Acta Materialia 309 (2026) 122101

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Full length article

Data-selective machine learning framework (DSML) for defect-aware, interpretable yield-strength prediction for LPBF-fabricated AlSi10Mg alloys

Jeong Ah Lee [a], Yeon Woo Kim [b], Takayoshi Nakano [c], Hyomoon Joo [d], Jeong Min Park [b,e,f,*], Hyoung Seop Kim [a,g,h,*]

[a] Department of Materials Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea
[b] Division of Nano Materials, Korea Institute of Materials Science (KIMS), Changwon, Gyeongnam 51508, Republic of Korea
[c] Division of Materials and Manufacturing Science, Graduate School of Engineering, The University of Osaka, Osaka, 565-0871, Japan
[d] Materials Processing Technology Development Team, Hyundai Motor Group, Uiwang, 16082, Republic of Korea
[e] Center for 3D Printing Process, Korea Institute of Materials Science (KIMS), Changwon 51508, Republic of Korea
[f] Department of Advanced Future Convergence Materials, Korea University, Seoul 02841, Republic of Korea
[g] Graduate Institute of Ferrous & Eco Materials Technology, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea
[h] Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai, 980-8577, Japan

ARTICLE INFO

Keywords:
Laser powder bed fusion (LPBF)
Al–Si–Mg alloy
Machine learning
Yield strength
Porosity
Symbolic regression

ABSTRACT

The yield strength of laser powder bed fusion (LPBF) alloys remains challenging to predict owing to defect-driven variability that cannot be captured by conventional near-dense empirical equations. Here, we develop a data-selective machine learning (DSML) pipeline that integrates data-driven black-box modeling with physics-informed white-box modeling through symbolic regression to derive a defect-aware, interpretable closed-form equation. A dual-source dataset was constructed, including 44 fully labeled datasets (process parameters, microstructural features, and mechanical properties), and 111 process-only datasets containing porosity data. The DSML framework identifies critical descriptors and then embeds a porosity sub-model into a closed-form yield-strength equation, explicitly capturing the influence of process-induced defects. Validation was performed via AlSi10Mg fabricated using LPBF under six distinct conditions. The results revealed that the porosity-aware white-box model achieves a coefficient of determination of 0.90 and mean absolute error (MAE) of 9.51 MPa, outperforming both the black-box predictor and a widely used cell-size–based empirical relation (MAE = 41.98 MPa). The recovered terms align with known mechanisms (effective load-bearing reduction by pores and boundary-mediated strengthening) and preserve dimensional consistency, enabling the construction of process–design maps for defect-aware optimization. By internalizing defect effects in an interpretable equation and performing rigorous validation against independent experimental conditions, this work provides a reproducible, physics-consistent route to determining process–structure–property relationships for LPBF AlSi10Mg and a scalable foundation for incorporating additional strengthening mechanisms in next-generation LPBF materials.

1. Introduction

Laser powder bed fusion (LPBF) is a rapidly advancing additive manufacturing (AM) technique, which facilitates the fabrication of complex three-dimensional components by selectively melting metal powders layer-by-layer using a high-energy laser beam [1–4]. Guided by computer-aided design models, LPBF enables the production of near-net-shape parts with exceptional geometric precision and material efficiency. Among LPBF-processed metals, eutectic Al–Si alloys, particularly AlSi10Mg alloy, have been widely employed in aerospace and automotive applications owing to their attractive combination of low density, high strength-to-weight ratio, corrosion resistance, as well as their good printability [5–8]. In particular, the LPBF-induced microstructure in Al–Si alloys, such as the unique dislocation network decorated by Si particles, acts as the additional boundaries to mobile dislocations during plastic deformation, thereby providing LPBF-fabricated Al–Si alloys with excellent strain-hardening capability [9–11].

Despite advances in process control, porosity remains a persistent challenge in many AM-fabricated components, arising from gas

* Corresponding authors.
E-mail addresses: jmpark@kims.re.kr (J.M. Park), hskim@postech.ac.kr (H.S. Kim).

https://doi.org/10.1016/j.actamat.2026.122101
Received 15 November 2025; Received in revised form 7 February 2026; Accepted 9 March 2026
Available online 10 March 2026

entrapment, keyholing, lack-of-fusion voids, unstable melt pools resulting from suboptimal processing parameters, and build inconsistencies [12–15]. Similar to other AM-fabricated metals, LPBF-fabricated Al–Si alloys are susceptible to internal porosity, which degrades tensile properties. Research [16,17] has shown that increased porosity constrains the nucleation of new voids, promotes rapid coalescence of existing pores, and accelerates fracture through void sheeting. Additionally, strain incompatibility at melt pool boundaries is a dominant fracture origin, and reducing the density of these boundaries can significantly enhance ductility. However, the direct influence of porosity on yield strength in LPBF-fabricated Al–Si alloys remains insufficiently characterized. Therefore, developing a direct predictive framework that comprehensively captures the influence of porosity on yield strength is essential.

Recent advances in data-driven machine learning (ML) have introduced a transformative paradigm that integrates the empirical accuracy of physical modeling with the predictive strength of advanced analytical techniques [18–22]. In particular, symbolic regression (SR) has emerged as a powerful tool for elucidating the internal structure of predictive models, enabling the identification of explicit and interpretable relationships between input features and target variables [23–26]. For instance, Lee et al. [27] applied SR to derive a yield-strength equation

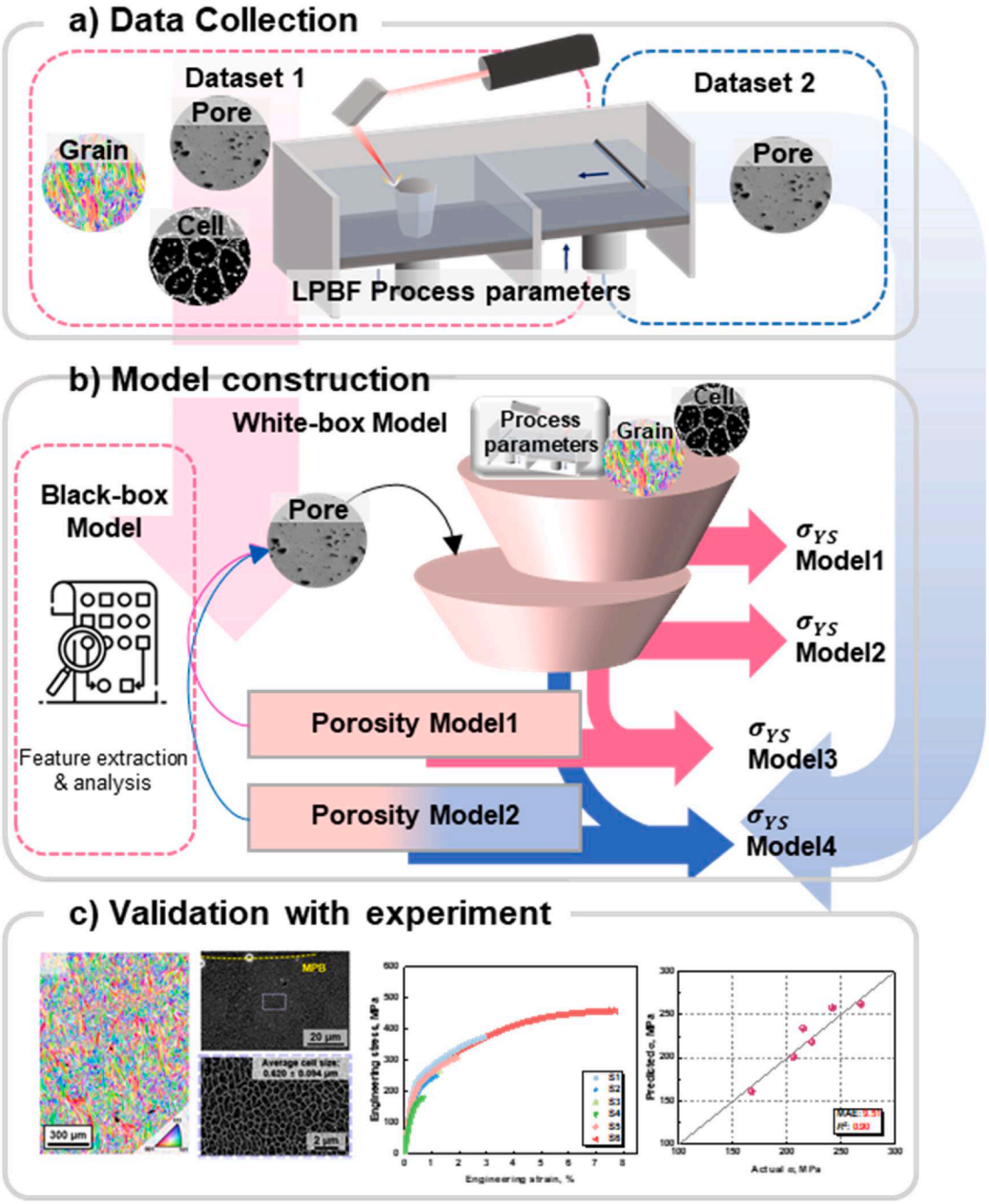

Fig. 1. Schematic of the process for applying DSML to predict yield strength. Process yield strength prediction is structured into three key stages: a) construction of a dual-source dataset; b) implementation of both black-box and four white-box ML models, with a focus on yield strength and porosity prediction; and c) model validation and comparison against empirical expression and experimental data.

based on the grain-boundary sliding mechanism, underscoring the importance of models that not only fit data accurately but also capture physical mechanisms beyond the training dataset. Similarly, Huang et al. [28] used SR in conjunction with optimized descriptors to model the thermal conductivity of polymers, yielding analytical expressions suitable for rapid preliminary screening of high thermal conductivity candidates. While conventional ML approaches based on "black-box" models are effective for predictive tasks, interpretable models that can clarify high-dimensional feature interactions offer greater transparency, reliability, and scientific insight.

Notably, the effectiveness of ML-driven process optimization in AM depends on the availability of large, high-quality datasets, which is often limited owing to low production volumes and the substantial cost of data collection [29–32]. Although numerical simulations can supplement experimental data, they are computationally intensive and often fail to capture accurate quantitative relationships between process parameters and material properties. Thus, recent studies have adopted various techniques to overcome the constraints imposed by small datasets. For instance, Pei et al. [33] optimized the squeeze-casting process for mine fuel tank partition castings using an ultra-small dataset of only 25 samples. Hollmann et al. [34] introduced the tabular prior-data fitted network, which outperformed conventional methods on datasets containing up to 10,000 samples. Tang et al. [35] used a U-Net-based model to optimize the scanning path in LPBF using only 30 samples. Despite these advances, most approaches still rely on complete datasets including compositional, microstructural, and property-related features, which are not always available. Furthermore, when data augmentation techniques are applied to compensate for small sample sizes, the resulting predictive accuracy often falls short of that achieved through experimental validation. Therefore, the development of customized ML strategies that ensure both predictive accuracy and interpretability is imperative for advancing data-driven process optimization in materials science.

Herein, this study was aimed at establishing a data-selective machine learning (DSML) framework for predicting the yield strength of LPBF-fabricated AlSi10Mg by identifying and leveraging the most informative data within a constrained dataset. A dual-source dataset was constructed, consisting of 44 fully labeled data points (Dataset 1) with complete microstructural and mechanical properties and 111 partially labeled data points (Dataset 2) containing only process parameters and porosity data. To effectively utilize both datasets, a predictive framework was developed, integrating the quantitative power of the larger dataset with the qualitative reliability of the fully labeled dataset. Fig. 1 schematically illustrates the proposed methodology: Dataset 1 includes LPBF process parameters along with microstructural features such as grain size, cell size, and porosity. Dataset 2 contains the same process parameters and porosity data but excludes grain and cell size. The DSML framework integrates two types of ML models: (1) a black-box model that captures the influence of porosity on yield strength, and (2) a transparent, interpretable white-box model that not only elucidates the role of porosity but also incorporates additional data sources to offer quantitatively accurate predictions. Comprehensive testing revealed that the proposed hybrid strategy demonstrates excellent predictive performance and can guide the optimization of mechanical properties of LPBF-fabricated alloys.

2. Experimental and methods

2.1. Data preprocessing

A total of 44 data combinations from published literature were compiled for Dataset 1 (details in Supplementary Data 1), encompassing key LPBF process parameters: laser power (*LP*), scan speed (*SS*), layer thickness (*LT*), hatch spacing (*HS*), preheating temperature (T_{pre}), and volumetric energy density (*VED*), calculated as

$$VED\left(J/mm^3\right) = \frac{LP}{HSxSSxLT}. \quad (1)$$

In addition, microstructural features such as cell size (d_{cell}), grain size (d_{grain}), porosity, and yield strength (σ_{YS}) were included to support comprehensive structure–property modeling. Only complete entries without missing values were used for both the black-box and white-box ML models, with most input parameters sourced from prior reports [9, 36,37]. Dataset 2 consisted of 111 data combinations derived from experimental measurements (Supplementary Data 2), including the same process parameters and porosity data but excluding microstructural and mechanical properties. To ensure high data quality, rigorous data preprocessing procedures were applied to both datasets, including handling of missing values and normalization, thereby facilitating reliable model training and evaluation. Although additional factors, such as dislocation density, Orowan strengthening from the Si phase, and other microstructural contributions to mechanical properties, could provide valuable insights, their inclusion was restricted owing to the limited dataset size. Incorporating a broader set of features would increase the dimensionality of the input space, potentially invoking the "curse of dimensionality" [38,39] without a proportional increase in data volume, thereby reducing learning efficiency and model interpretability. To balance model robustness and complexity, we focused on process parameters and primary microstructural features (d_{cell} and d_{grain}) as the most relevant inputs for yield-strength prediction.

2.2. ML modeling

The black-box ML models were implemented using the scikit-learn library [40], a widely adopted open-source Python-based ML framework. Rigorous feature engineering was performed to enhance the representation of input variables, and thus, the model performance. Model robustness and generalization were ensured through systematic 10-fold cross-validation, enabling reliable assessment of predictive capability across varying data partitions. Hyperparameters were optimized using an automated grid search strategy to efficiently determine the optimal model configurations. Detailed hyperparameter settings for each black-box model are provided in Supplementary Table S1. To ensure unbiased performance evaluation, the dataset was randomly divided into training and test sets, with 10% of the data held out as an independent test set. The test set was not used during model training or hyperparameter tuning. The remaining data were used for model development, within which 10-fold cross-validation was performed for model selection. Model performance is reported using MAE and R^2 evaluated on the held-out test set.

After training the black-box model and identifying the important features among the input parameters, the SR package *PySR* [41] was used to fit closed-form analytical expressions for the yield strength, which were then used in the white-box model. The fitting process is illustrated in Fig. 2a and b. First, as shown in Fig. 2 a, σ_{YS} Model 1 was developed using SR based on Dataset 1, excluding porosity as an input variable. Next, σ_{YS} Model 2 was constructed by including porosity in the SR, enabling direct comparison with Model 1 to quantify the effect of porosity on yield-strength prediction. To further refine the porosity term in Model 2, two porosity prediction models were introduced. Porosity Model 1 was trained only using Dataset 1. Porosity Model 2 was built using both Datasets 1 and 2, leveraging the larger dataset to enhance generalizability. As illustrated in Fig. 2b, the porosity expressions derived from these models were substituted back into the σ_{YS} Model 2 equation, yielding two new formulations: σ_{YS} Model 3, using Porosity Model 1, and σ_{YS} Model 4, using Porosity Model 2. Lastly, Models 3 and 4 were compared to evaluate how the inclusion of additional porosity data, despite the absence of microstructural and mechanical properties in Dataset 2, affects the accuracy and robustness of yield-strength prediction.

Fig. 2c presents the distributions of three input features (*LP, LT*, and

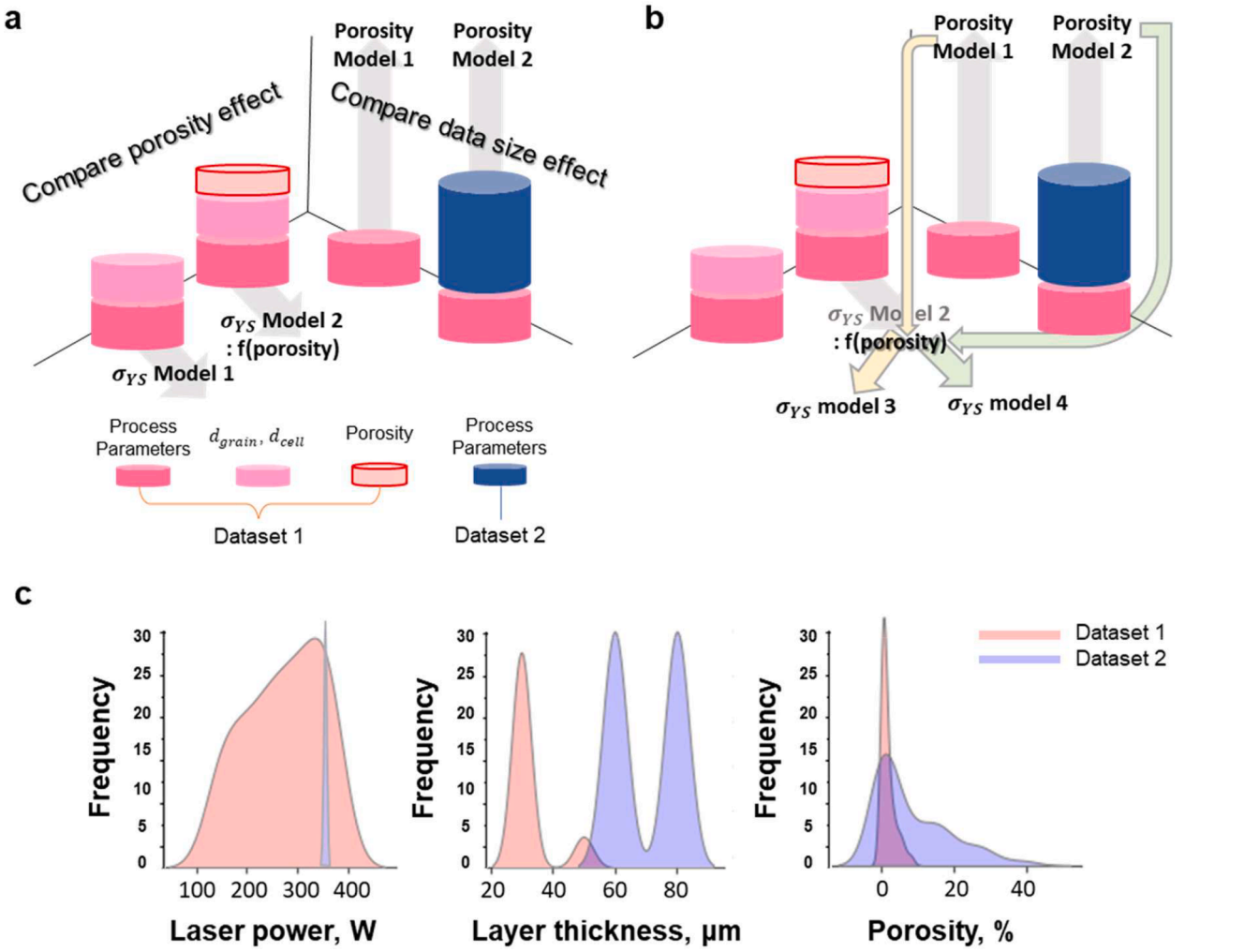

Fig. 2. Dataset composition distribution demonstrating correlations between input parameters and two datasets. a) Dataset composition used for σ_{YS} Models 1 and 2 and Porosity Models 1 and 2; b) Construction of Yield Strength Models 3 and 4 via Integration of Porosity Models 1 and 2; c) Distribution of all AlSi10Mg samples in the Dataset 1 marked in red and Dataset 2 marked in blue on the values of laser power, layer thickness, and porosity.

porosity) for Datasets 1 (red) and 2 (blue). These kernel density estimates highlight the distinct distribution patterns between the two datasets, supporting their integration in Porosity Model 2. For *LP*, Dataset 2 exhibits a highly concentrated distribution around a single power setting, whereas Dataset 1 spans a broader range. For *LT*, Dataset 1 primarily covers 20–60 μm, while Dataset 2 is concentrated around 45–90 μm, with minimal overlap. For porosity, Dataset 1 is narrowly distributed within the 0–10% range, whereas Dataset 2 extends over a wider range, providing greater diversity in the target variable. Relying solely on Dataset 2 would thus provide limited coverage of the critical parameter space, especially in poorly represented regions. Therefore, both Datasets 1 and 2 were incorporated into Porosity Model 2 to ensure broader coverage of the feature space and enhance model generalizability across a wider range of process conditions and porosity values. To further clarify, the multidimensional process-window coverage of Datasets 1 and 2 is explicitly compared in Supplementary Fig. S1, illustrating the expansion of the effective LPBF process space.

The optimal expressions were identified using the following steps: first, a diverse set of candidate expressions was generated across various predefined complexity levels. Complexity was quantitatively evaluated based on the number of mathematical operators, constants, and input variables in each expression. Model simplicity was determined as inversely related to complexity, whereas predictive performance was assessed using a loss function. Expressions exhibiting both high complexity and high loss were systematically discarded to balance accuracy and interpretability.

SR employs genetic algorithms to evolve analytical expressions through iterative exploration of the expression space. As this process is inherently exhaustive, the search space grows exponentially with the number of input variables and operators. To constrain the search and improve computational efficiency, the SR input space was restricted to a physically meaningful descriptor set, rather than the full combinational space of all possible variables. Within this restricted space, SR was performed using the same comprehensive descriptor set employed in the black-box models, including process parameters, microstructural features, and porosity-related variables. Specifically, for σ_{YS} Models 1 and 2, SR fitting was performed using an eight- or nine-dimensional process–structure descriptor set and a one-dimensional yield-strength output. In contrast, Porosity Models 1 and 2 were constructed using a reduced five-dimensional input space including process-related descriptors and a one-dimensional porosity target. The mathematical operations available during model construction included addition (+), subtraction (−), multiplication (×), and division (÷), as well as nonlinear functions such as logarithm (log), exponential (exp), and square root (sqrt). Each model was optimized over 3000 iterations. All datasets were treated uniformly during training to promote unbiased learning of process–porosity relationships, and no source-dependent weighting or domain adaptation techniques were applied.

2.3. Experimental procedure

Spherical powder of AlSi10Mg alloy (EOS GmbH, Germany) in a particle size range of 10–55 μm was used as the raw material for the LPBF process. To construct Dataset 2 and validate the performance of our prediction model, the cuboidal samples (10 mm x 10 mm x 10 mm) were fabricated on an Al6061 substrate using a commercial LPBF system (M290, EOS GmbH, Germany) at 350 W and 370 W, while varying scan speed (1200–4000 mm/s in 50 mm/s increments, including 1800 mm/s), hatch distance (100–120 μm), and layer thickness (30 and 60–80 μm), leading to a total of 228 build conditions. All printing was performed in an island pattern with a rotation angle 67° between the n^{th} and $(n+1)^{th}$ layers under an 99.7% argon atmosphere.

For microstructural analysis, the cross-section surface of the cuboidal samples was finely polished using SiC abrasive papers, followed by colloidal silica finishing. In order to characterize the grain size and morphologies, electron backscatter diffraction (EBSD) analysis was performed in field-emission scanning electron microscopy (FE-SEM, JSM-7900F, JEOL Ltd., Japan) system equipped with EBSD detector and TSL OIM Data Analysis 7 software. The LPBF-driven cell structure of the AlSi10Mg samples was characterized by SEM analysis on the sample surface etched using Keller's reagent.

For mechanical testing, plate-type tensile specimens (gauge length: 25 mm, width: 6 mm) following ASTM E8 standard were machined from the as-printed samples (100 mm × 15 mm × 5 mm; height × width × thickness). Note that uniaxial tensile tests were performed along the tensile direction parallel to the building direction. Tensile tests were conducted on an INSTRON 4204 (INSTRON Co., USA) at a strain rate of 10^{-3} s^{-1} at room temperature. An extensometer (Epsilon Co., USA) was used to measure accurate elongation of the tensile specimens.

3. Results and discussion

3.1. Performance of black-box models

To assess the influence of porosity on yield strength, the predictive accuracy of various black-box models was compared using mean absolute error (MAE) and the coefficient of determination (R^2) on the test set, defined using Eqs. (2) and (3), respectively. Additionally, advanced statistical techniques, including bootstrapping, were used to estimate the confidence intervals of the predictive models, offering a more robust evaluation of their reliability and generalization.

$$\mathrm{MAE} = \frac{1}{n}\sum |X_i - Y_i|, \tag{2}$$

$$R^2 = 1 - \frac{\sum (X_i - Y_i)^2}{\sum (\overline{Y} - Y_i)^2}, \tag{3}$$

where X_i, Y_i, and $\overline{Y}$ denote the predicted, actual, and mean of actual values, respectively.

We compared the yield-strength prediction performance of various black-box models, including the least absolute shrinkage and selection operator (LASSO), lasso least angle regression (LLAR), k-nearest neighbors (KNN), ridge regression (Ridge), Bayesian ridge (BR), linear regression (LR), and ensemble tree regressor (ETR) models, as shown in

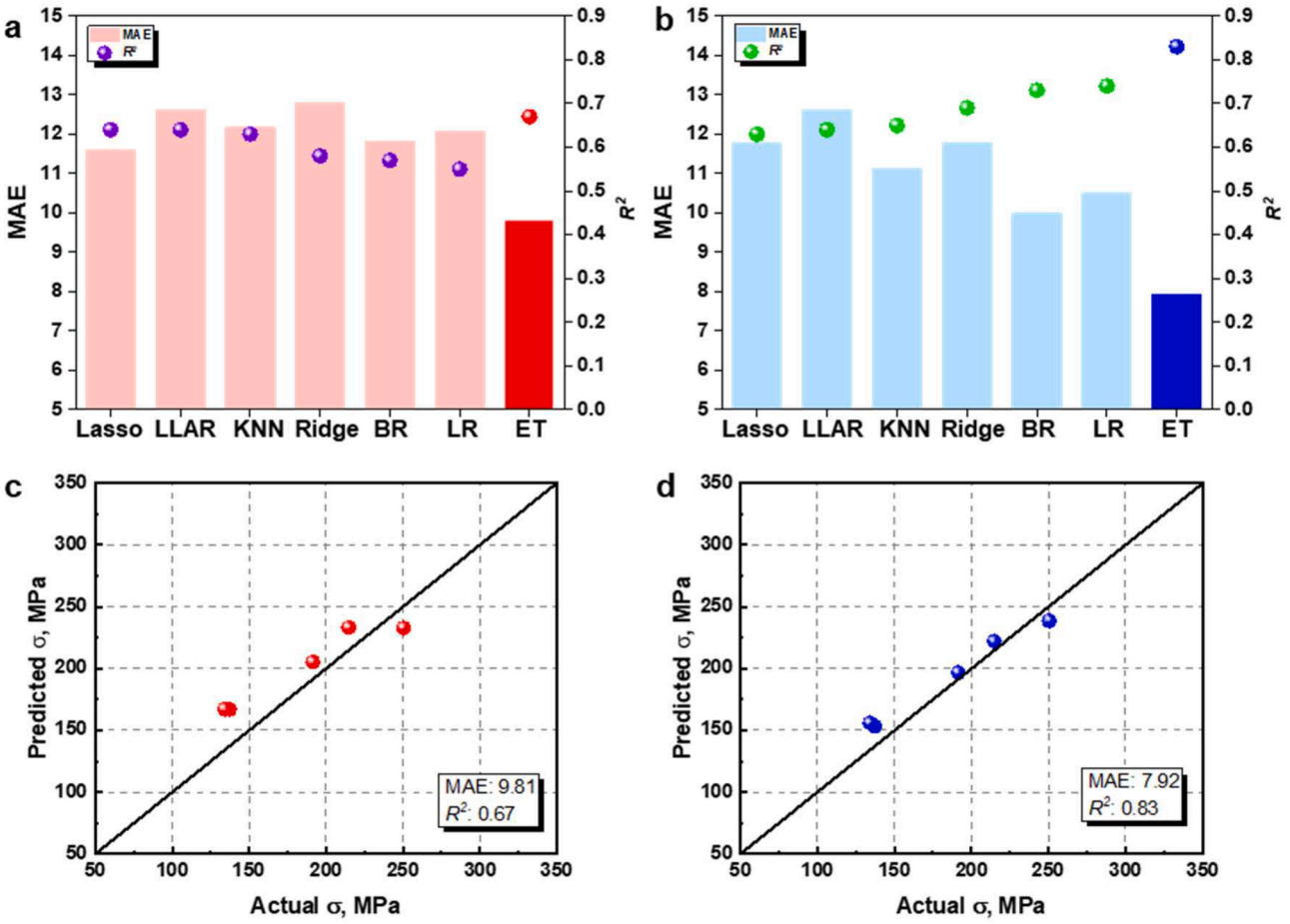

Fig. 3. Performance evaluation of black-box models for yield strength prediction. Mean absolute error (MAE) and coefficient of determination (R^2) comparison of black-box models trained (a) without porosity and (b) with porosity as an input feature. (c) Predicted versus experimental yield strengths for the ETR model on the test set, corresponding to (a) the porosity-excluded model and (b) the porosity-included model.

Fig. 3a and b. When the porosity is excluded (Fig. 3a), the ETR model outperforms all others (MAE: 10.56, R^2: 0.705), whereas the KNN model exhibits the lowest performance (MAE: 13.73, R^2: 0.662). When porosity is incorporated as an input feature (Fig. 3b), the predictive accuracy of all models improves significantly. The ETR model again exhibits the best performance (MAE: 6.32, R^2: 0.865), followed by LR and BR, confirming the importance of porosity in yield-strength prediction. These results highlight the effectiveness of ensemble models and critical role of porosity in improving predictive reliability for LPBF-fabricated AlSi10Mg alloys.

Fig. 3c and d show scatter plots of predicted versus actual yield strength for the test set. When porosity is excluded (Fig. 3c), the predictions deviate significantly from the line of unity, with MAE of 9.81 MPa and R^2 of 0.67, indicating moderate predictive accuracy. In contrast, when porosity is incorporated as an input (Fig. 3d), the predictions align well with the line of unity, indicating a strong correlation between predicted and actual values. The models achieve a lower MAE of 7.92 MPa and a higher R^2 of 0.83, reflecting enhanced predictive performance and generalizability. The generalization behavior of the ET model for both datasets was further evaluated using learning curves and validation curves, which are provided in Supplementary Figs. S2-S3.

3.2. Feature importance analysis using the ETR model

To evaluate the influence of individual descriptors on yield-strength prediction, the ETR model was used to perform feature importance

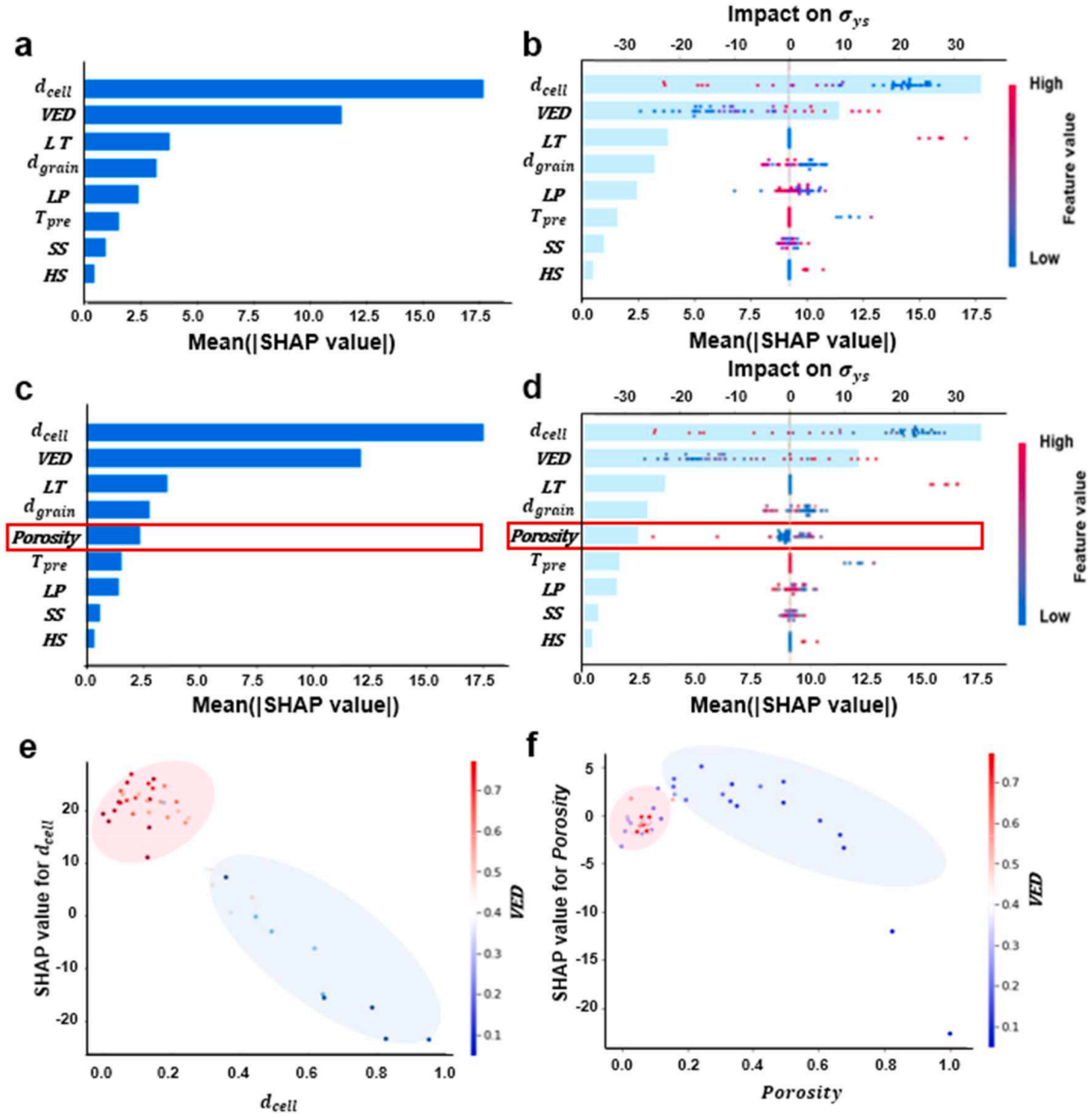

Fig. 4. Feature importance and SHAP analysis based on the ETR model. (a) Feature importance derived from the ETR model trained with porosity included, and (b) corresponding SHAP summary plot highlighting the impact of each feature on yield strength prediction.(c) Feature importance derived from the ETR model excluding porosity, and (d) corresponding SHAP summary plot.(e) SHAP interaction visualization between d_{cell} and *VED*, and (f) SHAP interaction between porosity and *VED*, illustrating the combined influence of these parameters on yield strength prediction.

analysis and interpretability assessment using Shapley additive explanations (SHAP) [42], as illustrated in Fig. 4a–d. SHAP values were used to quantify the contribution of each input feature to the model output, providing insights into how variations in specific descriptors affect the predicted yield strength. This approach allowed us to identify the most influential features and clarify both the direction and magnitude of their influence on the model predictions.

When porosity is excluded (Fig. 4a), feature importance analysis identifies d_{cell} and *VED* as the most influential predictors of yield strength, whereas *SS* and *HS* exhibit the lowest relative importance. To further investigate the individual effect of each feature, SHAP main values are presented as a summary plot (Fig. 4b). Smaller d_{cell} values are associated with higher predicted yield strength, attributable to the enhanced dislocation-blocking effect of the cellular boundaries. In LPBF-fabricated Al–Si alloys, these boundaries are typically enriched with Si-rich eutectic phases and exhibit high dislocation densities [43]. Although they possess low misorientation angles, cell boundaries can still hinder dislocation motion, particularly when decorated with segregated solute atoms or fine precipitates—analogous to the mechanism of grain-boundary segregation strengthening. Conversely, lower *VED* values correspond to reduced yield strength, likely owing to insufficient energy input during processing. Inadequate melting leads to poor fusion between layers, incomplete densification, and inadequate cellular-structure formation [44,45]. These deficiencies can result in increased porosity, weak interlayer bonding, and diminished dislocation density, ultimately reducing the yield strength of the as-built material [46,47]. Similarly, d_{grain} exhibits a negative correlation with yield strength, consistent with the role of grain boundaries in obstructing dislocation movement. In addition to microstructural descriptors, *LT* also exhibits a non-negligible influence on yield strength, as reflected in the SHAP analysis. An increase in *LT* is associated with a higher fraction of fine equiaxed grains and a reduction in the length of columnar grains, leading to an increased density of grain boundaries [48]. This microstructural refinement enhances grain-boundary strengthening and can partially offset strength degradation associated with other processing-induced defects. It should be noted, however, that although *LT* contributes meaningfully to yield strength prediction, the range of *LT* values explored in the dataset is relatively narrow compared to other descriptors. As a result, the SHAP dependence analysis for *LT* does not exhibit as clear a proportional or inverse trend as features with broader variability. These results indicate that when the porosity is not input, *VED* (among processing parameters) and cell size (among microstructural features) are the most critical factors influencing the yield strength of LPBF-fabricated AlSi10Mg alloys.

When porosity is included as an input feature (Fig. 4c), it is identified as one of the most influential predictors of yield strength, with a contribution comparable to that of d_{grain}. The strong weighting on porosity reflects its direct influence on mechanical performance: higher porosity acts as a stress concentrator, promoting early failure under load [49,50]. Its inclusion substantially alters the feature hierarchy observed in the porosity-excluded model (Fig. 4a), indicating that porosity effectively encapsulates critical defect-related information that would otherwise be distributed among multiple process parameters. To further examine the influence of each feature on the model output, SHAP values are presented as a summary plot (Fig. 4d). As porosity and yield strength do not exhibit a clear linear relationship in the training data, interaction analysis provides critical insights into how porosity influences the model predictions in combination with other features. Fig. 4e and f show SHAP interaction plots involving the two most salient factors, d_{cell} and porosity, each evaluated relative to *VED*, a key process parameter.

The SHAP interaction analysis between d_{cell} and *VED* (Fig. 4e) reveals that combinations of low d_{cell} and high *VED* result in the most positive SHAP values, indicating a synergistic enhancement in the predicted yield strength. This behavior is attributable to the formation of refined cellular microstructures under sufficient energy input, which promotes complete melting and uniform solidification [51,52]. These processing conditions promote grain-boundary strengthening and reduce defect-induced stress concentrations, thereby improving the resistance of the material to dislocation motion. In contrast, large d_{cell} values under low *VED* conditions correspond to strongly negative SHAP values, suggesting inadequate melting, increased porosity, and the development of coarse microstructures that are less effective at obstructing dislocation motion.

A similar analysis between porosity and *VED* (Fig. 4f) demonstrates that while higher *VED* typically enhances yield strength, by promoting melt pool stability and reducing lack-of-fusion defects, this positive effect is substantially diminished in the presence of increased porosity. High-porosity samples consistently exhibit strongly negative SHAP values, even under conditions of increased energy input, highlighting the dominant negative impact of porosity on yield strength. In contrast, the most favorable SHAP values are observed for samples combining low porosity with high *VED*, indicating that optimal yield strength is achieved when defect formation is rigorously suppressed.

In summary, these interaction plots highlight the necessity of optimizing multiple interdependent parameters rather than relying on single-variable adjustments. In particular, achieving high yield strength in LPBF-fabricated parts requires not only sufficient energy input but also effective suppression of defects, especially porosity.

3.3. Effect of porosity integration into SR-Based yield-strength models

To facilitate interpretable ML in yield-strength prediction, SR was employed to derive analytical expressions that reflect the underlying physical mechanisms. SR was chosen for its ability to generate transparent, closed-form equations, consistent with the objectives of explainable artificial intelligence. Unlike conventional black-box approaches, SR provides physically grounded models that not only offer predictive capability but also yield mechanistic insights into the governing deformation phenomena.

Applying SR to σ prediction without the porosity input feature (σ_{YS} Model 1) results in the following most accurate expression:

$$\sigma = 345 + \frac{d_{cell}}{LT} \times \left(-6.72 \times 10^{3}\right), \tag{4}$$

where d_{cell} and *LT* are expressed in micrometers. This expression reveals an inverse relationship between yield strength and cell size, corroborated by the SHAP analysis of the black-box model. Additional candidate models presented in Fig. 5a exhibit similar trends, reinforcing the central role of cell morphology in strengthening behavior.

Conversely, when porosity is included as a feature (σ_{YS} Model 2), SR identifies a more complex relationship with yield strength:

$$\sigma = 1.86 \times [LT - Porosity \times \{(Porosity - 5.65) + 0.11 \times (d_{cell} - HS) - 138.8\}]. \tag{5}$$

Here, in addition to d_{cell} and *LT*, porosity emerges as a critical determinant of σ, as evidenced by the consistent trends observed across alternative models presented in Fig. 5b. To ensure dimensional consistency, the Burgers vector of AlSi10Mg ($b = 0.286$ nm) was incorporated:

$$\sigma = 1.86 \times \left[\frac{\alpha}{b} LT - (100 \times Porosity) \times \left\{((100 \times Porosity) - 5.65) + 0.11 \times \frac{\alpha}{b}\left(10^{3} \times d_{cell} - HS\right) - 138.8\right\}\right], \tag{6}$$

where α denotes a constant parameter equaling 0.286×10^{-3}. Integrating SHAP values, Eq. (6) validates the continued inverse relationship between σ and d_{cell}, while explicitly integrating the dominant effect of porosity. Notably, even at constant grain or cell size, variations in porosity significantly modulate yield strength, reflecting the sensitivity of LPBF-fabricated materials to internal defects.

To evaluate the efficacy of each model in predicting the yield

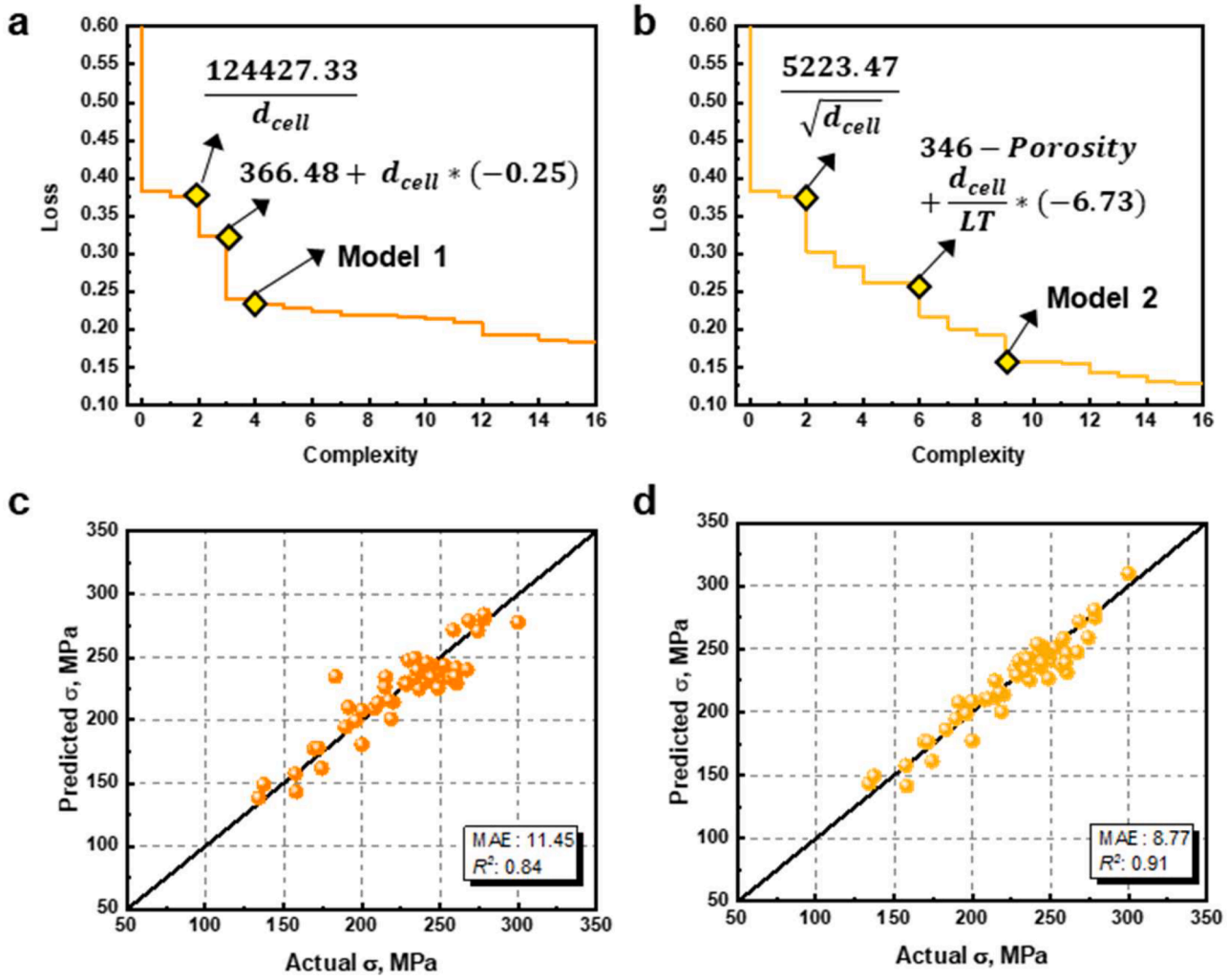

Fig. 5. Prediction results for the models versus the actual values of dataset 1. SR results of the variety of calculation loss of (a) models excluding the porosity feature and (b) models incorporating porosity as an input variable with the complexity. Prediction versus actual yield strength for (c) σ_{YS} Model 1 and (d) σ_{YS} Model 2.

strength, the predictive performances of the best-fitting models are shown in Fig. 5c and d. For the model excluding porosity (σ_{YS} Model 1), the derived expression achieves MAE of 11.45 and R^2 of 0.84, as indicated by the scatter plot of predicted versus actual values. In comparison, the model including porosity (σ_{YS} Model 2) achieves an improved MAE of 8.77 and R^2 of 0.91, indicating enhanced prediction fidelity. Notably, these expressions were selected directly from the complexity–error space and represent the most accurate interpretable equations discovered via SR.

Collectively, these results emphasize the superior performance of the porosity-informed SR model in capturing the multivariable complexity of yield-strength behavior. The transparent nature of the derived expressions enhances physical interpretability while retaining predictive accuracy. These findings underscore the efficacy of SR in developing compact, physically meaningful models, particularly in the context of AM, where defect-related variabilities critically influence mechanical performance.

3.4. Hierarchical SR for porosity-driven yield-strength prediction

To further improve the generalizability of the yield-strength prediction framework, the porosity component in the SR-derived model was refined through the development of dedicated porosity prediction equations. Two SR models were constructed to predict porosity: Porosity Model 1, trained exclusively on Dataset 1, and Porosity Model 2, developed using both Datasets 1 and 2. Although Dataset 2 lacked microstructural and mechanical property data, its inclusion enabled a broader representation of process-induced variability while increasing the overall data volume, thereby enhancing both the generalizability and statistical reliability of the porosity model.

The predictive expressions obtained from the two porosity models were then substituted into the original yield-strength model (σ_{YS} Model 2), resulting in two updated formulations: σ_{YS} Model 3, incorporating Porosity Model 1, and σ_{YS} Model 4, incorporating Porosity Model 2. This strategy enabled the evaluation of whether an expanded dataset, even with reduced feature fidelity, could improve yield-strength prediction through more accurate porosity estimation.

For Porosity Model 1, SR yields the following expression:

$$Porosity = \frac{SS}{0.22 \times LP^3}. \tag{7}$$

As illustrated in Fig. 6a, several symbolic expressions can be identified along the model complexity–loss curve for porosity prediction. Notably, the intermediate models, annotated with expressions involving *VED*, consistently demonstrate an inverse relationship between porosity and *VED*. This trend aligns with established physical understanding: increasing *VED*, either by increasing *LP* or decreasing *SS*, leads to more stable melt pool dynamics and enhanced consolidation, thereby

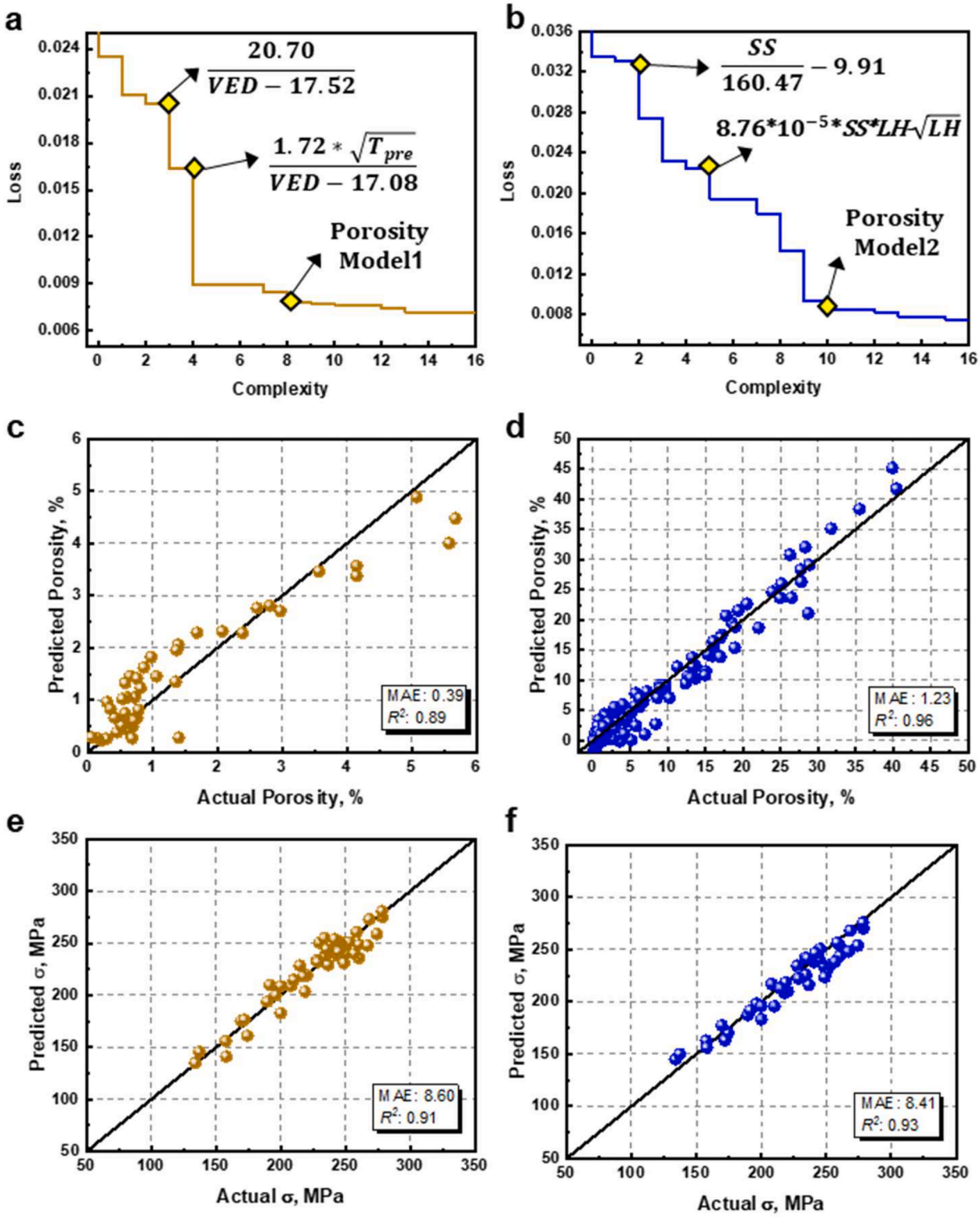

Fig. 6. SR model performance for porosity prediction and its integration into hierarchical σ_{YS} models. SR results of the variety of calculation loss of (a) porosity Model 1 and (b) porosity Model 2 with the complexity. Prediction versus actual yield strength for (c) porosity Model 1, and (d) porosity Model 2. Prediction versus actual yield strength for (e) σ_{YS} Model 3, and (f) σ_{YS} Model 4.

reducing porosity. Although Eq. (7), corresponding to the lowest-loss model (Porosity Model 1), does not explicitly include *VED*, it encodes a physically similar trend. Specifically, porosity increases with higher *SS* and decreases sharply with higher *LP*, mirroring the *VED* dependence defined in Eq. (1).

The coherence between Eq. (7) and *VED*-based models substantiates its physical validity and supports its selection as the optimal model based on the dual criteria of accuracy and interpretability. Moreover, the use of only two process parameters, *SS* and *LP*, enhances the model's simplicity, facilitating its practical implementation for process

optimization and control in laser-based AM.

When Dataset 2 is incorporated into the modeling framework, SR yields a more complex expression for porosity:

$$Porosity = (2.08 - 0.011 \times SS) \times \left(5.15 + \frac{SS \times LT}{LP \times (HS - 156.11)}\right). \quad (8)$$

This formulation, derived from the dataset combining microstructurally rich (Dataset 1) and process-only (Dataset 2) inputs, is associated with Porosity Model 2 and effectively reflects the interaction between *SS, HS, LP*, and *LT* (Fig. 6b). As shown in Fig. 6d, the model achieves an MAE of 1.23, moderately higher than that of Porosity Model 1 (MAE = 0.39, Fig. 6c). However, its R^2 (0.96) is considerably higher than that of Porosity Model 1 ($R^2 = 0.89$), indicating greater consistency and stability across a broader range of process conditions. This result suggests that the inclusion of Dataset 2, despite its lack of microstructural and mechanical detail, enhances the statistical reliability and generalizability of the porosity model.

To evaluate the influence of improved porosity estimation on yield-strength prediction, the two porosity models were embedded within the original σ_{YS} Model 2 framework, yielding σ_{YS} Model 3 (based on Porosity Model 1) and σ_{YS} Model 4 (based on Porosity Model 2). The predictive performance of these extended models is presented in Fig. 6e and f. Although both models demonstrate high predictive fidelity, σ_{YS} Model 4, constructed using the porosity prediction from the combined dataset, achieves a higher accuracy (MAE: 8.41, $R^2 = 0.93$) compared with σ_{YS} Model 3 (MAE = 8.60, $R^2 = 0.91$). These results suggest that the broader data coverage afforded by Dataset 2 compensates for the absence of explicit microstructural features, enabling more robust and generalizable yield-strength predictions.

To further enhance the practical applicability of the yield-strength prediction framework, the porosity component was refined through a hierarchical modeling strategy. The enhanced performance of σ_{YS} Model 4, which incorporates Porosity Model 2, underscores the benefit of leveraging process-only datasets to improve model generalizability in data-limited space. By embedding intermediate physical descriptors, such as porosity, into a hierarchical SR framework, the proposed methodology enables accurate and interpretable prediction of complex mechanical properties while maintaining physical relevance through explicit functional relationships with process variables. These findings emphasize the importance of data diversity and structured model decomposition in advancing predictive frameworks for AM, particularly for bridging process–structure–property relationships when micro-

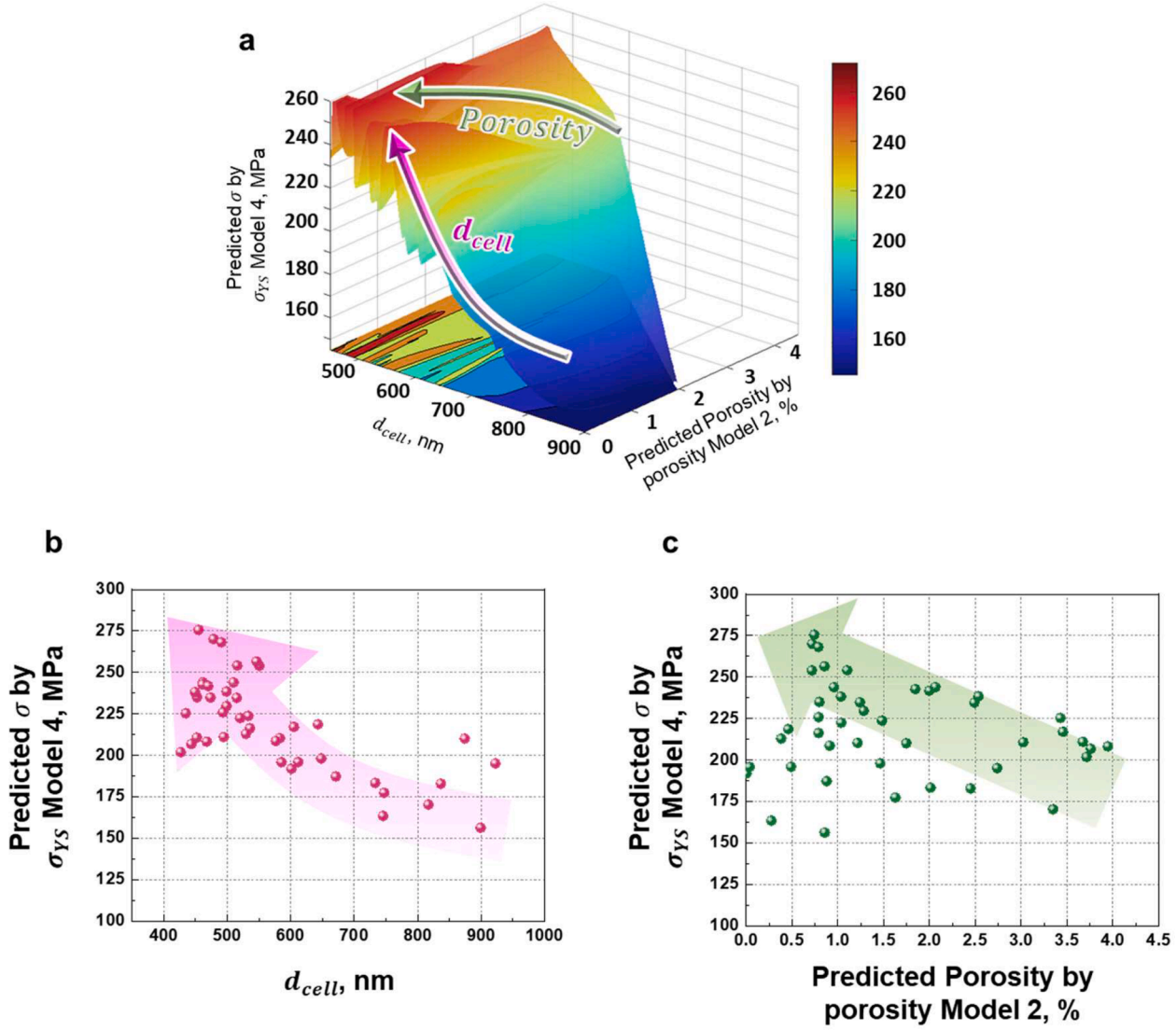

Fig. 7. Global response surface analysis of the coupled process–porosity–strength relationship in the final white-box model. (a) Three-dimensional response surface of yield strength predicted by σ_{YS} Model 4 as a function of d_{cell} and predicted porosity from Porosity Model 2. Relationship between yield strength predicted by σ_{YS} Model 4 and (b) d_{cell} and (c) predicted porosity from Porosity Model 2.

structural data are limited. Since porosity represents a dominant process-induced structural defect, the following section examines how its interaction with microstructural features governs yield strength through the global structure–porosity–strength response surface.

3.5. Global response surface stability of the structure–porosity–strength relationship

Fig. 7a presents the global response surface of yield strength obtained from the final coupled σ_{YS} Model 4, as a function of d_{cell} and predicted porosity, in which Porosity Model 2 is explicitly embedded. Because the σ_{YS} Model 4 requires a complete set of process and microstructural descriptors, Dataset 1 was used for this analysis.

Among the process-related variables included in σ_{YS} Model 4, *LT* and *HT* exhibit limited variability in Dataset 1, which restricts their suitability for constructing meaningful global response surfaces (Supplementary Fig. S4). In contrast, d_{cell} exhibits a sufficiently broad range and serves as a key structural descriptor that mediates the coupling between structure, porosity, and yield strength. Accordingly, the global linkage analysis is centered on the cell size–porosity–yield strength relationship.

As shown in Fig. 7a, both d_{cell} and porosity exhibit inverse correlations with the predicted yield strength. In particular, the effect of d_{cell} is pronounced: yield strength decreases systematically with increasing d_{cell}, reflecting microstructural coarsening that diminishes the strengthening contribution of solute-enriched cellular boundaries and associated dislocation networks, which otherwise impede dislocation motion. This trend is consistent with the feature importance identified in the black-box model through SHAP analysis and is quantitatively preserved in the white-box σ_{YS} Model 4, as further illustrated in Fig. 7b. The agreement between the black-box and white-box analyses confirms that the dominant microstructural influence identified statistically is retained in the interpretable analytical model.

By comparison, porosity exhibits an inverse relationship with yield strength, as shown in Fig. 7c. Although its correlation with yield strength is weaker than that associated with d_{cell}, the observed trend is consistent with the role of porosity as a stress concentrator that promotes premature yielding and failure under applied loading. Notably, unlike d_{cell}, porosity enters σ_{YS} Model 4 through a quadratic term rather than a linear contribution. Consequently, the resulting model captures the overall degrading influence of porosity on strength but does not lend itself to a similarly direct, linear interpretation.

Overall, this analysis demonstrates that the influence hierarchy inferred from black-box SHAP analysis is reliably reflected in the white-box DSML framework, while simultaneously enabling global inspection of structure–porosity–strength coupling and stability.

3.6. Experimental validation

To evaluate the generalizability and effectiveness of the proposed models, validation was performed using six distinct process parameter combinations listed in Table 1, all with *LP* of 370 W, except for S6, which had *LP* of 350 W, and T_{pre} of 100°C. All tensile tests were repeated at least three times to obtain reproducible results.

Table 1
LPBF process parameter combinations used for validation, and the corresponding yield-strength values obtained through tensile tests.

Sample	Printing Parameters			
	Scan Speed (mm/s)	Layer thickness (μm)	Hatch spacing (μm)	Volumetric energy density (J/mm^3)
S1	2100	60	100	29.37
S2	3300	60	100	18.69
S3	2700	60	120	19.03
S4	3300	80	100	14.02
S5	2100	80	120	18.35
S6	1800	30	100	64.81

To enable comparative analysis between the white-box and black-box models, the average grain and cell sizes were calculated, as illustrated in Fig. 8a1–f1 and Fig. 8a3–f3, respectively. The values in Fig. 8a3–f3 were extracted from the melt pool fine regions highlighted in Figs. 8 a2–f2. The yield strength was experimentally determined under varying process parameter conditions (Fig. 8*g*). Numerical values for each grain size and the corresponding yield strength are provided in Supplementary Table S2.

In addition to comparing the white-box and black-box models, the predictive capability of a conventional empirical equation used to quantify yield strength in LPBF-fabricated alloys based on cell size was also evaluated as a baseline model [37]. The empirical relationship can be expressed as

$$\sigma_{YS} = \sigma_i + \Delta\sigma_{ss} + \Delta\sigma_{GB} + \Delta\sigma_{Cell}, \tag{9}$$

where σ_i is the intrinsic friction stress of pure Al; $\Delta\sigma_{SS}$ represents solid solution strengthening owing to Si and Mg, calculated as

$$\sigma_{SS} = M\left(\frac{3}{8}\right)^{2/3}\left(\frac{1+\nu}{1-\nu}\right)^{4/3}\left(\frac{\omega}{b}\right)^{1/3}\mu|\varepsilon|^{4/3}c^{2/3} \tag{10}$$

$\Delta\sigma_{GB}$ denotes grain-boundary strengthening ($\frac{K_{GB}}{\sqrt{d_{grain}}}$); and $\Delta\sigma_{Cell}$ represents the strengthening contribution of cellular structures, defined as $\frac{K_{CB}}{\sqrt{d_{cell}}}$. The definitions and numerical values of all parameters used in Eq. (9) are summarized in Supplementary Table S3.

Fig. 9 presents a series of scatter plots comparing the predicted yield strength (σ) and experimentally measured values for three modeling approaches: black-box (Fig. 9a), white-box (Fig. 9b), and empirical equation (Fig. 9c) models. The white-box model demonstrates the strongest correlation with experimental data, with an R^2 of 0.90 and low MAE of 9.51, indicating the highest predictive fidelity. Although the black-box model exhibits a comparable MAE of 11.34, its R^2 is considerably lower (0.73), implying reduced predictive precision. The empirical equation, conventionally used to describe LPBF strengthening mechanisms under near-fully dense conditions (>99.9% relative density), shows substantial deviation, with a markedly increased MAE of 41.98, highlighting its limited ability to capture variability in samples with high porosity.

Notably, as shown in Fig. 9c, when porosity exceeds 6%, the empirical equation systematically overestimates the yield strength owing to its inability to account for porosity effects. This observation reveals a fundamental limitation of traditional empirical equations, which are primarily derived from wrought materials or near-fully dense LPBF samples with negligible porosity. In contrast, the white-box model explicitly integrates porosity as a governing variable, addressing one of the key distinctions between wrought and LPBF-processed alloys. Consequently, it accurately reflects the pronounced role of process-induced pores in diminishing mechanical performance, offering a more realistic and comprehensive representation of the strengthening mechanisms in AM-fabricated alloys.

In summary, this work highlights the inadequacy of conventional empirical equations in capturing the mechanical behavior of LPBF alloys with significant porosity and underscore the importance of mechanistically informed, interpretable modeling approaches. The white-box model bridges the gap between experimental observations and theoretical predictions, offering a reliable framework for extrapolating beyond existing datasets and enabling predictive modeling that reflects the intrinsic physical phenomena governing LPBF materials.

A key limitation of this study lies in its focus on porosity as the dominant descriptor of LPBF-induced variability. While porosity plays a critical role, other microstructural strengthening mechanisms, such as dislocation density, Si-rich phase distribution, and post-processing parameters like heat treatment, were not explicitly incorporated into the

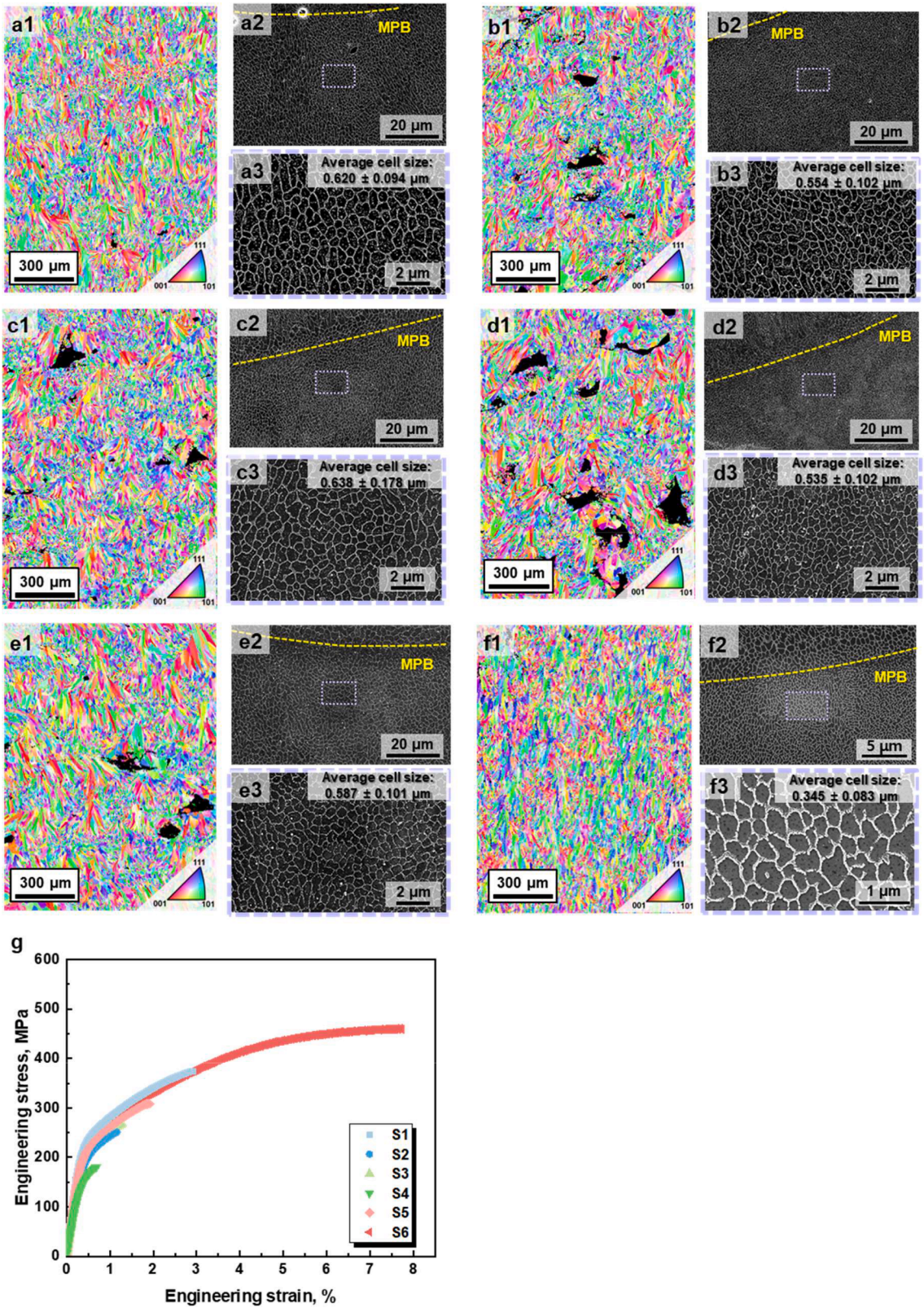

Fig. 8. Experiment results on grain, cell size, and engineering stress-strain curves of a multi-combination of LPBF-fabricated AlSi10Mg. (a1–f1) side view (plane parallel to the building direction) of the L-PBF processed specimens. (a2-f2) SEM image from the top view showing the cellular structures of specimens. The dashed yellow line distinguishes the MPB (melt pool boundary) region. (a3–f3) Magnified SEM image of the MP fine region. (g) Engineering stress–strain curves of six combinations.

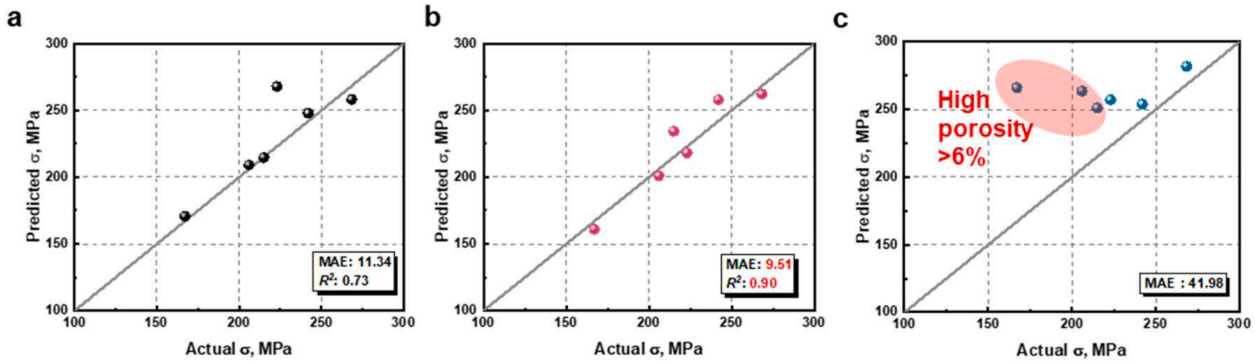

Fig. 9. Prediction results for the models versus the actual values of experimental LPBF-fabricated AlSi10Mg. Prediction results for the (a) black-box (ETR), (b) white box, and (c) empirical equation models.

proposed framework. Future work will thus expand the white-box model to include these variables, enabling a more comprehensive, mechanistically grounded understanding of yield-strength evolution. A model integrating these factors will likely provide deeper insights into the complex interplay among porosity, phase transformations, and dislocation behavior, leading to more accurate predictions for both single-phase and multiphase strengthened alloys.

4. Conclusions

In summary, this study developed a DSML framework for predicting the yield strength of LPBF-fabricated alloys by strategically combining black-box and white-box modeling approaches. A key feature of the DSML framework is its hierarchical data-selective strategy, in which process-only datasets are systematically integrated to refine porosity predictions prior to their incorporation into the yield-strength model. This selective data utilization mitigates the limitations imposed by scarce microstructural features, while simultaneously enhancing model generalizability across diverse processing conditions.

Through the integration of SR within the DSML framework, the white-box model produces transparent and highly accurate equations that quantitatively describe the coupled effects of porosity, microstructural descriptors, and process parameters on yield strength. The proposed model demonstrates exceptional predictive fidelity ($R^2 = 0.90$, MAE = 9.51), effectively balancing predictive accuracy with physical interpretability. Comparative evaluations against both black-box and conventional empirical models highlight the superior performance of the DSML framework, not only in terms of predictive precision but also in its ability to reveal the mechanistic role of porosity in controlling the mechanical behavior of LPBF-fabricated alloys.

Overall, the DSML framework offers a generalizable and interpretable pathway for predictive modeling in AM. By bridging the gap between experimental observations and theoretical principles, it provides a powerful tool for process optimization and defect control, thereby supporting the development of next-generation LPBF alloys with tailored performance and improved reliability.

Data availability

The source data of this research can be found in the Supplementary data 1,2 of this article.

CRediT authorship contribution statement

Jeong Ah Lee: Writing – original draft, Visualization, Validation, Software, Investigation, Data curation, Conceptualization. **Yeon Woo Kim:** Validation. **Takayoshi Nakano:** Investigation. **Hyomoon Joo:** Resources. **Jeong Min Park:** Writing – review & editing, Supervision, Resources. **Hyoung Seop Kim:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A5A1030054). Jeong Ah Lee was supported by the Basic Science Research Program “Fostering the Next Generation of Researcher” through the NRF funded by the Ministry of Education (RS-2025-25399900). This study was also supported by Fundamental Research Program “Development of Digital Twin Technology (PNKA590)” of the Korean Institute of Materials Science (KIMS), and by the Hyundai Motor Group.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2026.122101.

References

[1] D. Wang, L. Liu, J. Tang, Y. Liu, C. Wei, Z. Weng, J. Shao, H. Tan, W. Zhou, B. Neirinck, N. Gianfolcaro, Y. Yang, C. Han, Recent advances on additive manufacturing of heterogeneous/gradient metallic materials via laser powder bed fusion, Int. J. Extreme Manuf. 7 (2025) 062007.
[2] J. Xu, P. Kontis, R.L. Peng, J. Moverare, Modelling of additive manufacturability of nickel-based superalloys for laser powder bed fusion, Acta Mater 240 (2022) 118307.
[3] J.A. Lee, J. Park, M.J. Sagong, S.Y. Ahn, J.-W. Cho, S. Lee, H.S. Kim, Active learning framework to optimize process parameters for additive-manufactured Ti-6Al-4V with high strength and ductility, Nat. Commun. 16 (2025) 931.
[4] D. S, J. Sahu, A comprehensive review on laser powder bed fusion process of SS316L mechanical properties and post-processing, Met. Mater. Int (2025).
[5] S.Y. Ahn, J. Moon, Y.T. Choi, E.S. Kim, S.G. Jeong, J.M. Park, M. Kang, H. Joo, H. S. Kim, A precipitation-hardened AlSi10Mg alloy fabricated using selective laser melting, Mater. Sci. Eng. A 844 (2022) 143164.
[6] J. Suryawanshi, K.G. Prashanth, S. Scudino, J. Eckert, O. Prakash, U. Ramamurty, Simultaneous enhancements of strength and toughness in an Al-12Si alloy synthesized using selective laser melting, Acta Mater 115 (2016) 285–294.
[7] Z. Li, L. Cui, Y. Liu, F. Bu, Z. Shi, Y. Shao, D. He, Q. Cao, Enhancing mechanical properties of LPBF-fabricated AlSi10Mg welded joints via laser metal deposition combined with direct aging heat treatment, Met. Mater. Int. (2025).
[8] A. Ghasemi, E. Fereiduni, M. Balbaa, M. Elbestawi, S. Habibi, Unraveling the low thermal conductivity of the LPBF fabricated pure Al, AlSi12, and AlSi10Mg alloys through substrate preheating, Addit. Manuf. 59 (2022) 103148.
[9] B. Chen, S.K. Moon, X. Yao, G. Bi, J. Shen, J. Umeda, K. Kondoh, Strength and strain hardening of a selective laser melted AlSi10Mg alloy, Scr. Mater. 141 (2017) 45–49.
[10] Y.S. Eom, J.M. Park, J.-W. Choi, D.-J. Seong, H. Joo, Y.C. Jo, K.T. Kim, J.H. Yu, I. Son, Fine-tuning of mechanical properties of additively manufactured AlSi10Mg

alloys by controlling the microstructural heterogeneity, J. Alloys Compd. 956 (2023) 170348.
[11] H. Hyer, L. Zhou, A. Mehta, S. Park, T. Huynh, S. Song, Y. Bai, K. Cho, B. McWilliams, Y. Sohn, Composition-dependent solidification cracking of aluminum-silicon alloys during laser powder bed fusion, Acta Mater 208 (2021) 116698.
[12] S. Wang, J. Ning, L. Zhu, Z. Yang, W. Yan, Y. Dun, P. Xue, P. Xu, S. Bose, A. Bandyopadhyay, Role of porosity defects in metal 3D printing: formation mechanisms, impacts on properties and mitigation strategies, Mater. Today 59 (2022) 133–160.
[13] K. Zhang, Y. Chen, S. Marussi, X. Fan, M. Fitzpatrick, S. Bhagavath, M. Majkut, B. Lukic, K. Jakata, A. Rack, M.A. Jones, J. Shinjo, C. Panwisawas, C.L.A. Leung, P. D. Lee, Pore evolution mechanisms during directed energy deposition additive manufacturing, Nat. Commun. 15 (2024) 1715.
[14] K.-S. Park, H. Jin, Y.-A. Shin, A. Gupta, J.-K. Kim, H.-N. Rhee, Y.-J. Son, S.-H. Choi, A study on the effect of bulk porosity and surface roughness on the biocompatibility of TPMS gyroid structures fabricated via the SLM process, Met. Mater. Int. 31 (2025) 3455–3473.
[15] J. Reijonen, A. Revuelta, S. Metsä-Kortelainen, A. Salminen, Effect of laser focal point position on porosity and melt pool geometry in laser powder bed fusion additive manufacturing, Addit. Manuf. 85 (2024) 104180.
[16] C.M. Laursen, S.A. DeJong, S.M. Dickens, A.N. Exil, D.F. Susan, J.D. Carroll, Relationship between ductility and the porosity of additively manufactured AlSi10Mg, Mater. Sci. Eng 795 (2020) 139922. A.
[17] H. Chen, S. Patel, M. Vlasea, Y. Zou, Enhanced tensile ductility of an additively manufactured AlSi10Mg alloy by reducing the density of melt pool boundaries, Scr. Mater. 221 (2022) 114954.
[18] Z. Ren, L. Gao, S.J. Clark, K. Fezzaa, P. Shevchenko, A. Choi, W. Everhart, A. D. Rollett, L. Chen, T. Sun, Machine learning–aided real-time detection of keyhole pore generation in laser powder bed fusion, Science 379 (2023) 89–94.
[19] S. Sheikh, B. Vela, P. Honarmandi, P. Morcos, D. Shoukr, I. Karaman, A. Elwany, R. Arróyave, An automated computational framework to construct printability maps for additively manufactured metal alloys, Npj Comput. Mater. 10 (2024) 252.
[20] J.A. Lee, M.J. Sagong, J. Jung, E.S. Kim, H.S. Kim, Explainable machine learning for understanding and predicting geometry and defect types in Fe-Ni alloys fabricated by laser metal deposition additive manufacturing, J. Mater. Res. Technol. 22 (2023) 413–423.
[21] S. Ouyang, Q. Wu, R.E. Kim, J.H. Lee, J.A. Lee, H.S. Kim, Recent progress in developing creep-resistant Mg-Gd-based alloys, Met. Mater. Int. 31 (2025) 2513–2524.
[22] W.L. Ng, G.L. Goh, G.D. Goh, J.S.J. Ten, W.Y. Yeong, Progress and opportunities for machine learning in materials and processes of additive manufacturing, Adv. Mater 36 (2024) 2310006.
[23] Y. Wang, N. Wagner, J.M. Rondinelli, Symbolic regression in materials science, MRS Commun 9 (3) (2019) 793–805.
[24] N. Makke, S. Chawla, Interpretable scientific discovery with symbolic regression: a review Artif, Intell. Rev. 57 (2024) 2.
[25] S.M. Udrescu, M. Tegmark, A.I. Feynman, A physics-inspired method for symbolic regression sci, Adv 6 (2020) eaay2631.
[26] P.A.K. Reinbold, L.M. Kageorge, M.F. Schatz, R.O. Grigoriev, Robust learning from noisy, incomplete, high-dimensional experimental data via physically constrained symbolic regression, Nat. Commun. 12 (2021) 3219.
[27] J.A. Lee, R.B. Figueiredo, H. Park, J.H. Kim, H.S. Kim, Unveiling yield strength of metallic materials using physics-enhanced machine learning under diverse experimental conditions, Acta Mater 275 (2024) 120046.
[28] X. Huang, S. Ma, C.Y. Zhao, H. Wang, S. Ju, Exploring high thermal conductivity polymers via interpretable machine learning with physical descriptors, Npj Comput. Mater. 9 (2023) 1–14.
[29] S. Zhu, D. Sarıtürk, R. Arróyave, Accelerating CALPHAD-based phase diagram predictions in complex alloys using universal machine learning potentials: opportunities and challenges, Acta Mater 286 (2025) 120747.
[30] V. Menon, S. Das, V. Gavini, L. Qi, Atomistic simulations and machine learning of solute grain boundary segregation in Mg alloys at finite temperatures, Acta Mater 264 (2024) 119515.
[31] A.V. Kapustin, M.M. Abramova, N.Y. Anisimova, M.V. Kiselevskiy, N.A. Enikeev, Effect of elementary unit configurations on the mechanical performance of periodic lattice structures to architect porous scaffolds by FEM-driven additive manufacturing, Met. Mater. Int. 31 (2025) 2417–2431.
[32] A.M. D'Altri, M. Pereira, S. de Miranda, B. Glisic, Simulation-driven machine learning for real-time damage prognosis in masonry structures, Int. J. Mech. Sci. 289 (2025) 110055.
[33] X. Pei, J. Pei, H. Hou, Y. Zhao, Optimizing casting process using a combination of small data machine learning and phase-field simulations, Npj Comput. Mater. 11 (2025) 27.
[34] N. Hollmann, S. Müller, L. Purucker, A. Krishnakumar, M. Körfer, S.B. Hoo, R. T. Schirrmeister, F. Hutter, Accurate predictions on small data with a tabular foundation model, Nature 637 (2025) 319–326.
[35] H. Tang, H. Ren, Y. Xiong, Optimization of scanning path in laser powder bed fusion: training and application of deep learning model using small dataset, J. Intell. Manuf. (2025).
[36] Q. Liu, W. Chen, V. Yakubov, J.J. Kruzic, C.H. Wang, X. Li, Interpretable machine learning approach for exploring process-structure-property relationships in metal additive manufacturing, Addit. Manuf. 85 (2024) 104187.
[37] J. Kwon, Y.T. Choi, E.S. Kim, S.Y. Ahn, S.G. Jeong, M. Kang, H. Joo, J.M. Park, H. S. Kim, Effect of cell characteristics on mechanical properties of AlSi10Mg alloy fabricated by laser powder bed fusion, Mater. Sci. Eng. A 901 (2024) 146537.
[38] M. Verleysen, D. François, The curse of dimensionality in data mining and time series prediction, in: J. Cabestany, A. Prieto, F. Sandoval (Eds.), Comput. Intell. Bioinspired Syst, Springer, Berlin, Heidelberg, 2005, pp. 758–770.
[39] T. Poggio, H. Mhaskar, L. Rosasco, B. Miranda, Q. Liao, Why and when can deep-but not shallow-networks avoid the curse of dimensionality: A review, Int. J. Autom. Comput. 14 (2017) 503–519.
[40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine Learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[41] M. Cranmer, Interpretable machine learning for science with PySR and SymbolicRegression.Jl, (2023).
[42] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model prediction, Adv. Neural inf. Process. (2017).
[43] Z. Wu, N. Ye, L. Liu, W. Zhu, J. Mao, H. Chen, Y. Liu, J. Tang, Enhancing strength-ductility synergy of LPBF nano-zirconia strengthened AlSi10Mg composites via laser remelting, Mater. Sci. Eng. A 944 (2025) 148918.
[44] X. Wu, J. Liu, Y. Yang, J. Bai, C. Shuai, J. Buhagiar, X. Ning, Laser powder bed fusion of biodegradable magnesium alloys: process, microstructure and properties, Int. J. Extreme Manuf. 7 (2024) 022007.
[45] T. de Terris, O. Andreau, P. Peyre, F. Adamski, I. Koutiri, C. Gorny, C. Dupuy, Optimization and comparison of porosity rate measurement methods of Selective laser melted metallic parts, Addit. Manuf. 28 (2019) 802–813.
[46] W.H. Kan, L.N.S. Chiu, C.V.S. Lim, Y. Zhu, Y. Tian, D. Jiang, 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 (2022) 9818–9865.
[47] S.F. Nabavi, H. Dalir, A. Farshidianfar, A comprehensive review of recent advances in laser powder bed fusion characteristics modeling: metallurgical and defects, Int. J. Adv. Manuf. Technol. 132 (2024) 2233–2269.
[48] M. Liu, K. Wei, X. Zeng, High power laser powder bed fusion of AlSi10Mg alloy: effect of layer thickness on defect, microstructure and mechanical property, Mater. Sci. Eng. A 842 (2022) 143107.
[49] M. Zhang, C.-N. Sun, X. Zhang, P.C. Goh, J. Wei, D. Hardacre, H. Li, Fatigue and fracture behaviour of laser powder bed fusion stainless steel 316L: influence of processing parameters, Mater. Sci. Eng. A 703 (2017) 251–261.
[50] L. Afroz, M. Qian, J. Forsmark, Y. Li, M. Easton, R. Das, Fatigue life of laser powder bed fusion (L-PBF) AlSi10Mg alloy: effects of surface roughness and porosity, Prog. Addit. Manuf. 10 (2025) 2423–2441.
[51] Y. Geng, Z. Shan, J. Zhang, T. Wei, Z. Zhang, Densification, microstructure, mechanical properties, and thermal stability of high-strength Ti-modified Al-Si-Mg-Zr aluminum alloy fabricated by laser-powder bed fusion, Int. J. Miner. Metall. Mater. 32 (2025) 2547–2559.
[52] S. Wei, P. Wang, L. Zhang, U. Ramamurty, Grain morphologies in additively manufactured alloys: from solidification fundamentals to advanced microstructure control, J. Mater. Sci. Technol. 235 (2025) 133–145.