Materialia 20 (2021) 101216

Contents lists available at ScienceDirect

Materialia

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

Full Length Article

Bayesian automated weighting of aggregated DFT, MD, and experimental data for candidate thermodynamic models of aluminum with uncertainty quantification

Joshua J. Gabriel[a,*], Noah H. Paulson[a], Thien C. Duong[a], Chandler A. Becker[b], Francesca Tavazza[b], Ursula R. Kattner[b], Marius Stan[a]

[a] Argonne National Laboratory, Applied Materials Division, Lemont, IL, USA
[b] Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA

ARTICLE INFO

Keywords:
Uncertainty quantification
DFT
MD
Aluminum
Thermodynamics
Bayesian
Automated dataset weighting

ABSTRACT

Atomic-scale modeling methods such as density functional theory (DFT) and molecular dynamics (MD) can predict the thermodynamic properties of materials at a lower cost than experimental measurements. However, their regular usage in thermodynamic model construction is hampered by the lack of quantitative agreement with experimental measurements and the lack of uncertainty estimates on the data. To make regular usage of this atomistic simulation data, it is important to assess whether the atomistic simulation datasets, by themselves or in combination with experimental measurements, result in the same physics-informed models best supported by experimental measurements alone. In this work, models of aluminum thermodynamic properties are discussed using three data sources: atomistic calculations (DFT and MD), experiments, and a combination of atomistic calculations and experiments. The study shows that, after ensuring self-consistency in predicting key invariant points, both experimental measurements and atomistic calculations can significantly contribute to an optimal model.

1. Introduction

Atomic-scale methods such as density functional theory (DFT) and molecular dynamics (MD) have emerged as sources of thermodynamic property information, capable of predicting properties of phases that are difficult to measure. In addition, DFT and MD can assist the evaluation of phase stability in temperature and composition regimes where the uncertainty of experimental measurements is high [1,2]. Furthermore, DFT and MD have recently been used to refine thermodynamic models [3]. Despite the support for the usage of DFT and MD as tools to estimate thermodynamic properties such as enthalpy and heat capacity [4–11], confident selection of datasets is hampered because the uncertainties are rarely reported. Some steps have been taken towards quantifying uncertainties on DFT and MD data [12–15], including evaluations of the relative contributions of various types of uncertainty sources [16]. Using aluminum as an example, this study demonstrates that atomistic simulation datasets with estimated uncertainties can be used together with experimental measurements for developing optimal thermodynamic models for the heat capacity and the enthalpy of materials. This approach is of significant consequence, as it allows substantial expansion of the pool of data available for evaluating thermodynamic models or, in case of missing data, it provides a variety of avenues to obtain them, in addition to experiments.

Models of the heat capacity and enthalpy of a given material are defined by thermodynamic principles and parametrized using available datasets for specific materials [17]. Each dataset has its own reported uncertainty resulting from the experimental setup and the specific methodology. In this work, aluminum (Al) is used as an example because it is an important metal for numerous applications including the automotive, aerospace, and manufacturing industries. For example, Al is a critical element in commercially used alloys, either cast or 3D-printed. Thus, evaluating the uncertainty of thermodynamic models of pure aluminum would impact the uncertainty of alloys of aluminum, improving the alloy design process. In this work, an optimization approach making use of Bayesian inference for thermodynamics [18–20] and an automated weighting scheme [21] is used for developing models that include uncertainty. Atomistic simulation datasets are used to complement datasets from experimental measurements by providing more data samples in the liquid phase using MD as well as more heat capacity data samples for the solid phase from three different

* Corresponding author.
E-mail address: jgabriel@anl.gov (J.J. Gabriel).

https://doi.org/10.1016/j.mtla.2021.101216
Received 20 August 2021; Accepted 27 August 2021
Available online 30 August 2021

types of DFT-based approaches. Both atomistic simulation datasets and experimental measurement datasets used in this work include estimated uncertainties.

In the past, thermodynamic models of pure aluminum have made use of calorimetry measurements of the heat capacity [22, , –35] and the enthalpy of aluminum [36], with each dataset differing in the temperature range of measurement and the respective estimated uncertainties. There are many sources of uncertainty in the experimental measurements, including heating and sample preparation methods. A full discussion is outside the scope of this work but can be found for calorimetry in [37].

Past thermodynamic assessments [38] of calorimetry measurements have shown that aluminum reflects the properties of a Debye solid and that the heat capacity has both electronic and lattice contributions. More recently, an assessment of heat capacity data for solid aluminum for the third generation of SGTE (Scientific Group of Thermodata Europe) descriptions was made in the work of Bigdeli et al. [3] They used an Einstein term to model the Gibbs free energy at low temperatures and a polynomial function for high temperatures. Furthermore, Bigdeli et al. derived a heat capacity model from the Gibbs free energy expression and demonstrated the value of atomistic simulation data in evaluating the properties of superheated aluminum that collapses to the liquid state. The atomistic data was obtained via *ab initio* molecular dynamics (AIMD). The inclusion of atomistic simulation data in the model of the superheated solid overcame a kink at the solid to liquid transition that is present in past SGTE models. The past assessments [38] and models [39] of the heat capacity of solid and liquid aluminum had been developed using experimental measurements and did not include atomistic calculations because no atomistic data was available at the time.

The goal of this study is to develop a strategy to obtain the optimal model for the heat capacity and enthalpy of aluminum in its solid and liquid phases by using Bayesian inference with automatic weighting to analyze more datasets than previously considered including experimental and atomistic datasets. Specifically, we show how atomistic simulation data sources with estimated uncertainties can be selected for the purpose of complementing experimental measurements by providing more data samples at temperatures where measurements are sparse or missing. To this end, Section 2 describes the candidate models, the Bayesian inference approach, and the preparation of the experimental and atomistic simulation datasets with uncertainty estimates considered in this study. Section 3 presents the optimized models, developed by considering the datasets in the following sequence: (a) experimental measurements alone, (b) only atomistic simulation data, and (c) a comparable combination of atomistic simulation data and experiments. Finally, Section 4 provides a comparison of the optimal model presented in Section 3 with past and present assessments of aluminum thermodynamic properties. In addition, the impact on developing similar models for other elements and alloys is discussed.

2. Methods

All datasets generated in this work and the code that performs the Bayesian inference are available in the supplementary information online [40]. This section presents the candidate models, the Bayesian inference approach to select optimal models, and the preparation of individual datasets. Section 2.1 describes the candidate models and the Bayesian inference approach to weight the datasets and select an optimal model from the candidates. Section 2.2 presents the selection of data points from the experimental datasets considered in this work, followed by the DFT and MD methods to generate heat capacity and enthalpy data used in this work. The uncertainty in DFT and MD predicted properties can be divided into numerical/statistical uncertainties, parametric uncertainty, and model uncertainty [12,15,16]. In this work, parametric factors (for DFT) and numerical/statistical factors (for MD) are employed to estimate the uncertainties. Uncertainty estimates on the experimental data were made using the Guide to the Expression of Uncertainty in Measurement (GUM) [41]. Further, we note that these are estimates of the uncertainties which are refined in the Bayesian inference approach through automated weighting. Also, it was shown in Paulson et al. [19] that the uncertainty estimate has limited influence on the value of the rescaled uncertainty corresponding to the model with the maximum likelihood.

2.1. Candidate models and application of Bayesian automated weighting

2.1.1. Candidate models

In this work, we develop thermodynamic models with uncertainty quantification for aluminum using both heat capacity and enthalpy datasets. For the solid, two versions of the segmented regression model [42] for the heat capacity and enthalpy are considered as candidates: one using the Debye model (referred as Debye-SR) to capture low temperature (less than room temperature of 298.15 K) effects and the other using the Einstein model (referred as Einstein-SR) to capture low temperature effects on the heat capacity. The segmented regression model was formulated to capture all sources (electronic, phonon, and magnetic) of physical effects on the temperature dependence of the heat capacity. This model has been used recently in thermodynamic modeling to improve model agreement both with heat capacity datasets near 0 K as well as above 298.15 K [19,21,42]. It has the following general form:

$$C_p = C_V^{Low\,T} + C_{p,bcm} + C_{p,mag} \tag{2.1.a.1}$$

where the low temperature electronic and phonon effects on the heat capacity are captured by the heat capacity at constant volume, $C_V^{Low\,T}$, the high temperature effects are captured by the bent-cable model $C_{p,bcm}$, and magnetic effects are captured by the last term $C_{p,mag}$. The corresponding enthalpy model is determined by integrating the heat capacity with respect to temperature and referencing the enthalpy of pure substances as 0 at room temperature (298.15 K):

$$H - H_{298.15} = \int_0^T C_p\, dT \tag{2.1.a.2}$$

Two models are considered in this work for the low temperature effects on the heat capacity, $C_V^{Low\,T}$: the Debye model, and the Einstein model:

$$C_V^{Debye} = 9R\left(\frac{T}{\theta_D}\right)^3 \int_0^{\frac{\theta_D}{T}} \frac{x^4 e^x}{(e^x-1)^2} dx \tag{2.1.a.3}$$

$$C_V^{Einstein} = \frac{3R\left(\frac{\theta_E}{T}\right)^2 e^{\frac{\theta_E}{T}}}{\left(e^{\frac{\theta_E}{T}}-1\right)^2}. \tag{2.1.a.4}$$

where, the parameters for the Bayesian inference are θ_D and θ_E, which are the Debye temperature and Einstein temperature, respectively. R is the gas constant 8.314 J/mol.K. A previously described numerical Simpson's integral over 100 equally spaced temperature points was employed to evaluate the Debye term [19]. The physical difference between the Debye and Einstein model lies in how the heat capacity approaches zero. In the Debye model, the heat capacity approaches zero as a T^3 power law, while in the Einstein model the heat capacity decreases exponentially. The Debye model was formulated using the concept of phonons to describe vibration in solids, while the Einstein model assumed atoms to be harmonic oscillators with a single frequency.

The bent-cable model in Eq. (2.1.a.1) is a polynomial model designed to capture the variation in slope through the higher temperatures, which is exhibited by experimental measurements and calculations and has the following form:

$$C_{p,bcm} = \begin{cases} \beta_1 T, & T < \tau - \gamma \\ \beta_1 T + \beta_2 \frac{(T-\tau+\gamma)^2}{4\gamma}, & \tau-\gamma \le T < \tau+\gamma, \\ \beta_1 T + \beta_2 (T-\gamma), & T > \tau+\gamma \end{cases} \tag{2.1.a.5}$$

where the parameters β_1, β_2, τ, and γ enable the realization of three regions of variable slope with a smooth transition between them. The last term $C_{p,mag}$ captures the magnetic effects on the heat capacity and is neglected in this work because aluminum is non-magnetic. For brevity, the following two equations will be referenced to represent these candidate models for the solid (Debye-SR refers to Eq. (2.1.a.6) and Einstein-SR refers to Eq. (2.1.a.7):

$$C_{p,\ Debye-SR} = C_V^{Debye} + C_{p,bcm} \tag{2.1.a.6}$$

$$C_{p,\ Einstein-SR} = C_V^{Einstein} + C_{p,bcm} \tag{2.1.a.7}$$

For the liquid, the constant heat capacity model (Eq. (2.1.a.8)) is compared with the linear heat capacity model (Eq. (2.1.a.9)):

$$C_p = c_1 \tag{2.1.a.8}$$

$$C_p = c_0 + c_1 \mathrm{T}, \tag{2.1.a.9}$$

The present study aims to prepare the best combination of atomistic simulation dataset sources and experimental measurements to parameterize these thermodynamic models for aluminum.

2.1.2. Automated weighting of datasets

The datasets are automatically weighted using a Bayesian approach that makes use of the Multinest Monte Carlo algorithm [43] with 800 live points. This algorithm was recently applied to explore the model parameter space for thermodynamic models of hafnium [19] and aluminum [21,42]. This approach performs inference for two categories of parameters, namely, model parameters themselves (for example for Debye-SR in Eq. (2.1.a.6): Θ_D, β_1, β_2, τ, γ), and weights for each of the datasets considered for a given model. The weight for a given dataset is set as a ratio of the uncertainty estimate δ to a factor α. This factor enters the Bayesian inference as a Bayesian hyperparameter ϵ, and determines the weight assigned to the dataset:

$$\epsilon = \frac{\delta}{\alpha} \tag{2.1.b.1}$$

This factor α determines whether a given dataset's original uncertainty is scaled-up ($\alpha < 1$) or scaled-down ($\alpha > 1$). Scaled-up means that the uncertainty estimate δ was underestimated with respect to the considered model, while scaled-down means that the δ was overestimated. The level of weighting on a dataset is discussed in terms of the magnitude of the rescaled uncertainty. In other words, the higher the rescaled uncertainty, the less a dataset is weighted in the model. The prior distribution for the factor α is the exponential distribution with a mean of 1. For each model parameter, the prior distribution is a broad uniform distribution with a physically reasonable range informed by intuition (for example, the Debye temperature and Einstein temperature are set as positive quantities less than 700 K, and the constant term for the liquid heat capacity is set between 0 and 100) for each parameter as described in Paulson et al. [21] and the supplemental information [40]. The likelihood function for a given parameter set is computed for each dataset using two Student's t-distributions with $\nu = 2$ degrees of freedom, one for the heat capacity error ΔC_p and other for enthalpy error ΔH, with each centered at zero, scaled to the Bayesian hyperparameter of each dataset. The joint probability density for the likelihood function (probability of the data given the parameters $P(D|\Theta)$) is given as:

$$P(D|\theta) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{\left(\frac{\Delta C_p}{\epsilon}\right)^2}{\nu}\right)^{-\frac{\nu+1}{2}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{\left(\frac{\Delta H}{\epsilon}\right)^2}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{2.1.b.2}$$

The model evidence is computed as the logarithm of the marginal likelihood function, and models are compared using the Bayes' factor (expressed as the ratio of the marginal likelihood function of two competing model forms) as discussed in detail in Paulson *et al* [19,21,40]. To reproduce the findings of this work, the reader is referred to the supplementary datasets and codebase available online [40].

2.2. Thermodynamic property datasets

2.2.1. Preparation of datasets

In this paper, the data sources are referred to in the format of LNAyyyy (first three letters of the last name of the first author followed by the four-digit year of publication). Experimental data sources with at least 10 data points in the solid phase or 6 data points in the liquid phase are chosen in this work so that as many sources of data with comparable amounts of data are considered. It is of interest to understand the statistical contribution of each data source. Of these experimental measurement data sources, heat capacity for temperatures lower than 298 K is available from MAI1934 [22], KOK1937 [23], GIA1941 [24], HOP1962 [25], BER1968 [26], DOW1980 [27], POC1953 [28], ROR1962 [29], and PAR1961 [30]. Above 298 K, heat capacity data is available from HIR1955_1 and HIR1955_2 [31], SCH1970 [32], KRA1972 [33], ZOL1990 [34], and EAS1924 [35]. In particular, SCH1970 [32] and KRA1972 [33] report heat capacity data for liquid aluminum in addition to solid aluminum. MCD1967 [36] is a source of enthalpy data for both solid and liquid aluminum from calorimetry measurements in a titanium diboride crucible from 366 K up to 1647 K. McDonald [36] shows that data measured above 600 K has larger uncertainty than data measured at room temperature because of the partial reaction of aluminum with the crucible.

Each experimental dataset differs from the other in the number of points available and since each data point is measured in the same way, correlations exist within a dataset. To reduce the differing degrees of within-dataset correlation, 10 data points are selected from each dataset for the solid phase in the following manner: (i) if the dataset spans the entire temperature range for the solid (0 K to 933.15 K), one data point is selected at random from each 100 K interval of temperature from 0 K to 933.5 K, (ii) otherwise, 10 data points are chosen at random from the available temperature range of the dataset. Similarly, 6 data points are selected for each dataset for the liquid from a temperature of 933.5 K to 1600 K. Applying this approach, the datasets EAS1924, POC1953, and PAR1961 are not considered in this work because they have fewer than 7 data points each. ROR1962 and HOP1962 are datasets providing heat capacity data in the range of 1 K to 1.2 K. ROR1962 provides data points over 1.17 K to 1.19 K, while HOP1962 provides data over 1.08 K to 1.19 K. With the larger range and to avoid redundancy, HOP1962 is considered in this work and ROR1962 is not considered.

Melting points determined by atomistic simulation methods (MD) and experimental measurements show systematic discrepancies. To account for this discrepancy, we propose that a single melting point is considered across all datasets. To account for the difference in melting point between experimental datasets and MD datasets, data is excluded from the interval $T_m^E < T < T_m^P$ if $T_m^E < T_m^P$ and from $T_m^E > T > T_m^P$ if $T_m^E > T_m^P$, where T_m^E and T_m^P are the melting points according to experiment (T_m^E = 933.5 K) and the MD dataset, respectively. This approach avoids the inclusion of data, at a given temperature, which is solid according to the interatomic potential and liquid according to experimental measurements, or vice versa. Past CALculation of PHAse Diagrams (CALPHAD) assessments show that certainty in the invariant points is critical for a reliable assessment [2].

2.2.2. DFT generated heat capacity data

For the DFT data sources, a heat capacity dataset for solid aluminum is generated in this work (and labeled as GAB2021) and also extracted from literature (datasets labeled as GUA2019 [13], GRA2009 [44]).

For the DFT calculations in this work, the Vienna *ab initio* Simulations Package (VASP)[1] [45] and its projector augmented wave method [46] pseudopotential (version 5.4) of aluminum in the $3s^2\ 3p^1$ configuration was employed. The plane wave cutoff of 600 eV and a Γ - centered Monkhorst and Pack [47] mesh corresponding to a *k*-points density of 8000 per reciprocal atom were used. The 2 × 2 × 2 supercell of the 4 atom face-centered cubic (*fcc*) conventional cell of aluminum was used as a basis for density functional perturbation calculations to calculate the quasi-harmonic approximation (QHA) to the free energy as implemented in Phonopy [48]. A volume range of -5 to 5 % with respect to the equilibrium volume was used to calculate the phonon density of states, and hence derive the heat capacity C_p (T) under the QHA up to the melting point of 933.5 K. To obtain the equilibrium volume, energy minimization was performed before the density functional perturbation calculations, wherein the total energies were converged to 10^{-6} eV and forces to 0.01 eV/Å. The uncertainty with the DFT prediction for C_p (T) was estimated using the effect of the selected exchange correlation, which has been shown in the literature to be a key source of uncertainty for DFT computed properties [12,13] among others such as calculation convergence parameters and approximations to the electronic structure in form of pseudopotentials. Since heat capacity at constant pressure is a first derivative of the enthalpy, the epistemic uncertainty due to our choice of convergence parameter (*k*-point density) is considered to be less than 1% or on the order of 0.1 J/mol.K [49]. Duong et al. [50] showed that our choice of supercell size also causes an uncertainty on the order of 0.1 J/mol.K. Further, uncertainty in the DFT prediction can be a result of systematic bias due to approximations made by choosing a single exchange correlation functional. To minimize this bias, we employ the observation in literature that shows that local density approximation (LDA) and generalized gradient approximation (GGA) functionals are known to systematically underestimate and overestimate structural properties (such as atomic distances and cohesive energies) in metals such as Al [51,52]. This explains the finding in the literature that shows that the LDA and GGA bound experimental heat capacity data [44,53] for metals such as Cu, Ni, and Al. We note this approach of using the LDA and GGA to minimize the systematic bias is not universal to all materials, especially when both LDA and GGA overestimate or underestimate structural properties. We note that the approach of Guan et al. [13] to use the Bayesian Error Estimation Functional (BEEF) is attractive for materials when sources of systematic bias due to choice of any one functional is unknown. Using this approach to estimate uncertainty, the predicted value of C_p at a given temperature, $C_{p,DFT}(T)$, is estimated as the average of $C_{p,GGA}(T)$, *i.e.*, the DFT predicted value with the GGA of Perdew Burke Ernzerhof (PBE) [54], and $C_{p,LDA}(T)$, *i.e.*, the DFT predicted value with LDA parametrized by Perdew and Zunger [55]. The associated uncertainty $\sigma_{C_{p,DFT}}(T)$, is estimated as the absolute difference between $C_{p,GGA}(T)$ and $C_{p,LDA}(T)$:

$$C_{p,DFT}(T) = \frac{1}{2}\left[C_{p,LDA}(T) + C_{p,GGA}(T)\right] \tag{2.2.b.1}$$

$$\sigma_{C_{p,DFT}}(T) = \left| C_{p,GGA}(T) - C_{p,LDA}(T) \right| \tag{2.2.b.2}$$

The DFT data for C_p (T) with uncertainties described above is used as one source of atomistic simulation data and is referred as GAB2021 in this work. In addition, two more DFT datasets [13,44] for C_p (T) are extracted from the literature using the open source WebPlotDigitizer tool [56]: (i) dataset GUA2019 corresponding to Guan et al. [13] that made use of the Bayesian Error Estimation Functional (BEEF) in conjunction with the Debye model to calculate C_p (T) with Bayesian uncertainty estimates; and (ii) GRA2009 corresponding to Grabowski et al. [44] that made use of PBE to calculate C_p (T) from free energy calculations using *ab initio* molecular dynamics. An uncertainty of 0.1 J/mol.K was estimated on GRA2009 based on the tighter DFT input parameters such as k-point density chosen in Grabowski et al [44]. Among the three DFT datasets, GUA2019 shows larger uncertainty estimates than GAB2021 and GRA2009, especially at temperatures more than 600 K.

2.2.3. MD generated enthalpy data

For the MD datasets, 4 interatomic potentials are used to generate enthalpy datasets through MD simulations and considered as candidates because of the large variability in predicted melting points for aluminum [14]. In this work, we use a diverse candidate subset of 4 interatomic potentials as a representative example of use cases where an available interatomic potential is chosen for accuracy in predicting phase transitions and mechanical properties in aluminum and its alloys. We include interatomic potentials developed for alloys as representative of use cases where accuracy in these properties for modeling both aluminum and its alloy phases is desired, such as for developing thermodynamic phase diagrams. With this aim, three candidate Embedded Atom Method (EAM) interatomic potentials [57–59] designed for predicting properties of aluminum and its alloys are chosen from the NIST interatomic potential repository [60] according to the accuracy of their predicted melting points as tabulated in Zhu et al. [14]. To generate a diverse example candidate subset, the accuracy criterion was set as 150 K of the experimentally measured melting point of 933.5 K. Specifically, Sturgeon and Laird [57] was designed to correct for discrepancy in melting point, while Mishin et al. [58] was designed for mechanical properties of aluminum and Mishin [59] was designed for Ni-Al alloys. Additionally, a recently developed machine learning interatomic potential for Al-Mg alloys that predicts the melting point of aluminum as 918 K [61] is considered as representative of a different fitting strategy than the EAM potentials. We note that other potentials predict melting points closer to the experimental value of 933.5 K as tabulated in Zhu et al. [14]. The goal of this study, however, is to demonstrate how to choose from a diverse candidate subset for developing a thermodynamic property model with quantified uncertainty. Further, in this work, we do not judge the quality of any interatomic potential but introduce an approach on how to choose the one from a candidate subset that can be most compatible with experiments to expand the data pool available for optimizing a thermodynamic model.

The enthalpy for both the solid phase and liquid phase are determined using isothermal-isobaric molecular dynamics (NpT-MD) simulations, using Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [62] with a 1 fs timestep. For the EAM potentials, a 6912 atom cell (12 × 12 × 12 supercell of the 4 atom conventional cell) of aluminum is used following the procedure described in Becker et al [9]. For the machine learning force field, a smaller cell (864 atom 6 × 6 × 6 supercell) is considered because of the increased computational cost. Enthalpies of both solid and liquid aluminum are determined, for each of these potentials, over the range of 200 K to 1600 K. For each of the 4 potentials considered in this study, the block averaging method of Flyvbjerg and Petersen is used to estimate the ensemble averaged enthalpy and the associated statistical uncertainty [63,64]. This method estimates the uncertainty as the plateau of the square of the block-averaged uncertainty of the enthalpy over the molecular dynamics trajectory with respect to the reciprocal number of blocks. The plateau corresponded to a height ranging from 1 to 25 J^2/mol^2 for the squared uncertainty of the enthalpy of the solid (computed for temperatures 200 K to 900 K) and 80 to 120 J^2/mol^2 for the liquid (computed for temperatures 1000 K to 1600 K), each over a width of 4 to 8 blocks. This approach resulted in an estimated uncertainty of 1 to 5 J/mol for enthalpy of the solid and 9 to 11 J/mol for the liquid. Further, enthalpies of fusion are calculated and reported in the present study at the melting temperatures for each potential reported in Zhu et al. [14] and Zhang et al. [61].

[1] Commercial products are identified in this paper for reference. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology (NIST), nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Table 1
Optimal model parameters for heat capacity model of solid aluminum expressed as mean ± standard deviation.

Dataset Selection	Θ_D(K)	β_1(x 10^{-3} J/K)	β_2(x 10^{-3} J/K)	τ(K)	γ(K)
Experiments only	390.3 ± 0.9	1 ± 0.002	8 ± 0.4	176.9 ± 17.8	84 ± 44.5
Atomistic only	363.4 ± 5.7	-0.7 ± 0.6	9 ± 0.7	177.3 ± 16.6	102.6 ± 51.5
Aggregated	390.2 ± 0.8	1 ± 0.002	6 ± 0.3	150.9 ± 16.7	50.8 ± 40.1

Table 2
Optimal model parameters for heat capacity of liquid aluminum expressed as mean ± standard deviation.

Dataset Selection	c_1(J/mol.K)
Experiments only	31.2 ± 4.2
Atomistic only	34.7 ± 0.2
Aggregated	31.1 ± 0.2

3. Results

The results of this work are presented as follows: Section 3.1 presents the optimal model derived from experimental measurement datasets alone. Section 3.2 describes the selection of atomistic simulation datasets to best complement the experimental measurement datasets by presenting optimal models derived from atomistic simulation datasets alone. Finally, Section 3.3 presents the optimal model combining experimental measurement datasets and atomistic simulation datasets. In each section, the optimal model for the considered collection of datasets refers to the optimal model form out of the candidate model forms introduced in Section 2.1.1. The parameters for each of the models presented in Sections 3.1 to 3.3 are listed in Table 1 for the solid phase, and in Table 2 for the liquid phase. The optimal model presented in Section 3.3 is further compared to the literature and the similarities and differences are discussed in Section 4. Further, the differences in weighting of datasets will be discussed in Section 4.

3.1. Optimal property model from experimental measurements

Fig. 1 shows the mean heat capacity and enthalpy for solid and liquid aluminum for the optimal model form, when considering experimental data sets alone. The uncertainty on the model is expressed as the Bayesian 95% credible interval bands. The full-width of the Bayesian 95% credible interval increases from 0 J/mol.K to a maximum of 0.7 J/mol.K close to the melting point. Owing to the increased estimated uncertainty of data points beyond the melting point, the full-width maximum of the Bayesian 95% credible interval reaches 1 J/mol.K.

The optimal model form is selected from the candidates introduced in Section 2.1.1 (see Eq. (2.1.a)) for each phase. For the solid, between the Debye-SR (Eq. (2.1.a.6)) and Einstein-SR (Eq. (2.1.a.7)) models, the Bayes' factor (ratio of the marginal likelihood between the Debye-SR and Einstein-SR) is on the order of 1000. This Bayes' factor indicates a strong preference for the Debye-SR model form when considering these experimental measurement datasets. This is an expected result because aluminum is known to be a Debye solid from past assessments for this temperature range [38].

For the liquid, between the constant model (Eq. (2.1.a.8)) and linear model (Eq. (2.1.a.9)), the Bayes' factor (ratio of the marginal likelihood between the constant model and linear model) is 5. This lower Bayes' factor indicates a weaker preference of model form between the linear model and the constant model. Being the model with fewer parameters, the constant model is selected. Past assessments [38] and models [39] have used a constant value to approximate the heat capacity. Certain datasets [32,33] had been ignored in past assessments [38] because of their deviations from the rest of the data. In this work, samples from all these datasets are considered. As will be discussed further in Section 4, these datasets receive lower weights, which is essentially the choice made in past assessments.

3.2. Selection of atomistic simulation datasets

In this subsection, we present the selection of an optimal combination of atomistic simulation datasets for complementing the experimental measurement datasets presented in the previous subsection. Four different combinations of atomistic simulation datasets are considered. In each combination, one enthalpy dataset from MD (STU2000 [57], MIS1999 [58], MIS2004 [59], and ZHA2019 [61]) is combined with all three heat capacity datasets from DFT (GRA2009 [44], GUA2019 [13], and GAB2021 (this work)). For brevity, these models are referred as MIS1999, STU2000, MIS2004, and ZHA2019 according to the respective MD enthalpy datasets considered in each. All three DFT datasets are considered together in one combination because each dataset is the result of: (i) different model assumptions for the underlying DFT method, (ii) different approaches to calculating C_p, and (iii) different parametric uncertainty estimation approaches. GUA2019 was calculated using the Debye model and the Bayesian Error Estimation Functional, while GAB2021 made use of the quasi-harmonic approximation approach, and GRA2009 considers anharmonic and vacancy effects through *ab initio* MD. The classical MD enthalpy datasets on the other hand were computed using the same MD methodology and differ only in the fitting approach to create each underlying interatomic potential. To build an optimal model where atomistic simulation datasets complement experimental measurement datasets, we choose the combination that supports the same optimal model form for both solid and liquid aluminum.

For the solid, the Bayes' factor (ratio of the marginal likelihood between the Debye-SR and Einstein-SR) is on the order of 1000 for all four combinations, indicating a strong preference for the Debye-SR model. This is the same model form preferred by the experimental measurement datasets presented in Section 3.1. For the liquid, however, MIS1999 shows a Bayes' factor of 5 (ratio of the marginal likelihood of linear heat capacity to constant heat capacity), while STU2000, MIS2004, and ZHA2019 show a Bayes' factor of 109, 2440, and 3071, respectively. The low Bayes' factor of 5 for MIS1999 indicates that there is a weak preference for linear heat capacity. Being the simpler model with fewer parameters, the constant heat capacity model can be selected for the MIS1999 dataset. On the other hand, the higher Bayes' factor for STU2000, MIS2004, and ZHA2019 indicate that there is a strong preference for the linear heat capacity.

The difference in preferred model form may be associated with differences in how the liquids are represented by the interatomic potential models. For example, although STU2000 predicts a melting point close to experiment, it was not fit directly to properties of the liquid. Moreover, interatomic potentials may be optimized to reproduce properties of alloys and not only the elements, such as MIS2004 and ZHA2019. Fig. 2 compares the optimal heat capacity and enthalpy model forms for solid and liquid aluminum for the four combinations of atomistic simulation datasets. Comparing Fig. 2 (a, b) with (e, f) shows that the difference between MIS1999 and MIS2004 is large, especially in the liquid phase; MIS2004 was optimized for the Ni-Al alloy properties. On the other hand, comparing Fig. 2 (a, b) with (g, h) shows a smaller dif-

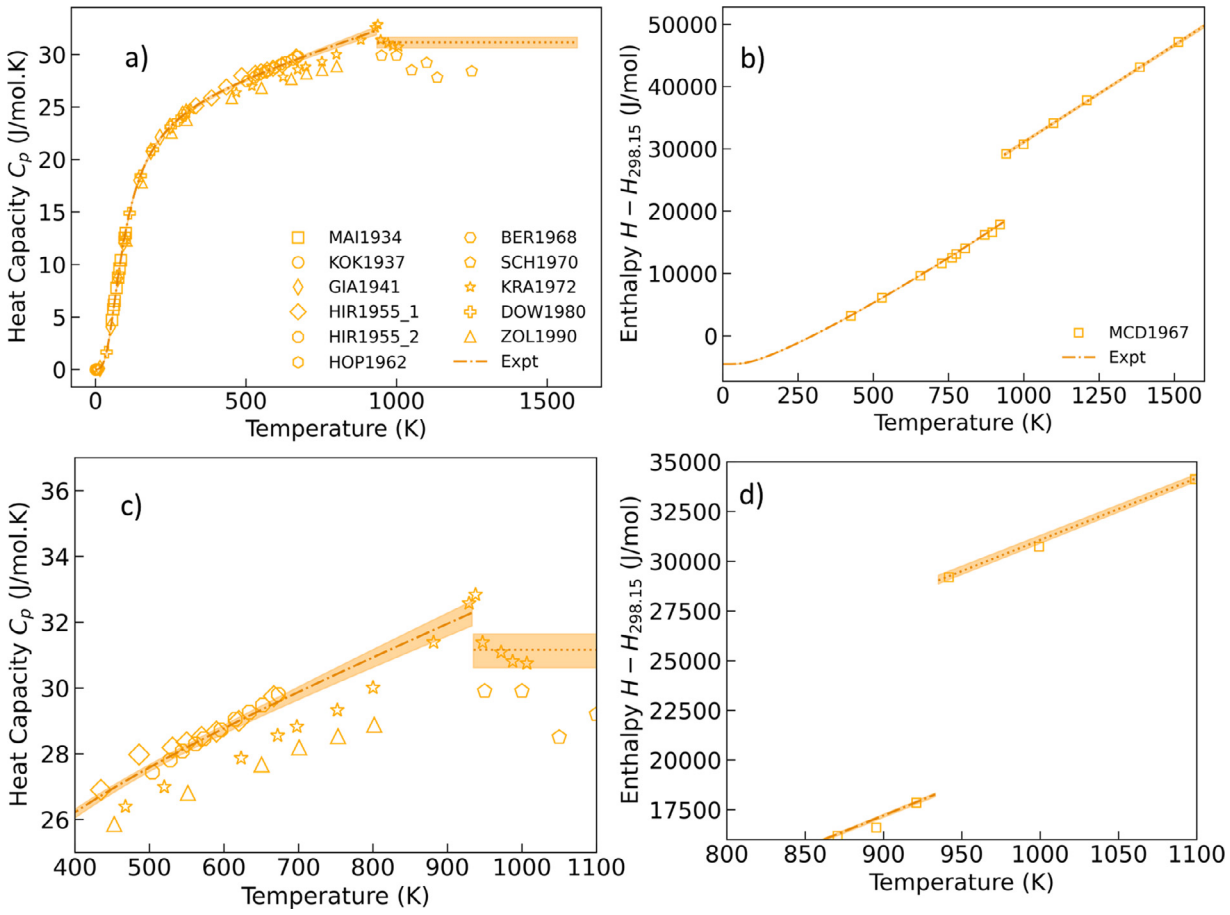

Fig. 1. shows the mean (a) heat capacity (zoomed in view in (c)) and (b) enthalpy (zoomed in view in (d) for solid (orange dashed dotted line) and liquid aluminum (orange, dotted line) when considering experimental data sets alone. The Bayesian 95% credible interval bands are also shown and increase with increasing temperaturate.

Table 3
Heat of fusion and melting temperature of aluminum from atomistic simulations using different interatomic potentials.

Interatomic Potential	Heat of Fusion (kJ/mol)	Melting Temperature (K)
MIS1999 [57]	9.76	1043.0
STU2000 [58]	9.33	932.5
MIS2004 [59]	9.01	871.0
ZHA2019 [61]	10.2	918.0

ference between MIS1999 and ZHA2019. ZHA2019 was fitted to some liquid phase configurations.

Fig. 3 additionally shows all atomistic simulation datasets considered in this work for each of the DFT computed heat capacity and the MD computed enthalpies. The major difference between various MD datasets resides in the published melting points [14] as tabulated in Table 3. Enthalpies of fusion are calculated and reported in the present study. Despite the discrepancy in melting point and differences in preferred model form, Fig. 3 shows that the different MD datasets agree with each other. This can be explained by the discrepancies in both the enthalpy of fusion and the melting point, partially canceling each other out. While the accuracy in melting point (in general phase transition point) is a reasonable criterion to select an interatomic potential for modeling the liquid phase (in general the phase after phase transition point on heating), it is not a sufficient criterion and additional checks are recommended. For material systems where the model form for the phase is well established, such as aluminum, checking whether the atomistic simulations support the same model form as the experimental measurement datasets is recommended before making use of the interatomic potential to model the phase. Hence, the model derived based on MIS1999 is selected to complement the experimental measurements from Section 3.1 towards building an aggregated model that considers both experimental measurement datasets and atomistic simulation datasets.

3.3. *Complementing the experimentally derived model with atomistic simulation datasets*

Fig. 4 presents the optimal model when the atomistic simulation datasets GUA2019, GAB2021, GRA2009, and MIS1999 are used together with the experimental measurements described in Section 3.1. This model is referred henceforth as the aggregated model. Further, Fig. 4 compares the aggregated model (green) with the models presented in Section 3.1 (orange, referred as experimental in text, and Expt. in the figure) and 3.2 (blue, referred as atomistic in text, and Atom. in the figure). Each model has an associated uncertainty described by the Bayesian 95% credible interval. The full-width of the Bayesian 95% credible interval is at most 1 J/mol.K for the experimental model, and 0.9 J/mol.K for the atomistic model and the aggregated model. The three models agree within the Bayesian 95% credible interval up to 400 K and then the aggregated model and atomistic model deviate by up to 3 J/mol.K from the experimental model presented in Section 3.1. Compared with the experimental model, the Bayesian 95% credible interval of the aggregated model is smaller, indicating a reduction in the overall model uncertainty. Furthermore, for the solid, the Debye-SR model form is once again preferred with a Bayes' factor on the order of 1000

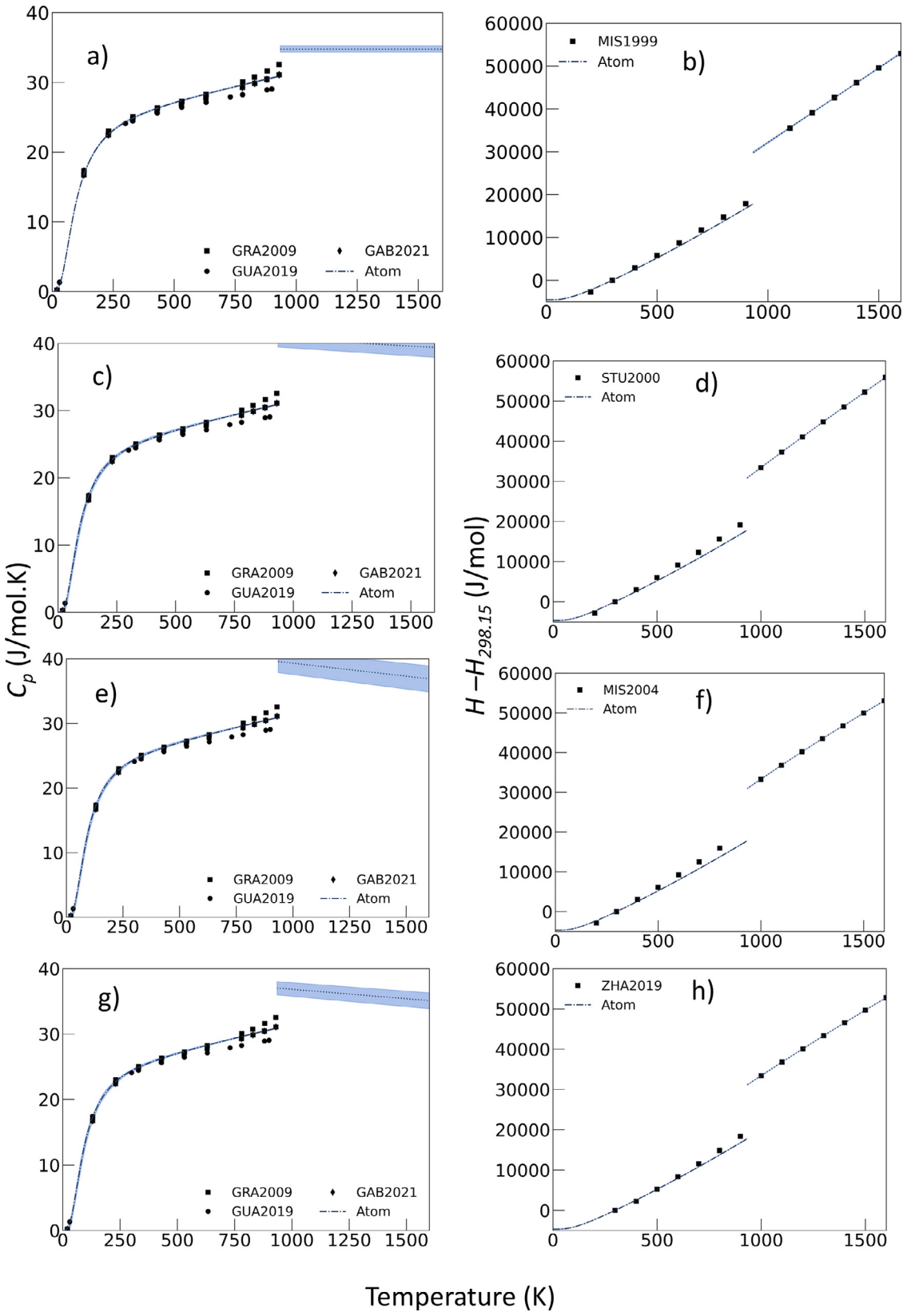

Fig. 2. shows the mean heat capacity model (left column: a, c, e, g) and enthalpy model (right column: b, d, f, h) for solid (blue, dashed dotted line) and liquid (blue, dotted line) aluminum when considering DFT heat capacity datasets with one MD dataset each, i.e, (a,b) MIS1999, (c,d) STU2000, (e, f) MIS2004, and (g,h) ZHA2019. The constant heat capacity model is preferred by the model developed using MIS1999.

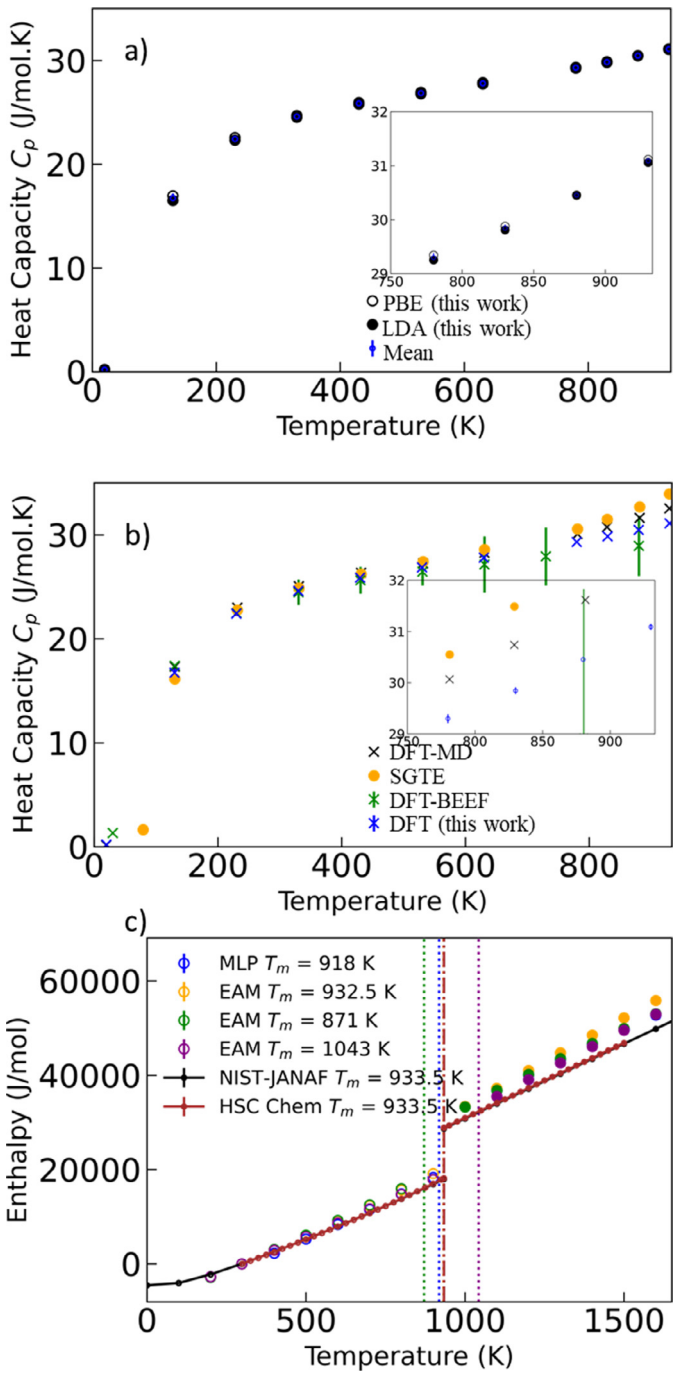

Fig. 3. Data with uncertainty from atomistic simulations in this work and others from literature. a) Estimation of uncertainty in Heat Capacity (C_p) from DFT PBE and LDA calculations (DFT-this work, is dataset GAB2021 in blue) compared with b) SGTE function (in orange, circles) and C_p from other DFT work: DFT-BEEF (is dataset GUA2019 in green) and DFT-MD (is dataset GRA2009 in black). Uncertainty estimates range from 0.5 J/mol.K to 2 J/mol.K. c) Enthalpy (H) derived from MD simulations, with 4 different interatomic potentials compared with NIST-JANAF (black) [65], and an assessment by HSC Chemistry (red). Vertical lines represent the melting points (T_a^{mp}) predicted by each potential or model. Experimental melting point is T_m^E = 933.5 K. To account for this difference in melting point, data is excluded from the interval $T_m^E < T < T_m^P$ if $T_m^E < T_m^P$ and from $T_m^E > T > T_m^P$ if $T_m^E > T_m^P$. Uncertainty estimates are between 1 and 10 J/mol.

with respect to the Einstein-SR model. For the liquid, the Bayes' factor (ratio of marginal likelihood of the linear heat capacity to the constant heat capacity) is approximately unity. This once again indicates a weak preference to linear heat capacity, and hence, being the simpler model, the constant heat capacity is selected.

Tables 2 and 3 show the differences in model parameters for the solid and liquid phase, respectively. For the solid, the Debye temperature is predicted to be 390 K when the model is developed from only experimental data or a combination of experimental data and atomistic data. This is in good agreement with a Debye temperature of 390 K at 298 K, which was obtained from the derivation of experimental thermal conductivity data [66]. The Debye temperature is predicted to be 366 K when only atomistic datasets are included. The underestimation of the Debye temperature can be explained by the difference in phonon density of states between what can be measured and what is calculated. In the liquid phase, the atomistic model predicts a heat capacity higher than the aggregated model by 2 J/mol.K.

Fig. 5 compares the original uncertainty estimate and rescaled uncertainty on each dataset in terms of the mean uncertainty and the associated Bayesian 95% credible interval for the experimental model (orange), atomistic model (blue), and for the aggregated model (green) described in this section. A lower rescaled uncertainty indicates a higher weight for the dataset, in each model. A complete tabulation of the Bayesian rescaling factor in addition to the rescaled uncertainties shown in Fig. 5 (their respective mean, 2.5th and 97.5th percentiles) is available in the supplemental spreadsheet under the sub-sheets "Cp_solid", "Cp_liquid", "H_solid", and "H_liquid" for the respective parts of each model. Among heat capacity datasets in the low temperature region of less than 10 K, the datasets HOP1962, KOK1937, and BER1968 have estimated uncertainties on the order of 10^{-5} J/mol.K and hence do not show on the plotting scale of Fig. 5. Upon rescaling the uncertainty, the supplementary spreadsheet "Cp_solid" shows that HOP1962 is weighted the most among HOP1962, KOK1937, and BER1968 in both the aggregated model as well as the experimental model. In the higher temperature region for the solid, that is, greater than 10 K, the supplementary spreadsheet shows that DOW1980 has the lowest rescaled uncertainty for both the experimental and aggregated models and is hence weighted the most. For the solid phase, all heat capacity experimental measurement datasets show rescaled uncertainty less than 1 J/mol.K. All the atomistic simulation heat capacity datasets, that is, GUA2019, GRA2009 and GAB2021 also show rescaled uncertainties of less than 1 J/mol.K. For the liquid phase, SCH1970 shows the highest rescaled uncertainty of 2 J/mol.K and is hence weighted the least. Among the enthalpy datasets, the mean rescaled uncertainty of the atomistic simulation dataset MIS1999 in the aggregated model is 2000 J/mol in the liquid phase, and 500 J/mol in the solid phase. The experimental measurement enthalpy dataset MCD1967 shows a maximum rescaled uncertainty of 100 J/mol in the aggregated model. Together, the supplementary spreadsheets and Fig. 5 show that although the atomistic simulation datasets are initially underestimated in their uncertainties, their rescaled uncertainties are within one order of magnitude as the experimental measurement datasets.

4. Discussion

The aggregated model derived in Section 3 is now compared in Fig. 6 with models published by the past assessment by Desai [38], an assessment generated by a commercial assessment software HSC Chemistry [67], and the 1991 SGTE function [39]. HSC Chemistry makes use of the thermodynamic databases and optimizes thermodynamic property models. For the case of aluminum, the HSC evaluation is based on the thermodynamic data for the solid from Burcat and Ruscic [68] and the liquid data from Burcat and Ruscic [68], Knacke et al.[69], and Belov et al.[70] The 1991 SGTE function was adopted in many commercial thermodynamic databases such as those distributed by Thermo-Calc [71]. The model developed in this work agrees within 3 J/mol.K over the

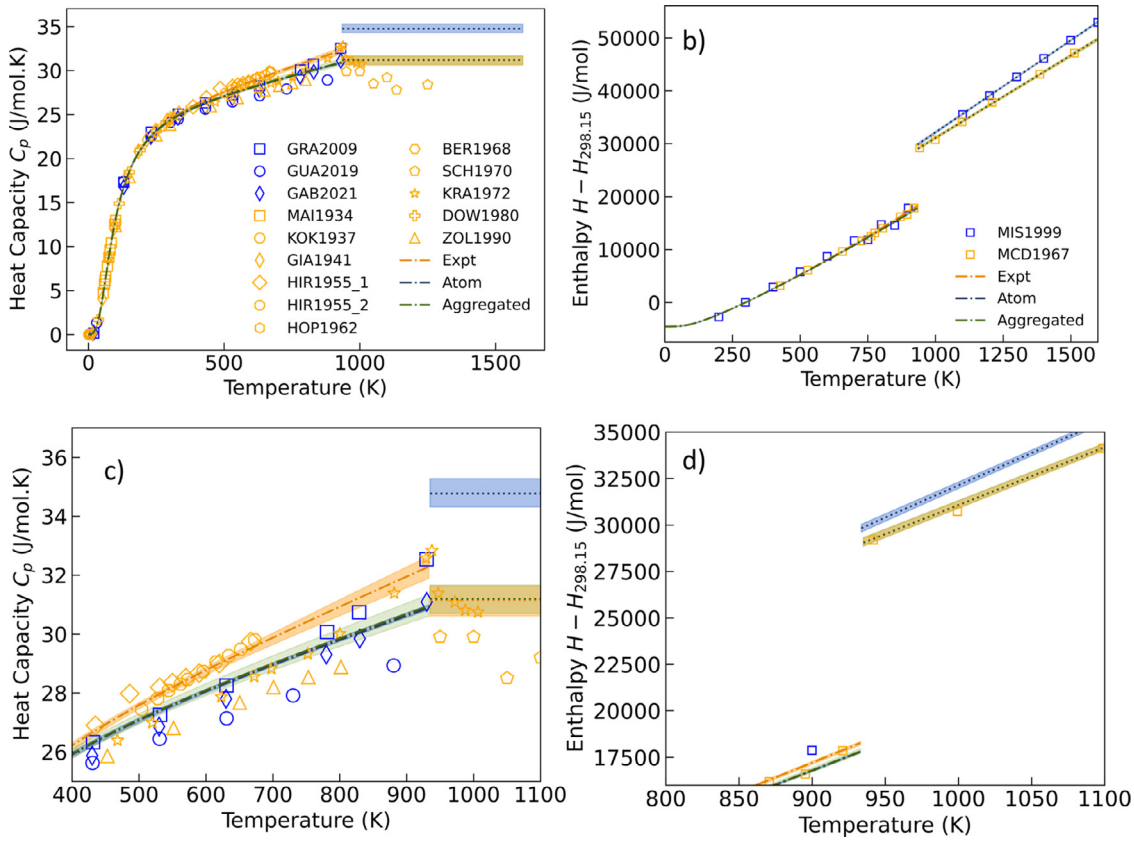

Fig. 4. Comparison of (a) Heat capacity (with zoomed view in c)) and corresponding (b) Enthalpy (with zoomed view in d)) for the aggregated model (green), atomistic model (blue), and experimental model (orange) for solid (dashed dotted line) and liquid (dotted line). Agreement between the three models within 95% CI is up to 400 K for Cp, while agreement within 95% CI for the enthalpy agreement is upto 600 K.

temperature range of 300 K to 1600 K with all these considered assessments. Both the models from HSC chemistry and the SGTE function are plotted from 300 K and above because the model form is not meant to extrapolate to 0 K. The model in this work deviates more than 2 J/mol.K above 600 K from the assessment by Desai and HSC Chemistry, but is within 1 J/mol.K of the SGTE function across the complete temperature range considered. We attribute these differences to two factors: first, the increased number of data sources considered in this work, *i.e.*, inclusion of DFT and MD data, compared to Desai, HSC Chemistry, and SGTE; and second, the increased uncertainty in heat capacity measurements above 600 K. The maximum deviation near the melting point is likely because the melting point is a point of discontinuity in the enthalpy of a pure substance. Hence, the derivative of enthalpy (heat capacity) goes to infinity or is at best undefined at the melting point. Further, it is likely that as measurements are made closer to the melting point, the measured heat capacity increases because of the increase in vacancies in aluminum. Furthermore, there is an increasing uncertainty of the measurement technique, for example, because of reaction of aluminum with the crucible at temperatures close to the melting point [36].

This work sought to examine the statistical contribution of all data sources to the model, correcting for the bias from differences in numbers of observations per dataset as well as differences in predicted invariant points. We also note that the aggregated model makes use of datasets which were excluded from past assessments as outliers because of deviations of more than 5 % with respect to the most trusted datasets at the time Desai made the assessment, such as KRA1972 and SCH1970. Nonetheless, these datasets receive low weights in the Bayesian approach, but still contribute to a more refined estimate of the heat capacity of liquid aluminum with uncertainty. The atomistic simulation datasets statistically contribute more in the solid phase than in the liquid phase. Choice of appropriate atomistic simulation datasets in combination with experimental measurements supports the same model form supported by experimental measurements alone. In addition, inclusion of the atomistic simulation datasets provides more data and reduces the uncertainty of the overall optimal model for the range of 0 to 1600 K. We note that inclusion of an interatomic potential that is optimized with liquid phase properties may receive a larger weight than the presently considered interatomic potentials. Furthermore, although the constant heat capacity model is preferred by experimental datasets for aluminum, a non-constant model for the heat capacity of the liquid phase is possible in materials such as found in the work of Becker et al. [9]. We also note that, in general, systematic bias can be caused by choosing a particular physics-based model for modeling the system because of the possibility of missing physics in the model. This can be remedied by modeling the bias as a Gaussian, such as been done for other models using Bayesian inference [72]. Although this was not performed in this work, segmented regression was chosen because literature shows that it captures all known physical effects on the heat capacity for aluminum [42], namely, both electronic and phonon contributions to the heat capacity at both low and high temperatures.

5. Conclusions

This study demonstrates that the inclusion of atomistic simulation data along with experimental measurements for the heat capacity and enthalpy of aluminum can support and improve the model forms used for the analysis of experimental datasets. The three models (experimental, atomistic and aggregated) agree within the Bayesian 95% credible

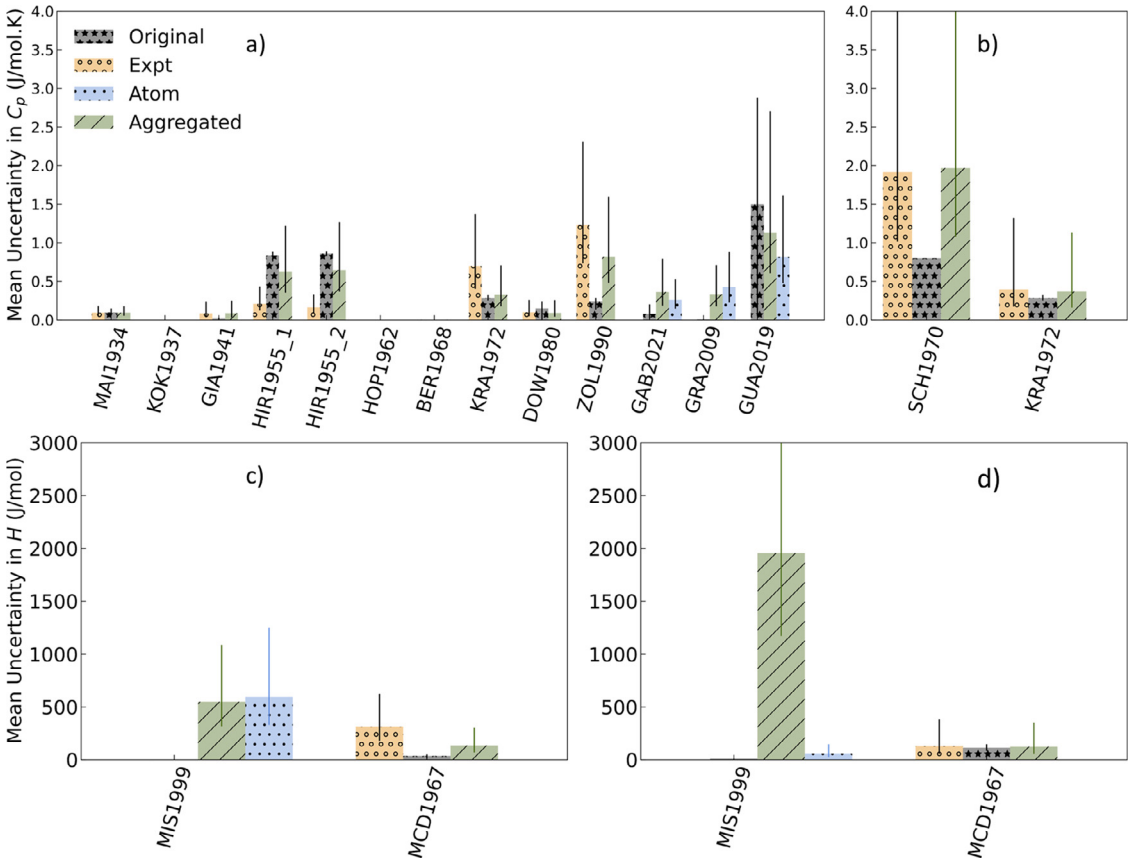

Fig. 5. Comparison of the mean original estimated uncertainty with the mean Bayesian rescaled uncertainty for (a, b) heat capacity datasets (in J/mol.K) and (c, d) enthalpy (in J/mol) considered in this work. In each subfigure, mean original uncertainty is solid (in black, web version). The mean rescaled uncertainty is shown as (i) circles (orange web version) for experimental model (Expt) (ii) dots (blue web version) for atomistic model, and (iii) hatched (green web version) aggregated model. Error bars indicate the Bayesian 95% credible interval on the mean uncertainty. All heat capacity datasets are rescaled by no more than an order of magnitude, while enthalpy datasets MIS1999 and MCD1967 are rescaled by more than an order of magnitude from their original uncertainties.

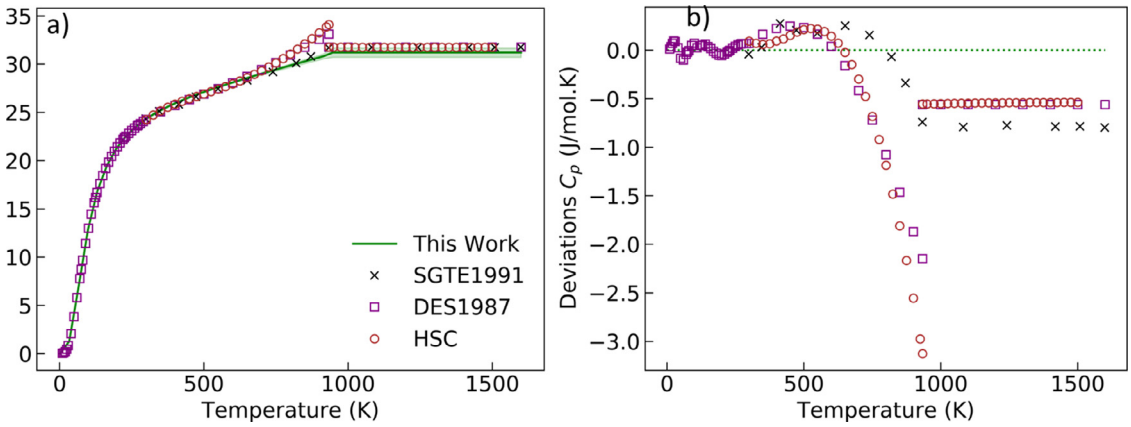

Fig. 6. Comparison of the optimal heat capacity model with past assessments by Desai, a commercial optimization software HSC Chemistry, and the SGTE1991 description in terms of (a) the predicted values and (b) the deviations, showing that agreement is within 3 J/mol.K.

interval up to 400 K and then the aggregated model and atomistic model deviate by less than 3 J/mol.K from the experimental model. The maximum full-width of the Bayesian 95% credible interval is 1 J/mol.K for the experimental model and is reduced to 0.9 J/mol.K for the aggregated model. When considered alone, certain atomistic simulation data can result in qualitatively different models compared to the experimental ones. Nevertheless, estimating and evaluating uncertainties allows for reliable selection of the optimal model. Moreover, the atomistic calculation results complement experimental measurements by providing data in regions of temperature and composition domains that are difficult to measure. In addition, the approach provides more data for model optimization, which results in models that are more reliable and predictable by providing quantified uncertainty.

Declaration of Competing Interest

None.

Acknowledgments

J.J.G., N.H.P. and M.S. gratefully acknowledge financial support from awards 70NANB14H012 and 70NANB19H005 from US Department of Commerce, National Institute of Standards and Technology as part of the Center for Hierarchical Materials Design (CHiMaD) in the Northwestern-Argonne Institute of Science and Engineering, and the Laboratory Directed Research and Development (LDRD) funding from Argonne National Laboratory, provided by the Director, Office of Science, of the US Department of Energy under Contract No. DE-AC02-06CH11357. We also gratefully acknowledge the computing resources provided by the Laboratory Computing Resource Center on the Bebop supercomputer at Argonne National Laboratory.

References

[1] M. Asta, D. de Fontaine, M. van Schilfgaarde, First-principles study of phase stability of Ti-Al intermetallic compounds, J. Mater. Res. 8 (10) (1993) 2554–2568.
[2] T.C. Duong, R.E. Hackenberg, A. Landa, P. Honarmandi, A. Talapatra, H.M. Volz, A. Llobet, A.I. Smith, G. King, S. Bajaj, A. Ruban, L. Vitos, P.E.A. Turchi, R. Arróyave, Revisiting thermodynamics and kinetic diffusivities of uranium-niobium with Bayesian uncertainty analysis, Calphad 55 (2) (2016) 219–230.
[3] S. Bigdeli, L.F. Zhu, A. Glensk, B. Grabowski, B. Lindahl, T. Hickel, M. Selleby, An insight into using DFT data for Calphad modeling of solid phases in the third generation of Calphad databases: a case study for Al, Calphad 65 (2019) 79–85.
[4] B. Sundman, H.J. Seifert, F. Aldinger, The Ringberg workshop 1996 on solution modeling, Calphad 21 (2) (1997) 139–141.
[5] B. Sundman, F. Aldinger, H.J. Seifert, The Ringberg workshop 1997 on the application of computational thermodynamics, Calphad 24 (1) (2000) 15–17.
[6] B. Burton, N. Dupin, S. Fries, G. Grimvall, A. Fernández Guillermet, P. Miodownik, A. Oates, V. Vinograd, Using *ab initio* calculations in the CALPHAD environment, Z. Metall. 92 (6) (2001) 514–525.
[7] M. Palumbo, B. Burton, A. Costa e Silva, B. Fultz, B. Grabowski, G. Grimvall, B. Hallstedt, O. Hellman, B. Lindahl, A. Schneider, P.E.A. Turchi, W. Xiong, Thermodynamic modeling of crystalline unary phases, Phys. Status Solidi B 251 (1) (2014) 14–32.
[8] P.E.A Turchi, I. Abrikosov, B. Burton, S. Fries, G. Grimvall, L. Kaufman, P. Korzhavyi, V. Manga, M. Ohno, A. Pisch, A. Scott, W. Zhang, Interface between quantum-mechanical-based approaches, experiments, and CALPHAD methodology, Calphad 31 (2007) 4–27.
[9] C.A. Becker, J. Ågren, M. Baricco, Q. Chen, S.A. Decterov, U.R. Kattner, J.H. Perepezko, G.R. Pottlacher, M. Selleby, Thermodynamic modelling of liquids: CALPHAD approaches and contributions from statistical physics, Phys. Status Solidi B 251 (1) (2014) 33–52.
[10] T. Hickel, U.R. Kattner, S.G. Fries, Computational thermodynamics: Recent development and future potential and prospects, Phys. Status Solidi B 251 (2014) 9–13.
[11] Z.K. Liu, First-principles calculations and CALPHAD modeling of thermodynamics, J. Phase Equilb. Diffus. 30 (5) (2009) 517–534.
[12] J.J. Gabriel, N.H. Paulson, T.C. Duong, F. Tavazza, C.A. Becker, S. Chaudhuri, M. Stan, Uncertainty quantification in atomistic modeling of metals and its effect on mesoscale and continuum modeling: a review, JOM 73 (1) (2021) 149–163.
[13] P.W. Guan, G. Houchins, V. Viswanathan, Uncertainty quantification of DFT-predicted finite temperature thermodynamic properties within the Debye model, J. Chem. Phys. 151 (24) (2019) 244702.
[14] L.-F. Zhu, J. Janssen, S. Ishibashi, F. Körmann, B. Grabowski, J. Neugebauer, A fully automated approach to calculate the melting temperature of elemental crystals, Comput. Mater. Sci. 187 (11) (2021) 110065.
[15] P. Honarmandi, R. Arróyave, Uncertainty quantification and propagation in computational materials science and simulation-assisted materials design, Integr. Mater. Manuf. Innov. 9 (1) (2020) 103–143.
[16] F. Cailliez, P. Pernot, Statistical approaches to forcefield calibration and prediction uncertainty in molecular simulation, J. Chem. Phys. 134 (5) (2011) 54124.
[17] H.L. Lukas, S.G. Fries, B. Sundman, Computational Thermodynamics: The Calphad Method, Cambridge University Press, Oxford, UK, 2007.
[18] M. Stan, B. Reardon, A Bayesian approach to evaluating the uncertainty of thermodynamic data and phase diagrams, Calphad 27 (3) (2003) 319–323.
[19] N.H. Paulson, E. Jennings, M. Stan, Bayesian strategies for uncertainty quantification of the thermodynamic properties of materials, Int. J. Eng. Sci. 142 (2019) 74–93.
[20] N.H. Paulson, B.J. Bocklund, R.A. Otis, Z.K. Liu, M. Stan, Quantified uncertainty in thermodynamic modeling for materials design, Acta Mater. 174 (2019) 9–15.
[21] N.H. Paulson, S. Zomorodpoosh, I. Roslyakova, M. Stan, Comparison of statistically-based methods for automated weighting of experimental data in CALPHAD-type assessment, Calphad 68 (2020) 101728.
[22] C.G. Maier, C.T. Anderson, The disposition of work energy applied to crystals, J. Chem. Phys. 2 (8) (1934) 513–527.
[23] J. Kok, W. Keesom, Measurements of the atomic heat of aluminium from 1.1 to 20 K, Phys. D 4 (9) (1937) 835–842.
[24] W.F. Giauque, P.F. Meads, The heat capacities and entropies of aluminum and copper from 15 to 300 K, J. Amer. Chem. Soc. 63 (7) (1941) 1897–1901.
[25] D.C. Hopkins, Thermodynamic Properties of Aluminum Near its Superconducting Critical Temperature, University of Illinois at, Urbana-Champaign, Illinois, 1962.
[26] W.T. Berg, Heat capacity of aluminum between 2.7 and 20 K, Phys. Rev. 167 (3) (1968) 583–586.
[27] D. Downie, J. Martin, An adiabatic calorimeter for heat-capacity measurements between 6 and 300 K. The molar heat capacity of aluminium, J. Chem. Thermodyn. 12 (8) (1980) 779–786.
[28] T. Pochapsky, Heat capacity and resistance measurements for aluminum and lead wires, Acta Metall. 1 (6) (1953) 747–751.
[29] D.C. Rorer, H. Meyer, R.C. Richardson, Specific heat of aluminum near its superconductive transition point, Z. Naturf. Pt A 18 (2) (1963) 130–140.
[30] W.J. Parker, R.J. Jenkins, C.P. Butler, G.L. Abbott, Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity, J. Appl. Phys. 32 (9) (1961) 1679–1684.
[31] K. Hirano, H. Maniwa, Y. Takagi, Specific heat measurements on quench-annealed Al, Cu and α-phase alloys of Cu, J. Phys. Soc. Jpn. 10 (1955) 909–910.
[32] U. Schmidt, O. Vollmer, R. Kohlhaas, Thermodynamische Analyse Kalorimetrischer Messungen an Aluminium und Wolfram im Bereich hoher Temperaturen, Z. Naturf. Pt A 25 (1970) 1258–1264.
[33] W. Kramer, J. Noelting, Anomalous specific heat and disorder of the metals indium, tin, lead, zinc, antimony, and aluminum, Acta Metall. 20 (12) (1972) 1353–1359.
[34] M. Zoli, V. Bortolani, Thermodynamic properties of FCC metals: Cu and Al, J. Phys. Condens. Matter 2 (3) (1990) 525–539.
[35] E.D. Eastman, A.M. Williams, T.F. Young, The specific heats of magnesium, calcium, zinc, aluminum and silver at high temperatures, J. Am. Chem. Soc. 46 (5) (1924) 1178–1183.
[36] R.A. McDonald, Enthalpy, heat capacity, and heat of fusion of aluminum from 366 K to 1647 K, J. Chem. Eng. Data 12 (1) (1967) 115–118.
[37] M.J. O'Neill, Measurement of specific heat functions by differential scanning Calorimetry, Anal. Chem. 38 (10) (1966) 1331–1336.
[38] P.D. Desai, Thermodynamic properties of aluminum, Int. J. Thermophys. 8 (5) (1987) 621–638.
[39] A.T. Dinsdale, SGTE data for pure elements, Calphad 15 (4) (1991) 317–425.
[40] N. Paulson, UnaryBayes, [Online Accessed May 2021]. URL: https://github.com/npaulson/UnaryBayes.
[41] Evaluation of measurement data – guide to the expression of uncertainty in measurement, 2008 [Online Accessed August 2020]. URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6
[42] I. Roslyakova, B. Sundman, H. Dette, L. Zhang, I. Steinbach, Modeling of Gibbs energies of pure elements down to 0 K using segmented regression, Calphad 55 (2) (2016) 165–180.
[43] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, D. Kocevski, X-ray spectral modelling of the AGN obscuring region in the CDFS: bayesian model selection and catalogue, Astron. Astrophys. 564 (2014) 1–25.
[44] B. Grabowski, L. Ismer, T. Hickel, J. Neugebauer, *Ab initio* up to the melting point: anharmonicity and vacancies in aluminum, Phys. Rev. B 79 (13) (2009) 134106.
[45] G. Kresse, J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (16) (1996) 11169–11186.
[46] P. Blöchl, Projector augmented-wave method, Phys. Rev. B 50 (24) (1994) 17953–17979.
[47] H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B. 13 (12) (1976) 5188–5192, doi:10.1103/PhysRevB.13.5188.
[48] A. Togo, I. Tanaka, First principles phonon calculations in materials science, Scr. Mater. 108 (2015) 1–5.
[49] J.J. Gabriel, F.Y.C. Congo, A. Sinnott, K. Mathew, T.C. Allison, F. Tavazza and R.G. Hennig, Uncertainty quantification for materials properties in density functional theory with *k*-point density, arXiv preprint arXiv:URL: 2001.01851 (2020). https://arxiv.org/abs/2001.01851.
[50] T. Duong, S. Gibbons, R. Kinra, R. Arróyave, *Ab-initio* aprroach to the electronic, structural, elastic, and finite-temperature thermodynamic properties of Ti_2AX (A = Al or Ga and X = C or N), J. Appl. Phys. 110 (2011) 093504.
[51] A. van de Walle, G. Ceder, Correcting overbinding in local-density-approximation calculations, Phys. Rev. B. 59 (23) (1999) 14992–15001.
[52] P. Janthon, S.M. Kozlov, F. Viñ, J. Limtrakul, F. Illas, Establishing the accuracy of broadly used density functionals in describing bulk properties of transition metals, J. Chem. Theory Comput. 9 (3) (2013) 1631–1640.
[53] L.F. Zhu, F. Körmann, A.V. Ruban, J. Neugebauer, B. Grabowski, Performance of the standard exchange-correlation functionals in predicting melting properties fully from first principles: application to Al and magnetic Ni, Phys. Rev. B 101 (14) (2020) 144108.
[54] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865–3868.
[55] J.P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Phys. Rev. B 23 (10) (1981) 5048–5079.
[56] A. Rohatgi, WebPlotDigitizer, [Online Accessed November 2020]. URL: https://automeris.io/WebPlotDigitizer.
[57] J.B. Sturgeon, B.B. Laird, Adjusting the melting point of a model system via Gibbs–Duhem integration: application to a model of aluminum, Phys. Rev. B 62 (22) (2000) 14720–14727.
[58] Y. Mishin, D. Farkas, M.J. Mehl, D.A. Papaconstantopoulos, Interatomic potentials for monoatomic metals from experimental data and *ab initio* calculations, Phys. Rev. B 59 (5) (1999) 3393–3407.
[59] Y. Mishin, Atomistic modeling of the γ and γ'-phases of the Ni-Al system, Acta Mater. 52 (6) (2004) 1451–1467.
[60] NIST Interatomic Potentials Repository, [Online Accessed August 2020]. URL: http://www.ctcms.nist.gov/potentials.

[61] L. Zhang, D.Y. Lin, H. Wang, R. Car, E W., Active learning of uniformly accurate interatomic potentials for materials simulation, Phys. Rev. Mater. 3 (2) (2019) 023804.
[62] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. 117 (1) (1995) 1–19.
[63] H. Flyvbjerg, H.G. Petersen, Error estimates on averages of correlated data, J. Chem. Phys. 91 (1) (1989) 461–466.
[64] Allen-Tildesley python examples, [Online Accessed August 2020]. URL: https://github.com/Allen-Tildesley/examples/blob/master/python_examples/error_calc.py.
[65] M.W. Chase, NIST-JANAF thermochemical tables, J. Phys. Chem. Ref. Data 9 (1998) 1–1551 Monograph No. 9.
[66] C. Ho, R. Powel, R. Lily, Thermal conductivity of the elements, J. Phys. Chem. Ref. Data 1 (2) (1972) 279–421.
[67] HSC Chemistry, [Accessed Online August 2020]. URL: https://hsc-chemistry.com
[68] A. Burcat, B. Ruscic, Update of third millennium ideal gas and condensed phase thermochemical database for combustion with updates from active thermochemical tables, Aerosp. Eng. Argonne Natl. Lab. Report ANL 05/20 and TAE960 Technion-llT, 2005.
[69] O. Knacke, O. Kubaschewski, K. Hesselman, in: Thermochemical Properties of Inorganic Substances, 2nd ed., Springer Verlag, Berlin, 1991, pp. 1114–2412.
[70] G. Belov, V. lorish, V. Yungman, IVTANTHERMO - database on thermodynamic properties and related software, Calphad 23 (2) (1999) 173–180.
[71] ThermoCalc Software, [Accessed Online August 2020]. URL: https://thermocalc.com/
[72] J. Brynjarsdóttir, A. O'Hagan, Learning about physical parameters: the importance of model discrepancy, Inverse Probl. 30 (2014) 114007.