ARL-TR-10145 • AUG 2025

Machine Learning Framework Development for Rapid Input Space Mapping for Laser Powder Bed Fusion

by Alexander Butler and Brandon McWilliams

ARL-TR-10145 ● AUG 2025

Machine Learning Framework Development for Rapid Input Space Mapping for Laser Powder Bed Fusion

Alexander Butler and Brandon McWilliams
DEVCOM Army Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED START DATE	3. DATES COVERED END DATE
August 2025	Technical Report	08/2023	01/2025

4. TITLE AND SUBTITLE
Machine Learning Framework Development for Rapid Input Space Mapping for Laser Powder Bed Fusion

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER	5e. TASK NUMBER	5f. WORK UNIT NUMBER

6. AUTHOR(S)
Alexander Butler and Brandon McWilliams

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
DEVCOM Army Research Laboratory ATTN: FCDD-RLA-MD Aberdeen Proving Ground, MD 21005	ARL-TR-10145

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

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Process optimization, alloy development, and other research efforts all possess broad input spaces that take considerable effort to characterize and understand. These lengthy and costly efforts hinder progress and knowledge application. ARL researchers developed an ML framework that aims to reduce the characterization and analytical burden of mapping a broad input space. This framework uses a greedy sampling algorithm and Latin hypercube sampling to assist with initial design of experiments and sample selection. The initial dataset is then prepared and used to train various ML models, using a hyperparameter grid search and k-fold cross validation. Using the trained models, process maps and correlation coefficients can be generated to better understand the given input space. The model can be queried to provide a predicted output for a given input combination or, conversely, provide a set of input combinations that would likely result in a given output. This data-driven framework is process and material agnostic, making it ideal for rapid implementation and transfer learning. The development and implementation of this framework is described for the laser powder bed fusion metal additive manufacturing process of aluminum alloy (AlSi10Mg).

15. SUBJECT TERMS
Sciences of Extreme Materials, laser powder bed fusion, process optimization, ML predictions, aluminum alloys, additive manufacturing

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	C. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UU	39

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Alexander Butler	(520) 691-5564

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

Contents

List of Figures

List of Tables

Acknowledgments

We thank Drs. Berend C. Rinderspacher (ARL) and Matthew Vaughn (Oakridge Associated Universities) for assistance in ML implementation and for their efforts in making this work possible.

1. Introduction

Metal laser powder bed fusion (LPBF) is an additive manufacturing (AM) process in which a layer of metal powder, of a set thickness, is evenly distributed across a build plate. A laser then selectively exposes the powder with a prescribed power, moving across the build plate at a set speed, with a set distance between previous scan lines. The build plate then lowers, another layer of powder is spread, and the laser exposes the next slice.[1] This iterative process yields a part surrounded in a bed of loose powder. The density and performance of the printed material are heavily dependent on the process parameters used.

The combinations of these process parameters yield a finite amount of energy input into the powder bed. This energy input to the system can be simplified by calculating the volumetric energy density, E_v[2] (in J/mm^3):

$$E_v = \frac{P}{v * h * t}, \tag{1}$$

where P is the laser power, v is the laser scan speed, h is the distance between scan lines (defined as hatch spacing), and t is the layer thickness.[2] E_v is an easy-to-track, widely used metric when conducting process optimization.

Optimizing process parameters for an AM process, and therefore obtaining an ideal energy density, is critical for mitigating defects. Material manufactured with process parameters with an E_v that is too high can cause material to vaporize from the molten material, which causes keyhole porosity.[3] Conversely, process parameters with an E_v that is too low can also produce porosity. Powder exposed using process parameters with a very low E_v can yield balling. An increase in E_v from the balling regime can produce lack of fusion (LOF) porosity, which is characterized by incomplete melting.[4] Finding process parameters that fall between these defect-forming regimes can help define your optimized processing window.[5]

Even when an ideal operating window, or range of E_v values, has been determined, it does not necessarily mean that all material manufactured within the optimal operating window will perform the same. Different combinations of process parameters, all yielding the same E_v, can produce different microstructures, thermal histories, and mechanical performances.[2] This means that although E_v is a useful metric to simplified process parameter combinations, it does not tell the whole story.

Research is being done to map processing parameters to structure and performance, but this is a time-consuming and restrictive process. Considerable effort is being done to map power and laser scanning speed to structure and performance, but many of these efforts do not span the entire input space possible or consider the influence of other processing parameters.[6] The gaps left in the parameter search can hold process parameter combinations that are more ideal than those investigated, but fully defining this space is almost impossible.

In addition, variables other than laser parameters or layer thickness can influence defect formation and the performance of the material. The direction of the powder distribution, argon gas flow, and spatial dependency all have an influence on the manufactured material. Moreover, this influence is not consistent across all LPBF systems because most manufacturers have different powder distribution systems, gas flow manifolds, and other nuances, which makes implementing corrections and tracking its contribution more difficult.[7] In situ monitoring has the potential to provide additional information from the build process that is not necessarily captured by processing parameters. However, this data comprises varying quality and resolution, and this usually generates a lot of data that is difficult to process and understand. For these reasons, many commercial in situ monitoring products implement ML to process the data. However, the processing of this data is usually done in a "black box" system, again, calling into question the quality and usefulness of the data.

ML can assist researchers with metal AM efforts; however, there are some barriers to progress. First, there is a knowledge gap between those familiar with the metal AM processes and the ML algorithms used. Trying to get these researchers to "speak the same language" can be difficult. Second, ML requires ample data. This becomes an issue with metal AM efforts, because manufacturing, post-processing, and characterizing data can be time consuming. Third, as the number of input variables increases, so does the required amount of data. This means that, if researchers want to incorporate all relevant process parameters and conditions, then the logistical burden is unreasonably large.

Many data sets and models are for a single material that is manufactured on a single machine. This opens the door for transfer learning; however, data sharing between entities is limited, with varying success, because data streams can have different qualities and formats, and many industry partners are not willing to share proprietary information.[8] Finally, there is a converse relationship between interpretability of a model and its accuracy. A model that easily conveys the relationship between input and output features may not be as accurate as a more complex and difficult-to-implement model.[9] This works aims to develop a framework for ML implementation, generalized to any research effort, with a low barrier of entry regarding the data required and ML expertise.

2. Experimental

To develop the ML framework, process optimization was conducted for LPBF of AlSi10Mg, with different target variables. Two cases were considered for this effort: porosity and tensile properties. These were treated independently because predicting both structure and performance output spaces could over generalize the models, especially with limited data sets, leading to high rates of prediction errors. The porosity case was investigated first, because a 1-D output space makes it easier to evaluate predictions. Preliminary success from the first case showed promise and tensile properties were investigated.

2.1 Sample Selection

2.1.1 Latin Hypercube Sampling (LHS)

An input space was defined for each case, defining bounds for the process parameters selected, from which input combinations were found. For most input features, the machine limits were used to set the minimum and maximum values, whereas the machine resolution defined the resolution of the generated input space. This allows almost every possible combination of the input space to be considered. Tables 1 and 2 outline the parameter grids.

Table 1. Input space for metallography cylinders.

Type	Name	Unit	Abbreviation	Min	Max	Resolution	Relation
Design	Laser power	W	P	50	370	1	. . .
Design	Hatch spacing	µm	h	30	400	5	. . .
Design	Laser speed	m/s	v	0.2	7	0.05	. . .
Design	Layer thickness	mm	t	0.03	0.08	0.01	. . .
Design	X position	mm	xPos	20	230	20	. . .
Design	Y position	mm	yPos	20	230	20	. . .
Calculated	Energy density	J/mm^3	E_v	12.5	110	1	$P/(h*t*v)$

Note: Max = maximum; Min = minimum

Table 2. Input space for tensile bars.

Type	Name	Unit	Abbreviation	Min	Max	Resolution	Relation
Design	Laser power	W	P	75	370	1	. . .
Design	Hatch spacing	µm	h	30	375	5	. . .
Design	Laser speed	m/s	v	0.35	6.5	0.05	. . .
Design	Layer thickness	mm	t	0.03	0.08	0.01	. . .
Design	X position	mm	xPos	20	230	20	. . .
Design	Y position	mm	yPos	20	230	20	. . .
Calculated	Energy density	J/mm^3	E_v	12.5	110	1	$P/(h*t*v)$

Note: Max = maximum; Min = minimum

Input variables can be designated as design, fixed, or calculated. The design variables are varied within the range provided. The fixed variables will not change between input combinations. Finally, the calculated variables are dependent on other input features, and the relation provided calculates a value for each input combination. A range of acceptable values for this calculated variable can be provided to restrict input combinations to a desired window of interest, without restricting any range of input variables. The name and unit for each input variable are provided to format the data. In addition, an abbreviation for each input variable is provided to enable the use of the relationships listed for calculated variables. For this case, the calculated variable E_v was used to restrict input combinations, in an attempt to restrict any input combinations that were too far outside the expected processing window. The resultant samples would most likely be unable to retain its integrity and threaten to fail each build. The bounds for the energy density criterion were determined based on the energy density of the original equipment manufacturer (OEM) process parameters and researchers' experience with previous optimization. After manufacturing the samples to evaluate porosity, it was determined that the ranges for some input features were too large, yielding some samples that threatened to fail builds. Some modifications were made to the definition of the input space for the tensile bars.

Typically, most metal AM process optimization and mapping only look at the power and scan velocity input space. To more holistically map the input space, laser power, hatch spacing, laser speed, and layer thickness were all considered. To account for any influence of build plate position on structure or performance of the samples, build plate X and Y positions were provided by the sampling generation. During the generation of parameter sets, no restrictions were set on repeating X and Y positions. In these cases, one sample was moved 20 mm in the positive X direction to avoid overlapping the samples. If another sample occupied this position, the sample was moved 20 mm in the negative Y direction instead. The data was edited to reflect the change in position.

LHS is a type of stratified sampling in which each input feature is broken up into subspaces. Within the intersection of multiple input feature's subspace, a point is chosen at random. This allows for an element of randomness while ensuring the whole range of a given input feature is captured.[10] The total number of samples generated by the LHS is dependent on how many samples will be down selected and tested. The goal is to have an excess of samples for the greedy sampling algorithm from which to choose. For both the porosity and tensile samples, an oversampling factor of 50 was used, which generated 4150 and 2400 parameter sets, respectively. These generated parameter sets were then filtered by the calculated energy density criterion, leaving only what was believed to be “printable” parameter sets that will not fail during the build process. The contour parameters—that is, the parameters that border the infill of the material—were set to the OEM-recommended parameters and were not changed for any of the samples. This was done to ensure that samples containing high porosity would maintain some integrity and prevent build failures.

2.1.2 Greedy Sampling Algorithm

Once the LHS script generates the parameter sets, a greedy sampling algorithm can be implemented to select the most dissimilar point, distributing selected data points efficiently. First, legacy data from previous efforts can be provided as seed points. It is not necessary to include legacy data, but it ensures points will only be selected in unknown regions of the input space. The normal distance is found between each legacy seed point and the generated point across all input variables.

With k seed points, the normal distance from a new input combination x_n to the seed points for a collection of *N* samples is as follows:

$$d_{nm}^{x} = \|x_n - x_m\|, m = 1, \dots, k; n = k + 1, \dots, N \ . \tag{2}$$

The minimum distance from a new input combination x_n is found for all previously selected points *k*:

$$d_n^x = \min_m d_{nm}^x, n = k + 1, \dots, N \ . \tag{3}$$

The input combinations with the largest d_n^x are selected as a point of interest. If no seed points are provided, the distance between every input combination, to every other input combination, is found; the input combination with the greatest distance, d_n^x, where $n = 1, \dots, N$, is then selected as the first seed point.[11] Eighty-three and forty-eight parameter sets were selected for the case of porosity and tensile properties, respectively, manufactured between two and four builds.

2.2 Data Collection

2.2.1 Feedstock and Manufacturing

Aluminum alloy, AlSi10Mg, samples were manufactured on the EOS M290, using reused EOS powder, sieved using a 90-µm mesh screen. The composition provided by EOS and size distribution measured via dynamic image analysis on the CAMAIZER X2 are shown in Table 3 and Figure 1, respectively.

Table 3. Elemental composition of AlSi10Mg powder.

Composition weight percentage of AlSi10Mg powder										
Si	**Fe**	**Cu**	**Mn**	**Mg**	**Ni**	**Zn**	**Pd**	**Sn**	**Ti**	**Al**
10.0	≤0.55	≤0.05	≤0.45	0.33	≤0.05	≤0.10	≤0.05	≤0.05	≤0.15	Balance

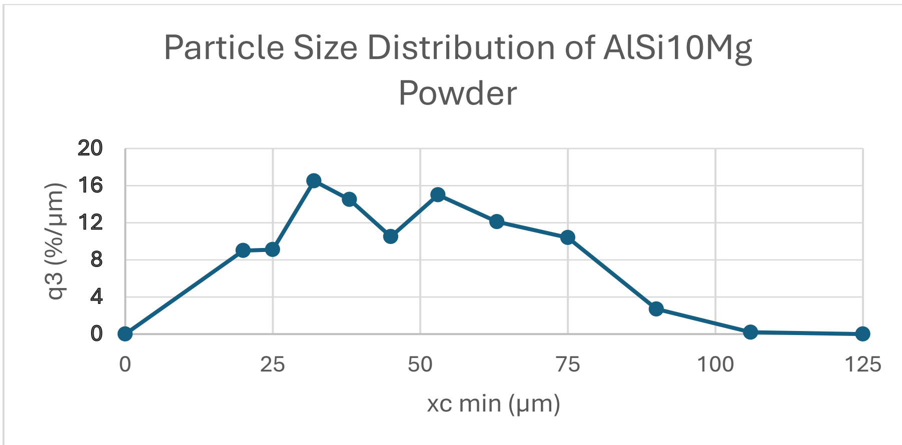

Figure 1. Sieved AlSi10Mg particle-size distribution.

Cylinders (15-mm-diameter and 25-mm-tall) were printed for metallography and porosity measurements, and 4-mm gauge subsize ASTM E8[12] tensile bars were printed for tensile testing. Both sample types were manufactured on support structures for easy removal from the build plate.

2.2.2 In situ Monitoring

The EOS M290 is equipped with two photodiodes: one coaxial and one off axis. These sensors record the intensity of light emitted during the build, as a function of position and time. A layer-wise average intensity was collected for each part during the builds. For the metallography cylinders, layers at the beginning of the build, which included support structures, were omitted as well as the last five layers, to remove any influence of the up-skin exposure parameters. For the tensile bars, the data was trimmed to only include layers within the gauge section of the tensile bar. This allows for a consistent cross section as a function of build height. For each sample, an average of all layer-wise averages was taken, creating a general intensity average over the course of the build, and was included as an input feature. Depending on the circumstances, this can be treated as an input or target variable, because the in situ data is a result of input parameters, but it can also reflect uncaptured data attributed to other factors. Here, parameter sets are generated without considering in situ data; then, after the samples are manufactured, the in situ data is collected and added to the data as an input feature.

2.2.3 Metallography and Image Analysis

The metallography cylinders were sectioned and polished to a 0.02-µm finish using colloidal silica. The samples were then imaged using the Keyence VHX7000 optical microscope. The images were cropped and converted to binary for image analysis using ImageJ. High amounts of porosity made polishing difficult, and a dilate and erosion step was used to reduce the influence of polishing artifacts on the binary images that would skew porosity calculations.

2.2.4 Mechanical Testing

The tensile samples were tested in the as-printed surface condition with no post build heat treatment. The gauge sections were painted with white and black spray paint to produce a speckled pattern. Following the procedures in ASTM E8/E8M,[12] the samples were tested on the Instron 5984 electromechanical load frame at a constant displacement rate of 1.2 mm/min. 3-D digital image correlation was used to obtain strain data. The ultimate tensile strength (UTS) and failure strain were calculated for each sample.

2.3 Regression Models Preparation and Setup

2.3.1 Data Preparation

Preparing the data for subsequent predictions is paramount for accurate, unbiased, and repeatable results.[13] In this case, all values of the input and output features are expected to be nonnegative. Therefore, to restrict all predictions to positive numbers, the first step is to perform a log transformation of the input and output spaces.

In the second step, it is important to standardize the data so that no input variable has a bias over another. The standardization, for a given input variable, adjusts the data so all the data points have a mean of zero and a standard deviation of one. The mean and variance for each input variable is stored to later transform predicted, normalized values to real values. For all input features (i), where x^n is the original value of an element and the standardized value is z_i^n ,

$$z_i^n = \frac{x_i^n - mean(x_i)}{standard\ deviation(x_i)} . \quad (4)$$

Next, stratification sampling was performed to produce more representative datasets and reduce any influence on the order of data. For example, if the data was ordered by increasing laser power, and the subsets of the data were selected by simply going down the list and portioning the data as such, there will be a drastic difference from the first to the last data subset. Stratification sampling ensures that the training and test sets are both representative of the input and output space.[14] The stratification sampling produced the training and test data sets as well as the training subsets used for k-fold cross validation.

The influence of the in situ monitoring data collected was unknown. Therefore, the training and optimization of the models were performed twice: once with an input data set excluding the in situ data and again including it. The error of each model was then compared to evaluate the usefulness of the in situ data collected in this fashion.

2.3.2 Models and Hyperparameters

Five regression models were used to predict the given output variable. A hyperparameter grid search was done for each model, across multiple hyperparameters. This ensures all models are optimized for the given case. Table 4 shows the models and hyperparameters. The exact values of each hyperparameter range and step are dependent on the case being predicted. The models were trained multiple times, expanding the search if the ideal hyperparameters were on the edge case, or increasing the resolution of the search by decreasing the step size. A step that is too small in the hyperparameters' grid search can lead to a local minimum in prediction errors, yielding an unoptimized model.

Table 4. Models with their respective hyperparameters.

Model	Hyperparameters
Ridge	Alpha fit_intercept solver
Gradient boosting	Alpha n_estimators max_depth
Random forest	n_estimators max_depth
Support vector	C kernel gamma
Neural net	Alpha activation solver

2.3.3 K-Fold Cross Validation

The number of folds used in cross validation is dependent on the number of data points. For many cases, tenfold is the standard practice. However, with a large input space and relatively few data points, the decision was made to use fivefold to ensure each fold could be representative of the whole dataset.

2.4 Model Predictions and Queries

The trained models with optimized hyperparameters were used to predict the training and test data, and mean squared errors (MSE) were found. In the case of a 1-D output space, like that of porosity, the model with the lowest MSE on the test data is the most accurate. In the case of an output space greater than 1 dimension, the MSE can be normalized between the models for a given target; the best performing model had a normalized MSE value of 0, and the worse performing model had a value of 1. Then, a weighted sum of the normalized MSE values were calculated to obtain a score for a given model across all target variables. The weighted sum allows the user to add priority to specific target variables of interest. The model with the lowest score is the ideal model to use. For a model *m* yielding an MSE for predicting a target *t*:

$$y_t^m = \frac{MSE_t^m - \min(MSE_t)}{\max(MSE_t) - \min(MSE_t)} . \tag{5}$$

For all normalized MSE (y_t^m) for a given model, with weights (w_t) for all targets (t), the score is as follows:

$$Score^m = \sum w_t y_t^m . \tag{6}$$

The best performing model with its optimized hyperparameters can then be used to predict the initial input space generated by the LHS script. The model can predict the outputs for the whole input space, producing predicted process maps, and can be used to calculate correlation coefficients that can be utilized to better understand the system. If this collection of predicted points is used, then the data can be queried for different information.

2.4.1 Best Case

To evaluate the model's ability to predict cases near the edge of the output space, the "best" cases were identified for both porosity and tensile property predictions. For porosity, the best result for a given parameter set was "no porosity" or a fully dense sample. Therefore, the samples predicted to yield the least amount of porosity were manufactured and characterized for evaluation. For tensile properties, the best result for a given parameter set is the strongest and highest elongation. Therefore, the parameter sets with high UTS and strain to failure were chosen. Because the tensile property case had multiple target variables, it makes it difficult to determine what is the "best"-performing parameter set. Therefore, a similar approach to determining the score of a model when predicting multiple target parameters was taken to evaluate the performance of the parameter sets. The predicted values across each target variable were normalized, then for a given process parameter set, taking a weighted sum and yielding a score for a given process parameter set. In this instance, because the goal is to maximize both target variables, the process parameters with the highest scores should be used.

2.4.2 Specific Case

To evaluate the model's ability to interpolate and predict target values within the bounds of the measured output space, arbitrary target values were set and process parameters that were predicted to perform similarly to those targets were selected. For the porosity cases, a target value of 20% porosity was selected. In addition, for tensile performance, a UTS of 350 MPa and failure strain of 0.03 were selected.

In both cases, the differences between the predicted and target values for all predicted elements within the initial LHS input space were found. For a 1-D output space, finding the parameter set that was predicted to perform closest to the target was simple: min (|predicted value − target value|).

For output spaces with more than one dimension, finding one parameter set closest to each of the target values is more difficult. For all output features, where t is the target value and x is the predicted value, the normalized difference is given below:

$$x_{diff}^{p} = t - x^{p} \tag{7}$$

$$z_{\mathrm{diff}}^{p} = \frac{x_{diff}^{p} - mean(x_{diff})}{standard\ deviation(x_{diff})}\,. \tag{8}$$

After a normalized distance is found for all target variables for a given parameter set with their associated predicted values, similar to producing the model score for multi-target predictions, a weighted sum can be taken. This yields a cumulative "score" of the tensile performance for each parameter set.

In the case of porosity, the 12 samples that were predicted to yield the least amount of porosity and the 6 samples that were the closest predictions to 20% porosity were selected. In the case of tensile properties, the 23 parameter sets with the highest cumulative score were selected. For targeting specific values, weights of 1 for both the UTS and failure strain were used to find the cumulative score for each parameter set, and the 12 parameter sets with the lowest values were selected. All samples were manufactured, characterized, and compared with the model predictions for validation.

3. Results

The findings presented below are the results of characterization efforts of both the porosity and tensile samples. In addition, the training and optimization of ML models, following data preparation steps, are described in Section 3.2. Finally, the methods of evaluating the performance of the models and their subsequent use for additional queries and validations are presented.

3.1 Characterization

The characterization of the porosity content and tensile performance is presented below. In summary, compared with the OEM-process parameters, more dense, stronger, and more ductile performance was achieved in the initial broad parameter search. In addition, during validation, in which the model was queried for the best-performing parameter sets, samples that achieved higher density than the OEM-process parameters were observed. However, during the validation step, the parameter sets tested could not achieve better tensile performance than the OEM-process parameters.

3.1.1 Porosity Analysis

An example of a build plate containing porosity samples is shown in Figure 2. Porosity cylinders were difficult to polish due to high porosity, which led to trapped liquid and particles, as well as material pullout due to weakly bonded material. The dilate and erosion steps on the binary image were needed to remove polishing artifacts. Various samples showed evidence of keyhole and LOF porosity, some of which were so severe that pores coalesced and created large, open cavities. The samples containing LOF porosity showed some order, highlighting that this type of porosity is highly driven heavily by geometric relationships between scan strategy and melt pool width and depth. Metrics for the resultant porosity are shown in Table 5, and micrographs displaying various amounts of porosity are shown in Figure 3.

Figure 2. Metallography cylinders as printed, before removal from support structures and build plate.

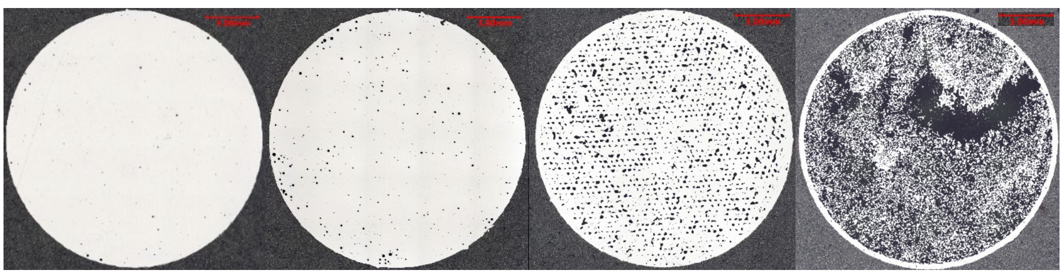

Figure 3. A subset of results from metallography of porosity cylinders.

Table 5. Porosity analysis results.

Statistical metric	Porosity (%)
OEM-optimized parameters	0.164
Minimum	0.093
Maximum	62.3
Average	17.3

3.1.2 Tensile Testing

Tensile testing produced a variety of mechanical responses, shown in Figure 4 and summarized in Table 6. Some samples performed like that of the OEM-optimized parameter sets. Some samples failed prematurely, resulting in large reductions in UTS and failure strain.

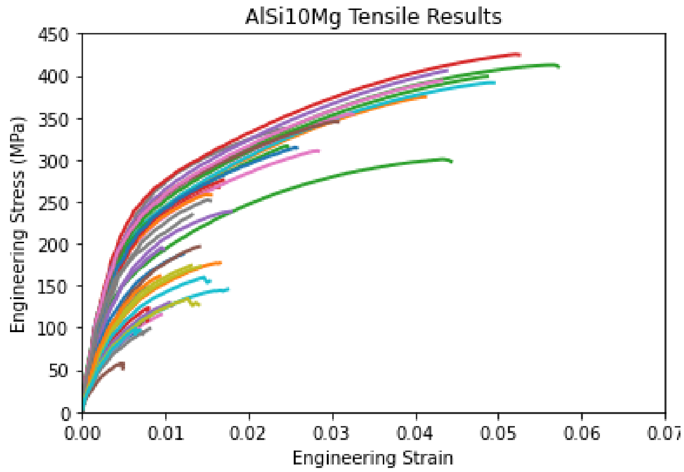

Figure 4. Results of tensile testing.

Table 6. Tensile testing results.

Statistical metric	UTS (MPa)	Failure strain
OEM-optimized parameters	411	0.061
Minimum	57.8	0.005
Maximum	436	0.098
Average	286	0.034

3.2 Model Training and Selection

3.2.1 Porosity Model

When predicting porosity, the models were trained and evaluated twice: once with a dataset that included in situ monitoring data and once omitting this data. The accuracy of the models is visualized in Figure 5 and detailed in Table 7. In most cases, by including the in situ monitoring data, the test MSE was appreciably reduced, with an average reduction of 17%. The neural net regression model, trained with data including the in situ data, yielded the lowest test MSE and had a test R^2 value of 0.958. These predictions are shown in Figure 6.

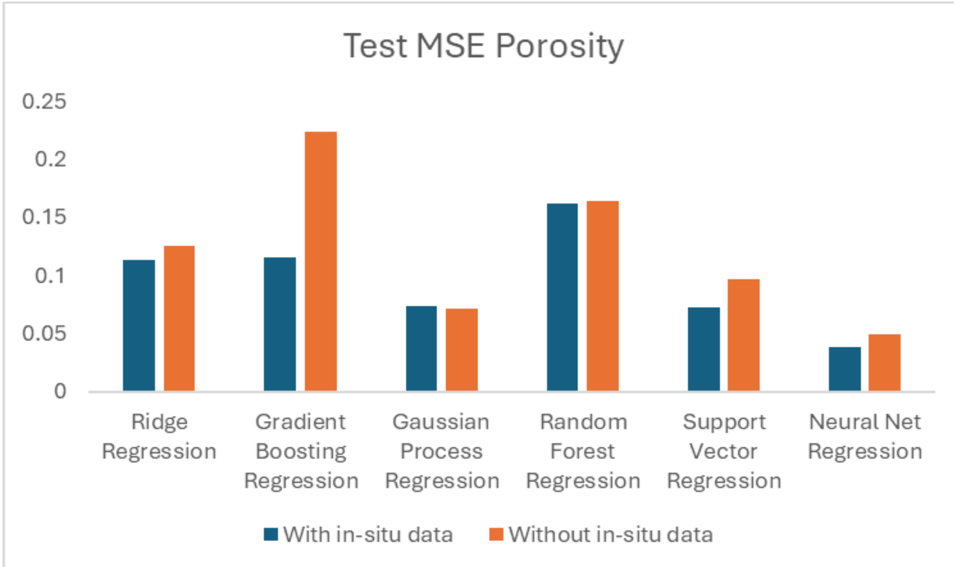

Figure 5. Training results for porosity predictions.

Table 7. Performance of models during training for porosity predictions.

Model	Test R^2		Test MSE		
	With in situ data	**Without in situ data**	**With in situ data**	**Without in situ data**	**%Change with in situ data**
Ridge regression	0.876	0.864	0.114	0.125	−9.1%
Gradient boosting Regression	0.874	0.756	0.116	0.224	−48.3%
Gaussian process Regression	0.920	0.756	0.074	0.072	2.2%
Random forest Regression	0.823	0.821	0.162	0.164	−1.3%
Support vector Regression	0.921	0.894	0.072	0.097	−25.6%
Neural net	0.958	0.946	0.039	0.050	−22.5%

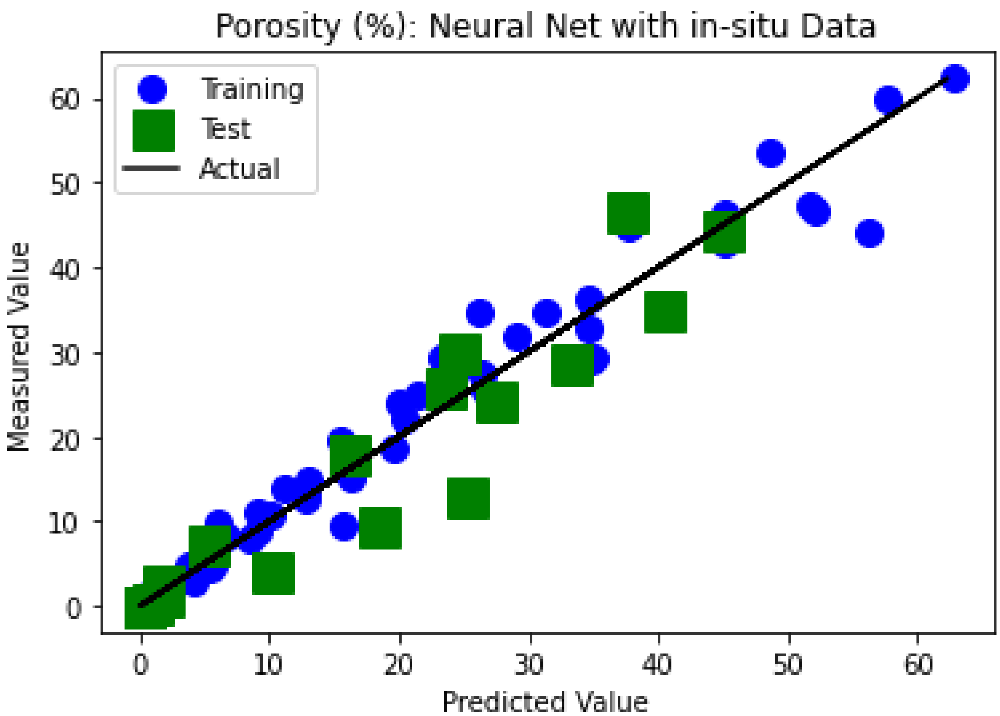

Figure 6. Results of best-performing model.

3.2.2 Tenile Properties Model

The performance of the models when predicting the UTS and failure strain are shown in Tables 8 and 9, respectively, with the weighted normalized scores shown in Table 10. In addition, the test MSE for each model and target are shown in Figures 7 and 8. When predicting the UTS and failure stain, the inclusion of in situ monitoring data (on average) decreased the test MSE by 9% and 15%, respectively. Using the method described previously, the test MSE values were normalized across each target; then, the weighted sums' normalized errors were found for each model. For this case, weights of 1 for each UTS and failure strain were used. As stated previously, if the use case calls for a particular emphasis on a certain target variable, the weights can be adjusted to ensure the model selected can accurately predict the emphasized target variables. In this case, the identification of the best performing model was easy, because the neural net regressor had the lowest test MSE for both target variables. The prediction of mechanical property targets was less accurate than the prediction of porosity content, and predictions of the failure strain were the least accurate.

In both the prediction of porosity and mechanical properties, the inclusion of in situ monitoring data increased the accuracy of predictions. However, for the prediction of the input space and queries, models trained without in situ data were used to generate desirable parameter sets from the result of the queries. Since the in situ data is an input variable, but it is collected after the build, it is not possible to fully utilize the models trained with in situ data, without first training a model to predict in situ results. To simplify this portion of the task, the models trained without in situ data was used. The models trained with in situ data and the subsequent collection of additional data are more likely to yield the most accurate results. A comparison of the measured and values predicted by the best-performing model are shown in Figures 9 and 10.

Table 8. Performance of models during training for UTS predictions.

UTS					
	Test R^2		Test MSE		
Model	**With in situ data**	**Without in situ data**	**With in situ data**	**Without in situ data**	**%Change with in situ data**
Ridge regression	0.719	0.689	0.249	0.276	−9.8%
Gradient boosting Regression	0.359	0.378	0.568	0.551	3.1%
Gaussian process Regression	0.722	0.672	0.247	0.290	−15.0%
Random forest Regression	0.261	0.216	0.655	0.695	−5.8%
Support vector Regression	0.503	0.508	0.440	0.436	1.0%
Neural net regression	0.825	0.758	0.155	0.214	−27.4%

Table 9. Performance of models during training for failure strain predictions.

Failure strain					
	Test R^2		Test MSE		
Model	**With in situ data**	**Without in situ data**	**With in situ data**	**Without in situ data**	**%Change with in situ data**
Ridge regression	0.584	0.584	0.365	0.364	0.1%
Gradient boosting Regression	0.474	0.352	0.461	0.568	−18.9%
Gaussian process Regression	0.583	0.581	0.365	0.367	−0.5%
Random forest regression	0.227	−0.050	0.677	0.920	−26.4%
Support vector regression	0.593	0.457	0.357	0.476	−25.1%
Neural net	0.601	0.515	0.350	0.425	−17.7%

Table 10. Weighted normalized score for each model.

Model	Weighted score
Ridge regression	0.233
Gradient boosting regression	1.166
Gaussian process regression	0.231
Random forest regression	2.000
Support vector regression	0.592
Neural net regression	0.000

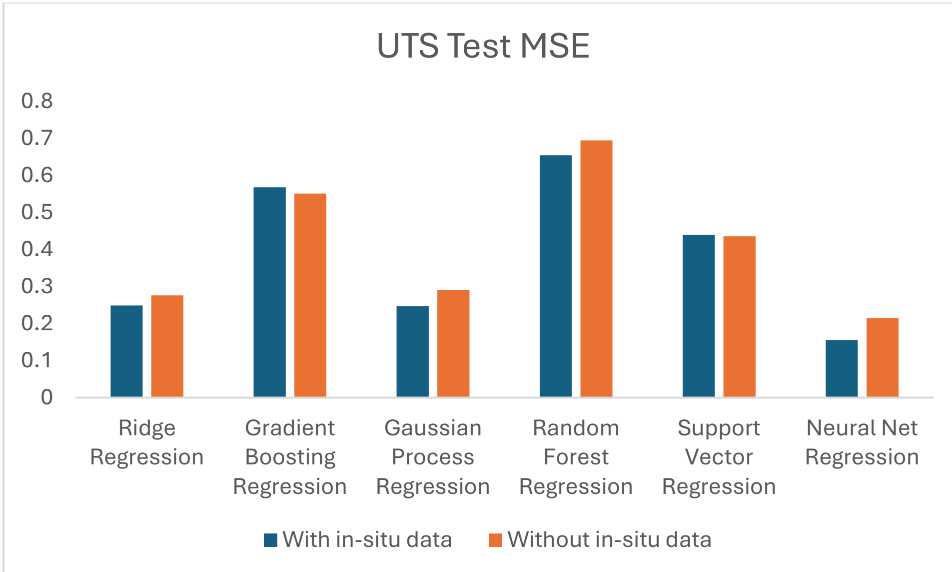

Figure 7. Training results for UTS predictions.

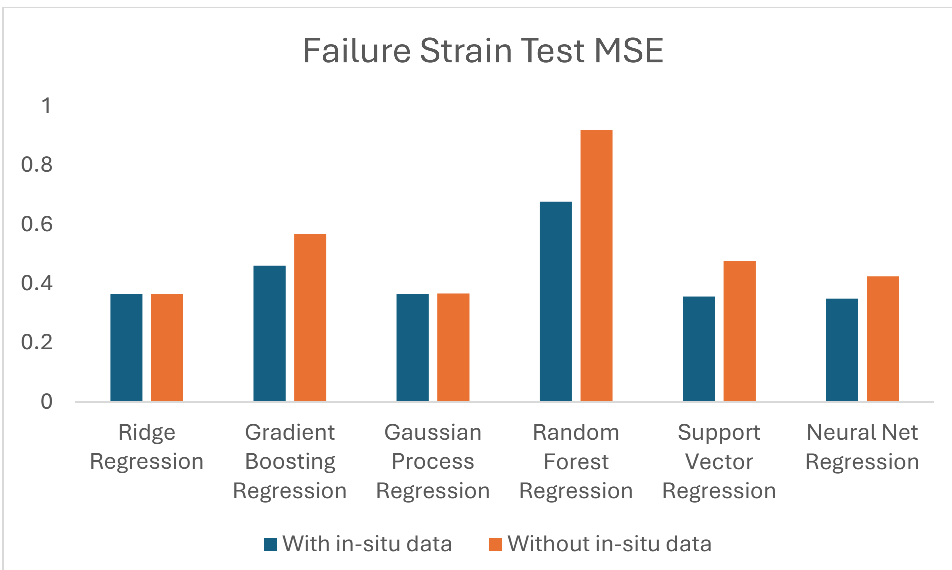

Figure 8. Training results for failure strain predictions.

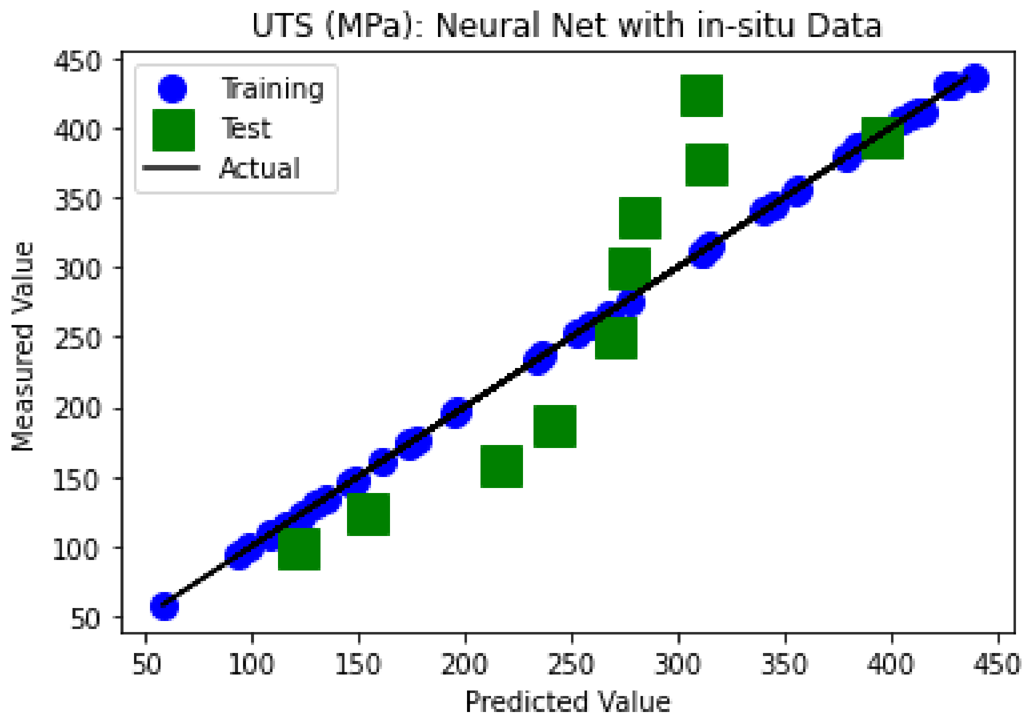

Figure 9. **Results of best-performing model when predicting UTS.**

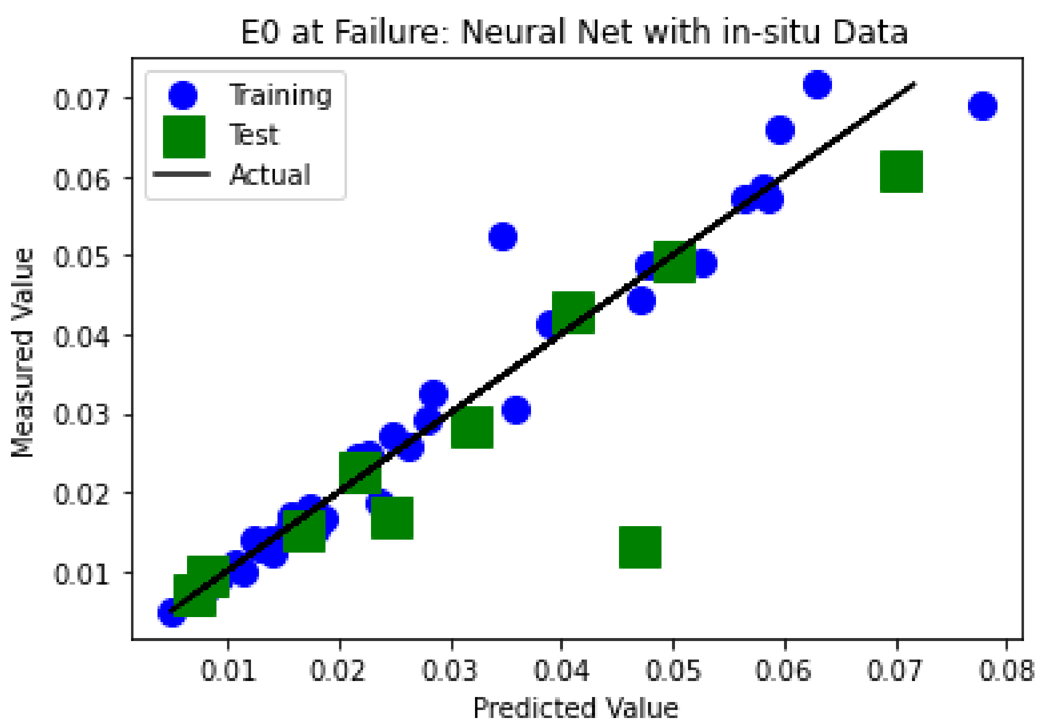

Figure 10. **Results of best-performing model when predicting strain at failure.**

3.3 Utilizing the Optimized Models

3.3.1 Process Mapping

The optimized neural net regression models were trained and used to predict the outcome of the oversampled input spaces initially generated in earlier steps by the LHS code. The process maps (shown in Figure 11) can be drawn between any two input variables for a single-target variable. In addition, correlation coefficients can be found for each input and target variable combination and used to create the heatmap in Figure 12, which can be used to better understand the relationships of this system.

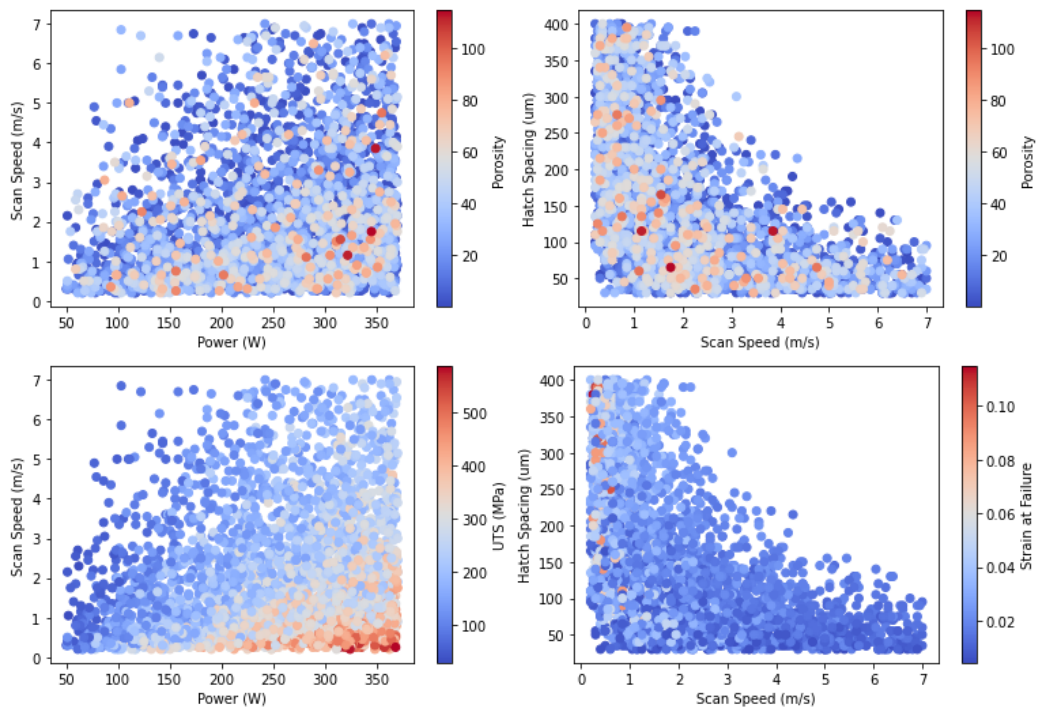

Figure 11. Generated process maps.

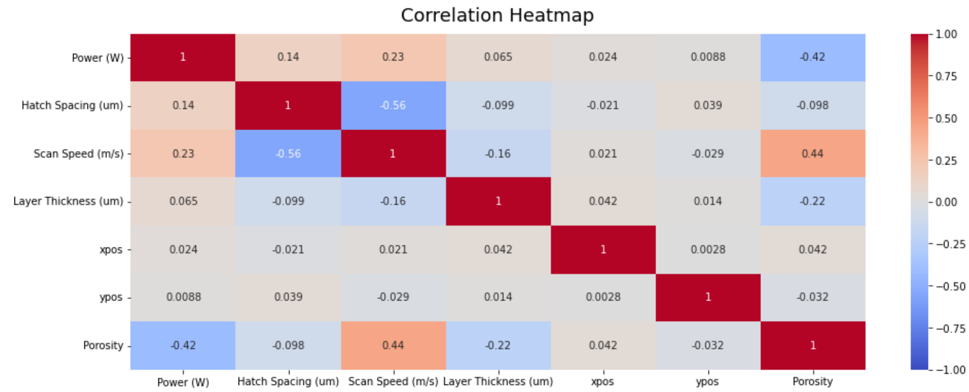

Figure 12. Correlation coefficient heatmap between input features and porosity.

3.3.2 Queries

For the best case, the predicted input space was sorted from least to greatest predicted porosity value, and the 12 parameter sets predicted to have the least porosity were selected, manufactured, and characterized. The most and least dense samples are shown in Figure 13. The statistics of the measured and predicted values are shown in Table 11. In addition, the measured versus predicted porosity values for each parameter set are shown in Figure 14. Of the samples printed, 6 of 12 were under 1% porosity; however, samples with larger amounts of porosity were evident, the maximum reaching 9.62% porosity, with an average porosity of 2.5%.

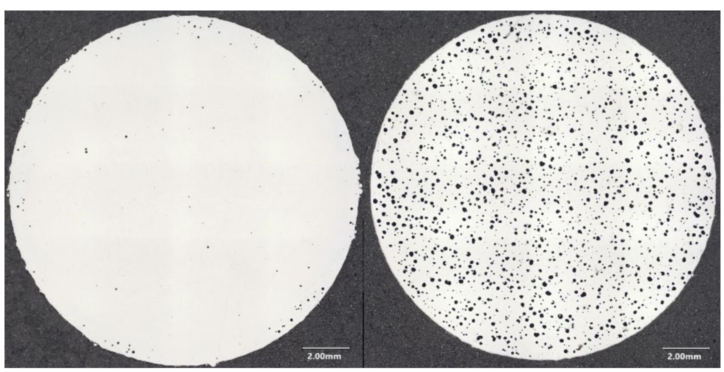

Figure 13. The least and most porosity from validation set for best case.

Table 11. Predicted vs. measured porosity for best case.

Statistical metric	Predicted porosity	Measured porosity
Minimum	0.05%	0.07%
Maximum	0.15%	9.62%
Average	0.11%	2.50%
Standard deviation	0.03%	3.55%

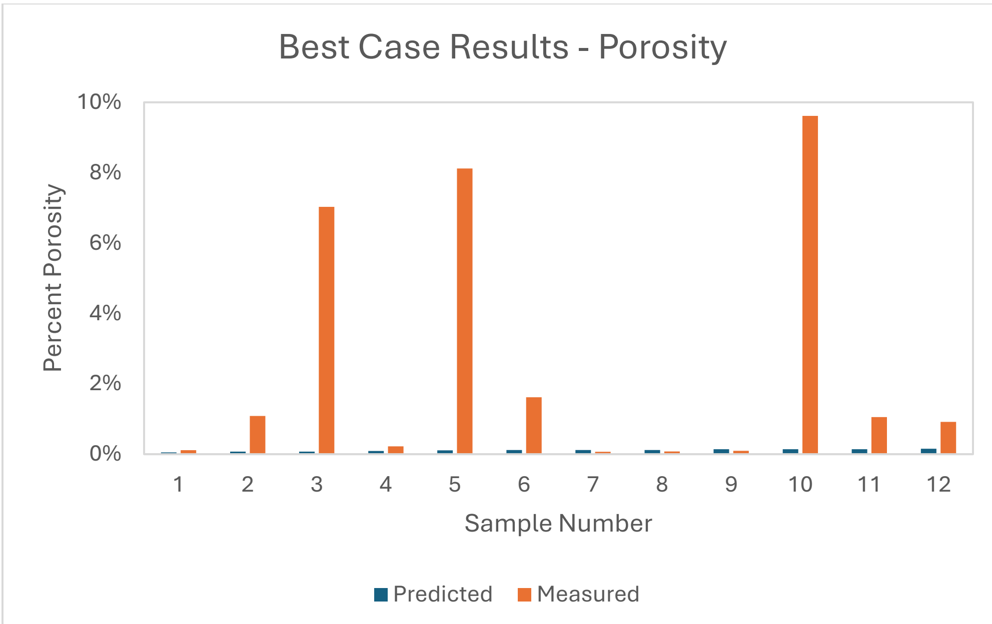

Figure 14. Porosity validation results for best case.

For the "specific" case, the absolute difference was taken between the target value of 20% and the predicted porosity value. The six closest samples were manufactured and characterized. The six samples yielded an average porosity of 25% with the target value within one standard deviation, with an average percent error of 5%. In addition, the sample closest to the target value had a measured porosity of 20.08%. The statistics for the predicted and measured porosity values are shown in Table 12. The measured versus predicted porosity with the target of 20% marked is shown in Figure 15.

Table 12. Predicted vs. measured porosity for specific case.

Statistical metric	Predicted porosity	Measured porosity
Minimum	19.98%	20.08%
Maximum	20.02%	35.14%
Average	20.00%	25.00%
Standard deviation	0.01%	5.47%
Average %error	. . .	5.01%

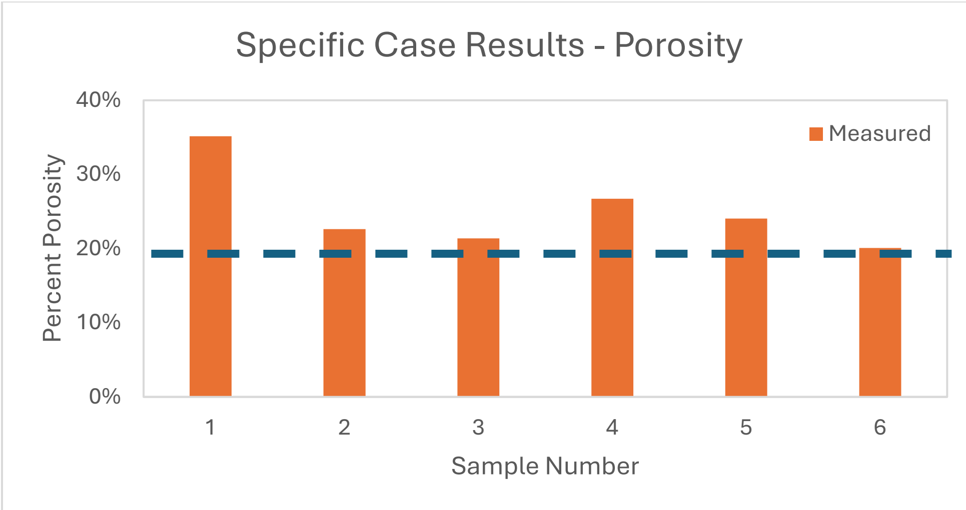

Figure 15. Porosity validation results for specific case.

The predicted mechanical properties were normalized for each target, and a weighted calculated score of each process parameter in the LHS input space was calculated. Again, weights of 1 were used for both the UTS and failure strain predictions. These scores were sorted from least to greatest, and the 23 highest scoring parameter sets were selected and tested.

Figure 16 shows the stress–strain curve for these samples, with the dashed lines indicating the UTS and failure strain of the OEM-optimized parameters. The performance of these 23 samples varied, some achieving near-ideal performance but not surpassing the performance yielded by the OEM-process parameters. Many parameters achieved a much lower UTS and failure strain than the OEM parameters. The mean absolute error for these predictions is high.

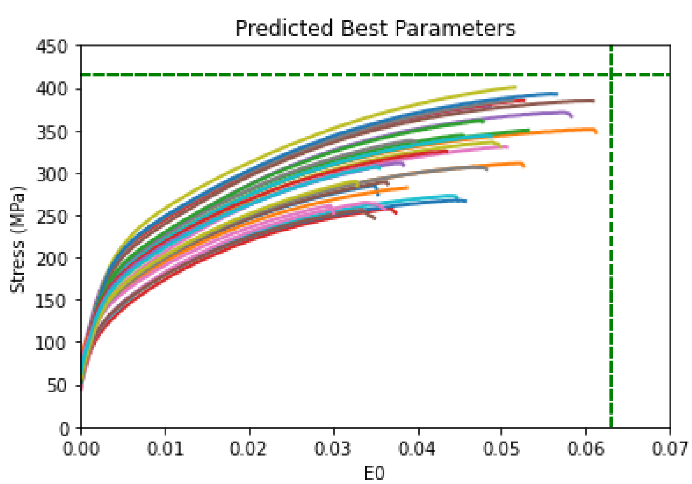

Figure 16. Tensile performance of samples from best case: green dashed lines represent the OEM-optimized parameter set performance.

The predicted input space was queried for the 12 parameter sets predicted to yield performance closest to the target 350 MPa UTS and 0.03 failure strain; again, with the weighted sum of the normalized difference between the target values and the predicted values calculated, as described in Equations 7 and 8. These process parameters were manufactured and tested. Figure 17 shows the stress–strain curves with targets denoted by dashed lines. Table 13 provides some statistics regarding the performance compared to the target values.

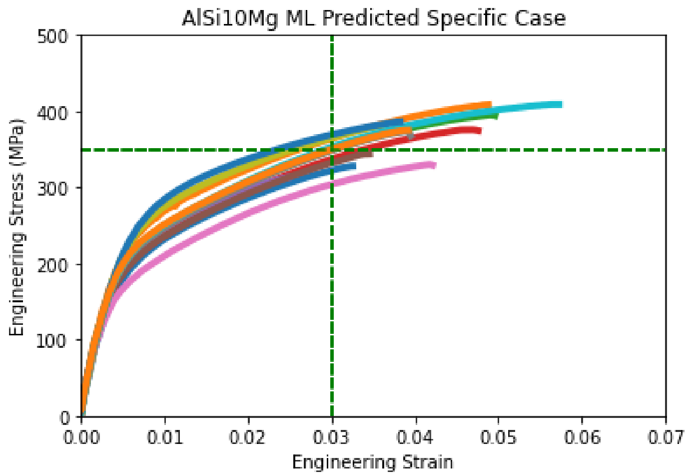

Figure 17. Tensile performance of specific case samples with green dashed lines indicating target values.

Table 13. Tensile results for specific case.

Statistical metric	UTS	Failure strain
Target	350 MPa	0.0300
Average	368.5 MPa	0.0409
Standard deviation	32.1 MPa	0.0088
Mean absolute percent error	8.85%	26.8%

The errors for these predictions were much less those in targeting the best performance, but the average UTS and failure strain were still outside one standard deviation.

4. Discussion

The model for predicting porosity was trained for a single-target variable, whereas the model for predicting tensile properties was generalized and evaluated across two target variables, possibly increasing the error for each respective target. Overall, the initial training of the models showed high degrees of accuracy, with the queries and validation of those models showing mixed but promising results.

4.1 In situ Monitoring

The inclusion of the in situ monitoring data improved predictions for both the porosity and mechanical property predictions. Predictions were not improved for every model tested. However, in models in which the in situ monitoring data did not improve predictions, the overall performance of that model was poor. The better predicting models seemed to respond well to the inclusion of the in situ data. The porosity samples and tensile bar gauges were uniform and symmetrical over a large portion of the build, and the very generalized value of mean layer-wise average was used for predictions. When this data collection method is used for complex shapes—where the cross-sectional area being melted is changing with build height—higher fidelity is needed.

4.2 Best-Case Predictions

The porosity predictions had relatively low errors compared with mechanical property predictions. This is most likely because in instances of porosity, the target value of 0% porosity was set. This is a realistic value for this system, and there were many data points near the target value. ML models are most effective when predicting data similar to that on which it is trained. Therefore, asking the model to predict process parameters that would yield low porosity was well within its ability. However, for mechanical properties, there were no practical limitations like those for the porosity target; this was simply the strongest and most ductile performance. This method meant the parameters selected based on their predicted outputs, most likely had higher uncertainty, which produced higher errors. If these results were fed back into the model and another round of samples were done, then the errors and uncertainty would probably decrease.

4.3 Specific Case Predictions

In the specific case predictions, the error for the porosity predictions were slightly higher than the best case. Overall, the target porosity was still within one standard deviation of the average porosity measured in these samples.

For tensile properties, the prediction errors were significantly less than the errors for predicting best case. Since the targets are well within the bounds of the output space, there is training data that is similar to the target values.

4.4 Sources of Error

Although the models generally performed well, there are some sources of error that could have contributed to a decrease in model performance, especially with the smaller datasets. For the porosity samples, a metallography and optical image analysis method for determining porosity content is highly influenced by sampling. Depending on where you cut your sample, areas could be revealed that have either more or less porosity than is the ground truth or the actual true porosity. In addition, with some of the higher porosity samples, the material pullout could have contributed to higher measured porosity than was native to the as-printed material. Moreover, the XY plane was observed during metallography. Depending on the size and shape of the melt pool, the porosity content can be anisotropic, meaning multiple planes may have to be observed to get an accurate porosity measurement.[15] Taking multiple slices of each sample, performing metallography and optical image analysis, and calculating an average porosity or measuring porosity via μ-CT, which looks at a larger volume of material, could yield more accurate porosity measurements and predictions.

A possible source of error that could decrease accuracy when predicting mechanical properties was for a given parameter set, a single tensile bar was manufactured and tested. This was done to reduce the testing burden and better track the position on the build plate of each sample. Similar to the issue with porosity, a single data point might contain some error and could produce additional scatter, which could be remedied if enough samples were tested to yield average properties for each parameter set. In addition, all tensile bars were tested with the as-printed surface condition, meaning that surface defects could have influenced the performance of the samples. Drastically different surface roughnesses were observed. If surface roughness was included as another target variable and minimized, then perhaps the maximum UTS and failure strain could increase. Moreover, increased porosity content in the tensile bars could cause randomness and variability, especially in the failure strain.

Finally, the size of a data set heavily influences the accuracy of an ML model. Here, relatively small datasets were used. Over time, as data are collected, the model can be retrained and fine-tuned to better predict the whole input space or specific regions of interest based on the data provided.

5. Conclusion

With limited data, we showed that various models can be optimized and evaluated to predict structure and performance of LPBF process parameters. These optimized models can then be utilized to provide informative process maps and relational metrics. The models can then be used to suggest process parameters that will yield a specific output. Validation of these performances showed some mixed results, but they were promising when predicting areas known well within the output space.

The inclusion of in situ monitoring data improved the accuracy of predictions; however, the current method of data collection is not sustainable for larger models or complex shapes. Our framework is amenable to any process, material, or target variable, and it is a useful tool that requires minimal data.

6. References

1. Sun S, Brandt M, Easton M. Powder bed fusion processes: an overview. Laser Additive Manufacturing. 2017;55–77.
2. Zhao R et al. On the role of volumetric energy density in the microstructure and mechanical properties of laser powder bed fusion Ti-6Al-4V alloy. Additive Manufacturing. 2022;51:102605.
3. DebRoy Tl, David SA. Physical processes in fusion welding. Reviews of Modern Physics. 1995;67(1):85.
4. Lindström V, Lupo G, Yang J, Turlo V, Leinenbach C. A simple scaling model for balling defect formation during laser powder bed fusion. Additive Manufacturing. 2023;63:103431.
5. Ahmed N, Barsoum I, Haidemenopoulos G, Abu Al-Rub RK. Process parameter selection and optimization of laser powder bed fusion for 316L stainless steel: a review. Journal of Manufacturing Processes. 2022;75:415–434.
6. Liu Q et al. Machine-learning assisted laser powder bed fusion process optimization for AlSi10Mg: new microstructure description indices and fracture mechanisms. Acta Materialia. 2020;201:316–328.
7. Moran TP, Warner DH, Soltani-Tehrani A, Shamsaei N, Phan N. Spatial inhomogeneity of build defects across the build plate in laser powder bed fusion. Additive Manufacturing. 2021;47:102333.
8. Liu S, Stebner AP, Kappes BB, Zhang X. Machine learning for knowledge transfer across multiple metals additive manufacturing printers. Additive Manufacturing. 2021;39:101877.
9. Liu J et al. A review of machine learning techniques for process and performance optimization in laser beam powder bed fusion additive manufacturing. Journal of Intelligent Manufacturing. 2023;34(8):3249–3275.
10. Shields MD, Zhang J. The generalization of Latin hypercube sampling. Reliability Engineering & System Safety. 2016;148:96–108.
11. Wu D, Lin C-T, Huang J. Active learning for regression using greedy sampling. Information Sciences. 2019;474:90–105.
12. ASTM E8/E8M-22. Standard test methods for tension testing of metallic materials. ASTM International; 2024.

13. Badillo S et al. An introduction to machine learning. Clinical Pharmacology & Therapeutics. 2020;107(4):871–885.

14. Sadaiyandi J, Arumugam P, Kumar Sangaiah A, Zhang C. Stratified sampling-based deep learning approach to increase prediction accuracy of unbalanced dataset. Electronics. 2023;12(21):4423.

15. Romano S, Abel A, Gumpinger J, Brandão, Beretta S. Quality control of AlSi10Mg produced by SLM: metallography versus CT scans for critical defect size assessment. Additive Manufacturing. 2019;28:394–405.

List of Symbols, Abbreviations, and Acronyms

1-/3-D	one-/three-dimensional
AM	additive manufacturing
ARL	Army Research Laboratory
CT	computed tomography
E_v	volumetric energy density
LHS	Latin hypercube sampling
LOF	lack of fusion
LPBF	laser powder bed fusion
ML	machine learning
MSE	mean squared errors
OEM	original equipment manufacturer
UTS	ultimate tensile strength

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