ARL-TR-10352 • MAY 2026

Accelerating Materials Performance Evaluation and Discovery: Prediction of Steel Structures for High-Throughput Design and Testing

by Heather Murdoch, Michael Rupinen, and Krista Limmer

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-10352 ● MAY 2026

Accelerating Materials Performance Evaluation and Discovery: Prediction of Steel Structures for High-Throughput Design and Testing

Heather Murdoch, Michael Rupinen, and Krista Limmer
DEVCOM Army Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
May 2026	Technical Report	January 2025	January 2026

4. TITLE AND SUBTITLE
Accelerating Materials Performance Evaluation and Discovery: Prediction of Steel Structures for High-Throughput Design and Testing

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER	5e. TASK NUMBER	5f. WORK UNIT NUMBER

6. AUTHOR(S)
Heather Murdoch, Michael Rupinen, and Krista Limmer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
DEVCOM Army Research Laboratory ATTN: FCDD-RLA-MF Aberdeen Proving Ground, MD 21005	ARL-TR-10352

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCIDs: Heather Murdoch, 0000-0001-7710-0577; Michael Rupinen, 0009-0009-9438-9902; Krista Limmer, 0000-0003-4775-5876

14. ABSTRACT
The Accelerating Materials Performance Evaluation and Discovery (AMPED) program was a 12-month cross-service program to build foundational elements for rapid, automated materials development and characterization. This report documents the ARL-led contributions to the AMPED program. Several modeling approaches to predict the categorical structure of a given steel based on composition and processing were evaluated with two datasets constructed from open literature. Significantly, these modeling approaches do not require commercial software and have low computational expense, enabling integration into accelerated design workflows including predictive modeling tools and guiding high-throughput experimental strategies. A new random forest classifier model was trained that was highly successful at predicting the structure of a steel, given composition and processing.

15. SUBJECT TERMS
steel, martensite, microstructure prediction, machine learning, high-throughput, Sciences of Extreme Materials

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UU	30

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Heather Murdoch	(520) 691-6293

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures iv

List of Tables v

Acknowledgments vi

1. Introduction 1

2. Dataset Creation 3

3. Martensite Kinetics 4

4. Microstructure Classifiers 7

4.1 Empirical CCT 7

4.2 Schaeffler 9

4.3 RF Classifier 10

4.4 Retraining RF Classifier 12

5. Application 15

6. Conclusions 16

7. References 18

List of Symbols, Abbreviations, and Acronyms 20

Distribution List 22

List of Figures

Figure 1. Predicted hardness values for exemplar steels in AMPED using the Tempered Hardness Optimization of Martensitic Alloy Steels Gaussian process regressor ML model[2] that was trained on only martensitic steels. 2

Figure 2. All the martensite fraction curves in the dataset. 3

Figure 3. Composition space of (a) martensite kinetics and (b) classifier datasets. Opacity of violin plot is a function of the number or steels in the dataset that include that element. Few contain W, which is mostly transparent, compared with almost all containing Mn, which is colored solid. 4

Figure 4. Comparison of experimental (Exp.) martensite fraction curves with the existing models. KM = Koistinen–Marburger prediction, T-C = Thermo-Calc prediction. 5

Figure 5. Comparison of existing models with our large dataset. 6

Figure 6. Predicted kinetics curves for AerMet100 and HiPO. 7

Figure 7. CCT curves for two different steels: (a) fully martensitic alloy and (b) mixed microstructure alloy. Solid lines are phase transformation onsets (dotted lines are transformation finish temperatures). Quench rates are labeled gray curves. 8

Figure 8. Predicted class of steels in the full classifier dataset using the empirical CCT curve as the predictive method/model: R^2 of 0.19. 9

Figure 9. (a) Full classification dataset plotted on the original Schaeffler diagram using the Cr and Ni equivalence of DeLong.[16] (b) Confusion matrix for the predictive capability of the Schaeffler diagram. 10

Figure 10. RF model predictions of (a) original model using C, Ni, Cr, and two heat-treatment steps as inputs for the original test set; (b) original model using C, Ni, Cr, and only the first heat-treatment step for the original test set; (c) model using inputs of (b) but predicting the original test set and new data; and (d) model of (b) but using Ni_{eq} and Cr_{eq} instead of Ni and Cr, predicting the original test set and new data. 11

Figure 11. Sum across classes of average SHAP values input features for RF classifier model trained on all elements and *ausT* and *qr* as input features across 100 test/train splits. Features are ranked by total impact across all classes. 13

Figure 12. SHAP values separated by class for RF classifier model trained on all elements and *ausT* and *qr* as input features across 100 test/train splits. Features are ranked by total impact across all classes. 14

Figure 13. RF classifier model trained using different input features on the full dataset: (a) DeLong[16] Ni/Cr equivalencies + C; (b) Lee[17] Ni/Cr equivalencies + C; (c) some elements; and (d) all elements. 15

Figure 14. Predicted phase map of Fe-xCr-yMn-0.05C with an austenitization temperature of 1000 °C and a quench rate of 100 °C/s using the classifier model with all elements as inputs. Color scale is based on the probability associated with the prediction of the class. 16

List of Tables

Table 1. Coefficients for Ni and Cr equivalencies. 12

Acknowledgments

Funding for Accelerating Materials Performance Evaluation and Discovery was provided by the Office of the Under Secretary of Defense for Research and Engineering Emerging Research Opportunities Seedlings program as part of the Applied Research for the Advancement of Science and Technology Priorities.

We thank Drs. Keith Knipling, David Rowenhorst, and Colin Stewart at the U.S. Naval Research Laboratory for austenitic steel data; and Drs. Blake Barnett, Efrain Hernández, and Philip Goins at ARL for stimulating discussions on data presentation.

1. Introduction

The Accelerating Materials Performance Evaluation and Discovery (AMPED) program was a 12-month cross-service program to build foundational elements for rapid, automated materials development and characterization. The AMPED program focused on assessment of small-scale tests to evaluate mechanical properties, namely shear punch testing (SPT) and a simplified high-temperature tensile testing unit (AeroForce Test System). The program's secondary objective was to extract maximal value from these high-throughput–capable testing methods using ML tools to accelerate development and deployment.

ML approaches can be leveraged both to connect raw data from small-scale testing to full-scale properties[1] and to integrate it into an accelerated design workflow. Unfortunately, not enough data was produced by the SPT testing during this program to evaluate its incorporation into predictive multitask ML models as planned. Therefore, this report focuses on potential components for a design workflow that would support rapid alloy development and guide high-throughput experimentation. The specific emphasis is on steel alloy development, as there are a wide range of steels of interest for defense applications. A family of austenitic steels developed at the U.S. Naval Research Laboratory (NRL) and a family of martensitic steels developed at the U.S. Army Combat Capabilities Development Command (DEVCOM) Army Research Laboratory (ARL) were used as exemplars during this program for both mechanical testing and design conditions.

ARL previously developed an alloy design ML framework focused on predicting properties of martensitic steels,[2] particularly hardness and impact toughness. Large datasets were assembled, including composition, processing history, and mechanical properties for use in training predictive ML models. These datasets were curated to contain only steels known to be—or strongly expected to be—fully martensitic. The mechanisms underpinning the composition–process–structure–property relationships in martensitic alloys are different than those in austenitic, ferritic, or mixed-microstructure steels; thus, the ML models trained on only martensitic alloys will not well predict other classes of steels. This is illustrated in Figure 1, where the ML model trained on martensitic steels was used to predict hardness for the two families of steel tested in AMPED: martensitic (green circles) and austenitic (blue triangles). The predictions for the martensitic group (which includes changes in both composition and processing history) are good and could be reasonably used to guide further design efforts. On the other hand, as expected, the predictions for the austenitic steels are unreliable.

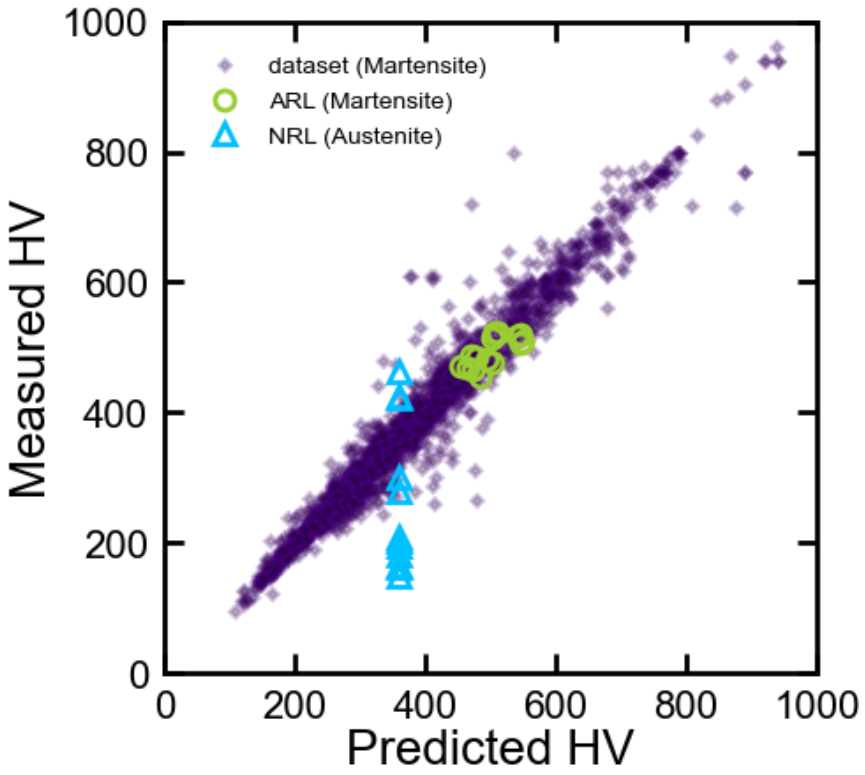

Figure 1. Predicted hardness values for exemplar steels in AMPED using the Tempered Hardness Optimization of Martensitic Alloy Steels Gaussian process regressor ML model[2] that was trained on only martensitic steels.

A robust alloy design workflow for steel ideally would not be limited to one microstructure family, particularly as a given composition could produce a wide range of microstructures depending on the processing method. For example, the cooling rate after austenitizing can control whether the material is fully martensitic, bainitic, pearlitic, ferritic, or some mixed structure. Other processing methods, such as quenching and partitioning, depend strongly on not only cooling rate, but also close control of the fraction of martensite versus austenite. The common methods to predict the base microstructure of a given steel are continuous cooling transformation (CCT) curves, time–temperature–transformation curves, martensite kinetics curves, and Schaeffler diagrams. These are commonly generated through some combination of empirical, semi-empirical, and thermodynamic models. More recently, ML models have been developed for a few of these approaches. In AMPED, two datasets were developed and used to examine current capabilities in predicting steel microstructure. One is specifically for the austenite to martensite transformation (martensite kinetics), and the second is for the full range of microstructural classes (steel classifier).

2. Dataset Creation

The martensite kinetics dataset compiled for this program contains 231 complete martensite fraction versus temperature curves. This data corresponds to 228 unique alloys, as 3 alloys feature multiple datasets derived from different starting temperatures. The breadth of temperatures and transformation rates in the dataset is shown in Figure 2. This dataset is more comprehensive than any previous dataset of this type; i.e., the work of van Bohemen and Sietsma[3] has 12 full curves whereas the Huyan et al.[4] model includes 46 alloys but does not include their curves in their report.

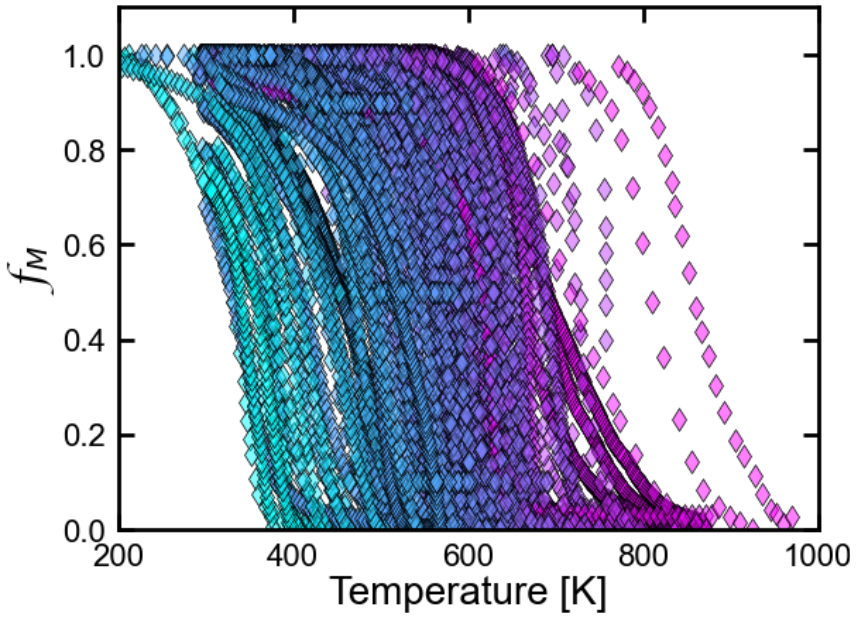

Figure 2. All the martensite fraction curves in the dataset.

The steel classifier dataset has 1148 entries of composition with austenitizing temperature (*ausT*) and quench rate (*qr*). Quench rates are generally not reported quantitatively, but instead qualitatively based on the quench medium. The open dataset used as the basis for this work[5] used 500 °C/s for water quench, 120 °C/s for oil quench, and 3 °C/s for air cooling. Each entry is labeled with the steel microstructure: austenite (A), ferrite (F), martensite (M), pearlite (P), and combinations of those. The original classifier dataset does not distinguish between pearlite and bainite, as both microstructures consist of ferrite and carbide. Thus, both pearlite and bainite are given as "P" for pearlite.

The composition spaces for the developed martensite kinetics and steel classifier datasets are shown in Figure 3. The wide composition space was intended to cover the breadth of steel compositions used in defense applications.

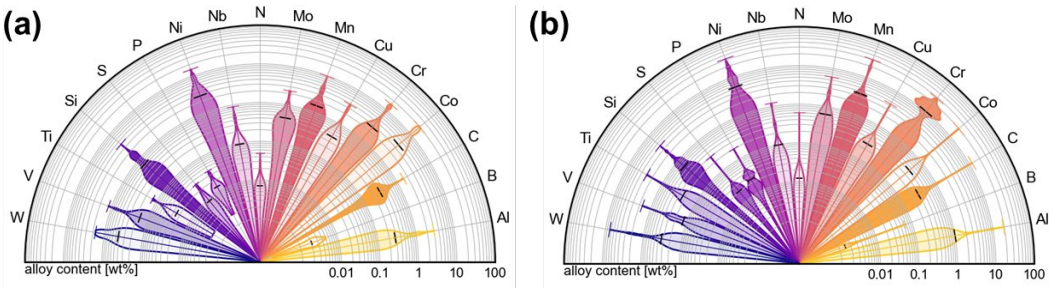

Figure 3. Composition space of (a) martensite kinetics and (b) classifier datasets. Opacity of violin plot is a function of the number or steels in the dataset that include that element. Few contain W, which is mostly transparent, compared with almost all containing Mn, which is colored solid.

3. Martensite Kinetics

Two empirical models are widely used to predict martensite kinetics: the Koistinen–Marburger[6] (KM) and Huyan et al.[4] models, the latter of which is integrated into the Thermo-Calc software (labeled as TC model). The KM model parameter originally only fit four experimental alloys but has been applied to a wider range of alloy systems with small modifications.[3,7] These modifications are often focused on adding fitting parameters to fit a particular dataset and are limited in providing a versatile kinetics model. The Huyan et al.[4] model has parameters fit to a dataset of 46 alloys and uses a thermodynamic driving-force calculation, which should increase the extensibility of the model; however, it requires access to commercial thermodynamic software to calculate. This software restriction limits the ease of integrating the model into other high-throughput frameworks.

The kinetic behavior of martensite formation is shown in Figure 4 for two of the martensitic steels tested in this program: a commercial alloy, AerMet100, and a modified high-performance (mod HP) alloy from the ARL-developed high-performance (HiPO) alloy family. The commercial AerMet100 steel requires a cryogenic quench to complete the martensitic transformation. This is demonstrated by the experimental data shown in Figure 4, where the transformation is slow and stalls out near room temperature. If the martensitic transformation is not completed through a cryogenic quench, the toughness of the material decreases.[8] Therefore, models predicting the toughness of steel during the design process would benefit from an assessment of martensite kinetics as an input feature. The martensite kinetics are also crucial for quenched and partitioned steels where the initial fraction of martensite is key in determining the austenite fraction and stability.

The HiPO family of alloys was partially designed to reduce the alloy content of AerMet100, which also affected the transformation kinetics. This is observed by the mod HP curve in Figure 4 (green pentagons), which completes the martensite

transformation well above room temperature, despite having essentially the same starting point as AerMet100. Both the KM and TC models fail to reproduce this behavior, erroneously predicting the AerMet100 to have a faster transformation than HiPO. Furthermore, neither model correctly describes the shape of either curve.

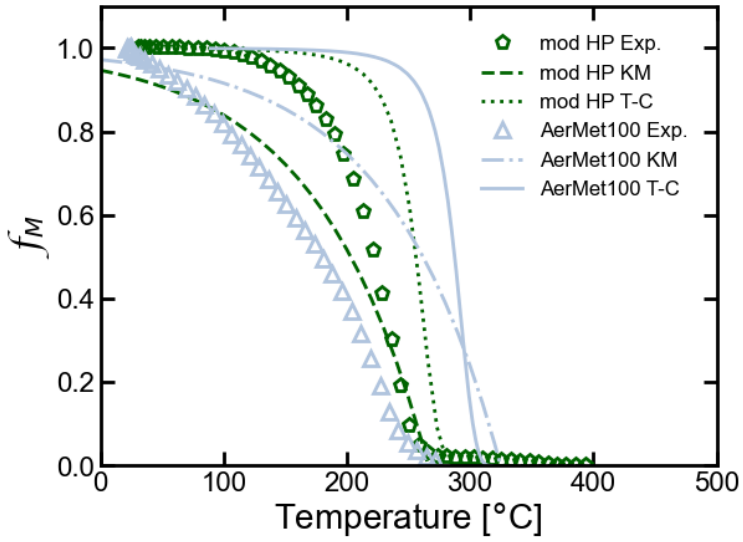

Figure 4. Comparison of experimental (Exp.) martensite fraction curves with the existing models. KM = Koistinen–Marburger prediction, T-C = Thermo-Calc prediction.

The existing KM and TC models were used to predict the kinetic behavior of the 231 alloys collected in the new dataset. The difference between the predicted fraction and actual martensite fraction is shown in Figure 5. Because the functional form of the KM equation is not bounded, it can result in unphysical negative predictions, as shown across the whole range of experimental fractions. The TC model is bounded, however the high concentration of predicted fractions greater than 0.8 for a wide range of measured fractions shows that the TC-predicted kinetics are generally too fast (as is also seen in Figure 4). Discarding the negative points, KM predictions have better error statistics than TC predictions (mean squared error of 0.081 vs. 0.105). This lends support for the development of an empirical model that does not require use of thermodynamic software, fit to this larger composition space and with an improvement to the functional form to introduce physical bounds.

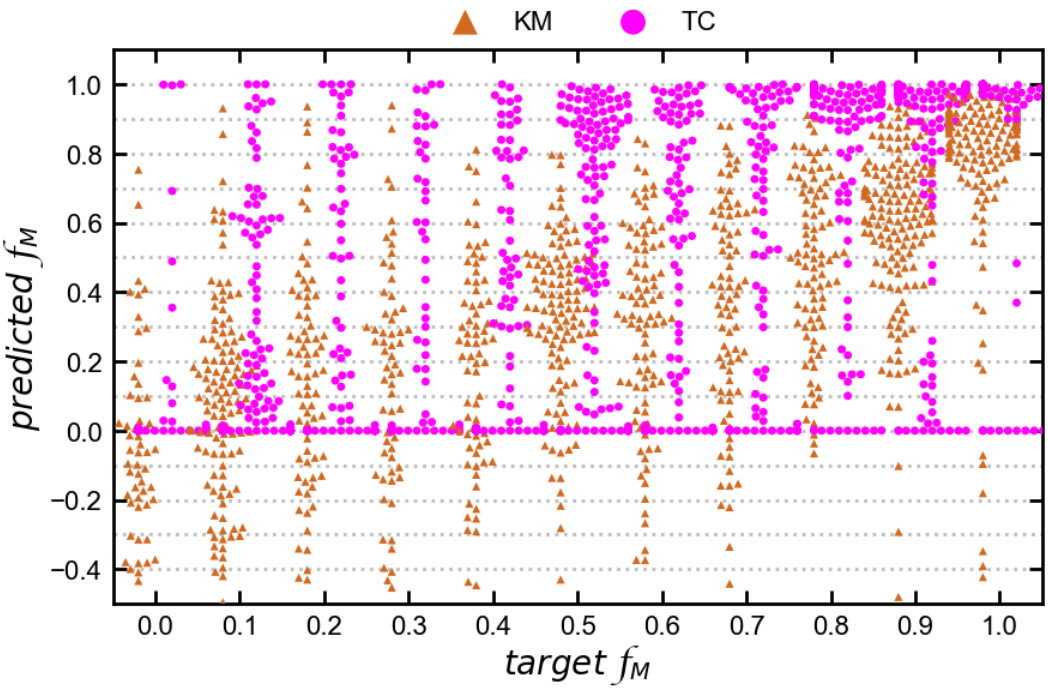

Figure 5. Comparison of existing models with our large dataset.

Symbolic regression with the PySR Python module[9] was used to develop a new functional form that is asymptotic to 0 and 1, covers the full shape of experimental curves (including the S shape), and penalizes overfitting:

$$f_M = 1 - \frac{1}{1+e^{\alpha T^2+\beta T+\gamma}} \quad (1)$$

Where T is temperature and α, β, and γ are functions of composition. Requiring the functional form to have the martensite start temperature as a feature degraded the performance of the predictions. Full details of the model will be published in a future report, including evaluation of several options for the coefficients. An example prediction for a model in the form of Equation 1, with only six elements used for the coefficients, is shown in Figure 6 with excellent agreement for HiPO and reasonable agreement for AerMet100. Most importantly, it reproduces the relative kinetics of the two alloys, unlike all previous models (even using only a fraction of the elements involved).

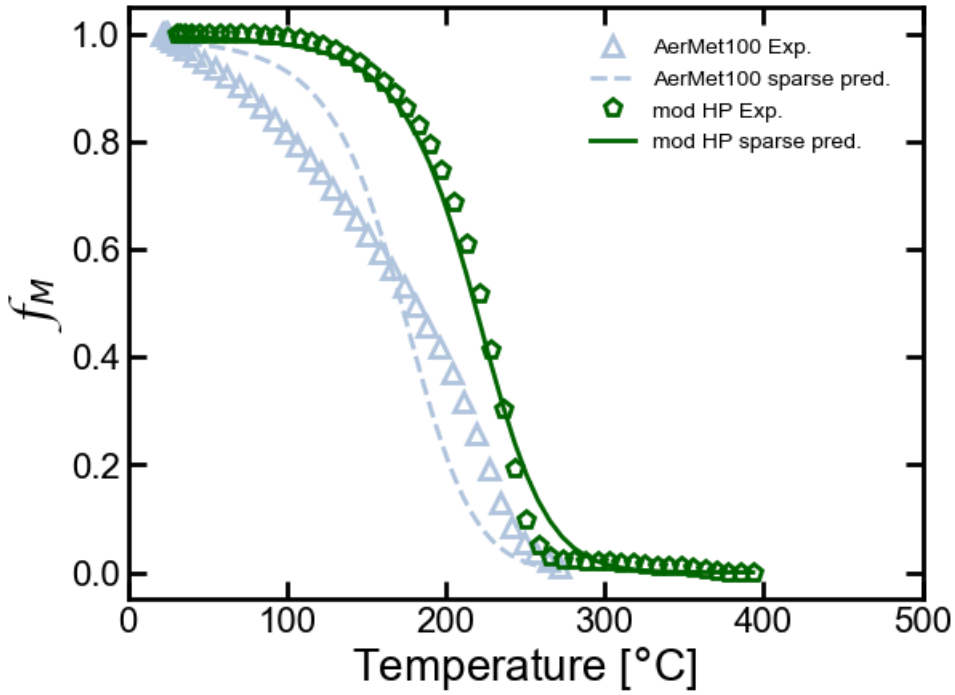

Figure 6. Predicted kinetics curves for AerMet100 and HiPO.

The martensite kinetics model only describes the case of decomposition upon cooling of austenite into martensite, with no other phases present. The next section describes various methods to predict whether austenite will decompose into martensite or other microstructures, e.g., ferrite, bainite, pearlite, or whether the steel will remain the same microstructure throughout the whole temperature range of processing. The steel classifier dataset compiled in this work will be used to evaluate the various predictive methods, including empirical CCT curve predictor, the Schaeffler diagram, and a random forest (RF) ML approach.

4. Microstructure Classifiers

4.1 Empirical CCT

Figure 7 shows examples of CCT curves for two alloys, with solid-colored lines representing the onset of formation of the given phase and gray curves indicating the cooling rate. Figure 7a shows an alloy that is martensitic across the entire range of cooling rates while Figure 7b shows an alloy that could form many different structures depending on cooling rate. In contrast to the alloy in Figure 7a, the alloy in Figure 7b would be difficult to cool fast enough to result in a completely martensitic structure; therefore, the microstructure will be a mixture of phases. For example, a 50 °C/s cooling rate on the alloy in Figure 7b would be a mixture of ferrite, bainite, and martensite.[10]

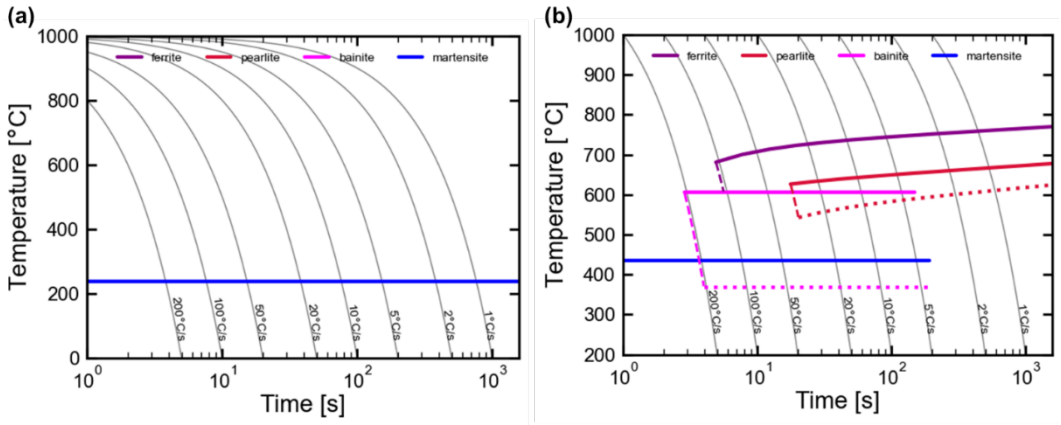

Figure 7. CCT curves for two different steels: (a) fully martensitic alloy and (b) mixed microstructure alloy. Solid lines are phase transformation onsets (dotted lines are transformation finish temperatures). Quench rates are labeled gray curves.

Miettinen et al.[10,11] constructed a set of empirical equations to describe the formation of martensite, ferrite, bainite, and pearlite from reference CCT curves. They predict the critical temperatures and rates for formation for the various microstructures using inputs of composition, austenitization temperature and time, and cooling rate. In a limited analysis, these agree with the predictions of thermodynamic software (e.g., JMatPro and TC) and are significantly faster and more accessible. However, a significant drawback is that they assume the austenite will transform into some combination of martensite, ferrite, bainite, and pearlite, but cannot directly identify which alloys will be austenitic. Austenitic alloys are potentially classifiable with the Miettinen et al. model[10,11] by recognizing unphysical predictions; alloy compositions far outside the space described by the equations result in predictions of extremely high rates, temperatures above the austenitizing temperature, or temperatures below zero for various transformations.

If we assume that a predicted transformation temperature of less than zero or greater than the austenitization temperature indicates the composition is austenitic, we arrive at the predictions shown in Figure 8 for the 1148 entries in the steel classifier dataset. The confusion matrix in Figure 8 shows the predicted versus true label for the class with the percentage of the predictions in that pair indicated, along with the number of items for that pair. The total number of steels in each true category is shown on the right-hand side of the matrix (i.e., there are 347 austenitic steels). The CCT diagrams generated from the Miettinen et al.[10,11] equations predict the phase mixtures of martensite–pearlite (MP) for 250 entries and ferrite–pearlite–martensite (FMP) for 26 entries, although these classifications do not exist in the dataset; therefore, those rows are blank.

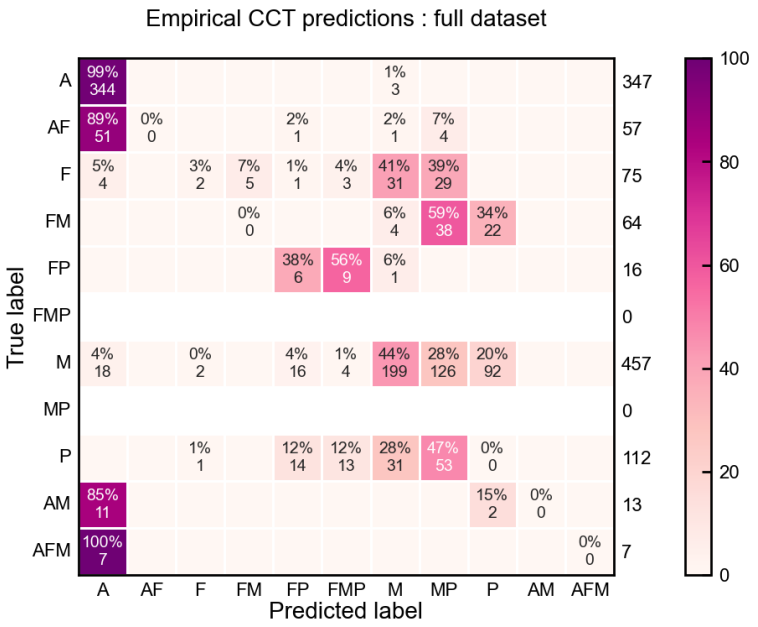

Figure 8. Predicted class of steels in the full classifier dataset using the empirical CCT curve as the predictive method/model: R^2 of 0.19.

From these predictions we see that the assumption predicts fully austenitic materials well. It also successfully predicts austenite for most of the mixed austenitic materials (austenite–ferrite [AF], austenite–martensite [AM], and austenite–ferrite–martensite [AFM]), although it does not differentiate between them based on the available criteria. Overall, the predictions are not accurate, so despite the appeal of this method to generate a complete CCT diagram in addition to a microstructure label with low computational expense, its poor predictive capability is not ideal for integration with a design or testing workflow.

4.2 Schaeffler

The Schaeffler diagram was originally developed for predicting microstructures in stainless steel welds[12] but has also been used in alloy design. The diagram axes of chromium (Cr) and nickel (Ni) equivalencies (Cr_{eq} and Ni_{eq}, respectively) attempt to account for ferrite- and austenite-stabilizing elements, respectively. The boundaries have been empirically determined but do align well with the underlying thermodynamics of phase stability and competition between austenite and ferrite.[13] Martensite formation is also related to the difference between austenite and ferrite stability.[14] However, it is still a rate-dependent process, which is not considered in the Schaeffler diagram. Bainite and pearlite formation are also not considered, although the precursor diagram of Strauss and Maurer included a ferrite–pearlite (FP) region.[15] Figure 9a shows the classification dataset plotted using the equivalence equations of DeLong,[16] which are the same as the original Schaeffler

equations with the addition of nitrogen to the Ni_{eq}. The original FP region is also plotted. For clarity, steels from the classification dataset that are pearlite only (112 values) are not shown in Figure 9a because there is no region of the diagram for them.

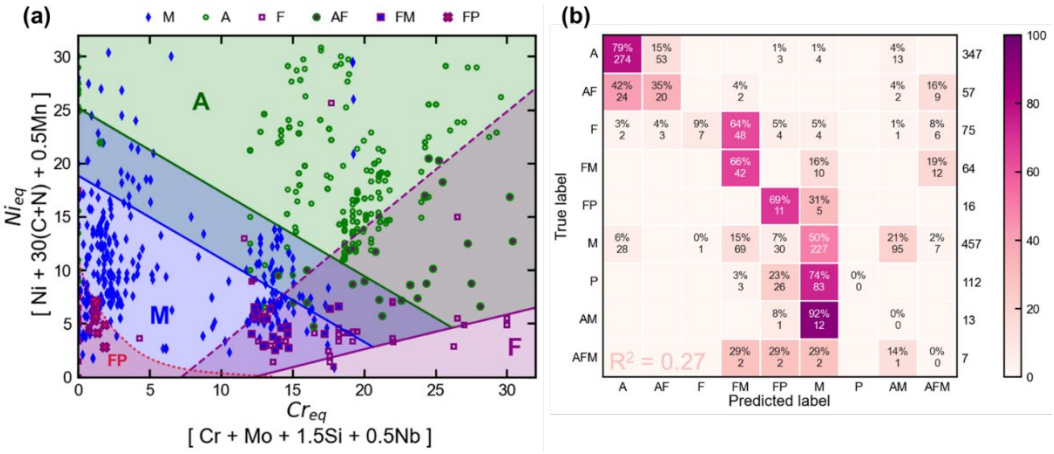

Figure 9. (a) Full classification dataset plotted on the original Schaeffler diagram using the Cr and Ni equivalence of DeLong.[16] (b) Confusion matrix for the predictive capability of the Schaeffler diagram.

A visual assessment of the diagram shows a fair amount of the steels in the correct region, illustrating that the concept of the equivalence equations is still a useful descriptor. Quantifying the predictive capability (Figure 9b) shows a low overall R^2 value (0.27). While the prediction at every point is not perfect, it does generally get some component of the class correct, i.e., predicting "M" or "AFM" for ferrite–martensite ("FM"). The lack of quench rate mentioned previously also means the accuracy of martensitic predictions may be limited. A more recent set of equivalency equations and regions[17] that accounts for more elements improves the ferrite prediction considerably but is significantly worse overall (not shown).

4.3 RF Classifier

Korotaev and Yanilkin[5] developed an RF classifier model from a dataset of 647 steels that included composition and a two-stage heat treatment, e.g., quench and temper. They reduced the number of elements used as input features from the original 18 to only the elements carbon (C), Ni, and Cr in addition to the heat-treatment descriptors (temperature and quench rate) through recursive feature elimination with cross-validation. This selection also aligns with the general concept of the Schaeffler diagram where Ni is an austenite stabilizer and Cr is a ferrite stabilizer. With these input features, their model achieved an excellent prediction for their test set with an R^2 of 0.94 compared with that of 0.76 for the same alloys using thermodynamic software (Thermo-Calc). The authors did not

describe all the parameters used for their RF model in their work so, in our efforts to reproduce the model with the same test/train split, we had a slightly higher R^2 of 0.96 (Figure 10a). As both a large portion of the larger dataset assembled for steel classification in this work is only a single-state heat treatment (e.g., austenitize and quench rather than austenitize and quench followed by a temper and quench), and 44% of the original dataset[5] has room temperature for the temper temperature (indicating that it was also a single-stage heat treatment), we retrain their model using only the first stage heat treatment as inputs. This makes a minor difference only in the prediction of the "FM" class (Figure 10b).

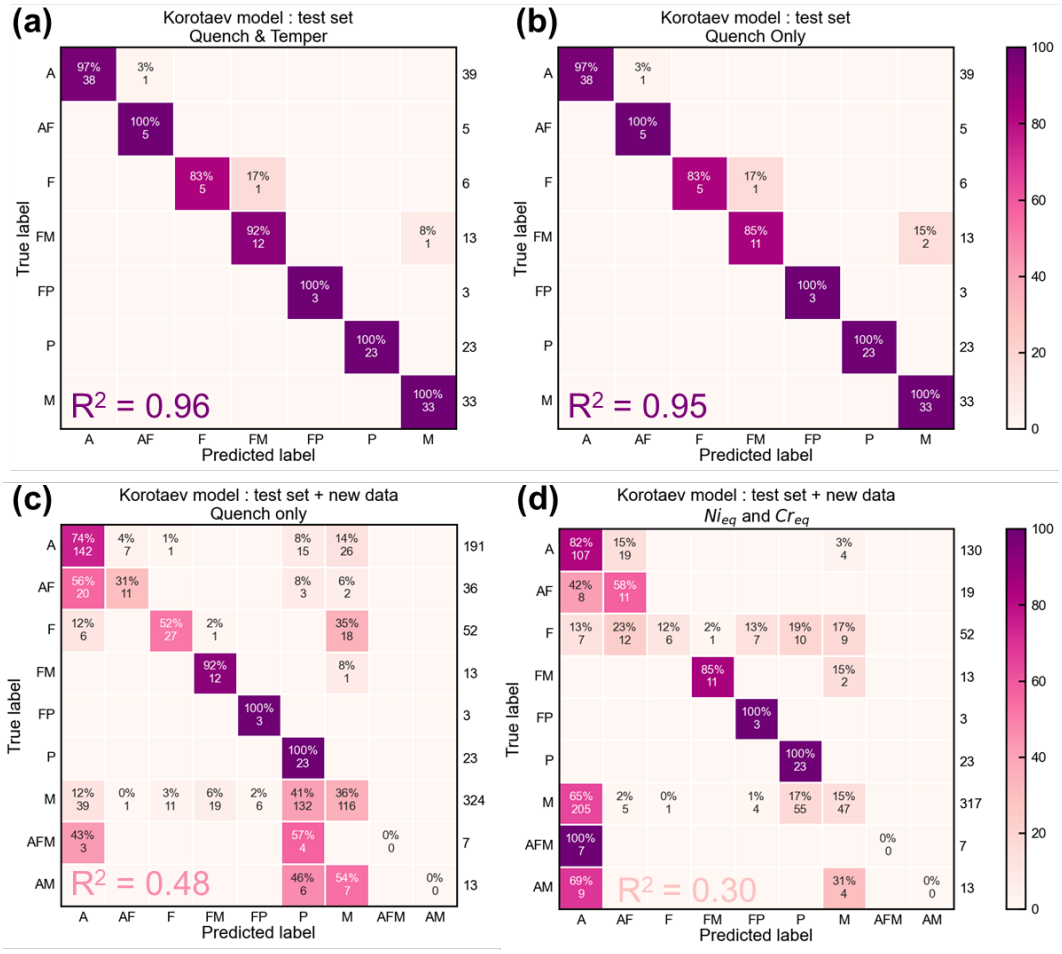

Figure 10. RF model predictions of (a) original model using C, Ni, Cr, and two heat-treatment steps as inputs for the original test set; (b) original model using C, Ni, Cr, and only the first heat-treatment step for the original test set; (c) model using inputs of (b) but predicting the original test set and new data; and (d) model of (b) but using Ni_{eq} and Cr_{eq} instead of Ni and Cr, predicting the original test set and new data.

The RF model trained following Korotaev and Yanilkin[5] (i.e., with only the original training set and choice of elemental inputs) is used to predict the class of the new classifier dataset (Figure 10c). The original Korotaev and Yanilkin training data did

not have any "AM" or "AFM" labels so it cannot successfully predict those new data points. There is also a significant drop in performance for all classes with new data, especially for martensite.

The introduction of additional steels to the dataset would be expected to affect the relative importance assigned to each element. A category of austenitic steels of interest to the Navy that were added to the dataset include those which aim to replace the Ni with manganese (Mn),[18] another austenite stabilizer that is less effective but cheaper. All the cases where the true class is "A" but the model of Korotaev and Yanilkin predicted "M" or "P" have high Mn and no Ni. This could be addressed by using Ni and Cr equivalencies in place of the original Ni and Cr, and the results are shown in Figure 10d. While this improved the austenite prediction, it made everything else worse. Clearly the model must be retrained, including new data in the training set and a new assessment of the essential elements for input features made.

4.4 Retraining RF Classifier

We retrained the RF classifier model with several different combinations of input features on the full classifier dataset, including the new data plus the original data from Korotaev and Yanilkin.[5] First, we use all elements as input features, which is likely overfitting. From the analysis of these inputs, we down-select a shorter list of elements. We also try the Ni and Cr equivalencies as inputs in place of the elements—both the DeLong[16] as discussed above, and a newer set by Lee et al.[17] that incorporate more elements. Both their coefficients are shown in Table 1.

Table 1. Coefficients for Ni and Cr equivalencies.

Ni_{Eq}	DeLong[16]	Lee[17]	Cr_{Eq}	DeLong[16]	Lee[17]
Ni	1	1	Cr	1	1
Mn	0.5	0.5	Mo	1	1.5
C	30	30	Si	1.5	2
N	30	25	Nb	0.5	1.75
Co	. . .	1	Al	. . .	3.5
Cu	. . .	0.3	V	. . .	5
			Ti	. . .	1.5
			W	. . .	0.75

The RandomForestClassifier from scikit-learn is used with the default 100 trees to replicate the assumed parameters of the original Korotaev and Yanilkin[5] model. The high number of trees limits scientific explainability, but we are aiming for high-throughput predictive capability.[19] The dataset was split into test and training sets 100 times. The same random_seed was used for both the test/train split and bootstrap. SHapley Additive exPlanations (SHAP) analysis[20] was used to examine

the effect of the various input features when using all elements. SHAP features are calculated by class.

The impact of the input features averaged over 100 test/train splits are shown in Figure 11, colored by class. Similar to the initial work by Korotaev and Yanilkin,[5] Ni, Cr, and C are highly important, and Mn is additionally significant in part due to the Ni-replacement austenitic steels, but it is also a strong influence on the martensite prediction presumably due to its enhancement of hardenability. A side-by-side comparison of a subset of classes in Figure 12 is more illustrative of the difference in impact of various features on each class. Some perplexing results emerge, such as aluminum (Al) having a stronger effect on martensite formation than the quench rate. A detailed analysis of the results relative to well-established effects of alloying element in the literature is outside the scope of the AMPED program where the focus is on high-throughput, but will be considered in the future.

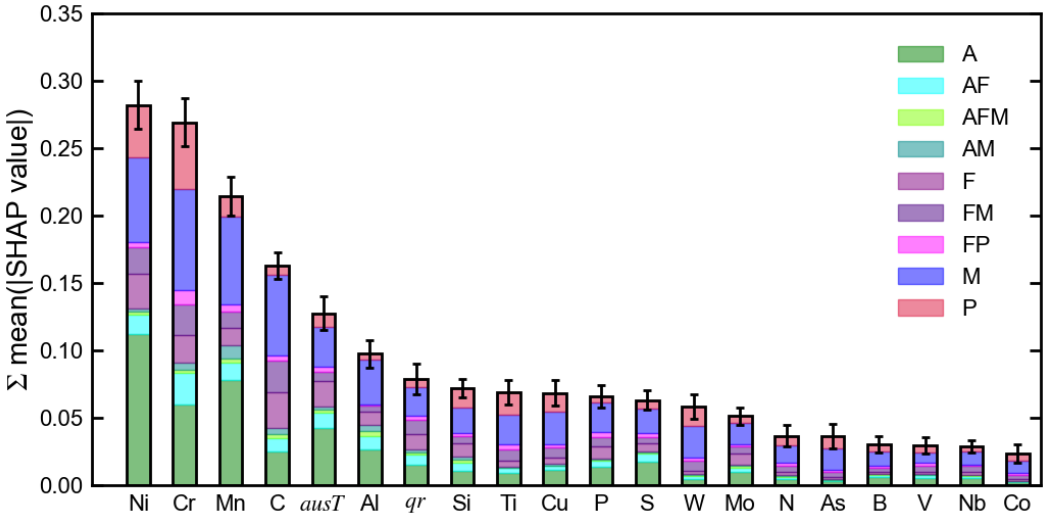

Figure 11. Sum across classes of average SHAP values input features for RF classifier model trained on all elements and *ausT* and *qr* as input features across 100 test/train splits. Features are ranked by total impact across all classes.

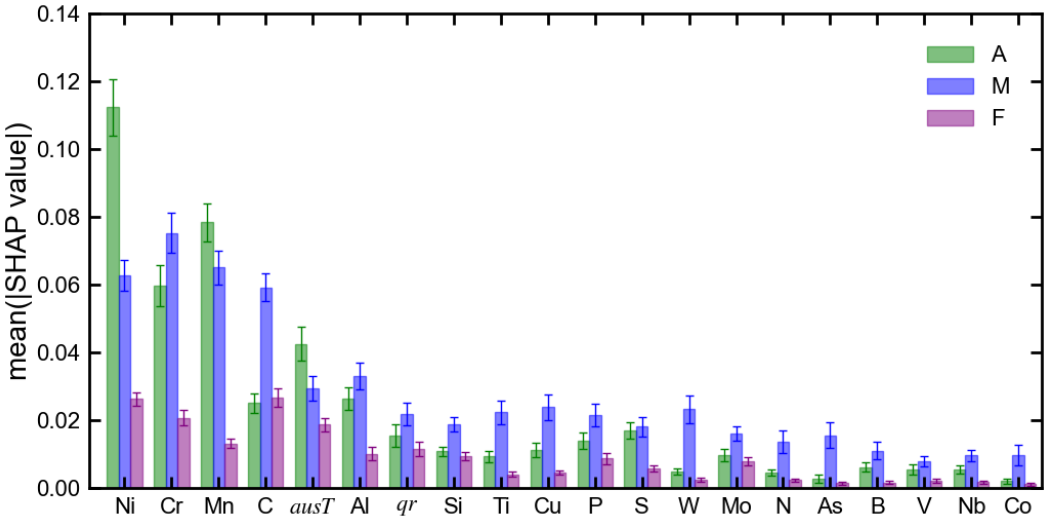

Figure 12. SHAP values separated by class for RF classifier model trained on all elements and *ausT* and *qr* as input features across 100 test/train splits. Features are ranked by total impact across all classes.

The predictions on the test set using various input features are shown in Figure 13. The test sets are slightly different for each case due to the variations across the 100 test/train split permutations. The "some elements" option for input features uses only the elements Ni, Cr, C, Mn, and Al in addition to the austenitization temperature and quench rate. The inputs for the DeLong[16] and Lee et al.[17] equivalencies include the Ni_{eq} and Cr_{eq} calculated from their respective equations in Table 1 in addition to C, and the austenitization temperature and quench rate. All four of these options do quite well.

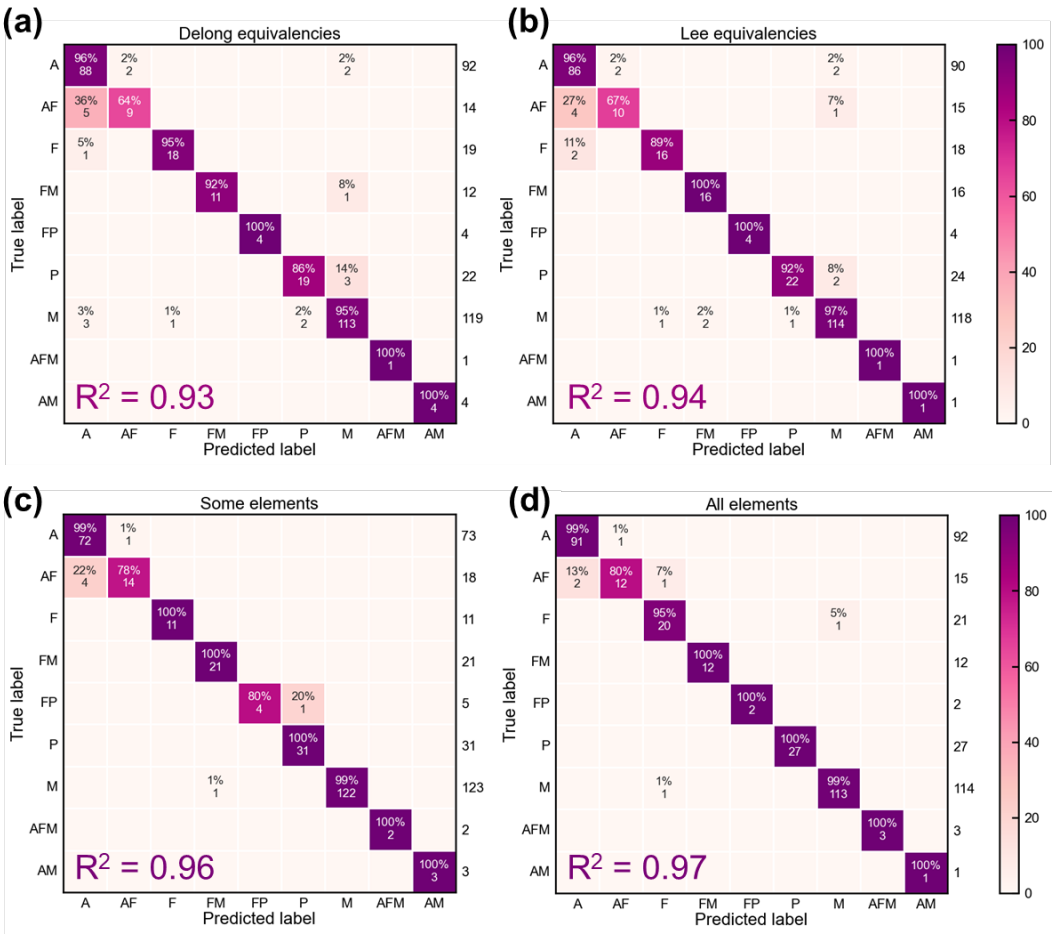

Figure 13. RF classifier model trained using different input features on the full dataset: (a) DeLong[16] Ni/Cr equivalencies + C; (b) Lee[17] Ni/Cr equivalencies + C; (c) some elements; and (d) all elements.

The trained classifier model also predicted correctly all ARL and NRL alloys referenced in the AMPED program that are not in the steel classifier dataset; this includes those plotted in Figure 1 plus some additional steels in the same families.

5. Application

In a high-throughput material design and testing campaign, the classifier model could be used to predict the steel class for a given composition and heat treatment and can direct which models, experiments, or experimental settings should be used next. Furthermore, the probability associated with the prediction can be used for uncertainty quantification.

An example composition space, with variation in Mn and Cr content for a C content of 0.05 wt% is shown in Figure 14. The austenitization temperature was set at 1000 °C and the quench rate at 100 °C/s (~oil quench). The composition space is

colored by the predicted steel class with contour levels based on the predicted class probability. The highest probabilities are the high Mn/Cr austenite prediction and the very low Mn/medium Cr martensite prediction. The probabilities are low (i.e., high uncertainty) at the boundaries between classes. This calculation comprising 1681 compositions took less than 0.1 s to complete on a standard laptop.

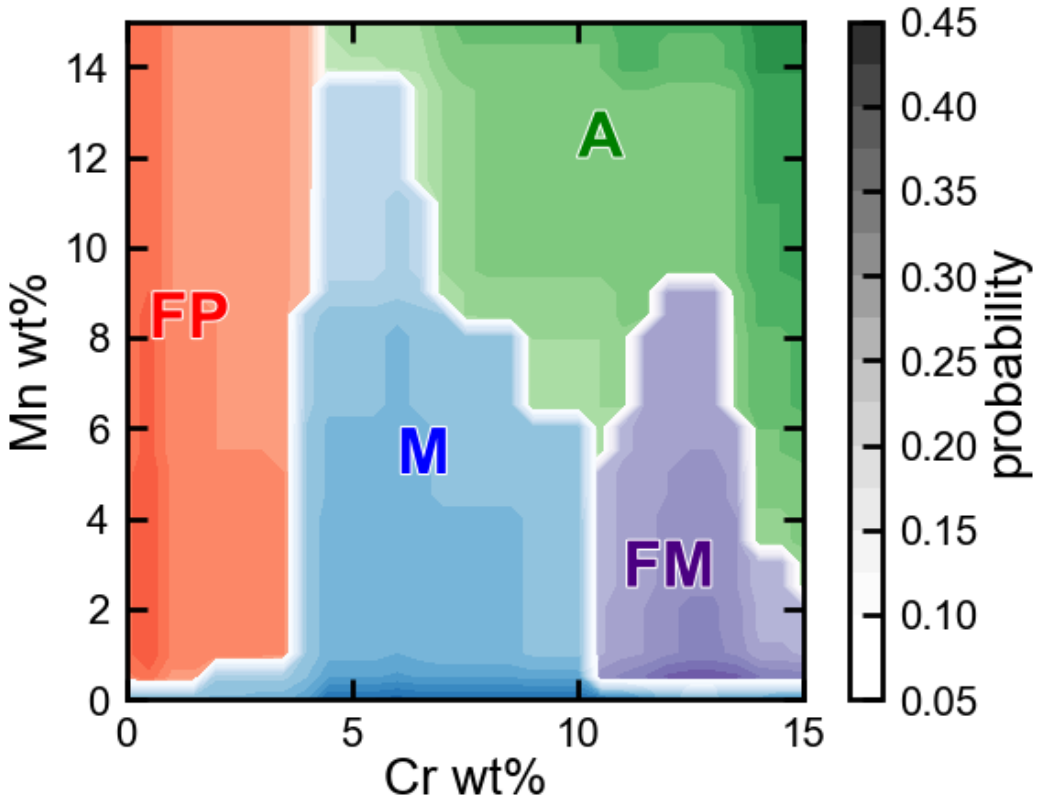

Figure 14. Predicted phase map of Fe-xCr-yMn-0.05C with an austenitization temperature of 1000 °C and a quench rate of 100 °C/s using the classifier model with all elements as inputs. Color scale is based on the probability associated with the prediction of the class.

6. Conclusions

Several modeling approaches to predict the structure of a given steel based on composition and processing were evaluated using a dataset constructed from open literature. Significantly, these modeling approaches do not require use of commercial software and have low computational expense. First, a significantly expanded dataset of martensite transformation kinetics was assembled and used to develop an improved martensite kinetics model. This model will aid in design of martensitic steels, particularly where retained austenite is a concern and could better inform industry design of quenched and partitioned steel. Secondly, a large steel classifier dataset was assembled and used to evaluate several methods to predict microstructure categories based on composition and heat treatment. The

predictive capability of existing methods, including empirically predicted CCT diagrams, the Schaeffler diagram, and a pretrained ML model were explored on the large classifier dataset and their respective limitations noted. Finally, a new RF classifier model was trained and was highly successful at predicting the structure. This predictive capability will be useful in accelerated design workflows, including integration with predictive modeling tools and guiding high-throughput experimental strategies.

7. References

1. Leclerc N, Khosravani A, Hashemi S, Miracle DB, Kalidindi SR. Correlation of measured load-displacement curves in small punch tests with tensile stress-strain curves. Acta Mater. 2021;116501.

2. Murdoch HA et al. Tempered hardness optimization of martensitic alloy steels. Integr Mater Manuf Innov. 2023;12:301–320.

3. van Bohemen SMC, Sietsma J. Effect of composition on kinetics of athermal martensite formation in plain carbon steels. Mater Sci Technol. 2009;25:1009–1012.

4. Huyan F, Hedström P, Höglund L, Borgenstam A. A thermodynamic-based model to predict the fraction of martensite in steels. Metall Mater Trans A. 2016;47:4404–4410.

5. Korotaev P, Yanilkin AV. Steels classification by machine learning and Calphad methods. Calphad. 2023;82:102587.

6. Koistinen DP, Marburger RE. A general equation prescribing the extent of the austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels. Acta Metall. 1959;7:59–60.

7. Lee S-J, Van Tyne CJ. A kinetics model for martensite transformation in plain carbon and low-alloyed steels. Metall Mater Trans A. 2012;43:422–427.

8. Gruber M et al. Microstructural effects on the toughness of a high Co–Ni steel. Adv Eng Mater. 2017;19:1600633.

9. Cranmer M. Interpretable machine learning for science with PySR and SymbolicRegression.jl. arXiv; 2023 May 2. arXiv:2305.01582. https://doi.org/10.48550/arXiv.2305.01582

10. Miettinen J et al. Optimization of CCT equations using calculated grain boundary soluble compositions for the simulation of austenite decomposition of steels. Metall Mater Trans B. 2019;50:2853–2866.

11. Miettinen J et al. Optimization of the CCT curves for steels containing Al, Cu and B. Metall Mater Trans B. 2021;52:1640–1663.

12. Schaeffler AL. Constitution diagram for stainless steel weld material. Metal Progress. 1949;56:680.

13. Brandi SD, Schön CG. A thermodynamic study of a constitutional diagram for duplex stainless steels. J Phase Equilib Diff. 2017;38:268–275.

14. Ghosh G, Olson GB. Kinetics of FCC→ BCC heterogeneous martensitic nucleation—I. The critical driving force for athermal nucleation. Acta Metall Mater. 1994;42:3361–3370.

15. Guiraldenq P, Hardouin Duparc O. The genesis of the Schaeffler diagram in the history of stainless steel. Metal Res Technol. 2017;114:613.

16. DeLong WT. Effect of nitrogen in stabilizing austenite phase. Metal Process. 1960;98.

17. Lee S, Chan-Young L, Young-Kook L. Schaeffler diagram for high Mn steels. J Alloys Compd. 2015;628:46–49.

18. Stewart CA, Fonda RW, Knipling KE. Mn-stabilized austenitic steel strengthened by nano-scale β-NiAl (B2), FCC-Cu, and carbides via ICME design. Scr Mater. 2021;200:113903.

19. Holm EA. In defense of the black box. Science. 2019;364:26–27.

20. Lundberg SM et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020;2:56–67.

List of Symbols, Abbreviations, and Acronyms

A	Austenite
AF	Austenite–Ferrite
AFM	Austenite–Ferrite–Martensite
Al	Aluminum
AM	Austenite–Martensite
AMPED	Accelerating Materials Performance Evaluation and Discovery
ARL	Army Research Laboratory
ausT	Austenitizing Temperature
C	Carbon
CCT	Continuous Cooling Transformation
Co	Cobalt
Cr	Chromium
Cu	Copper
DEVCOM	U.S. Army Combat Capabilities Development Command
F	Ferrite
FM	Ferrite–Martensite
FMP	Ferrite–Pearlite–Martensite
FP	Ferrite–Pearlite
HiPo	High Performance
KM	Koistinen–Marburger
M	Martensite
ML	Machine Learning
Mn	Manganese
Mo	Molybdenum
Mod HP	Modified High-Performance
MP	Martensite–Pearlite

N	Nitrogen
Nb	Niobium
Ni	Nickel
NRL	U.S. Naval Research Laboratory
P	Pearlite
qr	Quench Rate
RF	Random Forest
SHAP	SHapley Additive exPlanations
Si	Silicon
SPT	Shear Punch Testing
TC	Thermo-Calc
Ti	Titanium
V	Vanadium
W	Tungsten

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLB CB TECH LIB