**A Machine Learning Aided Interpretable Model for Yield Strength Prediction in Fe-based Martensitic and Austenitic Alloys**

Osman Mamun1, Madison Wenzlick2, 3, Jeffrey Hawk2, and Ram Devanathan1

1Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, USA

2Materials Performance Division, National Energy Technology Laboratory, 1450 Queen Avenue SW, Albany, OR 97321, USA

3Leidos Research Support Team, 1450 Queen Avenue SW, Albany, OR 97321, USA

**Abstract**

9-12% Cr ferritic-martensitic alloy (FMA) and austenitic 347H alumina forming alloy (AFA) have been receiving considerable attention due to their numerous applications in high temperature power generation industries. To design high strength steels with prolonged service life requires a thorough understanding of the long-term properties, e.g., creep rupture strength, rupture life, etc., as a function of the chemical composition and processing parameters that govern the microstructural characteristics. In this letter, the creep rupture strength of both 9-12 % Cr FMA and austenitic AFA has been parameterized using curated experimental datasets with a gradient boosting machine (GBM). The trained model has been cross validated against unseen test data and achieved high predictive performance in terms of correlation coefficient ( for 9-12% Cr FMA and for austenitic AFA) thus bypassing the need for time and cost prohibitive tensile test or physical theoretical calculations. Furthermore, the feature importance has been computed using the Shapley value analysis to understand the complex interplay of different features.

Austenitic AFA is used ubiquitously in power plants, mainly because of their excellent corrosion resistance properties, operating at high temperatures ) and pressures ( >)1–4; However, due to their inability to keep its crystal structure intact during cooling renders them extremely difficult to utilize for a long time5,6. In contrast, Martensitic steel offers higher resistance to creep at high temperature conditions, but are very prone to corrosive action in harsh environments7–9. For a fossil energy power plant, it is essential to have high operating temperature with prolonged alloy rupture life and rupture strength to ensure high thermodynamic efficiency, less carbon emission, and cost-effective operations. In order to push the temperature and pressure envelope of a modern fossil energy power plant, a systematic investigation of the chemical and mechanical factors affecting the rupture life and rupture strength is desired for two main reasons, 1. To have more control over the chemical composition and processing parameters that will yield the desired properties, and 2., to confidently assess the performance of a new alloy developed.

Experimental trial and error combined with physics based constitutive equations and/or computational CALPHAD based study are the two main approaches to develop predictive models for rupture life or rupture strength10; However, recently data science based approaches, such as machine learning, are emerging as a powerful tool to build accurate and reliable predictive model to make future predictions in several orders of magnitude shorter time than the traditional experimental or computational approaches10,11. Machine learning enables computer to find hidden latent rules in the data and later exploit them for making future prediction without any active human intervention12. In a recent assessment of the literature, data science method has been found to be more accurate than the physics-based method and/or thermodynamics-based models for the rupture life or rupture strength prediction13–16. However, the assessment also revealed a number of weaknesses in the existing ML models, e.g., 1. lack of dataset size, in particular a large amount of published data has not been used in building these models17, 2., the model accuracy is not high enough to make reliable predictions because of the use of inadequate algorithms, 3., an accurate interpretation of the developed model for the inverse design of novel alloy materials.

In this scientific work, a workflow combining machine learning with high quality experimental data is developed to construct an accurate predictive model for rupture strength prediction in 9-12% Cr FMA and Austenitic AFA. The workflow consists of the following steps:

1. Data preprocessing to convert the raw experimental data suitable for machine learning model training, e.g., removing features or instances with lots of missing values, imputation of the missing values using the mean of the rest of the values, and scaling the data using the mean and standard deviation of each feature.
2. Preliminary analysis of the data, e.g., correlation between different features and distribution of the data.
3. Training the machine learning algorithm using a 5-fold cross validation scheme.
4. Identification of the importance of different features on the machine learning model parameters.

The data used in this study were collected and compiled into a consistent and reliable set of data by the National Energy Technology’s (NETL) effort on Extreme Environment Materials (XMAT). XMAT is a collaborative undertaking between seven national labs with the goal of developing a suite of improved heat resistant alloys for various components in fossil energy power plants and to predict the long-term performance of these alloys, e.g., rupture life and rupture strength. By utilizing state-of-the-art computational materials modeling and cutting-edge experimental tools across the National Laboratories in conjunction with the industry partnership, XMAT is expected to accelerate the development and deployment of new heat resistant alloys for Fossil Energy applications. Two datasets, i.e., 9-12 % Cr FMA and Austenitic AFA, are used to build the predictive models as both are very important for different components of a fossil energy power plant. 9-12 % Cr FMA dataset contains 1203 data instances with 30 features and austenitic AFA dataset contains 823 data instances with 24 features after the preprocessing steps. In table 1., various aspects of both datasets, including feature names, description, unit, mean, and standard deviations etc. are tabulated. In figure 1, the Pearson correlation coefficient between the features are shown to identify collinear features which may lead to fitting artifacts that are not easy to disentangle.

A picture containing building, sitting, table, computer

Description automatically generatedA picture containing building, cage

Description automatically generated

**Figure 1:** Pearson correlation coefficients of all the features for the 9-12% Cr FMA (left) and austenitic AFA (right).

**Table 1:** Summary of the features present in both dataset including feature name, description, unit, mean, and standard deviation

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Feature name | Description | Unit | 9-12% Cr FMA | | Austenitic stainless steel | |
| Mean | Std. | Mean | Std. |
| Fe | Iron content |  |  |  |  |  |
| C | Carbon content |  |  |  |  |  |
| Cr | Chromium content |  |  |  |  |  |
| Mn | Manganese content |  |  |  |  |  |
| Si | Silicon content |  |  |  |  |  |
| Ni | Nickel content |  |  |  |  |  |
| Co | Cobalt content |  |  |  |  |  |
| Mo | Molybdenum content |  |  |  |  |  |
| W | Tungsten content |  |  |  |  |  |
| Nb | Niobium content |  |  |  |  |  |
| Al | Aluminum content |  |  |  |  |  |
| P | Phosphorous content |  |  |  |  |  |
| Cu | Copper content |  |  |  |  |  |
| Ti | Titanium content |  |  |  |  |  |
| Ta | Tantalum content |  |  |  |  |  |
| Hf | Hafnium content |  |  |  |  |  |
| Re | Rhenium content |  |  |  |  |  |
| V | Vanadium content |  |  |  |  |  |
| B | Boron content |  |  |  |  |  |
| N | Nitrogen content |  |  |  |  |  |
| O | Oxygen content |  |  |  |  |  |
| S | Sulfur content |  |  |  |  |  |
| Zr | Zirconium content |  |  |  |  |  |
| Homo | Homogenization | Yes/No |  |  |  |  |
| Normal | Normalization or Austenization heat treatment temperature |  |  |  |  |  |
| Temper1 | Temper heat treatment 1 |  |  |  |  |  |
| AGS No. | Austenitic grain size |  |  |  |  |  |
| TT\_Temp | Temperature |  |  |  |  |  |
| TT\_EL | Elongation to failure | % |  |  |  |  |
| TT\_RA | Reduction in area | % |  |  |  |  |

From the Pearson correlation coefficient analysis, no strong correlation is found amongst the features in the 9-12% Cr FMA dataset; However, some feature pairs in austenitic stainless steels dataset demonstrate strong correlation. Even though it is desirable to have low correlation amongst features present in the machine learning models, in this case it can be considered as a coincidence rather than correlation as those occurrences are only present for different chemical compositions.

To quantify the machine learning performance of different algorithms, correlation coefficient () of a linear fit of the predicted data and the actual data is used. To avoid overfitting in the regression modeling and to collect sufficient statistics about the model performance on unseen data, 5-fold cross validation is performed for each machine learning algorithm.

First, we train our model using a Gaussian Process (GP)18 regression with a combination of a radial basis function (RBF) and white kernel. The white kernel acts as a regularizer which accounts for noise by adding a constant to the diagonal elements of the covariance matrix. The scikit-learn Python package19 was used to train the model and log negative likelihood is used to optimize the kernel parameters. for testing set are 0.92 and 0.83, respectively, for the 9-12% Cr FMA and austenitic AFA. In Figure 2., the parity plots for the predicted strength and actual strength are shown for both the dataset for GP regression. The performance of austenitic stainless-steel model is worse than the 9-12% Cr FMA model, mainly due to the dataset size of the latter is about 1/3rd smaller and also it is missing some features that are present in the former. In Table 2., both the training and testing set performance for GP and all the subsequent algorithms are tabulated. From the table, it is evident that the model is not overfitting as the training and testing performance are quite similar; However, considering the performance of the state-of-the-art performance ( > 0.95) of yield strength prediction reported in the literature20 for similar dataset and the high variance of the , this performance is deemed as unsatisfactory for accurate prediction of rupture strength. It is worth noting here that, in that study, in addition to the chemical composition and processing parameters, synthetic alloy features generated using CALPHAD was incorporated to capture microstructural and phase transformation related information.

A picture containing dark, computer, smoke

Description automatically generatedA picture containing computer

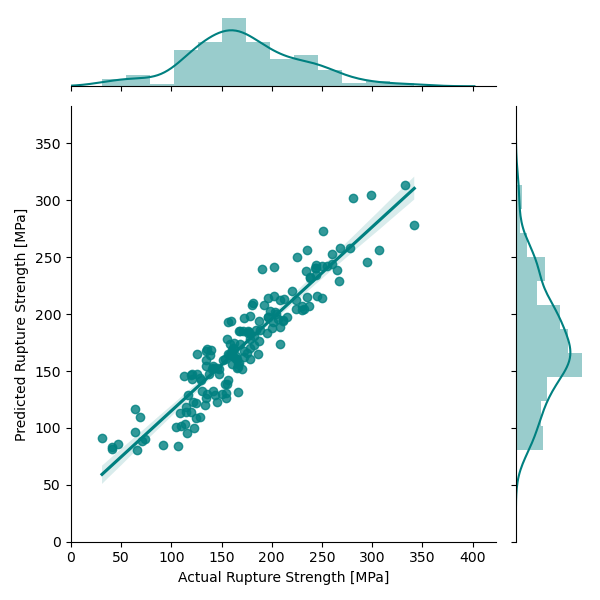
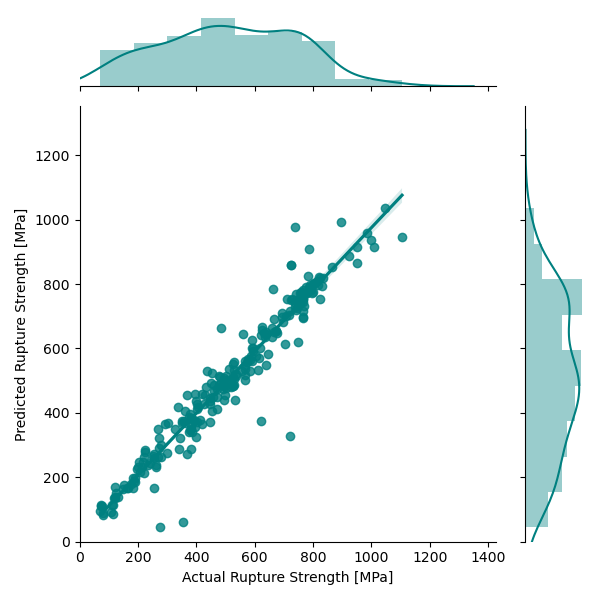
Description automatically generated

**Figure 1:** Parity plot for the testing data for GP regression for the 9-12% Cr FMA (left) and austenitic AFA (right).

**Table 2:** Summary of the machine learning model results in terms of correlation coefficient () for both the training and hold out testing set

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Gaussian Process | Neural Network | Gradient Boosting Machine |
| 9-12% Cr FMA | Training set |  |  |  |
| Testing set |  |  |  |
| Austenitic stainless steel | Training set |  |  |  |
| Testing set |  |  |  |

Next, a Neural Network (NN) was trained to map the underlying correlations between the features and the target properties, i.e., rupture strength. Keras Python package21 with TensorFlow backend22 was used to train the model. Two hidden layers, with 64 neurons each, were used in addition to the input layer and the output layer. For activation function, Rectified Linear Unit (ReLU) is used and the models were trained for 4000 epochs to ensure convergence. for testing set are 0.93 and 0.84, respectively, for the 9-12% Cr FMA and austenitic stainless steel dataset. In Figure 3., ML predicted rupture strength are plotted against the actual rupture strength. For the NN, similar performance was obtained as the GP, though the variance of is smaller, indicating the model performance is robust; However, considering the success of Gradient Boosting Machine (GBM)23 in several scientific articles and machine learning competition, a GBM based workflow was also built.



**Figure 3:** Parity plot for the testing data for NN regression for the 9-12% Cr FMA (left) and austenitic AFA (right).

Finally, a Gradient Boosting on Decision Trees (GBDT) was trained to predict the rupture strength of 9-12% Cr FMA and Austenitic. To train the model, CatBoost package with Python interface24 is used. for testing set are 0.98 and 0.95, respectively, for the 9-12% Cr FMA and austenitic stainless-steel dataset. In Figure 3., the parity plots of the actual and predicted data are illustrated. Based on the performance of the three algorithms, we conclude that GBDT is the best algorithm to build machine learning model conditioned on the chemical composition and processing related features for the accurate (mean of is high) and robust (variance of is low) prediction of rupture strength. Also, it is found that when the dataset volume is substantial, additional synthetic alloy features and/or intermediate model for Prior Austenitic Grain Size (PAGS) are not needed to faithfully map the underlying functional relation between the features and the rupture strength.

A picture containing computer, dark, computer

Description automatically generated

**Figure 4:** Parity plot for the testing data for GBDT regression for the 9-12% Cr FMA (left) and austenitic AFA (right).

Subsequently, Feature importance is investigated to quantify the effect of different features on the model performance for the GBDT algorithm using the Shapley value25,26 analysis. In Figure 5., feature importance plot for both the dataset is shown. For the 9-12% Cr dataset, Temperature, reduction in area, elongation are the most important features. As for the chemical composition, Carbon is the most important features, as expected, with a positive correlation indicating having more Carbon content increases the rupture strength of a martensitic steels. For austenitic AFA, Temperature, Austenitic grain size, elongation are the most important features. The positive correlation indicates to have a higher rupture strength in austenitic AFA, it is essential to have a large grain size of austenitic particles in order to reduce the grain boundaries which is usually responsible for crack propagation.

A picture containing drawing

Description automatically generatedA picture containing light

Description automatically generated

**Figure 5:** Importance of different features for the 9-12% Cr FMA (left) and austenitic AFA (right). The x-axis is plotted in terms of arbitrary unit to show the relative importance of different features, but the magnitude has no physical interpretation. Dark reddish color indicates negative correlation and dark bluish color indicates positive correlation.

In conclusion, a workflow is developed based on the GBDT algorithm to make efficient prediction of the rupture strength in 9-12% Cr FMA and austenitic AFA. Based on the Shapley value analysis, important features are identified and the findings are rationalized in light of the domain knowledge. By integrating these machine learning models into the existing alloy design strategy, significant acceleration can be gained for the identification of 9-12% Cr FMA or austenitic AFA materials with superior tensile strength properties.

**Associated content**

**Declaration of competing interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

This work was supported by the U.S. Department of Energy, Office of Fossil Energy, eXtreme environment MATerials (XMAT) consortium. This research used resources of the Pacific Northwest National Laboratory, which is supported by the U.S. Department of Energy.

**Supplementary materials**

Supplementary materials associated with this article can be found in the online version, at doi: 10.1016/j.scriptamat.2020.XX.XXX

**Disclaimer**

This manuscript was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express 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 therein 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 therein do not necessarily state or reflect those of the United States Government or any agency thereof.

**Graphical abstract:**

**A screenshot of a cell phone

Description automatically generated**

**Bibliography**

1. Chen, X. H., Lu, J., Lu, L. & Lu, K. Tensile properties of a nanocrystalline 316L austenitic stainless steel. *Scr. Mater.* **52**, 1039–1044 (2005).

2. Sourmail, T. Precipitation in creep resistant austenitic stainless steels. *Mater. Sci. Technol.* **17**, 1–14 (2001).

3. Yamamoto, Y. *et al.* Alloying effects on creep and oxidation resistance of austenitic stainless steel alloys employing intermetallic precipitates. *Intermetallics* **16**, 453–462 (2008).

4. Yamamoto, Y. *et al.* Creep-resistant, Al2O3-forming austenitic stainless steels. *Science (80-. ).* **316**, 433–436 (2007).

5. Bengochea, R., Lopez, B. & Gutierrez, I. Microstructural evolution during the austenite-to-ferrite transformation from deformed austenite. *Metall. Mater. Trans. A* **29**, 417–426 (1998).

6. Militzer, M., Mecozzi, M. G., Sietsma, J. & Van der Zwaag, S. Three-dimensional phase field modelling of the austenite-to-ferrite transformation. *Acta Mater.* **54**, 3961–3972 (2006).

7. Klueh, R. L. & Nelson, A. T. Ferritic/martensitic steels for next-generation reactors. *J. Nucl. Mater.* **371**, 37–52 (2007).

8. Klueh, R. L. *et al.* Ferritic/martensitic steels–overview of recent results. *J. Nucl. Mater.* **307**, 455–465 (2002).

9. Bischoff, J. *et al.* Corrosion of ferritic–martensitic steels in steam and supercritical water. *J. Nucl. Mater.* **441**, 604–611 (2013).

10. Vasudevan, M., Venkadesan, S., Sivaprasad, P. V & Mannan, S. L. Use of the Larson-Miller parameter to study the influence of ageing on the hardness of cold-worked austenitic stainless steel. *J. Nucl. Mater.* **211**, 251–255 (1994).

11. Ennis, P. J., Zielinska-Lipiec, A., Wachter, O. & Czyrska-Filemonowicz, A. Microstructural stability and creep rupture strength of the martensitic steel P92 for advanced power plant. *Acta Mater.* **45**, 4901–4907 (1997).

12. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* (2018) doi:10.1038/s41586-018-0337-2.

13. Shin, D., Yamamoto, Y., Brady, M. P., Lee, S. & Haynes, J. A. Modern data analytics approach to predict creep of high-temperature alloys. *Acta Mater.* **168**, 321–330 (2019).

14. Liu, Y. *et al.* Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning. *Acta Mater.* (2020).

15. Chatzidakis, S., Alamaniotis, M. & Tsoukalas, L. H. Creep Rupture Forecasting: A Machine Learning Approach to Useful Life Estimation. *Int. J. Monit. Surveill. Technol. Res.* **2**, 1–25 (2014).

16. Jiang, X. *et al.* A strategy combining machine learning and multiscale calculation to predict tensile strength for pearlitic steel wires with industrial data. *Scr. Mater.* **186**, 272–277 (2020).

17. Panchal, J. H., Kalidindi, S. R. & McDowell, D. L. Key computational modeling issues in Integrated Computational Materials Engineering. *CAD Comput. Aided Des.* (2013) doi:10.1016/j.cad.2012.06.006.

18. Williams, C. K. I. & Rasmussen, C. E. Gaussian processes for regression. in *Advances in neural information processing systems* 514–520 (1996).

19. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

20. Peng, J., Yamamoto, Y., Hawk, J. A., Lara-Curzio, E. & Shin, D. Coupling physics in machine learning to predict properties of high-temperatures alloys. *npj Comput. Mater.* **6**, 141 (2020).

21. Chollet, F. Keras: The python deep learning library. *ascl* ascl-1806 (2018).

22. Abadi, M. *et al.* Tensorflow: A system for large-scale machine learning. in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)* 265–283 (2016).

23. Song, K., Yan, F., Ding, T., Gao, L. & Lu, S. A steel property optimization model based on the XGBoost algorithm and improved PSO. *Comput. Mater. Sci.* **174**, 109472 (2020).

24. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. CatBoost: unbiased boosting with categorical features. in *Advances in neural information processing systems* 6638–6648 (2018).

25. Winter, E. The shapley value. *Handb. game theory with Econ. Appl.* **3**, 2025–2054 (2002).

26. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. in *Advances in neural information processing systems* 4765–4774 (2017).