



Understanding oxidation of Fe-Cr-Al alloys through explainable artificial intelligence

Indranil Roy , Bojun Feng, Subhrajit Roychowdhury, Sandipp Krishnan Ravi, Rajnikant V. Umretiya, Christopher Reynolds, Sayan Ghosh, Raul B. Rebak, Andrew Hoffman, General Electric Research, 1 Research Circle, Schenectady, NY 12309, USA

Address all correspondence to Andrew Hoffman at andrew.hoffman@ge.com

(Received 19 August 2022; accepted 28 November 2022; published online: 9 January 2023)

Abstract

The oxidation resistance of FeCrAl based on alloying composition and oxidizing conditions is predicted using a combinatorial experimental and artificial intelligence approach. A neural network (NN) classification model was trained on the experimental FeCrAl dataset produced at GE Research. Furthermore, using the SHapley Additive exPlanations (SHAP) explainable artificial intelligence (XAI) tool, we explore how the NN can showcase further material insights that are unavailable directly from a black-box model. We report that high Al and Cr content forms protective oxide layer, while Mo in FeCrAl creates thick unprotective oxide scale that is vulnerable to spallation due to thermal expansion.

Introduction

Throughout history, we have been mostly dependent on trial and error and serendipity^[1] for alloy design and discovery. Identifying new materials with specific property requirements is challenging. Although no computational approach can displace the ground truth of experiments, the time and cost associated with only experimental studies to guide us toward best alloy composition can retard significant progress. With the exponential increase in demand for new materials, data-driven modeling based on experimental dataset coupled with expert domain knowledge has flourished in the last decade,^[2] primarily from academic research. Although data-driven modeling (based on compositional features) has been applied to predict a variety of mechanical properties,^[3–6] corrosion and oxidation behavior are often considered a challenging problem due to the inherent complexity of these physical processes.^[7,8]

Although a consistent emphasis was allocated toward building models with high accuracy using ML models such as neural network, NN^[9–11] for different applications, interpretability of these models remains as a consistent limitation. One of the challenges of using data-driven modeling in material science is the scarcity of data which will be persistent until high-throughput experimental methods become ubiquitous. Therefore, along with building a good model for prediction, explainability of the results (even with smaller datasets) is essential to extract useful information to overcome the engineering challenges in materials discovery. The domain of explainable artificial intelligence (XAI) and interpretable ML (IML)^[9] was conceived to resolve this challenge. Methods and tools, such as SHapley Additive exPlanations (SHAP)^[5,12–15] and local interpretable model-agnostic explanations (LIME),^[11] were developed to extend predictive modeling to possess local level interpretability, so that the models can be verified, studied, and analyzed with

the support of domain knowledge. For this work, the SHAP approach is applied as the XAI tool to provide explainability to a black-box NN model. The NN model is chosen because of its effective predictive ability,^[16–18] properties like universal approximation theorem,^[19] and good compatibility with SHAP XAI tools. However, the primary focus of this paper is to boost the understanding of experimental data using XAI, but not to build best predictive modeling. SHAP applications have helped material scientists infer or recommend design rules, optimal compositions, and experimental parameters in alloy design and discovery. For example, Yan et al.^[19] used SHAP for designing high fatigue strength alloys, where SHAP analysis recommended increasing Cr and Mo concentration and decreasing tempering and normalizing temperature to achieve high fatigue strength. Additionally, both Xiong et al.^[5] and Yang et al.^[16] used SHAP to find important parameters for improving hardness of the high-entropy alloys (HEAs).

One such industrial goal is to optimize the elemental composition of FeCrAl alloys for its applications in nuclear environments. Because of the limitations of Zr-based alloys in loss of coolant accident (LOCA) scenarios, FeCrAl is being developed as nuclear fuel cladding material. In a nuclear reactor, the cladding material has to combat low temperature (~300–400°C) normal boiling water reactor (BWR) chemistry water and steam for long (>100 h) periods of time during normal operation conditions. Additionally, accident tolerant fuel cladding must have good resistance to high temperature (~1200°C) oxidation for short (~2–4 h) thermal excursions due to LOCA event.^[17] This poses the challenge of designing alloys with both high and low temperature oxidation resistance for short and long exposure, respectively. In the past, a consistent experimental effort, especially at high temperature,^[12,18,20,21] has been noticed to optimize FeCrAl composition for nuclear applications, but

a computational predictive modeling is rare. Previously, we have built predicative models for oxidation mass gain based on FeCrAl alloy composition and oxidation conditions.^[13] While performing such predications, one of the challenges faced was skewness or scarcity of the dataset. One approach to encounter such difficulty is to use classification instead of regression to understand oxidation behaviors. A material scientist can classify an alloy to be a good or bad oxidation-resistant alloy based on the amount of mass gained per unit surface area. Here, we apply NN classifier to segregate FeCrAl based on composition and oxidation condition, e.g., temperature and duration. On top of that, a black-box model such as a NN is augmented with the explainability of machine learning methods. We apply XAI which is still under development in computer science community and surely new to material science, to further extract useful material insights about FeCrAl oxidation.

Data and modeling

The experimental dataset used in this work is produced on FeCrAl alloy both from literature and ongoing experimental effort in GRC. In this work, only experiments from steam oxidation condition are included in model building and validation, due to the limited amount of data from air and prototypic hydrothermal corrosion conditions in BWRs. New steam data used in these studies were from model alloys that were first vacuum induction cast, then spin cast into water cooled Cu molds. Steam oxidation studies were carried out in Ar purged tube furnaces, and samples were brought up to the test temperature prior to adding steam. The details of the experimental test condition were provided in Ref.^[21] The wt% of Fe, Cr, Al, Mo, and Ni are selected as input features for predictive model building along with oxidation temperature and test duration. In addition, instead of formulating the problem as a regression problem to predict the experimental mass gain of alloys in our previous paper,^[8] the experimental data are categorized into three classes, spallation (mass gain < 0 g/cm² or mass loss), good oxidation resistance (mass gain between 0 and 5.0×10^{-4} g/cm²), and non-oxidation resistance (mass gain greater than 5.0×10^{-4} g/cm²). The reasons of having the classification problem are (1) the goal of the paper is to use XAI to explain what factors drive the FeCrAl alloys to have good/bad corrosion resistance, not to derive a predictive modeling; (2) the dataset is highly skewed and scarce, with the absolute values of mass gain/loss concentrate in the order of 0 to $\sim 10^{-4}$ g/cm² and $\sim 10^{-1}$ g/cm². This distribution of mass gain can be naturally treated as a classification problem and better for XAI purpose. For the choice of the threshold, it was decided to balance of the number of data points (for training and testing) in each class. In the end, by an 80–20% split of dataset, the training set contains 66 data points (9 spallation, 42 good, and 15 non-oxidation), and the test set contains 17 data points (5 spallation, 9 good, and 3 non-oxidation).

In addition, a 2-layer NN with 20 hidden unit and ReLU activation function in each layer is used to build the predictive

modeling using PyTorch.^[14] L2 norm is used to reduce overfitting because of the scarcity of data. The hyperparameters are chosen to attempt to reduce the number of weights to be trained, as well as to reduce overfitting, through cross validation, so that the model trained can have good training and testing accuracy, because of the limitation of number of available experimental data points. On the other hand, to build ML models in the materials science domain, a certain level of overfitting is fine, as every experimental result is important and time-consuming to obtain, and the model is expected to understand as much of the data distribution as possible. Especially for XAI purpose, reasonable level of overfitting can help the model to understand the trend of data distribution and explain our scarce data better. For deriving XAI explanation, Python SHAP package is used, which is naturally compatible with PyTorch model.

Results and discussion

The prediction accuracies of the NN model are 100% for training set and 88.2% (15/17) for test set. The confusion matrix of the test set is $\begin{bmatrix} 4 & 1 & 0 \\ 0 & 9 & 0 \\ 0 & 1 & 2 \end{bmatrix}$. The *F1* scores for spallation, good, and non-projective oxidation are 88.9%, 90%, and 80%. It is important to note that although a certain level of overfitting is expected, the test set accuracy is good, as only two data points are incorrectly classified. From the test set confusion matrix, 1 out of 5 spallation cases and 1 out of 3 non-oxidation resistances are classified incorrectly, which can be due to lack of training data for those two specific classes.

The average SHAP contribution of all the features is provided in Fig. 1(a). As mentioned earlier, SHAP values explain how each of the features locally impact the selection of three classes of oxidation resistance (both directionally and quantitatively). The SHAP value of each of the feature quantifies the numerical influence on the final prediction. A positive SHAP value indicates that the feature pushes the prediction toward the particular class and negative value deters the selection of the class. Apart from the directional information, the quantitative estimation of the SHAP values can precisely estimate how much the positive or negative influence is. Mo has the highest contribution in classifying a composition as a good oxidation-resistant alloy, while Al has the most influence on spallation and non-oxidation-resistant alloys. This may come as counter-intuitive if analyzed from a stable oxide forming element perspective. As the feature importance plot provides very little directional information, one gets no indication about the explainability of results from such plot. A scatter plot for the data points for all the classes is shown in Fig. 1(b–d) which indicate the directional feature importance. For good oxidation-resistant alloys, Mo concentration always contributes negatively, while Al primarily contributes positively, i.e., increases the chance of forming oxidation-resistant alloy [Fig. 1(b)]. As the blue and red dots are intertwined for Cr in Fig. 1(b), its contribution is inconclusive. Figure 1(c) clearly demonstrates the effect of Cr and Al in predicting non-oxidation-resistant alloy. Alloys with low Al and Cr wt% (blue dots) increase the chances

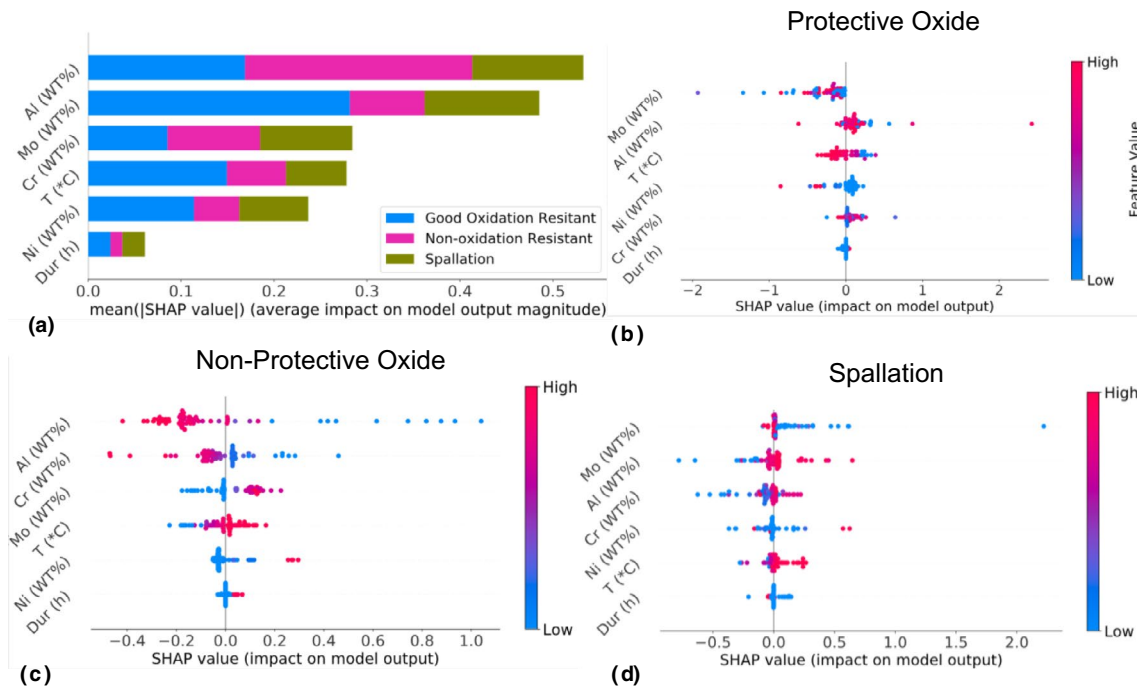


Figure 1. (a) Feature importance from SHAP values for good or protective, non-oxidation resistant (unprotective), and spallation classification, and (b–d) SHAP contribution of all features for every data point in the dataset.

of forming non-oxidation-resistant alloy, e.g., alloys with thick oxide scale. On the other hand, high Mo-containing alloys have higher chances of forming thick oxide scale as the red dots lie toward the right side in Fig. 1(c). In the experimental literature of FeCrAl oxidation, the addition of Al and Cr has been proven to form thin oxide scale at both high and low temperatures.^[21] If Mo is present, it forms a thick oxide scale at low temperature, making the alloy unprotective of oxidation.^[21] Such understandings are drawn from expensive experimental characterization like scanning electron microscopy (SEM) and transmission electron microscopy (TEM) of different FeCrAl compositions. Gaining similar insight from specific mass change data only is new to the community and can provide meaningful insight from inexpensive and high-throughput tests if XAI is used correctly. In terms of spallation [Fig. 1(d)], the absence of Mo (blue dots)

increases the change of oxide layer falling off. This is mildly contradictory, and no such direct evidence is obtained so far in the literature. That being said, FeCrAl alloys in the absence of reactive elements and Mo, especially with zero to very low Al content, are susceptible to spallation at high temperature oxidation as reported in Ref. ^[21]. The presence of high Ni concentration in FeCrAl will generally encourage low oxidation resistance and more changes of spallation, therefore, will not form good oxidation-resistant alloys. Ni is kinetically favored to form an oxide, but NiO is not a passivating layer as shown in the previous literature.^[15] Table I lists a summary of SHAP elemental contribution.

Next, we pull two alloys, one predicted correctly by the model [Fig. 2(a)], while the other was an incorrect prediction [Fig. 2(b)] to further analyze the SHAP contributions for each

Table I. Summary of SHAP contributions for elements of interest shown in Fig. 1.

Element of interest	Elemental concentration	Protective oxide contribution	Non-protective oxide contribution	Spallation of oxide contribution
Al	High	Positive	Positive	Positive/Neutral
	Low	Positive	Negative	Negative
Cr	High	Positive	Negative	Positive/Neutral
	Low	Positive	Positive	Negative
Mo	High	Negative	Positive	Neutral/Negative
	Low	Negative	Negative/Neutral	Positive
Ni	High	Negative	Positive	No Trend
	Low	Positive	Negative/Neutral	No Trend

deeply. According to the original dataset and the threshold of good oxidation-resistant alloy, they both form thin oxide scale as the mass gain is small (below $5.0 \times 10^{-4} \text{ g/cm}^2$). Steam oxidation of Fe–21Cr–5Al–3Mo at 900°C after 4 h is predicted correctly with 84.9% chance for forming protective scale, while the same composition tested at 1000°C for 2 h is predicted incorrectly with 78.8% chance of forming thick oxide scale. The SHAP contribution of all the components remains within $\pm 10\%$ except for Mo. Mo reduces the chances for both the cases. For the high temperature short duration test [Fig. 2(b)], SHAP negative contribution of Mo increases by 30%, but it has not decreased high enough to classify it as good oxidation material. Looking back at the experimental specific mass gain data [i.e., $5.75 \times 10^{-4} \text{ g/cm}^2$] for the wrong classification, we found it to be very close to the threshold [$5.0 \times 10^{-4} \text{ g/cm}^2$] as well. For the other wrongly classified case, although the ground truth class is spallation that is classified to be good, the mass change is a small negative value [$-5.2 \times 10^{-6} \text{ g/cm}^2$], which is very close to the threshold (0 g/cm^2). This emphasizes the importance of careful selection of threshold and care should be taken to select the threshold if a problem statement is changed from regression to classification.

The two-way interaction plot of two variables is important to understand the interaction of features, as shown for Al and temperature of steam oxidation in Fig. 3. Alloys without Al at high temperature tend to be poor oxidation resistant as the absence of Al contributes positively toward classification of unprotective oxide formation which can be observed from Fig. 3(e). As our dataset is not well distributed in the temperature and Aluminum range, we stay away from drawing further conclusive information. The purpose is to demonstrate the ability of SHAP to inform what the model is doing and not treat it as a black box.

Provided all the features remaining unaltered, effect of one parameter change is helpful to understand the effect of single feature. The effect of Al wt% changes from 0 to 7% for Fe–21Cr alloy for oxidation at 1300°C for 4 h, e.g., regular operating scenario [Fig. 4(a)] and 600°C for 100 h, e.g., accidental scenario [Fig. 4(b)]. At 600°C , all the alloys fall into

good oxidation-resistant alloy category, primarily due to the presence of high Cr (21 wt%) forming protective oxide. At high temperature, however, low Al alloys form thick oxide scale, and more than 6 wt% Al alloys tend to indicate spallation. At high temperature ($>1000^\circ\text{C}$), only Al oxide is protective, and its absence will create thick oxide scale, hence oxidation non-protective alloy. On the other hand, as the model alloy does not have Mo or any other reactive element, high Al-containing alloys are susceptible to spallation which has been experimentally reported.^[21] The optimal performance is seen for 3.5 to 6 wt% of Al. At low temperature, the Cr oxide is protective enough to make the alloy oxidation resistant even in the absence of Al.

Conclusion

Understanding the effect of alloy composition and environmental condition on oxidation behavior of alloys is a difficult problem. Given the large compositional space of FeCrAl alloys and its effectiveness as nuclear cladding materials over Zr-based alloys, data-driven modeling can guide material scientists to the optimized composition. The small number of experimental data points of target specific mass gain dataset, to begin with, poses a challenge that has been tackled with changing the problem into a classification model instead of a regression one. The proposed NN predicts the classes correctly for 15 alloys among 17 test cases. Using XAI, we explored the fundamental understanding that can only be obtained from expensive material characterization experiments. We found XAI to be extremely powerful in terms of understanding the basic physics if interpreted correctly. For the discussion before, the important conclusions drawn from XAI are proposed in Table II.

This is the first work of applying XAI to understand corrosion in any alloy system. Along with building predictive models, explaining the inner workings of the model is needed to understand the influence of different material features into the target property prediction. We understand that there are challenges about skewness or scarcity of the dataset, but the goal of this paper is to show the capability of XAI in terms

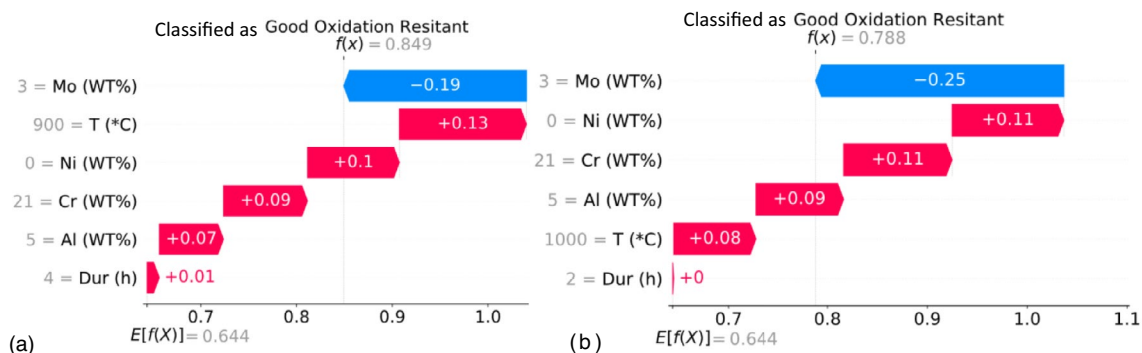


Figure 2. SHAP contribution of the features of Fe–21Cr–5Al–3Mo alloy after oxidation at (a) 900°C after 4 h classified correctly and (b) 1000°C after 2 h classified incorrectly.

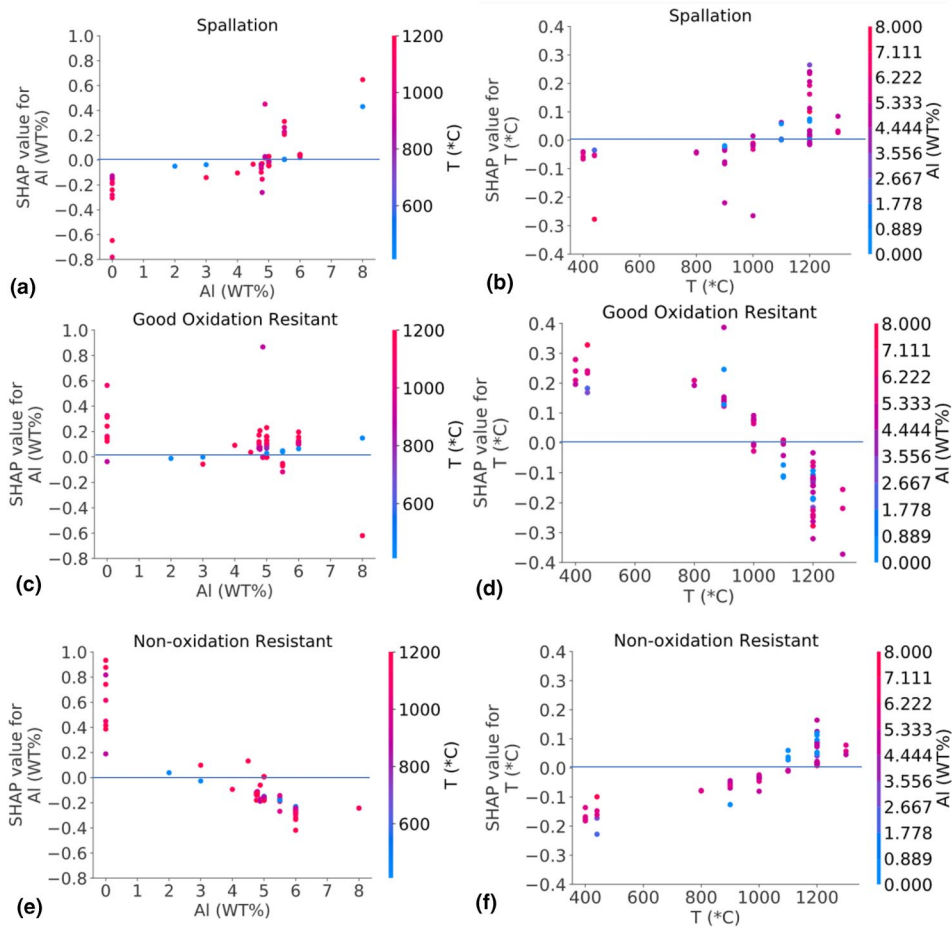


Figure 3. Effect of (a, c, e) Al wt% and (b, d, f) temperature (°C) on classifying FeCrAl as (a, b) spallation, (c, d) protective, and (e, f) non-protective scale forming alloy. The blue line indicates the zero SHAP contribution, above and below the line SHAP values are positive and negative, respectively.

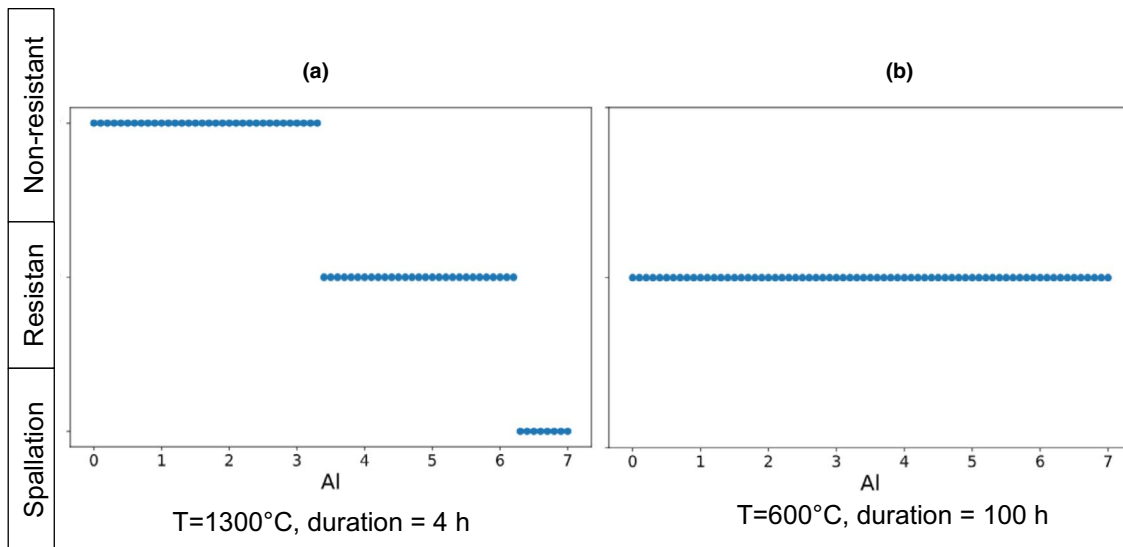


Figure 4. Prediction of Fe-21Cr-xAl (where x varies from 0 to 7 wt%) ability to resist oxidation at (a) 1300°C for 4 h and (b) 600°C 100 h.

Table II. Proposed contribution of different alloying elements of FeCrAl in forming protective oxide layer.

Element	Proposed explanation of SHAP contributions
Al	Al is primary passivator at high temperature and contributes to passivation synergistically with Cr at low temperatures. Al may contribute to spallation due to increasing coefficient of thermal expansion through stabilizing increased thermal vacancy concentration
Cr	Cr is primary passivator at low temperatures and contributes synergistically to passivation with Al at high temperatures. Cr does not significantly contribute to spallation behavior
Mo	Mo forms non-passive oxides which prevent dense passive oxide layers from forming. High Mo content may contribute to an increase in coefficient of thermal expansion leading to spallation of the oxide layer
Ni	Low concentration of Ni can contribute positively to passivation by acting as a catalyst for Cr and Al oxide formation. High Ni content contributes to non-passive less dense spinel formation

of understanding experimental data and help the design and discovery of new alloys. Once new experiments are available based on the study of XAI, new data will be incorporated to update the ML model and the XAI understanding of experimental data will be updated, and new experiments can be proposed, which the idea of active learning.

Acknowledgments

This work is supported by the US Department of Energy, National Nuclear Security Administration, under Award Number DE-NE0009047.

Data availability

The data supporting this study's findings are available from the corresponding author, AH, upon reasonable request.

Declarations

Conflict of interest

There is no conflict of interest.

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. J. Ling, M. Hutchinson, E. Antono, S. Paradiso, B. Meredig, High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates. *Integr. Mater. Manuf. Innov.* **6**, 207–217 (2017). <https://doi.org/10.1007/s40192-017-0098-z>
2. A. Roy, I. Roy, L.J. Santodonato, G. Balasubramanian, Data-guided feature identification for predicting specific heat of multicomponent alloys. *JOM* **74**, 1406–1413 (2022). <https://doi.org/10.1007/s11837-022-05183-6>
3. J. Xiong, T. Zhang, S. Shi, Machine learning of mechanical properties of steels. *Sci. China Technol. Sci.* **63**, 1247–1255 (2020). <https://doi.org/10.1007/s11431-020-1599-5>
4. Z. Deng, H. Yin, X. Jiang, C. Zhang, G. Zhang, B. Xu, G. Yang, T. Zhang, M. Wu, X. Qu, Machine-learning-assisted prediction of the mechanical properties of Cu–Al alloy. *Int. J. Miner. Metall. Mater.* **27**, 362–373 (2020). <https://doi.org/10.1007/s12613-019-1894-6>
5. J. Xiong, S.-Q. Shi, T.-Y. Zhang, Machine learning of phases and mechanical properties in complex concentrated alloys. *J. Mater. Sci. Technol.* **87**, 133–142 (2021). <https://doi.org/10.1016/j.jmst.2021.01.054>
6. S. Felix, S. Ray Majumder, H.K. Mathews, M. Lexa, G. Lipsa, X. Ping, S. Roychowdhury, T. Spears, In situ process quality monitoring and defect detection for direct metal laser melting. *Sci. Rep.* **12**, 8503 (2022). <https://doi.org/10.1038/s41598-022-12381-4>
7. A. Roy, M.F.N. Taufique, H. Khakurel, R. Devanathan, D.D. Johnson, G. Balasubramanian, Machine-learning-guided descriptor selection for predicting corrosion resistance in multi-principal element alloys. *NPJ Mater. Degrad.* **6**, 1–10 (2022). <https://doi.org/10.1038/s41529-021-00208-y>
8. I. Roy, S. Roychowdhury, B. Feng, S.K. Ravi, S. Ghosh, R. Umretiya, R.B. Rebak, D.M. Ruscitto, V. Gupta, A.K. Hoffman, Data-driven predictive modeling of FeCrAl oxidation. *MRS Adv.* (2022) (under review)
9. P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: a review of machine learning interpretability methods. *Entropy* **23**, 18 (2021). <https://doi.org/10.3390/e23010018>
10. R. Mitchell, E. Frank, G. Holmes, GPUtreeShap: massively parallel exact calculation of SHAP scores for tree ensembles. *PeerJ Comput. Sci.* **8**, e880 (2022). <https://doi.org/10.7717/peerj-cs.880>
11. “Why Should I Trust You?” | Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (n.d.). <https://dl.acm.org/doi/abs>. Accessed 25 July 2022
12. B.A. Pint, K.A. Terrani, R.B. Rebak, Steam oxidation behavior of FeCrAl cladding, in *Proceedings of the 18th International Conference on Environmental Degradation of Materials in Nuclear Power Systems—Water Reactors*. ed. by J.H. Jackson, D. Paraventi, M. Wright (Springer, Cham, 2019), pp.1451–1460. https://doi.org/10.1007/978-3-030-04639-2_96
13. Understanding Impacts of Chemistry on Oxidation of FeCrAl Alloys (2022). <https://doi.org/10.21203/rs.3.rs-1927179/v1>
14. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* (Curran

- Associates, Inc., 2019). <https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html>. Accessed 25 July 2022
15. I. Roy, P.K. Ray, G. Balasubramanian, Examining oxidation in β -NiAl and β -NiAl + Hf alloys by stochastic cellular automata simulations. *NPJ Mater. Degrad.* **5**, 1–9 (2021). <https://doi.org/10.1038/s41529-021-00202-4>
 16. C. Yang, C. Ren, Y. Jia, G. Wang, M. Li, W. Lu, A machine learning-based alloy design system to facilitate the rational design of high entropy alloys with enhanced hardness. *Acta Mater.* **222**, 117431 (2022). <https://doi.org/10.1016/j.actamat.2021.117431>
 17. R.B. Rebak, *Accident-Tolerant Materials for Light Water Reactor Fuels* (Elsevier, Amsterdam, 2020)
 18. K.A. Unocic, Y. Yamamoto, B.A. Pint, Effect of Al and Cr content on air and steam oxidation of FeCrAl alloys and commercial APMT alloy. *Oxid. Met.* **87**, 431–441 (2017). <https://doi.org/10.1007/s11085-017-9745-1>
 19. F. Yan, K. Song, Y. Liu, S. Chen, J. Chen, Predictions and mechanism analyses of the fatigue strength of steel based on machine learning. *J. Mater. Sci.* **55**, 15334–15349 (2020). <https://doi.org/10.1007/s10853-020-05091-7>
 20. V.K. Gupta, M. Larsen, R.B. Rebak, Utilizing FeCrAl oxidation resistance properties in water, air and steam for accident tolerant fuel cladding. *ECS Trans.* **85**, 3 (2018). <https://doi.org/10.1149/08502.0003ecst>
 21. A. Chikhalikar, I. Roy, H. Abouelella, R. Umretiya, A.K. Hoffman, M. Larsen, R.B. Rebak, Effect of aluminum on the FeCr(Al) alloy oxidation resistance in steam environment at low temperature (400 °C) and high temperature (1200 °C). *Corros. Sci.* **209**, 110765 (2022). <https://doi.org/10.1016/j.corsci.2022.110765>

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