



Data-Driven Study of Shape Memory Behavior of Multi-Component Ni–Ti Alloys in Large Compositional and Processing Space

Shreyas J. Honrao¹ · Othmane Benafan² · John W. Lawson³

Received: 2 June 2022 / Revised: 31 October 2022 / Accepted: 10 November 2022 / Published online: 8 December 2022
© ASM International 2022

Abstract Shape memory alloys have found wide-spread use in aerospace, automotive, biomedical, and commercial applications owing to their favorable properties and ease of operation. Binary NiTi, in particular, is known for its remarkable shape memory properties, mechanical strength, ductility, corrosion resistance, and biocompatibility. These properties can be further enhanced and better controlled through alloying NiTi with ternary, quaternary, and higher-order elements. Recently, researchers at NASA have compiled an extensive database of shape memory properties of materials, including over 8000 multi-component Ni–Ti alloys containing 37 different alloying elements. Using the Ni–Ti dataset, we train machine learning models to explore shape memory behavior of Ni–Ti alloys over a large compositional and processing space. The models predict transformation temperatures, hysteresis, and transformation strain, with low mean absolute errors of 14.8 °C, 7.2 °C, and 0.36%, respectively. We use these models to map trends and learn relationships between shape memory behavior and different parameters in the input design space. They can be used to make predictions for any multi-component alloy, without need for additional training. The combination of an extensive experimental dataset and

accurate learning models, together, make our approach highly suitable for the discovery and design of new alloys with targeted properties.

Keywords SMA · NiTi · Machine learning · Transformation temperature · Hysteresis · Transformation strain

Introduction

Shape Memory Alloys (SMAs) have gained popularity over the last few decades owing to their excellent actuation capabilities, simple design, and their small size and weight requirements [1, 2]. A range of SMA systems including Ni–Ti, Cu–Al–Ni, Cu–Zn–Al, and Fe–Mn–Si have been studied for their use in sensors, actuators, and dampers, of which binary NiTi has garnered the most attention [3, 4]. NiTi is known for its stability of transformation temperatures, superior mechanical strength, ductility, corrosion resistance, and biocompatibility, making it an excellent choice for several aerospace, automotive, commercial, and biomedical applications [5–9]. However, pure NiTi only allows for a narrow range of operation, and other ternary and quaternary additions are necessary to access higher operating temperatures, larger transformation strains, or smaller hysteresis [10–13]. The most common alloying additions to NiTi include Pd, Pt, Hf, Au, Cu etc. Ternary systems such as Ni–Ti–Zr have emerged as promising alternatives due to their reduced cost and lower weight [14]. Besides, other constituents such as Cr, Mn, Fe, Co, V, etc. have also been added to NiTi to control its properties.

For their use in most applications, it is vital to understand the transformation behavior of SMAs. The martensitic transformation temperature of these alloys is the most

✉ Shreyas J. Honrao
shreyas.j.honrao@nasa.gov

✉ John W. Lawson
john.w.lawson@nasa.gov

¹ KBR Inc., Intelligent Systems Division, NASA Ames Research Center, Moffett Field, CA 94035, USA

² Materials and Structures Division, NASA Glenn Research Center, Cleveland, OH 44135, USA

³ Intelligent Systems Division, NASA Ames Research Center, Moffett Field, CA 94035, USA

critical as it dictates the temperature window for the particular application. Transformation strain, which is calculated as the strain recovery due to the martensite-to-austenite transformation upon heating, determines the ability of the alloy to provide work output. A low hysteresis is essential to increase the energy conversion efficiency in engineering applications; a direct correlation between large hysteresis and poor functional fatigue properties has been established [15].

Besides performing actual experiments, there are also two computational approaches generally used to study transformation behavior in SMAs: *ab initio* approaches based on thermodynamic integration or the P_4 method, [16, 17] and classical molecular dynamics approaches that require the development of interatomic potentials [18, 19]. Of the two, *ab initio* approaches are more accurate. However, they are still limited in their present capabilities and tend to be computationally expensive. Most engineering alloys are composed of three or more elements; first principles calculations of such alloys are limited to extremely small unit cell sizes (< 1000 atoms) and timescales (< 1 ns). Further, although they provide a good estimation of transformation temperatures, these techniques cannot be used to estimate hysteresis or transformation strain. They also do not account for processing history, even though it is well known that many alloys do not exhibit shape memory behavior without prior thermo-mechanical processing [5]. As a result, they are uncondusive to large systematic studies over wide compositional and processing spaces. For the current task, we take a different approach of using a dataset of previously compiled experimental measurements and training machine learning (ML) models to predict the shape memory behavior of new alloys.

ML has the potential to greatly accelerate the search for novel and useful materials by effectively learning underlying structure–processing–property relationships from existing data, and efficiently scanning high-dimensional composition and processing spaces using these learned relationships. Many different materials properties like formation enthalpies, adsorption energies, elastic moduli, band gaps, diffusion barriers, glass forming ability, etc. have been predicted accurately using experimental and computational databases [20–28]. Most ML studies combine physics-inspired feature sets with learning techniques like tree-based ensemble methods, neural networks, or support vector machines, to predict accurate properties from the underlying composition or structure. It is postulated that the development of such physically motivated features has improved the predictive ability of ML models by at least two orders of magnitude in the past few years [29].

Lately, there has been great interest in using ML approaches to model relationships between the

composition and properties of SMAs. In a recent paper, Liu et al. used a physics-informed ML approach to predict the mean transformation temperature and thermal hysteresis using a dataset of about 500 Ni–Ti–Hf alloys with Hf in the 0–30 atomic % range [30]. In another work, using just 53 Ni–Ti–Fe–Pd–Cu alloys in their training set, Xue et al. developed a three-feature model to also predict transformation temperatures [31]. Trehern et al. used both composition and processing features to build a materials informatics framework to identify SMA chemistries and associated thermo-mechanical treatments that result in narrow transformation range and hysteresis [32]. In each study, the compositional space being explored was restricted to three to five elements. This limits the predictive capability of the trained models to alloy compositions within that space exclusively.

In contrast, we use a dataset of over 8000 multi-component Ni–Ti alloys containing 37 different alloying elements to train ML models for predicting transformation temperatures, hysteresis, and transformation strain. Although predictions are made for all four transformation temperatures—Martensite Start (M_S), Martensite Finish (M_F), Austenite Start (A_S), and Austenite Finish (A_F)—we mostly refer to the A_F model in this work, given that the performance of the other three models is also quite similar. Hysteresis is defined as the difference between the A_F and M_S temperatures ($A_F - M_S$). Instead of training a separate model, individual predictions from the A_F and M_S models are used for predicting hysteresis. Confidence intervals are computed to accompany model predictions and assist experimental validation. Together, these models allow us to study trends and learn relationships between shape memory behavior and different parameters in the input design space. To the best of our knowledge, no other current approach is capable of predicting SMA transformation behavior over such a wide range of compositions and processing conditions.

Dataset

The two main requirements for a good ML model are a representative dataset and an effective learning algorithm. For the current task, we use the extensive shape memory materials database compiled by Benafan et al. [33], containing experimentally measured shape memory properties mined from the literature. The database contains over 8000 Ni–Ti alloys, with compositions ranging from binary NiTi to quinary systems, which constitutes our training dataset. Fig. S1 of the Supplementary Information shows a histogram of the number of components for alloys within the dataset. In addition to the four transformation temperatures, hysteresis, and transformation strain, the dataset also

contains columns with information about alloy compositions, melting techniques, heat treatments, cycling temperatures, cycle number, heating and cooling rates, and applied stress, all of which can be used as inputs to the ML models. While popular systems like Ni–Ti–Pd, Ni–Ti–Pt, Ni–Ti–Hf, and Ni–Ti–Cu are more heavily represented in the dataset, those with less common constituents like Al, Nb, Ta, Sn, etc. are also present. A boxplot of the distribution of alloys by element is shown in Fig. 1. Although the figure provides a rough estimate of the range of A_F temperatures accessible through the addition of various elements, it should be noted that only alloys within the dataset are represented. By training ML models on such a vast dataset, predictions for any multi-component alloy within the large compositional and processing space being explored are possible.

To prepare the data for ML, a cleaning step is first performed to exclude alloys with missing compositions or transformation temperatures, and to eliminate duplicate entries. A few alloys with similar compositions and heat treatments, but different measured temperatures, remain. These are likely separate measurements of the same alloy by the same or different group, and all such entries are retained. Additionally, outliers are not discarded even though the possibility of erroneous measurements, reporting, and data collection always persists. We are left with 7988 alloys with at least one reported transformation temperature. 5211 of them have all four temperatures reported, while 7305 have M_S , 5390 have M_F , 6032 have A_S , and 6513 have A_F temperatures. In order to maximize data utilization for learning purposes, individual ML models are trained for each of the four temperatures. Hysteresis is defined as $A_F - M_S$, and there are 5961 alloys for which both values are reported. Transformation strains are only measured for 2350 alloys, which are used for training the corresponding ML model.

Feature Selection

Next, input features are identified from the available columns to predict transformation behavior. These can be divided into three groups: (i) composition related features, (ii) thermo-mechanical processing-related features, and (iii) test parameters.

Composition

For the purpose of representing compositions of alloys, the percentage of elements in the alloy are used as inputs. For example, a $Ni_{25}Pd_{25}Ti_{50}$ alloy is represented with 25, 25, and 50 for Ni, Pd, and Ti and 0 for every other element. We also tried using composition-weighted element property features based on the MAGPIE descriptor set [28] but did not notice any improvements in performance.

Thermo-Mechanical Processing

Although compositions of Ni, Ti, and other alloying elements largely determine transformation behavior of SMAs, many alloys do not even exhibit shape memory effect without prior thermo-mechanical processing. Thus, processing history of alloys is as important as composition for studying transformation behavior. Processing primarily includes the melting or preparation method as well as the subsequent thermo-mechanical treatments performed on the alloy.

Alloys can be prepared through many techniques—vacuum induction melting (VIM), vacuum arc remelting (VAR), sputtering, melt spinning, powder metallurgy, etc. Different preparation techniques result in alloys with slightly different end compositions, eventually affecting the mechanical and functional properties of the final alloy. For example, while the starting and end compositions are quite similar for VAR, large variances are possible with

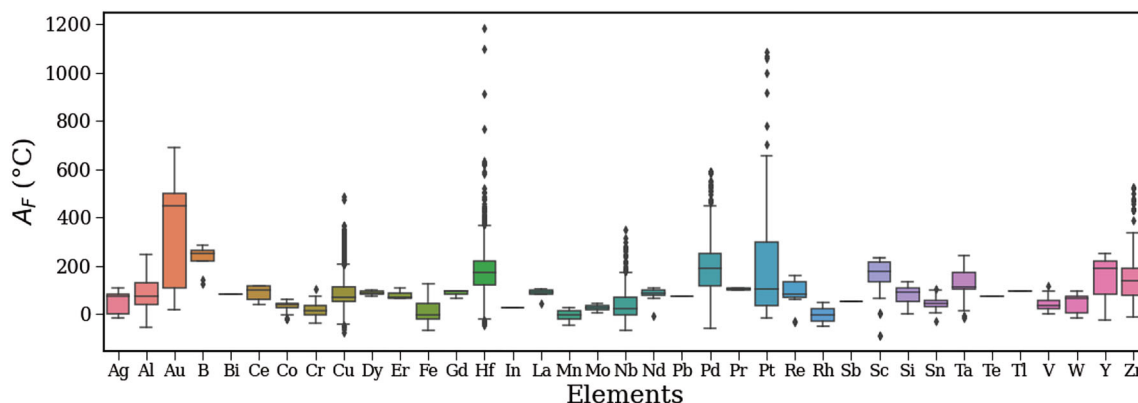


Fig. 1 Boxplot showing the distribution of Ni–Ti alloys in the dataset by alloying element and A_F temperatures. In all, 37 total alloying elements are represented in the dataset

VIM. Other impurities like C, N, O, etc. can also be introduced into the melt, depending on the choice of technique. To account for these variations, one-hot encoded features representing different techniques are included as inputs to the ML models. Thus, an alloy prepared using VIM will have a value of 1 for VIM, and 0 for all other techniques. To limit the total number of features, and to avoid sparse inputs, only the five most popular techniques from the dataset have corresponding one-hot encoded features. Alloys prepared using other techniques, or without a reported technique, are encoded using a 1 in the “Others” column.

Following the melting stage, SMAs typically undergo up to three different heat treatments: homogenization, solutionizing, and aging. Homogenization at a temperature below the solidus, followed by quenching, allows for a uniform chemical composition throughout the solid solution. Quenching helps prevent the formation of undesirable precipitates in the alloy. Often times, a solutionizing step is performed to put precipitates that may have formed back into the solid solution. Finally, aging below the solvus, performed for non-stoichiometric alloys, introduces useful second-phase precipitates such as Ti_3Ni_4 , H-phase, or P-phase in Ni–Ti alloys. The size and distribution of secondary particles can change depending on the exact aging conditions, resulting in different transformation behaviors. The three heat treatments are accounted for in our models through the introduction of six features—homogenization temperature and time, solutionizing temperature and time, aging temperature and time.

Test Parameters

Inputs grouped under test parameters include the applied stress, cycling temperatures, cycle number, and heating and cooling rates. The applied stress value indicates the magnitude of external stress, when present, and is used as one of the inputs to the ML models. Cycling temperatures and heating or cooling rates are tested as possible features; however, they show no appreciable effect on the predictive ability and are excluded from the final models. Lastly, even though it is known that the thermal response of an alloy drifts to lower temperatures as a function of cycle number [34], very few alloys in the dataset have a reported cycle number, making it an ineffective input for our ML approach.

This leaves us with a total of 52 selected features as input for training the ML models. These are listed in Table 1. Although latent factors like internal stress, short range order, size and distribution of precipitates, etc. actually control transformation behavior of SMAs, we can only capture their influence indirectly, through input variables like composition, preparation technique, heat

treatments, and applied stress. Further, since original data are collected from multiple publications, all of these values may not always be reported. For features with missing values, appropriate imputations are performed. Where no heat treatment was carried out or reported, we use the room temperature value of 25 °C and 0 h for temperature and time, respectively. Similarly, where an applied stress was not reported, we assume a stress-free measurement at 0 MPa.

Learning Algorithm and Model Assessment

Several different learning algorithms have been used to tackle materials science problems over the last few years, including Gaussian process regression, support vector regression, random forest regression, gradient boosting regression, artificial neural networks, etc. Each has its own advantages and applicability domains. Here, we use a tree-based ensemble algorithm called Extremely Randomized Trees (ExRT) [35].

Ensemble methods combine predictions from several base estimators to reduce bias and variance, thus, improving performance. They are extremely effective for small- and medium-sized datasets that have a mix of categorical and continuous features spanning various scales. Ensemble algorithms also do not require scaling of input features. ExRT is a bagging technique, similar to random forest regression, where the optimum splits for the individual trees are chosen at random to further limit overfitting and, thus, improve prediction accuracy. We use the scikit-learn [36] implementation of ExRT for our current work. Although algorithms like ExRT are capable of interpolating extremely well in high-dimensional input feature spaces, their extrapolative capacities are known to be unsatisfactory. One must be careful when extending these models outside their range of applicability.

After training the ML models, their predictive abilities are analyzed by performing 10-fold cross validation (CV) over the entire dataset. All CV errors are computed on unseen data only and then averaged over 25 randomized splits of the dataset. The mean absolute error (MAE) is our preferred scoring metric as it penalizes all errors equally. In contrast, the root-mean-squared error (RMSE) penalizes larger errors more and is, therefore, more sensitive to outliers, which are not excluded in this work. In addition to the MAE, we also report the cross-validated R^2 values for each model.

Along with predicting transformation temperatures, hysteresis, and transformation strain, these models are also equipped to compute confidence intervals to accompany predictions. Since data acquired to train ML models are usually limited and fail to capture the true distributions,

Table 1 List of 52 input features selected for training the ML models

| Compositions | Processing conditions | Test parameters |
|-------------------------------------|--------------------------------------|----------------------|
| Ni atomic % | Preparation technique [\times 6] | Applied stress (MPa) |
| Ti atomic % | Homogenization temp. ($^{\circ}$ C) | |
| Alloy elem. atomic % [\times 37] | Homogenization time (hrs) | |
| | Solutionizing temp. ($^{\circ}$ C) | |
| | Solutionizing time (hrs) | |
| | Aging temp. ($^{\circ}$ C) | |
| | Aging time (hrs) | |

uncertainties are introduced in the model parameters which propagate to the predictions. We use the jackknife approach [37] to compute uncertainties across all ML models using the MAPIE [38] python package which is based on techniques introduced by Barber et al. [39]. Using the confidence intervals along with actual predictions allows us to select the best alloys for experimental validation. All prediction figures in this work show one sigma or 68% confidence intervals.

Permutation feature importance analysis (PFI) [40] helps us to explain outputs from the ML models. PFI measures how the model R^2 changes when a single input feature is randomly permuted; a higher score reflects a higher dependence of the model on the corresponding feature. We use the rfpimp [41] python package to calculate PFI for the models. Since PFI is affected by the presence of highly correlated features, similar features are clustered into groups as discussed in “Results and Discussion.”

Results and Discussion

Using the selected input features, ExRT models are trained to predict transformation temperatures (M_S , M_F , A_S , A_F), hysteresis, and transformation strain. The MAE and R^2 for each model are computed by averaging 10-fold CV errors over 25 randomized splits of the dataset. These are reported in Table 2. The MAE for the M_S (13.6 $^{\circ}$ C), M_F (15.9 $^{\circ}$ C), A_S (15.6 $^{\circ}$ C), and A_F (14.8 $^{\circ}$ C) models are in a similar range, as are their R^2 values (0.92–0.95). For hysteresis, computed indirectly using the A_F and M_S predictions, the MAE is 7.2 $^{\circ}$ C. The average error for the transformation strain model is 0.36%. Fig. S2 of the Supplementary Information shows parity plots for each model. The inherent noise in the experimental data results in lower R^2 values and some scatter around the diagonal, particularly for the hysteresis and transformation strain models.

After initial training and evaluation, the ML models are subsequently employed to study the effects of composition and processing on shape memory behavior. First, we explore how the transformation temperature (A_F) and

Table 2 MAE and R^2 values for the different ML models reported as 10-fold cross-validation errors that are averaged over 25 randomized splits of the dataset

| Property | Training set | MAE | R^2 |
|---|--------------|------|-------|
| Martensite start temperature (M_S) | 7305 | 13.6 | 0.95 |
| Martensite finish temperature (M_F) | 5390 | 15.9 | 0.92 |
| Austenite start temperature (A_S) | 6032 | 15.6 | 0.94 |
| Austenite finish temperature (A_F) | 6513 | 14.8 | 0.94 |
| Hysteresis ($A_F - M_S$) ^a | 5961 | 7.2 | 0.65 |
| Transformation strain | 2350 | 0.36 | 0.81 |

^aMAE for the hysteresis model is computed using the predictions from the A_F and M_S models

hysteresis for stoichiometric ternary alloys, where either Ni or Ti concentration is fixed at 50 atomic %, varies with increasing ternary element concentration. To isolate the effects of composition on transformation behavior, input parameters for processing are largely fixed while making the predictions. VAR is selected as the preparation technique, with a homogenization temperature and time of 1050 $^{\circ}$ C for 24 h, no solutionizing or aging, and no applied stress. Figures 2 and 3 show predictions for Ni–Ti–Pd and Ni–Ti–Hf. Similar predictions for Ni–Ti–Pt and Ni–Ti–Zr are available as Figs. S3 and S4 in the Supplementary Information. Here, Pd (Pt) replaces Ni in NiTi, whereas Hf (Zr) substitutes Ti. The actual A_F and hysteresis predictions are fit to fourth degree polynomials for smoothness and better visualization (dotted lines). The computed confidence intervals are one standard deviation (68%) away and shown as shaded regions around the dotted line. There is good agreement between predictions and experimental data (solid triangles) for the four ternary systems.

The A_F predictions for Ni–Ti–Pd in Fig. 2a show a slight dip around 8–10 atomic %, before increasing again, representing the change in phase from B19' (monoclinic) to B19 (orthorhombic). [42] This change also manifests itself in the hysteresis curve in Fig. 2b, which shows decreasing values in the B19' region and then a flattening out to very low hysteresis temperatures in the orthorhombic

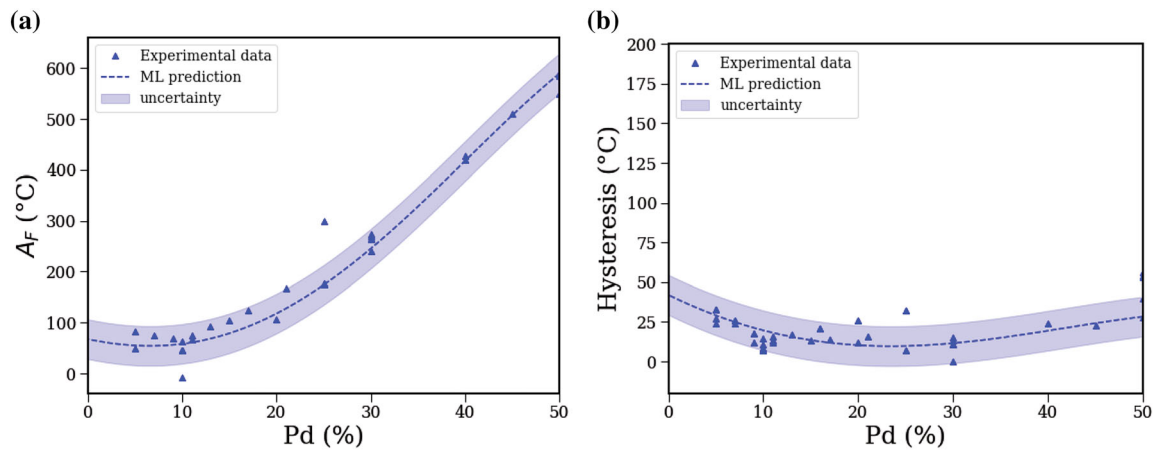


Fig. 2 ML predictions for A_F and hysteresis of stoichiometric Ni–Ti–Pd alloys. Dotted lines are fourth degree polynomials fit to actual model predictions. Shaded regions represent confidence intervals at one standard deviation (68%)

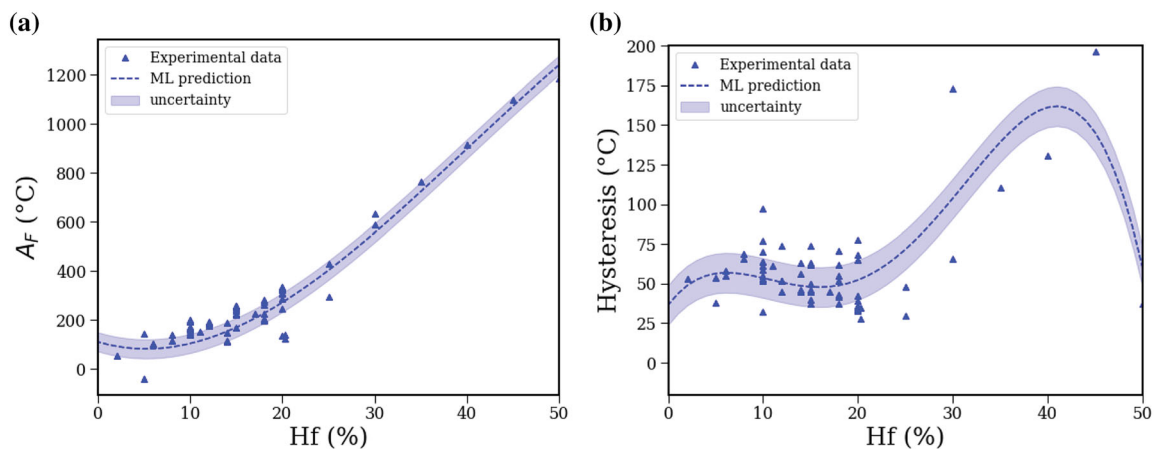


Fig. 3 ML predictions for A_F and hysteresis of stoichiometric Ni–Ti–Hf alloys. Dotted lines are fourth degree polynomials fit to actual model predictions. Shaded regions represent confidence intervals at one standard deviation (68%)

B19 phase. [11] The same B19' to B19 phase change is observed in Ni–Ti–Pt and Ni–Ti–Au ternaries as well [5]. In contrast, Fig. 3b shows consistently higher hysteresis temperatures for Ni–Ti–Hf, evidence of B19' presence, and a distinct uptrend between 20% and 40–45% Hf. Since the dataset contains only a handful of high Hf alloys, they carry a larger weight in the model and have a greater influence on the shape of the prediction curve. It is interesting to note that although the B19' phase is associated with higher hysteresis, it permits for larger transformation strains and work outputs compared to B19.

In addition to common ternary elements like Pd, Pt, and Hf, the effects of other alloying additions such as Cr, Fe, Co, to NiTi are also studied. These elements are typically added in smaller quantities (up to 3 atomic %) and the available experimental data for these systems are fewer and noisier. Hence, a simple linear fit is used to represent model predictions for these systems. Figure 4a reveals that only Ni–Ti–Cr experimental data lie perfectly on the prediction

line, whereas other alloys show plenty of scatter. All three elements, however, seem to lower A_F , and can, thus, be useful additions for tuning transformation temperatures of Ni–Ti alloys.

We make similar predictions for ternary Ni–Ti alloys containing Ag, Sn, Nb, and Ta. These elements are typically present in the 0–10 atomic % range and seem to have minimal effect on the A_F temperatures when compared to binary NiTi. Supplementary Information Fig. S5 shows the linear fit of ML predictions for these systems. While the A_F temperatures are within the same narrow range for all four, their hysteresis values show slightly larger variation. Ag, in particular, can potentially be a useful additive for lowering the hysteresis in Ni–Ti alloys, while keeping the A_F constant.

Besides manipulating the compositions of alloying elements, another way to control transformation behavior in SMAs is through the addition of excess Ni. Adding just 0.1% Ni to near-stoichiometric Ni–Ti alloys can lower

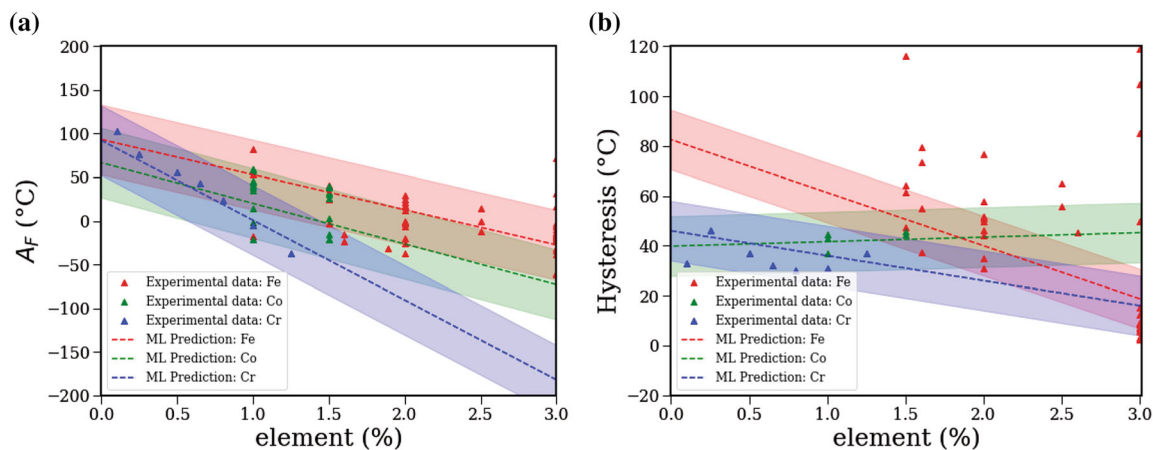


Fig. 4 ML predictions for A_F and hysteresis of stoichiometric Ni–Ti–Fe, Ni–Ti–Co, and Ni–Ti–Cr alloys. Dotted lines are linear fits to actual model predictions. Shaded regions represent confidence intervals at one standard deviation (68%)

transformation temperatures by as much as 20 °C [43]. A comparison between predictions for stoichiometric and slightly off-stoichiometric Ni–Ti–Pd and Ni–Ti–Hf alloys is shown in Fig. 5. There is a clear downward displacement of the blue Ni-rich curve for both systems. However, alloys with more than 1–2% excess Ni are known to show much lower transformation strains and are, thus, far less attractive for actuation applications. In contrast, the addition of excess Ti has minimal effect on the transformation temperatures. Figure 5 shows that A_F slightly increases, or remains constant, for Ti-rich alloys. Excess Ti, however, can lead to the formation of detrimental oxides and other phases in the alloy, degrading workability [44, 45]. Hence, Ti-rich alloys are rarely used in practice even when transformation temperatures are within the desired range.

The effects of excess Ni can be counteracted through the aging heat treatment. For non-stoichiometric alloys, aging helps precipitate secondary phase particles that are rich in

Ni, into the austenite B2 matrix. The size and distribution of the precipitates are controlled by adjusting the aging temperature and time. Coherent nanometer-sized precipitates can impede dislocations, lead to strengthening, and thus, increase transformation temperatures. The effect is compounded because the precipitates also contribute to the removal of excess Ni from the alloy. Ti_3Ni_4 in binary NiTi, H-phase [46, 47] in Ni–Ti–Hf and Ni–Ti–Zr alloys, and P-phase [48] in Ni–Ti–Pd and Ni–Ti–Pt alloys, are all examples of secondary phases. In slightly Ni-rich Ni–Ti–Hf alloys, the precipitation of H-phase particles in the matrix is found to impart excellent mechanical strength and functional ability to the alloy. [49]

Figure 6 shows the effects of aging on non-stoichiometric Ni–Ti–Pd and Ni–Ti–Hf alloys. While A_F predictions for slightly Ni-rich alloys aged at 550 °C (Pd) and 650 °C (Hf) are generally higher compared to unaged alloys, the differences are smaller than experimentally

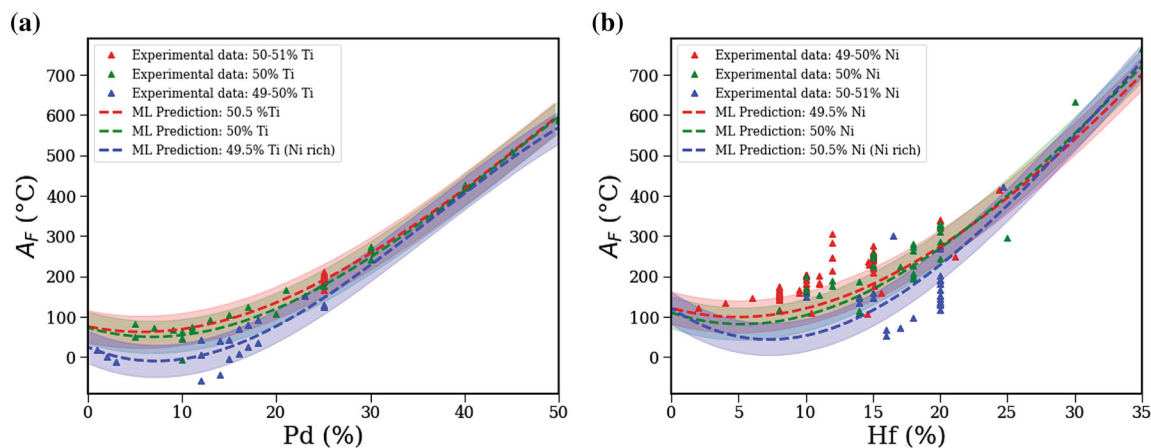


Fig. 5 Comparing the effects of Ni and Ti additions on A_F predictions of Ni–Ti–Pd and Ni–Ti–Hf alloys. Slightly Ni-rich alloys (blue) have lower transformation temperatures whereas alloys with

excess Ti (red) have similar, or slightly higher, transformation temperatures compared to stoichiometric alloys (green) (Color figure online)

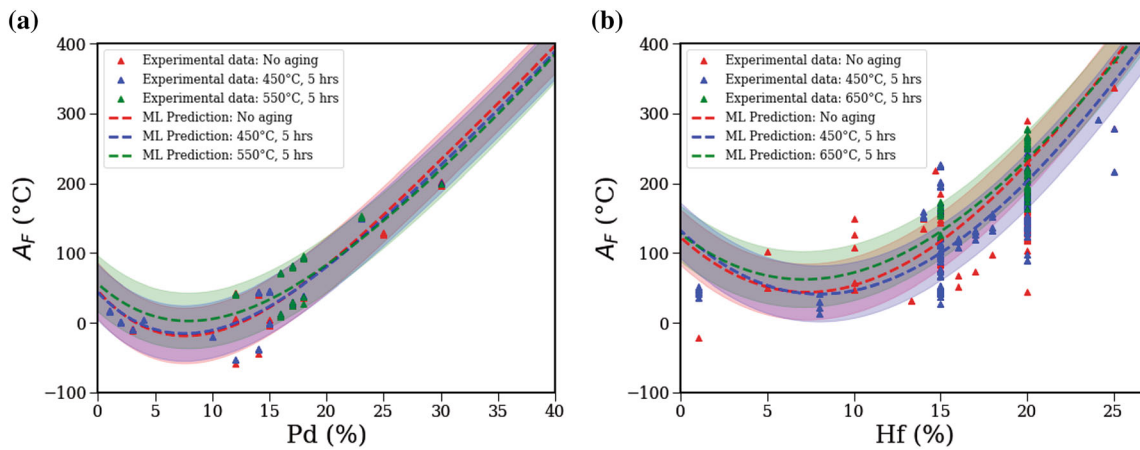


Fig. 6 Effects of aging on A_F predictions for non-stoichiometric Ni–Ti–Pd (Ti = 49.5%) and Ni–Ti–Hf (Ni = 50.5%) alloys. Aging at 550 °C for Pd and 650 °C for Hf (green) increases the transformation temperatures of slightly Ni-rich alloys (Color figure online)

observed. In Ni–Ti–Pd alloys, the P-phase shows unusual precipitation behavior and does not follow the predictable aging trend observed in Ni–Ti–Pt alloys [50]. In fact, some Ni-rich formulations of Ni–Ti–Pd show no precipitates even after aging. The P-phase can also dissolve back into the solid solution at relatively lower temperatures (close to 550 °C). In the case of Ni–Ti–Hf alloys, there is huge scatter in the experimental data itself, which adds to the uncertainty of predictions. Nonetheless, aging at 650 °C clearly shows the expected trend, where A_F is higher compared to unaged alloys. Further work is currently being performed to study how the effects of aging and other heat treatments can be more effectively captured through the ML models.

Next, we explore how an external applied stress affects transformation temperatures. Figure 7a shows that A_F generally increases with higher applied stress in Ni–Ti–Pd alloys. When plot in a different way (Fig. 7b), we see this relationship between A_F temperatures and applied stress is in fact linear. There is good agreement between model

predictions and experimental data at all three Pd compositions. While the 20 and 30% Pd lines have positive slopes, the 40% line appears almost flat, indicating that the effect of applied stress might diminish for high Pd alloys. Such plots are extremely valuable for designing alloys that need to operate at multiple loading conditions, where knowledge of the stress-temperature sensitivity is crucial. Additionally, by using the slope at a given composition, one can determine the zero stress equivalents of the transformation temperatures.

We also consider the effects of applied stress on transformation strain. Figure 8 shows a plot of the predicted transformation strain versus applied stress for Ni–Ti–Pd. Here, too, fourth degree polynomials are used to fit actual predictions. We see that transformation strain shows a linear dependence on applied stress at low values, before flattening out as the applied stress increases. The uncertainty intervals appear wider due to the relatively small dataset available for training. Additionally, plenty of scatter is observed in the experimental data, although most

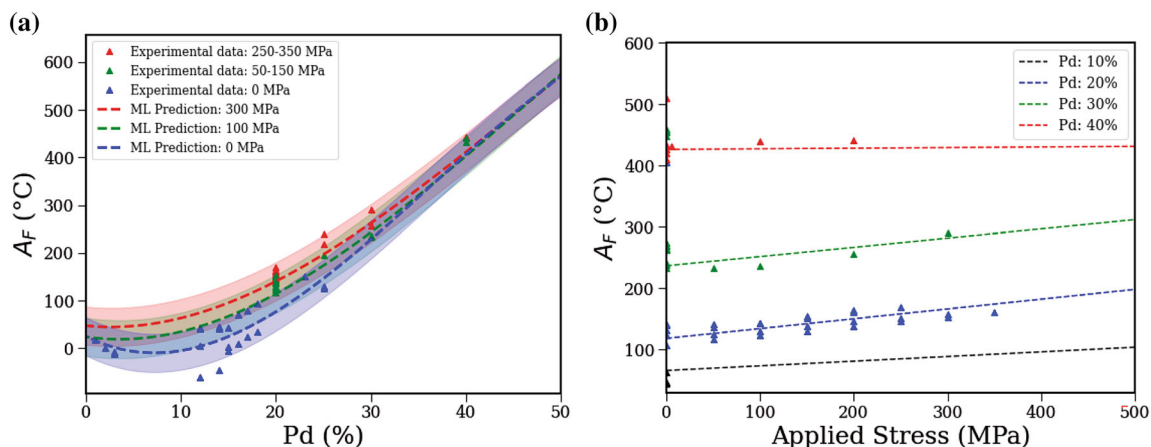


Fig. 7 Effects of applied stress on A_F predictions for Ni–Ti–Pd. A_F increases linearly with increasing applied stress

experimental strain values still lie within the shaded confidence intervals surrounding predictions.

The predicted transformation strain versus applied stress plot for Zr (Supplementary Information Fig. S6) also shows large confidence intervals as a result of scatter within the experimental data. The transformation strains for ternary Ni–Ti–Zr alloys appear to be smaller, on average, when compared to Pd alloys. However, the same trend of decreasing transformation strains with higher Zr content is observed. Thus, while addition of ternary elements like Pd and Zr contribute to higher transformation temperatures in

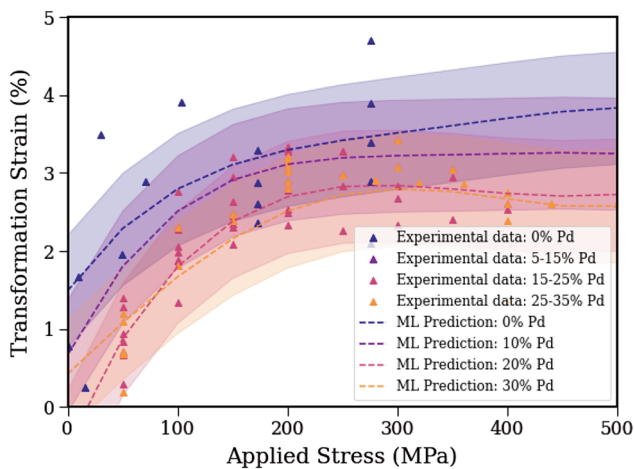


Fig. 8 Effects of applied stress on transformation strain predictions for Ni–Ti–Pd. Transformation strain increases linearly at lower values, before flattening out at higher applied stress

these SMAs, they come at the cost of lower transformation strain, and thus, work output.

The dependence of transformation strain on composition, heat treatment, and other test parameters is more complex and not so obvious from the limited data available to us within the database. Since work output is directly related to the transformation strain, understanding these correlations will prove critical while designing new alloys for commercial applications. Future work includes supplementing the database with additional mechanical testing data generated within our own labs, to study these relationships.

A major advantage of our current approach is the ability to make predictions for alloys with any number of components. Although the ExRT models are trained on the predominantly ternary data available within the dataset, they can also be used to make predictions for higher component alloys. Figure 9 shows A_F and hysteresis predictions for quaternary Ni–Ti–Pd–Pt and Ni–Ti–Hf–Zr alloys computed using these models, with no additional training. Each circle represents an alloy at a particular composition. The color of the circle indicates the predicted A_F whereas its size is indicative of the predicted hysteresis. Such plots are useful for screening alloys. By selecting appropriate filters for transformation temperature and hysteresis, the design space can be quickly narrowed down to a handful of compositions which can then be tested during the experimental validation step. Adding the validation results back into the dataset creates an active learning loop that allows us to iteratively explore the massive space and optimize target properties using far

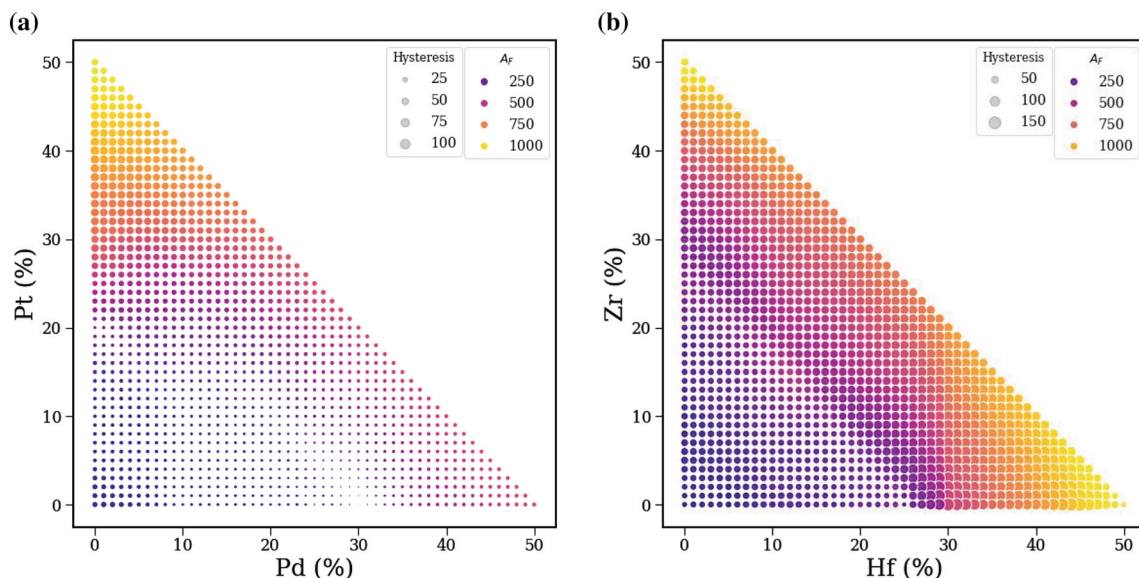


Fig. 9 ML predictions for A_F and hysteresis of quaternary Ni–Ti–Pd–Pt and Ni–Ti–Hf–Zr alloys. Color indicates the predicted A_F whereas size is indicative of the predicted hysteresis

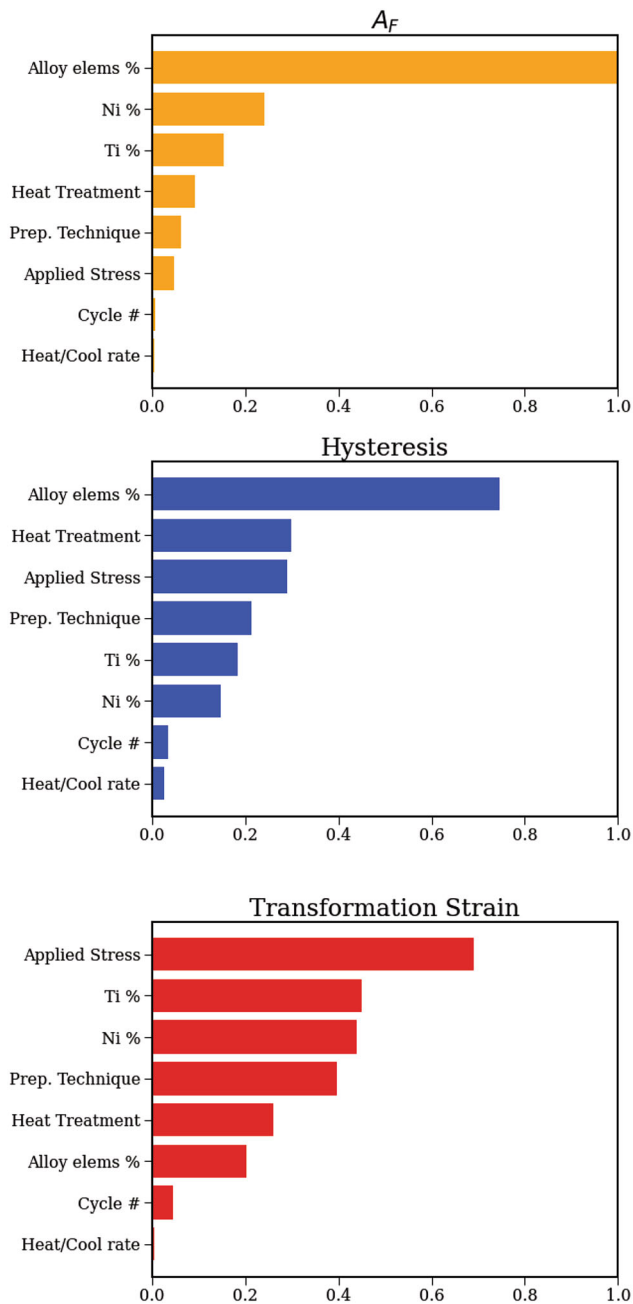


Fig. 10 Permutation feature importance (PFI) plots for the A_F , hysteresis, and transformation strain ML models showing the decrease in R^2 upon random permutation of individual features. Higher values indicate a larger impact on model predictions

fewer experiments. The same models can also be used for five, six, and higher component alloy predictions.

Finally, PFI is used to study the importance of different features in the ML models. Figure 10 shows how randomly permuting input features affects overall performance for the A_F , hysteresis, and transformation strain prediction tasks. Since PFI is affected by the presence of highly correlated features, we perform clustering of related inputs

shown in Table 1. The effects of all 37 alloying elements are grouped under a single column called “Alloy Elems %.” Similarly, the six preparation technique columns are together visualized as “Prep. Technique,” and the six heat treatment features—homogenization temperature and time, solutionizing temperature and time, aging temperature and time—as “Heat Treatment.” Besides, we include the original Ni atomic %, Ti atomic %, and Applied Stress features. Although heating and cooling rate or cycle number are not included as inputs to the final models, their relative importance is also shown in the figure. It is clear that alloy chemistry has the largest impact on predictions of A_F , followed by heat treatment. For hysteresis, thermo-mechanical processing and applied stress are relatively more important, besides composition of alloying elements. Transformation strain predictions largely depend on the applied stress, and simply changing the external stress will have the largest effect on predictions. Knowledge of the most important features can help guide the design of future alloys.

Here, it is important to discuss the advantages and drawbacks of the current approach. Although the large dataset makes predictions for any multi-component alloy possible, it also means that the choice of input features needs to be general. In their recent work, Liu et al. predicted the mean transformation temperature for ternary Ni–Ti–Hf compositions and showed that the use of kinetics-informed process features that are specific to Ni–Ti–Hf, improved predictive performance. [30] The diversity of alloy chemistries in our dataset prevents the use of such composition-specific features. Additionally, our dataset is compiled from measurements performed and reported by different groups. The differences in equipment, sample purity, heat treatment procedures, measurement practices, and data standards introduce scatter in the data and additional uncertainty in model predictions. For example, the choice of melting technique determines which impurities are introduced into the alloy; such small changes in composition can have large effects on transformation behavior. Further, an alloy could be heat treated in several non-standard ways (multiple annealing or aging steps, different temperatures and time ranges, etc.) or have incomplete data reported. In all such cases, the true processing history of the alloy cannot be accurately captured through the input features. Lastly, many engineering alloys also undergo repeated thermal cyclic tests to achieve two-way shape memory effect, referred to as “training.” While training can have a significant effect on the transformation behavior of SMAs, it is not accounted for in the current dataset. [51] Although these drawbacks can lead to a slight loss of accuracy, the wider confidence intervals that result can provide hints about the scatter in the data, allowing designers to make informed decisions knowing the

uncertainty. Still, the low overall MAE achieved by our models confirms the usefulness of such an approach. The added benefit of exploring much larger design spaces makes it even more promising for materials discovery.

Summary

In summary, we developed a data-driven approach to study transformation behavior in multicomponent Ni–Ti SMAs. Using a dataset of about 8000 Ni–Ti alloys containing 37 different alloying elements, machine learning models were trained to predict transformation temperatures (M_S , M_F , A_S , A_F), hysteresis ($A_F - M_S$), and transformation strain. Through feature engineering, 52 composition, thermo-mechanical processing, and test parameter related inputs were identified, which helped achieve low MAE on all learning tasks. The trained models were employed to study trends, learn correlations, and make predictions for new higher-order alloy systems. Uncertainty intervals and feature importance were also computed to assist with experimental validation and to guide future alloy design. The current approach makes it possible to explore vast compositional and processing spaces using the same models, making it ideal for rapid discovery of new SMAs. We are not aware of any other approaches capable of predicting SMA transformation behavior over such a wide range of compositions and processing conditions.

Future work will involve experimentally validating promising alloy compositions identified using the models. The validation results will be fed back into the dataset to create an active learning loop that will result in improved predictions with each iteration.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40830-022-00405-x>.

Acknowledgments This work was performed at NASA Ames Research Center under Prime Contract No. 80ARC020D0010. We acknowledge funding from the NASA Aeronautics Research Mission Directorate's (ARMD) Transformational Tools and Technologies (TTT) project. We also acknowledge Joshua Stuckner for his useful suggestions.

Data Availability All data used in this work is publicly available at https://shapememory.grc.nasa.gov/shape_memory_alloys. The code for generating the ML models can be made available upon request from the corresponding authors.

References

- Otsuka K, Ren X (1999) Recent developments in the research of shape memory alloys. *Intermetallics* 7:511–528
- Jani JM, Leary M, Subic A, Gibson MA (2014) A review of shape memory alloy research, applications and opportunities. *Mater Des* 56:1078–1113
- Dasgupta R (2014) A look into Cu-based shape memory alloys: present scenario and future prospects. *J Mater Res* 29:1681–1698
- Otsuka H, Yamada H, Maruyama T, Tanahashi H, Matsuda S, Murakami M (1990) Effects of alloying additions on Fe–Mn–Si shape memory alloys. *ISIJ Int* 30:674–679
- Otsuka K, Ren X (2005) Physical metallurgy of Ti–Ni-based shape memory alloys. *Prog Mater Sci* 50:511–678
- Duerig T, Pelton A, Stöckel D (1999) An overview of nitinol medical applications. *Mater Sci Eng A* 273:149–160
- Stöckel D (1990) Shape memory actuators for automotive applications. *Mater Des* 11:302–307
- Hartl DJ, Lagoudas DC (2007) Aerospace applications of shape memory alloys. *Proc Inst Mech Eng G: J Aerosp Eng* 221:535–552
- Biesiekierski A, Wang J, Gepreel MA-H, Wen C (2012) A new look at biomedical Ti-based shape memory alloys. *Acta Biomater* 8:1661–1669
- Ma J, Karaman I, Noebe RD (2010) High temperature shape memory alloys. *Int Mater Rev* 55:257–315
- Bigelow GS, Padula SA, Garg A, Gaydos D, Noebe RD (2010) Characterization of ternary NiTiPd high-temperature shape-memory alloys under load-biased thermal cycling. *Metall Mater Trans A* 41:3065–3079
- Bigelow G, Garg A, Padula SA, Gaydos D, Noebe R (2011) Load-biased shape-memory and superelastic properties of a precipitation strengthened high-temperature Ni_{50.3}Ti_{29.7}Hf₂₀ alloy. *Scripta Mater* 64:725–728
- Sehitoglu H, Patriarca L, Wu Y (2017) Shape memory strains and temperatures in the extreme. *Curr Opin Solid State Mater Sci* 21:113–120
- Bigelow G, Benafan O, Garg A, Noebe R (2021) Effect of Hf/Zr ratio on shape memory properties of high temperature Ni_{50.3}Ti_{29.7}(Hf, Zr)₂₀ alloys. *Scripta Mater* 194:113623
- Grossmann C, Frenzel J, Sampath V, Depka T, Eggeler G (2009) Elementary transformation and deformation processes and the cyclic stability of NiTi and NiTiCu shape memory spring actuators. *Metall Mater Trans A* 40:2530–2544
- Haskins JB, Malmir H, Honrao SJ, Sandoval LA, Lawson JW (2021) Low temperature mechanical instabilities govern high-temperature thermodynamics in the austenite phase of shape memory alloy constituents: Ab initio simulations of NiTi, NiZr, NiHf, PdTi, and PtTi. *Acta Mater* 212:116872
- Kadkhodaei S, van de Walle A (2018) First-principles calculations of thermal properties of the mechanically unstable phases of the PtTi and NiTi shape memory alloys. *Acta Mater* 147:296–303
- Ko W-S, Grabowski B, Neugebauer J (2015) Development and application of a Ni–Ti interatomic potential with high predictive accuracy of the martensitic phase transition. *Phys Rev B* 92:134107
- Mutter D, Nielaba P (2010) Simulation of structural phase transitions in NiTi. *Phys Rev B* 82:224201
- Honrao SJ, Anthonio BE, Ramanathan R, Gabriel JJ, Hennig RG (2019) Machine learning of ab-initio energy landscapes for crystal structure predictions. *Comput Mater Sci* 158:414–419
- Honrao SJ, Xie SR, Hennig RG (2020) Augmenting machine learning of energy landscapes with local structural information. *J Appl Phys* 128:085101
- Toyao T, Suzuki K, Kikuchi S, Takakusagi S, Shimizu K-I, Takigawa I (2018) Toward effective utilization of methane: machine learning prediction of adsorption energies on metal alloys. *J Phys Chem C* 122:8315–8326

23. Furmanchuk A, Agrawal A, Choudhary A (2016) Predictive analytics for crystalline materials: bulk modulus. *RSC Adv* 6:95246–95251
24. Pilania G, Wang C, Jiang X, Rajasekaran S, Ramprasad R (2013) Accelerating materials property predictions using machine learning. *Sci Rep* 3:1–6
25. Pilania G, Mannodi-Kanakthodi A, Uberuaga B, Ramprasad R, Gubernatis J, Lookman T (2016) Machine learning bandgaps of double perovskites. *Sci Rep* 6:1–10
26. Wu H, Lorensen A, Anderson B, Witteman L, Wu H, Meredig B, Morgan D (2017) Robust FCC solute diffusion predictions from ab-initio machine learning methods. *Comp Mat Sci* 134:160–165
27. Honrao SJ, Yang X, Radhakrishnan B, Kuwata S, Komatsu H, Ohma A, Sierhuis M, Lawson JW (2021) Discovery of novel Li SSE and anode coatings using interpretable machine learning and high-throughput multi-property screening. *Sci Rep* 11:1–14
28. Ward L, Agrawal A, Choudhary A, Wolverton C (2016) A general-purpose machine learning framework for predicting properties of inorganic materials. *NPJ Comput Mater* 2:1–7
29. von Lilienfeld OA, Müller K-R, Tkatchenko A (2020) Exploring chemical compound space with quantum-based machine learning. *Nat Rev Chem* 4:347–358
30. Liu S, Kappes BB, Amin-ahmadi B, Benafan O, Zhang X, Stebner AP (2021) Physics informed machine learning for composition-process-property design: shape memory alloy demonstration. *Appl Mater Today* 22:100898
31. Xue D, Xue D, Yuan R, Zhou Y, Balachandran PV, Ding X, Sun J, Lookman T (2017) An informatics approach to transformation temperatures of NiTi-based shape memory alloys. *Acta Mater* 125:532–541
32. Trehern W, Ortiz-Ayala R, Atli K, Arroyave R, Karaman I (2022) Data-driven shape memory alloy discovery using artificial intelligence materials selection (AIMS) framework. *Acta Mater* 228:117751
33. Benafan O, Bigelow GS, Young AW (2020) Shape memory materials database tool—a compendium of functional data for shape memory materials. *Adv Eng Mater* 22:1901370
34. Zarnetta R, Takahashi R, Young ML, Savan A, Furuya Y, Thienhaus S, Maaß B, Rahim M, Frenzel J, Brunken H, Chu YS, Srivastava V, James RD, Takeuchi I, Eggeler G, Ludwig A (2010) Identification of quaternary shape memory alloys with near-zero thermal hysteresis and unprecedented functional stability. *Adv Funct Mater* 20:1917–1923
35. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63:3–42
36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in python. *J Mach Learn Res* 12:2825–2830
37. Miller RG (1974) The jackknife—a review. *Biometrika* 61:1–15
38. Taquet V, Martinon G. A Scikit-learn-compatible module for estimating prediction intervals. <https://github.com/scikit-learn-contrib/>.
39. Barber RF, Candès EJ, Ramdas A, Tibshirani RJ (2021) Predictive inference with the jackknife+. *Ann Stat* 49:486–507
40. Altmann A, Toloşi L, Sander O, Lengauer T (2010) Permutation importance: a corrected feature importance measure. *Bioinformatics* 26:1340–1347
41. Parr TP, Turgutlu K. Feature importances for scikit-learn machine learning models. <https://github.com/parrt/random-forest-importances>.
42. Benafan O, Bigelow GS, Garg A, Noebe RD (2019) Viable low temperature shape memory alloys based on Ni-Ti-Hf formulations. *Scr Mater* 164:115–120
43. Frenzel J, Wieczorek A, Opahle I, Maaß B, Drautz R, Eggeler G (2015) On the effect of alloy composition on martensite start temperatures and latent heats in Ni-Ti-based shape memory alloys. *Acta Mater* 90:213–231
44. Bertheville B, Bidaux JE (2005) Alternative powder metallurgical processing of Ti-rich NiTi shape-memory alloys. *Scr Mater* 52:507–512
45. Safaei K, Nematollahi M, Bayati P, Dabbaghi H, Benafan O, Elahinia M (2021) Torsional behavior and microstructure characterization of additively manufactured NiTi shape memory alloy tubes. *Eng Struct* 226:111383
46. Han X, Wang R, Zhang Z, Yang D (1998) A new precipitate phase in a TiNiHf high temperature shape memory alloy. *Acta Mater* 46:273–281
47. Yang F, Coughlin D, Phillips PJ, Yang L, Devaraj A, Kovarik L, Noebe RD, Mills M (2013) Structure analysis of a precipitate phase in an Ni-rich high-temperature NiTiHf shape memory alloy. *Acta Mater* 61:3335–3346
48. Kovarik L, Yang F, Garg A, Diercks D, Kaufman M, Noebe R, Mills M (2010) Structural analysis of a new precipitate phase in high-temperature TiNiPt shape memory alloys. *Acta Mater* 58:4660–4673
49. Benafan O, Bigelow G, Garg A, Noebe R, Gaydos D, Rogers R (2021) Processing and scalability of NiTiHf high-temperature shape memory alloys. *Shape Mem Superelasticity* 7:109–165
50. Coppa AC, Kapoor M, Noebe R, Thompson GB (2015) The compositional stability of the P-phase in Ni–Ti–Pd shape memory alloys. *Intermetallics* 67:56–62
51. Luo HY, Abel EW (2007) A comparison of methods for the training of NiTi two-way shape memory alloy. *Smart Mater Struct* 16:2543

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.