Readme – Supplementary data

Data set title: Supplementary Dataset for: *Identifying process windows of aluminum alloys in laser powder bed fusion using machine learning*

DOI of data set: https://doi.org/10.15480/882.15323

Handle of data set: https://hdl.handle.net/11420/56023

Involved persons:

Maxim Kuehne (1)*
maxim.kuehne@tuhh.de

Bastian Bossen (1)

Oleg Kristanovski (2)

Sebastian Jäger (2)

Ingomar Kelbassa (1)

(1) Institute for Industrialization of Smart Materials, Hamburg University of Technology, Harburger Schloßstraße 28, 21079 Hamburg, Germany

(2) Fehrmann Materials X GmbH, Stenzelring 19, Hamburg, 21107, Germany

** corresponding author*

Context: This dataset is supplementary material to the publication with the title " *Identifying process windows of aluminum alloys in laser powder bed fusion using machine learning*" of the International Congress on Applications of Lasers & Electro-Optics (ICALEO) 2025, 13 – 16 October 2025, Orlando/USA.

The corresponding paper investigates a novel machine learning (ML) based approach for the identification of process windows for aluminum based alloys. The metal additive manufacturing (AM) sector is limited by the scarcity of certified materials. Developing new materials is costly and time-consuming due to the complexity of AM processes where numerous parameters impact the part's quality. The goal of this study is using simulation and empirical data of AlSi10Mg and an alloy from the Al5000 series (AlMgty80) to train ML models to predict part density in laser powder bed fusion of metals. The models are used to predict process windows for a new alloy from the Al7000 series (AlZnty) to identify the potential of recognising process windows prior to empirical testing. By integrating simulated data, material influences are indirectly accounted for in the dataset. Various ML algorithms were trained and evaluated using root mean squared error (RMSE), R^2, and qualitative process map assessments. Results show that ML models based on AlSi10Mg and AlMgty80 data accurately predict the density for the respective materials, achieving a RMSE around 1% and R^2 close to 0.9. Predictions for AlZnty showed larger errors (RMSE 3 – 7 %, R^2 near zero), yet the qualitative evaluation of process maps show that high-density regions align with experimental data due to the model systematically underestimating the density. Overall, the methodology demonstrates the ability to identify process windows for novel alloys using simulations and ML models based on similar alloys, despite limitations in density prediction accuracy. Future work will involve training models including AlZnty data to iteratively enhance accuracy. Furthermore, limited data availability continues to challenge ML in AM, highlighting the need for extensive databases.

Content of this data set:

The content of the data set is listed below in the same order as the files are stored in the folder *Supplementary_Data.zip*. Files are highlighted with bold font and a short description is given.

- **Supplementary_Data_README.pdf**: This readme file
- Supplementary_Data.zip: ZIP-Folder containing supplementary data
  - 01_data_analysis
    - **01_Histograms of each input feature.png:** Histograms illustrating the statistical distribution of each input feature within the training dataset (AlSi10Mg + AlMgty80).
    - **02_Scatter plot of input variables against the target variable.png:** Scatter plots visualizing the relationship between individual input features and the target variable for the training dataset (AlSi10Mg + AlMgty80).
    - **03_Histograms of each input feature_transferability analysis.png:** Histograms illustrating the statistical distribution of each input feature within transferability analysis dataset (AlZnty).
    - **04_Scatter plot of input variables against the target variable_transferability analysis.png:** Scatter plots visualizing the relationship between individual input features and the target variable for the transferability analysis dataset (AlZnty).
    - **05_Heat maps of Spearman correlation coefficients before and after feature selection.png:** Summary of heat maps visualizing the spearman correlation coefficient prior to and after feature selection.
    - **06_Feature and Permutation Importance before feature selection.png:** Two bar diagrams showing the feature importance on the left and the permutation importance on the right.
    - **07_SHAP analysis of feature importance and impact.png:** A bar diagram and a beeswarm plot based on the SHAP analysis to visualize the feature importance and impact.
  - 02_hyperparamter
    - **01_Investigated hyperparameter ranges and selected values after random search.pdf:** Table that summarizes the investigated parameter ranges and lists the selected values used for the evaluation run of the models.
  - 03_model_training
    - **01_Residual plots of ML models after training without and with alloy_name.png:** Summary of residual plots for all model configurations of the AlMgty80+AlSi10Mg model. One plot for each model (RFR, MLP, XGBoost), for each configuration (with and without alloy_name) and for the investigated subset (training, test).
    - **02_AlZnty_Parity plots of ML models after training with alloy_name with and without AlZnty data in the training set.png:** Parity plots of predictions for the AlZnty dataset for all considered models (RFR, MLP, XGBoost) trained with alloy_name for different training data sets (including either no AlZnty data points or 70 AlZnty data points).
    - **03_AlZnty_Residual plots of ML models after training with alloy_name with and without AlZnty data in the training set.png:** Residual plots of predictions for the AlZnty dataset for all considered models (RFR, MLP, XGBoost) trained with alloy_name for different training data sets (including either no AlZnty data points or 70 AlZnty data points).