RESEARCH ARTICLE | SEPTEMBER 26 2023

Data-driven density prediction of AlSi10Mg parts produced by laser powder bed fusion using machine learning and finite element simulation

Special Collection: Proceedings of the International Congress of Applications of Lasers & Electro-Optics (ICALEO 2023)

Bastian Bossen ; Maxim Kuehne ; Oleg Kristanovski ; Claus Emmelmann

J. Laser Appl. 35, 042023 (2023)
https://doi.org/10.2351/7.0001141

Journal of Laser Applications ARTICLE pubs.aip.org/lia/jla

Data-driven density prediction of AlSi10Mg parts produced by laser powder bed fusion using machine learning and finite element simulation

Cite as: J. Laser Appl. 35, 042023 (2023); doi: 10.2351/7.0001141
Submitted: 26 June 2023 · Accepted: 1 September 2023 ·
Published Online: 26 September 2023

Bastian Bossen,[1] Maxim Kuehne,[1] Oleg Kristanovski,[2] and Claus Emmelmann[1]

AFFILIATIONS

[1]Institute of Laser and Systems Technology, Hamburg University of Technology, Harburger Schloßstraße 28, Hamburg 21079, Germany
[2]Fehrmann Materials X GmbH, Stenzelring 19, Hamburg 21107, Germany

Note: Paper published as part of the special topic on Proceedings of the International Congress of Applications of Lasers & Electro-Optics 2023.

ABSTRACT

Powder bed fusion of metals using laser beam (PBF-LB/M) is a commonly used additive manufacturing process for the production of high-performance metal parts. AlSi10Mg is a widely used material in PBF-LB/M due to its excellent mechanical and thermal properties. However, the part quality of AlSi10Mg parts produced using PBF-LB/M can vary significantly depending on the process parameters. This study investigates the use of machine learning (ML) algorithms for the prediction of the resulting part density of AlSi10Mg parts produced using PBF-LB/M. An empirical data set of PBF-LB/M process parameters and resulting part densities is used to train ML models. Furthermore, a methodology is developed to allow density predictions based on simulated meltpool dimensions for different process parameters. This approach uses finite element simulations to calculate the meltpool dimensions, which are then used as input parameters for the ML models. The accuracy of this methodology is evaluated by comparing the predicted densities with experimental measurements. The results show that ML models can accurately predict the part density of AlSi10Mg parts produced using PBF-LB/M. Moreover, the methodology based on simulated meltpool dimensions can provide accurate predictions while significantly reducing the experimental effort needed in process development in PBF-LB/M. This study provides insights into the development of data-driven approaches for the optimization of PBF-LB/M process parameters and the prediction of part properties.

Key words: additive manufacturing, PBF-LB/M, AlSi10Mg, machine learning, density prediction, finite element method

I. INTRODUCTION

Additive manufacturing (AM) has been a rapidly growing market and is projected to maintain this trajectory in the coming years.[1,2] Out of the various AM processes, powder bed fusion of metals using laser beam (PBF-LB/M) is one of the most industrialized and well established AM technologies for creating metal parts in various key industries such as aerospace or medical applications.[1,3] However, the process is known to be very sensitive with over 150 factors influencing the resulting part quality.[4] Because of this complex nature and unknown correlations between part quality and process parameters, empirical full-factor or at least partial-factor design of experiments are still the norm in AM process development.[5] Since machine time as well as powder feedstock and labor are oftentimes scarce and expensive, this drives costs for process and material developments and also for maintaining high-quality parts in production. Due to this costly and time-consuming development and qualification, as of 2022, Srinivasan *et al.*[6] identified 141 alloys that have been investigated in the literature in metal additive manufacturing—97 of these attributable to the PBF-LB/M process. The potential alloy options available

in conventional manufacturing exceed this number by some orders of magnitude[7] and highlight the need for the reduction of costly empirical studies.

One approach to reduce the empirical effort in process development is the use of numerical simulations to gain insights into the underlying effects such as residual stresses, temperature distribution, or melt flow dynamics during the PBF-LB/M process. These simulations can be divided into micro-, meso-, or macroscale depending on the scale they cover, and each scale has its own drawbacks and advantages.[8–10] One of the most prominent approaches is the finite element method (FEM) for numerical analysis of the PBF-LB/M process. While numerical approaches have been yielding valuable information to understand and optimize the process, the computational effort involved, depending on the simulation scale and accuracy, can be very high.[8,11] Thus, while reducing the empirical effort via numerical approaches is possible, it is in many cases, not economical.

Another increasingly considered approach is the use of artificial intelligence (AI) and machine learning (ML) methods and algorithms to optimize the PBF-LB/M process.[12–15] In these approaches, models are developed that utilize statistical principles to examine existing data sets, identify hidden patterns in the data, and then allow generalized projections for unknown situations and data points. These models can predict, among other things, meltpool characteristics,[16,17] tensile properties,[18] or porosity[19,20] by using input data such as optical or acoustic sensor data, design information, or process parameters. Optimization of the PBF-LB/M process with the goal of maximizing density during initial material qualification is one of the most common approaches.[21,22] The use of ML for such optimization is reported by Park *et al.*[23] for Ti-6Al-4V with respect to relative part density and surface roughness. Their model was trained on 2048 empirical data points and yielded a high accuracy with an error of only 0.94% and 0.4% for density and surface roughness, respectively, when tested.[23] Furthermore, in a study by Gor *et al.*,[24] different ML approaches are tested to predict the relative part density of the PBF-LB/M 316L specimen based on 40 empirical data points. Their approach yielded an accuracy of 96.01% in the predicted versus actual value.

So, while these approaches are highly effective at predicting certain critical processes or quality quantities, they are highly dependent on both, the amount of data as well as the quality and preparation of data that is used to train the models.[25] Additionally, the data need to be gathered first, which is again costly, and further depending on the application of the ML, might negate its purpose. Thus, the approach used in this study aims to tackle this problem through a combination of ML and numerical methods to predict the relative density of PBF-LB/M parts. The addition of simulation data can provide the ML model with physical information about the process which can exceed the information of process parameters alone. One goal of including simulation data in the model can be to create an accurate model with less empirical data points since additional information is provided. Second, the model can more easily generalize between, e.g., different machines or materials by adjusting machine or material properties in the simulation and simulate the process for a specific machine and material.

Such a model can be a powerful tool in process optimization of the PBF-LB/M process. It can be used to improve existing process parameters, e.g., in regard to productivity, by predicting faster process parameters while maintaining a defined minimum density. Furthermore, such a model can be used to transfer parameters for a specific material from one machine to another by incorporating individual machine and process parameters such as laser diameter or process gas. Lastly, by adjusting the material properties in the simulation, such an approach can also be used to reduce the empirical effort for the development of new materials, because the ML model can get material specific inputs via simulation to inform and accelerate experimental approaches.

II. METHODOLOGY

The basic methodology used in this study is shown in Fig. 1. Two ML models for relative part density prediction are developed for AlSi10Mg, a material commonly used in PBF-LB/M. The models are trained via process parameters, empirical density measurements, as well as simulated meltpool geometries. The meltpool geometry data are generated using FEM simulation. The simulation is validated via comparison with empirical meltpool measurements and yields meltpool dimension data for single and double tracks. Single and double hatch meltpool dimensions are considered.

Based on the input parameters, the models are trained and validated via the hold-out method. To analyze the effect of including simulated meltpool data on the model's performance, one

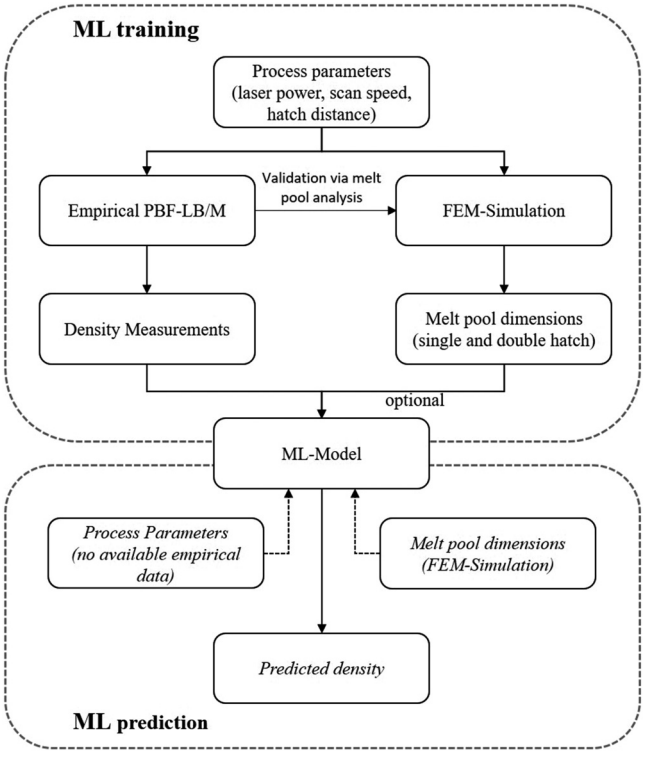

FIG. 1. Schematic methodology for the development of a ML model for density prediction based on simulated meltpool geometries and empirical density data.

model is trained based on the process parameters and the second model with process parameters and additional meltpool dimensions. The output of the models is the relative part density based on the provided inputs as shown in Fig. 1.

The goal of the methodology is the development of a prediction model for relative part density for unknown parameter combinations in the PBF-LB/M process. Future use of the ML model can be to optimize existing process parameters, e.g., in regard to productivity, by predicting the relative density of a broad parameter window based on unknown parameters and simulation data. For this purpose, 64 additional parameter combinations are simulated and the relative part density is predicted by the trained ML-model. Based on the predicted density results, new parameter combinations that yield a high predicted density and productivity can be tested directly from the ML model. Thus, the empirical effort for process optimization can be potentially reduced. Furthermore, due to the use of simulation data and resulting additional physical information, the methodology can potentially also be applicable when transferring existing materials to new machines or developing new materials. For this purpose, the ML model can be trained continuously during process development and predict suitable process windows that result in high-density parts for the subsequent experiments, thus, also potentially reducing the experimental effort.

A. Empirical data

The data for the empirical relative density measurements are based on the work of Wischeropp *et al.*[26,27] Here, density cubes are fabricated from AlSi10Mg using an AconityLAB PBF-LB/M system (Aconity3D GmbH, Herzogenrath, Germany). The process parameters of the 54 sample cubes are shown in Table I. The relative density of the samples is measured optically via micrographs and reaches 62.5%–99.9% with a median of 99.5%. Additionally, empirically generated single track meltpool measurements from the same work are taken for simulation validation. This data set was chosen because it is largely homogenous in terms of machine, powder used, time it was conducted, etc. This eliminates the potential statistical error and other deviations from other process influences that might increase the error and need for larger amounts of data in the proposed methodology.

The whole data set is displayed in Fig. 2. The data point with the lowest relative density of 62.5% is excluded from the set since the gap to the next data point at 85.4% was deemed too large for a ML approach and might decrease the model's accuracy.

TABLE I. Process parameters for empirical density data generation used in this study from Ref. 26.

Factor	Level
Laser power (W)	300, 400, 500, 700, 800
Scanning speed (mm/s)	1000, 1500, 2000
Hatch distance (mm)	0.10, 0.15, 0.20
Layer thickness (mm)	0.05
Temp. build-platf. (°C)	200

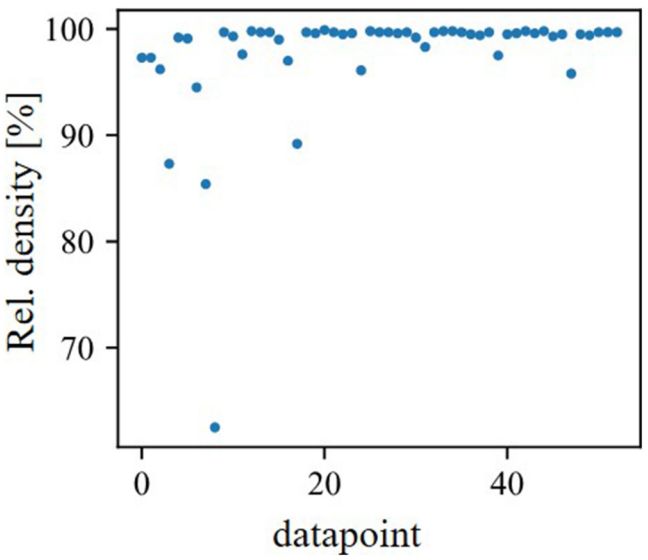

FIG. 2. Full data set for relative density used in this study from Wischeropp *et al.* (Ref. 26).

B. FEM simulation

The simulation model used here is based on a previous work by the authors.[28] A thermal FEM simulation of the PBF-LB/M process in COMSOL MULTIPHYSICS 5.2A (Comsol Multiphysics GmbH, Göttingen, Germany) is conducted. The general simulation setup is shown in Fig. 3. It consists of one powder layer where the powder is assumed to be a homogenous volume on a bulk base material. The energy source is simulated using a Gaussian heat source that is moved on the first layer with a defined scan strategy and process parameters such as laser diameter, laser power, or scan speed. The thermal material properties needed for the simulation are based on the work of Wischeropp[27] and, additionally, on their own measurements for AlSi10Mg.

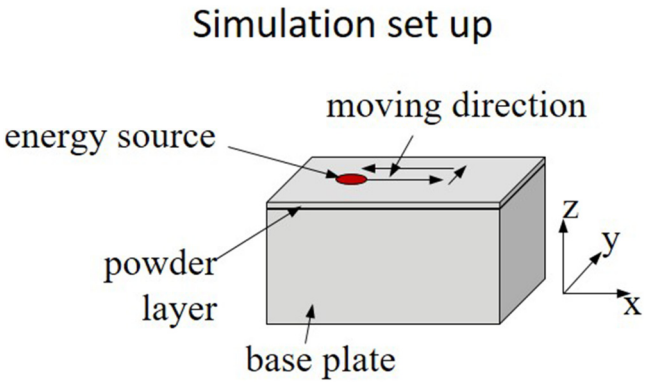

FIG. 3. Schematic drawing of the simulation set up with bulk material and a homogeneous powder layer. Indicated is the laser path for two hatches.

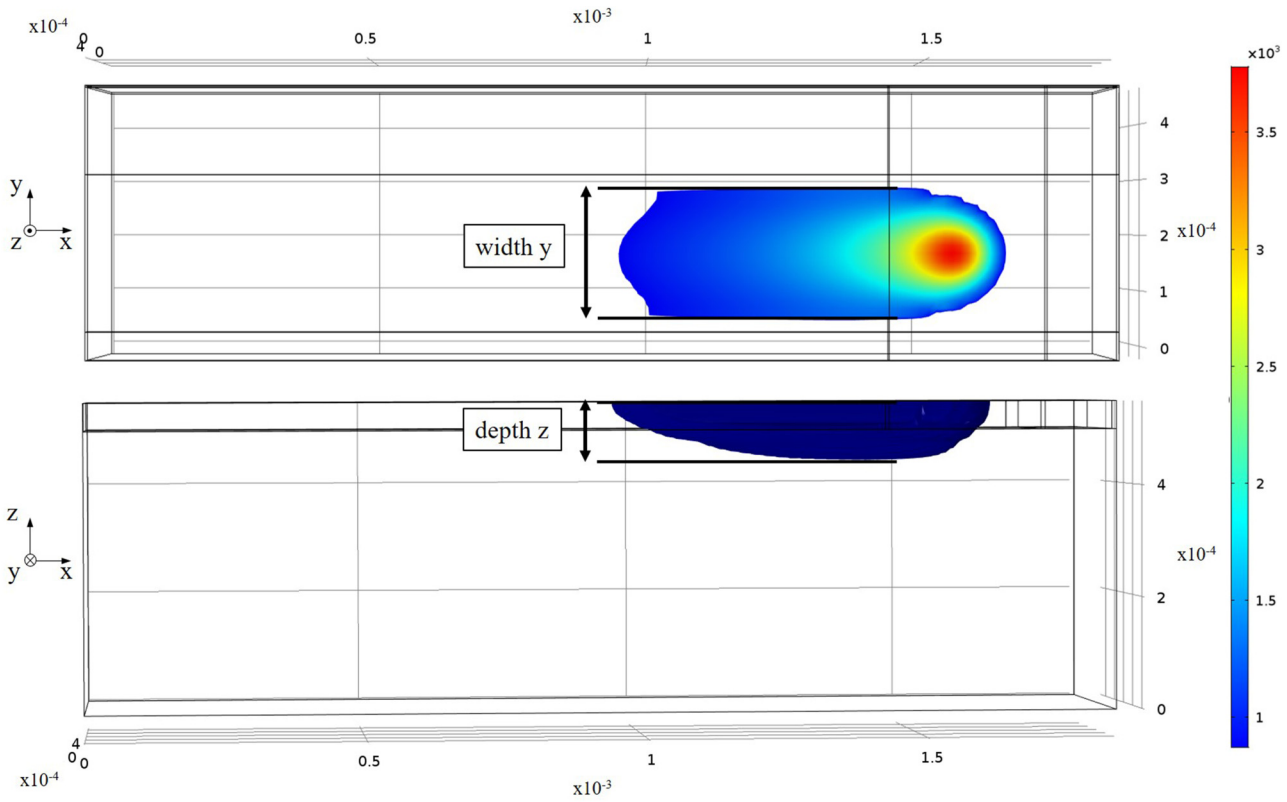

FIG. 4. Simulated meltpool width and depth of AlSi10Mg in the PBF-LB/M process using COMSOL MULTIPHYSICS.

The validation of the numerical simulation for AlSi10Mg is conducted by comparing empirical meltpool measurements with simulated meltpool dimensions. The simulation error is displayed by plotting the relative error (1) over linear energy density (LED),

$$\text{Rel. error } (\%) = \frac{y_{measured} - y_{simulated}}{y_{simulated}} \times 100. \quad (1)$$

For data generation, the meltpool geometries of the 54 empirical process parameter combinations from Table I were simulated with a hatch length of 1.5 mm for each hatch. Figure 4 shows the evaluation of the meltpool width and depth via the FEM simulation for the first meltpool by filtering elements with $T > T_{liq}$.

By simulation of multiple tracks, the effect of hatch distance as a process parameter can be simulated, which can have a direct influence on part density.[29] Figure 5 shows the relationship between lack of fusion and evaporation induced porosity due to varying the hatch distance. Hatch distances that are too large lead to lack of fusion pores since the energy in the overlap area is not sufficient to completely melt the powder layer. On the other hand, if the hatch distance is too narrow, gas pores are created by resulting high temperatures in the overlap zone leading to evaporation of the material or parts of the alloy. This relation can be indirectly represented in the simulation by the temperature field and resulting meltpool in the overlap zone of two meltpools.

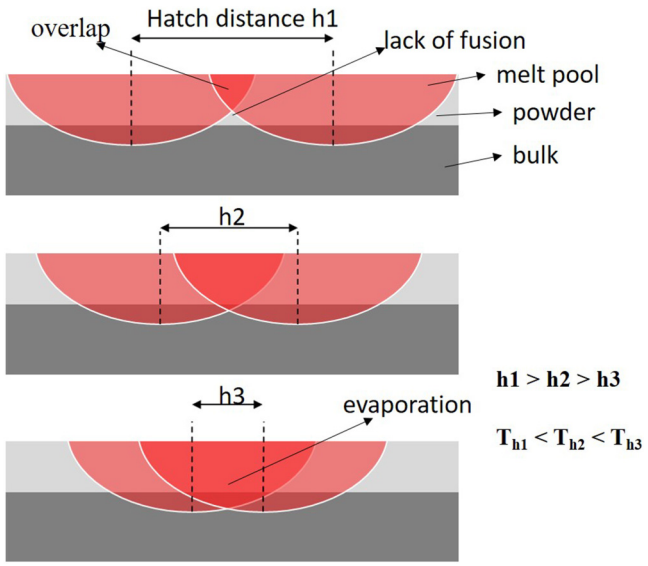

FIG. 5. Illustration of varying hatch distances in the PBF-LB/M process and its relation to lack of fusion and evaporation induced porosity.

The simulated meltpool dimensions for all timesteps are shown in Fig. 6 for increasing laser energy at constant scan speed and hatch distance for two parameter combinations. In the simulation, the first ~45 timesteps represent the first hatch, then follow the turning of the energy source until it moves back to the next hatch. It can be observed that for laser powers equal to or below 500 W, the first meltpool width and depth reach a stable state, sharply increase at the turning point, and then decrease over time again for the second melt track. For high energies, no stable state arises in the meltpool width and, additionally, levels out at ~0.5 mm, which is a limitation of the simulation. Based on this finding, two timesteps for data generation are defined. The meltpool width and depth of the first hatch are indicated as y_1 and z_1, respectively. Furthermore, for the second hatch, y_2 and z_2 are defined. The main reason for selecting these timesteps is using a mostly stable meltpool that changes little over time.

C. Machine learning

A ML model comparison from a previous work by the authors showed that a multilayer perceptron (MLP) approach had the highest performance and accuracy for the type and amount of data.[28] Based on this, an MLP approach is followed in this work. Since the data set in this study contains information about input and output, a supervised learning algorithm is considered. The ML model is implemented in Python 3.8 with the libraries pandas and scikit-learn.[25,30] The data set for training and evaluation of the MLP is based on the presented 54 empirical relative density values, associated process parameters, and simulation data. The models are trained with and without the inclusion of simulation data to analyze the impact of data augmentation with simulation outputs. Both training approaches are analyzed and compared to each other to evaluate the effect of including simulation data in the training. Here, 70% of the data sets are used for training and 30% are retained for later evaluation. Due to the small amount of data,

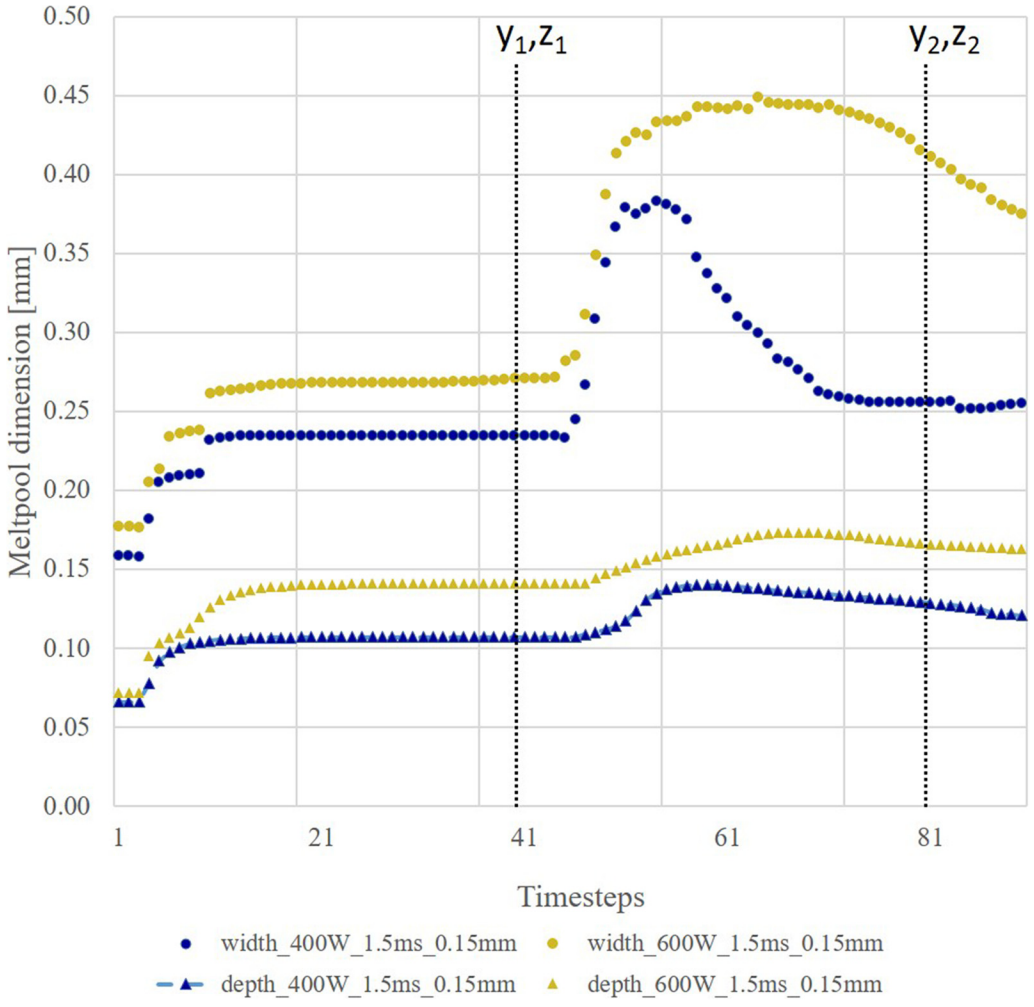

FIG. 6. Simulated meltpool width and depth over consecutive simulation timesteps with indicated data output as ML-input parameters.

TABLE II. Simulated PBF-LB/M process parameters for relative density prediction.

Factor	Level
Laser power (W)	285, 470, 655, 840
Scanning speed (mm/s)	0.95, 1.33, 1,77, 2.10
Hatch distance (mm)	0.09, 0,13, 0.18, 0.22
Layer thickness (mm)	0.05

the data set is split using stratified shuffle split[25] based on energy density. This ensures that data points from each energy density range are present in the training and test sets and, thus, avoiding sampling bias. A grid search algorithm is implemented to optimize hyperparameters to improve accuracy and minimize overfitting as well as underfitting.

The output of the prediction model is the process parameter dependent relative part density. The relative density values are displayed in a process map with the inverse normalized hatch width $1/h^*$ over the normalized energy density E^* according to Herzog *et al.*,[31]

$$h^* = \frac{h}{r_L}, \tag{2}$$

$$E^* = \frac{a \cdot P_L}{2 \cdot v_L \cdot \rho \cdot c_p \cdot (T_{liq} - T_0)}. \tag{3}$$

The inverse normalized hatch distance h^* (2) is calculated via the hatch distance h and the laser radius r_L and E^* (3) via absorption α, laser energy P_L, scan speed v_L, density ρ, specific heat capacity c_p, as well as the initial temperature T_0 and melting temperature T_{liq}. Due to its normalized nature, this proposed representation of data has the advantage of being more generalized and thus allowing comparison between different machines and materials more easily.

The evaluation of the MLP is conducted, on the one hand, via the qualitative comparison of the predicted and the empirically measured relative density data points. On the other hand, the model error is described by the commonly used root mean square error (RMSE) for test and training data. It describes the averaged squared distance between the predicted value $\hat{y}$ and the real value y of a set of data points m.[32] Due to statistical variance, the model is trained 50 times and the mean RMSE is used to determine the training and test data errors,

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y} - y_i)^2}. \tag{4}$$

Moreover, both prediction models are analyzed and compared in terms of model accuracy in terms of their proposed input parameters. To further prove the validity of the approach of including simulated meltpool data, the correlation between meltpool dimension data with relative density is analyzed.

Lastly, to test the plausibility and usability of the ML-model for unknown parameters, 64 PBF-LB/M parameter combinations are simulated and their relative density is predicted in a process map. The parameter window is shown in Table II. The parameters are chosen by extending the experimental limits by 5% for laser power and scan speed, resulting in circa 10% in normalized energy density and 10% for hatch distance. Two levels in between the limits are included at equal distances.

III. RESULTS

A. Simulation validation

For simulation validation, the empirically measured meltpool widths and depths from Ref. 27 are compared to the simulated meltpool geometries generated by the FEM simulation. Figure 7 shows the relative error of the simulated versus the empirical widths and depths, respectively, over the LED.

For the simulation of the meltpool width, a good agreement with the empirical measured values can be achieved. For most of the investigated measurements, an error smaller than 10% is reached with only a few outliers reaching a relative error of up to 20%. The simulation of the meltpool depth, on the other hand, shows a larger error and a clear tendency to underestimate the real meltpool depth. Here, deviations up to −45% are observed. This is especially visible for lower linear energy densities. For larger linear energy densities, the error decreases. At the same time, the width is underestimated. Thus, it is likely that the overall energy in the system is underestimated. The average error for the meltpool width is −0.42% and approximately −27% for the depth. The model error can be explained by inaccuracies in the material model that is used in this study. Additionally, for computational and numerical reasons, simplifications to the simulation have to be made, and, thus, some physical effects such as evaporation or meltpool flow effects are not represented here. Considering these points, the simulation error is still small and it is deemed sufficiently accurate for further use as data input for ML.

B. Empirical density data

Based on the process parameters in Table I, Fig. 8 shows the resulting process map with which the prediction models are validated. Also indicated are schematic borders indicating lack of fusion and evaporation induced porosity.

In this process map, a clear zone can be observed, in which, parts with at least 99% relative density can be produced. The empirical data also show a rapidly falling density for decreasing energy densities indicating a lack of fusion effects. However, for evaporation porosity, this distinction can only be clearly shown for $1/h^* = 0.70$. The two remaining hatches show little to no decreasing relative part density with increasing energy density. Thus, this part of the data set is likely not ideal for the ML approach, because there is no clear boundary visible. However, overall, the data set should be suitable for the training of a ML predictive model. Lastly, it should be noted that some data points are superimposed in this representation due to having the same energy density at the same hatch distance. Therefore, these data points are shown with increasing marker size.

C. ML density prediction

First, the data set is studied to identify correlations between the meltpool dimensions and the relative density. For this purpose, the relation between relative density and simulated meltpool width

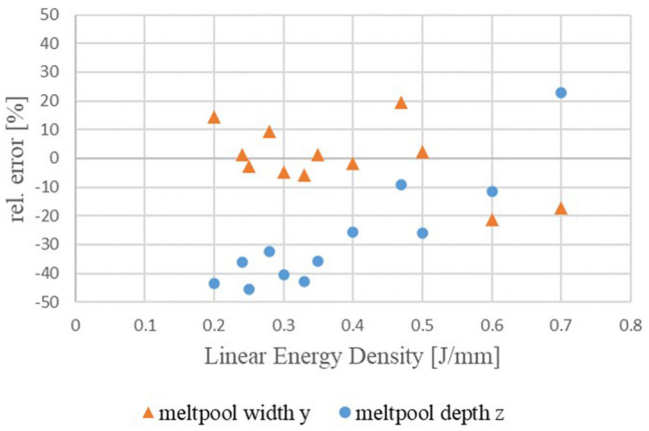

FIG. 7. Relative error of meltpool width y and depth z of simulation data compared to single track empirical measurements.

and depth is displayed in Fig. 9. A correlation between the data is visible. On the one hand, for small meltpool widths and depths, the relative density trends toward lower values, hence indicating a lack of fusion porosity. On the other hand, for large meltpool depths, lower density is shown, indicating evaporation porosity. Meltpool geometries between those regions yield high density. However, there are some outliers present in the data, e.g., at meltpool width y_2 at 0.44 mm width, four different relative density values, and also for depth, there are data points that do not follow this trend. This noise might be due to larger errors in the simulation model for higher energy densities.

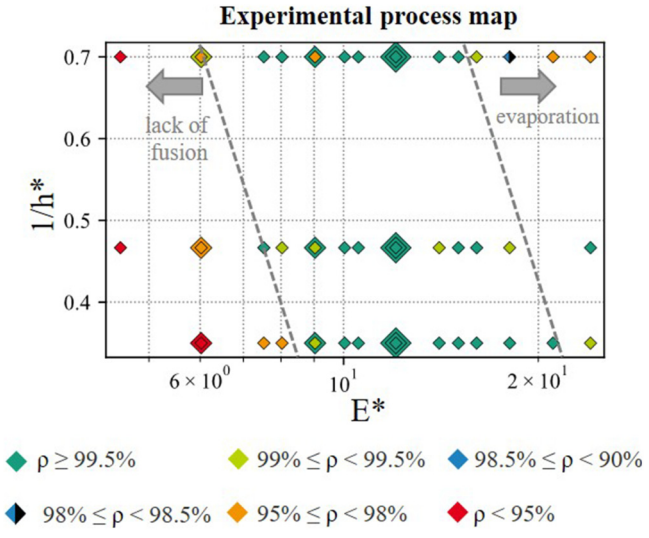

FIG. 8. Experimental PBF-LB/M process map of AlSi10Mg density cubes with schematically indicated lack of fusion and evaporation porosity.

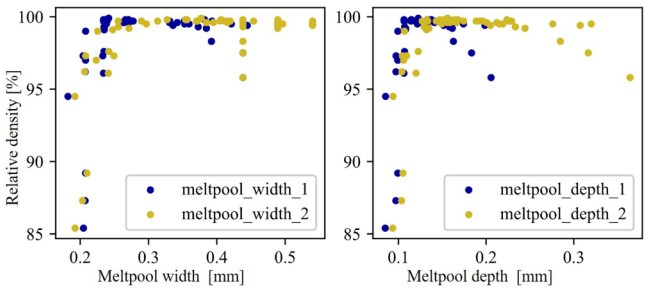

FIG. 9. Relative density against simulated meltpool widths y_1, y_2, and depths z_1, z_2.

Afterward, the models are trained on empirical as well as additional simulation data. They are evaluated with the hold-out method using 30% of the data as test data. The results of the model's prediction ability are shown in Fig. 10. Each training and test data point is plotted against the model's prediction. The diagonal line shows the area where the predicted values match the empirical ones, and consequently, the proximity of the data points to this line is a measure of the accuracy of the predictive model. The model trained with only process parameters shows a good fit with training and test data. The predictions for both training and test data vary homogeneously around the optimum, indicating that no overfitting occurs. The test data predict the empirical reference to a high degree, especially for relative densities above 98%.

For one data point, the prediction underestimates the measured relative density by approximately 4%. The model trained with additional simulated meltpool data behaves similarly overall. The predicted relative density values scatter slightly around the optimum and, thus, are also not considered overfit. The test data predict empirical measurements to a high degree, except for the same data point that is again underestimated to a similar degree. Thus, the accuracy of both models can be considered very high and, consequently, are suitable for the prediction of the relative density. However, both models show inaccuracies for at least one data point with a low relative density. This might indicate inaccuracies at the lower limit of the relative density due to the limited data in that area.

To further evaluate the performance of each model, the average training and test errors are calculated via RMSE after 50 training runs, as shown in Table III. Both models show small training and test errors below 1% for relative density. However, including simulated meltpool data increases the test error by 21%, resulting in a decrease in accuracy, compared to using only process parameters and empirical relative density as the input.

For further evaluation of the relative density prediction, process maps are shown in Fig. 11 for both models and the empirical data set for comparison. Indicated is also, schematically, the boundary for lack of fusion and evaporation porosity. Training data and test data are shown in Figs. 11(a) and 11(b), respectively. Similar to Fig. 8, superimposed data points are shown with increasing marker size. Comparing the empirical process map to the predicted process map by the ML model trained without simulation

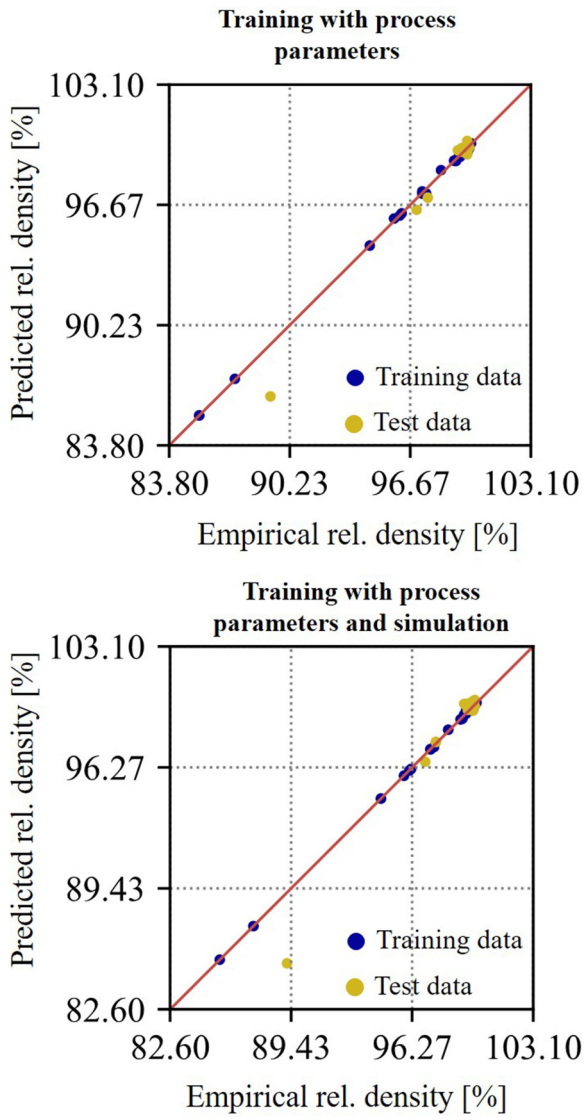

FIG. 10. Predicted against empirically measured relative density of training and test data for different input parameters.

data, the model has small deviations, especially in the test data, within 0.5%–1% of relative density. It can correctly identify reduced relative density in both the lack of fusion and the evaporation zone. Due to the data set that is used, evaporation is not as clearly distinguishable as the lack of fusion porosity. This is likely due to the reason that the empirical data do not include sufficient data in that region.

The addition of simulated meltpool data shown in Fig. 11.2 largely shows similar results. Here, the model is also able to predict evaporation and lack of fusion porosity, as well as the region where almost fully dense samples can be manufactured. It also does show over- and underestimations within 0.5%–1% of relative density.

TABLE III. Average training and generalization RMSE for both models after 50 iterations.

ML-input	Training error RMSE (n = 50) (rel. density %)	Generalization error RMSE (n = 50) (rel. density %)
P_L, v_L, h, ρ	0.67	0.75
P_L, v_L, h, ρ y_1, y_2, z_1, z_2	0.65	0.91

Furthermore, unknown process parameter combinations (see Table II) are used for the melt pool simulation and tested with the trained ML models. In Fig. 12, the results are displayed. The figures show that both the region of low energy densities, which results in increased porosity, and the region that yields high-density samples can be observed in the process map, similar to the empirical data. For the region indicated with evaporation porosity, both models show slightly different results. While the model trained on process parameters exclusively is mostly in accordance with the empirical data, the model trained with additional simulation data tends to overestimate this region, especially for lower $1/h^*$ values. This can be a result of decreasing correlation for higher energies and larger meltpools with a relative density that can be supported by Fig. 9. Additionally, the simulation validation also shows increasing simulation error for high linear energy densities, thus supporting this further.

However, overall, both models show good agreeance with the empirical and unknown data sets, as supported by the presented data and can, within their limitations, be used for density prediction with high accuracy.

IV. DISCUSSION

A. Model performance and use of simulation data

The results of this work show that ML models based on different input variables allow for accurate density prediction for the PBF-LB/M process of AlSi10Mg samples. This can be shown despite a comparatively small quantity of data. Both models have training and generalization errors well below 1% of relative density and, hence, show a strong ability to predict density within the used data set. One of the reasons for the high predictive ability of the presented models could be the high dependency of the inputs and outputs. The process parameters investigated in this study have a clear physical causality to the temperature profile of the process as well as the resulting meltpool and ultimately the relative density of the samples. Thus, there naturally is a high degree of correlation between the investigated inputs and outputs, which is also shown in Fig. 9 for simulated meltpool geometries. This can decrease the amount of data needed to find predictive correlations. Additionally, the data set used in this study is highly constant in terms of other influences such as machine, powder, or ambient conditions. Using a larger data set that introduces influence factors such as different powder lots, powder storage regimes, and machines will possibly increase the amount of data needed for sufficient correlations and accurate predictions. Similarly, introducing more quality outputs

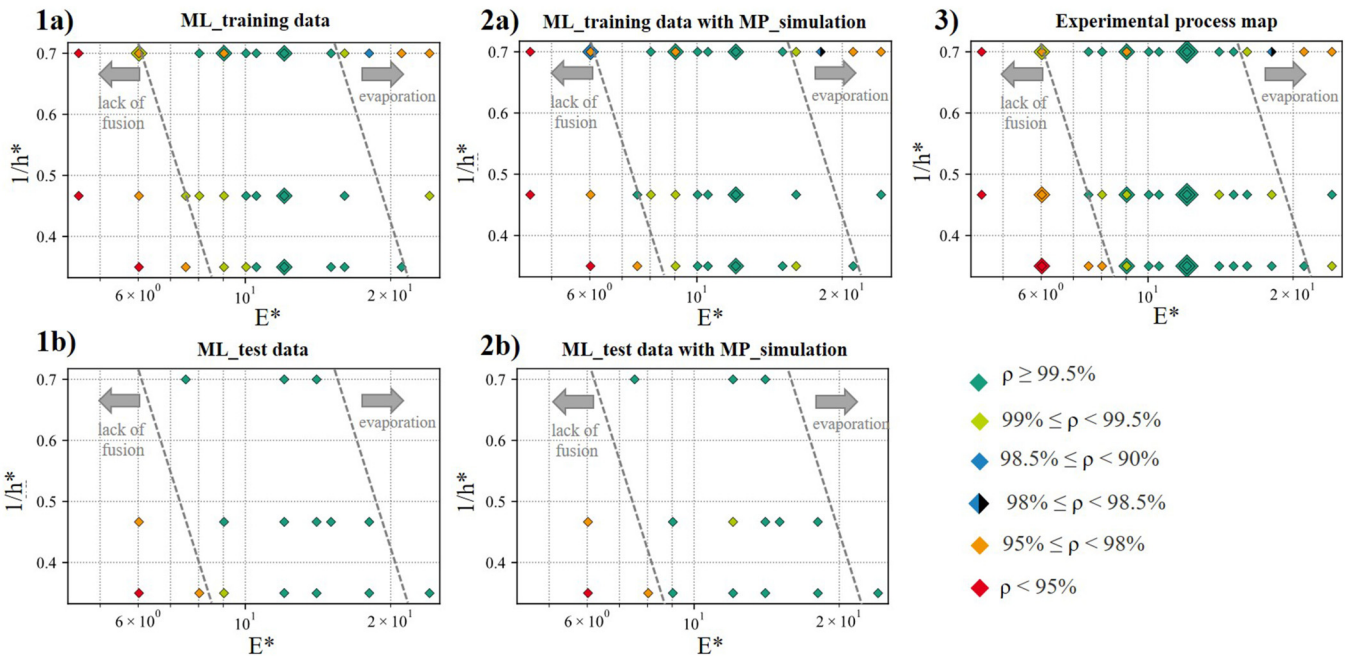

FIG. 11. ML density prediction (1) without simulation data, (2) with simulation data, and (3) experimental process map.

such as surface roughness or mechanical properties, subsequently, also would likely increase the amount of data needed.

Against the initial hypothesis, the use of simulation data, however, did not improve the model's accuracy. While still showing a strong predictive ability with a generalization error of only 0.91% in terms of RMSE, the error is 21% higher than when simulation data are excluded as an input parameter. This is especially evident for the high-energy region where evaporation porosity is expected, but not accurately predicted for some $1/h^*$ values. This is further supported by testing 64 unknown parameter combinations where the error can also be especially observed within that region for the ML model trained, including the simulation data. There can be several reasons for this behavior. First, the simulation used might be too simplified and not describe critical physical processes sufficiently. Therefore, the error increases while not providing enough accurate physical information about the process. This can be further supported by the simulation validation. Here, the error clusters similarly for meltpool width and depth until a linear energy density of 0.4 W/mm. For higher energy densities, the simulation error increases significantly for both cases. The correlation of meltpool geometry and relative density additionally shows slightly diverging results for large meltpool dimensions. The subsequent simulation data could, therefore, be difficult to process for the chosen ML approach.

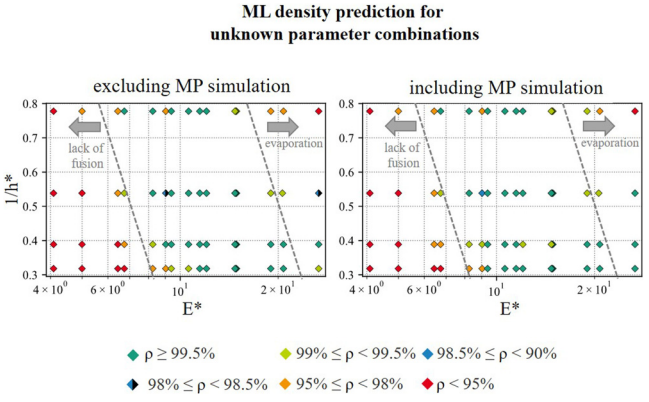

FIG. 12. Prediction of relative density for unknown parameter combinations for both ML models.

Second, the amount of overall data might be insufficient. The data shown has a statistical error and the majority of data points are in the high-density region. This might have a negative impact on the addition of simulation data with the goal of increasing the model's generalization. Additionally, increasing the number of parameters used as inputs from 3 to 7, when within a limited data set of 54 data points, could also increase the error within the MLP.

Overall, this study shows that both training approaches have a comparatively small model error. However, as for many ML-prediction models, there are limits, especially at the boundaries of the trained model. Thus, the use of such a model should be done with a critical evaluation of the prediction in mind. Also, the accuracy of the model can be improved by continuously adding experimental results. Within the presented limitations, the ML models can predict the relative density for unknown parameter

combinations to a high degree. Additionally, evaluating the generalization ability between machines, and for new materials in the future, is feasible.

B. Application of the ML models in process development

The presented approach and developed prediction models can have a significant impact on process development methodologies. The high accuracy and use of simulation data as the input parameter enable different use cases for the presented ML models.

The optimization of the existing parameters is one main use case for both ML models presented. As illustrated in Fig. 12, the relative density of unknown parameters can be predicted by the models. This can be used to optimize the productivity of process parameters without the need of a full factorial design of the experiment. The ML model can predict the relative density of various process parameter combinations. The most productive ones with a minimum requirement for relative part density can be identified and validated empirically. The empirical measurements can then be used to train the model further and predict another set of optimized parameters to be tested. If the model has a good degree of generalization and accuracy, this approach can save valuable machine and operator time, as well as analysis costs during process optimization.

When using simulation data as an input, additional use cases for the prediction model can be realized. PBF-LB/M machines differ in various ways, e.g., in their beam shape, focus diameter, or process gas used. A ML model trained only with process parameters and resulting quality output has poor ability for generalization, if applied on a different machine without the existing data. By using simulation data as the model input, information about the process for different machines can be included without the need for an initial empirical data. The simulation can simulate the process for specific machines by adjusting machine parameters in the simulation such as beam profile, beam diameter, or process gas. Therefore, when including simulation in the training data, e.g., when optimizing process parameters for existing alloys for a new machine, the ML model trained with simulation data has the possibility to generalize between different machines. For this purpose, parameter combinations can be simulated for the new machine, specifically to predict an initial parameter window to be tested. This can accelerate the process development by reducing the empirical effort.

Similarly, in the development of new materials, the inclusion of simulation data can be of value if the material properties are known. By adjusting the material parameters in the simulation, new physical information for the PBF-LB/M process can be generated before experiments are conducted. These data can be used to predict an initial parameter window for processing. Subsequently, the predicted parameters are tested empirically and the results are again used to train the model to improve the prediction during process development for the new material. This approach potentially reduces the empirical effort compared to a completely empirical process development.

V. CONCLUSION

In this work, a ML model was developed that can predict the resulting relative part density for AlSi10Mg in the PBF-LB/M process. The training of the model is based on 54 empirical data points containing process parameters and relative component density. In addition, simulated meltpool width and depth are calculated via FEM simulation. Subsequently, the ML model was trained using the process parameters as input and the relative part density as output. Furthermore, another model was trained additionally using the simulated meltpool dimensions. The hold-out approach was used for training, whereby the model was trained using 70% of the data and the remaining 30% of the data were retained as test data for the evaluation of the generalization error. It has been shown that the generalization error of both models is 0.75% and 0.91% in relative density, respectively. The trained models can then make predictions based on process parameters and/or simulated meltpool geometries.

Both models are able to identify and predict high-density regions with sufficient accuracy and low error. Low-density regions could be delineated in most cases. For isolated process parameter combinations, the models show an error, especially for evaporative porosity. This effect can be explained by insufficient delineation in the empirical data set, low data volume, and simulation error. It could possibly be corrected by broadening the parameter field, increasing the data volume, and improving the simulation.

In conclusion, it has been demonstrated that the proposed methodology of training a model with both empirical and optionally additional simulation data can result in accurate prediction models for relative part density in PBF-LB/M.

VI. OUTLOOK

The presented models and their potential as well as their limitations offer a starting point for future evaluation. To help the accuracy and generalization of the model, a bigger and more diverse data set should be considered. This is especially true for the fringe regions of a process map that result in decreased density. Furthermore, the simulation model needs to be improved, especially for high-energy regions, since it might be the cause for a decreased ML model accuracy.

A deep evaluation and validation of the model's use cases is still outstanding. This includes process optimization for existing machine parameters and also the proposed greater generalization ability due to the inclusion of the simulation data. The latter can be evaluated, e.g., by testing the generalization when transferring the existing process parameters to another PBF-LB/M machine or when qualifying new alloy candidates for the process. Lastly, the methodology can also be potentially applied to other quality criteria such as surface roughness, dimensional accuracy, or mechanical properties.

ACKNOWLEDGMENTS

The presented results are based on the work being conducted in the DigitAl project. The project is funded by the Federal Ministry of Education and Research (BMBF) with Grant No. 01IS21052C. Furthermore, the authors would like to thank Sebastian Jäger and Bhuiyan S. M. Ebna Hai from Fehrmann Material X GmbH for the insightful discussions.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Bastian Bossen: Conceptualization (equal); Data curation (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Software (supporting); Writing – original draft (lead). **Maxim Kuehne:** Conceptualization (equal); Data curation (equal); Investigation (equal); Methodology (equal); Project administration (supporting); Software (lead); Writing – original draft (equal); Writing – review & editing (equal); **Oleg Kristanovski:** Conceptualization (supporting); Funding acquisition (equal); Project administration (equal); Supervision (supporting); Writing – review & editing (supporting). **Claus Emmelmann:** Project administration (equal); Supervision (lead).

REFERENCES

[1]M. Munsch, M. Schmidt-Lehr, and E. Wycisk, "Additive manufacturing report 2019, Hamburg," 2019.
[2]T. Wohlers, R. I. Campbell, O. Diegel, J. Kowen, N. Mostow, and I. Fidan, "Wohlers report 2022 3D printing and additive manufacturing: Global State of the Industry, Washington, DC" 2022.
[3]BIS Research, "Global metal 3D printing market (2020): Focus on type, application, and country level analysis—Analysis & forecast, 2019–2025" (2020).
[4]O. Rehme, *Cellular Design for Laser Freeform Fabrication* (Cuvillier Verlag, Göttingen, Germany, 2010), Vol. 4.
[5]M. Elsayed, M. Ghazy, Y. Youssef, and K. Essa, "Optimization of SLM process parameters for Ti6Al4V medical implants," Rapid Prototyp. J. **25**, 433–447 (2019).
[6]D. Srinivasan and K. Ananth, "Recent advances in alloy development for metal additive manufacturing in gas turbine/aerospace applications: A review," J. Indian Inst. Sci. **102**, 311–349 (2022).
[7]D. Beckers and G. Graf, "Effiziente Qualifizierung neuer Legierungen für die additive Fertigung," Lightweight Des. **12**, 48–51 (2019).
[8]W. E. King, A. T. Anderson, R. M. Ferencz, N. E. Hodge, C. Kamath, S. A. Khairallah, and A. M. Rubenchik, "Laser powder bed fusion additive manufacturing of metals; physics, computational, and materials challenges," Appl. Phys. Rev. **2**, 041304 (2015).
[9]C. Meier, R. W. Penny, Y. Zou *et al.*, "Thermophysical phenomena in metal additive manufacturing by selective laser melting: Fundamentals, modeling, simulation and experimentation," arXiv:1709.09510 (2017).
[10]G. Parivendhan, P. Cardiff, T. Flint, Ž. Tuković, M. Obeidi, D. Brabazon, and A. Ivanković, "A numerical study of processing parameters and their effect on the melt-track profile in laser powder bed fusion processes," Addit. Manuf. **67**, 103482 (2023).
[11]M. J. Ansari, D.-S. Nguyen, and H. S. Park, "Investigation of SLM process in terms of temperature distribution and melting pool size: Modeling and experimental approaches," Materials **12**, 1272 (2019).
[12]L. Meng, B. McWilliams, W. Jarosinski, H.-Y. Park, Y.-G. Jung, J. Lee, and J. Zhang, "Machine learning in additive manufacturing: A review," J. Mater. **72**, 2363–2377 (2020).
[13]C. Wang, X. P. Tan, S. B. Tor, and C. S. Lim, "Machine learning in additive manufacturing: State-of-the-art and perspectives," Addit. Manuf. **36**, 101538 (2020).
[14]J. Qin, F. Hu, Y. Liu *et al.*, "Research and application of machine learning for additive manufacturing," Addit. Manuf. **52**, 102691 (2022).
[15]J. Liu, J. Ye, D. Silva Izquierdo *et al.*, "A review of machine learning techniques for process and performance optimization in laser beam powder bed fusion additive manufacturing," J. Intell. Manuf. **1–4**, 77 (2022).
[16]G. Tapia, S. Khairallah, M. Matthews, W. E. King, and A. Elwany, "Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel," Int. J. Adv. Manuf. Technol. **94**, 3591–3603 (2018).
[17]L. Meng and J. Zhang, "Process design of laser powder Bed fusion of stainless steel using a Gaussian process-based machine learning model," J. Mater. **72**, 420–428 (2020).
[18]M. Kusano, S. Miyazaki, M. Watanabe, S. Kishimoto, D. S. Bulgarevich, Y. Ono, and A. Yumoto, "Tensile properties prediction by multiple linear regression analysis for selective laser melted and post heat-treated Ti-6Al-4V with microstructural quantification," Mater. Sci. Eng. A **787**, 139549 (2020).
[19]D. S. Ertay, S. Kamyab, M. Vlasea, Z. Azimifar, T. Ma, A. D. Rogalsky, and P. Fieguth, "Toward sub-surface pore prediction capabilities for laser powder bed fusion using data science," J. Manuf. Sci. Eng. **143**, 071016 (2021).
[20]G. O. Barrionuevo, J. A. Ramos-Grez, M. Walczak, and C. A. Betancourt, "Comparative evaluation of supervised machine learning algorithms in the prediction of the relative density of 316L stainless steel fabricated by selective laser melting," Int. J. Adv. Manuf. Technol. **113**, 419–433 (2021).
[21]H. O. Psihoyos and G. N. Lampeas, "Density-based optimization of the laser powder bed fusion process based on a modelling framework," Alloys **2**, 55–76 (2023).
[22]A. Dareh Baghi, S. Nafisi, R. Hashemi *et al.*, "A new approach to empirical optimization of laser powder bed fusion process for Ti6Al4V parts," J. Mater. Eng. Perform. **61**, 844 (2023).
[23]H. S. Park, D. S. Nguyen, T. Le-Hong, and X. Van Tran, "Machine learning-based optimization of process parameters in selective laser melting for biomedical applications," J. Intell. Manuf. **33**, 1843–1858 (2022).
[24]M. Gor, A. Dobriyal, V. Wankhede, P. Sahlot, K. Grzelak, J. Kluczyński, and J. Łuszczek, "Density prediction in powder bed fusion additive manufacturing: Machine learning-based techniques," Appl. Sci. **12**, 7271 (2022).
[25]A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. (O'Reilly, Beijing, 2019).
[26]T. M. Wischeropp, H. Tarhini, and C. Emmelmann, "Influence of laser beam profile on the selective laser melting process of AlSi10Mg," J. Laser Appl. **32**, 22059 (2020).
[27]T. M. Wischeropp, *Advancement of Selective Laser Melting by Laser Beam Shaping* (Springer Berlin Heidelberg, Berlin, 2021).
[28]M. Kuehne, K. Bartsch, B. Bossen, and C. Emmelmann, "Predicting melt track geometry and part density in laser powder bed fusion of metals using machine learning," Prog. Addit. Manuf. **8**, 47–54 (2023).
[29]W. H. Kan, L. N. S. Chiu, C. V. S. Lim, Y. Zhu, Y. Tian, D. Jiang, and A. Huang, "A critical review on the effects of process-induced porosity on the mechanical properties of alloys fabricated by laser powder bed fusion," J. Mater. Sci. **57**, 9818–9865 (2022).
[30]F. Pedregosa, G. Varoquaux, A. Gramfort *et al.*, "Scikit-learn: Machine learning in Python," arXiv:1201.0490 (2012).
[31]D. Herzog, K. Bartsch, and B. Bossen, "Productivity optimization of laser powder bed fusion by hot isostatic pressing," Addit. Manuf. **36**, 101494 (2020).
[32]A. V. Joshi, *Machine Learning and Artificial Intelligence* (Springer, Cham, 2020).