculations.

The configuration of the $Ni_{34}XZ$ supercell is determined by examining all possible structures determined by the ATAT code [18], indicating that there are 5 independent structures when X and Z elements on the same (111) plane; see one of the structures in *Figure 2* (a), and 8 independent structures when X and Z elements stay at different (111) planes; see one of the structures in *Figure 2* (b). The alloying element X (purple atom) is fixed and the numbers indicate equivalent sites when substituting the second alloying element Z.

Taking $Ni_{34}FeX$ as an example, *Figure 3* shows the relative energies in terms of the total 13 independent configurations, where the structures S1 to S5 represent different alloying site scenarios when X and Z are at the same layer, while the structures D1 to D8 represent different alloying site scenarios when X and Z are at the different layer. For example, the structure S1 means the second alloying element Z is substituted at site 1 in *Figure 2* (a) and D3 means the second alloying element Z is substituted at site 3 in *Figure 2* (b). *Figure 3* suggests that the configurations of S1, S2, D1, and D7 process a relatively higher energy, ranging from 0.06 to 0.34 eV per supercell, while the other 9 configurations process a relatively low energy. The S3 configuration is one of the lowest energy configurations for most elements, except for the systems $Ni_{34}FeAl$ and $Ni_{34}FeMn$ showing the D2 configuration is the lowest energy configuration. Since the configurations S3 and D2 are both one of the maximum entropy structures[19] with alloying element X and Y separate the most, the energy difference is small for the systems $Ni_{34}FeAl$ (0.0046 eV/supercell) and $Ni_{34}FeMn$ (0.0086 eV/supercell). In addition, the examination shows the energy difference effect on ideal shear strength calculation of the S3 and D2 configuration is negligible. Since the Ni based alloy solid solution should be the mixture of the lowest energy configurations, therefore, the configuration S3 was selected in the present work to perform calculations of ideal shear strength.

2.3 CALPHAD modeling of ideal shear strength

The CALPHAD approach has been extended to model any properties related to individual phases, such as elastic properties [20], molar volume [21], diffusion coefficient [22], and stacking fault energy [23] in multicomponent systems. In the present work, the CALPHAD approach is used in

modeling ideal shear strength of dilute multicomponent Ni-based alloys. The CALPHAD modeling of phase-related properties can be written as follows [23]:

$$\boldsymbol{\phi} = {}^{0}\boldsymbol{\phi} + \Delta\boldsymbol{\phi} \qquad \textit{Eq. 3}$$

$${}^{0}\boldsymbol{\phi} = \sum_{\mathbf{i}} \mathbf{x}_{\mathbf{i}}^{\mathbf{0}}\, \boldsymbol{\phi}_{\mathbf{i}} \qquad \textit{Eq. 4}$$

where ϕ is a property of a multicomponent system and ${}^{0}\phi$ represents linear mixing of individual elements' properties in terms of their mole fraction (x_i^0). Lastly, $\Delta\phi$ represents the interaction among alloying elements, written as:

$$\Delta\phi = \Delta\phi_{conf} + \sum_{i} x_i x_j \sum_{j>i} \sum_{L=0} {}^{L}\phi_{ij}\,(x_i - x_j)^{L} + \sum_{i} \sum_{j>i} \sum_{k>j} x_i x_j x_k \phi_{ijk} \qquad \textit{Eq. 5}$$

where $\Delta\phi_{conf}$ is usually described by ideal atomic configurational entropy, which is ignored in the present work. ${}^{L}\phi_{ij}$ is the L^{th}-order binary interaction, and ϕ_{ijk} the ternary interaction parameter. The superscript L of $(x_i - x_j)$ indicates its power, and i, j, and k denotes different species in multicomponent system. The present work aims to evaluate all the binary parameters ${}^{0}\phi_{ij}$ and the ternary parameters ϕ_{ijk} with input from first-principles calculations.

The CALPHAD modeling approach as shown in *Eq. 5* was adopted to describe the ideal shear strength by following equation:

$$\tau_{id} = {}^{0}\tau + \Delta\tau = \sum_{i} x_i^0\, \tau_i + \sum_{i} x_i x_j \sum_{j>i} {}^{0}\phi_{ij} + \sum_{i} \sum_{j>i} \sum_{k>j} x_i x_j\, x_k \phi_{ijk} \qquad \textit{Eq. 6}$$

where τ_{id} is ideal shear strength of multicomponent system, ${}^{0}\phi_{ij}$ is zero order of binary fitting parameter and ϕ_{ijk} is ternary fitting parameter, and x_j, x_k represent mole fractions of alloying element X and Z respectively. Since we are investigating the dilute Ni-based system, ${}^{0}\tau$ is

assumed equal to the ideal shear strength of pure Ni, and the mole fraction of Ni (x_i) in *Eq. 6* is assumed to be 1. Therefore, all the binary parameters and the ternary parameters were calculated using the following equation:

$$\tau_{id} = \tau_0^{Ni} + \sum_j x_j \, {}^0\phi_{Ni,X} + \sum_j \sum_k x_j x_k \phi_{Ni,X,Z} \quad \text{with } \tau_0^{Ni} = 5.09 \text{ GPa} \qquad \textit{Eq. 7}$$

where τ_0^{Ni} is ideal shear strength of pure Ni, ${}^0\phi_{Ni,X}$ is zero order of binary fitting parameter between Ni and alloying element X, and $\phi_{Ni,X,Z}$ is ternary fitting parameter among Ni, alloying element X and Z. x_j, x_k represents mole fraction of alloying element X and Z respectively.

2.4 Correlation analysis and feature selection

To understand connections between physical informed features of alloying elements and the resulting ideal shear strength, the variation of ideal shear strength was examined by correlation analysis and feature selection algorithms. The major material features were chosen based on the features table in our previous study regarding the effect of 26-alloying elements on ideal shear strength [5], including atomic, periodic, Elastic, thermodynamic, lattice, and electronic properties. In addition to the material specific properties, the present work includes extra features, which are DFT-calculated features and combination of materials properties feature based on some surrogate models and criteria. For example, 'Gb2pih' in the feature table represents the feature of $\frac{Gb}{2\pi h}$, which equals to ideal shear strength based on the work of Frenkel [24]. 'DFTBvG' represents the ratio between bulk modulus and shear modulus based on DFT-based predictions. 'Cauchy' is 'Cauchy pressure' defined by the difference of elastic constants C_{12} and C_{44}. These two features are usually the intrinsic ductility criteria indicating whether a material is ductile or brittle [25]. *Table 6* in the appendix lists all the features and explanations.

Feature selection methods can be categorized into filter type methods and wrapper type methods [26]. Filter type methods are separate from any machine learning algorithms and are based on statistical tests of correlation. For example, the coefficient of determination of linear fitting (i.e., the R^2) is a filter type method. In addition to the coefficient of determination, the present work includes several other filter methods, such as the maximal information coefficient (MIC), F-tests, and feature scores from a regression Relief algorithm. MIC evaluates the linear or non-linear correlation between two variables, conceptualized by a partitioning grid separating the scatterplot of two variables [32]. F-tests statistically test the variance of two variables, extending the R^2 method. The relief algorithm quantifies correlations by adjusting feature weights based on their ability to explain the variance of targeted features [27].

In wrapper type methods, a subset of features will be used to train a machine learning model. A metric of model performance, for example the mean squared error (MSE), will determine whether features are added into or removed from the feature subset. Both the forward and the backward feature selection modes can be used in the wrapper methods. Forward means features are sequentially added into a null set and backward means features are sequentially removed from a full-feature set. All filter-type method and wrapper-type method are coded by MATLAB (version R2020b used herein) [28]. A rational quadratic gaussian process regression (GPR) model was used in the wrapper methods based on our tests to run the forward and backward feature selection, with the 5-fold cross validation used to determine model performance, i.e., the MSE value. 1000 iterations in wrapper method were used in the present work to guarantee the converge of feature

selections and the higher iteration times do not significantly change the frequency of the selected features.

3 Results and discussion

3.1 First-principles calculations

In the present work, the calculated ideal shear strength of pure Ni is 5.09 GPa, which agrees well with the previous results using the same method (5.13 GPa) [5]. A probable explanation of the 0.04 GPa difference comes from the different size of the supercell used in the calculation. Compared with 12-atom supercell in previous study, a larger supercell with 36 atoms was used in the present work. The Table 1 shows the relaxed total energy of different size of pure Ni when reaching to the ideal shear strength. Compared with the smaller supercell size, a larger supercell processes a higher relaxed energy. In the present study, all the results consistently show that a higher relaxed energy is associated with a lower ideal shear strength. The ideal shear strength could be experimentally measured by such as nanoindentation [29] and micropillar compression method [30], which results approach to DFT-based predictions, at least in similar magnitudes. For instance, compared to the ideal shear strength of pure Ni by first-principles calculations, the nanoindentation result of pure Ni is around 8 GPa [31], and the difference could be explained by the complicated triaxial stress state during the test.

Table 2 summarized the $Ni_{34}XZ$ ideal shear strength values, which fall in the range of 4.16 to 5.24 GPa. It suggests that $Ni_{34}CoCo$ has the highest ideal shear strength, but $Ni_{34}NbNb$ has the lowest. Additionally, there is a clear trend in ideal shear strength from the highest to the lowest when alloying elements into dilute Ni alloys, i.e., Co > Mn~Fe > Cr > Al > Ti > Mo > Si > Nb. By setting ideal shear strength of pure Ni as reference, Co, Mn, and Fe can be considered as the strengthening alloying elements for dilute multicomponent Ni alloys, while Cr, Al, Ti, Mo, Si, and Nb can be considered as the softening alloying elements.

Figure 4 shows one example ($Ni_{34}CoNb$) of ideal shear strength and total energy evolution with different applied shear displacements. It shows the increase of shear stress with increasing shear strain before reaching the maximum point, i.e., the ideal shear strength. Besides shear stress, the slope of total energy with respect to the displacement continuously increases when approaching the ideal shear strength. Since the force definition is $\mathrm{F} = -\frac{\partial \mathrm{E}}{\partial \mathrm{r}}$, the maximum stress occurs at the point of largest slope in *Figure 5*. After reaching the maximum point, the stress sharply drops. When the total energy reaches the highest point, the stress of the system is zero. *Figure 6* shows the calculated shear stress versus shear strain relationships of $Ni_{34}CoCo$ (blue circles), $Ni_{34}CoNb$ (green triangles) and $Ni_{34}NbNb$ (purple squares). The red symbols represent the ideal shear strength among these three systems. The ideal shear strength of $Ni_{34}CoNb$ is roughly equal to the average ideal shear strength of the $Ni_{34}CoCo$ and $Ni_{34}NbNb$ system. The interaction parameters will be quantified in the CALPHAD result discussion. The calculation results also suggests that a higher ideal shear strength occurs at a lower shear strain.

Table 3 summarizes the results of ideal shear strength of $Ni_{35}X$, $Ni_{34}X_2$ from the present work and $Ni_{11}X$ results from previous work [5]. It is worth mentioning that there is only one independent

configuration of $Ni_{35}X$ and $Ni_{11}X$ when one alloying element on the top layer of the supercell as shown in Table 3. The ideal shear strength difference among $Ni_{35}X$, $Ni_{34}X_2$ and $Ni_{11}X$ could be attributed to the alloying elements' concentration difference. *Figure 7* shows relative ideal shear strength as function of X alloying elements concentration with the reference being the ideal shear strength of pure Ni (5.09 GPa), indicating that the ideal shear strength roughly linearly correlates with alloying element concentration (at least in dilute Ni-based alloy), but the trend splits in terms of the effect of alloying elements. The only exception is the ideal shear strength of $Ni_{11}Mo$ increase as the Mo concentration increase. A likely explanation is that the ideal shear trend of $Ni_{11}Mo$ from alloying element concentration effect only works well at dilute Ni-based alloy, with Mo concentration under around 6%. When the concentration higher than 6%, the ideal shear strength will fluctuate rather than keep decreasing. Higher concentration of strengthen alloying elements (Co, Mn, Fe) will lead to higher ideal shear strength. However, higher concentration of softening alloying elements (Cr, Al, Ti, Mo, Si, Nb) will lead to lower ideal shear strength of the alloys. The largest decrease of ideal shear strength is due to the addition of Nb, while the largest increase is caused by adding Co. The split threshold is atomic volume of host element, pure Ni. The atomic volume of Co, Mn and Fe all are close or smaller than pure Ni; whereas the others are all larger than pure Ni. It is also suggested that the ideal shear strength determined by the first-principles calculation has negative correlation with alloying elements' atomic volume, i.e., the alloying elements with larger atomic volume decrease the ideal shear strength of the alloy. These results can be explained by the total relaxed energy change when applying the displacement on the shear plane. Compared with reference element, pure Ni, alloying elements with larger atomic volume causes lower total relaxed energy, which always lead to the higher ideal shear strength in the

present work. The relationship between ideal shear strength and alloying elements' atomic volume will be quantified the in the correlation analysis discussion.

3.2 CALPHAD modeling of ideal shear strength

The CALPHAD method evaluates all the binary parameters ${}^{0}\phi_{ij}$ and the ternary parameters ϕ_{ijk} with input from first-principles calculations. Based on CALPHAD approach as shown in *Eq. 7*, the binary interactions are calculated as follows:

$$\tau_{id} = \tau_0^{Ni} + {}^{0}\phi_{Ni,X}x_i = 5.09 + {}^{0}\phi_{Ni,X} * \frac{2}{36} \qquad \textit{Eq. 8}$$

Table 4 summarizes the results of all binary interactions from 9 systems ($Ni_{34}X2$). The binary interaction of the alloying element Co, Mn and Fe is positive, which means these elements have strengthen effect on ideal shear strength of dilute Ni-based alloys. The sign of ${}^{0}\phi_{Ni,X}$ indicates the strengthen or softening effect on ideal shear strength and the absolute value of ${}^{0}\phi_{Ni,X}$ indicates the degree of strengthening or softening. Agreed with the analysis of previous investigation, it follows the conclusion that larger alloying atomic volume is associated with a smaller ideal shear strength. Furthermore, to determine the ternary interaction parameter, based on the pure Ni and binary interaction parameter, we can obtain the ternary interaction $\phi_{Ni,X,Z}$ using the following equation:

$$\tau_{id} = \tau_0^{Ni} + {}^{0}\phi_{Ni,X}x_j + {}^{0}\phi_{Ni,Z}x_k + \phi_{Ni,X,Z}x_j x_k$$
$$= 5.09 + \frac{1}{36} * \left({}^{0}\phi_{Ni,X} + {}^{0}\phi_{Ni,Z} \right) + \left(\frac{1}{36}\right)^2 * \phi_{Ni,X,Z} \qquad \textit{Eq. 9}$$

Table *5* shows the results of the ternary interaction parameters $\phi_{Ni,X,Z}$ of all 45 $Ni_{34}XZ$ systems. Similar to the binary interaction parameter, the sign of $\phi_{Ni,X,Z}$ indicated positive or negative effect on the ideal shear strength from the interaction between alloying elements X and Z and the absolute value of $\phi_{Ni,X,Z}$ indicated the degree of the strengthening or softening effect. For most cases, the ternary interaction parameter $\phi_{Ni,X,Z}$ is negative, which means the ideal shear strength of dilute Ni-based alloy will decrease due to the ternary interaction. Some large ternary interaction parameters are observed in the

Table 5. The large ternary interaction parameters of Nb-Co, Nb-Mn, and Nb-Fe can be explained by the large atomic volume difference between two alloying elements, but it is hard to find a clear trend of all these ternary interaction $\phi_{Ni,X,Z}$ fitting results only in terms of their atomic volumes. For example, a relatively large interaction value is also found for Nb and Cr related interactions, such as Cr-Al and Cr-Si, but the atomic volumes difference is not significant, especially for the Cr-Al system. These large ternary interactions parameter in Cr-Al and Cr-Si can possibly be understood from structure similarity principles, since each element has a different standard state crystal structure at room temperature: Cr (BCC), Al (FCC), Si (diamond cubic). Therefore, the ternary interaction discrepancy can be majorly attribute to the atomic volumes difference and the standard state crystal structures difference.

Another interesting finding is non-linear effect is examining the present work. *Figure 8* shows a promising prediction between DFT-predicted ideal shear strengths and linear combination predictions based on the binary system ($Ni_{34}X2$), with R^2 value of 0.977. The linear combination ideal shear strength of $Ni_{34}XZ$ is defined as followed:

$$\tau_{id}^{linear}(Ni_{34}XZ) = 0.5 * \tau_{id}(Ni_{34}X_2) + 0.5 * \tau_{id}(Ni_{34}Z_2) \qquad Eq.10$$

Deviation from τ_{id}^{linear} and τ_{id}^{DFT} indicates nonlinear effects not captured in the above equation. The result suggests that in the dilute Ni-based alloys, the ternary interaction does not significantly change ideal shear strength. The ideals shear strength of Ni-based dilute ternary system could be estimated by the linear combined of binary systems.

3.3 Correlational analyses based on elemental properties

Table 7 shows the correlational analysis result of alloying element X and Z variations in ideal shear strength. We want to infer what physical features, or combination of features, could best explain the relationship between ideal shear strength and alloying element properties. There is a general agreement between the filter methods that among the most important features are the DFTv0 (atomic volume based on DFT-based calculations). *Figure 9* shows the ideal shear strength versus the average atomic volumes of alloying elements X and Z, suggesting that, in general, larger atomic volumes of alloying elements associating with lower calculated ideal shear strength. This trend could be explained by total energy variation when applying displacement on shear plane. Previous investigation shows relaxed energy decreased with increasing alloying element atomic volumes.

In addition to the filter-type selected features, the wrapper method indicates DebyeT (Debye temperature), Radius_vDW(Van der Waals atomic radius), USFE (Unstable stacking fault energy) are the common features by using forward and backward feature selection. The unstable stacking fault energy is selected 436 times for the forward selection method and 846 times for the backward selection method. The strong association between unstable stacking fault energy and ideal shear strength may be explained by the information is carried by the unstable stacking fault energy when applying the shear deformation on the supercell. This specific correlation cannot solely support by this general correlation, but in statistically perspective it gives us some directions when we want to predict ideal shear strength by using intrinsic materials properties.

4 Summary

The present work predicts ideal shear strength of dilute multicomponent Ni-based alloys determined by first-principles calculations. There is a clear trend of ideal shear strength from highest to lowest when alloying elements into dilute Ni alloys, i.e., Co>Mn~Fe>Cr>Al>Ti>Mo>Si>Nb. Larger atomic volume of alloying elements are associated with lower calculated ideal shear strengths. Trends of concentration of alloying elements splits according to atomic volume difference between host element (pure Ni). The CALPHAD method evaluate all the binary parameters ${}^{0}\phi_{ij}$ and the ternary parameters ϕ_{ijk} with input from first-principles calculations. CALPHAD method quantified the binary and ternary interaction parameters of ideal shear strength and supports the conclusion that larger atomic volume is associated with a smaller ideal shear strength. Additionally, ideal shear strength of Ni-based dilute ternary system could be estimated by the linear combination of binary systems. Filter type and wrapper type feature selection methods generally agree that the atomic volume is the most important feature when calculating the ideal shear strength. Some other atomic properties like Debye temperature, van der Waals atomic radius, and unstable stacking fault energy also show a strong connection with ideal shear strength. These intrinsic material properties can be the interesting directions to predict ideal shear strength of the materials in the future.

5 CRediT authorship contribution statement

Shuang Lin: Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. Shun-Li Shang: Conceptualization, Methodology, Supervision, review & editing, Funding acquisition. John D. Shimanek: Conceptualization, review & editing. Yi Wang: Conceptualization, Methodology. Allison M. Beese: Conceptualization, Project administration, Resources, Writing – review & editing, Funding acquisition. Zi-Kui Liu: Conceptualization, Project administration, Resources, Writing – review & editing, Funding acquisition.

6 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.

7 Acknowledgements

The present work was financially supported by the U. S. Department of Energy (DOE) via award no. DE-FE0031553. Computations were performed partially on The Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer, and partially on the resources of XSEDE supported by NSF via grant no. ACI-1548562.

8 Figures

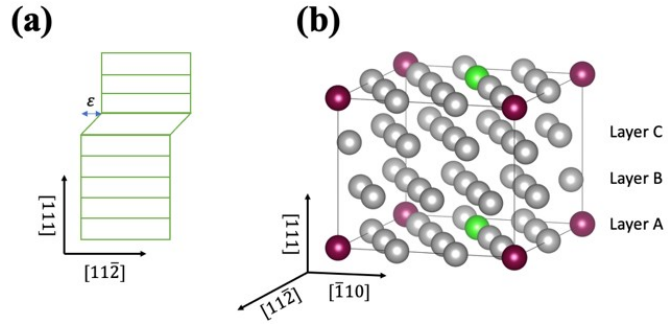


Figure 1. Schematic representation of (a) alias shear deformation, where ε is the magnitude of shear strain (b) Rotated orthorhombic FCC lattice with 36-atom supercell, showing Ni atom (gray) and alloying atoms X (purple) and Z (green). As lattice vectors shown, three {111} layers and 12 atoms on each layer is employed.

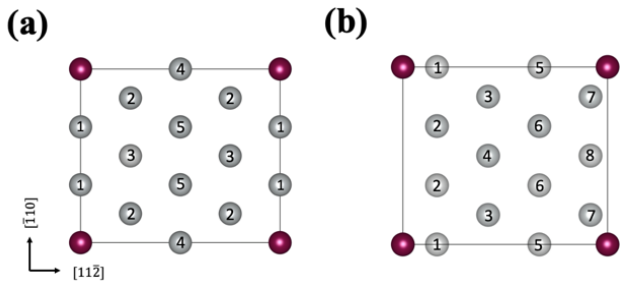


Figure 2. Schematic representation of independent configurations. First alloying element (purple atom) is fixed. The numbers indicate equivalent sites when substituting the second alloying element. (a) Five independent configurations when alloying elements at the same layer. (b) Eight independent configurations when alloying elements at the different layers.

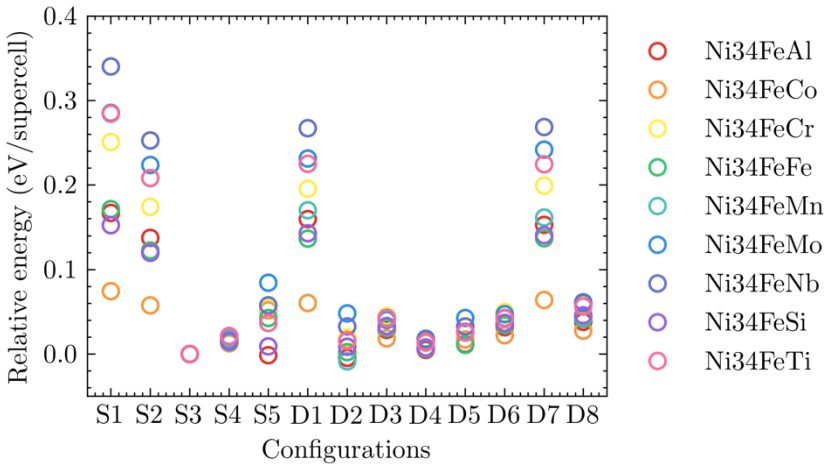


Figure 3. The result of the 13 independent structures' relative energy in terms of lowest energy structure (S3), where S1 to S5 represent different alloying site scenarios when X and Z are at the same layer, i.e. Figure 2 (a), while D1 to D8 represent different alloying site scenarios when X and Z are at the different layer i.e. Figure 2 (b). The plot suggests that S1, S2, D1 and D7 have a relatively higher energy, however structure S3, one of the maximum entropy structures, has the lowest energy.

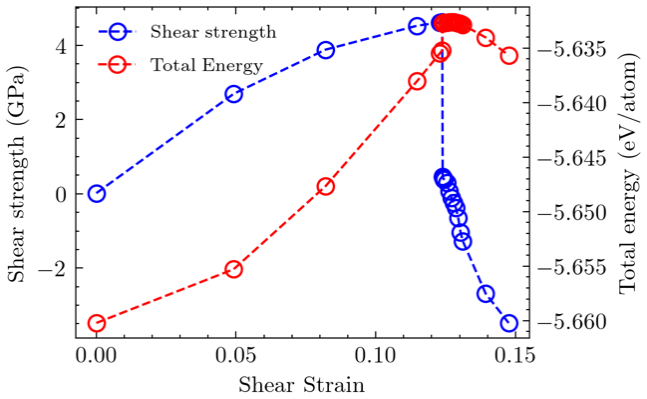


Figure 4. The result of total energy and shear strength with imposing displacement of $Ni_{34}CoNb$ system determined by first-principles calculations. It shows the increase of shear stress with increasing shear strain before reaching the maximum point, i.e., the ideal shear strength. Besides shear stress, the slope of total energy with respect to the displacement continuously increases when approaching the ideal shear strength.

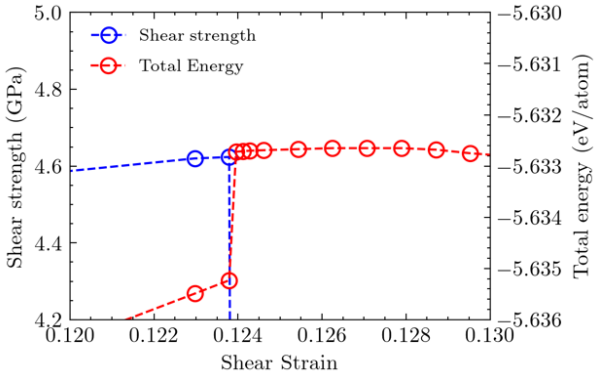


Figure 5. The zoom-in area shows the result of $Ni_{34}CoNb$ when reaching to the ideal shear strength point. The ideal shear strength occurs at the point of largest slope of total energy.

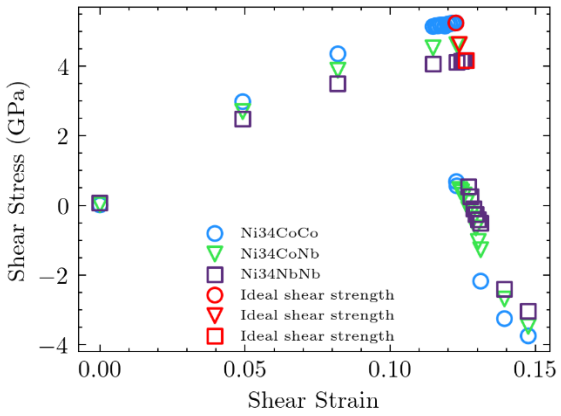


Figure 6. Calculated shear strength vs imposing shear strain for systems of Ni34CoCo (blue circles), Ni34CoNb (green triangles) and Ni34NbNb (purple squares). The red symbols represent the highest value of the shear strength, which are the ideal shear strength among these three systems. The ideal shear strength of Ni34CoNb is roughly equal to the average ideal shear strength of the Ni34CoCo and Ni34NbNb system and a higher ideal shear strength typically occurs at a lower shear strain.

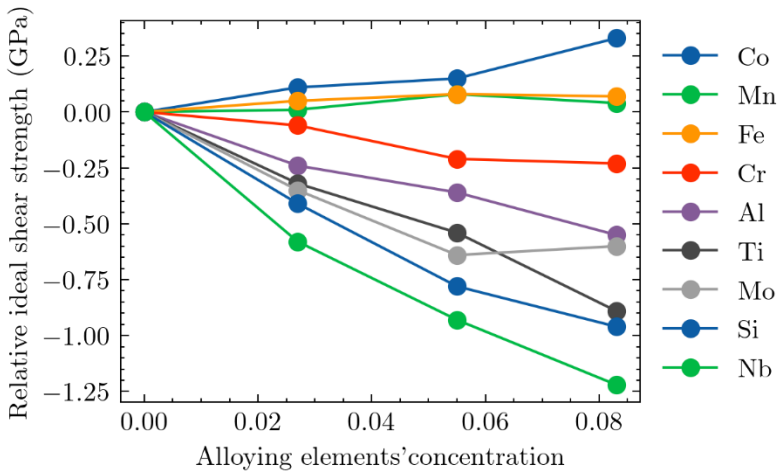


Figure 7. Relative ideal shear strength as function of X alloying elements concentration, where taking the ideal shear strength of pure Ni as reference. Higher concentration of strengthen alloying elements (Co, Mn, Fe) will lead to higher ideal shear strength. However, higher concentration of softening alloying elements (Cr, Al, Ti, Mo, Si, Nb) will lead to lower ideal shear strength of the alloys. The only exception is ideal shear trend of $Ni_{11}Mo$ from alloying element concentration effect only works well at dilute Ni-based alloy, with Mo concentration under around 6%.

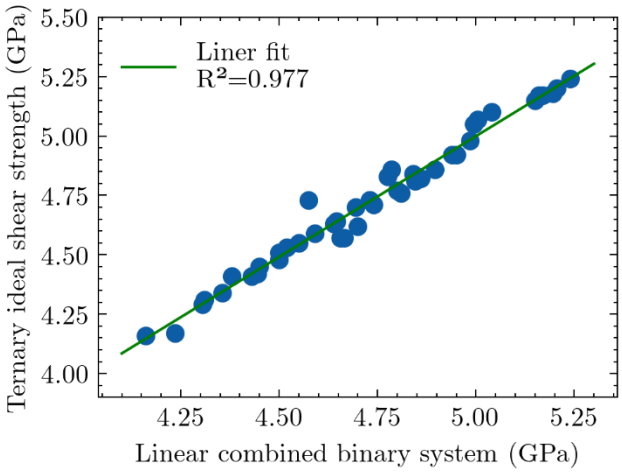


Figure 8. Linear correlation between DFT ternary predicted ideal shear strength and linear combined from binary system, suggesting that in the dilute Ni-based alloys, the ternary interaction does not significantly change ideal shear strength. The ideals shear strength of Ni-based dilute ternary system could be estimated by the linear combined of binary systems.

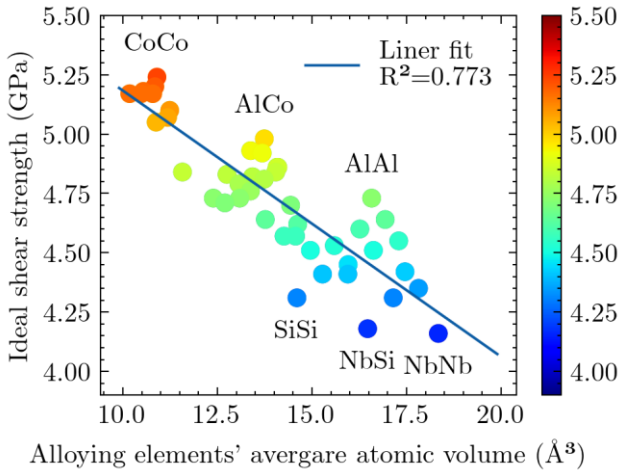


Figure 9. Ideal shear strength as function of the linearly combined volume of alloying elements X and Z. The blue line shows the linear fit result with R^2= 0.773.

9 Tables

Table 1. Comparison results of pure Ni with different supercell size. Typically, larger supercell processes a higher relaxed energy at the ideal shear strength point. This deviation causes a lower ideal shear strength

Supercell size	Relaxed Energy at the ideal shear strength point (eV/atom)	Ideal shear strength (GPa)
12-atom	-5.4409	5.13
36-atom	-5.4406	5.09

Table 2. Ideal shear strength summary of $Ni_{34}XZ$ determined by the first-principles calculations. The ideal shear strength ranging from 4.16 to 5.24 GPa with a clear trend in ideal shear strength from the highest to the lowest when alloying elements into dilute Ni alloys, i.e., Co > Mn~Fe > Cr > Al > Ti > Mo > Si > Nb.

Co	Mn	Fe	Cr	Al	Ti	Mo	Si	Nb	
5.24	5.20	5.18	5.10	4.98	4.87	4.83	4.83	4.62	Co
	5.17	5.17	5.02	4.92	4.84	4.76	4.71	4.57	Mn
		5.17	5.05	4.94	4.82	4.79	4.73	4.57	Fe
			4.88	4.88	4.70	4.64	4.73	4.51	Cr
				4.73	4.64	4.60	4.53	4.42	Al
					4.57	4.52	4.41	4.35	Ti
						4.45	4.41	4.31	Mo
							4.32	4.18	Si
								4.16	Nb

Table 3.Schematic representation of the shear plane view from {111} plane. The ideal shear strength summary of $Ni_{35}X$, $Ni_{34}X_2$ and $Ni_{11}X$[5].

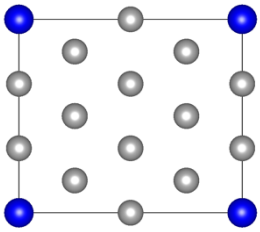
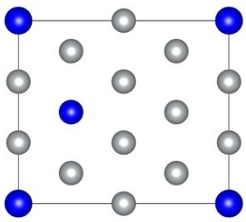
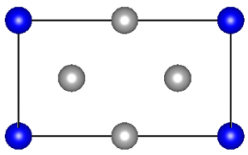
	$Ni_{35}X$ 	$Ni_{34}X_2$ 	$Ni_{11}X$ 
X element concentration	2.7%	5.5%	8.3%
Co	5.20	5.24	5.46
Mn	5.10	5.17	5.17
Fe	5.14	5.17	5.20
Cr	5.03	4.88	4.90
Al	4.85	4.73	4.58
Ti	4.77	4.55	4.24
Mo	4.74	4.45	4.53
Si	4.68	4.31	4.17
Nb	4.51	4.16	3.91

Table 4. Calculated the binary interaction parameters ${}^{0}\phi_{Ni,X}$ using *Eq. 8*.

X	Co	Mn	Fe	Cr	Al	Ti	Mo	Si	Nb
$\phi_{Ni,X}$	3.10	1.80	1.80	-3.44	-6.09	-9.27	-11.15	-13.64	-16.37

Table 5. Calculated the ternary interaction parameters ${}^{0}\phi_{Ni,X}$ using *Eq. 9*.

Co	Mn	Fe	Cr	Al	Ti	Mo	Si	Nb	
0.00	-5.81	-31.90	46.32	-13.31	-55.09	-35.78	72.16	-100.84	Co
	0.00	-4.23	64.80	-34.43	-35.19	-61.98	-40.01	-119.18	Mn
		0.00	28.43	-27.64	-64.82	-30.72	-12.42	-117.27	Fe
			0.00	73.35	-23.77	-32.64	175.12	-7.64	Cr
				0.00	-8.78	10.32	10.49	-37.77	Al
					0.00	22.29	-32.75	-6.58	Ti
						0.00	42.64	2.90	Mo
							0.00	-75.27	Si
								0.00	Nb

Table 6. Correlation features and corresponding explanations

	Feature	Explanation
Atomic	Radius_Coval	Covalent radius (pm) based on the collections of Wolfram Mathematica; see "ElementData" in Ref. [32].
	Radius_vDW	Van der Waals atomic radius (pm) [32,33].
	V0_Miedema	Atomic volume (cm^3/mol) used in the Miedema model [34].
	Mass	Mass of pure elements
.Periodic	Group	Group of pure elements in the periodic table.
	M_Num2	Mendeleev number MN2, starting bottom left and moving up then to the right. [35]
	Number	Atomic number of pure elements in the periodic table
	Period	Period of pure elements in the periodic table.
Elastic	B	Bulk modulus (GPa) of pure elements based on [36,37]. Note that elastic properties of fcc Sr were taken from [38].
	G	Shear modulus (GPa) of pure elements based on [36,37].
	Y	Young's modulus (GPa) of pure elements based on [36,37].
.Thermodynamic	BoilingT	Boiling temperature (K) [39].
	Ele_Conduc	Electrical conductivity of metals in $(ohm\text{-}cm)^{-1}$ [40].
	Heat_Capacity	Heat capacity at 298 K (J/kg-mol·K) [40].
	Heat_Fusion	Heat of fusion at 298 K (J/mol) [40].
	Heat_Sublimation	Heat of sublimation (J/mol) at 298 K [40].

	MeltingT	Melting temperature (K) based on the collections by Kittel [40].
	S298	Standard entropy (J/mol.K) at 298 K [41].
	Therm_Conduc	Thermal conductivity at 300 K (W cm^{-1} K^{-1}) [32,40].
	VaporHeat	Vaporization heat (kJ/mol) based on the collections of Wolfram Mathematica; see "VaporizationHeat" in Ref. [32].
Lattice	CohEnergy	Cohesive energy (eV/atom) collected by Kittel [40].
	DebyeT	Debye temperature (K) collected by Kittel [40].
	Va_Acti_FCC	Predicted vacancy activity energy of pure elements in the fcc structure, with the vacancy formation energy adopted for those with unstable fcc structures (i.e., Ge and La) [42].
	Va_Form_FCC	Predicted vacancy formation energy of pure elements in fcc structure [42].
Electronic	Electron_Affinity	Electron affinity (eV) [33].
	EleDensity_Miedema	Electron density at the boundary of Wigner-Seitz cell used in the Miedema model [34].
	EleNeg_Miedema	Electronegativity (Volt) used in the Miedema model [34].
	EleNeg_Pauling	Electronegativity (dimensionless) on the Pauling scale [32,33][13,16].
	Ion_Pot_1	The first ionization potential (eV) [39].
	Ion_Pot_2	The second ionization potential (eV) [39].
	Ion_Pot_3	The third ionization potential (eV) [39].

	MaxR_Ele_in_Solid	Maximum range (mm) of electrons in solid elements for electron energy of 15 keV [39].
	No_Spectral_lines	Number of spectral lines of the elements [39].
	PPot_radius_s	Nonlocal pseudopotential radius (a.u.) for the s orbital [43].
	PPot_radius_p	Nonlocal pseudopotential radius (a.u.) for the p orbital [43].
	NsVal	Number of filled s-shell valence electron states.
	NpVal	Number of filled p-shell valence electron states.
	NdVal	Number of filled d-shell valence electron states.
	NfVal (removed when correlation)	Number of filled f-shell valence electron states.
	Nval	Number of filled valence electron states.
	NsUnfill	Number of unfilled s-shell valence electron states.
	NpUnfill	Number of unfilled p-shell valence electron states.
	NdUnfill	Number of unfilled d-shell valence electron states.
	NfUnfill (removed when correlation)	Number of unfilled f-shell valence electron states, not applicable here.
	Nunfill	Number of unfilled valence electron states.
DFT	DFTC11	Elastic constant C11
DFT	DFTC12	Elastic constant C12

	DFTC44	Elastic constant C44
	DFTC13	Elastic constant C13
	DFTC33	Elastic constant C33
	DFTv0	Equilibrium volume
	DFTb	Burgers vector
	DFTpoisson	Poisson ratio
	USFE	Unstable stacking fault energy [12]
	WorkFunc	Electronic work function, energy (or work) required to withdraw an electron completely from a metal surface.
	DFTBh	Voigt-Reuss-Hill approach shear modulus [44].
	DFTGh	Voigt-Reuss-Hill approach calculated shear modulus [44].
Features combination	DFTBvG	The ratio between bulk modulus and shear modulus [45].
	DFTGvb	The ratio between bulk modulus and Burgers vector
	DFTGV	Shear modulus G multiply Volume
	Gb2pih	Ideal shear strength empirical model by Frenkel (Gb/2πh)[24] .
	BGsq	Squart (B/G)
	Cauphy	Cauphy pressure which is difference between elastic constant C12 and C44, i.e., C12-C44 [25].

Table 7. Result of filter-type and wrapper type features ranking.

Index	Ftest		rRelief		R^2		MIC		forward		backward	
1	DFTv0	25.31	DFTv0	0.07	DFTv0	0.77	DFTv0	0.77	DFTv0	964	DebyeT	969
2	DFTGvb	16.31	Radius_vDW	0.07	V0_Miedema	0.69	CohEnergy	0.67	DebyeT	955	Radius_vDW	958
3	V0_Miedema	16.22	V0_Miedema	0.07	DFTGvb	0.64	Heat_Sublimation	0.67	Va_Activa_fcc	728	USFE	846
4	DFTb	16.07	BoilingT	0.06	CohEnergy	0.54	DFTb	0.62	Nval	573	DFTv0	721
5	Heat_Fusion	14.78	Heat_Fusion	0.05	DFTb	0.54	DFTGvb	0.60	Radius_vDW	513	NdVal	695
6	CohEnergy	12.82	Mass	0.05	VaporHeat	0.52	V0_Miedema	0.58	USFE	436	EleDensity_Miedema	630
7	Mass	11.66	G_wiki	0.05	DFTC44	0.52	DFTpoisson	0.55	V0_Miedema	356	V0_Miedema	612
8	Number	11.31	DFTC12	0.04	Heat_Sublimation	0.51	VaporHeat	0.53	DFTb	335	BoilingT	580
9	Heat_Sublimation	10.88	Y_wiki	0.04	DFTpoisson	0.50	Heat_Fusion	0.51	VaporHeat	326	Ion_Pot_2	561
10	No_Spectral_lines	10.87	Number	0.04	Nval	0.50	Mass	0.51	Heat_Sublimation	325	PPot_radius_p	553
11	Nval	9.26	Heat_Sublimation	0.04	BGsq	0.48	DFTC44	0.49	G_wiki	296	DFTC33	551
12	Radius_vDW	8.15	Period	0.04	DFTBvG	0.46	Radius_vDW	0.48	Group	286	Nval	545
13	DebyeT	7.91	CohEnergy	0.04	DebyeT	0.36	Number	0.48	M_Num2	282	NUnfill	544
14	M_Num2	7.72	Ion_Pot_1	0.04	Electron_Affinity	0.34	BGsq	0.47	MeltingT	279	WorkFunc	541
15	VaporHeat	7.64	VaporHeat	0.03	DFTGh	0.34	Nval	0.46	PPot_radius_p	278	Y_wiki	529
16	DFTpoisson	7.44	DFTC13	0.03	NUnfill	0.34	DFTC12	0.45	Electron_Affinity	258	Ion_Pot_3	515
17	DFTC44	7.44	DebyeT	0.03	BoilingT	0.33	DFTBvG	0.45	NUnfill	179	VaporHeat	501
18	BoilingT	7.25	Ion_Pot_3	0.03	EleNeg_Miedema	0.33	DebyeT	0.44	PPot_radius_s	174	DFTC11	485
19	DFTGh	7.12	MeltingT	0.03	Radius_vDW	0.32	BoilingT	0.44	WorkFunc	153	Therm_Conduc	484
20	DFTC12	7.06	MaxR_Ele_in_Solid	0.03	NdVal	0.31	No_Spectral_lines	0.43	EleDensity_Miedema	144	DFTBh	482
21	Y_wiki	6.77	NsVal	0.03	No_Spectral_lines	0.30	Ion_Pot_1	0.43	EleNeg_Miedema	141	Group	472
22	DFTBvG	6.55	NsUnfill	0.03	Ion_Pot_3	0.28	NUnfill	0.42	NpVal	126	DFTC12	470
23	BGsq	6.55	DFTGvb	0.02	MeltingT	0.24	PPot_radius_s	0.41	Y_wiki	123	NpVal	467
24	NdVal	6.16	USFE	0.02	NsVal	0.20	PPot_radius_p	0.41	DFTpoisson	111	M_Num2	451

25	Period	5.99	PPot_radius_s	0.02	NsUnfill	0.20	M_Num2	0.40	Ele_Conduc	97	B_wiki	423
26	G_wiki	5.72	Radius_Coval	0.02	Ion_Pot_2	0.20	DFTC11	0.39	DFTC33	86	PPot_radius_s	374
27	Gb2pih	5.40	Ion_Pot_2	0.02	Heat_Capacity	0.18	DFTC33	0.39	Cauchy	85	Ion_Pot_1	362
28	Electron_Affinity	4.98	DFTBvG	0.02	Y_wiki	0.16	Radius_Coval	0.38	NpUnfill	82	MaxR_Ele_in_Solid	361
29	NsVal	4.78	NpVal	0.02	DFTBh	0.15	Y_wiki	0.37	Va_Form_fcc	77	NpUnfill	350
30	NsUnfill	4.78	Va_Activa_fcc	0.02	Radius_Coval	0.15	NdVal	0.37	Radius_Coval	73	Va_Form_fcc	343
31	MeltingT	4.67	NdVal	0.02	DFTC12	0.15	DFTGh	0.37	BoilingT	72	EleNeg_Miedema	313
32	USFE	4.55	S298	0.02	Gb2pih	0.14	USFE	0.37	DFTGvb	72	No_Spectral_lines	290
33	Radius_Coval	4.54	PPot_radius_p	0.01	DFTC33	0.14	MaxR_Ele_in_Solid	0.35	NdUnfill	70	CohEnergy	285
34	Heat_Capacity	4.30	BGsq	0.01	DFTC13	0.13	G_wiki	0.35	Ion_Pot_1	68	G_wiki	281
35	Group	4.18	Ele_Conduc	0.01	G_wiki	0.13	EleNeg_Miedema	0.34	NsUnfill	68	Heat_Sublimation	273
36	NUnfill	4.13	Nval	0.01	EleDensity_Miedema	0.13	Ion_Pot_3	0.34	Ion_Pot_3	64	NdUnfill	269
37	EleNeg_Miedema	3.86	NpUnfill	0.01	DFTC11	0.13	Electron_Affinity	0.33	NsVal	64	Mass	266
38	Ion_Pot_3	3.68	Heat_Capacity	0.01	DFTGV	0.13	DFTGV	0.33	CohEnergy	63	NsUnfill	261
39	PPot_radius_s	3.54	DFTBh	0.01	Ele_Conduc	0.12	MeltingT	0.32	NdVal	61	NsVal	260
40	DFTGV	3.19	DFTb	0.01	NpVal	0.12	Va_Form_fcc	0.32	DFTC44	60	Number	230
41	PPot_radius_p	2.96	M_Num2	0.01	Cauchy	0.10	DFTBh	0.31	EleNeg_Pauling	59	Cauchy	209
42	DFTC33	2.68	NdUnfill	0.00	Heat_Fusion	0.09	NsVal	0.30	MaxR_Ele_in_Solid	58	Radius_Coval	202
43	EleDensity_Miedema	2.62	Va_Form_fcc	0.00	MaxR_Ele_in_Solid	0.09	NsUnfill	0.30	BGsq	37	Heat_Fusion	185
44	Ion_Pot_1	2.27	Group	0.00	Mass	0.08	Period	0.30	Ion_Pot_2	35	Period	175
45	DFTC11	2.25	B_wiki	0.00	Number	0.07	Ion_Pot_2	0.30	Therm_Conduc	33	Ele_Conduc	159
46	NpVal	1.94	NUnfill	0.00	Ion_Pot_1	0.07	DFTC13	0.28	B_wiki	29	DFTC13	159
47	Ion_Pot_2	1.91	EleNeg_Miedema	0.00	PPot_radius_s	0.07	Gb2pih	0.27	DFTC11	27	Electron_Affinity	118
48	Ele_Conduc	1.82	EleNeg_Pauling	0.00	NpUnfill	0.06	NpVal	0.27	Gb2pih	18	DFTpoisson	112

49	MaxR_Ele_in_Solid	1.74	Therm_Conduc	0.00	Period	0.05	NpUnfill	0.27	Period	17	MeltingT	106
50	Cauchy	1.72	DFTC44	-0.01	USFE	0.04	NdUnfill	0.27	DFTBh	17	DFTGh	98
51	NpUnfill	1.71	DFTpoisson	-0.01	PPot_radius_p	0.04	Group	0.27	Mass	11	Gb2pih	84
52	WorkFunc	1.67	EleDensity_Miedema	-0.01	Va_Form_fcc	0.03	Heat_Capacity	0.26	DFTBvG	11	DFTGvb	75
53	NdUnfill	1.39	DFTC11	-0.01	Therm_Conduc	0.02	Cauchy	0.25	Heat_Fusion	10	BGsq	61
54	DFTC13	1.37	Electron_Affinity	-0.02	Va_Activa_fcc	0.01	EleDensity_Miedema	0.25	DFTGV	9	DFTC44	57
55	DFTBh	1.18	DFTGV	-0.02	WorkFunc	0.01	S298	0.24	Number	7	DFTb	38
56	Va_Form_fcc	0.96	DFTGh	-0.02	NdUnfill	0.01	Ele_Conduc	0.21	No_Spectral_lines	7	DFTBvG	38
57	S298	0.70	DFTC33	-0.02	EleNeg_Pauling	0.00	Therm_Conduc	0.21	DFTC12	4	DFTGV	29
58	B_wiki	0.61	Cauchy	-0.04	Group	0.00	Va_Activa_fcc	0.20	Heat_Capacity	3	Va_Activa_fcc	26
59	Therm_Conduc	0.28	WorkFunc	-0.04	M_Num2	0.00	B_wiki	0.19	DFTGh	3	EleNeg_Pauling	17
60	EleNeg_Pauling	0.15	Gb2pih	-0.04	S298	0.00	EleNeg_Pauling	0.19	S298	2	Heat_Capacity	3
61	Va_Activa_fcc	0.05	No_Spectral_lines	-0.05	B_wiki	0.00	WorkFunc	0.19	DFTC13	1	S298	3

10 References

[1] S. Ogata, J. Li, S. Yip, Ideal Pure Shear Strength of Aluminum and Copper, Science (1979). 298 (2002) 807–811. https://doi.org/10.1126/science.1076652.

[2] S. Ogata, J. Li, N. Hirosaki, Y. Shibutani, S. Yip, Ideal shear strain of metals and ceramics, Phys Rev B Condens Matter Mater Phys. 70 (2004) 1–7. https://doi.org/10.1103/PhysRevB.70.104104.

[3] J.D. Shimanek, S. Qin, S.L. Shang, Z.K. Liu, A.M. Beese, Predictive Crystal Plasticity Modeling of Single Crystal Nickel Based on First-Principles Calculations, JOM. (2022). https://doi.org/10.1007/s11837-022-05175-6.

[4] F.R.N. Nabarro, Fifty-year study of the Peierls-Nabarro stress, Materials Science and Engineering A. 234–236 (1997) 67–76. https://doi.org/10.1016/s0921-5093(97)00184-6.

[5] J.D. Shimanek, S.-L. Shang, A.M. Beese, Z.-K. Liu, Insight into ideal shear strength of Ni-based dilute alloys using first-principles calculations and correlational analysis, Comput Mater Sci. 212 (2022) 111564. https://doi.org/10.1016/j.commatsci.2022.111564.

[6] L. Kaufman, H. Bernstein, Computer calculation of phase diagrams. With special reference to refractory metals, (1970). https://inis.iaea.org/search/search.aspx?orig_q=RN:2004171 (accessed August 4, 2022).

[7] Z.K. Liu, First-principles calculations and CALPHAD modeling of thermodynamics, J Phase Equilibria Diffus. 30 (2009) 517–534. https://doi.org/10.1007/s11669-009-9570-6.

[8] S.L. Shang, J. Shimanek, S. Qin, Y. Wang, A.M. Beese, Z.K. Liu, Unveiling dislocation characteristics in N i3Al from stacking fault energy and ideal strength: A first-principles study via pure alias shear deformation, Phys Rev B. 101 (2020). https://doi.org/10.1103/PhysRevB.101.024102.
[9] M. Jahnátek, J. Hafner, M. Krajčí, Shear deformation, ideal strength, and stacking fault formation of fcc metals: A density-functional study of Al and Cu, Phys Rev B Condens Matter Mater Phys. 79 (2009). https://doi.org/10.1103/PhysRevB.79.224103.
[10] V. Shankar, K. Bhanu Sankara Rao, S.L. Mannan, Microstructure and mechanical properties of Inconel 625 superalloy, Journal of Nuclear Materials. 288 (2001) 222–232. https://doi.org/10.1016/S0022-3115(00)00723-6.
[11] M. Rahman, W.K.H. Seah, T.T. Teo, The Machinability of Incone1 718 Materials Processing Technology, 1997.
[12] S.L. Shang, W.Y. Wang, B.C. Zhou, Y. Wang, K.A. Darling, L.J. Kecskes, S.N. Mathaudhu, Z.K. Liu, Generalized stacking fault energy, ideal strength and twinnability of dilute Mg-based alloys: A first-principles study of shear deformation, Acta Mater. 67 (2014) 168–180. https://doi.org/10.1016/j.actamat.2013.12.019.
[13] T. Bučko, J. Hafner, J.G. Ángyán, Geometry optimization of periodic systems using internal coordinates, Journal of Chemical Physics. 122 (2005). https://doi.org/10.1063/1.1864932.
[14] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys Rev B. 54 (1996) 11169–11186. https://doi.org/10.1103/PhysRevB.54.11169.
[15] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys Rev B. 59 (1999) 1758–1775. https://doi.org/10.1103/PhysRevB.59.1758.
[16] J.P. Perdew, Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Phys Rev B. 45 (1992) 13244–13249. https://doi.org/10.1103/PhysRevB.45.13244.
[17] M. Methfessel, A.T. Paxton, High-precision sampling for Brillouin-zone integration in metals, Phys Rev B. 40 (1989) 3616–3621. https://doi.org/10.1103/PhysRevB.40.3616.
[18] A. van de Walle, Multicomponent multisublattice alloys, nonconfigurational entropy and other additions to the Alloy Theoretic Automated Toolkit, CALPHAD. 33 (2009) 266–278. https://doi.org/10.1016/j.calphad.2008.12.005.
[19] S. Shang, Y. Wang, W.Y. Wang, H. Fang, Z.K. Liu, Low energy structures of lithium-ion battery materials Li(Mn xNixCo1-2x)O2 revealed by first-principles calculations, Appl Phys Lett. 103 (2013). https://doi.org/10.1063/1.4817763.
[20] Z.K. Liu, H. Zhang, S. Ganeshan, Y. Wang, S.N. Mathaudhu, Computational modeling of effects of alloying elements on elastic coefficients, Scr Mater. 63 (2010) 686–691. https://doi.org/10.1016/j.scriptamat.2010.03.049.
[21] B. Hallstedt, N. Dupin, M. Hillert, L. Höglund, H.L. Lukas, J.C. Schuster, N. Solak, Thermodynamic models for crystalline phases. Composition dependent models for volume, bulk modulus and thermal expansion, Calphad. 31 (2007) 28–37. https://doi.org/10.1016/j.calphad.2006.02.008.
[22] J. Andersson, J. Ågren, Models for numerical treatment of multicomponent diffusion in simple phases, J Appl Phys. 72 (1992) 1350–1355. https://doi.org/10.1063/1.351745.
[23] S. Shang, Y. Wang, Y. Du, M.A. Tschopp, Z.K. Liu, Integrating computational modeling and first-principles calculations to predict stacking fault energy of dilute multicomponent

Ni-base alloys, Comput Mater Sci. 91 (2014) 50–55. https://doi.org/10.1016/j.commatsci.2014.04.040.
[24] J. Frenkel, Zur Theorie der Elastizit€tsgrenze und der Festigkeit kristallinischer K€rper, Zeitschrift F€r Physik. 37 (1926) 572–609. https://doi.org/10.1007/BF01397292.
[25] O.N. Senkov, D.B. Miracle, Generalization of intrinsic ductile-to-brittle criteria by Pugh and Pettifor for materials with a cubic crystal structure, Sci Rep. 11 (2021) 10–13. https://doi.org/10.1038/s41598-021-83953-z.
[26] A.E. Isabelle Guyon, An Introduction to Variable and Feature Selection, Procedia Comput Sci. 94 (2016) 465–472.
[27] M. ROBNIK SIKONJA MarkoRobnik, friuni-ljsi IGOR KONONENKO IgorKononenko, Theoretical and Empirical Analysis of ReliefF and RReliefF, Mach Learn. 53 (2003) 23–69.
[28] The Mathworks Inc., MATLAB - MathWorks, Www.Mathworks.Com/Products/Matlab. (2020).
[29] D. Lorenz, A. Zeckzer, U. Hilpert, P. Grau, H. Johansen, H.S. Leipner, Pop-in effect as homogeneous nucleation of dislocations during nanoindentation, Phys Rev B Condens Matter Mater Phys. 67 (2003). https://doi.org/10.1103/PhysRevB.67.172101.
[30] M.B. Lowry, D. Kiener, M.M. Leblanc, C. Chisholm, J.N. Florando, J.W. Morris, A.M. Minor, Achieving the ideal strength in annealed molybdenum nanopillars, Acta Mater. 58 (2010) 5160–5167. https://doi.org/10.1016/j.actamat.2010.05.052.
[31] D. Lorenz, A. Zeckzer, U. Hilpert, P. Grau, H. Johansen, H.S. Leipner, Pop-in effect as homogeneous nucleation of dislocations during nanoindentation, Phys Rev B. 67 (2003) 172101. https://doi.org/10.1103/PhysRevB.67.172101.
[32] Wolfram version 11.3, Wolfram Mathematica: Modern Technical Computing, Wolfram - 11.3. (2019) version 11.3. https://www.wolfram.com/mathematica/.
[33] PubChem Database, Periodic Table of Elements - PubChem, (2019).
[34] R.F. Zhang, S.H. Zhang, Z.J. He, J. Jing, S.H. Sheng, Miedema Calculator: A thermodynamic platform for predicting formation enthalpies of alloys within framework of Miedema's Theory, Comput Phys Commun. 209 (2016) 58–69. https://doi.org/10.1016/j.cpc.2016.08.013.
[35] P. Villars, K. Brandenburg, M. Berndt, S. LeClair, A. Jackson, Y.H. Pao, B. Igelnik, M. Oxley, B. Bakshi, P. Chen, S. Iwata, Binary, ternary and quaternary compound former/nonformer prediction via Mendeleev number, J Alloys Compd. 317–318 (2001) 26–38. https://doi.org/10.1016/S0925-8388(00)01410-9.
[36] M. Winter, WebElements, The University of Sheffield. (2011).
[37] Ansys/Granta Materials Database, Azom materials public datapages, (2020). http://www.azom.com/; https://www.grantadesign.com/industry/products/data/.
[38] M.S. Anderson, C.A. Swenson, D.T. Peterson, Experimental equations of state for calcium, strontium, and barium metals to 20 kbar from 4 to 295 K, Phys Rev B. 41 (1990) 3329–3338. https://doi.org/10.1103/PhysRevB.41.3329.
[39] G. V. Samsonov, Handbook of the Physicochemical Properties of the Elements, Springer, New York, 1968. https://doi.org/10.1007/978-1-4684-6066-7_1.
[40] C. Kittel, Introduction to Solid State Physics, John Wiley & Sons, Inc., Hoboken, NJ, 2005.
[41] A.T. Dinsdale, SGTE data for pure elements, Calphad. 15 (1991) 317–425. https://doi.org/10.1016/0364-5916(91)90030-N.

[42] S.L. Shang, B.C. Zhou, W.Y. Wang, A.J. Ross, X.L. Liu, Y.J. Hu, H.Z. Fang, Y. Wang, Z.K. Liu, A comprehensive first-principles study of pure elements: Vacancy formation and migration energies and self-diffusion coefficients, Acta Mater. 109 (2016) 128–141. https://doi.org/10.1016/j.actamat.2016.02.031.
[43] A. Zunger, Systematization of the stable crystal structure of all AB-type binary compounds: A pseudopotential orbital-radii approach, Phys Rev B. 22 (1980) 5839–5872. https://doi.org/10.1103/PhysRevB.22.5839.
[44] G. Simmons and H. Wang, Single crystal elastic constants and calculated aggregate properties, Mass., M.I.t. Press. (1971).
[45] S.F. Pugh, XCII. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 45 (1954) 823–843. https://doi.org/10.1080/14786440808520496.

Machine learning-enabled identification of micromechanical stress and strain hotspots predicted via dislocation density-based crystal plasticity simulations

Adnan Eghtesad[a], Qixiang Luo [a], Shun-Li Shang [a], Ricardo Lebensohn[b], Marko Knezevic[c], Zi-Kui Liu [a] and Allison M. Beese [a,d,*]

[a] Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA 16802, United States
[b] Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87544, USA
[c] Department of Mechanical Engineering, University of New Hampshire, Durham, NH 03824, USA
[d] Department of Mechanical Engineering, Pennsylvania State University, University Park, PA 16802, United States

Abstract

The present work uses a full-field crystal plasticity model with a first principles-informed dislocation density (DD) hardening law to identify the key microstructural features correlated with micromechanical fields localization, or hotspots, in polycrystalline Ni. An ensemble learning approach to machine learning interpreted with Shapley additive explanation was implemented to predict nonlinear correlations between microstructural features and micromechanical stress and strain hotspots. Results reveal that regions within the microstructure in the vicinity of the grain boundaries, higher Taylor and Schmid factors, and high intergranular misorientations, are more prone to being micromechanical hotspots. Additionally, under combined loading, intergranular misorientations are more responsible than Schmid factor in formation of stress hotspots while Schmid factors take precedence under high plastic strain localizations. The present work demonstrates a successful integration of physics-based crystal plasticity with DD-based hardening into machine learning models to reveal the microscale features responsible for the formation of local

*Corresponding author at: Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA 16802, United States. E-mail address: beese@matse.psu.edu (A.M. Beese).

stress and strain hotspots within the grains and near the grain boundaries, as function of applied deformation states, grain morphology/size distribution, and microstructural texture, providing insights into micromechanical damage initiation zones in polycrystalline metals.

Keywords: Machine learning; ensemble learning; microstructure; hotspots; crystal plasticity

1 Introduction

The formation of micromechanical hotspots during loading triggers damage initiation and accumulation, e.g., void nucleation, grain boundary sliding, shear band formation (Eghtesad and Knezevic, 2021a; Orme et al., 2016; Qin and Beese, 2021). Therefore, to aid in the understanding of microstructural origins of failure, as well as to enable microstructural design for superior properties, it is critical to understand the microstructural characteristics that contribute to the formation of micromechanical hotspots, and if and how these change with microstructure and applied loading state. A quantitative description of plastic deformation and subsequent damage incubation in crystalline materials requires knowledge of how crystallographic texture, grain morphology, and grain boundary (GB) character of polycrystalline microstructures alter the localization of micromechanical fields inside the grains and near the GBs.

Crystal plasticity (CP) models facilitate the modeling of microstructure-sensitive elasto-plastic deformation based on the mechanical response of crystalline grains (Roters et al., 2010). Among CP models that capture explicit grain-grain interactions and spatial gradients, the crystal plasticity fast Fourier transform (CPFFT) (Lebensohn et al., 2012a) and the crystal plasticity finite element method (CPFEM) (Roters et al., 2011) are predominant in the literature. Of these two, CPFFT is the more computationally efficient formulation (Eghtesad et al., 2018; Eghtesad and Knezevic, 2020a; Lebensohn and Rollett, 2020), especially with recent developments enabling high-performance computing (HPC) and graphics processing unit (GPU) hardware acceleration.

While HPC has improved the efficiency of CP modeling, simulations of very large datasets are still time-consuming. To address this, recent research has enabled the integration of machine learning (ML) into

microstructure-property linkage applications (Batra, 2021; Gao et al., 2022; Pilania, 2021; Rodrigues et al., 2021; Veasna et al., 2023). Minaroodi et al. proposed a deep neural network (DNN) surrogate model for the quantification of stress fields in anisotropic microstructures (Mianroodi et al., 2021). Cecen et al. implemented a data driven approach using a convolutional neural network (CNN) for efficient microstructure-property linkage (Cecen et al., 2018). Yang et al. proposed an artificial intelligence (AI) based approach within a conditional generative adversarial neural network (cGAN) to find correlations between a composite microstructure and its micromechanical response (Yang et al., 2021). Pandey et al. reported an ML based surrogate method for predictions of texture evolution under uniaxial tension within crystal plasticity (Pandey and Pokharel, 2021).

Several studies have explored the formation of stress hotspots within polycrystalline microstructures. Rollett et al. investigated stress hotspots under uniaxial tension in Cu using the CPFFT model with Voce hardening (Rollett et al., 2010a). Chief findings of a study of strain localization under rolling conditions were that strain concentrations occur at triple junctions or quadruple points and then interconnect with further straining to create shear bands that extend across the polycrystalline structure(Ardeljan et al., 2015). Particularly, the triggering strain hotspots occurred at junctions of grains with dissimilar reorientation propensities, while cold spots were formed vice-versa, i.e., at junctions of grains with similar reorientation trends. Donegan et al. used convolutional neural networks based on microstructural images to predict stress localization during the thermoelastic response of particulate microstructures (Donegan et al., 2019). Fatigue stress hotspots in polycrystalline Cu were explored by a combination of high resolution EBSD (HR-EBSD) and CPFEM by Wan et al (Wan et al., 2016). In a recent study (Mangal and Holm, 2018a), the formation of stress hotspots in polycrystalline Cu was investigated by integrating ML techniques and the CPFFT model with a phenomenological Voce hardening law

(Lebensohn et al., 2012a). The methodology was based on grain-wise averaging of the stress fields and intergranular misorientations of neighboring grains. The most relevant microstructural features related to the stress hotspots formed under uniaxial loading were identified using a Least Absolute Shrinkage and Selection Operator (LASSO) linear regression criterion (Ranstam and Cook, 2018).

A grain-wise averaging method reduces the complexity of micromechanical variations and local gradients in the vicinity of grain boundaries. In most practical applications, materials are subjected to complex multiaxial loading conditions, which affect the intragranular fields in ways that are lost during homogenization. Additionally, it has been shown that the local distribution of stresses and strains within a microstructure is heavily dependent on the hardening law (Patil et al., 2021). Phenomenological models such as Voce, in contrast to the physics-based dislocation density hardening law used here, underestimate the heterogeneity of spatial distributions by introducing spurious grain-wise homogenizations.

To identify the dominant microstructural features responsible for stress and strain hotspots, the present work adopts a CPFFT model with a physics-based dislocation density (DD) hardening model informed by density functional theory (DFT) (Eghtesad et al., 2022). A set of microstructures varying in crystallographic texture and grain morphology was generated, using the software DREAM3D (Groeber and Jackson, 2014a), and deformed under a range of applied deformation conditions. Machine learning techniques were then applied to identify the microstructural features most strongly associated with local stress and strain hotspots in pure polycrystalline Ni.

2 Methods

Fig. 1 illustrates the ML-based identification of microstructural features correlated to hotspots. The CPFFT model allows for the quantification of micromechanical fields as a function of applied deformation. Microstructural RVEs used in this study were generated using DREAM3D software (Groeber and Jackson, 2014a) with high-resolution (128 voxels in each direction) in order to accurately capture the spatial gradients within grains and near grain boundaries. Hotspots were defined as locations in which the value for the field of interest exceeded 95% of the mean value of the field. The microstructural features and resulting fields vary as function of crystallographic texture, grain morphology, grain boundary (GB) features, and applied deformation.

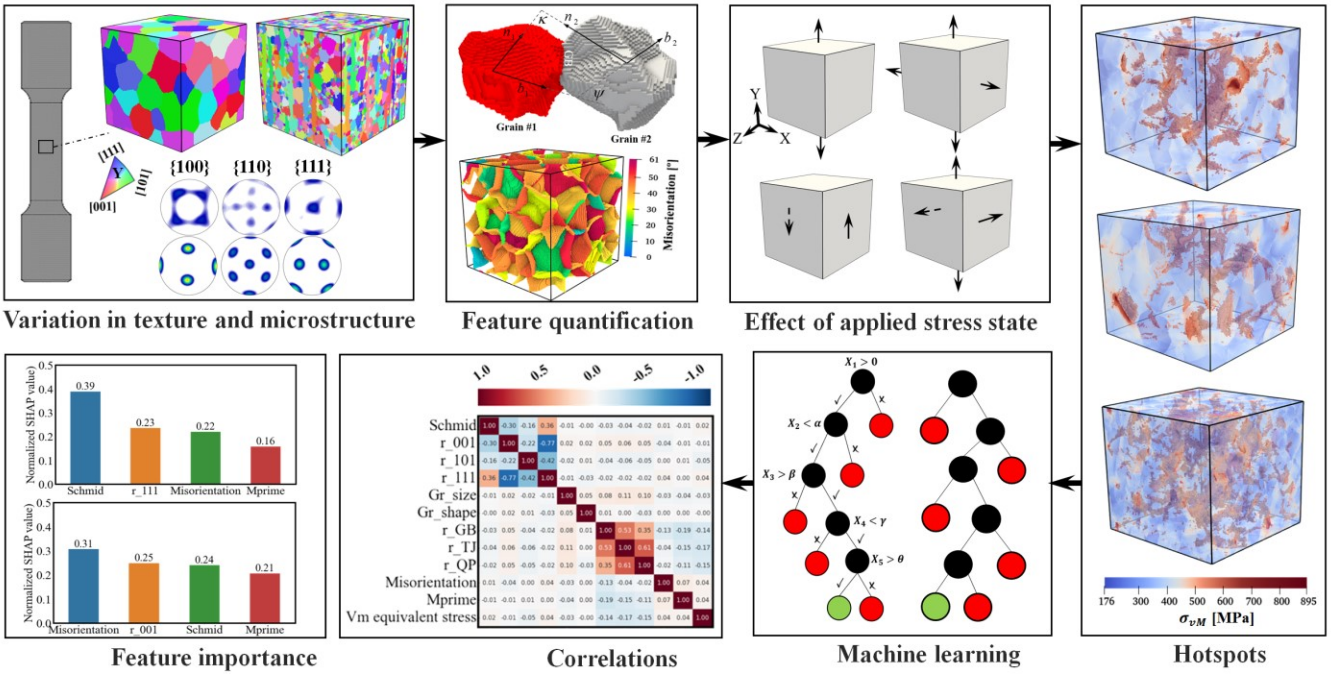


Fig. 1. Schematic of method for machine learning identification of micromechanical features responsible for hotspot formation used in the present study.

2.1 CPFFT and DFT-based dislocation density hardening law

This study used a DFT-informed DD hardening law within the CPFFT framework, which mitigates the uncertainties involved in calibration to a macroscopic flow response obtained from experimental measurements (Eghtesad et al., 2022)[†]. The CPFFT model uses a power law constitutive formulation that defines the plastic strain rate, $\dot{\boldsymbol{\varepsilon}}^{\mathbf{p}}(\mathbf{x})$, as function of Cauchy stress, $\boldsymbol{\sigma}(\mathbf{x})$, through a superposition of shear rates on *N* slip systems (Asaro, 1983):

$$\dot{\boldsymbol{\varepsilon}}^{\mathbf{p}}(\mathbf{x}) = \sum_{s=1}^{N} \mathbf{P}^{\mathbf{s}}(\mathbf{x})\ \dot{\gamma}^{s}(\mathbf{x}) = \dot{\gamma}_0 \sum_{s=1}^{N} \mathbf{P}^{\mathbf{s}}(\mathbf{x}) \left(\frac{\lfloor \mathbf{P}^{\mathbf{s}}(\mathbf{x}) : \boldsymbol{\sigma}(\mathbf{x}) \rfloor}{\tau_c^s(\mathbf{x})} \right)^{n} sign(\mathbf{P}^{\mathbf{s}}(\mathbf{x}) : \boldsymbol{\sigma}(\mathbf{x})), \qquad \textbf{Eq. 1}$$

$$\mathbf{P}^{\mathbf{s}}(\mathbf{x}) = \frac{1}{2}(\mathbf{b}^{\mathbf{s}} \otimes \mathbf{n}^{\mathbf{s}} + \mathbf{n}^{\mathbf{s}} \otimes \mathbf{b}^{\mathbf{s}}), \qquad \textbf{Eq. 2}$$

where $\dot{\gamma}^{s}(\mathbf{x})$, $\dot{\gamma}_0 = 0.001\ s^{-1}$ and $n = 20$ are the shear rate, reference shear rate, and power law viscoplastic exponent, respectively. $\tau_c^s(\mathbf{x})$ denotes the slip resistance, or critical resolved shear stress (CRSS). The geometry of a slip system *s* is defined by the Burgers vector, $\mathbf{b}^s$, and slip plane normal, $\mathbf{n}^s$, with the tangent vector being $\mathbf{t}^s = \mathbf{b}^s \times \mathbf{n}^s$. The family of 12 slip systems for face centered cubic (FCC) metals is given by $\{\bar{1}11\}\langle 110\rangle$.

The current slip resistance τ_c^s is written as the sum of contributions from the initial CRSS, τ_0^s, the contribution of mobile forest dislocations, τ_{forest}^s, and the slip resistance arising from sessile debris due

[†] In the following notation, tensors are denoted by bold letters while scalars are italic and not bold. The dot and tensor products between the two tensors are denoted by "·" and "⊗", respectively

to the kink-jog dislocation interactions at later stages of hardening, τ_{debris}, as follows (Beyerlein and Tomé, 2008; Eghtesad and Knezevic, 2021b, 2020b; Zecevic et al., 2016):

$$\tau_c^s = \tau_0^s + \tau_{forest}^s + \tau_{debris}. \qquad \textbf{Eq. 3}$$

Details on the DFT-informed DD hardening formulation describing the individual parameters and DFT calculations can be found in (Eghtesad et al., 2022).

2.2 Microstructural features

The microstructural features studied in the present work include crystallographic orientations of individual grains, grain morphology, GB features, Schmid factor, intergranular misorientation, and slip transmission factor, or Luster-Morris parameter, m′ (Bayerschen et al., 2016; Luster and Morris, 1995). The role of the Luster-Morris parameter in determining the likelihood of slip transfer across grain interfaces was found important in selecting active slip systems and predicting texture evolution (Riyad et al., 2021). Fig. 2 shows these features and the considered fields, which are discussed individually in following sections.

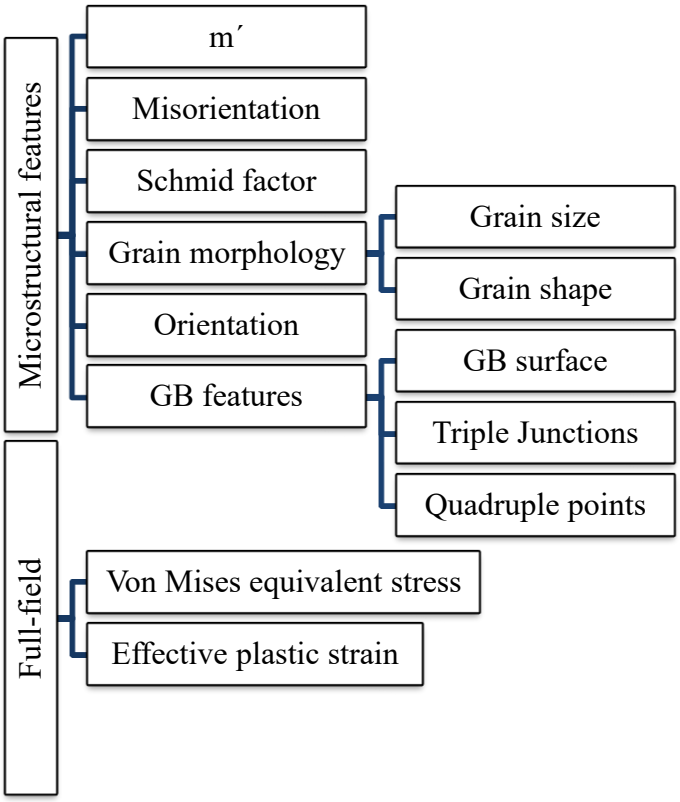


Fig. 2. Microstructural features and fields considered in this study.

2.2.1 Crystal orientations

The orientations of individual grains can be described using different conventions, including the transformation matrix composed of direction cosines (Kitayama et al., 2013), Bunge-Euler angles (Bunge, 2013), Miller indices (Frank, 1965; Schwarzenbach, 2003), axis-angle pairs (Kocks et al., 1998), Rodriguez vector (Neumann, 1991), and quaternions (Takahashi et al., 1985). The components of the axis-angle pair, $\boldsymbol{\theta} = \theta\hat{\boldsymbol{r}}$**,** are given as:

$$\theta = \cos^{-1}\left(\frac{\sum_i Q_{ii} - 1}{2}\right), \quad \textbf{Eq. 4}$$

$$r_1 = \frac{Q_{23} - Q_{32}}{2\sin\theta}, r_2 = \frac{Q_{31} - Q_{13}}{2\sin\theta}, \; r_3 = \frac{Q_{12} - Q_{21}}{2\sin\theta}, \quad \textbf{Eq. 5}$$

where, θ is the rotation angle, $\hat{\boldsymbol{r}}$ is the unit direction vector, and **Q** is the transformation matrix, which converts the coordinates of the crystal to the sample frame, and is defined as:

$$\mathbf{Q} = \begin{bmatrix} \cos\phi_1\cos\phi_2 - \sin\phi_1\cos\Phi\sin\phi_2 & -\cos\phi_1\sin\phi_2 - \sin\phi_1\cos\Phi\sin\phi_2 & \sin\phi_1\sin\Phi \\ \sin\phi_1\cos\phi_2 + \cos\phi_1\cos\Phi\cos\phi_2 & -\sin\phi_1\sin\phi_2 + \cos\phi_1\cos\Phi\cos\phi_1 & -\cos\phi_1\sin\Phi \\ \sin\Phi\sin\phi_2 & \sin\Phi\cos\phi_2 & \cos\Phi \end{bmatrix}, \quad \textbf{Eq. 6}$$

where, ϕ_1, Φ and ϕ_2 are the Bunge-Euler angles.

The Euclidean distance between a pair of orientations is calculated using their corresponding unit vectors, $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$. The distance between an arbitrary orientation **P**=**Q** (ϕ_1, Φ, ϕ_2) and a crystal direction *[hkl]* as shown in Fig. 3, can be described as:

$$r^{\mathbf{P}-[hkl]} = \sqrt{\left(r_1^{\mathbf{P}} - r_1^{[hkl]}\right)^2 + \left(r_2^{\mathbf{P}} - r_2^{[hkl]}\right)^2 + \left(r_3^{\mathbf{P}} - r_3^{[hkl]}\right)^2}, \quad \textbf{Eq. 7}$$

where [*hkl*] represents each of the three selected directions of [001], [101], and [111], and $r_i^{\mathbf{P}}, i = 1,3$ are the unit direction vectors of orientation **P**=**Q** (ϕ_1, Φ, ϕ_2), with [100], [010] and [001] directions parallel to X, Y and Z axes of sample coordinates. Fig. 3b shows the Euclidean distances between individual grains within an example RVE microstructure.

2.2.2 Schmid factor

Schmid's law (Schmid and Boas, 1950) describes the relationship between applied stress, σ, and the resolved shear stress, τ, on a crystallographic slip system as follows:

$$\tau = \sigma \cos\lambda \cos\xi = \sigma m, \qquad \textbf{Eq. 8}$$

where $\cos\lambda \cos\xi = m$ is the Schmid factor, λ is the angle between the loading direction and Burger's vector, and ξ is the angle between the loading direction and slip plane normal. Fig. 3d illustrates the distribution of Schmid factor with respect to the Y axis within an example microstructure.

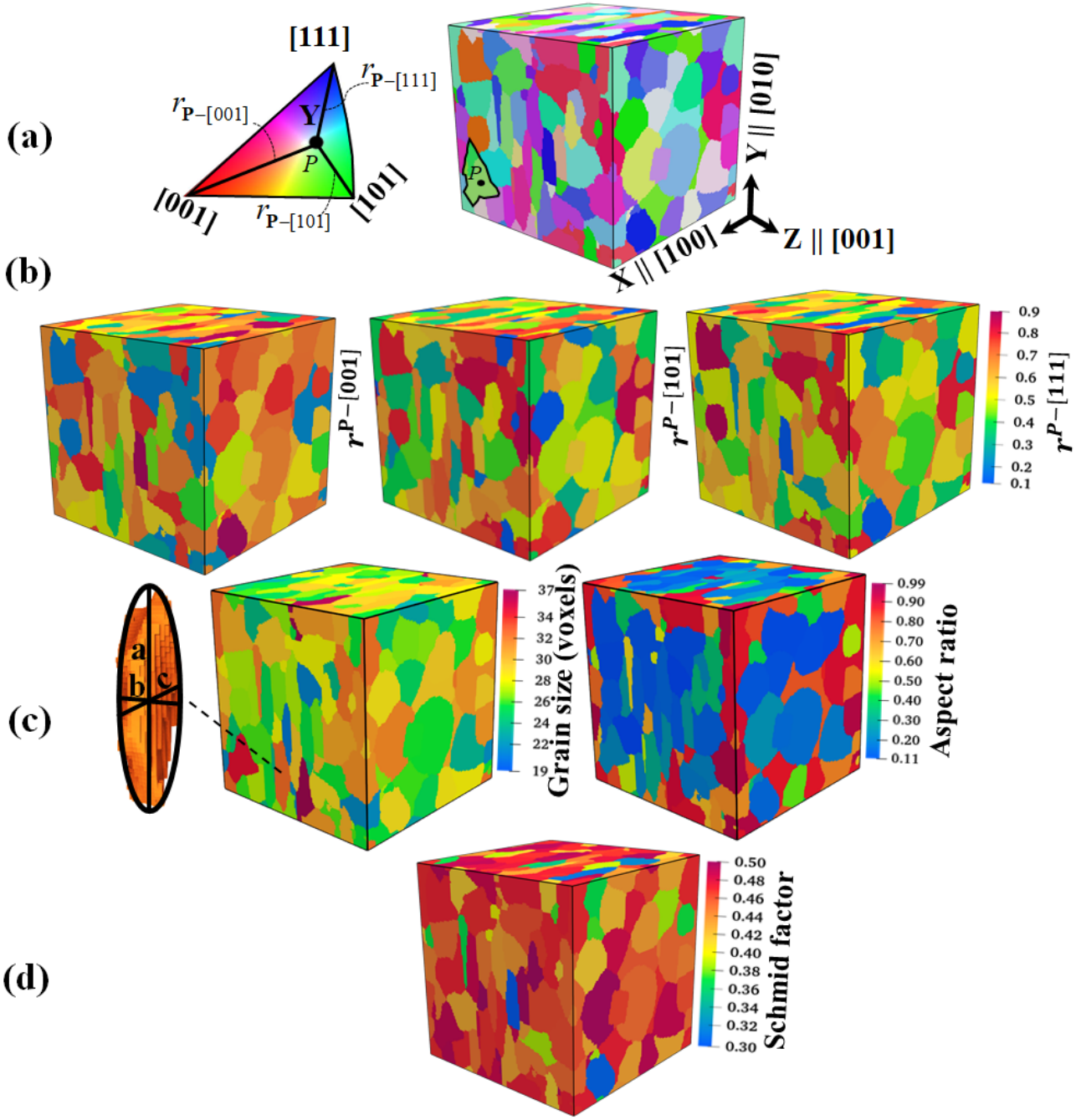


Fig. 3. (a) Euclidean distances of an orientation P=Q ($\boldsymbol{\phi_1}, \boldsymbol{\Phi}, \boldsymbol{\phi_2}$) with respect to the directions [001], [101] and [111], the corners of a standard inverse pole figure triangle correspond to Y axis, (b) Euclidean distances within an example microstructure for individual grains, (c) representation of an elongated grain with ellipsoid and quantifications of grain size and grain aspect ratio, (d) Schmid factor distribution for the entire microstructure RVE.

2.2.3 Grain morphology

The grain morphology and the distribution of grain sizes and aspect ratios within an example RVE are quantified by considering the elongated grains as ellipsoids with cord lengths a, b, and c. Realization of ellipsoids from the voxelized grains come from the DREAM3D pipeline "Find Feature Shapes", where, the second-order moments of each feature are calculated to determine the ellipsoid principal axis lengths and directions (Groeber and Jackson, 2014b). Grain size is defined as the equivalent sphere diameter (ESD), $ESD = \sqrt{a^2 + b^2 + c^2}$. Grain shape is defined by the minimum aspect ratio, $min(\frac{b}{a}, \frac{c}{a})$ as shown in Fig. 3c, with *a* as the major axis.

2.2.4 Grain boundary character

GB character is defined by five independent crystallographic and geometrical parameters, three of which form the lattice intergranular misorientations (ϕ_1, Φ and ϕ_2) while the other two represent the GB plane normal (Beladi et al., 2020). In this study, we also use the Euclidean distances between voxels within the grain and the location of GB geometrical features such as GB planes, lines, and vertices and slip transmission compatibility factor, described in the next subsections.

2.2.5 Grain boundary geometrical features

A GB consists of three major features: (i) surfaces where two adjacent grains meet, (ii) triple junctions where three grains meet, and (iii) quadruple points at the intersections of four grains (Gottstein et al., 2010; Gottstein and Shvindlerman, 2006, 2005; Rios and Glicksman, 2015; Zhao et al., 2011). Fig. 4a shows an illustration of the three types of GB features for an arbitrary grain within an example microstructure. The distances of each voxel within a RVE to the nearest GB features, r_{GB}, r_{TJ}, and r_{QP} were calculated, with these values for an example RVE shown in Fig. 4b.

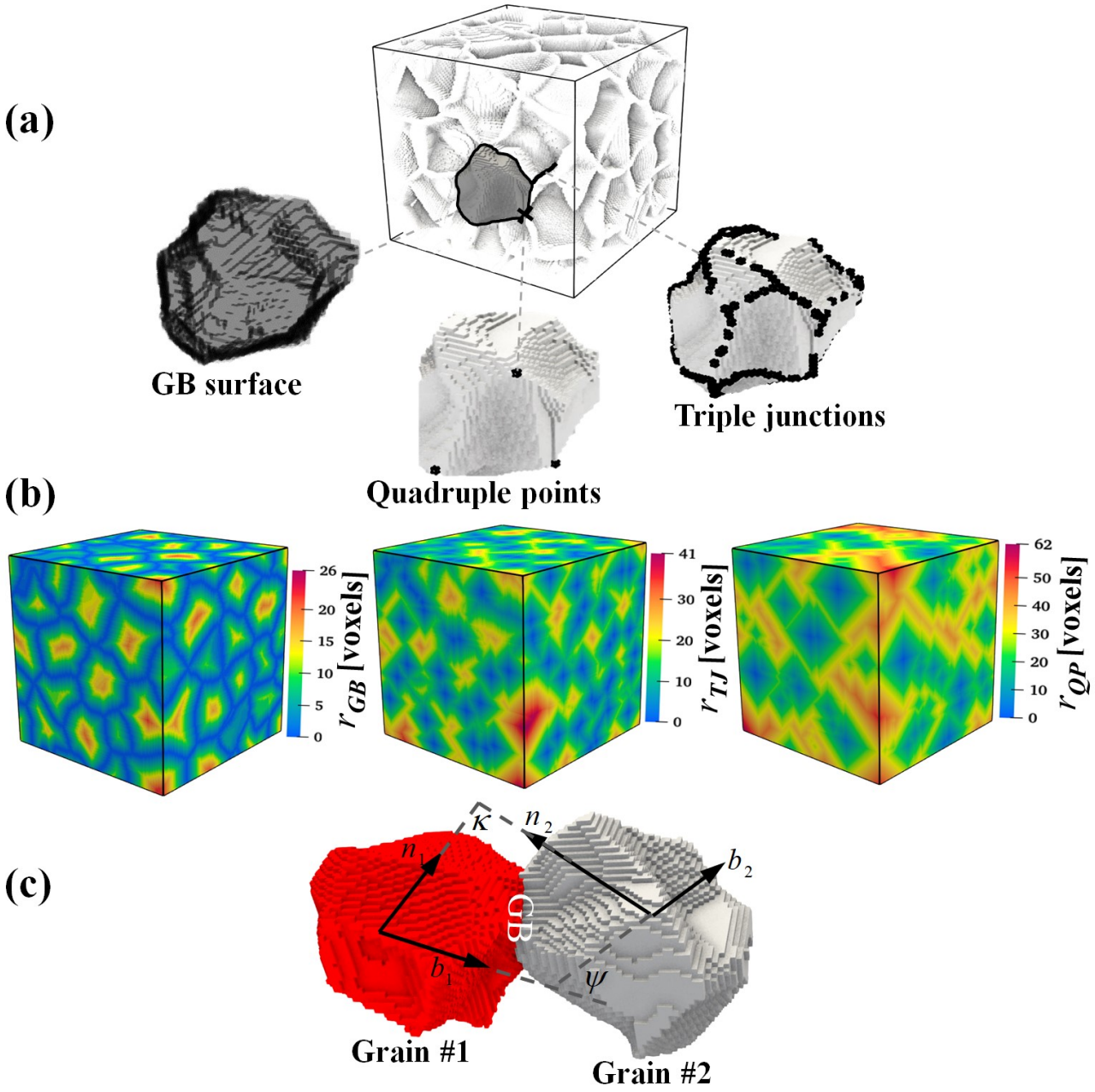


Fig. 4. (a) Illustration of GB features with surface, triple junctions, and quadruple points for a grain within an RVE, (b) distances between GB features to voxels within individual grains for a sample RVE, (c) Schematic illustrating the angles for Luster-Morris parameter between two adjacent grains.

2.2.6 Intergranular misorientation and slip transmission compatibility factor

Recent studies indicate that hotspots are more likely to initiate at the high angle grain boundaries with GB intergranular misorientations ranging from 45º to 60º (Bieler et al., 2009; Eghtesad and Knezevic, 2021c). The intergranular misorientation between two adjacent grains α and β is defined as follows:

$$\Delta Q_{ij}^{\alpha\beta} = \left(Q_{ij}^{\alpha}\right)^{-1} Q_{ij}^{\beta}, \qquad \textbf{Eq. 9}$$

The scalar rotation angle $\theta^{\alpha\beta} = \cos^{-1}\left(\frac{\sum_i \Delta Q_{ii}^{\alpha\beta} - 1}{2}\right)$ between the two adjacent grains at the GB is defined using the axis-angle pair convention in Eq. 4.

The Luster-Morris parameter, m', is a measure of the ability of slip transmission between two adjacent grains, given as (Alizadeh et al., 2020):

$$m' = \cos\kappa \cos\psi, \qquad \textbf{Eq. 10}$$

where κ and ψ are the angles between the slip plane normals and slip directions of two adjacent grains, as illustrated in Fig. 4c. A value of $m' = 1$ indicates co-planarity of slip systems between the two grains and a fully transmissible GB, while $m' = 0$ indicates the grain boundary is impenetrable, which would eventually result in dislocation pileups and field localizations near the GBs.

While it is possible to calculate a single value of intergranular misorientation and m' per grain by averaging over neighboring grains (Mangal and Holm, 2018a), here we use two methods to account for the heterogeneity of these quantities along the GB segments like in(Knezevic et al., 2014). In the first method, as shown in Fig. 5a, the intragranular misorientation is obtained by inversely correlating the weighted average to the Euclidean distances with respect to the GB surfaces as follows:

$$\theta_{\text{intragranular}} = \left(\frac{ESD - r_{GB}}{ESD}\right) \times \frac{\sum_{i=1}^{N}(S_{GB,i} \times \theta)}{\sum_{i=1}^{N} S_{GB,i}} \quad \textbf{Eq. 11}$$

where $S_{GB,i}$ is the area of the GB surface for each grain neighbor, and N is the total number of neighbors of the grain. While this method allows quantification of intragranular misorientation and m', as shown in Fig. 5b-c, it obscures the individual characteristics of the GBs corresponding to different neighbors. Thus, in the second method, the intergranular misorientation and m' for each GB was computed as shown in Fig. 5d and e. In this method, a pair of neighbor grains with different grain IDs are identified as grain boundaries through a search algorithm applied to the individual voxels defining the grains. Once the grain boundary voxels are identified, their corresponding intragranular misorientation and m' are calculated using Eq. 9 and Eq. 10.

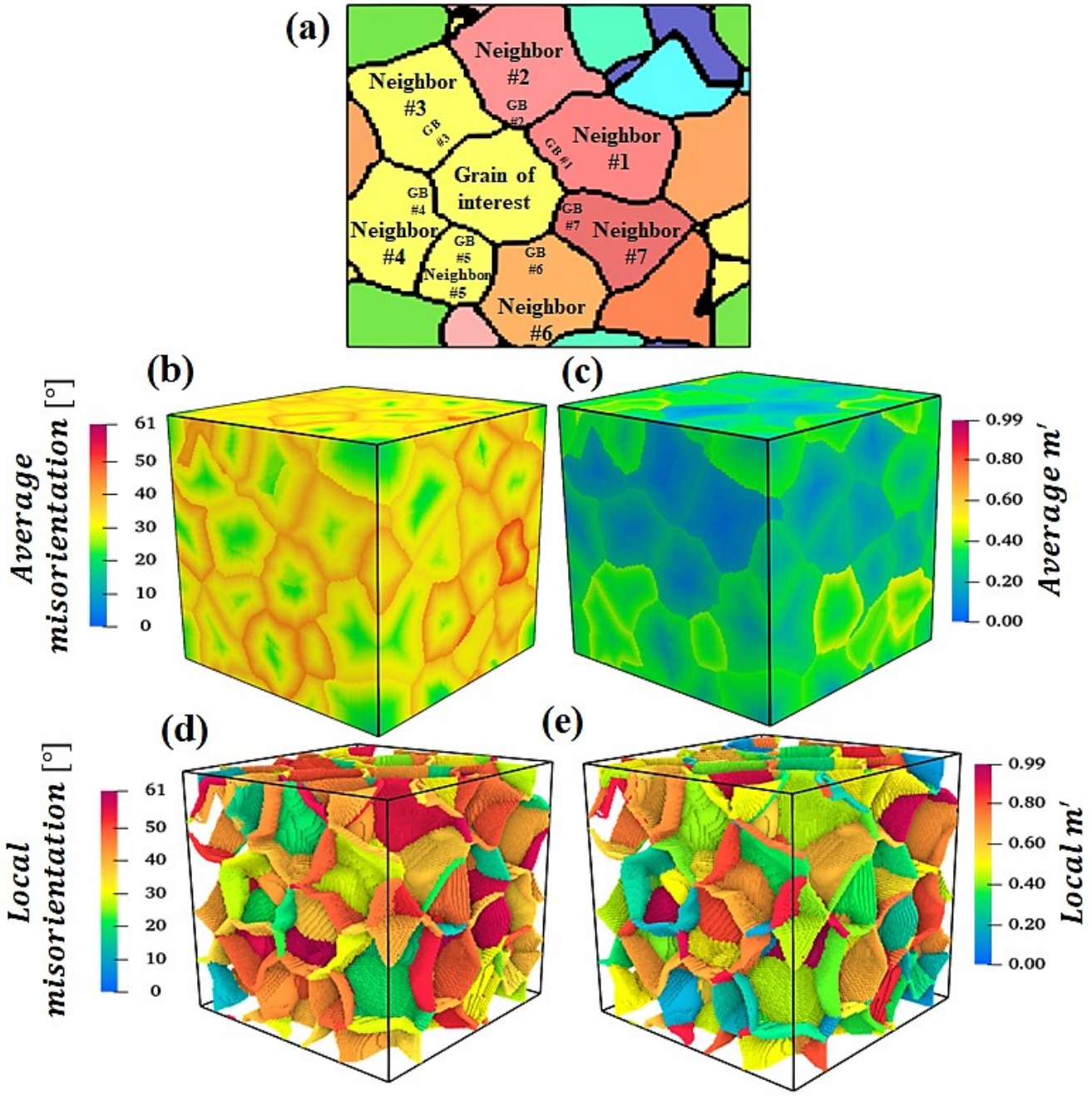


Fig. 5. (a) Neighbor grains and shared GB areas used for calculation of a weighted average misorientation angle and Luster-Morris parameter, $\boldsymbol{m'}$, (b) weighted gradient misorientation, (c) weighted gradient $\boldsymbol{m'}$, (d) local intergranular misorientation at GBs, and (e) local $\boldsymbol{m'}$ at GBs for an example RVE.

2.3 Correlation analysis

2.3.1 Pearson correlation coefficient

The Pearson correlation coefficient (PCC), r^{PCC}, measures the strength of linear association between the two variables, where $r^{PCC} = \pm 1$ indicate a strong positive/negative linear correlation and $r^{PCC} = 0$ implies no linear correlation. The PCC for a pair of data represented in form of vectors **X** and **Y**, with length *N*, is written as (Jebli et al., 2021):

$$r^{PCC} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})\ \sum_{i=1}^{N}(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}\ \sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}}, \quad \textbf{Eq. 12}$$

where, $\bar{X}$ and $\bar{Y}$ are the mean values of vectors **X** and **Y**.

2.3.2 Ensemble learning

Ensemble learning-based ML algorithms (Krawczyk et al., 2017; Sagi and Rokach, 2018; Zhang and Ma, 2012), such as Bagging (Breiman, 1996), AdaBoost (Solomatine and Shrestha, 2004), XGBoost (Chen and Guestrin, 2016), and random forest (Breiman, 2001; Ho, 1995), enable accurate construction of nonlinear data associations, by combining the predictions of multiple simpler ML models with a trade-off between accuracy and computational cost (LeDell, 2015; Sagi and Rokach, 2018). Among ensemble learning methods, random forest offers an ensemble hybridization training approach, improving the overfitting drawbacks present in its alternatives (Ardabili et al., 2019; Sagi and Rokach, 2018).

In the present study, we use the random forest regression ensemble learning method based on decision trees as shown in Fig. 6. As shown later, microstructural features present strongly nonlinear correlations

with stress and strain hotspots; thus, random forest regression, which can identify arbitrarily complex functions and correlations, is a suitable choice for identifying the correlations between hotspots and the microstructural data provided here.

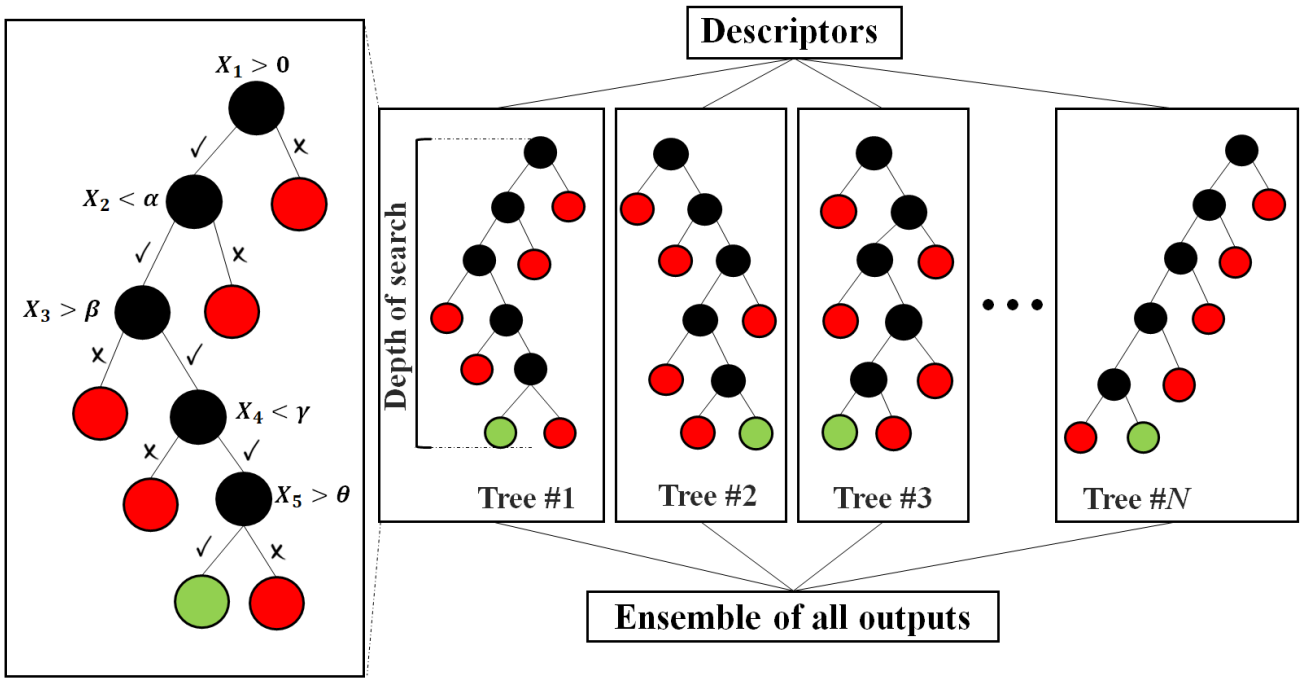


Fig. 6. Schematic of random forest ensemble learning. Final decisions are made based on predictions of all decision trees.

2.4 Shapley additive explanations

While ML models such as random forest offer high levels of predictive performance, their inner workings are often inexplicable except in the most abstract sense, which hinders the physical interpretation of their conclusions. Several interpretation methods have recently been proposed to illuminate the black-box nature of ML models. These methods are either specific to one ML model (i.e., model-specific) or generic for all ML models (i.e., model-agnostic). Among the model-agnostic

methods, Shapley additive explanations (SHAP) (Lipovetsky and Conklin, 2001; Lundberg and Lee, 2017; Merrick and Taly, 2020; Rozemberczki et al., 2022; Shapley et al., 1953) are an extension of the local interpretable model-agnostic explanations (LIME) (Ribeiro et al., 2016) within complex ML models such as random forests. SHAP uses a network of graphs to identify the features that alter the probability of a prediction and provides visualizable patterns that determine those predictions. Shapley values provide an interpretation of feature importance by satisfying three major mathematical properties of a permutation:(i) additivity: the overall net importance of features is directly related to sum of importance of those features, (ii) symmetry: importance of any two features that contribute equally to the predictions are the same, and (iii) dummy: SHAP value is zero in the case of zero contribution (Winter, 2002).

An explainable model ζ with feature vector x can be written in form of additive feature attributions as follows:

$$\zeta(x) = \omega_0 + \sum_{i=1}^{N} \omega_i x_i, \qquad \textbf{Eq. 13}$$

where ω_i is the weight of feature i among N features. In the SHAP model, ω_i is defined as follows (Gómez-Ramírez et al., 2020; Rodríguez-Pérez and Bajorath, 2019; Rozemberczki et al., 2022):

$$\omega_i = \frac{1}{N!} \sum_{S \subseteq N} S!\,(N - S - 1)!\,[f(S \cup \{x_i\}) - f(S)], \qquad \textbf{Eq. 14}$$

where $N!$ is all possible permutations of the total number of features, S is a subset of features, $S!$ is all possible permutations of the subset of features, and $f(S)$ is the prediction corresponding to a subset of features S.

Since Shapley additive explanations are permutation-based, predicted correlations are prone to bias due to preferential selection of features within the tree search algorithm (Mangal and Holm, 2018a; Strobl et al., 2008). To address such issue, we use a sequential feature selection (SFS) wrapper method (Chen and Chen, 2015; El Aboudi and Benhlima, 2016; Hall and Smith, 1999; Wang et al., 2015; Xiao et al., 2007) with an iterative K-fold cross-validation (Fushiki, 2011; Refaeilzadeh et al., 2009) to minimize bias in selection of features responsible for the formation of hotspots.

2.5 Feature selection and wrapper method

Wrapper-based SFS methods (Chen and Chen, 2015; El Aboudi and Benhlima, 2016; Hall and Smith, 1999; Wang et al., 2015; Xiao et al., 2007) are ML algorithms that use a sequential procedure to select the most important feature combinations contributing to an output. SFS methods are classified as either forward selection or backward elimination. Forward selection is based on adding features to a null set, while in backward elimination features are removed from the full set of features. One feature is selected at each step, and the next feature is chosen under the condition that the overall performance is improved. The performance is evaluated by K-fold cross-validation (Fushiki, 2011; Refaeilzadeh et al., 2009) to avoid the uncertainties arising from the data distribution. The final output of the SFS method is a combination of features that result in the best model performance - i.e., the highest accuracy/score in the regression model. Fig. 7 is a flowchart of a forward selection SFS using the random forest model. The wrapper method starts with adding features to a null list, swapping features out to find the next feature at that step that improves the model's cross validation score. Here, the wrapper result rankings were constructed based on the relative frequency of individual features occurring in the list over 100 repetitions.

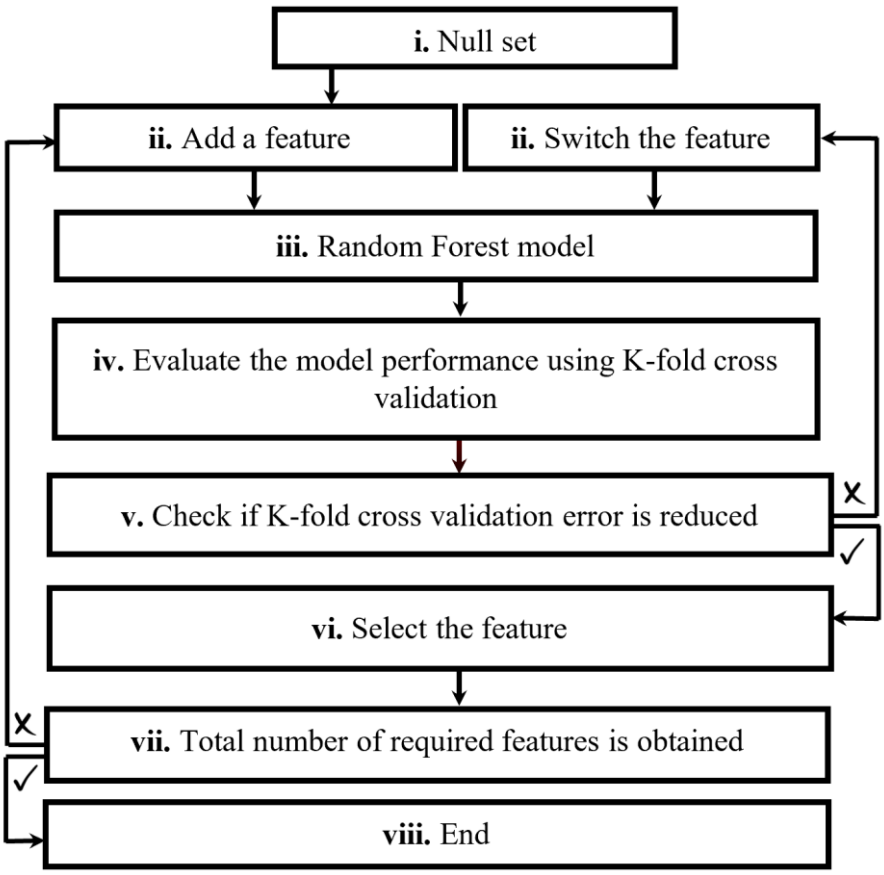


Fig. 7. Flowchart of forward sequential feature selection. The procedure starts with a null set. Features are added to the list and their correlations with hotspots are predicted using a random forest model. The random forest model performance is evaluated using K-fold cross-validation. A final set of features is obtained based on error minimization by K-fold cross-validation.

3 Results and discussion

While recent studies have identified microstructural features related to the formation of hotspots under uniaxial loading for equiaxed and untextured grains (Mangal and Holm, 2018b), this study expands on prior work by examining variations in applied deformation state, grain morphology, and texture.

3.1 Micromechanical deformation in polycrystalline Ni under complex deformation states

The CPFFT model with DFT informed DD hardening law was used to simulate the flow response of pure Ni polycrystal with a grain size of 2μm, using the DD hardening parameters and properties from DFT-based calculations provided in Table 1 (Eghtesad et al., 2022).

Table 1. DD hardening parameters and DFT calculations for pure Ni

Quantitiy	Description	Value	Calculation method
$C_{11}[GPa]$		275.5	DFT
C_{12} $[GPa]$	Elastic constants	160.1	DFT
C_{44} $[GPa]$		126.4	DFT
$b_{\langle 11\bar{2}\rangle}$ $[nm]$	Partial Burges vector	0.14	DFT
μ $[GPa]$	Shear modulus	92.3	DFT
Γ_{USFE} $[Jm^{-2}]$	Unstable stacking fault energy	0.28	DFT
Γ_{ISFE} $[Jm^{-2}]$	Stable stacking fault energy	0.13	DFT
H_{Va}^{F} $[EV]$	Vacancy formation energy	1.65	DFT
g	Normalized activation energy	0.01	DFT
q	Rate of debris formation	5.40	DFT
$\tau_0^S[MPa]$	Initial slip resistance	76.0	Calibration
$k_1[m^{-1}]$	Hardening rate	9.0e7	Calibration
$D[MPa]$	Drag stress	308	Calibration

To investigate the impact of applied macroscopic loading condition on micromechanical hotspots, we subjected the RVEs to the macroscale deformation states listed in Table 2. Note that a mixed boundary conditions are used to enforce zero stress for lateral directions under uniaxial loading (Lebensohn et al., 2012b). The resultant von Mises equivalent stress fields are provided in Fig. 8d-e showing significant variation in field localizations within the microstructure.

Table 2. Applied deformation states and corresponding deformation gradient tensors in CPFFT.

Applied deformation states	Deformation gradient tensor
Uniaxial tension	$\begin{bmatrix} \frac{-\dot{\varepsilon}^{app}}{2} & 0 & 0 \\ 0 & \dot{\varepsilon}^{app} & 0 \\ 0 & 0 & \frac{-\dot{\varepsilon}^{app}}{2} \end{bmatrix}$
Biaxial tension	$\begin{bmatrix} \dot{\varepsilon}^{app} & 0 & 0 \\ 0 & \dot{\varepsilon}^{app} & 0 \\ 0 & 0 & -2\dot{\varepsilon}^{app} \end{bmatrix}$
Simple shear	$\begin{bmatrix} 0 & \dot{\varepsilon}^{app} & 0 \\ \dot{\varepsilon}^{app} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Combined tension/shear #1	$\begin{bmatrix} \frac{-\dot{\varepsilon}^{app}}{2} & \dot{\varepsilon}^{app} & 0 \\ \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & 0 \\ 0 & 0 & \frac{-\dot{\varepsilon}^{app}}{2} \end{bmatrix}$
Combined tension/shear #2	$\begin{bmatrix} \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & 0 \\ \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & 0 \\ 0 & 0 & -2\dot{\varepsilon}^{app} \end{bmatrix}$
Combined tension/shear #3	$\begin{bmatrix} \frac{-\dot{\varepsilon}^{app}}{2} & \dot{\varepsilon}^{app} & 0 \\ \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} \\ 0 & \dot{\varepsilon}^{app} & \frac{-\dot{\varepsilon}^{app}}{2} \end{bmatrix}$
Combined tension/shear #4	$\begin{bmatrix} \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & 0 \\ \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} & \dot{\varepsilon}^{app} \\ 0 & \dot{\varepsilon}^{app} & -2\dot{\varepsilon}^{app} \end{bmatrix}$

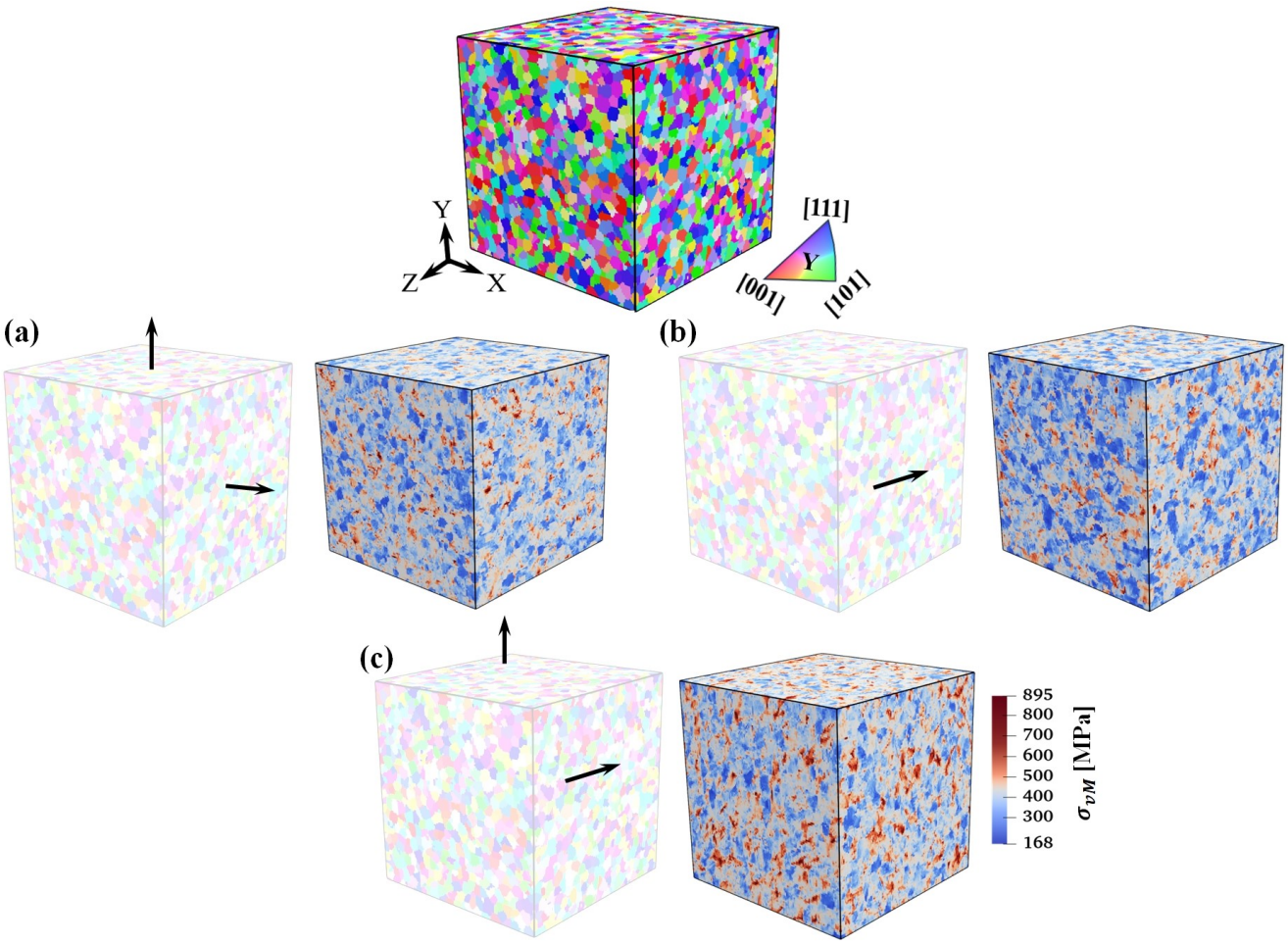


Fig. 8. Variation in von Mises equivalent stress as function of deformation state to a macroscopic strain of 2.5% for RVEs with equiaxed grains deformed under: (a) biaxial tension, (b) simple shear, and (c) combined tension/shear.

3.2 Identification of stress and strain hotspots

The CPFFT model enables quantification of micromechanical fields within the microstructure throughout the deformation history. While stress hotspots give useful insights into fracture and failure criteria in elastic regimes, they obscure the failure and dissipation effects resulting from localization of plastic strain. This becomes more important under multiaxial loading in ductile materials, or in fatigue

or creep, for which accumulated plastic strain plays a significant role in determining the failure behavior of the material (Chaboche, 2003). In this work, in addition to the stress hotspots studied in literature (Mangal and Holm, 2018b; Rollett et al., 2010a), we also consider the strain localizations, which are important for shear banding-type failure (Ardeljan et al., 2015)

Herein, instead of grain-wise average of the fields, as reported in (Mangal and Holm, 2018a), a local (i.e., voxel-based) thresholding is used to identify the hotspot locations and account for the gradients of the fields near the GBs. After obtaining the full-field information from CPFFT simulations extracted at 2.5% total strain, the stress and strain hotspot locations were identified as those exceeding 95% of the mean value of the field of interest. Fig. 9 shows the distribution of hotspots for von Mises equivalent stress and effective plastic strain after applying the thresholding.

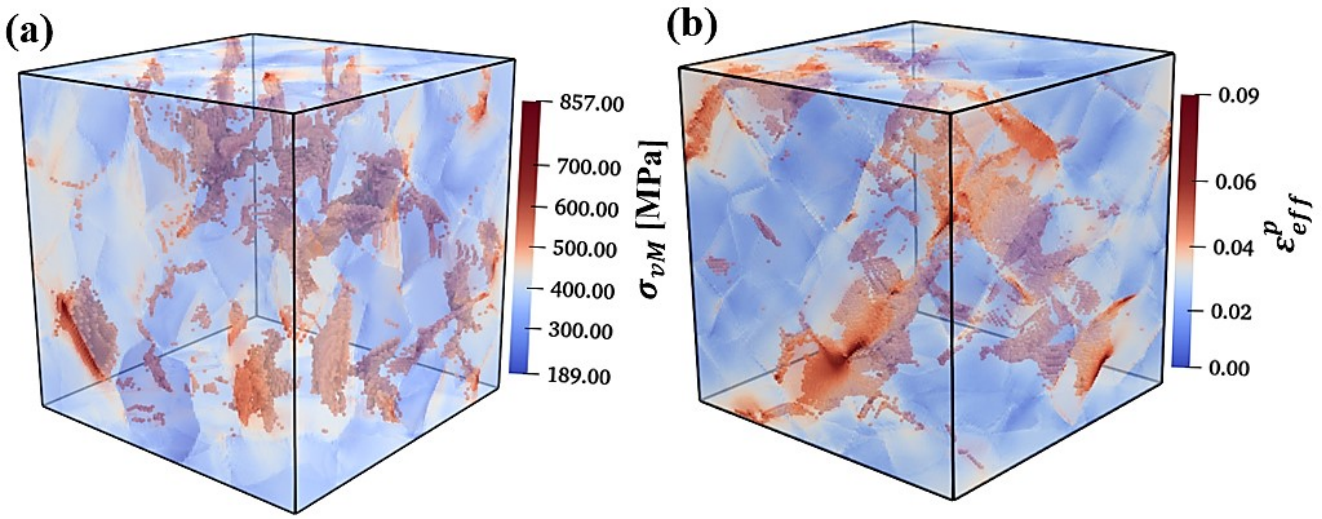


Fig. 9. Distribution of full-field values in pure Ni under uniaxial tension to 2.5% macroscopic strain: (a) von Mises equivalent stress and (b) effective plastic strain.

3.3 Variation in texture and grain morphology

The present work considers the effects of both textured and untextured (i.e., randomly distributed grain orientations) RVEs. Common FCC textures were investigated, including Goss, Brass, Cube and Copper (Butrón-Guillén et al., 1994; Jonas, 2009; Sarma et al., 2004). Fig. 10 shows the pole figures, plotted with MTEX software (Bachmann et al., 2010), associated with the types of texture used in this study with non-uniform textures showing strong intensities of preferred orientations up to 15.0 for {100} components.

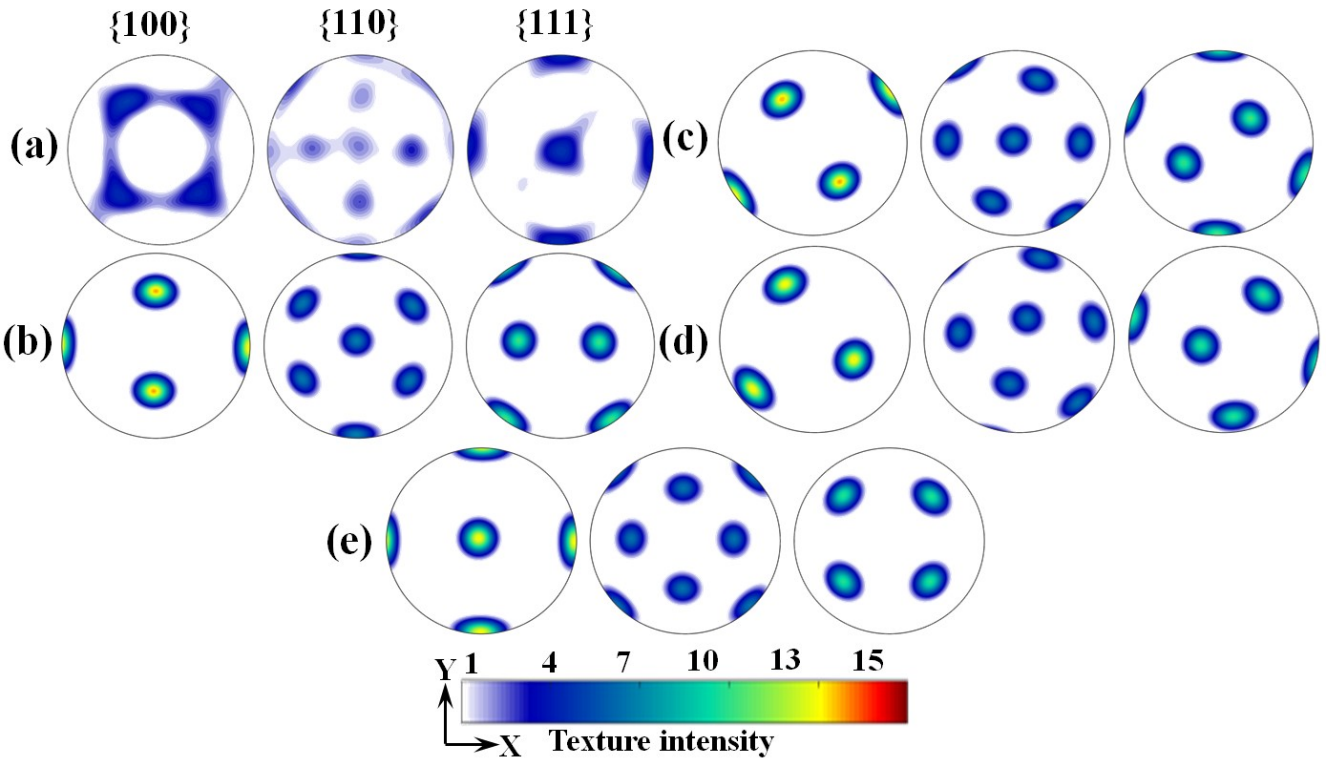


Fig. 10. Texture variants for FCC, (a) uniform, (b) Goss, (c) Brass, (d) Copper and (e) Cube.

To consider the effects of grain morphology, microstructures with a constant number of voxels (128^3), varying from equiaxed grains to elongated grains and mixed equiaxed and elongated grains were

generated as shown in Fig. 11a-c with distribution of equivalent sphere diameter (ESD) and aspect ratio of the elongated grains shown in Fig. 12. Corresponding von Mises equivalent stress fields obtained after an applied uniaxial deformation of 2.5% strain, reveal significant variations in stress fields and hotspot locations.

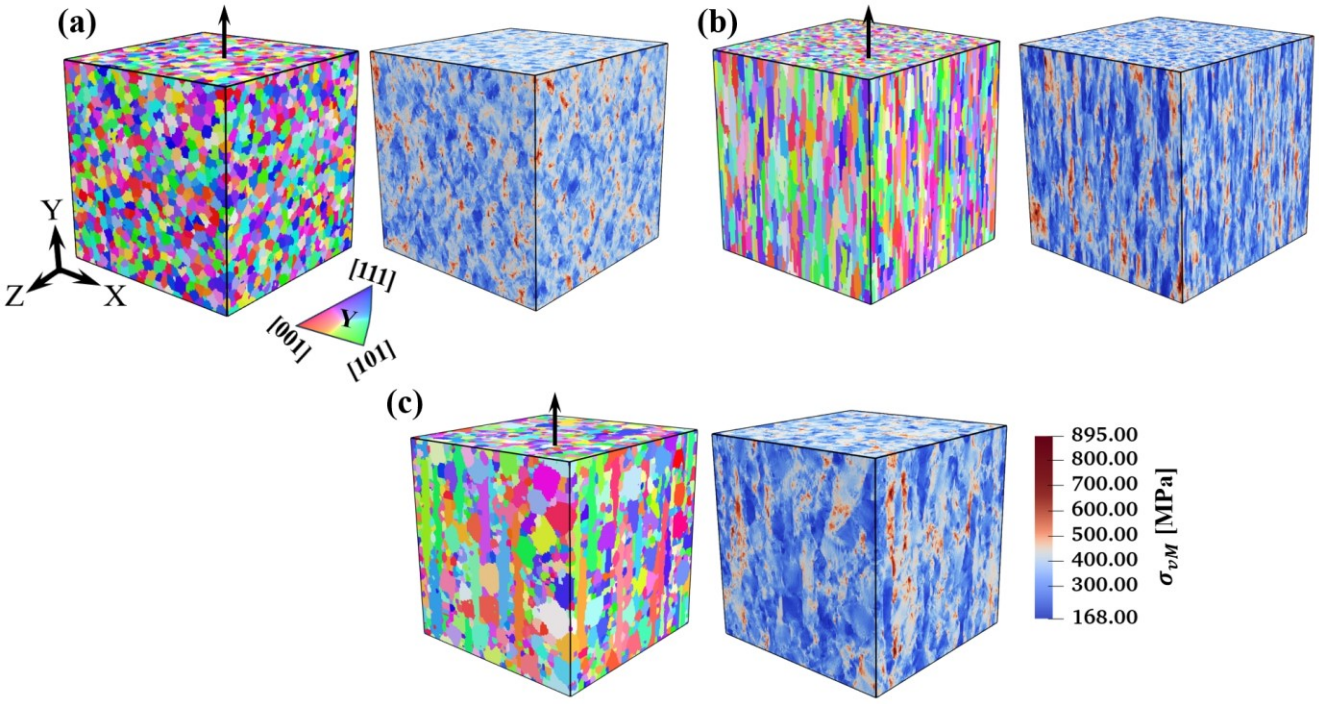


Fig. 11. Variation in von Mises equivalent stress as function of grain morphology in RVEs with uniform texture deformed under uniaxial tension to 2.5% macroscopic strain with: (a) equiaxed grains, (b) elongated grains, and (c) mixed equiaxed/elongated grains.

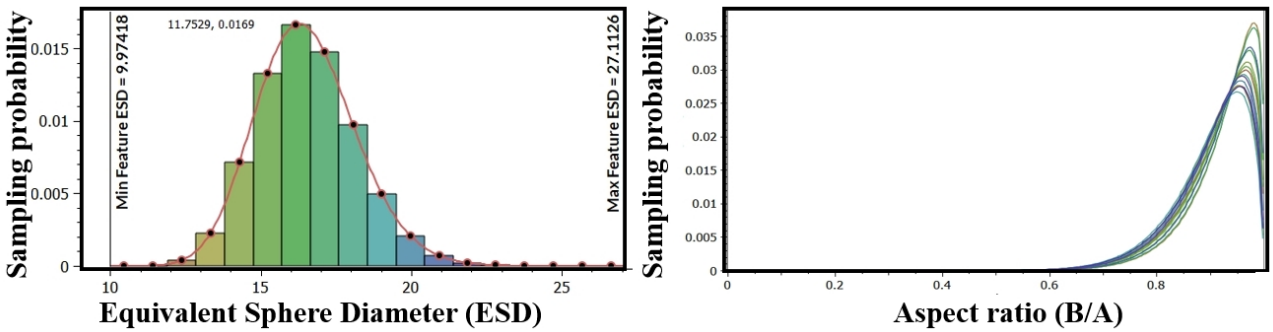


Fig. 12. Distribution of equivalent sphere diameter (ESD) and aspect ratio of the elongated grains for the microstructure with mixed equiaxed/elongated grains.

3.4 Correlation heatmaps

The degree of correlation between microstructural features and hotspots is evaluated in terms of PCC correlation metrics. Variations in correlation intensities for an equiaxed RVE with uniform texture under uniaxial tension are given in Fig. 13. More results for stress hotspots as a function of grain morphology, texture, and deformation state are provided in the supplementary material (Fig A1). Cells with higher non-zero absolute values indicate stronger correlations. Notable correlations are observed with misorientation, m', Euclidean distances from the GB features, and Euclidean distances to the directions [001], [101] and [111]. Note that the Euclidean distances from the GB features are more appropriate in the case of RVE with a small number of grains, as individual grains are represented by significantly more voxels. Correlations with misorientation and Schmid factors are positive, indicating that stress hotspots occur at grains with higher Schmid factors and larger misorientations.

The conclusion related to correlations between misorientation and hotspots differs from what was found in (Mangal and Holm, 2018a) due to different representations of misorientation. While (Mangal and Holm, 2018a) used an average of misorientations of neighbor grains and assigns that to the entire grain, the present work considers the local intergranular misorientation that accounts for the nonuniformity across the GBs with dissimilar grain neighborhood. Negative correlations with Euclidean distances from the GB features confirm the fact that hotspots form near the GBs, which agrees with the findings reported in (Mangal and Holm, 2018a; Rollett et al., 2010b).

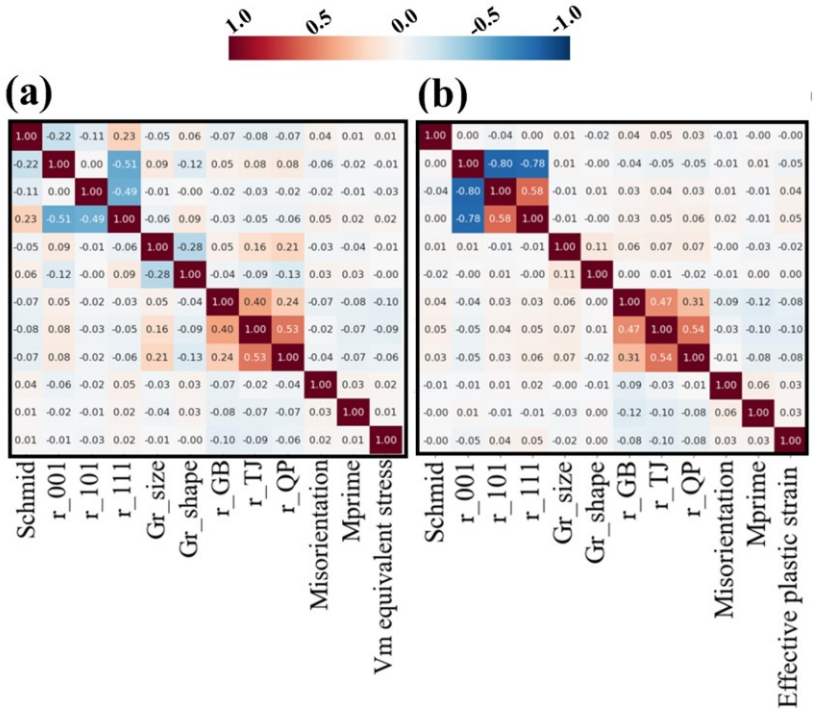


Fig. 13. PCC correlation heatmaps for microstructural features hotspots under uniaxial tension for RVEs with equiaxed grains and uniform texture: (a) Von Mises equivalent stress and (b) effective plastic strain.

While these heatmaps provide useful insights into linear positive/negative correlations, values far from unity indicate the possibility of nonlinear correlations between microstructural features and

micromechanical fields. In (Mangal and Holm, 2018a), a linear LASSO regularization regression provided by FeaLect method (Tibshirani, 1996; Zare, 2015) was used to select the most important features. In the present work, we selected a random forest model within an SFS algorithm for feature selection to identify nonlinear correlations and the relative importance of microstructural features contributing to the formation of hotspots.

3.5 SHAP-based feature importance and random forest-based feature selection

The SFS function "SequentialFeatureSelector()" from Scikit-learn using random forest regression as the ML core model and 5-fold cross-validation were used in the present study. The accuracy of random forest regression with all features included was evaluated as a function of the number of decision trees and the depth of tree search. The random forest regression function "RandomForestRegressor()" from Scikit-learn (Pedregosa et al., 2011) was used to find the nonlinear correlations between microstructural features and stress hotspots. The random forest model reached a maximum accuracy of R^2 = 87% and root mean square error of *RMSE*=0.09 by using 75 trees and a search depth of 35 as shown in Fig. 14. Next, the optimum number of features resulting in the same accuracy was found by iteratively running the SFS wrapper and selected features using a 5-fold cross-validation, varying from 1 to 10, saturating after 4 features. A maximum score of R^2 = 87% was obtained by using a subset of four selected features.

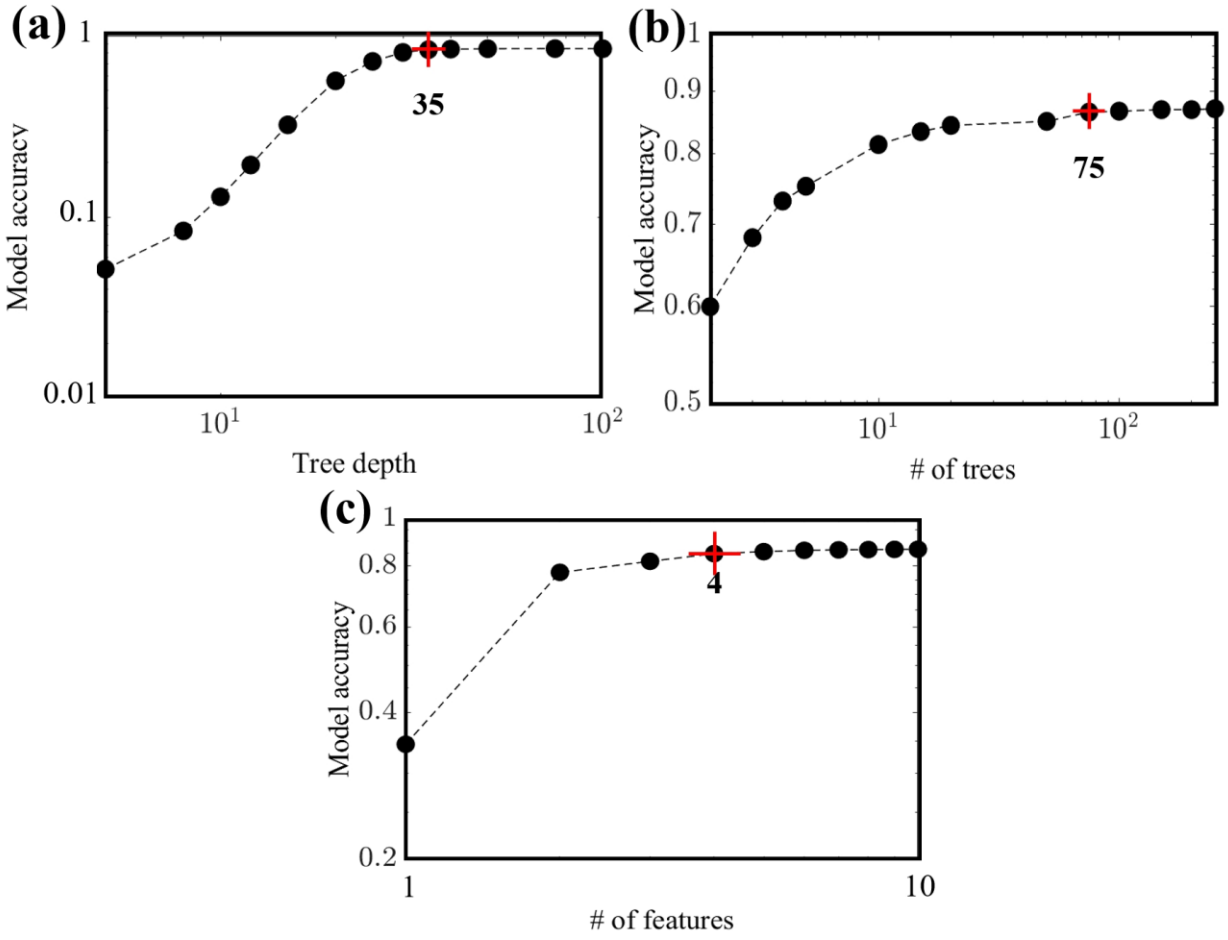


Fig. 14. Accuracy of random forest model as a function of (a) tree depth, (b) the total number of decision trees, and (c) accuracy of SFS as function of number of features selected by 5-fold cross validation. The model reaches a maximum accuracy score of R^2 = 87% and *RMSE*=0.09.

It is shown in the literature that the built-in permutation-based feature importance offered by random forests is prone to bias due to the preferential selection of the features (Mangal and Holm, 2018a). To avoid bias, the tree SHAP explainer, "shap.explainers.Tree()", from the SHAP package (Štrumbelj and Kononenko, 2014) was used to interpret the random forest regression predictions using normalized SHAP values. Fig. 15 shows the relative feature importance in formation of von Mises equivalent stress

and effective plastic strain hotspots for an RVE with uniform texture and equiaxed grains under uniaxial tension, biaxial tension and shear. More results as a function of grain morphology, texture, and mixed deformation states are provided in the supplementary material (Fig. A2 and Fig. A3).

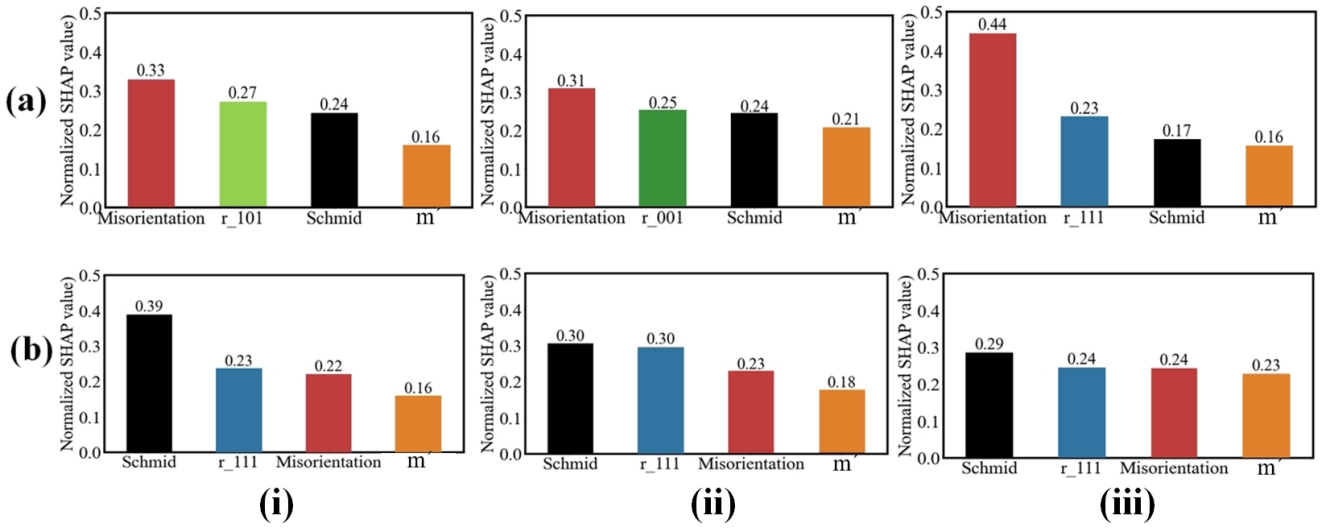


Fig. 15. Relative importance of top features described by normalized SHAP values responsible for the formation of hotspots:(a) von Mises equivalent stress, and (b) effective plastic strain as a function of deformation state: (i) uniaxial tension, (ii) biaxial tension, and (iii) simple shear.

Evaluating the SHAP values for a combination of stress and strain hotspots together, reveals that Schmid factor, intergranular misorientation, m′, and Euclidean distances to IPF corners with [111] being dominant are frequently selected as the top four contributing features in the formation of hotspots with m′ having the least effect on both stress and strain hotspots. Note that crystals with orientation closer to [111] directions exhibit higher Taylor factors and are stiffer grains. It is shown that the deviation of the Taylor factor of a given grain with its first neighbors are important in influencing localizations (Knezevic et al., 2014). Intergranular misorientation is found to be the top feature responsible for stress concentrations while Schmid factor is the top feature responsible for strain localizations. Therefore, regions with high misorientations are candidates for higher stress localization while locations with high

Schmid factors are more prone to micromechanical damage accumulation caused by plastic strain. It is also important to recognize that misorientation effects become more important in stress localizations under pure shear loading (much higher importance for this feature is observed under shear loading in Fig. 15 (a)(iii)). Additionally, Schmid factor becomes more dominant in strain localizations under uniaxial loadings, and its relative importance compared to misorientation and Euclidean distance to [111] directions reduces, with all four top features contributing nearly equally in simple shear. In conclusion, the positive correlation between hotspots and intergranular misorientation, and the fact that intergranular misorientations occur at GBs, suggest that harder grains with the highest Schmid factor and intergranular misorientation, as well as regions near the GBs, are likely locations for damage nucleation and failure.

In summary, we find that the roles of m′ and intergranular misorientations are more important in governing stress/strain localizations than those of grain size, we find that under uniaxial tension, m' and intergranular misorientations are more important in governing stress/strain localizations than grain size, contrary to the findings reported in (Mangal and Holm, 2018b). Intergranular misorientations are also found to be more important than Schmid factor which is another point of contrast. This may be due to the fact that the present work uses a physics-based DD hardening law, while a phenomenological hardening was used in (Mangal and Holm, 2018b), and it has been shown that the local distribution of stresses and strains are strongly dependent on the selected hardening law (Patil et al., 2021) with phenomenological laws obscuring the heterogeneity of spatial distributions and introducing spurious grain-wise homogenizations. Additionally, the present study uses local definitions of grain boundary descriptors that may be obscured if a grain neighborhood averaging scheme is used as in (Mangal and Holm, 2018b).

In addition to the above insights pertaining to uniaxial loading conditions, our work reveals the important influence of non-uniaxial loading conditions on the formation of both stress and strain hotspots. Intergranular misorientations become much more important than Schmid factor in identification of stress hotspots where inelastic deformations under multiaxial loadings, including shear, are more dominant. Schmid factor, intergranular misorientation, m′, and Euclidean distances to IPF corners with [111] all become important under the condition of damage growth and failure resulted by high levels of accumulated plastic strains. Grain size also becomes more important in cases of combined tension/shear loading and high levels of accumulated plastic strains (see supplementary Fig. A4).

4. Summary and conclusions

In this work, we study the role that various microstructural features play in the formation of stress and strain hotspots using a combination of physics-informed CPFFT simulations and ML techniques for data analysis. The microstructures used for simulations varied in grain structure/morphology and crystallographic texture. The intergranular misorientation and slip transmission were quantified locally to describe heterogeneities in the vicinity of grain boundaries. To evaluate the effects of applied deformation state on the formation of hotspots, a range of loading configurations with varying combination of tension and shear were studied. An ensemble learning random forest regression model was used to establish the nonlinear correlations between the microstructural features and hotspots with an accuracy of R^2 = 87% and $RMSE$ = 0.09. To determine the most critical features corresponding to stress and strain hotspots, the sequential feature selection wrapper method was used, with the relative importance of the top selected features quantified by Shapley additive explanations. The present study identifies the microstructural features responsible for stress and strain localizations that will be followed

by micromechanical damage growth and failure. ML analysis of the data obtained from the stress and strain hotspot locations for a wide range of microstructures deformed under complex loading conditions reveals the following findings:

- Schmid factor, intergranular misorientation, m′, and Euclidean distance of a point to the [111] directions were identified as the top four contributing features in the formation of stress and strain hotspots.
- Regions within the microstructure with the following characteristics were identified to be more prone to coincide with micromechanical stress and strain hotspots:
    - Near grain boundaries,
    - Crystals with higher Taylor and Schmid factors,
    - Crystals with high intergranular misorientations.
- Intergranular misorientations are more responsible than Schmid factor in formation of stress hotspots for all deformation states.
- Schmid factors take priority over the intergranular misorientations under high levels of plastic strain localization for all deformation states.
- Grain size becomes important with respect to the top features above only under combined tension and shear and at high levels of accumulated plastic strains.

4 Data availability

Data together with MATLAB and Python scripts supporting the findings of this study are provided in the supplementary materials.

5 Acknowledgements

This research is supported by the U.S. Department of Energy (DOE), National Energy Technology Laboratory (NETL), Award Number DE-FE0031553. RAL acknowledges support from Los Alamos National Laboratory's Laboratory-Directed Research & Development (LDRD) Program.

References

Alizadeh, R., Peña-Ortega, M., Bieler, T.R., LLorca, J., 2020. A criterion for slip transfer at grain boundaries in Al. Scr Mater 178, 408–412.

Ardabili, S., Mosavi, A., Várkonyi-Kóczy, A.R., 2019. Advances in machine learning modeling reviewing hybrid and ensemble methods, in: International Conference on Global Research and Education. Springer, pp. 215–227.

Ardeljan, M., Knezevic, M., Nizolek, T., Beyerlein, I.J., Mara, N.A., Pollock, T.M., 2015. A study of microstructure-driven strain localizations in two-phase polycrystalline HCP/BCC composites using a multi-scale model. Int J Plast 74, 35–57.

Asaro, R.J., 1983. Crystal plasticity.

Bachmann, F., Hielscher, R., Schaeben, H., 2010. Texture analysis with MTEX–free and open source software toolbox, in: Solid State Phenomena. Trans Tech Publ, pp. 63–68.

Batra, R., 2021. Accurate machine learning in materials science facilitated by using diverse data sources.

Bayerschen, E., McBride, A.T., Reddy, B.D., Böhlke, T., 2016. Review on slip transmission criteria in experiments and crystal plasticity models. J Mater Sci 51, 2243–2258.

Beladi, H., Ghaderi, A., Rohrer, G.S., 2020. Five-parameter grain boundary characterisation of randomly textured AZ31 Mg alloy. Philosophical Magazine 100, 456–466.

Beyerlein, I.J., Tomé, C.N., 2008. A dislocation-based constitutive law for pure Zr including temperature effects. Int J Plast 24, 867–895. https://doi.org/10.1016/J.IJPLAS.2007.07.017

Bieler, T.R., Eisenlohr, P., Roters, F., Kumar, D., Mason, D.E., Crimp, M.A., Raabe, D., 2009. The role of heterogeneous deformation on damage nucleation at grain boundaries in single phase metals. Int J Plast 25, 1655–1683.

Breiman, L., 2001. Random forests. Mach Learn 45, 5–32.

Breiman, L., 1996. Bagging predictors. Mach Learn 24, 123–140.

Bunge, H.-J., 2013. Texture analysis in materials science: mathematical methods. Elsevier.

Butrón-Guillén, M.P., Jonas, J.J., Ray, R.K., 1994. Effect of austenite pancaking on texture formation in a plain carbon and a Nb microalloyed steel. Acta metallurgica et materialia 42, 3615–3627.

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

Chaboche, J.L., 2003. 2.04 - Damage Mechanics, in: Milne, I., Ritchie, R.O., Karihaloo, B. (Eds.), Comprehensive Structural Integrity. Pergamon, Oxford, pp. 213–284. https://doi.org/https://doi.org/10.1016/B0-08-043749-4/02085-1

Chen, G., Chen, J., 2015. A novel wrapper method for feature selection and its applications. Neurocomputing 159, 219–226.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Donegan, S.P., Kumar, N., Groeber, M.A., 2019. Associating local microstructure with predicted thermally-induced stress hotspots using convolutional neural networks. Mater Charact 158, 109960.

Eghtesad, A., Knezevic, M., 2021a. A full-field crystal plasticity model including the effects of precipitates: Application to monotonic, load reversal, and low-cycle fatigue behavior of Inconel 718. Materials Science and Engineering: A 803, 140478. https://doi.org/10.1016/J.MSEA.2020.140478

Eghtesad, A., Knezevic, M., 2021b. A full-field crystal plasticity model including the effects of precipitates: Application to monotonic, load reversal, and low-cycle fatigue behavior of Inconel

718. Materials Science and Engineering: A 803, 140478. https://doi.org/10.1016/J.MSEA.2020.140478

Eghtesad, A., Knezevic, M., 2021c. A full-field crystal plasticity model including the effects of precipitates: Application to monotonic, load reversal, and low-cycle fatigue behavior of Inconel 718. Materials Science and Engineering: A 803, 140478.

Eghtesad, A., Knezevic, M., 2020a. High-performance full-field crystal plasticity with dislocation-based hardening and slip system back-stress laws: Application to modeling deformation of dual-phase steels. J Mech Phys Solids 134, 103750. https://doi.org/10.1016/J.JMPS.2019.103750

Eghtesad, A., Knezevic, M., 2020b. High-performance full-field crystal plasticity with dislocation-based hardening and slip system back-stress laws: Application to modeling deformation of dual-phase steels. J Mech Phys Solids 134, 103750. https://doi.org/10.1016/J.JMPS.2019.103750

Eghtesad, A., Shimanek, J.D., Shun-Li, S., Lebensohn, R., Knezevic, M., Liu, Z.-K., Beese, A.M., 2022. Density functional theory-informed dislocation density hardening within crystal plasticity: Application to modeling deformation of Ni polycrystals. Comput Mater Sci.

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

El Aboudi, N., Benhlima, L., 2016. Review on wrapper feature selection approaches, in: 2016 International Conference on Engineering & MIS (ICEMIS). IEEE, pp. 1–5.

Frank, F.C., 1965. On Miller–Bravais indices and four-dimensional vectors. Acta Crystallogr 18, 862–866.

Fushiki, T., 2011. Estimation of prediction error by using K-fold cross-validation. Stat Comput 21, 137–146.

Gao, C., Min, X., Fang, M., Tao, T., Zheng, X., Liu, Y., Wu, X., Huang, Z., 2022. Innovative materials science via machine learning. Adv Funct Mater 32, 2108044.

Gómez-Ramírez, J., Ávila-Villanueva, M., Fernández-Blázquez, M.Á., 2020. Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. Sci Rep 10, 1–15.

Gottstein, G., Shvindlerman, L.S., 2006. Grain boundary junction engineering. Scr Mater 54, 1065–1070.

Gottstein, G., Shvindlerman, L.S., 2005. A novel concept to determine the mobility of grain boundary quadruple junctions. Scr Mater 52, 863–866.

Gottstein, G., Shvindlerman, L.S., Zhao, B., 2010. Thermodynamics and kinetics of grain boundary triple junctions in metals: Recent developments. Scr Mater 62, 914–917.

Groeber, M.A., Jackson, M.A., 2014a. DREAM.3D: A Digital Representation Environment for the Analysis of Microstructure in 3D. Integr Mater Manuf Innov 3, 56–72. https://doi.org/10.1186/2193-9772-3-5

Groeber, M.A., Jackson, M.A., 2014b. DREAM. 3D: a digital representation environment for the analysis of microstructure in 3D. Integr Mater Manuf Innov 3, 56–72.

Hall, M.A., Smith, L.A., 1999. Feature selection for machine learning: comparing a correlation-based filter approach to the wrapper., in: FLAIRS Conference. pp. 235–239.

Ho, T.K., 1995. Random decision forests, in: Proceedings of 3rd International Conference on Document Analysis and Recognition. IEEE, pp. 278–282.

Jebli, I., Belouadha, F.-Z., Kabbaj, M.I., Tilioua, A., 2021. Prediction of solar energy guided by pearson correlation using machine learning. Energy 224, 120109.

Jonas, J.J., 2009. Transformation textures associated with steel processing, in: Microstructure and Texture in Steels. Springer, pp. 3–17.

Kitayama, K., Tomé, C.N., Rauch, E.F., Gracio, J.J., Barlat, F., 2013. A crystallographic dislocation model for describing hardening of polycrystals during strain path changes. Application to low carbon steels. Int J Plast 46, 54–69. https://doi.org/10.1016/J.IJPLAS.2012.09.004

Knezevic, M., Drach, B., Ardeljan, M., Beyerlein, I.J., 2014. Three dimensional predictions of grain scale plasticity and grain boundaries using crystal plasticity finite element models. Comput Methods Appl Mech Eng 277, 239–259.

Kocks, U.F., Tomé, C.N., Wenk, H.-R., 1998. Texture and anisotropy: preferred orientations in polycrystals and their effect on materials properties. Cambridge university press.

Krawczyk, B., Minku, L.L., Gama, J., Stefanowski, J., Woźniak, M., 2017. Ensemble learning for data stream analysis: A survey. Information Fusion 37, 132–156.

Lebensohn, R.A., Kanjarla, A.K., Eisenlohr, P., 2012a. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. Int J Plast 32, 59–69.

Lebensohn, R.A., Kanjarla, A.K., Eisenlohr, P., 2012b. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. Int J Plast 32, 59–69.

Lebensohn, R.A., Rollett, A.D., 2020. Spectral methods for full-field micromechanical modelling of polycrystalline materials. Comput Mater Sci 173, 109336. https://doi.org/https://doi.org/10.1016/j.commatsci.2019.109336

LeDell, E.E., 2015. Scalable ensemble learning and computationally efficient variance estimation. University of California, Berkeley.

Lipovetsky, S., Conklin, M., 2001. Analysis of regression in game theory approach. Appl Stoch Models Bus Ind 17, 319–330.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30.

Luster, J., Morris, M.A., 1995. Compatibility of deformation in two-phase Ti-Al alloys: Dependence on microstructure and orientation relationships. Metallurgical and Materials Transactions A 26, 1745–1756.

Mangal, A., Holm, E.A., 2018a. Applied machine learning to predict stress hotspots I: Face centered cubic materials. Int J Plast 111, 122–134.

Mangal, A., Holm, E.A., 2018b. Applied machine learning to predict stress hotspots I: Face centered cubic materials. Int J Plast 111, 122–134.

Merrick, L., Taly, A., 2020. The explanation game: Explaining machine learning models using shapley values, in: International Cross-Domain Conference for Machine Learning and Knowledge Extraction. Springer, pp. 17–38.

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, 1–10.

Neumann, P., 1991. Representation of orientations of symmetrical objects by Rodrigues vectors. Textures and Microstructures 14, 53–58.

Orme, A.D., Chelladurai, I., Rampton, T.M., Fullwood, D.T., Khosravani, A., Miles, M.P., Mishra, R.K., 2016. Insights into twinning in Mg AZ31: A combined EBSD and machine learning study. Comput Mater Sci 124, 353–363.

Pandey, A., Pokharel, R., 2021. Machine learning based surrogate modeling approach for mapping crystal deformation in three dimensions. Scr Mater 193, 1–5.

Patil, C.S., Chakraborty, S., Niezgoda, S.R., 2021. Comparison of full field predictions of crystal plasticity simulations using the Voce and the dislocation density based hardening laws. Int J Plast 147, 103099.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python. the Journal of machine Learning research 12, 2825–2830.

Pilania, G., 2021. Machine learning in materials science: From explainable predictions to autonomous design. Comput Mater Sci 193, 110360.

Qin, S., Beese, A.M., 2021. Identification of stress state dependent fracture micromechanisms in DP600 through representative volume element modeling. Int J Mech Sci 194, 106209.

Ranstam, J., Cook, J.A., 2018. LASSO regression. Journal of British Surgery 105, 1348.

Refaeilzadeh, P., Tang, L., Liu, H., 2009. Cross-validation. Encyclopedia of database systems 5, 532–538.

Ribeiro, M.T., Singh, S., Guestrin, C., 2016. " Why should i trust you?" Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1135–1144.

Rios, P.R., Glicksman, M.E., 2015. Grain boundary, triple junction and quadruple point mobility controlled normal grain growth. Philosophical Magazine 95, 2092–2127.

Riyad, I.A., Feather, W.G., Vasilev, E., Lebensohn, R.A., McWilliams, B.A., Pilchak, A.L., Knezevic, M., 2021. Modeling the role of local crystallographic correlations in microstructures of Ti-6Al-4V using a correlated structure visco-plastic self-consistent polycrystal plasticity formulation. Acta Mater 203, 116502.

Rodrigues, J.F., Florea, L., de Oliveira, M.C.F., Diamond, D., Oliveira, O.N., 2021. Big data and machine learning for materials science. Discov Mater 1, 1–27.

Rodríguez-Pérez, R., Bajorath, J., 2019. Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. J Med Chem 63, 8761–8777.

Rollett, A.D., Lebensohn, R.A., Groeber, M., Choi, Y., Li, J., Rohrer, G.S., 2010a. Stress hot spots in viscoplastic deformation of polycrystals. Model Simul Mat Sci Eng 18, 074005.

Rollett, A.D., Lebensohn, R.A., Groeber, M., Choi, Y., Li, J., Rohrer, G.S., 2010b. Stress hot spots in viscoplastic deformation of polycrystals. Model Simul Mat Sci Eng 18, 74005.

Roters, F., Eisenlohr, P., Bieler, T.R., Raabe, D., 2011. Crystal plasticity finite element methods: in materials science and engineering. John Wiley & Sons.

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, 1152–1211. https://doi.org/10.1016/j.actamat.2009.10.058

Rozemberczki, B., Watson, L., Bayer, P., Yang, H.-T., Kiss, O., Nilsson, S., Sarkar, R., 2022. The Shapley Value in Machine Learning. arXiv preprint arXiv:2202.05594.

Sagi, O., Rokach, L., 2018. Ensemble learning: A survey. Wiley Interdiscip Rev Data Min Knowl Discov 8, e1249.

Sarma, V.S., Eickemeyer, J., Mickel, C., Schultz, L., Holzapfel, B., 2004. On the cold rolling textures in some fcc Ni–W alloys. Materials Science and Engineering: A 380, 30–33.

Schmid, E., Boas, W., 1950. Plasticity of crystals.

Schwarzenbach, D., 2003. Note on Bravais–Miller indices. J Appl Crystallogr 36, 1270–1271.

Shapley, L.S., Kuhn, H., Tucker, A., 1953. Contributions to the Theory of Games. Ann Math Stud 28, 307–317.

Solomatine, D.P., Shrestha, D.L., 2004. AdaBoost. RT: a boosting algorithm for regression problems, in: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541). IEEE, pp. 1163–1168.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 1–11.

Štrumbelj, E., Kononenko, I., 2014. Explaining prediction models and individual predictions with feature contributions. Knowl Inf Syst 41, 647–665.

Takahashi, Y., Miyazawa, K., Mori, M., Ishida, Y., 1985. Quaternion representation of the orientation relationship and its application to grain boundary problems. Trans. Jpn. Inst. Met. 27, 345–352.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

Veasna, K., Feng, Z., Zhang, Q., Knezevic, M., 2023. Machine learning-based multi-objective optimization for efficient identification of crystal plasticity model parameters. Comput Methods Appl Mech Eng 403, 115740.

Wan, V.V.C., Cuddihy, M.A., Jiang, J., Maclachlan, D.W., Dunne, F.P.E., 2016. An HR-EBSD and computational crystal plasticity investigation of microstructural stress distributions and fatigue hotspots in polycrystalline copper. Acta Mater 115, 45–57. https://doi.org/10.1016/J.ACTAMAT.2016.05.033

Wang, A., An, N., Chen, G., Li, L., Alterovitz, G., 2015. Accelerating wrapper-based feature selection with K-nearest-neighbor. Knowl Based Syst 83, 81–91.

Winter, E., 2002. The shapley value. Handbook of game theory with economic applications 3, 2025–2054.

Xiao, Y., Hua, J., Dougherty, E.R., 2007. Quantification of the impact of feature selection on the variance of cross-validation error estimation. EURASIP J Bioinform Syst Biol 2007, 1–11.

Yang, Z., Yu, C.-H., Buehler, M.J., 2021. Deep learning model to predict complex stress and strain fields in hierarchical composites. Sci Adv 7, eabd7416.

Zare, H., 2015. FeaLect: Scores Features for Feature Selection. URL: https://cran. r-project. org/package= FeaLect.

Zecevic, M., Korkolis, Y.P., Kuwabara, T., Knezevic, M., 2016. Dual-phase steel sheets under cyclic tension–compression to large strains: Experiments and crystal plasticity modeling. J Mech Phys Solids 96, 65–87. https://doi.org/https://doi.org/10.1016/j.jmps.2016.07.003

Zhang, C., Ma, Y., 2012. Ensemble machine learning: methods and applications. Springer.

Zhao, B., Gottstein, G., Shvindlerman, L.S., 2011. Triple junction effects in solids. Acta Mater 59, 3510–3518.