

Full length article

# Machine learning superalloy microchemistry and creep strength from physical descriptors

Patrick L. Taylor<sup>a,b,\*</sup>, Gareth Conduit<sup>a</sup><sup>a</sup> Cavendish Laboratory, University of Cambridge, CB3 0HE, UK<sup>b</sup> Department of Mechanical Engineering, Imperial College, SW7 2BX, UK

## ARTICLE INFO

## Keywords:

Superalloys  
Machine learning  
Gaussian process regression  
Phase composition  
Microstructure  
Creep strength  
CALPHAD

## ABSTRACT

We propose an element-agnostic set of descriptors to model superalloy properties with Gaussian process regression. Furthermore, we develop a correction method to deliver the best and most physical predictions for microchemistry in multi-phase alloys. The model's performance in predictions is confirmed for superalloy microchemistry, microstructure, and strength properties. When including new, unseen elements in the test data, the models still give good predictions; such extrapolations into new chemical-space would be impossible with component-based descriptors.

## 1. Introduction

Hume-Rothery first developed his eponymous set of rules in 1935 [1, 2]. They describe whether any two elements could alloy together to form a stable solid solution. For substitutional alloys there are four rules: they concern the similarity of atomic radius, crystal structure, valency, and electronegativity. Further developments to solubility rules include the Pettifor scale and its 2016 proposed update by Glawe et al. [3,4]. The Hume-Rothery rules highlight the opportunity for fully element-agnostic models of phase composition that would enable them to be applied to any material, even those containing elements not before considered. However, the Calculation of Phase Diagram (CALPHAD) methodology – which has become the industry standard approach due to the promulgation of software such as ThermoCalc – typically relies on thermodynamic models that have been constructed element-by-element, limiting their broad applicability [5–7].

The same duality exists in applications of machine learning to alloys. Some researchers have adopted the approach of using alloy components directly as descriptors in their ML models. Such models have been used to model Ni-based superalloy microstructures, and design superalloys for a diverse range of applications [8–13]. Other researchers have taken the approach of mapping alloy components to physical descriptors based on domain knowledge [14–19]. This strategy greatly improves models trained on small datasets, whereas models trained on suitably large datasets were already able to self-encode the physical descriptions of a system in their latent space [20], see Hart et al. (2021) for a thorough review [21].

Superalloys are precipitation hardened—their important bulk properties, such as high-temperature creep and yield strength, derive from their two-phase microstructure [22–26]. Hence, accurate prediction of the relative phase fraction and phase compositions is a crucial first-step towards property prediction. However, each generation of superalloys has typically included additional elements [27], the most recent being the inclusion of ruthenium [28–32]. The trend of additional elements improving properties is naturally extended by high-entropy superalloys (HESA), in which the principles of high-entropy alloy design are applied to the superalloy matrix phase [33–37]. Machine learning models that better transfer their inherent knowledge of alloy chemistry to new composition-space could accelerate the further development of superalloys.

We propose a physics-inspired set of descriptors to describe superalloys that is element-agnostic. Gaussian process regression machine learning is then used to predict superalloy properties. For phase composition we build on previous work [10] to develop and demonstrate a probabilistic correction method. In addition to interpolative scoring of the models in the usual cross-validated manner (via a withheld, randomly selected, test dataset), models are also scored on their ability to extrapolate to datasets containing alloy components not seen during training. Finally, creep strength models of Ni superalloys are developed using similar physics-inspired descriptor sets.

## 2. Computational method

In this section we detail the computational method developed to predict phase behaviour. We first describe the physical descriptor-set

\* Corresponding author at: Department of Mechanical Engineering, Imperial College, SW7 2BX, UK.

E-mail address: [p.taylor23@imperial.ac.uk](mailto:p.taylor23@imperial.ac.uk) (P.L. Taylor).

<https://doi.org/10.1016/j.commsci.2023.112265>

Received 10 January 2023; Received in revised form 10 May 2023; Accepted 12 May 2023

Available online 23 May 2023

0927-0256/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

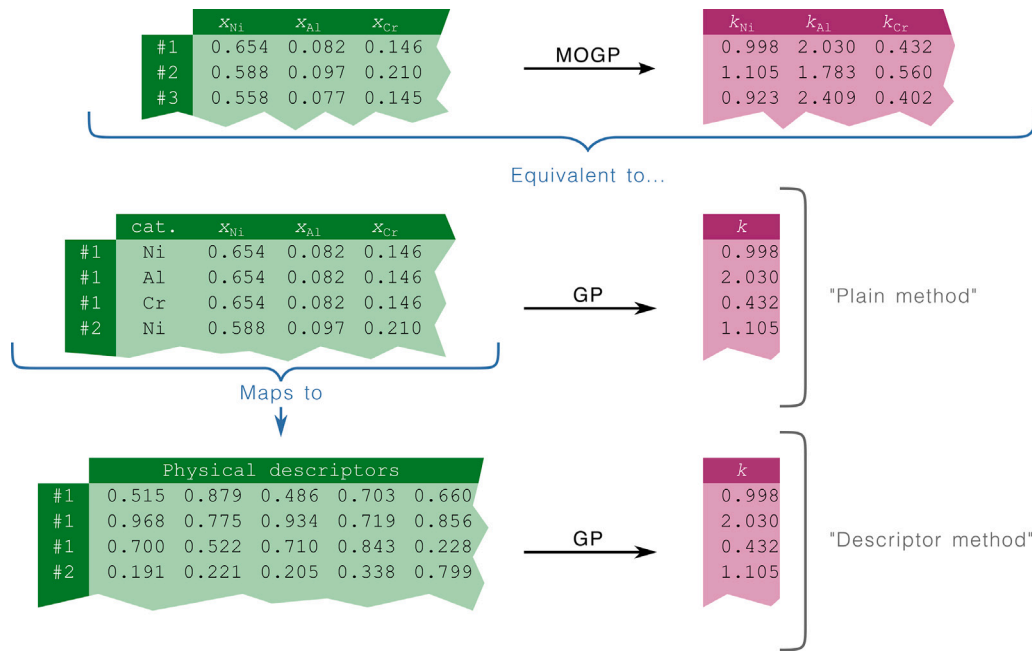


Fig. 1. An overview of how the Gaussian process regression models in this work were formulated.

used as input features to our machine learning models, then proceed to give details of the Gaussian Process Regression models. Next, we discuss some particular problems with representing phase microchemistry data in machine learning models, and the novel correction method developed to address them. Finally, we describe our approach to physical descriptor GPR models of creep strength.

### 2.1. Data representation

Our first task was to predict superalloy microchemistry. This is described by the phase fraction,  $f^\phi$ , and the corresponding composition of each phase,  $x_i^\phi$ . Following other authors and our previous work, we adopt partitioning coefficients in place of phase composition,  $k_i^\phi = x_i^\phi/x_i$  [8,10,28,32,38–41]. The partitioning coefficients give a simpler encoding of physical information—for a given element, its partitioning coefficients across many different alloys are typically more similar than its various phase component percentages. In order to use variables that are more normally distributed, we use a further transformation of  $f^\phi$ ,  $k_i^\phi$  from their respective intervals to the real line:

$$p_i^\phi = \ln\left(x_i^\phi/x_i\right) \quad ,$$

$$q^\phi = \operatorname{arctanh}\left(2f^\phi + 1\right) \quad .$$

As well as being necessary mappings, they also capture the physical symmetry:  $f^\phi \rightarrow 1 - f^\phi \Leftrightarrow q^\phi \rightarrow -q^\phi$  and  $k^\phi \rightarrow k^{\phi^{-1}} \Leftrightarrow q^\phi \rightarrow -q^\phi$ .

The dataset used in this work is an updated version of that in Ref. [10], containing 123 entries with complete microstructure data (matrix and precipitate phase composition and fractions) [42]. The phase composition dataset is presented in the conventional form, with the nominal alloy compositions and heat treatments being the inputs to our model, and the phase fractions and respective compositions being the output. However, in this work we “reshape” the dataset into a single-target format. This is represented graphically in Fig. 1 as a two-step process. First the dataset is reshaped so that for a single-phase ML model, each partitioning coefficient is considered to be a separate output, and the input gains an additional column labelling the output element. Note that in the GPR framework, this is equivalent to a multi-output Gaussian process (MOGP) [43]. We will refer to this model as the “plain method”. Secondly, the data is transformed into a physical representation, which is a function of composition and label,

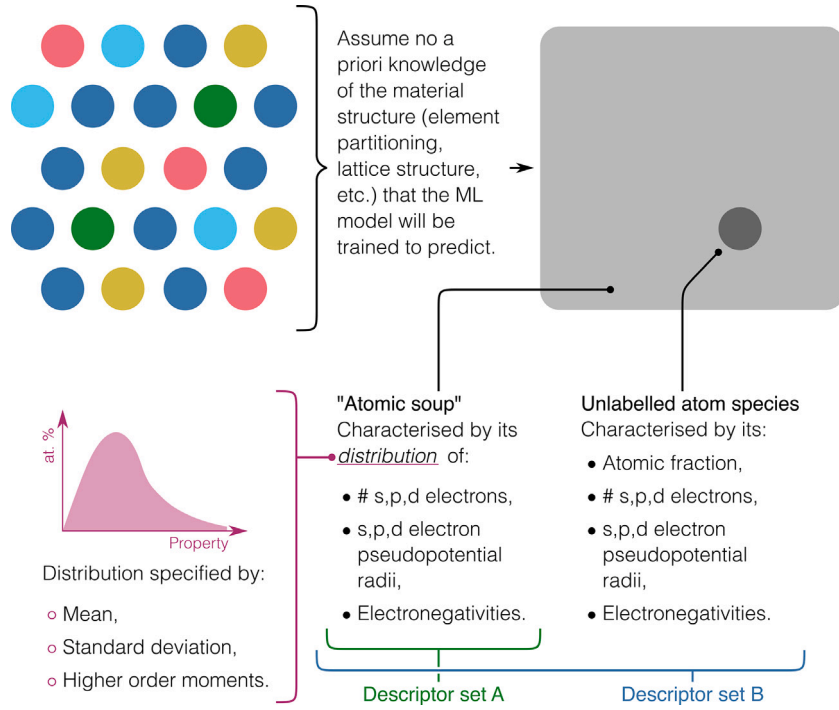
but does not explicitly preserve the label as a feature. This means that predictions can be made for labels (i.e. components) not present in the training data—so long as the physical descriptors are carefully chosen. We will refer to this model as the “descriptor method”.

### 2.2. Physical descriptors

To fully understand the phase behaviour we must predict both the phase fraction, and also the partitioning of each element between the phases. The model for each property requires different inputs: for the phase fraction model, inputs were composition and heat treatment data; for the partitioning coefficient model, each entry gains a further feature corresponding to the component-label of the target, see Fig. 1 and Section 2.1. In our model, both this label and the composition are converted into element-agnostic descriptors. The most natural choice of physical descriptors would be based on the lattice structure and atomic arrangement of an alloy, but these are a priori unknown. Instead, a given alloy – for which, in our case, we only know the nominal composition – can be thought of an “atomic soup”, see Fig. 2. Inspired by ab initio electronic structure methods used in physics and chemistry, the atomic soup can be represented by the distribution of the constituent atoms’ electronic properties. Said distributions can be approximately specified by their mean, standard deviation; and higher order moments if necessary. This is how the descriptors (descriptor set A in Fig. 2) for modelling phase fraction were formulated. For the GPR model of partitioning coefficients, the atomic species labels were also transformed to physical descriptors (Fig. 1 and Section 2.1), chosen following a similar logic (descriptor set B in Fig. 2).

Our choice of descriptors bore a strong similarity to the popular descriptor-set MAGPIE [44]. Those used by Ling et al. [45] and Liu et al. [46] to model nickel superalloy properties are also similar. Much like these descriptor sets, we found electronic structure inspired descriptors to be especially powerful. Data used to construct our descriptors was taken from Refs. [47,48].

Alongside its nominal composition, the precipitation heat treatments applied to an alloy will also affect its final composition and properties [10]. Up to three heat treatment stages were used as input features for each alloy; each comprising a treatment temperature ( $T_j$ ) and time ( $\tau_j$ ), for a total of six features. The theory of Ostwald ripening



**Fig. 2.** How physics-based descriptors were selected for this work. Descriptor sets A and B were used for different aspects of microstructure modelling. Descriptor set A was used to model phase fraction, B was used to model the partitioning coefficients.

gives a theoretical dependence on heat treatment time and temperature for an alloy's phase compositions [49–52]:

$$\langle \Delta x_i^\phi \rangle^3(t) \sim \frac{e^{-\varepsilon_i/k_B T}}{T} t \quad ,$$

where  $\varepsilon_i$  is an activation energy associated with the diffusion of species  $i$ . Our model's targets are the log partitioning coefficients; heat treatment times in our database varied over several orders of magnitude (from around 15 mins to 1000 h). Putting this altogether, we propose representing  $m_{\text{HT}}$ -stage heat treatments with the following descriptors:

$$h'_j = \begin{cases} \sum_{k=1}^j \ln(\tau_k) - \ln(T_k) , & j \leq m_{\text{HT}} \\ 1/T_j , & m_{\text{HT}} < j \leq 2m_{\text{HT}} \end{cases} \quad (1)$$

This retains the same total number of descriptors as in the initial input. In Section 3.3 we show that this set of heat treatment descriptors accurately capture the evolution of the alloys' microstructures.

### 2.3. Gaussian process regression

A large number of physical descriptors were proposed for use in our GPR model: much more than the number of input columns in our dataset (18 for our dataset with component and heat treatment inputs). To ensure that the GPR models were selecting the minimal set of relevant features during hyperparameter optimisation, an automatic relevance determination (ARD) Matérn kernel was used [53]. In cases where the ARD kernel lengthscale  $l$  for some of the features was found to be very large,  $l \gtrsim 10^3$ , small improvements in the model's score could be obtained by retraining the model without using said features at all. The GPyTorch library was used for GPR [54]. Both the L-BFGS and ADAM algorithms were tested for hyperparameter optimisation, with L-BFGS being found to give both better results for a small trade-off in training time.

### 2.4. Probabilistic correction to phase compositions

There are three physical constraints that alloy phase compositions – i.e. the output predictions of the GPR models for phase chemistry –

must obey [5,10]:

$$\sum_i x_i^\phi = 1 \quad , \quad (2)$$

$$\sum_\phi f^\phi = 1 \quad , \quad (3)$$

$$\sum_\phi f^\phi x_i^\phi = x_i \quad , \quad (4)$$

where there are  $m$  phases labelled  $\phi$  and  $n$  components labelled  $i$ . Eqs. (2) & (3) are hard constraints that the total concentration and total phase fractions must each sum to unity. Eq. (4) can be interpreted as a soft constraint representing an assumption of minimal material loss in the forging process. Gaussian process regression does not present an obvious way to impose such constraints—in Ref. [10] they were imposed on the final model outputs. In particular, a Bayesian approach was taken to apply constraint Eq. (4) via a correction to the predicted phase fraction. In this work, the same approach is extended to apply a simultaneous correction to the phase fractions  $f^\phi$  and the phase compositions  $x_i^\phi$ . The output of the GPR models for phase fraction and phase composition are taken to be independent Gaussian processes, which for each alloy gives a prior:

$$\exp \left[ -\frac{1}{2} (\mathbf{q} - \hat{\mathbf{q}})^T \Sigma_q (\mathbf{q} - \hat{\mathbf{q}}) \right] \prod_\phi \exp \left[ -\frac{1}{2} (\mathbf{p}^\phi - \hat{\mathbf{p}}^\phi)^T \Sigma_{p^\phi} (\mathbf{p}^\phi - \hat{\mathbf{p}}^\phi) \right] \quad , \quad (5)$$

and a likelihood relating to the soft constraint, where  $\tau_i$  is a tolerance for 'allowed' component loss, and  $\sigma_i^{(3)}$  is an estimated uncertainty on the sum on the LHS of Eq. (4)

$$\prod_i \exp \left[ -\frac{1}{2} \left( \frac{x_i - \sum_\psi f^\psi x_i^\psi}{\tau_i x_i + \sigma_i^{(3)}} \right)^2 \right] \quad . \quad (6)$$

Combining these to give a posterior, expanding the exponent of the likelihood to quadratic order in  $\Delta p_i^\phi = p_i^\phi - \hat{p}_i^\phi$ ,  $\Delta q^\phi = q^\phi - \hat{q}^\phi$ , and finally completing the square gives a new Gaussian probability distribution. This is equivalent to maximising the log-posterior probability with

respect to the corrections to  $\hat{q}$  and  $\hat{p}$ . This maximisation can be carried out subject to hard constraints for Eqs. (2) & (3). Doing so produces a correction to each of  $p_i^\phi$  and  $q^\phi$ , as well as a new, valid covariance for a given prediction, which in turn yields the associated uncertainties.

### 2.5. Creep strength modelling

We compiled a dataset of creep strength properties for single crystal Ni superalloys. Entries were drawn from academic literature and commercial databases [42]. A substantial contribution to the dataset was from the open source creep rupture life dataset compiled by Liu et al. [46] (266 entries). Database entries included multiple properties characterising creep strength, including elongation at rupture (82 entries), time to 1% creep (79), and minimum secondary creep rate (55). There were significantly more entries available for creep rupture life (388), reflecting the importance of this particular property as the primary metric of creep strength [23,55]. For this reason we solely focused on modelling creep rupture life. Each entry had two corresponding experimental conditions, a temperature and applied stress. Other authors have found that direct ML models of the creep rupture life are more effective than modelling the Larson–Miller parameter [18]. We adopted the same approach, which meant both experimental conditions were included as input features.

Three approaches were taken to create a GPR model for creep strength:

- Plain composition descriptors (directly analogous to the plain method described above).
- Physical descriptor set derived from the input composition only (analogous to the descriptors used for modelling phase fraction).
- Physical and metallurgical descriptors derived both from the input composition and from the fitted GPR microstructure model.

Atomistic microstructural properties, in combination with precipitate morphology, are known to determine the physical mechanisms by which creep occurs [22,23]. This motivated the use of the third model. In this model the precipitate fraction predicted by the microstructure model was used directly as a descriptor. Other derived descriptors were used: the lattice misfit between the two phases was calculated using the Vegard coefficients [56–58]. The matrix phase stacking fault energies were approximated using fcc and hcp formation energies for each element [23,41,59,60]. The formation energies were calculated via density functional theory using the PBEsol functional [61]. The mean and standard deviation of melting points for elements in the precipitate phase were included as proxies for the  $\gamma'$  solvus [62,63]. The mean interdiffusivity metric and the mean metal d-level metrics used in the alloys-by-design procedure were also used as descriptors [57,64–67]. The necessary microstructure predictions were made using the GPR model described in the previous sections, trained on the full microstructure database excluding any entries overlapping with the creep rupture life dataset.

This GPR model incorporated more explicit high-level domain knowledge than any of the other models, which was only possible because the physical mechanisms that govern creep deformation have been well studied by metallurgists. However, the descriptors of this type that were used are not an exhaustive list, and for this reason physics-based descriptors that did not use explicit metallurgical domain knowledge were also used in this model (see Fig. 8).

## 3. Results

With the data curated, descriptors selected, and machine learning formalism in place, we are now well-positioned to test the performance of our machine learning algorithms. We first study the performance of microstructure prediction, before studying the creep strength.

To compare the performance of the proposed physical-descriptor model to the plain model that uses composition features, we performed

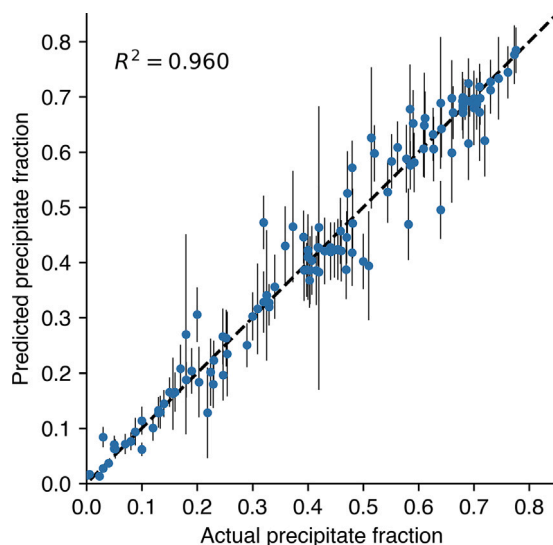


Fig. 3. Predicted precipitate fractions from the descriptor model versus the actual values. The vertical bars are the model's uncertainties.

two rounds of tests on the microstructure data: firstly on blind validation data where all elements had been present in the training data, and secondly for extrapolating to new materials that contain fresh elements not present in the training dataset.

### 3.1. Performance when all elements available

We first test the performance of the physical descriptor model when information about all elements is available at training. Two models were trained: one that uses physical descriptors as inputs and a second that inputs the plain composition. We carried out ten-fold cross validation on the available microstructure data (123 database entries). For each fold, the model was trained via log-likelihood maximisation over the training dataset, then predictions were made on the withheld validation set. These results were combined to give a single set of predictions for the entire dataset. The correction method from Section 2.4 was applied in the same way to both models.

The root mean squared error (RMSE) was used to compare the two models due to its ease of interpretation for percentage-like properties. The RMSE for each element and fraction component of the microstructure is given in Table 1. The physical descriptor model was significantly better than the plain composition model for predicting phase fraction, improving the RMSE from 6.3% to 4.4%. The predicted versus actual phase fraction is plotted in Fig. 3, which confirms not only the quality of predictions, but also the accuracy of the uncertainty estimates. The phase fraction is the most crucial measure of a superalloy's microstructure given its physical influence on yield and creep strength. The descriptor-based model shows a smaller improvement for phase composition predictions when compared with the plain method, although it does still achieve the same or better RMSE for almost every element.

We also confirmed the selected kernel hyperparameters. A Matérn kernel with smoothness parameter  $\nu = 2.5$  gave the best results for the log partitioning coefficient models and the phase fraction model. The fact that a smoother kernel gives better predictions when using physical descriptors is because the models are 'forced' to find the best physical descriptors rather than relying on the kernel's complexity to fit the data. In turn, the combination of a simpler, smoother model with a superior 'understanding' of chemistry allows the model to extrapolate well.

**Table 1**

Comparison of the root mean squared error (RMSE) for different components between the GPR model using descriptors and that using plain composition.

		Phase frac.(at. %)	Composition (at. %)					
			Ni	Cr	Co	Al	Ti	Heavy els.
Phase 1	Descriptor method	–	4.8	4.8	3.4	1.7	1.2	0.6
	Plain method	–	5.4	5.2	3.2	1.6	1.3	0.6
Phase 2	Descriptor method	4.4	3.8	2.7	2.7	2.4	1.5	0.5
	Plain method	6.3	3.8	3.0	3.0	2.4	2.0	0.6

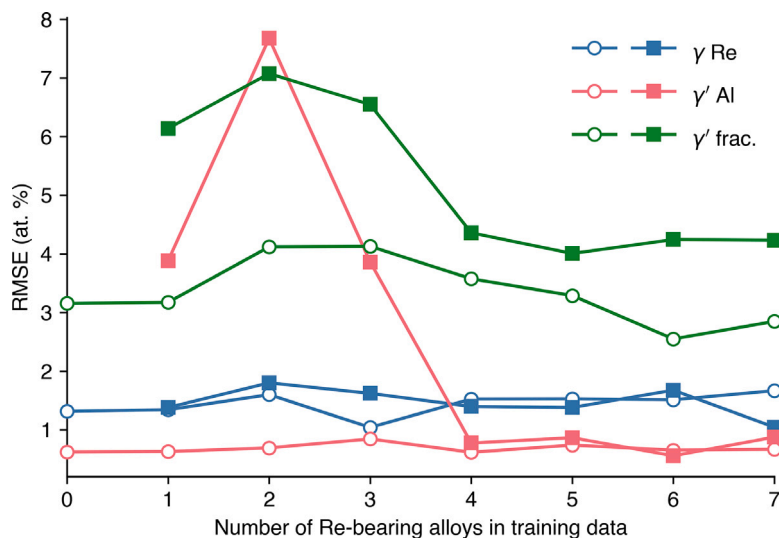


Fig. 4. Predictions on a test dataset of Re bearing superalloys made by models trained on a dataset of 85 superalloys without Re plus a number of alloys with Re (indicated on the x-axis). The points with open-circles are the physical descriptor model and with squares are the plain composition descriptor model.

### 3.2. Extrapolative predictions in composition-space

To test whether descriptors improve the ability of the model to extrapolate in composition-space, we adopted the simple approach of testing our GPR model on superalloy families outside the training dataset. We focus on two categories of alloys: Re/Ru-bearing Ni-superalloys, and high-entropy superalloys (HESA).

The addition of Re and Ru to commercial superalloy compositions was one of the key innovations of the most recent generations of conventionally developed single-crystal superalloys [23,68–70]. Here we repeat the training process and start from a database comprising 88 historic superalloys that contain neither Re nor Ru. Of the data on contemporary superalloys containing Re and Ru, 6 were randomly selected to form a test set, and the remaining were added incrementally to the training set so as to expose the importance of adding fresh alloys during a research project. Fig. 4 shows the RMSE vs. number of Re-bearing alloys in the training data. The initial predictions by the physical descriptor model are extremely impressive, giving about the same RMSE on the test data as it attains with any amount of Re-bearing alloy data. When a few data-points for Re alloys are added, the plain-descriptor model overfits so gives a higher RMSE than the physical descriptor model. For the crucial prediction of the  $\gamma'$  phase fraction, this improvement persists up to the maximum number of additional training data points.

High-entropy superalloys are a more recent development in alloy design, combining the design principles of high-entropy alloys and precipitation-strengthened alloys [71,72]. HESAs typically contain Fe as one of their entropy of mixing-boosting components. For a review on HESAs see Ref. [35]. HESAs are a prime target for the physical descriptor model owing to the large number of element permutations that cannot all be represented in the training dataset. The GPR model was trained on the full conventional superalloy database, that did not

contain a single HESA entry, nor any entries containing Fe. The model was then tested on the HESA data collected by Zhang et al. [71] that includes Fe-bearing superalloys, with results shown in Fig. 5. The results are overall in excellent agreement, capturing the behaviour of the elements known to the model well, and notably making good predictions for Fe, which would have been impossible with the plain descriptor model. For both experimental HESAs, the RMSE for  $\gamma'$  phase elements was 1.8% and for  $\gamma$  phase elements was 7.0%. Once again, predictions for the  $\gamma'$  phase are better than for the  $\gamma$  phase, reflecting the stronger influence of physical factors on this phase's formation. Such a prediction would be simply impossible to make with a plain composition descriptor model.

### 3.3. Heat treatments

To assess how well our heat treatment descriptors (Eq. (1)) capture the evolution of microstructure during a heat treatment, we test our model against two sets of experimental data [73,74]. In each case, the physical descriptor model was retrained on the full database excluding the respective set of test alloys.

The first test case was a commercial superalloy SRR99, aged under a variety of conditions. Our model captured both the qualitative and quantitative trends in the evolution of the  $\gamma'$  fraction, see Fig. 6(a). Surprisingly, it exhibited both the largest uncertainty and error for heat treatment A, the un-aged specimen. This could be reflective of a spurious correlation in the training data, since data for commercial superalloy microstructures are typically presented for fully heat treated specimens, whereas “experimental” alloy compositions are less frequently fully heat treated.

The second test case was for an experimental five-component superalloy, where each specimen was aged at 760 °C for increasingly long durations, shown in Fig. 6(b). Our model captured not only the

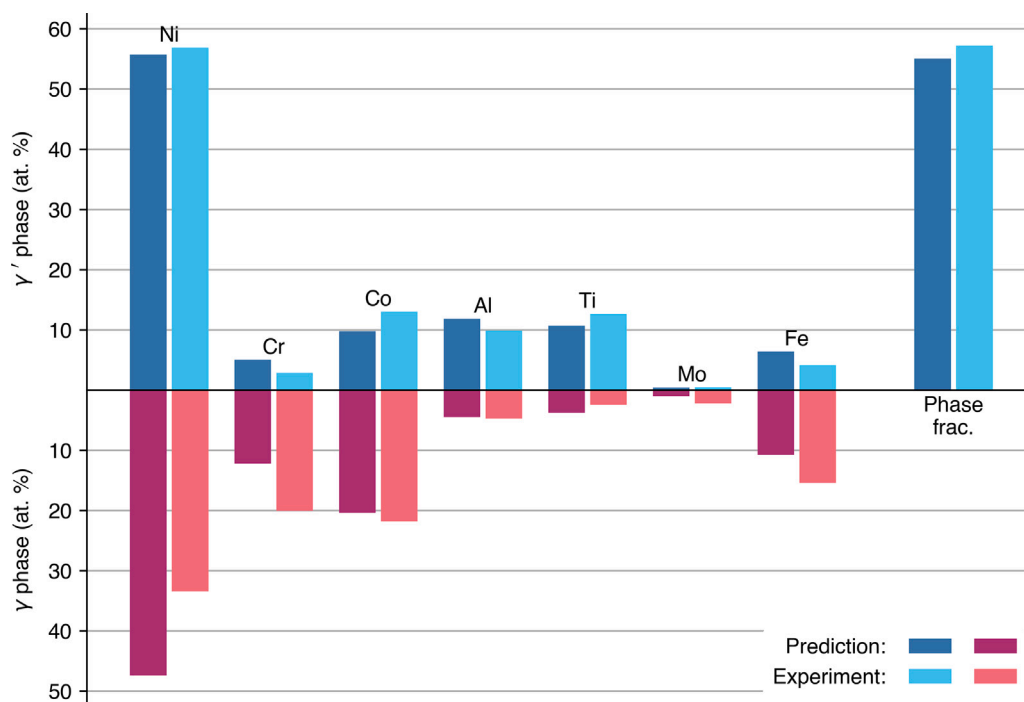


Fig. 5. Predictions for the phase composition of a high-entropy superalloy (HESA) [71] from the physical descriptor model.

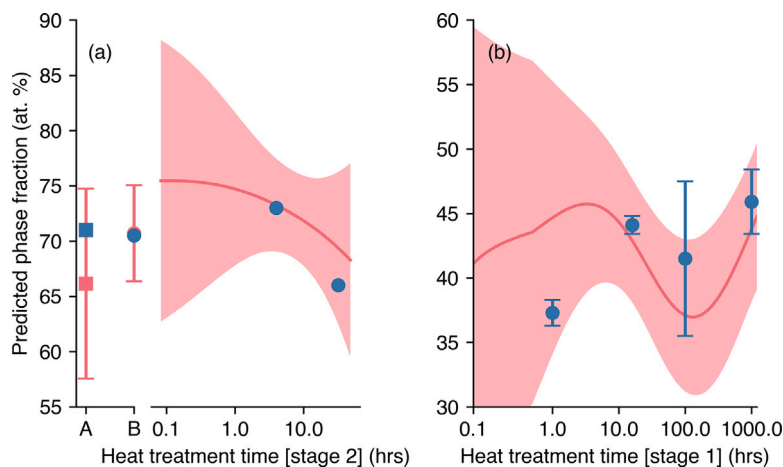


Fig. 6. Predicted (red points/series) and experimental (blue points) data for phase fraction evolution with applied heat treatments in two superalloys. The vertical bars and shaded region give the standard uncertainty in the model's predictions. (a) the commercial superalloy SRR99. Point A is for a sample that has only had a solid solution heat treatment applied, point B has undergone a single precipitation heat treatment at 870 °C, and all following data points have undergone the same 870 °C heat treatment preceded by a 1080 °C heat treatment for the time indicated [73]. (b) an experimental superalloy. Each alloy has undergone a single precipitation heat treatment at 760 °C for the number of hours indicated [74].

trend towards an increased  $\gamma'$  phase fraction with treatment time, but also the observed decrease in precipitate fraction at intermediate ageing times. It captured this qualitative trend despite the fact that it is opposite to that observed for the alloy SRR99; which furthermore was one of the only alloys in the training data with more than two different heat treatments applied to the same composition. Our new proposed physical description of heat treatments, Eq. (1), outperformed other descriptors for both test datasets, including a plain time and temperature descriptor and other Ostwald ripening-based descriptors, as well as giving better results for the five-fold cross-validation testing described above.

### 3.4. Creep rupture life model

GPR models for the creep rupture life of single crystal superalloys were trained. The accuracy of each model was assessed with ten-fold cross validation in the manner described in Section 3.1. As the creep rupture life data spans multiple orders of magnitude,  $R^2$  scores are given for both the actual values and log values. An ARD Matérn kernel was used in the Gaussian processes; but for the creep models, unlike those for microstructure, the optimal kernel smoothness parameter was found to be  $\nu = 1.5$ . The decreased smoothness of the kernel is likely because creep rupture life data spans multiple scales and testing regimes. For example, it is known that creep in single crystal

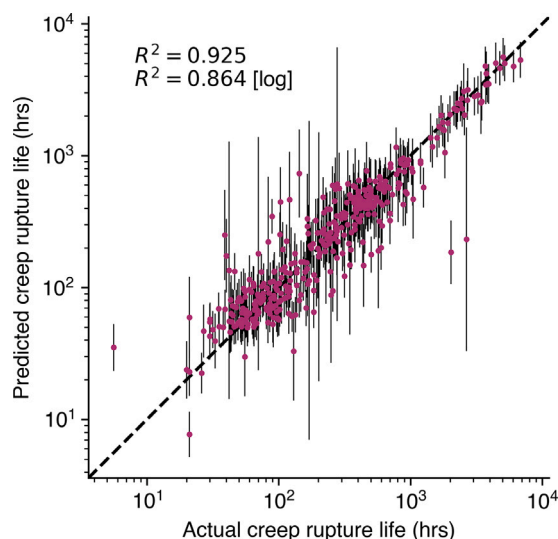


Fig. 7. Predicted creep rupture life from a descriptor model including microstructure-based descriptors. The vertical bars are the model's uncertainties.

Table 2

A summary of the coefficients of determination for GPR models of creep rupture life using different descriptors.

Descriptors	$R^2$
Physics + metallurgy-based descriptor method	0.864
Physics-based descriptor method	0.856
Plain method	0.840

superalloys occurs by two main mechanisms, dislocation and diffusion creep, with the former mechanism dominating at low temperatures and the latter at high temperatures.

Firstly, a plain composition descriptor GPR model was trained, which achieved  $R^2 = 0.840$ . Next, a model using a similar physics-based descriptor set to that used for microstructure, i.e. using descriptors calculated solely from each alloy's nominal overall composition, was trained. This delivered an improvement on the plain descriptor method, achieving  $R^2 = 0.856$  (see Table 2). This physics-based descriptor set was then further refined to include more high-level domain knowledge, summarised in Fig. 8. Metallurgy-based descriptors developed in Section 2.5 were added, some of which were calculated using phase compositions and fractions predicted by the pre-trained microstructure model described in the preceding sections. This model achieved a further improvement of  $R^2 = 0.864$  (Fig. 7).

The relevance of various descriptors in this model as determined by the ARD kernel lengthscales are given in Fig. 8. Many of the expected important descriptors are found to be relevant: the precipitate fraction, the overall mean interdiffusivity, and the mean melting point of the precipitate phase are selected. However, other descriptors that were expected to be highly relevant are not selected, including the lattice misfit, mean metal d-level, and the matrix phase stacking fault energy (SFE). The mean metal d-level is a metric used to estimate susceptibility to TCP phase formation during creep: commercial single-crystal superalloys have long been designed to avoid this behaviour, which will limit the influence of this descriptor for a model fitted to a dataset of largely commercial superalloys. The matrix SFE descriptor is unlikely to be sufficiently accurate: our estimate used an average of the difference between zero-point formation energies. This neglects thermal effects and uses an approximation to the SFE in fcc alloys. The lattice misfit is likely found to be irrelevant for both of those reasons. Of the non-metallic trace elements, the B and C at. % are relevant, whereas Y is not. The same heat treatment descriptors used for the microstructure modelling were again used, but the Ostwald ripening inspired descriptors

that were key to capturing microstructure evolution with ageing are not found to be relevant (although it should be noted that they indirectly enter the model via the predicted precipitate fraction descriptor). Three physics-based descriptors calculated from the nominal composition are found to be relevant. They are three descriptors that are also selected by the microstructure GPR model, so they may be relevant due to their influence on microstructure morphology, which is not explicitly modelled.

#### 4. Conclusion

In this work we have proposed a set of physical descriptors for composition and heat treatment for use in machine learning models of superalloy microstructure. A model using physical descriptors outperforms a model using plain composition descriptors when making *interpolative* predictions, see Table 1. Furthermore, when making *extrapolative* predictions, the model significantly outperforms the plain descriptor model, notably not suffering from such a severe overfitting effect (Fig. 4). Moreover, it can also make predictions for alloys containing elements that were not even present in its training dataset (Fig. 5). Such predictions are completely impossible to make using a plain descriptor model since they are not *element-agnostic*. This means our model can make useful predictions for cutting-edge superalloys, such as high-entropy superalloys or superalloys containing new heavy elements.

As well as standing up on its own, our model has a number of advantages over traditional CALPHAD. Previously identified benefits include the ability to easily retrain the GPR model, incorporate non-equilibrium features such as heat treatments, and most importantly its inherent quantification of uncertainties [10]. The model presented in this work captures both qualitative and quantitative effects of ageing on microstructure, see Fig. 6(a–b). A further benefit of our model is that it is possible to easily incorporate computational data, whilst still treating it as distinct from empirical data. This can be achieved by simply adding an extra feature to the inputs: a binary descriptor encoding the method by which an entry has been obtained. The process of fitting the Gaussian process will determine how relevant the computational composition data is to predicting 'true' compositions.

In addition to the microstructure model, we developed a GPR model for the creep rupture life of single crystal nickel superalloys. This model builds on that developed for microstructure in two ways. Firstly, it explicitly incorporates predictions of the microstructure model as descriptors, and finds them to be relevant to making predictions, providing a direct use-case for the usefulness of our GPR approach. Secondly, it expands on the domain knowledge paradigm by making use of descriptors that encapsulate metallurgical theories and principles. By incorporating such high-level domain knowledge, the ARD lengthscales of the fitted Gaussian process can be used to infer greater physical understanding about the properties it models. With a coefficient of determination of  $R^2 = 0.864$  for the log creep rupture lives, this model is itself useful for making predictions of this key strength property of single crystal superalloys. Like the microstructure model, its predictions include uncertainties, a crucial feature for its intended usage in alloy design [10,75,76].

Currently our model only calculates partitioning into two pre-determined phases, a use-case we have identified as most critical to superalloy design. Other authors have constructed machine learning models to classify alloys by phase stability [19]. This points towards development of a fully-fledged, generic, and probabilistic approach to the Calculation of Phase Diagrams problem. Beyond alloys, this approach could be extended to phase separation in other systems including polymer/polymer, polymer/filler, and aqueous two-phase mixtures [77–80].

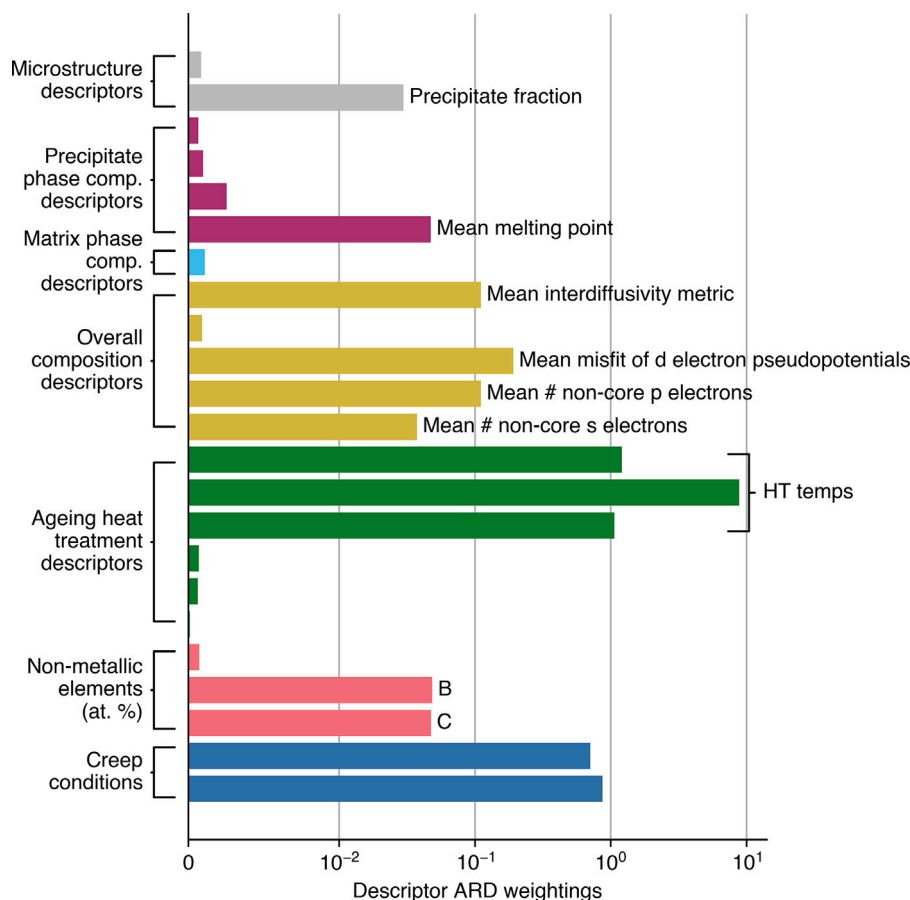


Fig. 8. Feature weightings (inverse of ARD lengthscales) determined for a creep rupture life model trained on the full dataset.

### CRedit authorship contribution statement

**Patrick L. Taylor:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft. **Gareth Conduit:** Conceptualization, Methodology, Writing – review & editing, Supervision.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Gareth Conduit reports a relationship with Intellegens Limited that includes: board membership.

### Data availability

The data is cited and a download link is included in the manuscript.

### Acknowledgements

Patrick Taylor acknowledges the financial support of EPSRC and an ICASE award from Dassault Systèmes UK. Gareth Conduit acknowledges the financial support of the Royal Society. We would also like to thank Dr. Victor Milman and Dr. Alexander Perlov at Dassault Systèmes UK for their feedback on the manuscript. There is Open Access to this paper and data available at <https://www.openaccess.cam.ac.uk/>.

### References

- [1] W. Hume-Rothery, H.M. Powell, On the theory of super-lattice structures in alloys, *Z. Kristallogr.* 91 (1–6) (1935) 23–47, <http://dx.doi.org/10.1524/ZKRI.1935.91.1.23>, URL <https://www.degruyter.com/document/doi/10.1524/zkri.1935.91.1.23/html>.
- [2] W. Hume-Rothery, *Atomic Theory for Students of Metallurgy*, 3d rev. reprint, in: Institute of Metals. Monograph and report series, no. 3, Institute of Metals, London, 1955.
- [3] D.G. Pettifor, A chemical scale for crystal-structure maps, *Solid State Commun.* 51 (1) (1984) 31–34, [http://dx.doi.org/10.1016/0038-1098\(84\)90765-8](http://dx.doi.org/10.1016/0038-1098(84)90765-8).
- [4] H. Glawe, A. Sanna, E.K. Gross, M.A. Marques, The optimal one dimensional periodic table: a modified Pettifor chemical scale from data mining, *New J. Phys.* 18 (9) (2016) 093011, <http://dx.doi.org/10.1088/1367-2630/18/9/093011>, URL <https://iopscience.iop.org/article/10.1088/1367-2630/18/9/093011> <https://iopscience.iop.org/article/10.1088/1367-2630/18/9/093011/meta>.
- [5] U.R. Kattner, The CALPHAD method and its role in material and process development, *Tecnol. Metalurgia Materiais E Mineração* 13 (1) (2016) 3–15, <http://dx.doi.org/10.4322/2176-1523.1059>, URL <http://tecnologiamm.com.br/doi/10.4322/2176-1523.1059>.
- [6] J.O. Andersson, T. Helander, L. Höglund, P. Shi, B. Sundman, Thermo-Calc & DICTRA, computational tools for materials science, CALPHAD, *Comput. Coupling Phase Diagr. Thermochem.* 26 (2) (2002) 273–312, [http://dx.doi.org/10.1016/S0364-5916\(02\)00037-8](http://dx.doi.org/10.1016/S0364-5916(02)00037-8).
- [7] S. Bajaj, A. Landa, P. Söderlind, P.E. Turchi, R. Arráyave, The U-Ti system: Strengths and weaknesses of the CALPHAD method, *J. Nucl. Mater.* 419 (1–3) (2011) 177–185, <http://dx.doi.org/10.1016/j.jnucmat.2011.08.050>.
- [8] H. Harada, H. Murakami, Design of Ni-base superalloys, in: T. Saito (Ed.), *Computational Materials Design*, Springer, Berlin, Heidelberg, 1999, pp. 39–70, [http://dx.doi.org/10.1007/978-3-662-03923-6\\_2](http://dx.doi.org/10.1007/978-3-662-03923-6_2).
- [9] Z. Qin, Z. Wang, Y. Wang, L. Zhang, W. Li, J. Liu, Z. Wang, Z. Li, J. Pan, L. Zhao, F. Liu, L. Tan, J. Wang, H. Han, L. Jiang, Y. Liu, Phase prediction of Ni-base superalloys via high-throughput experiments and machine learning, *Mater. Res. Lett.* 9 (1) (2021) 32–40, [http://dx.doi.org/10.1080/21663831.2020.1815093/SUPPL\\_FILE/TMRL\\_A\\_1815093\\_SM0670.JPG](http://dx.doi.org/10.1080/21663831.2020.1815093/SUPPL_FILE/TMRL_A_1815093_SM0670.JPG), URL <https://www.tandfonline.com/doi/abs/10.1080/21663831.2020.1815093>.



- [10] P.L. Taylor, G. Conduit, Machine learning predictions of superalloy microstructure, *Comput. Mater. Sci.* 201 (2022) 110916, <http://dx.doi.org/10.1016/j.COMMATSCI.2021.110916>.
- [11] F. Tancret, H.K. Bhadeshia, D.J. MacKay, Design of a creep resistant nickel base superalloy for power plant applications: Part 1 - Mechanical properties modelling, *Mater. Sci. Technol.* 19 (3) (2003) 283–290, <http://dx.doi.org/10.1179/026708303225009788>, URL <https://www.tandfonline.com/action/journalInformation?journalCode=ymst20>.
- [12] B.D. Conduit, N.G. Jones, H.J. Stone, G.J. Conduit, Design of a nickel-base superalloy using a neural network, *Mater. Des.* 131 (2017) 358–365, <http://dx.doi.org/10.1016/j.matdes.2017.06.007>.
- [13] B.D. Conduit, T. Illston, S. Baker, D.V. Duggappa, S. Harding, H.J. Stone, G.J. Conduit, Probabilistic neural network identification of an alloy for direct laser deposition, *Mater. Des.* 168 (2019) 107644, <http://dx.doi.org/10.1016/j.matdes.2019.107644>.
- [14] L. Asensio Dominguez, R. Goodall, I. Todd, Prediction and validation of quaternary high entropy alloys using statistical approaches, 31, (10) 2015, pp. 1201–1206, <http://dx.doi.org/10.1179/1743284715Y.0000000019>, URL <https://www.tandfonline.com/doi/abs/10.1179/1743284715Y.0000000019>.
- [15] F. Tancret, I. Toda-Caraballo, E. Menou, P.E.J. Rivera Díaz-Del-Castillo, Designing high entropy alloys employing thermodynamics and Gaussian process statistical analysis, *Mater. Des.* 115 (2017) 486–497, <http://dx.doi.org/10.1016/J.MATDES.2016.11.049>.
- [16] D. Xue, D. Xue, R. Yuan, Y. Zhou, P.V. Balachandran, X. Ding, J. Sun, T. Lookman, An informatics approach to transformation temperatures of NiTi-based shape memory alloys, *Acta Mater.* 125 (2017) 532–541, <http://dx.doi.org/10.1016/J.ACTAMAT.2016.12.009>.
- [17] E. Menou, I. Toda-Caraballo, P.E.J. Rivera-Díaz-del Castillo, C. Pineau, E. Bertrand, G. Ramstein, F. Tancret, Evolutionary design of strong and stable high entropy alloys using multi-objective optimisation based on physical models, statistics and thermodynamics, *Mater. Des.* 143 (2018) 185–195, <http://dx.doi.org/10.1016/J.MATDES.2018.01.045>.
- [18] O. Mamun, M. Wenzlick, J. Hawk, R. Devanathan, A machine learning aided interpretable model for rupture strength prediction in Fe-based martensitic and austenitic alloys, *Sci. Rep.* 11 (1) (2021) 1–9, <http://dx.doi.org/10.1038/s41598-021-83694-z>, URL <https://www.nature.com/articles/s41598-021-83694-z>.
- [19] K. Lee, M.V. Ayyasamy, P. Delsa, T.Q. Hartnett, P.V. Balachandran, Phase classification of multi-principal element alloys via interpretable machine learning, *Npj Comput. Mater.* 8 (1) (2022) 1–12, <http://dx.doi.org/10.1038/s41524-022-00704-y>, URL <https://www.nature.com/articles/s41524-022-00704-y>.
- [20] R.J. Murdock, S.K. Kauwe, A.Y.T. Wang, T.D. Sparks, Is domain knowledge necessary for machine learning materials properties? *Integr. Mater. Manuf. Innov.* 9 (3) (2020) 221–227, <http://dx.doi.org/10.1007/s40192-020-00179-z>, URL <https://link.springer.com/article/10.1007/s40192-020-00179-z>.
- [21] G.L. Hart, T. Mueller, C. Toher, S. Curtarolo, Machine learning for alloys, *Nat. Rev. Mater.* 6 (8) (2021) 730–755, <http://dx.doi.org/10.1038/s41578-021-00340-w>, URL <https://www.nature.com/articles/s41578-021-00340-w>.
- [22] M. Durand-Charre, *The Microstructure of Superalloys*, CRC Press, Boca Raton, FL, USA, 1997.
- [23] R.C. Reed, *The Superalloys Fundamentals and Applications*, Cambridge University Press, Cambridge, UK ; New York, 2006.
- [24] D.J. Crudden, A. Mottura, N. Warnken, B. Raesinia, R.C. Reed, Modelling of the influence of alloy composition on flow stress in high-strength nickel-based superalloys, *Acta Mater.* 75 (2014) 356–370, <http://dx.doi.org/10.1016/j.actamat.2014.04.075>.
- [25] E. Fleischmann, M.K. Miller, E. Affeldt, U. Glatzel, Quantitative experimental determination of the solid solution hardening potential of rhenium, tungsten and molybdenum in single-crystal nickel-based superalloys, *Acta Mater.* 87 (2015) 350–356, <http://dx.doi.org/10.1016/j.actamat.2014.12.011>.
- [26] M. Dodaran, A.H. Eftefagh, S.M. Guo, M.M. Khonsari, W.J. Meng, N. Shamsaei, S. Shao, Effect of alloying elements on the  $\gamma'$  antiphase boundary energy in Ni-base superalloys, *Intermetallics* 117 (2020) <http://dx.doi.org/10.1016/j.intermet.2019.106670>.
- [27] C.T. Sims, A history of superalloy metallurgy for superalloy metallurgists, in: *Superalloys*, 1984.
- [28] S. Tin, A.C. Yeh, A.P. Ofori, R.C. Reed, S.S. Babu, M.K. Miller, Atomic partitioning of ruthenium in Ni-based superalloys, *Tech. rep.*, *Superalloys 2004: 10th International Symposium on Superalloys*, 2004.
- [29] R.C. Reed, A.C. Yeh, S. Tin, S.S. Babu, M.K. Miller, Identification of the partitioning characteristics of ruthenium in single crystal superalloys using atom probe tomography, *Scr. Mater.* 51 (4) (2004) 327–331, <http://dx.doi.org/10.1016/j.scriptamat.2004.04.019>.
- [30] R.A. Hobbs, L. Zhang, C.M. Rae, S. Tin, The effect of ruthenium on the intermediate to high temperature creep response of high refractory content single crystal nickel-base superalloys, *Mater. Sci. Eng. A* 489 (1–2) (2008) 65–76, <http://dx.doi.org/10.1016/j.msea.2007.12.045>.
- [31] N. Tsuno, K. Kakehi, C.M. Rae, R. Hashizume, Effect of ruthenium on creep strength of Ni-base single-crystal superalloys at 750 °C and 750 MPa, *Metall. Mater. Trans. A* 40 (2) (2009) 269–272, <http://dx.doi.org/10.1007/s11661-008-9744-6>.
- [32] X.G. Wang, J.L. Liu, T. Jin, X.F. Sun, The effects of ruthenium additions on tensile deformation mechanisms of single crystal superalloys at different temperatures, *Mater. Des.* 63 (2014) 286–293, <http://dx.doi.org/10.1016/j.matdes.2014.06.009>.
- [33] T. Tsao, Y. Chang, K. Chang, J. Yeh, M. Chiou, S. Jian, C. Kuo, W. Wang, H. Murakami, Developing new type of high temperature alloys—High entropy superalloys, *International Journal of Metallurgical & Materials Engineering* 1 (1) (2015) <http://dx.doi.org/10.15344/2455-2372/2015/107>.
- [34] J. Chen, X. Zhou, W. Wang, B. Liu, Y. Lv, W. Yang, D. Xu, Y. Liu, A review on fundamental of high entropy alloys with promising high-temperature properties, *J. Alloys Compd.* 760 (2018) 15–30, <http://dx.doi.org/10.1016/j.jallcom.2018.05.067>.
- [35] M. Joele, W.R. Matizamhuka, A review on the high temperature strengthening mechanisms of high entropy superalloys (HESA), *Materials* 14 (19) (2021) <http://dx.doi.org/10.3390/MA14195835>, URL <https://pmc/articles/PMC8510092/>, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8510092/>.
- [36] M. Detroits, P.D. Jablonski, S. Antonov, S. Li, Y. Ren, S. Tin, J.A. Hawk, Design and thermomechanical properties of a  $\gamma'$  precipitate-strengthened Ni-based superalloy with high entropy  $\gamma$  matrix, *J. Alloys Compd.* 792 (2019) 550–560, <http://dx.doi.org/10.1016/j.jallcom.2019.04.054>.
- [37] A.M. Manzoni, U. Glatzel, New multiphase compositionally complex alloys driven by the high entropy alloy approach, *Mater. Charact.* 147 (2019) 512–532, <http://dx.doi.org/10.1016/j.matchar.2018.06.036>.
- [38] H. Harada, K. Ohno, T. Yamagata, T. Yokokawa, M. Yamazaki, Phase calculation and its use in alloy design program for nickel-base superalloys, *Tech. rep.*, *Superalloys 1988*, 1988.
- [39] A.P. Ofori, C.J. Humphreys, S. Tin, C.N. Jones, A TEM study of the effect of platinum group metals in advanced single crystal nickel-base superalloys, *Tech. rep.*, *Superalloys 2004: 10th International Symposium on Superalloys*, 2004.
- [40] G.E. Fuchs, B.A. Boutwell, Modeling of the partitioning and phase transformation temperatures of an as-cast third generation single crystal Ni-base superalloy, *Mater. Sci. Eng. A* 333 (1–2) (2002) 72–79, [http://dx.doi.org/10.1016/S0921-5093\(01\)01825-1](http://dx.doi.org/10.1016/S0921-5093(01)01825-1).
- [41] S. Ma, L. Carroll, T.M. Pollock, Development of  $\gamma$  phase stacking faults during high temperature creep of Ru-containing single crystal superalloys, *Acta Mater.* 55 (17) (2007) 5802–5812, <http://dx.doi.org/10.1016/j.actamat.2007.06.042>.
- [42] P. Taylor, *Superalloy database: strength and microstructure*, 2022, <http://dx.doi.org/10.17863/CAM.92063>, URL <https://www.repository.cam.ac.uk/handle/1810/344679>.
- [43] M.A. Álvarez, L. Rosasco, N.D. Lawrence, Kernels for vector-valued functions: a review, *Foundations and Trends in Machine Learning* 4 (3) (2011) 195–266, <http://dx.doi.org/10.1561/22000000036>, URL <https://arxiv.org/abs/1106.6251v2>.
- [44] L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials, *Npj Comput. Mater.* 2 (1) (2016) 1–7, <http://dx.doi.org/10.1038/npjcompumats.2016.28>, URL <https://www.nature.com/articles/npjcompumats201628>.
- [45] J. Ling, E. Antono, S. Bajaj, S. Paradiso, M. Hutchinson, B. Meredig, B.M. Gibbons, Machine learning for alloy composition and process optimization, in: *Proceedings of the ASME Turbo Expo*, 6, American Society of Mechanical Engineers Digital Collection, 2018, <http://dx.doi.org/10.1115/GT2018-75207>, URL <https://citration.com>.
- [46] Y. Liu, J. Wu, Z. Wang, X.G. Lu, M. Avdeev, S. Shi, C. Wang, T. Yu, Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning, *Acta Mater.* 195 (2020) 454–467, <http://dx.doi.org/10.1016/j.actamat.2020.05.001>.
- [47] J.T. Waber, D.T. Cromer, Orbital radii of atoms and ions, *J. Chem. Phys.* 42 (12) (2004) 4116, <http://dx.doi.org/10.1063/1.1695904>, URL <https://aip.scitation.org/doi/abs/10.1063/1.1695904>.
- [48] *Electronegativity*, in: *CRC Handbook of Chemistry and Physics*, 3rd electronic ed., CRC Press: Taylor & Francis, Boca Raton, Fla., 2021.
- [49] I.M. Lifshitz, V.V. Slyozov, The kinetics of precipitation from supersaturated solid solutions, *J. Phys. Chem. Solids* 19 (1–2) (1961) 35–50, [http://dx.doi.org/10.1016/0022-3697\(61\)90054-3](http://dx.doi.org/10.1016/0022-3697(61)90054-3).
- [50] J.E. Morral, G.R. Purdy, Particle coarsening in binary and multicomponent alloys, *Scr. Metall. Mater.* 30 (7) (1994) 905–908, [http://dx.doi.org/10.1016/0956-716X\(94\)90413-8](http://dx.doi.org/10.1016/0956-716X(94)90413-8).
- [51] T. Philippe, P.W. Voorhees, Ostwald ripening in multicomponent alloys, *Acta Mater.* 61 (11) (2013) 4237–4244, <http://dx.doi.org/10.1016/J.ACTAMAT.2013.03.049>.
- [52] M. Mostafaei, S.M. Abbasi, M.A. Mostafaei, Improvement of  $\gamma'$  coarsening model in high  $\gamma'$  volume fraction Ni-base superalloys containing different Ta/W ratio, *J. Alloys Compd.* 885 (2021) 160938, <http://dx.doi.org/10.1016/J.JALLCOM.2021.160938>.
- [53] D.K. Duvenaud, *Automatic model construction with Gaussian processes (Ph.D. thesis)*, University of Cambridge, Cambridge, 2014.
- [54] J.R. Gardner, G. Pleiss, D. Bindel, K.Q. Weinberger, A.G. Wilson, GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration, *Adv. Neural Inf. Process. Syst.* 2018-December (2018) 7576–7586, <http://dx.doi.org/10.48550/arxiv.1809.11165>, URL <https://arxiv.org/abs/1809.11165v6>.

- [55] W. Xia, X. Zhao, L. Yue, Z. Zhang, Microstructural evolution and creep mechanisms in Ni-based single crystal superalloys: A review, *J. Alloys Compd.* 819 (2020) 152954, <http://dx.doi.org/10.1016/J.JALLCOM.2019.152954>.
- [56] J.X. Zhang, J.C. Wang, H. Harada, Y. Koizumi, The effect of lattice misfit on the dislocation motion in superalloys during high-temperature low-stress creep, *Acta Mater.* 53 (17) (2005) 4623–4633, <http://dx.doi.org/10.1016/J.ACTAMAT.2005.06.013>.
- [57] R.C. Reed, T. Tao, N. Warnken, Alloys-By-Design: Application to nickel-based single crystal superalloys, *Acta Mater.* 57 (19) (2009) 5898–5913, <http://dx.doi.org/10.1016/j.actamat.2009.08.018>.
- [58] S. Neumeier, J. Ang, R.A. Hobbs, C.M. Rae, H.J. Stone, Lattice misfit of high refractory ruthenium containing nickel-base superalloys, *Adv. Mater. Res.* 278 (2011) 60–65, <http://dx.doi.org/10.4028/WWW.SCIENTIFIC.NET/AMR.278.60>, URL <https://www.scientific.net/AMR.278.60>.
- [59] T. Chandra, M. Ionescu, D. Mantovani, Influence of stacking fault energy on creep mechanism of a single crystal nickel-based superalloy containing Re, *Mater. Sci. Forum* 706–709 (2012) 2474–2479, <http://dx.doi.org/10.4028/www.scientific.net/MSF.706-709.2474>, URL <https://www.scientific.net/MSF.706-709.2474>.
- [60] S. Tian, X. Zhu, J. Wu, H. Yu, D. Shu, B. Qian, Influence of temperature on stacking fault energy and creep mechanism of a single crystal nickel-based superalloy, *J. Mater. Sci. Technol.* 32 (8) (2016) 790–798, <http://dx.doi.org/10.1016/j.jmst.2016.01.020>.
- [61] J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, L.A. Constantin, X. Zhou, K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, 2007, <http://dx.doi.org/10.1103/PhysRevLett.100.136406>, <http://arxiv.org/abs/0711.0156>.
- [62] T. Grosdidier, A. Hazotte, A. Simon, Precipitation and dissolution processes in  $\gamma/\gamma'$  single crystal nickel-based superalloys, *Mater. Sci. Eng. A* 256 (1–2) (1998) 183–196, [http://dx.doi.org/10.1016/S0921-5093\(98\)00795-3](http://dx.doi.org/10.1016/S0921-5093(98)00795-3).
- [63] P. Caron, High  $\gamma'$  solvus new generation nickel-based superalloys for single crystal turbine blade applications, in: *Superalloys, 2000*, pp. 737–746.
- [64] N. Yukawa, M. Morinaga, Y. Murata, H. Ezaki, S. Inoue, High performance single crystal superalloys developed by the d-electrons concept, in: *Superalloys, 1988*.
- [65] R.C. Reed, A. Mottura, D.J. Crudden, Alloys-by-design: Towards optimization of compositions of nickel-based superalloys, in: *Superalloys, 2016*, pp. 15–23.
- [66] Y.T. Tang, C. Panwisawas, J.N. Ghoussoub, Y. Gong, J.W. Clark, A.A. Németh, D.G. McCartney, R.C. Reed, Alloys-by-design: Application to new superalloys for additive manufacturing, *Acta Mater.* 202 (2021) 417–436, <http://dx.doi.org/10.1016/j.actamat.2020.09.023>.
- [67] C.E. Campbell, W.J. Boettinger, U.R. Kattner, Development of a diffusion mobility database for Ni-base superalloys, *Acta Mater.* 50 (4) (2002) 775–792, [http://dx.doi.org/10.1016/S1359-6454\(01\)00383-4](http://dx.doi.org/10.1016/S1359-6454(01)00383-4).
- [68] A. Sato, H. Harada, T. Yokokawa, T. Murakumo, Y. Koizumi, T. Kobayashi, H. Imai, The effects of ruthenium on the phase stability of fourth generation Ni-base single crystal superalloys, *Scr. Mater.* 54 (9) (2006) 1679–1684, <http://dx.doi.org/10.1016/J.SCRIPTAMAT.2006.01.003>.
- [69] D. Argence, C. Vemault, Y. Desvallées, D. Fournier, MC-NG: A 4th generation single-crystal superalloy for future aeronautical turbine blades and vanes, in: *Superalloys, 2000*.
- [70] S. Walston, A. Cetel, R. MacKay, D. Duhl, R. Dreshfield, Joint development of a fourth generation single crystal superalloy, Tech. rep., NASA, 2004, URL <http://www.sti.nasa.gov>.
- [71] L. Zhang, Y. Zhou, X. Jin, X. Du, B. Li, Precipitation-hardened high entropy alloys with excellent tensile properties, *Mater. Sci. Eng. A* 732 (2018) 186–191, <http://dx.doi.org/10.1016/J.MSEA.2018.06.102>.
- [72] F. Zheng, G. Zhang, X. Chen, X. Yang, Z. Yang, Y. Li, J. Li, A new strategy of tailoring strength and ductility of CoCrFeNi based high-entropy alloy, *Mater. Sci. Eng. A* 774 (2020) 138940, <http://dx.doi.org/10.1016/J.MSEA.2020.138940>.
- [73] R. Schmidt, M. Feller-Kniepmeier, Effect of heat treatments on phase chemistry of the Nickel-Base superalloy SRR 99, *Metall. Trans. A* 23 (3) (1992) 745–757, <http://dx.doi.org/10.1007/BF02675552>.
- [74] A.J. Goodfellow, E.I. Galindo-Nava, K.A. Christofidou, N.G. Jones, C.D. Boyer, T.L. Martin, P.A. Bagot, M.C. Hardy, H.J. Stone, The effect of phase chemistry on the extent of strengthening mechanisms in model Ni-Cr-Al-Ti-Mo based superalloys, *Acta Mater.* 153 (2018) 290–302, <http://dx.doi.org/10.1016/J.ACTAMAT.2018.04.064>.
- [75] E.J. Martin, V.R. Polyakov, L. Tian, R.C. Perez, Profile-QSAR 2.0: Kinase virtual screening accuracy comparable to four-concentration IC50s for realistically novel compounds, *J. Chem. Inf. Model.* 57 (8) (2017) 2077–2088, <http://dx.doi.org/10.1021/acs.jcim.7b00166>.
- [76] B.W. Irwin, J.R. Levell, T.M. Whitehead, M.D. Segall, G.J. Conduit, Practical applications of deep learning to impute heterogeneous drug discovery data, *J. Chem. Inf. Model.* 60 (6) (2020) 2848–2857, <http://dx.doi.org/10.1021/acs.jcim.0c00443>, URL <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.0c00443>.
- [77] J. Zhu, X. Lu, R. Balieu, N. Kringos, Modelling and numerical simulation of phase separation in polymer modified bitumen by phase-field method, *Mater. Des.* 107 (2016) 322–332, <http://dx.doi.org/10.1016/J.MATDES.2016.06.041>.
- [78] Y.C. Li, C.P. Wang, X.J. Liu, Thermodynamic assessments of binary phase diagrams in organic and polymeric systems, *CALPHAD* 33 (2) (2009) 415–419, <http://dx.doi.org/10.1016/J.CALPHAD.2008.12.007>.
- [79] Y. Li, Thermodynamic and kinetic study of spinodal phase separation in heptane-phenol system, *CALPHAD* 50 (2015) 113–117, <http://dx.doi.org/10.1016/J.CALPHAD.2015.05.004>.
- [80] A. Salabat, S. Tiani Moghadam, M. Rahmati Far, Liquid–liquid equilibria of aqueous two-phase systems composed of TritonX-100 and sodium citrate or magnesium sulfate salts, *CALPHAD* 34 (1) (2010) 81–83, <http://dx.doi.org/10.1016/J.CALPHAD.2009.12.004>.