Journal of Applied Physics

Estimating the thermal insulating performance of multi-component refractory ceramic systems based on a machine learning surrogate model framework

Cite as: J. Appl. Phys. **127**, 215104 (2020); https://doi.org/10.1063/5.0004395
Submitted: 12 February 2020 . Accepted: 16 May 2020 . Published Online: 02 June 2020

D. P. Santos, P. I. B. G. B. Pelissari, R. F. de Mello, and V. C. Pandolfelli

Journal of Applied Physics ARTICLE scitation.org/journal/jap

Estimating the thermal insulating performance of multi-component refractory ceramic systems based on a machine learning surrogate model framework

Cite as: J. Appl. Phys. **127**, 215104 (2020); doi: 10.1063/5.0004395
Submitted: 12 February 2020 · Accepted: 16 May 2020 ·
Published Online: 2 June 2020

View Online Export Citation CrossMark

D. P. Santos,[1,2,a)] P. I. B. G. B. Pelissari,[1,2] R. F. de Mello,[3] and V. C. Pandolfelli[1]

AFFILIATIONS

[1]Department of Materials Engineering, Federal University of São Carlos, São Carlos, São Paulo 13565-905, Brazil
[2]HiTemp Technological Solutions, São Carlos, São Paulo 13560-251, Brazil
[3]Department of Computer Science, Institute of Mathematics and Computer Science, University of S ao Paulo, São Carlos, São Paulo 13566-590, Brazil

Note: This paper is part of the special collection on Machine Learning for Materials Design and Discovery
[a)]**Author to whom correspondence should be addressed:** dns.prado.s@gmail.com

ABSTRACT

Predicting the insulating thermal behavior of a multi-component refractory ceramic system could be a difficult task, which can be tackled using the finite element (FE) method to solve the partial differential equations of the heat transfer problem, thus calculating the temperature profiles throughout the system in any given period. Nevertheless, using FE can still be very time-consuming when analyzing the thermal performance of insulating systems in some scenarios. This paper proposes a framework based on a machine learning surrogate model to significantly reduce the required computation time for estimating the thermal performance of several multi-component insulating systems. Based on an electric resistance furnace case study, the framework estimated the feasibility and the final temperature of nearly 1.9×10^5 insulating candidates' arrangements with reasonable accuracy by simulating only an initial sample of 2.8% of them via FE. The framework accuracy was evaluated by varying the initial sample size from ≈0.9% to 8% of total combinations, indicating that 3%–5% is the optimal range in the case study. Finally, the proposed framework was compared to the evolutionary screening procedure, a previously proposed method for selecting insulating materials for furnace linings, from which it was concluded that the machine learning framework provides better control over the number of required FE simulations, provides faster optimization of its hyperparameters, and enables the designers to estimate the thermal performance of the entire search space with small errors on temperature prediction.

NOMENCLATURE

ERF	Electric resistance furnace
ESP	Evolutionary screening procedure
FE	Finite elements
FN	False negative
FP	False positive
GA	Genetic algorithm
LS	Learning set
MAE	Mean-absolute error
ML	Machine learning
MLPc	Multi-layer perceptron classifier
MLPr	Multi-layer perceptron regressor
RMSE	Root mean squared error
RRc	Ridge regression classifier
TN	True negative
TN	True positive
XGB	Extreme gradient boosting

I. INTRODUCTION

Industrial processes involving high temperatures usually require insulation linings containing more than one layer of refractory ceramic materials in order to optimize costs, overall thicknesses, and insulating efficiency.[1] All those require critical decision steps related to selecting a proper set of commercial products to design and build the roofs, walls, and floors of heating chambers. Designers need to calculate the thermal performance of those linings based on their properties so they can compare costs and benefits among the candidates.[2] Nevertheless, assessing combinations of multiple material layers leads to solving a computationally intensive dynamical system,[3] which relies on initial conditions, heat propagation, and transient behaviors. In practical terms, the temperature profile along the insulating lining will be a consequence of the interaction of thermal properties of each component material, which responds differently to temperature variations.[4]

Finite element (FE) modeling has been used to calculate the heat transfer and resultant temperature distribution for complex scenarios by using computer simulations.[5,6] However, depending on the object structural complexities, physical phenomena to be modeled (which may increase the number and complexity of model equations), mesh refinement, total processing time and other factors, the FE model can be quite time-consuming, thus hindering a full analysis of the thermal performance when there are various likely combinations of materials.[7]

One may take advantage of other computer techniques in conjunction with FE to reduce the time required to estimate the temperature distributions of a vast number of insulating systems.[8–10] As an example, an evolutionary screening procedure (ESP) based on multi-objective genetic algorithms (GAs) and FE was capable of presenting 100 near-optimal trade-off lining systems by simulating $\approx$3.8% of all the possible configurations.[7] Although this methodology has proven to be useful from the perspective of material selection, other performance-equivalent but faster surrogate strategies should be taken into account to substitute the FE model.[11]

A powerful collection of computational tools is available within the area of machine learning (ML).[12] For instance, supervised learning algorithms have been frequently used to make data-driven predictions in materials science domains.[13–15] They theoretically converge to the best as possible classifier/regression function mapping input variables (X) into output ones (Y), according to samples obtained from some joint probability distribution $P(X \times Y)$.[16] The reason they are referred to as "supervised" is that given some training set with known values provided for $(x, y) \in X \times Y$, the algorithm "learns" those relations, and is capable of generalizing (or extrapolating) known patterns to unseen inputs still inferring their outputs. A careful description of several mechanisms, theories, and algorithms that support the statistical learning theory and ML foundations can be found elsewhere.[16–19]

In this context, this paper introduces an ML framework designed and developed to predict the thermal performance of electric resistance furnace (ERF) linings under different configurations, all based on candidate materials obtained from a previous study.[7] The framework consists of a two-stage supervised learning: (i) initially, a classification model is used to predict whether a combination of materials would fail, i.e., if the temperature profile along the insulating lining would exceed its maximum working temperatures (thus operating as a problem constraint) and (ii) next, a regression model is used to predict the final external temperature of the feasible lining systems (thus working as the objective function). The previously developed FE model[7] was applied to carry out thermal simulations of nearly 2.8% of all combinations of candidate materials as an initial learning sample. Afterward, the ML framework was applied to learn the non-linear relations mapping the furnace linings (according to their attributes: thermal conductivity, specific heat capacity, density, thickness, and maximum working temperature) into their respective thermal responses.

This study aims to present a faster and effective alternative for inferring the thermal responses of insulating systems in an attempt to simplify a multi-objective materials selection optimization. Unlike a previous study,[7] in which GA was capable of finding a near-Pareto[20] front by calculating the performance of around 3.8% of the combinations, the machine learning framework herein proposed can be used to estimate the behavior of all feasible system settings, thus providing an approximation for the real Pareto curve. The pros and cons of each methodology were discussed and, additionally, the effect of the initial set size in the predictability of the proposed model was evaluated in the range of 0.9% to 8% of all likely material configurations subject to the required constraints.

II. MATERIALS AND METHODS

A. Electric resistance furnace case study

This work is based on the case study of an electric resistance furnace (ERF) optimization, originally built up in a previous study.[7] Moreover, 121 products were selected as potential candidates to build the insulating system of a laboratory ERF with wall dimensions of $300 \times 300\,\text{mm}^2$ operating up to $1600\,^\circ$C for ceramic thermal processing purposes. From each product data sheet, the following pieces of information were collected: product form (rigid or flexible); thermal properties (thermal conductivity, specific heat capacity, and maximum operating temperature); thickness; density; and price (in US\$/m^2). The objective is to investigate which combination of insulating products would minimize both the furnace external temperature and the insulation cost such that its total thickness remain in the range of 40–200 mm and respecting the individual materials maximum operating temperature.

B. Insulating systems' datasets

Each insulating system comprises a sequence of products in the inner, middle, and outer layers, respecting the following building rules: (i) in the inner layer, only rigid materials with a maximum working temperature above $1600\,^\circ$C can be used; (ii) in the middle one, only materials withstanding a maximum working temperature higher than $1180\,^\circ$C are allowed; and (iii) at the outer layer, only those with a maximum working temperature below $1260\,^\circ$C were accepted. Under these constraints, 37 products could be used in the inner layer, 105 in the middle, and 49 in the outer one so that a total of 1.9×10^5 insulating systems were, at first, likely to be used.

In order to create the data examples for the insulating system, as a training set for the supervised machine learning algorithms, 5538 products combinations (around 2.83% of the total possibilities) were randomly chosen according to a single criterion: the number of times a product appeared in a layer should be approximately equal to all other products in that same layer such that the final set consisted of a proportionally balanced distribution of product combinations. From now on, this will be called the Learning Set (LS).

LS was then simulated by the FE model to obtain two pieces of information: (i) check whether the insulating system respected the constraints of maximum working temperatures and (ii) calculate the lining external temperature after the heating process. Then, both were added to LS as response attributes of the respective lining product combination.

C. Feature engineering

The next step was to convert product names into quantifiable variables, which could be interpreted by the ML procedure. Each insulating product was described using nine new variables: the thermal conductivity value for six equally spaced temperature points, the specific heat capacity multiplied by the material density, the product thickness, and the maximum working temperature. All data supporting our findings are available within the Appendix of Ref. 7. The authors collected the materials' properties from several commercial data sheets available from their manufacturers and prices either from different suppliers or after directly quoting with the manufacturers which were then converted to US dollars by square meter for normalization purposes.

At this point, LS consisted of 5538 rows by 29 columns (9 input attributes per insulating product plus 2 response ones). In order to improve the learning capability, each LS value was re-scaled into the range $[-1, 1]$. Additionally, the principal component analysis[16] was applied as a dimensionality reduction tool, for which 99.9% of the data variance was explained by the selected eigenvectors.

The same procedure was repeated to create another collection of insulating systems for testing the trained ML algorithms. To obtain consistent results, a much larger quantity of examples, 17 924 combinations of insulating products, was randomly chosen while respecting a single criterion: they were not present in the LS. From now on, this collection will be referred to as testing set. Datasets of sizes varying in the range of 0.9%–8.0% also followed this procedure in order to measure their influences on the predictability of the ML model.

D. Machine learning framework

The framework for the learning task was divided into two steps. (i) The first, referred to here as the classification task, a classification algorithm should learn to determine whether, during the heating process, a combination of insulating products would respect the temperature constraints across all three layers. The whole system fails if any of the materials become hotter than its maximum allowed temperature. (ii) The second step, referred to here as the regression task, a regression algorithm should learn to predict the external temperature after the whole heating procedure.

For the classification task, LS was split into training and validation samples with a proportion of 80% to 20%, respectively. An ensemble strategy was developed: initially, a multi-layer perceptron classifier (MLPc) algorithm and an extreme gradient boosting (XGB) algorithm for random forests predicted the failure probabilities of each insulating system; afterward, based on the probabilities resulting from both algorithms, a ridge regression classifier (RRc) algorithm was trained to produce the final result, i.e., whether the system would fail or not.

For the regression task, in turn, only the systems that did not fail in the classification task were considered. Those were also split into training and validation samples with the same proportion of 80% and 20%, respectively. A single multi-layer perceptron regressor (MLPr) algorithm was used to predict the external temperature at the final moment of the heating curve.

Various strategies were previously tested to select this set of algorithms comprising the ML framework. Each one of the selected models was configured using a set of empirical hyperparameters obtained from a grid search strategy for this specific case study. The main criterion to choose the model framework and its respective hyperparameters was based on the accuracy of the training set in the classification task, which was complemented by both assessment criteria, i.e., the root mean-squared error (RMSE) and the mean-absolute error (MAE) of the training set, to support the regression task.

III. RESULTS AND DISCUSSION

A. Predictive analysis

As described in Sec. II, in order to illustrate the ML framework operation, the thermal performance of around 2.8% of all likely insulating systems (within the initial list of candidate products) was calculated by a FE model, which had been previously validated with experimental data.[7] Based on these initial results, the ML framework was applied to estimate the external temperature of unseen lining combinations, capable of carrying out assessments in a much shorter time.

1. Failure prediction

For the classification task of failure prediction, the procedure consisted of an ensemble method which predicted the failure probability using the following parameterization: (i) initially, the lining failure probability was estimated by two distinct classifiers: an MLPc with back-propagation error, 23 nodes at the hidden layer using the Adam solver (stochastic optimization) and the Relu activation function, and an XGB with 1662 trees along three layers and (ii) next, an RRc with an alpha parameter of 20 determined whether the systems would fail or not.

At the first step of failure probability prediction, 4430 examples were used as the training sample and 1108 as the validation one. At the second step, the former validating sample was split into two equally sized samples with 554 examples each, one for training and another for validating the RRc. Table I lists the classification prediction accuracies for the validation examples of the ensemble method. The proposed procedure was capable of predicting (with nearly 98.4% of accuracy) when a system would fail based on 554 unseen examples.

TABLE I. Confusion matrix for the classification task using the ML framework on the validation set. Accuracy based on 554 examples: 98.38%.

	Predicted		
Actual	False	True	Total
False	287	7	294
True	2	258	260
Total	289	265	554

In the validation sample, one can measure the quality of the model by quantifying the number of false positives (FPs), true positives (TPs), false negatives (FNs), and true negatives (TNs). The insulating systems which got correct predictions about failing or not during the heating procedure are represented, respectively, by TP and TN. In this context, there are two kinds of missclassifications that designers should care about: the FNs, which refer to systems that the FE analysis classified as failed, but the ML framework classified as not failed, and the FPs, where FE classified as not failed, but the ML framework classified as failed.

In order to prevent the effect of FNs, the designer should simulate the desired materials combinations to make sure they would not fail under the application procedure. FPs present a higher risk because systems that might be useful will not have their external temperatures estimated by the model, so designers could end up ignoring good insulating solutions.

Considering the validation sample, the model predicted that 265 insulating systems would fail (True). However, seven of these would actually withstand the heating. By the formula %FP = FP/(FP + TP), this makes a total of 2.64% of FPs, thus resulting in a minor and acceptable risk.

2. External temperature prediction

After carrying out the classification task, the next step is to proceed with the regression one. From the initial, 5538 lining systems

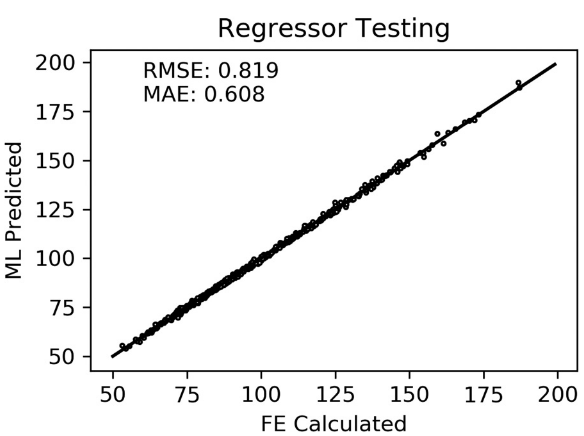

FIG. 1. Comparison between the external temperatures according to the FE calculation and to the ML framework on a given sample. Root mean-squared errors (RMSEs) and mean-absolute errors (MAEs) were taken into account.

comprising the LS, only 2916 configurations did not fail in the FE analysis. They were split into 2332 examples for training and 584 for validating the model consisting of an MLPr with back-propagated errors, 24 hidden nodes, using the limited-memory Broyden–Fletcher–Goldfarb–Shano solver and the logistic activation function. Figure 1 compares the machine learning estimation of the external temperatures and the respective FE results for the same sample.

The MLPr errors were very small and reproduced very well according to the FE calculations with less than 1 °C difference on average. Hence, the proposed machine learning model shows evidence of being a fine tool to estimate the furnace linings' external temperatures. However, a more robust proof of concept should be evaluated in order to draw conclusions about the proposed methodology.

3. Framework test on large unseen set of examples

In order to test the ML framework generalization, i.e., its capacity of predicting the thermal performance of unseen combinations of materials linings, the same learning procedure of Secs. II B–II D was repeated on samples from the whole LS and then applied to predict the failure and estimate the external temperatures of a much larger data pool. This test set consisted of 17 924 unseen random insulating systems, corresponding to nearly 9.4% of all combinations.

As shown in Table II, the classification task was concluded with 97.08% accuracy in failure prediction such that it has a generalization error of 1.38% when compared to the validation sample. It means that the precision obtained when training and validating are expected to be reduced by nearly 1.38% when applying this model in large sets of unseen examples. Such a small error makes the validation precision of the classification task trustworthy. The percentage of FPs was 3.34%, which could still be considered a small risk of information loss for the proposed ML framework.

Next, 9065 systems classified as not failing had their external temperatures evaluated at the regression task. Figure 2 shows the regression errors for the testing set, which remained significantly low. When compared to the validation sample, the test RMSE was only 0.003 greater and the MAE was even 0.04 smaller. These results show that the generalization error is almost null such that the precision obtained during training and validating the regression model is highly trustworthy and that the whole framework generalized well.

B. Insulating systems' multi-objective optimization

A multi-objective task involving the thermal performance of refractory insulating materials becomes easily solvable by using the validated ML model, no matter how large the dataset is.

TABLE II. Confusion matrix for the classification task using the ML framework on the testing set. Accuracy computed on 17 924 examples: 97.08%.

	Predicted		
Actual	False	True	Total
False	9065	288	9353
True	235	8336	8571
Total	9300	8624	17924

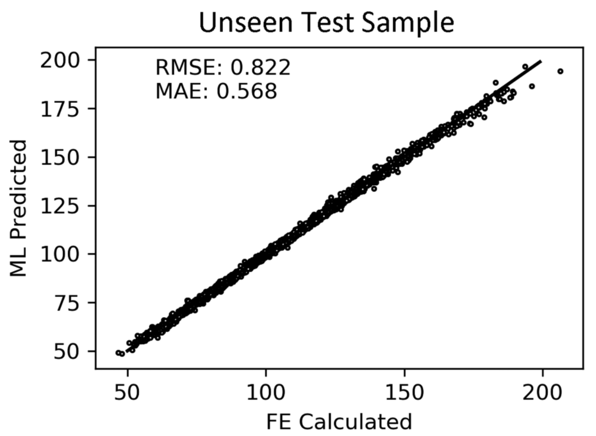

FIG. 2. Comparison between the external temperatures provided by the FE calculation and by the ML framework using the testing set. Root mean-squared errors (RMSEs) and mean-absolute errors (MAEs) were taken into account.

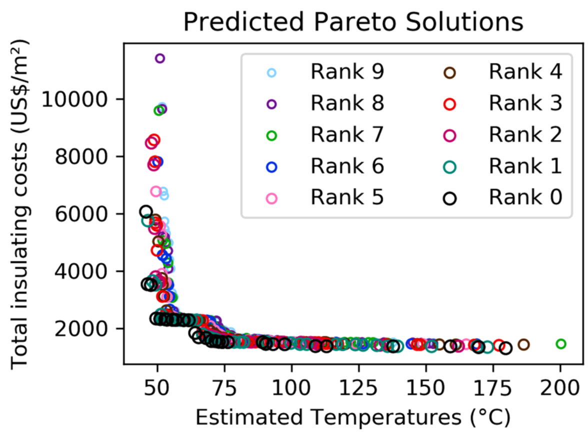

FIG. 4. Trade-off of insulating systems considering the thermal performance and total costs for the ten best Pareto ranks.

As an example, in order to check the trade-off between thermal performance and total costs for each one of the 190 365 likely insulating systems comprising the case study, price data for the candidate insulating products were considered as an additional system attribute, whereas the final external temperature was estimated by the proposed framework. Figure 3 shows a two-dimensional histogram that highlights the frequency distribution of the systems in terms of both criteria.

Because there are many possibilities, it is necessary to filter only the most potential insulating systems considering the objective of reducing both costs and external temperatures. For that, the dominance concept was applied to draw Pareto curves to analyze the whole set of feasible solutions into hierarchical ranks such that lower ranks contained better solutions in terms of costs and benefits than the higher ones. This means that for each solution at some rank k, there is at least one solution at rank $k-1$ that holds any of the following: it is cheaper for the same thermal insulating performance, it is a better insulator for the same price, or it is better in both objectives. Figure 4 shows the solutions belonging to the 10 best Pareto ranks, and Fig. 5 shows the Pareto fronts for each one of those ranks.

These results indicate that the proposed ML framework is useful to carry out multi-objective optimization in scenarios with a great number of options. While the FE simulation of all likely outcomes would take weeks or even months, using the surrogate model it took less than 10 min to generate the bi-objective trade-off in an Intel Core i3 with 8 GB of RAM memory.

When comparing the ML framework results to the Evolutionary Screening Procedure (ESP),[7] where multi-objective GA was used in order to make the ERF optimizations, certain points should be highlighted. (i) Using ML requires an initial

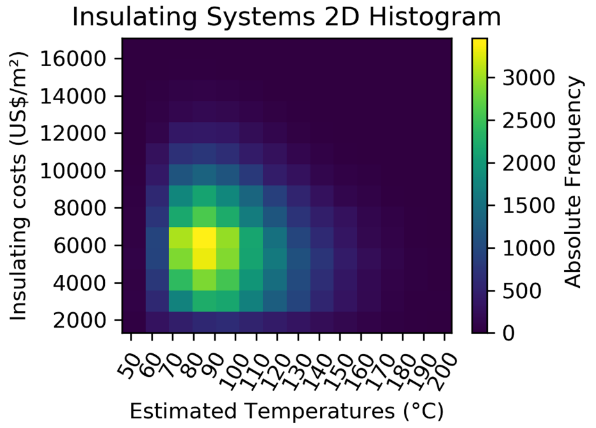

FIG. 3. Frequency distribution of the whole set of likely insulating systems in terms of external temperature and materials' costs.

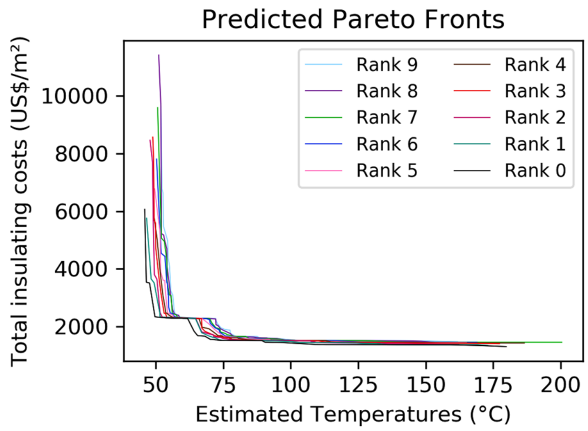

FIG. 5. Approximated Pareto fronts for the ten best Pareto ranks.

balanced dataset with calculated outputs so the model could learn the non-linear relations among variables such that the size of this initial set is set up by the designers, whereas the ESP strategy will simulate as many insulating systems as necessary to meet the established stop criteria. Hence, the former provides more control over the number of FE simulations and the latter does not demand such a decision. (ii) ML is somewhat more deterministic because its accuracy measure does not vary much when using the same input dataset such that its hyperparameters could be optimized in order to improve predictability during training. On the other hand, ESP presents substantial variations within the same hyperparameters, making their convergence to optimal values difficult for specific cases. (iii) ML procedure enables the designers to estimate the thermal performance of the whole search space of possibilities with small regression errors, except for the FPs eliminated in the classification task. The ESP mechanism, in turn, provides only the thermal performance of the best trade-off solutions found along iterations, but in most cases, it misses some of the Pareto optimal solutions.

C. Prediction performance dependence on the LS size

In order to better understand the impact of the initial learning set (LS) size on the performance of the proposed ML framework, new sets were generated, with sizes equal to 0.96%, 1.93%, 2.90%, 3.87%, 4.84%, 5.80%, 6.78%, and 7.75%, sampled from a uniform probability distribution applied on the original dataset. The training steps described in Sec. II were repeated seven times for each LS and tested over 17 924 unseen examples. The classification accuracy and both the percentages of true positives (%TP = 1−%FP) and true negatives (%TN = 1−%FN) are shown in Fig. 6 while the regression errors are depicted in Fig. 7, all of them as functions of the initial sample size in their respective LS.

For the ERF case study, the failure prediction accuracy rate grows when increasing the initial sample size from 0 to values around 5%, after which the accuracy seems to stabilize at a plateau.

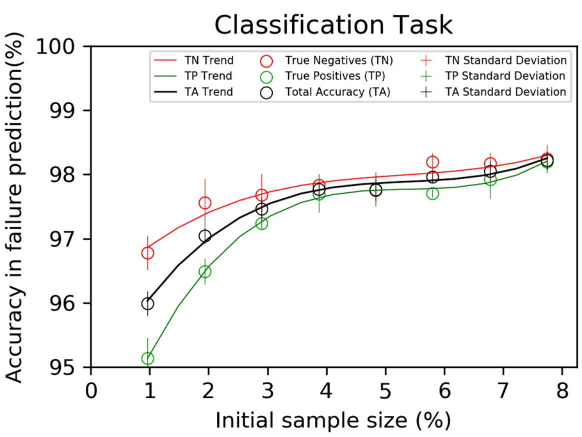

FIG. 6. Impact of the initial LS size on the classification task accuracy when applied to test sets with unseen data (each corresponding to nearly 9.4% of all combinations).

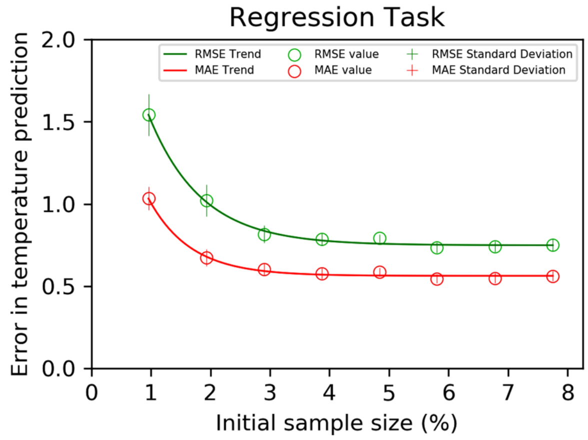

FIG. 7. Impact of the initial LS size on the regression task errors (RMSE and MAE) when applied to test sets with unseen data (each corresponding to nearly 9.4% of all combinations).

Between 7% and 8%, the accuracy begins to rise slightly. However, there seems to be an inherent misclassification error of at least 2% when using initial sample sizes in the range of 3%–7%.

As discussed before, part of this error is related to FPs, which, in other words, represent systems that might withstand the temperature conditions but were discarded by the procedure. In this range, the percentage of FPs is always greater than the percentage of FNs, and the difference gets bigger as the initial sample size gets smaller than 3%. This makes the classification task the major drawback for the proposed ML framework.

The regression task, in turn, has a great performance for all initial sample sizes of at least 3%, presenting an average RMSE smaller than 1. Still, it does not improve much for larger LS, remaining approximately constant for both RMSE and MAE values.

Hence, furnace designers would be advised to choose an initial sample size in the range of nearly 3% up to around 5% of total combinations to guarantee reasonable performance predictability. Their final decision should take into account the balance between the risk of losing information with FPs and the available amount of time for carrying out FE simulations.

IV. CONCLUSIONS AND SUMMARY

In this paper, an ML framework was proposed to predict possible thermal failure and estimate the thermal performance of a large number of multi-component insulating systems as a surrogate model for finite elements (FEs). By initially simulating only nearly 2.8% of the likely electric resistance furnace (ERF) configurations for the training set, the proposed framework had a generalization error of 1.38% in failure prediction and almost 0 during thermal performance prediction, with a precision of 97.08% at the first step and a RMSE of 0.822 at the second one. Hence, it was possible to estimate the temperatures of all the other systems with high precision in a shorter time than it took to simulate the training set. It was also shown that with this surrogate model strategy, any multi-objective optimization of furnace linings considering thermal performance as a criterion could have its time demand considerably reduced, regardless of the complexity of the FE model or the number of insulating systems to be evaluated.

A major drawback inherent to the proposed procedure is the misclassification in the failure prediction stage, mainly with false positives (FPs) which are discarded and, therefore, are not considered in multi-objective trade-offs. For the specific case of the electric resistance furnace, it was found out that the optimal range of the LS size is between ≈3% and ≈5% of uniformly sampled examples, in which the designers should consider the balance between the available time for carrying out the FE analysis and the acceptable risk of reducing the failure classification accuracy (in the range of 97%–98%).

The proposed ML framework was compared to a genetic algorithm procedure for electric resistance furnace lining optimization, from which it was concluded that: (i) the present proposal is more predictable in terms of computational costs, as it allows furnace designers to choose the number of FE simulations *a priori*; (ii) the hyperparameters from the ML framework models could be optimized without requiring new FE simulations; and (iii) unlike the GA method, the proposed framework could be used to evaluate the thermal performance of the entire search space of insulating systems.

ACKNOWLEDGMENTS

This study was financed and/or supported by CAPES—Coordination of Superior Level Staff Improvement (Brasil—Finance Code 001), by CNPq—National Council for Scientific and Technological Development (Grant No. 169129/2017-9), FAPESP—Sao Paulo Research Foundation (Grant No. 2017/16044-8) and F.I.R.E—Federation for International Refractory Research and Education. The authors gratefully acknowledge Dr. V. Salvini for the electric resistance furnace experiments and data provided.

DATA AVAILABILITY

The data that support the findings of this study are available within the Appendix of the article "Materials selection of furnace linings with multi-component refractory ceramics based on an evolutionary screening procedure," Ref. 7.

REFERENCES

[1] A. Bahadori, *Thermal Insulation Handbook for the Oil, Gas, and Petrochemical Industries*, 1st ed. (Gulf Professional Publishing, 2014).

[2] D. Mansuy and C. Gehin, "Thermal aspects of the application of refractories," Metall. Res. Technol. **80**, 935–945 (1983).

[3] M. Ashby, Y. Bréchet, D. Cebon, and L. Salvo, "Selection strategies for materials and processes," Mater. Des. **25**, 51–67 (2004).

[4] C. Demuth, J. Hubálková, M. Mendes, F. Ballani, D. Trimis, and S. Ray, "Prediction of effective thermal conductivity of refractory materials at high temperatures based on synthetic geometry generation," J. Ceram. Sci. Technol. **7**, 183–192 (2016).

[5] F. Roters, P. Eisenlohr, L. Hantcherli, D. Tjahjanto, T. Bieler, and D. Raabe, "Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications," Acta. Mater. **58**, 1152–1211 (2010).

[6] R. Quiza, O. López-Armas, and J. Davim, "Finite element in manufacturing processes," in *Springerbriefs in applied sciences and technology* (Springer-Verlag, Berlin, Heidelberg, 2012), Chap. 2, pp. 13–37.

[7] D. P. Santos, P. I. B. G. B. Pelissari, B. S. de Oliveira, D. R. Leiva, R. F. de Mello, and V. C. Pandolfelli, "Materials selection of furnace linings with multi-component refractory ceramics based on an evolutionary screening procedure," Ceram. Int. **46**, 4113–4125 (2020).

[8] S. Datta, *Materials Design Using Computational Intelligence Techniques* (CRC Press and Taylor & Francis Group, 2016), pp. 1–158.

[9] S. Datta and J. P. Davim, *Computational Approaches to Materials Design: Theoretical and Practical Aspects*, 1st ed. (IGI Global, Hershey, PA, 2016).

[10] *Optimization in Industry: Present Practices and Future Scopes*, 1st ed., edited by S. Datta and J. P. Davim (Springer International Publishing, 2019).

[11] F. Giudice, G. Fargione, R. Caponetto, and G. L. Rosa, "Modeling and optimization of multi-component materials selection and sizing problem," in *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications* (Sage Publishing, 2019), pp. 1–19.

[12] C. M. Childs and N. R. Washburn, "Embedding domain knowledge for machine learning of complex material systems," MRS Commun. **9**, 806–820 (2019).

[13] H. Liu, Z. Fu, K. Yang, X. Xu, and M. Bauchy, "Machine learning for glass science and engineering: A review," J. Non-Cryst. Solids **4**, 100036 (2019).

[14] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," Npj Comput. Mater. **5**, 83 (2019).

[15] R. K. Vasudevan, K. Choudhary, A. Mehta, R. Smith, G. Kusne, F. Tavazza, L. Vlcek, M. Ziatdinov, S. V. Kalinin, and J. Hattrick-Simpers, "Materials science in the artificial intelligence age: High-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics," MRS Commun. **9**, 821–838 (2019).

[16] R. F. de Mello and M. A. Ponti, *Machine Learning: A Practical Approach on the Statistical Learning Theory* (Springer, 2018).

[17] R. F. de Mello, M. D. Ferreira, and M. A. Ponti, "Providing theoretical learning guarantees to deep learning networks," arXiv:1711.10292 (2017).

[18] U. von Luxburg and B. Schölkopf, "Statistical learning theory: Models, concepts, and results," in *Handbook of the History of Logic* (North Holland, 2011), pp. 651–706.

[19] V. N. Vapnik, *The Nature of Statistical Learning Theory* (Springer, 2000).

[20] S. Luke, *Essentials of Metaheuristics*, 2nd ed. (Lulu, 2013), see http://cs.gmu.edu/sean/book/metaheuristics/.