

Screening of heritage data for improving toughness of creep-resistant martensitic steels

Amit K. Verma^a, Wei-Heng Huang^{a,c,1}, Jeffrey A. Hawk^b, Laura S. Bruckman^{a,c}, Roger H. French^{a,c}, Vyacheslav Romanov^d, Jennifer L.W. Carter^{a,*}

^a Department of Material Science and Engineering, Case Western Reserve University, Cleveland, OH, 44106, USA

^b National Energy Technology Laboratory, Albany, OR, 97321-2198, USA

^c SDLE Research Center, Case Western Reserve University, Cleveland, OH, 44106, USA

^d National Energy Technology Laboratory, Pittsburgh, PA, 15236-0940, USA

ARTICLE INFO

Keywords:

Martensitic Steels

Metal design

Plasticity

Data science

ABSTRACT

Data science techniques were used to quantify the effect of alloying additions on the tensile behavior of martensitic steels. The effort was undertaken to exploit the heritage data to establish the next experimental design space for the class of 9–12 wt% Cr steels for the application of turbine rotors with an operating temperature of 650 °C and above. Linear, lasso, and multivariate multiple regression models were utilized to identify which alloying elements contribute towards strength and ductility. Visualization techniques such as t-distributed stochastic neighbor embedding and pair-wise element specific comparisons were utilized to explore information gaps that exist within the data. The study found that tantalum, recently added to improve the creep rupture lifetime, does not show any effect on tensile properties. All combined, the results suggest that the low tempering temperature has compensated for the low alloying additions in the past, therefore, new experiments are needed to isolate the effects of tempering temperature from those of individual elements.

1. Introduction

Energy initiatives to lower carbon dioxide (CO₂) emissions are pushing the adoption of advanced ultra-supercritical (A-USC) power plants with increased steam temperature [1]. The increase in temperature requires development of new martensitic steel alloys, in the 9–12 wt% Cr class, for various applications such as turbine and boiler components for target temperatures ≥ 650 °C. These applications demand alloys with sufficient creep rupture strength and/or fracture toughness with adequate tensile strength and ductility to avoid crack formation and sustain a lifetime of 10⁵ h at 100 MPa or above. Alloy development programs in last decades in Japan and United States have relied heavily on intuitive trial-and-error approaches [2], and have resulted in alloys that can achieve operating temperature of 630 °C for turbine rotors and 610 °C for boilers. Additionally, these programs have resulted in a curated database of alloy compositions and properties that can be explored and exploited for future developments.

In this paper, a data-driven approach has been taken to identify the effect of alloying elements on the tensile behavior of martensitic steels for the application of turbine rotors operating at 650 °C and above. As

turbine components are relatively larger in size than boiler components [3], the application demands greater tensile toughness. The aim is to exploit the heritage data and guide the next phase of experiments by gaining insights from previously completed experiments to effectively reduce the time and cost for materials discovery. While there are numerous studies for creep strengthening mechanisms of 9–12 wt% Cr martensitic steels [4,5], high temperature tensile deformation mechanisms of 9–12 wt% Cr martensitic steels are not fully explored [6].

The 9–12 wt% Cr martensitic steels show temperature-dependent tensile deformation mechanisms, with athermal dislocation motion at low temperature and diffusion mediated dislocation motion at high temperatures [7]. Recent studies by Wang et al. [6,8], have correlated this transition with changes in the load partitioning among the three phases: Fe-matrix, carbides ($M = \text{Cr, Mo, W, Fe}$) and carbonitrides ($M = \text{V, Nb, Ta}$ and $X = \text{C, N}$), with temperature. At temperatures significantly below the austenitic start temperature ($\sim 0.65 A_{c1}$ [9] (~ 450 °C) and below), load transfers from Fe-matrix to precipitates, thereby strengthening the material. As deformation progresses, dislocation density increases resulting in strain hardening, followed by the formation of dislocation walls and cells [8]. Void nucleation may result

* Corresponding author.

E-mail address: jennifer.w.carter@case.edu (J.L.W. Carter).

¹ Current address: Department of Statistics, Feng Chia University, Taichung, Taiwan.

from de-cohesion at the precipitate/matrix interface and/or by the breaking of precipitates [10]. At high temperatures ($\sim 0.65 A_{c1}$ and above), dislocations in the Fe-matrix can bypass precipitates by diffusion-assisted climbing resulting in limited load transfer from matrix to precipitates. As deformation progresses, the dislocation mean free path increases and dynamic recovery coefficient decreases with the annihilation of free dislocations [11]. The nature of dislocation shifts from mainly edge type at room temperature to mainly screw type at high temperatures.

Toughness is a measure of the ability of a material to absorb energy without fracture, and is a product of strength and ductility. Therefore, both properties can be exploited to enable high tensile toughness. However, often materials suffer from strength-ductility trade-off and this behavior has been seen in martensitic steels [12]. For example, at room temperature severe plastic deformation (SPD) accompanied by post annealing increases strength by increasing the dislocation density and reducing the effective grain size, but reduces the work hardening capability and thus the overall ductility [13,14]. Although, recent developments [15] show that inter-pass annealing during SPD techniques can improve the work hardening capability. The ability to use SPD techniques for complex geometries and large parts is limited, thus requiring alternate design approaches.

Several strategies that had been proposed to obtain a combination of high strength and good ductility focuses on 1) increasing the number of potential deformation mechanisms [16,17], or 2) adding a soft phase such as metastable austenite grains embedded in a highly dislocated martensite matrix, and inducing a TRIP effect [18,19]. These strategies are beyond the scope of this work as the analysis is limited by the data collected and does not allow tailoring of strengthening mechanism. Most of the literature that encompasses the class of 9–12 wt% Cr martensitic steels, focuses on strength while retaining ductility, as increases in temperature with A-USC standards decreases strength rapidly while increasing ductility. This gives us an opportunity to explore the given data with a focus on strength while maintaining ductility to evaluate the possible solutions for next phase of experiments.

2. Methods

Data used for the analysis comes from two sources: 1) National Energy Technology Laboratory (NETL) – Albany, an agency of the United States Government; and 2) MatNavi database provided by National Institute for Materials Science (NIMS), Japan [20–28]. The collected dataset contains a total of 869 tensile tests results at different test temperatures (ranging from room temperature to 800 °C) and processing parameters such as composition and heat treatment. The compositional space comprises 18 different elements, with 79 unique compositions. Using t-distributed stochastic neighbor embedding (or t-SNE) cluster analysis [29] (subsection 3.1) and historical naming conventions [30], alloy sub-classes were assigned to group similar compositions (Table 1). The alloy specifications ensure a fully tempered martensitic microstructure by balancing the austenite and ferrite stabilizers. The heat treatment process comprises a solution-annealing treatment (or normalization), a quench to form martensite, followed by a tempering treatment.

2.1. Visualization using t-SNE: Cluster labels

The t-distributed stochastic neighbor embedding maps the similarities between objects of N-dimensional space as a Gaussian distribution and provides a 2-dimensional visual projection by converting the N-dimensional Gaussian distribution to a 2-dimensional t-distribution [29]. The goal is to find the mapping that minimizes the differences between Gaussian distribution and 2-dimensional t-distribution over all data points. The technique uses the Barnes-Hut approximation [31], and models the probability distribution of neighbors around each data point. For this work, the R language [32] (R) package developed by

Table 1
Composition and heat treatment ranges for the alloy sub-classes of the class of 9–12 wt% Cr martensitic steels used in the analysis. Main alloying additions are shown with units in wt%. Elements such as C, Si, Al, P, N, S, and remaining Fe content are not included in the table. Heat treatment temperature ranges are presented as TPR for tempering temperature (°C), and NML for normalization temperature (°C). NA represents that some of the observations within the sub-class do not contain the element.

Sub-Class	Cr	Mn	Ni	Mo	W	Co	Cu	Nb	V	Ta	B	NML	TPR
(△) 12Cr	11.64–12.38	0.42–0.71	0.15–0.45	0.04–0.21			0.03–0.21					950–980	630–750
◇ 9Cr–1Mo	8.50–9.15	0.38–0.50	0.047–0.120	0.93–1.05			0.02–0.08					900–970	720–860
(□) 9Cr–2Mo	9.04	0.5	0.11	2.03			0.02	0.006	0.01			950	800
(⊗) 9Cr–1.5Mo–V–Nb–B	9.11–9.29	0.05–0.06	0.08–0.09	1.54–1.57				0.055–0.059	0.26–0.29		0.0082–0.0088	1038–1221	610–732
(□) 9Cr–1.5Mo–1.3Co–V–Nb–B	9.28–9.31	0.31–0.34	0.15	1.51–1.54		1.25–1.33	0.02 (NA)	0.050–0.052	0.19–0.20		0.0091–0.0096	1095–1100	660–700
(○) 10.5Cr–1Mo–1W–V–Nb	9.94–10.50	0.41–0.55	0.67–0.77	0.97–1.05	0.005 (NA)	0.02 (NA)	0.02 (NA)	0.041–0.050	0.17–0.19		0.050	1050	660–690
(▽) 12Cr–1Mo–1W–Co–V–Cu	11.00–12.90	0.55–0.87	0.64–0.90	0.91–1.22	0.90–1.13	0.04–0.15	0.03–0.15		0.21–0.30			1030–1050	630–660
(⊕) 9Cr–1Mo–V–Nb	8.31–8.74	0.35–0.49	0.04–0.28	0.89–0.99			0.012–0.032 (NA)	0.048–0.09	0.185–0.23			1045–1060	760–790
(■) 9Cr–Mo–1.8W–V–Nb–B	9.26–9.50	0.41–0.42	0.13–0.17	0.36–0.42	1.67–1.74			0.057–0.062	0.16–0.19		0.002	1070–1100	780
(●) 10.5Cr–Mo–2W–V–Nb–Cu–B	10.65–10.73	0.60–0.63	0.32–0.36	0.33–0.38	1.87–1.97		0.85–0.97	0.050–0.056	0.19–0.22		0.0024–0.0039	1050	770
(▲) 12Cr–Mo–1.8W–V–Nb–Cu–B	12.10	0.59	0.33	0.34	1.82		0.82	0.06	0.19		0.003	1050	790
(⊙) 9Cr–3W–3Co–V–Nb–Ta–B	9.08	0.51	0.18–0.19	0.10	894–2.914	2.926–2.942	0.003–0.032	0.062–0.065	0.202–0.208	0.003–0.349		0.0101–0.0121	1150
(⊙) (Cr)–(Mo)–0.5W–(Co)–V–Nb–B–(Ta)	9.11–10.74	0.288–0.473	0.20–0.31	0.502–1.464	0.428–0.596	0.012–8.226	0.002–0.295 (NA)	0.049–0.062	0.190–0.237	0.003–0.361 (NA)		0.0075–0.0113	700

Jesse H. Krijthe [33] is applied. Within the designed function, the main parameter “perplexity” controls the number of nearest neighbors considered when matching the original and fitted distributions for each point. For the inherently similar groups within the compositional space, t-SNE algorithm was applied with a perplexity factor of three [34].

2.2. Change-point detection

The change-points are coordinated where two different linear relationships (i.e., with different slopes) between the same variables converge. The **R** segmented package [35] uses the maximum likelihood to fit a piecewise model where the method constrain the linear segments to be (nearly) continuous [36]. The function takes a generic linear model and a best-guess estimate of the change point as its arguments, and iteratively varies these parameters to get the best fit by minimizing the overall error while keeping the segments continuous. For the given dataset, the initial change-point best guess was 450 °C, where the tensile deformation mechanism changes.

2.3. Contributors: subset selection

The subset selection approach involves identifying the subset of predictors that are contributors to the maximum variation in the response, assuming a linear relationship between predictors and the response. For this study, predictors are the set *x*: material composition, heat treatment, and test temperature; the responses are the set *y*: yield strength (YS), ultimate tensile strength (UTS), reduction in area (RA), and elongation to failure (El). The **R** leaps package [37] performs an exhaustive search of all the subsets to find the best subset of the predictors in *x*, for predicting an individual response *y* using a branch-and-bound algorithm [38]. The algorithm returns the best subset of predictors for a fixed subset size, starting with a single predictor to the maximum number of available predictors in the set *x*. For the study, adjusted- R^2 (adj- R^2) and Akaike information criterion (AIC) were used to define the best subset. The subset selection process does not penalize for model complexity, therefore, a subsequent study of trade-off between bias and variance is required to guide the selection of number of contributors needed to minimize over-fitting and under-fitting [39].

2.4. Lasso regression

In 1970, Hoerl and Kennard [40] showed that coefficient estimates based on linear regression suffer if predictors are dependent. In response, they introduced “ridge trace” to show these effects and to obtain model estimates with smaller mean square error by controlling the inflation and general instability associated with least square estimates. Later (1996), Tibshirani [41] introduced lasso regression with “lasso penalty” (or “least absolute shrinkage and selection operator”), to shrink some coefficients and make other coefficients zero. This method retained the features of ridge regression and introduced an automatic contributor selection effect, which makes process less susceptible to high variance [39,42]. The **R** glmnet package developed by Friedman et al. [42,43], was utilized. It performs lasso regression using cyclical coordinate descent methods [44].

2.5. Multivariate multiple regression

Multivariate multiple regression (MMR) models build upon the foundation of linear regression model, by allowing more than one response variable in the set *y* to be simultaneously described by the same set of predictors in the set *x* [45,46]. MMR models follow a similar contributor selection process via a step-wise exhaustive search, although, constrained by AIC for all responses *y*, simultaneously. As the selection process accounts for multiple responses, a selected predictor at different steps may or may not strongly correlate with an individual response and is reflected in the β coefficients of the resulting models.

The **R** qtlmt package [47] (function **mStep**) performs the exhaustive search using AIC [48] values. These models allow the comparison of contributors across different response variables, and are useful where responses are dependent (i.e., strength-ductility trade-off).

2.6. Outliers: Cook's distance

Cook's distance [49] is a method of defining outliers in a multi-contributor model space and is computed with respect to a given regression model. It computes the influence exerted by each observation *i* on the predicted outcome by measuring the change in fitted values for all observations with and without the presence of observation *i* in the regression model. The **R** standard ‘stats’ package [32] was used to calculate Cook's distance.

2.7. Model validation

The validation of regression models was conducted by separating the dataset into train and test dataset. The train and test dataset was selected randomly, where 80% data is assigned for the training of the model and rest 20% for testing of the model. A model is assessed by taking the ratio of the errors for test:train dataset, where a value close to one indicates that the test error is close to training error and the model is not overfitting the data. The random selection of train and test dataset was repeated 1000 times to generate a distribution of test:train error ratio. While this ratio is a strong indicator of overfitting, it does not provide any information about the underfitting. The underfitting was assessed during the subset selection process through increase in adj- R^2 with each addition of predictor.

3. Results and analysis

3.1. Cluster labels

The evaluation of t-SNE cluster analysis for the processing variables (composition and heat treatment) yielded 11 groups (Fig. 1). Of the 11 groups, two groups where alloy sub-class labels overlap on t-SNE map were identified: **9Cr-2Mo** (⊞) with **9Cr-1Mo** (◇); and **12Cr-Mo-2W-V-Nb-Cu-B** (▲) with **10.5Cr-Mo-1.8W-V-Nb-Cu-B** (●). In both pairs, composition differs with respect to a single alloying element. One new sub-class label (**Cr**) **(Mo)0.5W(Co)VNbB(Ta)** (•) was created to accommodate a group of compositions that did not fall into a previously defined sub-class. The alloy compositions within the (**Cr**) **(Mo)0.5W(Co)VNbB(Ta)** (•) sub-class on t-SNE map is governed by high variations in **Ta**, **Co**, and **Cu**. The visualization, together with allocation of sub-classes, highlights the limited processing space, where 91 unique combinations of composition and heat treatment cluster as 11 distinct groups.

3.2. Change-point behavior of tensile data

The visualization of tensile response variables as a function of test temperature (Fig. 2a) indicates a linear change-point behavior. The change-point phenomenon is associated with a transition in dominant deformation mechanism. The legends in Fig. 2a highlight the change-point temperature for each response for this unique composition. Due to changes in deformation mechanisms, data separation into low vs high temperature regions is needed prior to conduct training of regression models. The separation of data can be done using the calculated change-points, although separation of data using change-points for different responses would result in different training sets and inhibit the comparison of contributors among responses. Since mechanistically there is one change-point for a unique composition for all four responses, and the differences (Fig. 2a) emerge due to variation in data on either side of change-point, a single change-point for a unique composition can be considered for the separation of data. The residuals

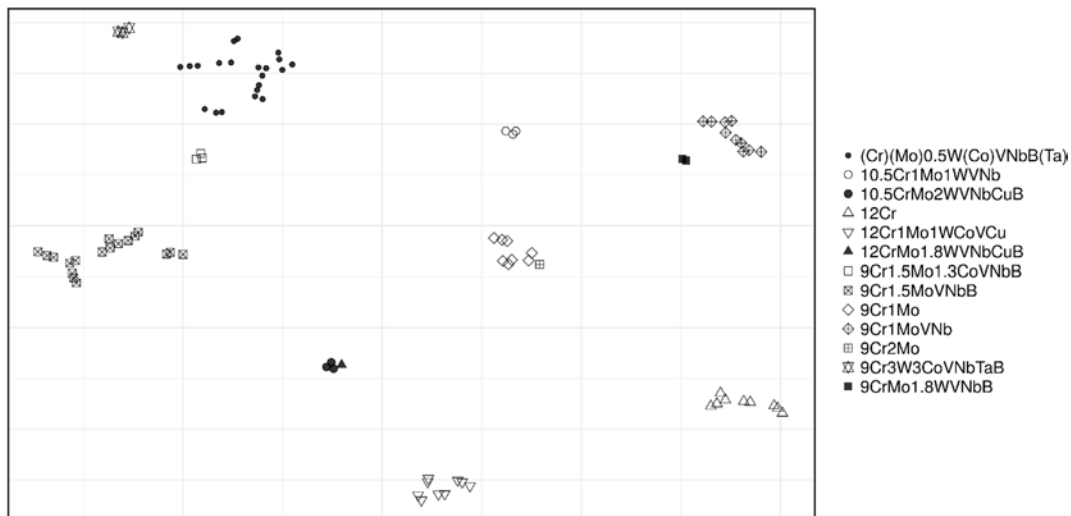


Fig. 1. Visual projection of 20-dimension processing space into 2-dimensional space using t-SNE [29,31,33] shows that the 91 unique combinations of composition and heat treatment fill the design space in a sparse manner. 11 unique clusters are identified; ten of which include previously identified sub-class labels (with two clusters containing two different sub-class labels) and one new sub-class defined for this paper. Twenty dimensions include composition (18 elements) and heat treatment (Table 1). Note: measuring the distances or angles between points in above plot does not give any quantitative information about the data.

(difference between predicted and observed values) (Fig. 2b) for each fit of the four tensile property variables show UTS change point model as the best fit (least deviation from the solid horizontal line) among all four fits for a single composition within **9Cr–1Mo–V–Nb** (\oplus) sub-class. This is consistent, but not visualized, for all other compositions. This suggests that the change-point calculated using the UTS versus temperature curve is stable and least subject to errors due to nonlinearity on either side of change-point (i.e., RA and El fit). Therefore change-points calculated from UTS versus temperature curves for all 79 compositions led to the separation of data (Fig. 3) into high (open triangles) and low (filled circles) temperature regions.

3.3. Strain hardening

The ratio of UTS to YS is a measure for the degree of strain hardening (SH), where values 1.4 are considered relatively high SH, and those < 1.2 indicate relatively low SH [50]. The calculation of SH

indicates that **9Cr–1Mo** (\diamond) sub-class as inherently different than the rest of the sub-classes. A visualization (Fig. 4a) of the change point behavior of the SH for the **9Cr–1Mo** (\diamond) and **9Cr–2Mo** (\boxplus) (as a surrogate for all other sub-classes) shows that the SH mechanisms as a function of temperature are different. There could be a variety of explanations for this difference in behavior, for example it could indicate that tensile tests for **9Cr–1Mo** (\diamond) sub-class were conducted at a different strain rate, or it could indicate a difference in microstructural evolution, but without additional information these theories are conjectures. At low temperatures, SH increases with temperature for **9Cr–1Mo** (\diamond), whereas it decreases for **9Cr–2Mo** (\boxplus) and other sub-classes (Fig. 4b). Symmetric differences are observed for high temperature region. As **9Cr–1Mo** (\diamond) sub-class is inherently different, it is removed from training/testing of models.

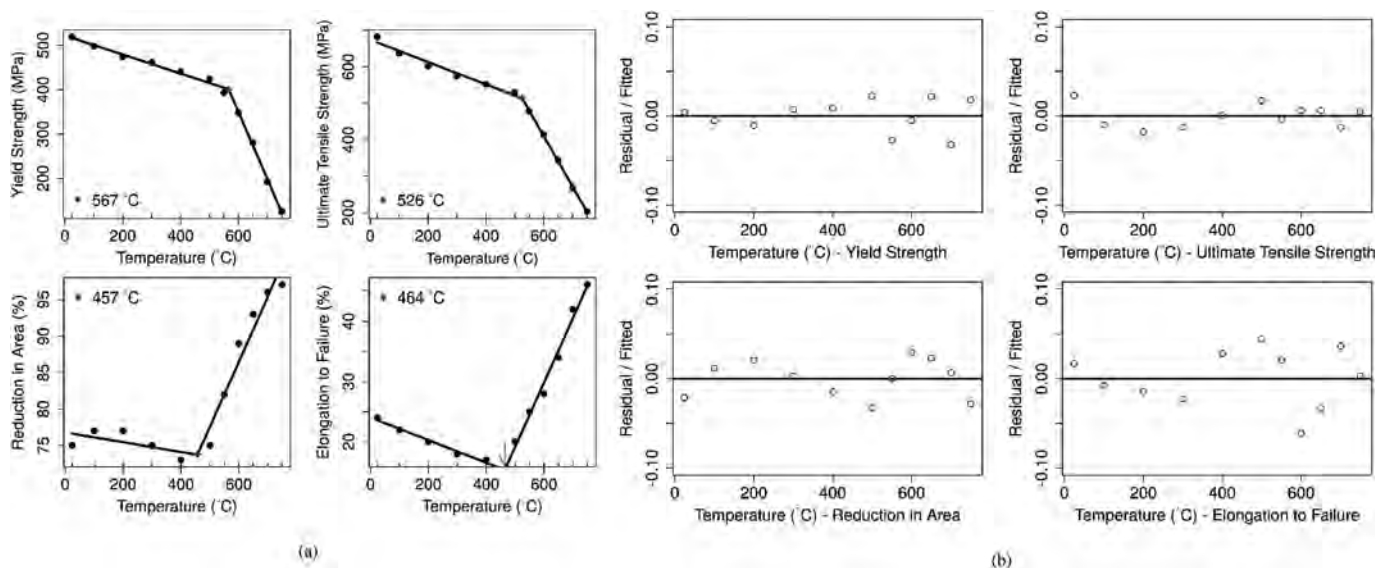


Fig. 2. (a) Variation in tensile test outcomes with test temperature for one alloy composition in **9Cr–1Mo–V–Nb** sub-class. The visualization highlights a change-point phenomenon associated with mechanism change with increase in temperature. Legend in each plot represents a calculated change-point temperature. (b) Residuals after fitting a change-point model for each response. Residuals were scaled with respect to fitted values for comparison among different models.

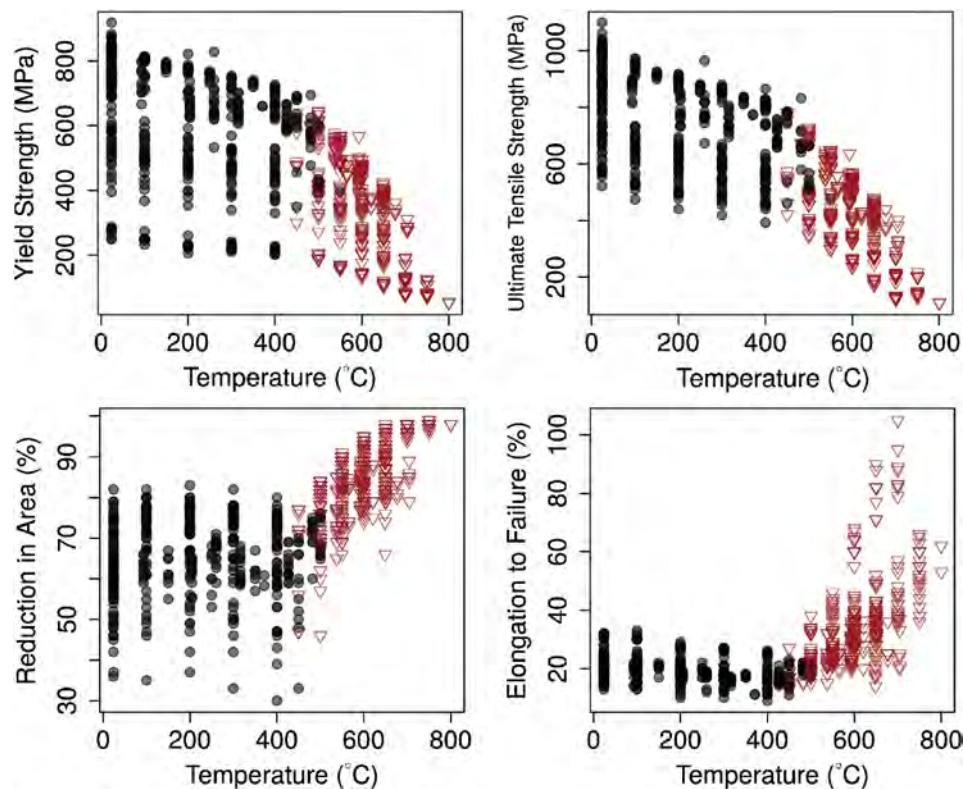


Fig. 3. Variation in tensile test outcomes with test temperature for all 869 observations shows properties as a function of composition and heat treatment. The figure highlights how all sub-classes exhibit the change point behavior, indicated by the low (filled circles) and high (open triangles) temperature regions.

3.4. Lasso corrected contributors

After the separation of the data into low and high temperature regimes, the spread in an observed tensile property at a particular test temperature in Fig. 3 is indicative of variations in composition and heat treatment that collectively form the 20-dimensional predictor (or processing) space. The t-SNE cluster analysis (Fig. 1) has shown limited variation in this high-dimensional space, which indicates that there are potentially multiple dependencies that exist between predictors. The cluster analysis, without *a priori* knowledge, indicates predictor dependencies that are not surprising to a materials engineer. Since the composition governs the stable thermodynamic ranges of different phases, it put constraints on subsequent selection of heat treatment temperatures. In other words, this suggests that tempering and normalization temperatures might be dependent on composition. Similarly, balancing of austenite and ferrite stabilizers for fully martensitic structure puts constraints on individual elements [51]. Therefore, lasso regression was applied before contributor selection to identify dependent predictor pairs, and pick one of the dependent predictor and discard the other. The contributor selection was done on the remaining independent predictors, providing an open design space where heat treatment and composition can be explored together for better properties.

3.4.1. Heat treatment

The uni-response models of tempering temperature (TPR) and normalization temperature (NML), presented in Table 2, were the result of lasso regression, followed by subset selection on the composition predictor set. The TPR model has a low adj- R^2 value (0.71), which suggests that TPR selection (design) is not strongly correlated with composition variables. The evaluation of Cook's distance and visualization of residuals [52] for the model highlighted one observation from 9Cr-2Mo (⊞) sub-class as an outlier with high tempering temperature. The NML model's high adj- R^2 value (0.92) highlights that NML is strongly

correlated with composition and therefore NML might be a dependent design variable. The NML model showed three outliers from 9Cr-1.5Mo-VNb-B (⊠) sub-class, which contain 12 observations, where 9 observations have the same composition while NML is different for each observation. All three outliers belong to these 9 observations, where one outlier has the highest value of the range and the other two have the two lowest values. The outliers highlight that a composition and its resulting NML temperature might have been chosen independently for a previous study, although high adj- R^2 values suggests otherwise for rest of the sub-classes in the dataset.

3.4.2. Tensile test responses

To identify the contributors governing the tensile behavior, lasso regression followed by subset selection was applied. Subsequently, the evaluation of Cook's distance and visualization of residuals for the different models highlighted 5 outliers: 3 from (Cr) (Mo)0.5W(Co) VNbB(Ta) (•), and 2 from 9Cr-1.5Mo-1.3Co-VNb-B (□) sub-class. All 5 of them exhibit relatively high strength compared to the rest of the data with moderate ductility. The RA high temperature model highlighted 7 observations from 12Cr-1Mo-1W-Co-V-Cu (▽), which were misclassified as a result of using the UTS versus temperature curve for separation of data. The five outliers were removed, the mislabeled observations were reclassified as low temperature responses, and contributors were re-evaluated for both temperature regions (tabulated in Table 3). Of all the models, El model for both regions shows a low adj- R^2 value.

3.5. MMR: Constrained contributor selection

The tensile data, after removing outliers in previous sections (i.e., section 3.3: 9Cr-1Mo (◇) sub-class, and section 3.4.2: reclassified 12Cr-1Mo-1W-Co-V-Cu (▽) observations from high temperature to low temperature), were used for MMR constrained contributor selection (tabulated in Table 4). The constrained selection process resulted in few

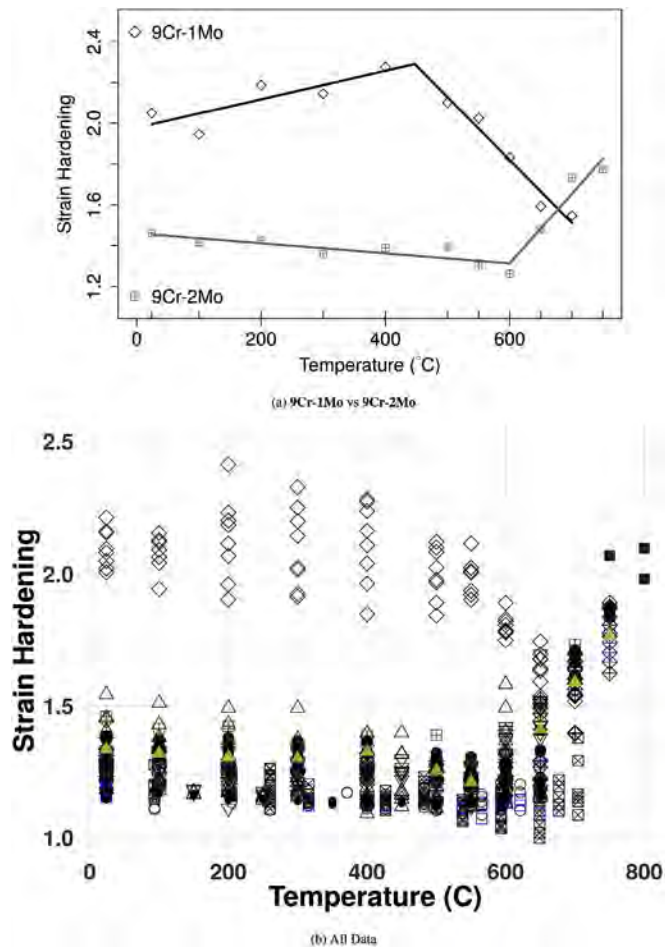


Fig. 4. (a) Strain hardening as a function of test temperature for a single alloy composition in 9Cr-1Mo sub-class with 9Cr-2Mo sub-class. (b) Strain hardening for different sub-classes as a function of temperature. Different shapes represent sub-classes as shown in Fig. 1 and Table 1. The figure highlights that 9Cr-1Mo sub-class is inherently different than the rest of the data.

Table 2

Rank-ordered contributors for tempering (TPR) and normalization (NML) temperature; with # representing the number of observations.

	#				adj-R ²
TPR	76	N	- Ni		0.71
NML	73	B	V	- P	0.92

Table 3

Rank-ordered contributors for tensile test outcomes for low (428 observations) and high (326 observations) temperature regions as determined by uni-response lasso regression coupled with subset selection. Temp, TPR, and NML represent test temperature, tempering temperature, and normalization temperature, respectively.

Low								adj-R ²
YS	- TPR	Ni	- Temp	V	Ta	Co	- Mn	0.93
UTS	- TPR		- Temp	V	Ta	Co	Ni	0.93
RA	- Ni	- B		- V	- P	- Co	TPR	0.82
El	- Ni	- Temp	TPR	- V	NML			0.61
High								adj-R ²
YS	- Temp	NML	- TPR	- Si				0.92
UTS	- Temp	NML	- TPR	Ni				0.91
RA	Temp	- B		- Ni	- V			0.79
El	Temp	- V	TPR	- N	Mn			0.69

Table 4

Rank-ordered contributors for tensile test outcomes for low (428 observations) and high (326 observations) temperature regions using multivariate multiple response model. “+” and “-” sign indicates that the coefficient is positive and negative, respectively. Temp, TPR, and NML represent test temperature, tempering temperature, and normalization temperature, respectively.

Low	Temp	C	P	TPR	Ni	B	Ta	V	Co	NML	adj-R ²
YS	-		-	-	+	+	+	+	+		0.93
UTS	-	+		-	+	+	+	+	+		0.94
RA	+		-	+	-	-		-	-		0.83
El	-		+	-	-					+	0.61
High	Temp	B	Ni	Mn	V	TPR	P				adj-R ²
YS	-	+	+		+	-	-				0.93
UTS	-	+	+	+	+	-	-				0.93
RA	+	-	-		-						0.79
El	+	+		+	-	+	+				0.67

differences in comparison with lasso corrected contributors (Table 3), where most contributors overlapped for each respective model and similar adj-R² values were observed. For high temperature region, the differences in the MMR constrained model highlight the NML dependency on B and V (Table 2).

3.6. Validation

The validation of linear models presented in Table 3 was done by partitioning the data into train and test dataset and looking at the ratio of errors. Table 5 shows that the ratio of test to training error is close to one for all models, except for El models where they have low adj- values.

4. Discussion

The motivation behind this study is to provide a new composition with an appropriate heat treatment route for the next generation of the class 9–12 wt% Cr martensitic steels for turbine rotors by addressing the strength-ductility trade-off. This can be translated to 1) identifying contributors which govern strength and ductility, 2) identifying contributors which govern strength but not ductility, or vice-versa, and 3) identifying whether these contributors should be decreased or increased (qualitative estimates) and by how much (quantitative estimates). In the results section, first two points were presented, the discussion section will focus on last point and move towards quantitative measures of individual elements while evaluating the credibility of results against domain knowledge and statistical measures.

4.1. Data limitations

The limitations of the data for statistical analysis must be discussed prior to using the models for quantitative design choices. Most of the

Table 5

Ratio for test to training error for different models, computed over 1000 different random allocations of test and training datasets. # Test/Training represents the number of observations in test/training dataset. Temp represents the low and high temperature regions.

Response	Temp	# Training	# Test	adj-R ²	ratio
YS				0.93	1.04 ± 0.19
UTS	Low	342	86	0.93	1.04 ± 0.21
RA				0.82	1.01 ± 0.18
El				0.61	0.91 ± 0.13
YS	High			0.92	1.04 ± 0.22
UTS	High	260	66	0.91	1.03 ± 0.19
RA				0.79	1.00 ± 0.18
El				0.69	0.96 ± 0.17

historical design choices, which dictate the scope of the data, were made in an attempt to maximize the creep lifetime of martensitic steels, while other properties such as toughness and ductility were secondary. This suggests that the data may be inadequate for data-driven modeling of the tensile strength-ductility trade-off as the data negative space might be hiding valuable information. Also, the historical data from multiple independent sources does not strictly follow any statistical assumptions (e.g., homoscedasticity). For example, the fact that all observations with **Ta** went through homogenization treatment [53], presents a clear case of data convolution. Similar cases can be found across the available data, however identifying the bias in the data allows for interpretation of the models resulting in design insights. For example, understanding the deformation mechanisms and bias and in the existing space might lead to reasons to run physics-based microstructure studies in the negative space to remove bias and test additional hypotheses. The next few sections discuss what the data-driven models suggest about the mechanisms and responses.

4.2. Measure of ductility

The ability of a material to accommodate plastic deformation without breaking can be expressed through two measures: percent reduction in area (% RA) and percent elongation to failure (% El). The % RA is the ratio of the change in the sample area at the fracture point to the initial sample area, and is therefore a measure of the highest plastic strain. In contrast, % El is calculated as the ratio of the change in gauge length to the initial gauge length. Even though test standards specify the gauge length as a multiple of the test section diameter or width, the initial gauge length is dictated by the arbitrary dimensions of the extensometer used to measure the change in sample length. For example, American Society for Testing and Materials (ASTM) specifies a ratio of gauge length to diameter (L_i/d_i) equal to four for specimens with round cross sections. Although, a ratio of L_i/d_i equal to five has been employed internationally [54]. Since the data under study comes from multiple countries and laboratories, with no information about gauge length, the comparison of El data is restricted. Low adj. value for El models (Table 3, Table 4) may be the result of this mismatch. The RA is not affected by the arbitrary selection of L_i/d_i , and is therefore a more fundamental measure of ductility. Conceptually, this has implications for the community of researchers developing databases for future studies. As a community, we need to update the mechanical testing standards to comply with the needs of the database developers. This could be requiring the reporting of % RA, or the reporting of the necessary metadata about the chosen initial gauge length as dictated by the equipment. For the rest of the paper, only the % RA models are considered as a stable predictor of the ductility for design insights.

4.3. Strength-ductility trade-off

The strength-ductility trade-off has been captured in Tables 3 and 4 distinctly. The comparison of ranked contributors between YS and RA for the low temperature region highlights contributors such as **Ni**, **V**, and **Co**, where each of the contributors exemplifies the trade-off between strength and ductility. Similar comparison for the high temperature region highlights contributors such as **Ni**, **V**, and **B**. The comparison also highlights contributors which do not follow the commonly observed trade-off between strength and ductility. In the low temperature region, addition of **Ta** only increases strength but does not contribute to ductility. Similarly, in the high temperature region, reducing TPR independently increases strength without changing ductility. While there are few contributors which alter strength without affecting ductility, there are no contributors which alters ductility without affecting strength. Further examination of these contributors against domain knowledge is needed for design insights. The differences between Tables 3 and 4 also highlight the benefit of lasso corrected contributors over MMR constrained selection. For example, in

the low temperature region, TPR is the first ranked contributor for YS, and it is the last for RA; while this information is missing from MMR constrained selection.

4.4. Tensile strengthening mechanism

In general, metallic alloys derive their strength from the interaction of dislocations with obstacles such as other dislocations, boundaries, solutes, and particles. Therefore, strength can be defined as the cumulative sum of individual obstacle strengthening mechanisms [55]. Considering the wide range of compositions within the study, the same element could contribute to different obstacle mechanisms for different sub-classes and test environments. For example, **Mo** contributes to strength as a solute obstacle at low temperatures, but precipitates out at higher temperatures and contribute as a particle obstacle [34]. Therefore, the ranked contributors in Tables 3 and 4 represent an average behavior of each element. The following sections discuss ranked contributors using different obstacle mechanisms listed above.

4.4.1. Contribution of grain boundaries

The tempered martensitic microstructure of the class 9–12 wt% **Cr** steels is a hierarchical structure with martensitic laths at smallest scale, followed by packet boundaries, and prior-austenite grain boundaries (PAGBs). In general, packets within a prior austenitic grain are considered a single grain and a linear dependency exists between PAGBs size and packet size ($\sim 1/3$ of PAGB size) [56]. However, multiple methods (explored in Ref. [12]) exist to further refine the packets by breaking up the alignment of laths resulting in different blocks of laths with similar orientation. Morito et al. [56], established that these blocks exhibit a Hall-Petch type effect on the YS at room temperature. For example, high NML increases PAGBs size [57], although Fig. 5a indicates no such dependence, indicating that YS might be the function of block size, not PAGB size within the studied dataset. Further, both Fig. 5a and b show no dependence between NML and YS/RA, indicating why NML is not a strong contributor in Tables 3 and 4. Although, NML do show up for the YS model in Table 3 due to its dependency with **B** and **V** (Table 2).

The martensitic transformation results in the formation of sub-boundaries and an inherent dislocation density, and the transformation is strongly dependent on the amount of **C** and **N** in the alloys. The amount of **C** and **N** is strongly tied to packet size, where packet size decreases with increase in **N** and **C** content [58]. Nitrogen has a greater hardening effect on austenite than **C** which suppresses microstructure growth in high-**N** alloys resulting in low dislocation density [58]. Table 2 captures this dependency, where TPR and **N** show a strong positive correlation, suggesting that low dislocation density due to **N** prior to tempering may correlate with a need for an increased tempering temperature as recovery is slower for microstructures with lower dislocation density (i.e., less driving force). An increase in tempering temperature results in an increase in lath width [57] and toughness [59] at the expense of strength (Fig. 5c) (Table 3/Table 4).

4.4.2. Solute strengthening

Solute atoms generate a local stress field due to a mismatch in atomic size and shear modulus. In 9–12 wt% **Cr** steels, the largest stress fields are created by the **Mo** and **W** solute atoms, although other elements such as **Cr**, **Mn**, & **Ni** (high lattice mismatch), and **Si** (modulus mismatch) provide additional solute strengthening [55]. Contrary to the literature, **Mo** and **W** do not show up in Table 3 or Table 4. This is explained by Fig. 6a where 12-**Cr** (\triangle) sub-class and 12**Cr**-**Mo**-1.8**W**-**VNb**-**B**-**Cu** (\blacktriangle) sub-class have equal YS at room temperature, indicating that other variables such as low NML and low TPR compensate for low alloying additions for 12-**Cr** (\triangle) sub-class. A similar correlation is observed, but not shown, with **Cr** content where low TPR compensate for the absence of alloying additions, explaining the absence of **Cr** from Tables 3 and 4. This reflects that the design choices

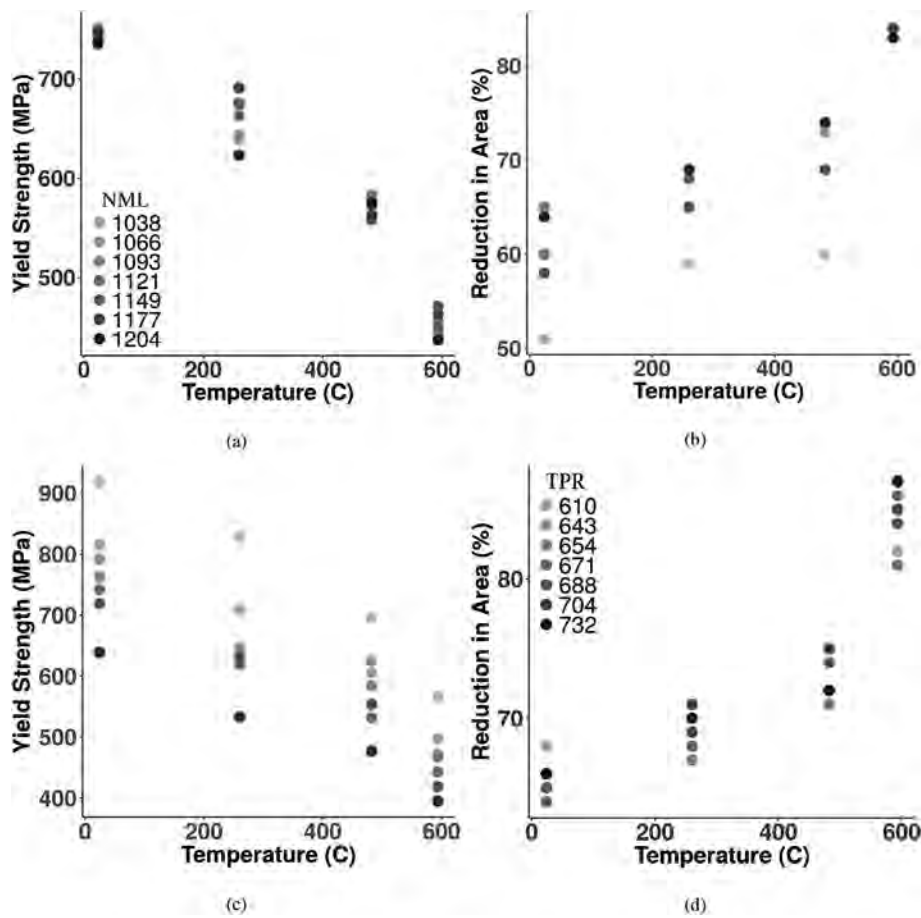


Fig. 5. (a)–(b) Effect of normalization temperature (NML) on yield strength (YS) and reduction in area (RA) for 9Cr-1.5Mo-V-Nb-B (⊠) sub-class. Figure highlights why NML is not a strong contributor to YS and RA. Legends indicate different NML values in °C. (c)–(d) Effect of tempering temperature (TPR) on YS and RA for 9Cr-1.5Mo-V-Nb-B (⊠) sub-class. Figure highlights that increase in TPR reduces YS consistently at all test temperatures, while it has no consistent effect on RA. Legends indicate different TPR values in °C.

were made in an attempt to maximize the creep lifetime, and not to evade the strength-ductility trade-off.

Other solutes such as Ni, Mn and Si do show up in Table 3 or Table 4, although Mn and Si are not strong contributors and their coefficients are negative. The negative coefficient for Mn in Table 3 can be attributed to its dependency on Ni (correlation coefficient of 0.65) and a similar distribution of data as of Mo, W, and Cr with YS. On the other hand, Si's negative coefficient is associated with its dependency on low alloying additions and high impurities (Fig. 6b). Silicon was added to decrease the oxidation rate for 9-Cr steels above 600 °C, where Cr content is insufficient for oxidation resistance [60,61], and also as an impurity getter [62].

4.4.3. Precipitate strengthening

The strengthening via precipitates is due to pinning of dislocations (Orowan strengthening), and it depends on mean inter-particle spacing and particle radius. Small atoms such as B, C and N predominantly form carbides and carbonitrides, instead of remaining in solution as interstitials. Carbon affects microstructural properties in Fe-C steels, where dislocation density increases with increasing C content [63] due to the decreasing block and lath width [58]. Boron reduces carbide spacing via formation of a finer distribution of more thermally stable carboboride [64,65], and effectively works to strengthen precipitate-matrix interfaces (Table 4). Although, high temperature normalization (Table 2) is needed to dissolve borides and to promote the segregation of B to the grain boundaries [66] for the enrichment of carbides. Nitrogen is important, as nitrides are more stable than carbides and borides [67], although excessively strong nitride-forming elements, e.g., Al [68] and B [5] can result in and, respectively, at the expense of ($M = V, Nb, Ta$) carbonitrides. The amount of C and N determines the amount of precipitates, thereby the amount of carbide formers in solid

solution. The dependencies between C/N and carbide forming elements eliminate C/N from Tables 3 and 4, although they determine various aspects of final microstructure.

Both Tables 3 and 4 capture V as a ranked contributor, although Nb is absent. This is due to the lack of variation for Nb in all observations (0.05–0.07 wt%), except for 9Cr-1Mo-VNb (⊕) sub-class (Nb > 0.07 wt%). Tantalum looks promising to increase strength without affecting ductility in the low temperature region (Table 3/Table 4). Although, a simple scatter plot (Fig. 7) shows that there is no significant change in YS or RA at room temperature with Ta concentration, which illustrates that Ta-rich carbides do not affect strength or ductility (same is observed at 650 °C). Only ~15% of the data contains compositions with Ta, therefore the association between Ta and YS (in Table 3/Table 4) may be the result of the presence and absence of Ta. Klueh et al. [69], concluded that Ta plays an important role in refining the PAGBs because most of the Ta remains in solid solution after heat treatment, with very small amount of Ta-rich carbides when Ta is 0.07 wt%. Recent studies [70,71] have extended this range to 0.078–0.28 wt% Ta. These results highlight that the PAGBs size does not change with additions beyond 0.15 wt% Ta and most of the Ta precipitates, with increase in particle size with amount of Ta. Further, no significant difference in mechanical properties were found with precipitation of Ta-C carbides at the expense of other particles [70,71]. This literature corroborate with our interpretation that the contribution of Ta to the YS is due to the presence/absence of Ta within the dataset.

In the low temperature region, Co shows up as a moderate contributor (Table 3/Table 4), although literature suggests that Co does not increase YS by solute strengthening as its atomic size and shear modulus are similar to those of Fe atoms [55]. This suggests that Co could be contributing as a precipitate strengthener [72], as precipitates govern strength at low temperatures.

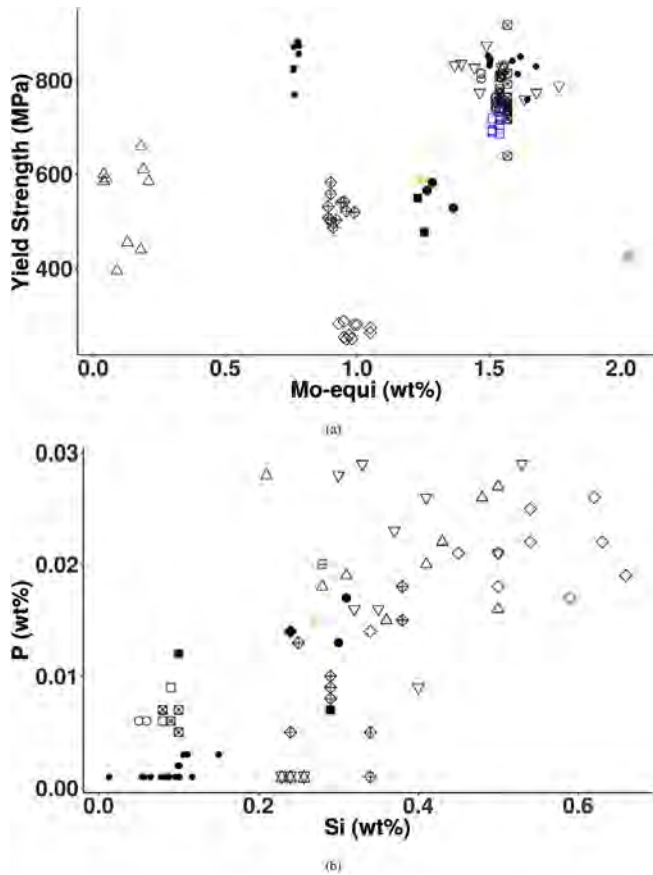


Fig. 6. (a) Variation in yield strength (YS) with Mo-equivalent (Mo + 0.5*W) for room temperature data (23.9 °C). Figure shows the distribution of data with respect to Mo and W, highlighting why Mo/W is not a ranked contributor in Table 3/Table 4. (b) Variation in P with Si. Figure shows that Si increases with the amount of impurity P, leading to a negative coefficient in Table 3. Different shapes represent different sub-classes from Fig. 1/Table 1.

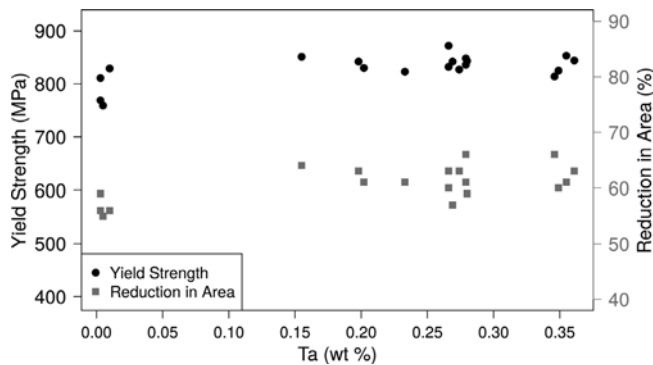


Fig. 7. Variation in Yield Strength (YS) and Reduction in Area (RA) with Ta for observations where Ta is present for room temperature data. Figure highlights that there is no significant change in YS or RA with Ta content.

4.4.4. Summary of tensile strengthening mechanism

The assessment of the models against domain knowledge under subsection 4.4 provided insights about the potential mechanisms each contributor present in Table 3/Table 4 imposes on the system, and highlighted bias that was not captured. In summary, low tempering temperature, high dislocation density, solute atoms (Ni) and fine precipitates (governed by V & B) are desirable for high strength, although solute atoms and precipitates reduces ductility. Low tempering temperature does not effect ductility, therefore, it can be exploited for increasing strength while maintaining ductility. To assess the effect of

Mo, W, Cr, Nb and Mn, a new design space needs to be explored; though literature indicates that these elements increase strength through multiple mechanisms. The dependency of normalization temperature on B, V, and P (Table 2), along with Fig. 5, highlights the similarities between Tables 3 and 4, and shows how B and V determine the final properties. The addition of Ta showed no change in YS or RA. The amount of C and N govern dislocation density and sub-grain structure, therefore, special attention is needed to assess their optimized amount.

4.5. Impurities

Phosphorus (P) and sulfur (S) tend to segregate to regions of high lattice misfit, or to PAGBs in martensitic steels [73], causing impurity-induced embrittlement. With low test temperature, as the matrix becomes stronger, a tendency towards intergranular failure increases with P concentration [74], and a PAGB grain size limits the segregation over a smaller boundary area per unit volume [75], resulting in a greater concentration of impurities at PAGBs. The traditional way to avoid this embrittlement involves reducing the impurity concentration and adding substitutional solutes (i.e., Si, Al) which pull the impurities from the matrix [62]. The former can be observed on average within the data studied (Table 3, Table 4), where the effect is stronger in the low temperature region. The YS on average is significantly higher when P content is ≤ 0.003 wt%, in comparison to the rest of the data where P is ≥ 0.005 wt% (mean value of 0.007 wt%).

4.6. New design space

Collectively, the insights and bias highlighted in sections 4.3, 4.4, and 4.5, along with domain study, highlight that a new design space is needed for the quantitative estimates of alloying elements for the application of turbine rotors. The difference between the correlations that are captured in Table 3/Table 4, and those that are suggested by domain study, reflects the presence of high-dimensional dependencies that exist among predictors/contributors and further supports the need for new design space. The historical design space was tailored for maximizing the creep rupture life, therefore, analysis captures the contributors which maximize creep rupture life [34] for YS/RA too (i.e., V, B, Co), reflecting the problem of multi-objective optimization [76].

An important heuristic that can be derived from the analysis is that dislocation density, a function of martensitic transformation and subsequent tempering treatment, determines the final strength of the material at high temperatures. Within the studied dataset, the effect of individual elements (e.g., Mo, W, Cr, Mn) on martensitic transformation, ergo on dislocation density, was masked by the often contradictory effect of tempering treatment on the dislocation density. Therefore to deconvolute the design space, the effect of tempering treatment and the effect of alloying additions on dislocation density needs to be separated. This could be done by designing the input space where tempering temperature and alloying additions are varied orthogonally/independent of each other.

Two compositions that are proposed in the literature for high toughness for an operating temperature of 630 °C are as follows: MTR10A (10Cr-0.7Mo-1.8W-3Co-VNb-B) and HR1200 (11Cr-2.6W-3Co-Ni-VNb-B) [2]. Both compositions have V, Nb, and B, which agree with the results obtained through this study, although the role of Co, Mo, and W, are not captured within the studied dataset. Once the effect of alloying additions and tempering treatment is separated, the same compositional space should be explored for new compositions, specially the effect of W, Mo, and Co within the (0–3) wt% range. The addition of Ni with the high Cr content supports the design objective of balancing austenite and ferrite stabilizers, while the absence of Ta from both compositions supports the results shown in Fig. 7.

5. Conclusions

The data driven approach successfully identified the major contributors towards tensile properties, and highlighted the sparse nature of the design space considered in the data. The visualization using t-SNE showed that the 91 unique combinations of composition and heat treatment do not contain sufficient variability among all 20 predictors resulting in several observed dependencies. The bias invoked in the alloy design space to maximize creep rupture life supervised the subset selection process and provided major contributors for YS and RA. The comparison of Lasso corrected contributors and MMR contributors highlighted the dependency between predictors, and provided a more flexible design space. The visualization of various attributes further elucidated the bias and showed dependencies among elements. The study highlights the limitations of the design space considered in the past and suggests that the deconvolution of the effect of tempering temperature and alloying additions on dislocation density is needed. This could be done by designing the input space where tempering temperature and alloying additions (especially W, Mo, and Co) are varied orthogonally/independent of each other.

Disclaimer

This project was funded by the Department of Energy (grant DE-FE0028685), National Energy Technology Laboratory (DOE-NETL), an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Research data

Part of the data used for the analysis is the property of DOE-NETL, an agency of the United States Government, and is confidential. The rest of the data comes from NIMS, Japan, and can be accessed directly from https://mits.nims.go.jp/index_en.html. Specific details of spreadsheets from NIMS are included in references (20–28).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.msea.2019.138142>.

References

- [1] J. Shingledecker, R. Purgert, P. Rawls, Current status of the US DOE/OCDO A-USC materials technology research and development program (2013. 10), Proceeding the 7th International Conference on Advances in Materials Technology for Fossil Power Plants, 2013, pp. 41–52 URL https://www.asinternational.org/search/-/journal_content/56/10192/05440G/PUBLICATION.
- [2] F. Abe, Progress in creep-resistant steels for high efficiency coal-fired power plants, 040804–040804–21, J. Press. Vessel Technol. 138 (4) (2016), <https://doi.org/10.1115/1.4032372> URL <https://doi.org/10.1115/1.4032372>.
- [3] R. Viswanathan, J. Nutting, Advanced Heat Resistant Steels for Power Generation: Conference Proceedings 27–29 April 1998 San Sebastian, IOM Communications, Spain, 1999 (google-Books-ID: xeNIQgAACAAJ).
- [4] K. Maruyama, K. Sawada, J.-i. Koike, Strengthening mechanisms of creep resistant tempered martensitic steel, ISIJ Int. 41 (6) (2001) 641–653, <https://doi.org/10.2355/isijinternational.41.641>.
- [5] F. Abe, Precipitate design for creep strengthening of 9% Cr tempered martensitic steel for ultra-supercritical power plants, Sci. Technol. Adv. Mater. 9 (1) (2008) 013002, <https://doi.org/10.1088/1468-6996/9/1/013002> URL <http://stacks.iop.org/1468-6996/9/i=1/a=013002>.
- [6] L. Wang, M. Li, J. Almer, In situ characterization of Grade 92 steel during tensile deformation using concurrent high energy X-ray diffraction and small angle X-ray scattering, J. Nucl. Mater. 440 (1) (2013) 81–90, <https://doi.org/10.1016/j.jnucmat.2013.04.063> URL <http://www.sciencedirect.com/science/article/pii/S0022311513006818>.
- [7] F. Abe, R. Viswanathan, T.-U. Kern (Eds.), Creep Resistant Steels, Woodhead Publishing, Cambridge, UK, 2008.
- [8] L. Wang, M. Li, J. Almer, Investigation of deformation and microstructural evolution in Grade 91 ferritic-martensitic steel by in situ high-energy X-rays, Acta Mater. 62 (2014) 239–249, <https://doi.org/10.1016/j.actamat.2013.10.003> URL <http://www.sciencedirect.com/science/article/pii/S1359645413007519>.
- [9] Chapter 2, Fundamentals of the heat treating of steel, ASM Handbook, Volume 4A: Steel Heat Treating Fundamentals and Processes - ASM International, 2006 URL https://www.asinternational.org/search/-/journal_content/56/10192/05344G/PUBLICATION.
- [10] B.A. Senior, F.W. Noble, B.L. Eyre, The nucleation and growth of voids at carbides in 9 Cr-1 Mo steel, Acta Metall. 34 (7) (1986) 1321–1327, [https://doi.org/10.1016/0001-6160\(86\)90019-2](https://doi.org/10.1016/0001-6160(86)90019-2) URL <http://www.sciencedirect.com/science/article/pii/0001616086900192>.
- [11] M. Li, L. Wang, J.D. Almer, Dislocation evolution during tensile deformation in ferritic-martensitic steels revealed by high-energy X-rays, Acta Mater. 76 (2014) 381–393, <https://doi.org/10.1016/j.actamat.2014.05.026> URL <http://www.sciencedirect.com/science/article/pii/S1359645414003747>.
- [12] J.J.W. Morris, Z. Guo, C.R. Krenn, Y.-H. Kim, The limits of strength and toughness in steel, ISIJ Int. 41 (6) (2001) 599–611, <https://doi.org/10.2355/isijinternational.41.599> URL https://www.jstage.jst.go.jp/article/isijinternational1989/41/6/41_6_599/article-char/en.
- [13] Z.Q. Fan, T. Hao, S.X. Zhao, G.N. Luo, C.S. Liu, Q.F. Fang, The microstructure and mechanical properties of T91 steel processed by ECAP at room temperature, J. Nucl. Mater. 434 (1) (2013) 417–421, <https://doi.org/10.1016/j.jnucmat.2012.12.009> URL <http://www.sciencedirect.com/science/article/pii/S0022311512006666>.
- [14] M. Song, C. Sun, Z. Fan, Y. Chen, R. Zhu, K.Y. Yu, K.T. Hartwig, H. Wang, X. Zhang, A roadmap for tailoring the strength and ductility of ferritic/martensitic T91 steel via thermo-mechanical treatment, Acta Mater. 112 (2016) 361–377, <https://doi.org/10.1016/j.actamat.2016.04.031> URL <http://www.sciencedirect.com/science/article/pii/S1359645416302889>.
- [15] T. Hao, H. Tang, G. Luo, X. Wang, C. Liu, Q. Fang, Enhancement effect of inter-pass annealing during equal channel angular pressing on grain refinement and ductility of 9Cr1Mo steel, Mater. Sci. Eng. A 667 (2016) 454–458, <https://doi.org/10.1016/j.msea.2016.04.098> URL <http://www.sciencedirect.com/science/article/pii/S0921509316305056>.
- [16] I. Miyazaki, T. Furuta, K. Oh-ishi, T. Nakagaki, S. Kuramoto, A. Shibata, N. Tsuji, Overcoming the strength–ductility trade-off via the formation of a thermally stable and plastically unstable austenitic phase in cold-worked steel, Mater. Sci. Eng. A 721 (2018) 74–80, <https://doi.org/10.1016/j.msea.2018.02.075> URL <http://www.sciencedirect.com/science/article/pii/S0921509318302934>.
- [17] L. Sun, T.H. Simm, T.L. Martin, S. McAdam, D.R. Galvin, K.M. Perkins, P.A.J. Bagot, M.P. Moody, S.W. Ooi, P. Hill, M.J. Rawson, H.K.D.H. Bhadeshia, A novel ultra-high strength maraging steel with balanced ductility and creep resistance achieved by nanoscale beta-NiAl and Laves phase precipitates, Acta Mater. 149 (2018) 285–301, <https://doi.org/10.1016/j.actamat.2018.02.044> URL <http://www.sciencedirect.com/science/article/pii/S1359645418301526>.
- [18] B.B. He, B. Hu, H.W. Yen, G.J. Cheng, Z.K. Wang, H.W. Luo, M.X. Huang, High dislocation density-induced large ductility in deformed and partitioned steels, Science 357 (6355) (2017) 1029–1032, <https://doi.org/10.1126/science.aan0177> URL <http://science.sciencemag.org/content/357/6355/1029>.
- [19] Z. Li, K.G. Pradeep, Y. Deng, D. Raabe, C.C. Tasan, Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off, Nature 534 (7606) (2016) 227–230, <https://doi.org/10.1038/nature17981> URL <https://www.nature.com/articles/nature17981>.
- [20] K. Yadi, NIMS creep data sheet (JIS SUS 403-B, 12cr, bar), Tech. Rep. (1994) 13B, NIMS, Tokyo.
- [21] H. Irie, NIMS creep data sheet (JIS STBA 26, 9cr-1mo, tube), Tech. Rep. (1997) 19B, NIMS, Tsukuba.
- [22] H. Irie, NIMS creep data sheet (1cr-0.5mo-0.25v and 12cr-1mo-1w-0.25v, bolting material), Tech. Rep. 44 (1997) NIMS, Tsukuba.
- [23] H. Irie, NIMS creep data sheet (JIS SUH 616-B, 12cr-1mo-1w-0.3v, bar), Tech. Rep. (1998) 10B, NIMS, Tsukuba.
- [24] S. Matsuoka, NIMS creep data sheet (KA-STBA 27, 9cr-2mo, tube), Tech. Rep. (2005) 46A, NIMS, Tsukuba.
- [25] NIMS creep data sheet, Tech. Rep. (2012) 48A, NIMS, Japan.
- [26] NIMS creep data sheet (KA-SUS 410 J3, 12cr-2w-0.4mo-1cu-Nb-V), Tech. Rep. (2013) 51A, NIMS, Japan.
- [27] NIMS Creep Data Sheet (KA-SUS 410j3 DTB, 12cr-2w-0.4mo-1cu-Nb-V (Tube)), Tech. Rep. 52A, NIMS, Japan (2013).
- [28] NIMS creep data sheet (9cr-1mo-V-Nb), Tech. Rep. (2014) 43A, NIMS, Japan.
- [29] L. v. d. Maaten, G. Hinton, Visualizing Data using t-SNE, J. Mach. Learn. Res. 9 (Nov) (2008) 2579–2605 URL <http://www.jmlr.org/papers/v9/vandermaten08a.html>.
- [30] F. Masuyama, History of power plants and progress in heat resistant steels, ISIJ Int. 41 (6) (2001) 612–625, <https://doi.org/10.2355/isijinternational.41.612> URL https://www.jstage.jst.go.jp/article/isijinternational1989/41/6/41_6_612/article.
- [31] L. v. d. Maaten, Accelerating t-SNE using tree-based algorithms, J. Mach. Learn. Res. 15 (2014) 3221–3245 URL <http://jmlr.org/papers/v15/vandermaten14a>.

- html.
- [32] R. R Core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017 URL <https://www.R-project.org/>.
 - [33] J.H. Krijthe, Rtsne: T-Distributed Stochastic Neighbor Embedding Using Barnes-Hut Implementation, R Package Version 0.13, (2015) URL <https://github.com/jkrijthe/Rtsne>.
 - [34] A. K. Verma, J. A. Hawk, L. S. Bruckman, R. H. French, V. Romanov, J. L. W. Carter, Mapping Multivariate Influence of Alloying Elements on Creep Behavior for Design of New Martensitic Steels, *Metall and Mat Trans* Adoi:10.1007/s11661-019-05234-9. URL <https://doi.org/10.1007/s11661-019-05234-9>.
 - [35] V.M. Muggeo, segmented: an r package to fit regression models with broken-line relationships, *R. News* 8 (1) (2008) 20–25 URL <https://cran.r-project.org/doc/Rnews/>.
 - [36] V.M.R. Muggeo, Estimating regression models with unknown break-points, *Stat. Med.* 22 (19) (2003) 3055–3071, <https://doi.org/10.1002/sim.1545> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.1545>.
 - [37] T. L. Based on Fortran Code by Alan Miller, Leaps: Regression Subset Selection, R Package Version 3.0, (2017) URL <https://CRAN.R-project.org/package=leaps>.
 - [38] P.M. Narendra, K. Fukunaga, A branch and bound algorithm for feature subset selection, *IEEE Trans. Comput.* 9 (C-26) (1977) 917–922, <https://doi.org/10.1109/TC.1977.1674939> URL <https://www.infona.pl/resource/bwmeta1.element.ieee-art-000001674939>.
 - [39] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, second ed., second ed., Springer Series in Statistics, Springer-Verlag, New York, 2009 URL www.springer.com/us/book/9780387848570.
 - [40] A.E. Hoerl, R.W. Kennard, ridge regression: biased estimation for nonorthogonal problems, *Technometrics* 12 (1) (1970) 55, <https://doi.org/10.2307/1267351> URL <https://www.jstor.org/stable/1267351?origin=crossref>.
 - [41] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. Ser. B* 58 (1) (1996) 267–288 URL <http://www.jstor.org/stable/2346178>.
 - [42] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, *J. Stat. Softw.* 33 (1) (2010) 1–22 URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929880/>.
 - [43] N. Simon, J. Friedman, T. Hastie, R. Tibshirani, Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent, *Journal of Statistical Software* 39 (5). doi:10.18637/jss.v039.i05. URL <http://www.jstatsoft.org/v39/i05/>.
 - [44] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani, Pathwise coordinate optimization, *Ann. Appl. Stat.* 1 (2) (2007) 302–332, <https://doi.org/10.1214/07-AOAS131> URL <https://projecteuclid.org/euclid.aos/1196438020>.
 - [45] R.A. Johnson, D.W. Wichern, *Applied Multivariate Statistical Analysis*, sixth ed., Pearson, Upper Saddle River, N.J., 2007.
 - [46] L. Breiman, J.H. Friedman, Predicting multivariate responses in multiple linear regression, *J. R. Stat. Soc. Ser. B* 59 (1) (1997) 3–54, <https://doi.org/10.1111/1467-9868.00054> URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00054>.
 - [47] R. Cheng, Qtlmt: Tools for Mapping Multiple Complex Traits, R Package Version 0.1-6, (2017) URL <https://CRAN.R-project.org/package=qtlmt>.
 - [48] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control* 19 (6) (1974) 716–723, <https://doi.org/10.1109/TAC.1974.1100705>.
 - [49] R.D. Cook, Detection of influential observation in linear regression, *Technometrics* 19 (1) (1977) 15–18, <https://doi.org/10.2307/1268249> URL <http://www.jstor.org/stable/1268249>.
 - [50] N.E. Dowling, *Mechanical Behavior of Materials*, fourth ed., Pearson, Boston, 2012.
 - [51] S.H. Ryu, J. Yu, A new equation for the Cr equivalent in 9 to 12 pct Cr steels, *MMTA* (6) (1998) 1573–1578 URL <https://doi.org/10.1007/s11661-998-0080-7> <https://link.springer.com/article/10.1007/s11661-998-0080-7>.
 - [52] D.C. Montgomery, E.A. Peck, G.G. Vining, *Introduction to Linear Regression Analysis*, John Wiley & Sons, google-Books-ID, 2012 0yR4KUL4VDkC.
 - [53] P.D. Jablonski, J.A. Hawk, Homogenizing advanced alloys: thermodynamic and kinetic simulations followed by experimental results, *J. Mater. Eng. Perform.* 26 (1) (2017) 4–13, <https://doi.org/10.1007/s11665-016-2451-3> URL <https://doi.org/10.1007/s11665-016-2451-3>.
 - [54] N.E. Dowling, Chapter 4: mechanical testing: tension test and other basic tests, *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue*, fourth ed., Pearson, Boston, 2013oCLC: ocn730404288.
 - [55] R.A. Barrett, P.E. O'Donoghue, S.B. Leen, A physically-based high temperature yield strength model for 9cr steels, *Mater. Sci. Eng. A* 730 (2018) 410–424, <https://doi.org/10.1016/j.msea.2018.05.086> URL <http://www.sciencedirect.com/science/article/pii/S092150931830755X>.
 - [56] S. Morito, H. Yoshida, T. Maki, X. Huang, Effect of block size on the strength of lath martensite in low carbon steels, *Mater. Sci. Eng. A* 438–440 (2006) 237–240, <https://doi.org/10.1016/j.msea.2005.12.048> URL <http://www.sciencedirect.com/science/article/pii/S0921509306006058>.
 - [57] P.J. Ennis, A. Zielinska-Lipiec, O. Wächter, A. Czyska-Filemonowicz, Microstructural stability and creep rupture strength of the martensitic steel P92 for advanced power plant, *Acta Mater.* 45 (12) (1997) 4901–4907, [https://doi.org/10.1016/S1359-6454\(97\)00176-6](https://doi.org/10.1016/S1359-6454(97)00176-6) URL <http://www.sciencedirect.com/science/article/pii/S1359645497001766>.
 - [58] S. Morito, Y. Iwami, T. Koyano, T. Ohba, Effect of solution carbon and nitrogen on the microstructural size and crystallography of lath martensite in Fe–N and Fe–C alloys, *Mater. Trans.* 57 (3) (2016) 227–232, <https://doi.org/10.2320/matertrans.MB201503> URL https://www.jstage.jst.go.jp/article/matertrans/57/3/57_MB201503/article.
 - [59] P. Yan, Z. Liu, H. Bao, Y. Weng, W. Liu, Effect of tempering temperature on the toughness of 9cr–3w–3co martensitic heat resistant steel, 1980–2015, *Mater. Des.* 54 (2014) 874–879, <https://doi.org/10.1016/j.matdes.2013.09.017> URL <http://www.sciencedirect.com/science/article/pii/S0261306913008698>.
 - [60] T. Ishitsuka, Y. Inoue, H. Ogawa, Effect of silicon on the steam oxidation resistance of a 9%Cr heat resistant steel, *Oxid. Metals* 61 (1–2) (2004) 125–142, <https://doi.org/10.1023/B:OXID.0000016280.81734.3f> URL <http://link.springer.com/article/10.1023/B:OXID.0000016280.81734.3f>.
 - [61] A.M. Huntz, V. Bague, G. Beauplé, C. Haut, C. Sévéc, P. Lecour, X. Longaygue, F. Ropital, Effect of silicon on the oxidation resistance of 9% Cr steels, *Appl. Surf. Sci.* 207 (1) (2003) 255–275, [https://doi.org/10.1016/S0169-4332\(02\)01505-2](https://doi.org/10.1016/S0169-4332(02)01505-2) URL <http://www.sciencedirect.com/science/article/pii/S0169433202015052>.
 - [62] C.L. Briant, S.K. Banerji, Intergranular failure in steel: the role of grain-boundary composition, arXiv:<https://doi.org/10.1179/imtr.1978.23.1.164>, *Int. Met. Rev.* 23 (1) (1978) 164–199, <https://doi.org/10.1179/imtr.1978.23.1.164> URL <https://doi.org/10.1179/imtr.1978.23.1.164>.
 - [63] S. Morito, J. Nishikawa, T. Maki, Dislocation density within lath martensite in Fe–C and Fe–Ni alloys, *ISIJ Int.* 43 (9) (2003) 1475–1477, <https://doi.org/10.2355/isijinternational.43.1475> URL https://www.jstage.jst.go.jp/article/isijinternational/43/9/43_9_1475/article.
 - [64] H. Semba, F. Abe, Alloy design and creep strength of advanced 9%Cr USC boiler steels containing high concentration of boron, *Energy Materials* 1 (4) (2006) 238–244, <https://doi.org/10.1179/174892406X173611> URL <https://doi.org/10.1179/174892406X173611>.
 - [65] F. Abe, M. Tabuchi, M. Kondo, S. Tsukamoto, Suppression of Type IV fracture and improvement of creep strength of 9cr steel welded joints by boron addition, *Int. J. Press. Vessel. Pip.* 84 (1) (2007) 44–52, <https://doi.org/10.1016/j.ijpvp.2006.09.013> URL <http://www.sciencedirect.com/science/article/pii/S0308016106001633>.
 - [66] T. Horiuchi, M. Igarashi, F. Abe, Improved utilization of added B in 9cr heat-resistant steels containing W, *ISIJ Int.* 42 (Suppl) (2002) S67–S71, https://doi.org/10.2355/isijinternational.42.Suppl_S67 URL https://www.jstage.jst.go.jp/article/isijinternational/42/Suppl/42_Suppl_S67/article.
 - [67] H.L. Schick, *Thermodynamics of Certain Refractory Compounds*, Academic Press, 1976 (google-Books-ID: nygFtwEACAAJ).
 - [68] H. Magnusson, R. Sandström, Influence of aluminium on creep strength of 9–12% Cr steels, *Mater. Sci. Eng. A* 527 (1) (2009) 118–125, <https://doi.org/10.1016/j.msea.2009.07.060> URL <http://www.sciencedirect.com/science/article/pii/S0921509309008156>.
 - [69] R.L. Klueh, D.J. Alexander, M. Rieth, The effect of tantalum on the mechanical properties of a 9cr–2w–0.25v–0.07ta–0.1c steel, *J. Nucl. Mater.* 273 (2) (1999) 146–154, [https://doi.org/10.1016/S0022-3115\(99\)00035-5](https://doi.org/10.1016/S0022-3115(99)00035-5) URL <http://www.sciencedirect.com/science/article/pii/S0022311599000355>.
 - [70] X. Zhai, S. Liu, Y. Zhao, Effect of tantalum content on microstructure and tensile properties of CLAM steel, *Fusion Eng. Des.* 104 (2016) 21–27, <https://doi.org/10.1016/j.fusengdes.2016.01.016> URL <http://www.sciencedirect.com/science/article/pii/S0920379616300163>.
 - [71] Z.X. Xia, C. Zhang, N.Q. Fan, Y.F. Zhao, F. Xue, S.J. Liu, Improve creep properties of reduced activation steels by controlling precipitation behaviors, *Mater. Sci. Eng. A* 545 (2012) 91–96, <https://doi.org/10.1016/j.msea.2012.03.004> URL <http://www.sciencedirect.com/science/article/pii/S0921509312003437>.
 - [72] L. Helis, Y. Toda, T. Hara, H. Miyazaki, F. Abe, Effect of cobalt on the microstructure of tempered martensitic 9cr steel for ultra-supercritical power plants, *Mater. Sci. Eng. A* 510–511 (2009) 88–94, <https://doi.org/10.1016/j.msea.2008.04.131> URL <http://www.sciencedirect.com/science/article/pii/S092150930801424X>.
 - [73] K.J. Harrelson, S.H. Rou, R.C. Wilcox, Impurity element effects on the toughness of 9cr–1mo steel, *J. Nucl. Mater.* 141 (1986) 508–512, [https://doi.org/10.1016/S0022-3115\(86\)80091-5](https://doi.org/10.1016/S0022-3115(86)80091-5) URL <http://www.sciencedirect.com/science/article/pii/S0022311586800915>.
 - [74] C. Naudin, J. Frund, A. Pineau, Intergranular fracture stress and phosphorus grain boundary segregation of a Mn–Ni–Mo steel, *Scr. Mater.* 40 (9) (1999) 1013–1019, [https://doi.org/10.1016/S1359-6462\(99\)00069-X](https://doi.org/10.1016/S1359-6462(99)00069-X) URL <http://linkinghub.elsevier.com/retrieve/pii/S135964629900069X>.
 - [75] A.V. Nikolaeva, Y.A. Nikolaev, Y.R. Kevorkyan, Grain-boundary segregation of phosphorus in low-alloy, *Steel Times* 91 (1) (2001) 534–542 URL <https://doi.org/10.1023/A:1012482419952>.
 - [76] A.M. Gopakumar, P.V. Balachandran, D. Xue, J.E. Gubernatis, T. Lookman, Multi-objective optimization for materials discovery via adaptive design, *Sci. Rep.* 8 (1) (2018) 3738, <https://doi.org/10.1038/s41598-018-21936-3> URL <https://www.nature.com/articles/s41598-018-21936-3>.