Journal of Materials Research and Technology 39 (2025) 2621–2632

Contents lists available at ScienceDirect

Journal of Materials Research and Technology

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

A machine learning method for predicting the elongation to failure of Al–Si alloy in high pressure die casting combining experiment and modeling

Hongjian Wu [a], Xuexia Song [a], Jingzhou Lu [a], Kun Dou [a,b,*], Wanlin Wang [a], Yijie Zhang [c], Zhongyun Fan [d]

[a] School of Metallurgy and Environment, Central South University, Changsha, 410083, Hunan, China
[b] Xiangjiang Laboratory, Changsha, 410205, Hunan, China
[c] Institute for New Materials, Dongliang Aluminium Co., Ltd, Huzhou, 313008, Zhejiang, China
[d] Brunel Centre for Advanced Solidification Technology (BCAST), Brunel University London, Kingston Lane, Uxbridge, UB8 3PH, United Kingdom

ARTICLE INFO

Keywords:
High pressure die casting
Machine learning
XGBoost
CPO
SHAP analysis
Solidification

ABSTRACT

High-pressure die casting (HPDC) of Al–Si alloys faces challenges in improving mechanical properties and early failure due to process-induced defects. Traditional quality assessment experiments and modeling methods are costly and lack predictive capability. This study proposed a machine learning (ML) method integrating numerical simulations and experimental data to predict elongation. The data were extracted from experiments and well validated mathematical models of the HPDC process and underwent preprocessing such as standardization and feature selection. The performance of twelve common ML models was evaluated, among which the eXtreme Gradient Boosting (XGBoost) and Gradient Boosting (GB) algorithms performed the best. By using the Crested Porcupine Optimizer (CPO) and Bayesian Optimization for hyperparameter optimization, the accuracy of the models was further improved. The R^2 value of the CPO-XGBoost model reached the optimal value of 0.882. The SHapley Additive exPlanations (SHAP) analysis revealed that total shrinkage volume dominantly reduced elongation (El), while die temperature and pouring temperature negatively affected El. Validation with independent experiments demonstrated that the difference between the predicted values and the measured values was within 5 %. This work establishes a novel methodology combining physics-based simulations and experiments to comprehensively predict and analyze the failures of HPDC Al–Si alloys from multiple perspectives. It can be further extended to the prediction of various types of alloys or multiple mechanical properties.

1. Introduction

The HPDC process has been widely employed in manufacturing alloy components with low surface roughness and high dimensional accuracy due to its high production efficiency and exceptional capability for forming complex geometries [1,2]. However, the process is significantly affected by factors such as high pressure and speed mold filling and solidification shrinkage, which often lead to internal defects including porosity, shrinkage and inclusions, thereby compromising their ability to meet high-performance requirements [3,4]. In the HPDC process, the elongation of as-cast products usually tends to have large variabilities [5–7], which are closely related to the solidification and defects formation during the process. The obvious variation in elongation can have a profound negative influence on the reliability of the cast component, leading to unpredictable early failures of the materials. Hence, understanding the potential interrelations between elongation and HPDC process parameters is crucial in controlling the variabilities of the elongation and stabilizing the HPDC process.

To ensure product quality, conventional quality control systems typically rely on multidimensional detection methods encompassing computed tomography inspection, tensile mechanical property testing, and microstructural analysis [8,9]. Nevertheless, these approaches present limitations such as substantial equipment investment costs, prolonged inspection cycles, and heavy reliance on manual operations, which collectively increase production costs and constrain optimization efficiency in manufacturing processes. Numerous researchers [10–14] have investigated the influence of mechanisms of HPDC solidification processes and microstructural phase evolution on casting quality, aiming to overcome these limitations while enhancing both production efficiency and product quality. However, current methodologies remain

* Corresponding author. School of Metallurgy and Environment, Central South University, Changsha, 410083, Hunan, China.
E-mail address: Kun.Dou@csu.edu.cn (K. Dou).

https://doi.org/10.1016/j.jmrt.2025.10.001
Received 26 June 2025; Received in revised form 14 September 2025; Accepted 1 October 2025
Available online 1 October 2025

insufficient to comprehensively elucidate the complex effects of HPDC process conditions on final product quality. In recent years, ML has demonstrated significant potential in the fields of materials computation and analysis, owing to their robust data mining and modeling capabilities [15,16]. Rai et al. [17] developed a Neural Network-based Casting Process (NN-CastPro) model, which utilized inlet melt temperature, mold initial temperature, inlet first phase velocity and inlet second phase velocity as input features to predict filling time, solidification time and porosity. The results demonstrated that the NN-CastPro model achieved accurate predictions while requiring substantially less computational time compared to finite element analysis software. Yarlagadda et al. [18] integrated process simulation software with neural networks trained using domain expertise to predict filling time based on parameters such as melt temperature, die temperature, injection pressure, and casting weight. Soundararajan et al. [19] employed neural network models to predict the ultimate tensile strength (UTS) and yield strength (YS) of aluminum alloy tensile bars produced by gravity casting, achieving correlation coefficients of 0.95 and 0.96, respectively. Zheng et al. [20] adopted artificial neural network (ANN) algorithms to analyze correlations between HPDC process parameters (e.g., mold temperature, pouring temperature and injection velocity) and surface defects.

This study establishes a mathematical model of the entire HPDC process to acquire data on filling time, solidification time, total shrinkage volume, temperature at fill time, air entrainment, oxides indicators and so on. These data are integrated with experimental measurements to predict the El of Al–Si alloy castings. To address challenges associated with small sample machine learning, feature selection methods are applied, and various algorithms including Random Forest (RF), XGBoost, GB, Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), K-Nearest Neighbors (KNN), Support Vector Regression (SVR), and ANN are employed to predict elongation. Hyperparameter optimization algorithms are implemented to enhance the prediction accuracy of these models, while SHAP analysis is utilized to visualize the importance of different parameters on the mechanical properties of HPDC Al–Si alloys. Finally, the predictive accuracy of the model is validated by comparing experimental measurements with model predictions. More importantly, the core methodology established in this work is integrating physics-based simulations, experimental characterization, machine learning modeling, presents a versatile framework. While demonstrated here for elongation prediction, this framework holds significant potential for application to predicting other critical quality metrics in HPDC, such as yield strength, ultimate tensile strength, or specific defect characteristics.

2. Modeling and experiment

2.1. Data acquisition from experiments and modelling

To predict the elongation of the HPDC components, a series of factors are considered, as shown in Fig. 1. Both the experimental process parameters (piston velocities (V1, V2, V3), intensification pressure (IP), pouring temperature, die temperature) and the solidification characteristics that are difficult to obtain directly from experiments such as filling time, solidification time, total shrinkage volume, temperature at fill time, air entrainment, oxides indicator, are determined by the FEM model, and are utilized as supplementary input features, while elongation measured from tensile tests served as the target output. Preprocessing of the collected experimental and simulated data was performed, followed by regression prediction analysis using machine learning methods. The key experimental process parameters used in this research are obtained from the HPDC machine's control panel, as is indicated in Fig. 2.

The numerical simulation of a cold chamber HPDC machine was implemented based on the Finite Element Method (FEM) using the ProCAST software platform, primarily involving three critical steps: finite element mesh generation, discretized computation and post processing. The three-dimensional geometric model of the HPDC process, constructed using CAD software, adopted a semi-symmetric structure to significantly improve simulation efficiency while ensuring computational accuracy. The machine configuration and the corresponding FEM model are shown in Fig. 3. Key parameters of the model and validation methodologies are referenced from the authors' prior research [21–24], as shown in Table 1.

According to the machine configuration in Fig. 3(a), to start the experiment, a typical HPDC process is carried out, which include: 1) Pouring of the A356 melt (added 0.5 % Mn) into the shot sleeve, 2) Moving of the piston towards the biscuit to inject the melt into the die region as well as the instant cooling and solidification of the melt, 3) Continual movement of the piston towards the biscuit under the intensification pressure to ensure adequate feeding of the melt to avoid excessive porosities in the as-cast components. The experimental

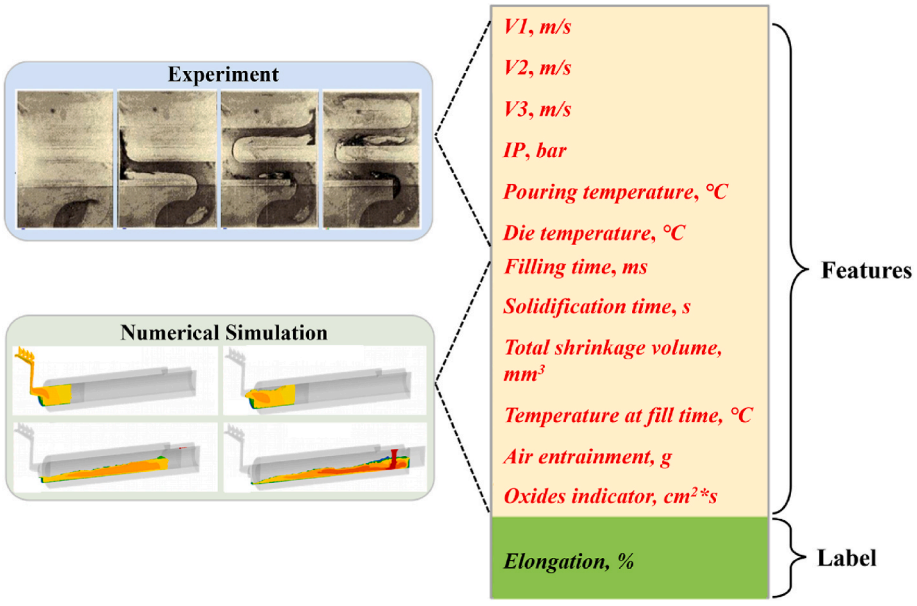

Fig. 1. Data acquisition from experiments and numerical simulations.

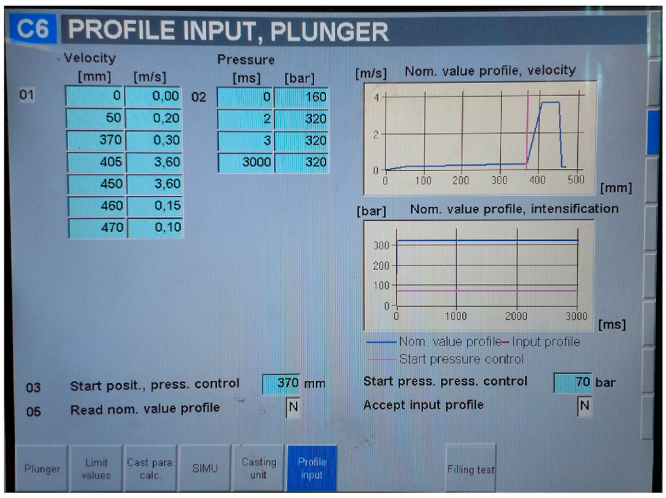

Fig. 2. Control panel of the HPDC machine indicating process parameters such as velocity and pressure.

procedures after the HPDC process involved aging the tensile specimens at room temperature for 24 h after demolding, followed by internal porosity characterization using an X-ray computed tomography system. A casting containing eight round tensile specimens was prepared in each high-pressure die casting, as shown in Fig. 3 (b). The gauge diameter and gauge length of the tensile specimens were prepared as 6.35 mm and 55 mm, respectively. Quasi-static tensile testing of the specimens was conducted on a universal testing machine at a loading rate of 1 mm/min until fracture. For each tensile sample position, 4 tests were conducted for repeatability of the elongation. For analyzing pore morphology and distribution characteristics, standard metallographic preparation procedures were employed, with multiscale characterization performed using optical microscopy (OM).

Based on the above experiments and FEM modelling, all 12 features and corresponding label mentioned in Fig. 1 can be obtained and summarized in Table 2. Fig. 4 presents a comprehensive feature analysis matrix, which includes scatter plots, individual feature distributions, and a correlation heatmap. The scatter plots highlight the pairwise relationships among process parameters. Due to the practical constraints in the industrial experiments, features such as V1, V2, and V3 exhibit a skewed distribution. The scatter plots of these features indicate a strong linear relationship. The correlation heatmap reveals several important patterns. V1 and V2 show a strong positive correlation, indicating potential redundancy. Both V1 and V2 also exhibit strong negative correlations with total shrinkage volume and pouring temperature and filling time are positively correlated. While multicollinearity exists between a few features, it can be addressed through appropriate feature selection strategies to ensure model stability and interpretability.

Table 1
The key parameters of process and boundary conditions of FEM model.

Key parameters	Value	Key parameters	Value
Alloy type	A356	Initial average die surface temperature	150 °C
Liquidus temperature	614 °C	Die blowing end time	55 s
Solidus temperature	553 °C	Ambient Air heat transfer coefficient	20 W/m^2·K
Pouring temperature	680 °C	Ambient temperature	25 °C
Part ejection time	20 s	Die spray cooling heat transfer coefficient	500 W/m^2·K
Shot sleeve temperature	180 °C	Spray temperature	20 °C
Shot sleeve length	480 mm	Die air-blow cooling heat transfer coefficient	300 W/m^2·K
Die and sleeve material	H13 steel	Air-blow temperature	20 °C

Table 2
Summary of feature and label values.

Feature and label	Values		
	Min	Max	Mean
V1, m/s	0.2	0.4	0.323
V2, m/s	0.3	0.6	0.485
V3, m/s	2.2	4.2	3.371
IP, bar	100	320	264.377
Pouring temperature, °C	670.1	679.7	674.584
Die temperature, °C	123.2	166.6	144.028
Filling time, ms	8	23	15.553
Solidification time, s	0.93	1.04	0.989
Total shrinkage volume, mm^3	0.051	2.774	0.243
Temperature at fill time, °C	593.1	624	611.704
Air entrainment, g	0.002	0.009	0.004
Oxides indicator, cm^2·s	2.61	3.46	3.025
Elongation, %	2.08	15.8	11.7

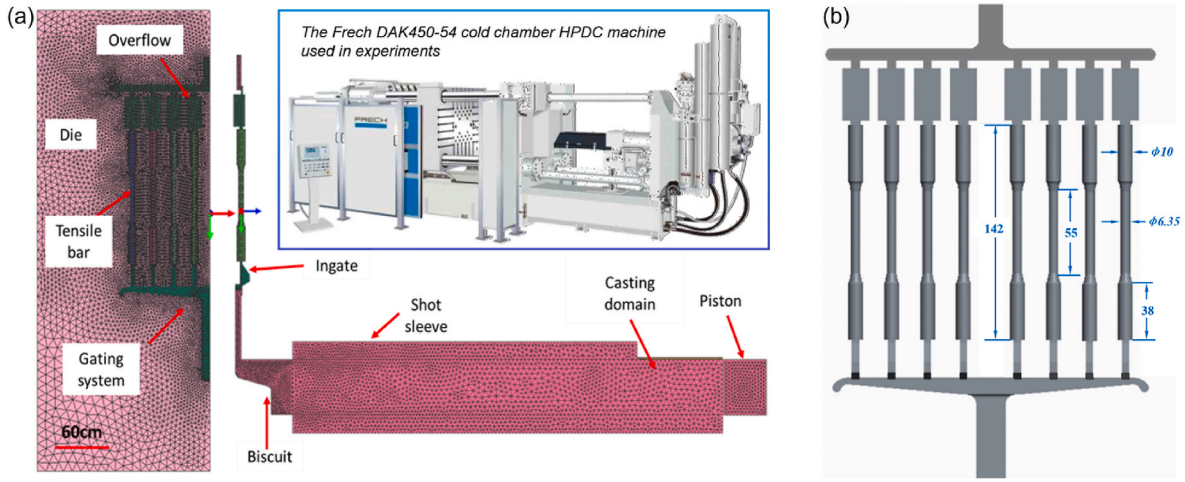

Fig. 3. (a) Configuration and FEM model of the cold chamber HPDC machine used in this research and (b) geometry of the tensile samples (dimensions in mm) prepared by HPDC.

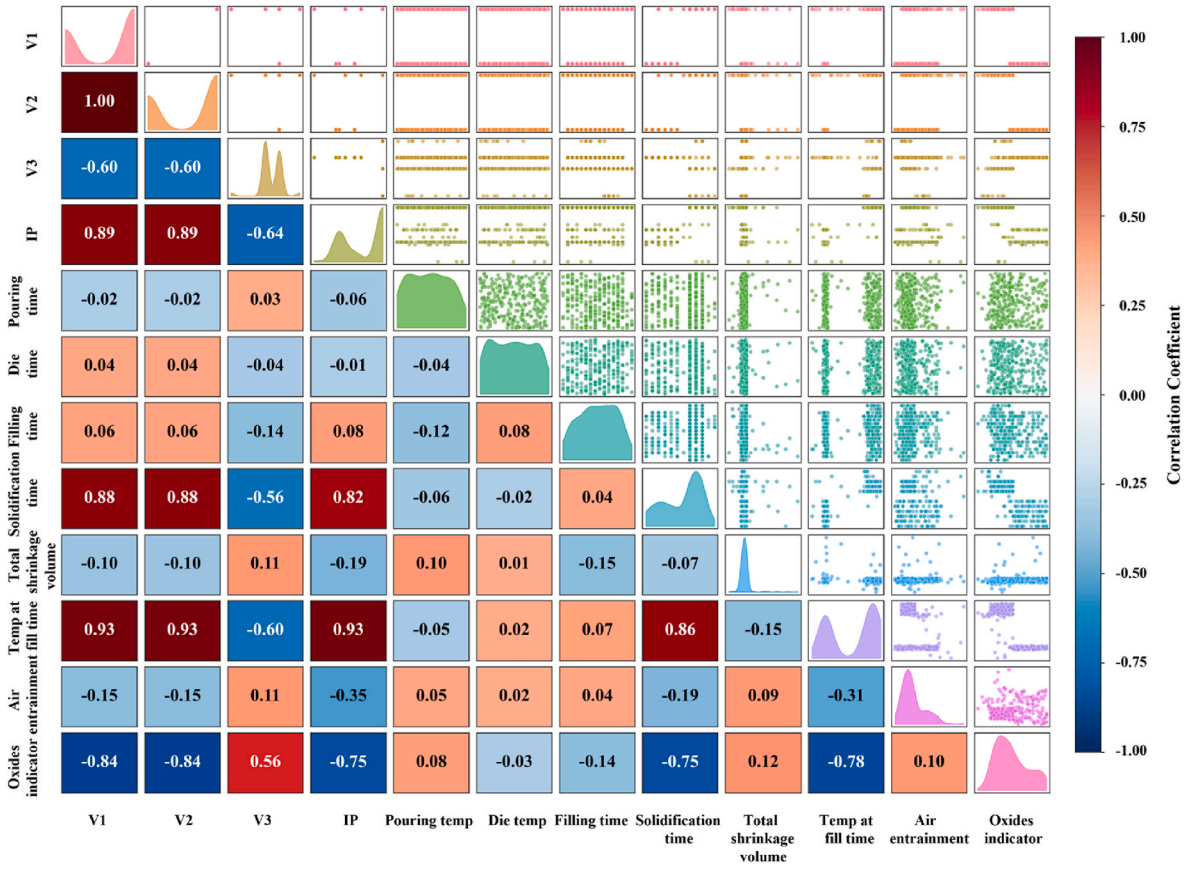

Fig. 4. A comprehensive feature analysis matrix: scatter plots, feature distributions and a correlation heatmap.

2.2. Machine learning algorithm

RF, XGBoost, LightGBM, GB, CatBoost [25,26], KNN [27], SVR [28], and ANN [29,30] are commonly used to predict the mechanical properties of alloys. These approaches facilitate the exploration of the influence of input features on alloy performance and can even be combined with optimization algorithms to design novel materials with enhanced properties.

RF, XGBoost, LightGBM, GB, and CatBoost are all improved variants of tree-based algorithms. For regression tasks, RF operates by averaging predictions from multiple decision trees (CART trees). Its ensemble prediction formula is expressed as:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \tag{1}$$

where $h_t(x)$ denotes the prediction of the t-th tree, and T represents the total number of trees.

The XGBoost algorithm iteratively optimizes its objective function, which combines a loss function and regularization terms, through an additive modeling framework. The total loss consists of prediction error and regularization components, formulated as:

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \sum_{k=1}^{t} \Omega(f_k) \tag{2}$$

$$\Omega(f_k) = \gamma T_t + 1 \Big/ 2\lambda \sum_{j=1}^{T_t} \omega_j^2 \tag{3}$$

$$\omega_j^* = -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i + \lambda} \tag{4}$$

where $\mathscr{L}^{(t)}$ is the total loss at the t-th iteration, $L\left(y_i, \hat{y}_i\right)$ is the loss function for the i-th sample, $\hat{y}_i^{(t-1)}$ denotes the cumulative prediction from the first t-1 iterations, $f_t(x_i)$ represents the output of the t-th tree for the i-th sample, $\Omega\left(f_k\right)$ is the regularization term for the k-th tree. γ is the minimum gain threshold for leaf node splitting, T_t is the number of leaf nodes in the k-th tree, λ is the L2 regularization coefficient for leaf weights ω_j, ω_j denotes the weight of the j-th leaf node, I_j represents the set of samples assigned to the j-th leaf node.

The KNN algorithm predicts outcomes based on local similarity. It employs Euclidean distance (Eq. (5)) or Manhattan distance (Eq. (6)) to measure spatial proximity between samples. Weighted KNN formulas (Eqs. (7) and (8)) are applied for prediction.

$$d\left(x_i, x_j\right) = \sqrt{\sum_{m=1}^{p} \left(x_{i,m} - x_{j,m}\right)^2} \tag{5}$$

$$d\left(x_i, x_j\right) = \sum_{m=1}^{p} \left|x_{i,m} - x_{j,m}\right| \tag{6}$$

$$\hat{y} = \frac{\sum_{i=1}^{k} \delta_i y_i}{\sum_{i=1}^{k} \delta_i} \tag{7}$$

$$\delta_i = \frac{1}{d(x, x_i)^p} \tag{8}$$

where δ_i denotes the weight of the i-th neighbor, and p represents the power function exponent for distance weighting.

The SVR algorithm addresses the primal optimization problem by maximizing the classification margin while minimizing misclassification penalties. For nonlinear classification, the primal problem is trans-

formed into a dual problem using kernel functions $K(xi,xj)$ to implicitly map data into a higher-dimensional space:

$$\min_{\omega,d,\xi} 1/2||\theta||^2 + C\sum_{i=1}^{n}\xi_i \quad s.t. \quad y_i\left(\theta^T x_i + d\right) \geq 1-\xi_i, \xi_i \geq 0 \tag{9}$$

$$\max_{\alpha} \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{ij}\alpha_i\alpha_j y_i y_j K\left(x_i, x_j\right) \quad s.t. \quad 0 \leq \alpha_i \leq C, \sum_{i=1}^{n}\alpha_i y_i = 0 \tag{10}$$

where θ is the hyperplane normal vector, d is the bias term, ξi are slack variables, C is the penalty parameter, and αi are Lagrange multipliers.

ANN algorithm computes outputs through forward propagation (Eq. (11)), where each layer's activation is derived from a linear transformation (weight matrix W and bias b) followed by a nonlinear activation function. The cross-entropy loss function is employed to quantify the discrepancy between the predicted probability distribution and the true distribution. Gradients of the loss with respect to the weights are computed via the chain rule, and the weights are iteratively updated in the direction opposite to these gradients to progressively reduce the loss.

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}, a^{(l)} = \sigma\left(z^{(l)}\right) \tag{11}$$

$$\mathscr{L}_{nn} = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log\widehat{y}_i + (1-y_i)\log(1-y_i)\right] \tag{12}$$

$$W^{(l)} \leftarrow W^{(l)} - \eta\frac{\partial\mathscr{L}_{nn}}{\partial W^{(l)}} \tag{13}$$

where $z^{(l)}$ is the linear output of the l-th layer, $W^{(l)}$ and $b(l)$ are the weight matrix and bias vector of the l-th layer, $a^{(l)}$ is the activation value, and σ is the activation function.

For optimization algorithms, Bayesian Optimization (BO) is grounded in Bayesian theorem and employs probabilistic models to model and infer the objective function. Its primary steps include constructing a probabilistic model, calculating the acquisition function, selecting the next sampling point, and iterating the optimization process. This framework enables BO to effectively balance exploration and exploitation, achieving optimal solutions with fewer function evaluations.

The CPO algorithm, proposed by Abdel-Basset et al. [31] in 2024, simulates four defense behaviors of crested porcupines, where the first and second defense strategies represent the exploratory behavior of CPO, while the third and fourth strategies correspond to its exploitative behavior. The algorithm introduces a cyclic population reduction technique, which enhances convergence speed and population diversity. Compared to other optimizers, CPO exhibits superior performance across most benchmark test functions.

Given the superior convergence speed and diversity preservation performance demonstrated by CPO on benchmark test functions, as well as its potential for handling complex nonlinear optimization problems, this study selected CPO as the primary hyperparameter optimization algorithm and conducted comparative verification with the widely used BO. The key algorithm parameters were set as population size = 30, maximum number of iterations = 200. The optimization process terminated upon reaching the maximum iteration count. The objective function for CPO and BO was the minimization of the root mean squared error (RMSE) estimated via 3-fold cross-validation on the training set.

To evaluate the accuracy of regression model training, three commonly used evaluation metrics are employed: RMSE, mean absolute error (MAE), and coefficient of determination (R^2). The corresponding calculations are shown as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{p}_i - p_i\right)^2} \tag{14}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{p}_i - p_i\right| \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\widehat{p}_i - \overline{p}\right)^2}{\sum_{i=1}^{n}\left(p_i - \overline{p}\right)^2} \tag{16}$$

where n, p_i and $\widehat{p}_i$ denote the number of data, real and predicted performance values, respectively. Furthermore, p is the average of the real performance values.

By quantifying each feature's marginal contribution to prediction outcomes, SHAP analysis [32] interprets feature-specific influences on final model predictions. The computation of SHAP values evaluates all possible feature subsets, which fundamental equation is shown as:

$$\varphi_i = \sum_{S\subseteq N\setminus\{i\}}\frac{|S|!(|N|-|S|-1)!}{|N|!}\left(v(S\cup\{i\}) - v(S)\right) \tag{17}$$

where φ_i denotes the SHAP value for feature, S represents a feature subset, N is the complete feature set, $v(S)$ is the model output with feature subset S, and $v(S\cup\{i\})$ is the output when feature i is added to S.

The computational procedure involves: (1) enumerating all possible feature subsets, (2) calculating each feature's marginal contribution across these subsets, and (3) averaging contributions to determine feature importance. SHAP analysis was characterized by its consistency, fairness, local-global interpretability, and model compatibility.

2.3. Construction procedure for ML model

This study was conducted using an initial dataset comprising 313 samples, which included twelve input features and one target output variable. An ensemble feature importance evaluation approach was implemented by integrating RF, Mutual Information (MI), and Lasso. This integration leverages the strengths of each method: RF captures nonlinear interactions between parameters; MI compensates for the potential insensitivity of tree-based models to weaker nonlinear relationships; and Lasso enforces sparsity through L1 regularization, preventing interference from multicollinearity. The normalized importance scores from each method were aggregated using weights of RF: MI: Lasso = 0.35:0.35:0.3 to generate a consolidated feature relevance ranking. To determine the optimal feature subset size, the top k features (k = 1, 2, …, 12) from this ranking list were used. The performance of models trained on these subsets was evaluated using 5-fold cross-validation on the training set. As illustrated in Fig. 5 (b), using the XGBoost model as an example, model performance reached its optimal state when k = 5. Further increasing the number of features leads to performance degradation due to overfitting. And decreasing the number of features results in a sharp decrease in model performance, which is attributable to underfitting.

The refined dataset was standardized and partitioned into training and testing sets at a 0.8:0.2 ratio. Base models including RF, LightGBM, XGBoost, CatBoost, KNN, and ANN and so on, were evaluated based on R^2, RMSE, and MAE metrics. The top two performing models were selected for hyperparameter optimization using the BO and CPO algorithms to enhance their generalization capabilities. Feature importance analysis was employed to visualize the impact of different features on the elongation of HPDC Al–Si alloys. Finally, the accuracy of the prediction results of the ML model is verified using experimental data. The flowchart of experimental procedure is shown in Fig. 5(a).

Model training, hyperparameter optimization and evaluation were performed on a computational workstation equipped with an Intel Xeon Gold 6234 CPU and an NVIDIA RTX 5080 GPU. All machine learning models were implemented using Python 3.9.7. The BO was performed using the 'BayesSearchCV' function from the scikit-optimize library and the CPO was implemented based on the description in references. The SHAP analysis was conducted using the 'shap' library. The ANN model was accelerated using the GPU for training. Tree-based models and other algorithms were primarily executed on the CPU. Hyperparameter optimization of the XGBoost model using CPO took approximately 45 min to

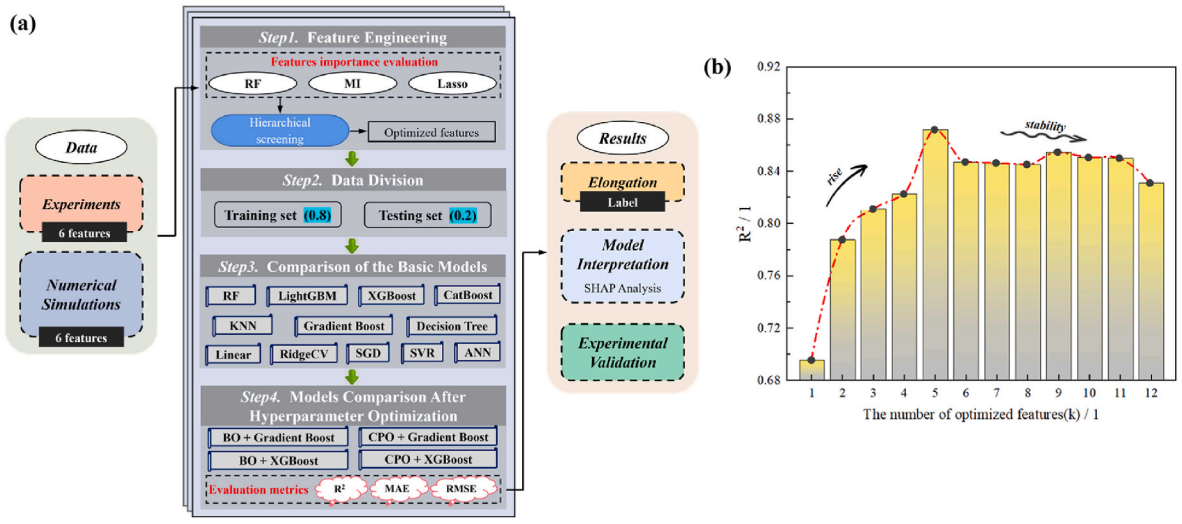

Fig. 5. (a) Flowchart of the ML model construction procedure, (b) the R2 of model with different numbers of optimized features.

evaluate 100 iterations. The final training time for the optimized CPO-XGBoost model was under 4 s.

3. Results and discussion

3.1. Solidification and defects formation

Before constructing the machine learning model for elongation prediction, the datasets needed to be clearly validated, especially regarding the features that are obtained via the finite element modelling. Hence in this section, a series of validations are performed regarding the solidification and defects formation in the HPDC process. Regarding the thermal profile of the die in the HPDC process, the actual temperature distribution during the experiment was measured and visualized via infrared camera, as is indicated in Fig. 6 (a) while the corresponding simulated temperature distribution is shown Fig. 6 (b). The thermal field from experiment and modelling are quite similar in both distribution and magnitude, as shown in Fig. 6 (c). This provides proof for the reliability of the model.

Based on thermal modelling, the defect formation, such as total shrinkage porosities, air entrainment and oxides in the HPDC components are further modeled based on POROS model and GAS model in ProCAST [33], with the run parameter JUNCTION = 15. To further validate this case, the actual cast components are obtained. The corresponding positions, as indicated by the red dot dash line in Fig. 7 (a), were cut, polished, etched and observed using OM. It can be clearly seen

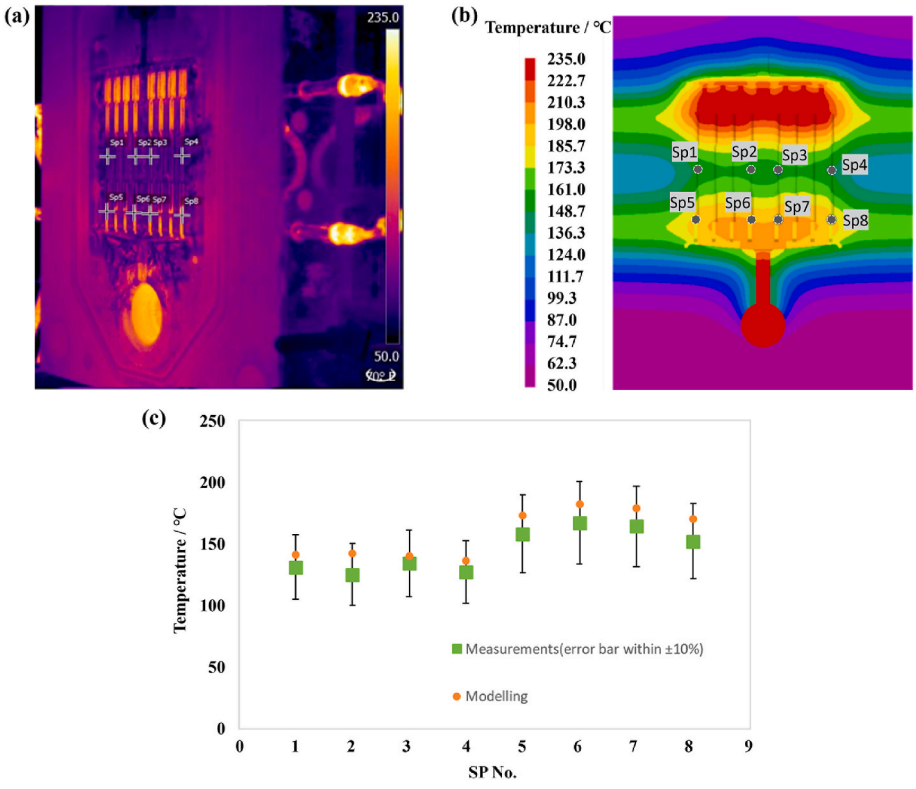

Fig. 6. The thermal profiles for the die piece: (a) measurement, (b) modelling and (c) verification of the accuracy.

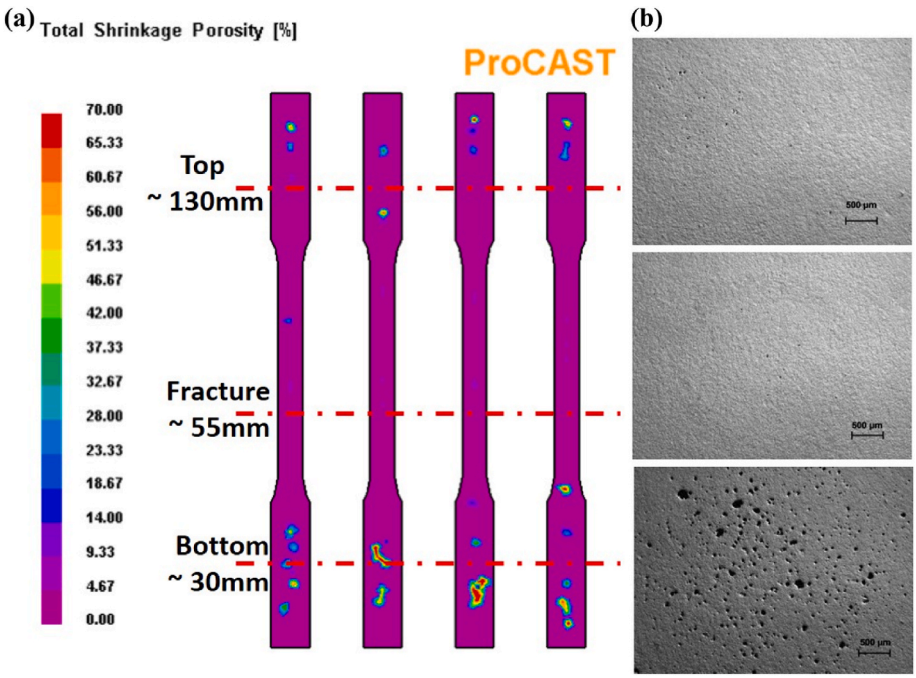

Fig. 7. Comparison of shrinkage porosities for the as-cast HPDC components between (a) modelling and (b) OM images at specific locations indicated by the red dot dash lines on modelling results.

that for the locations where the modelling indicated a larger value in the total shrinkage porosity legend, obvious porosity can be found in actual OM results, as shown in Fig. 7 (b). Hence, the POROS 1 model is valid for defect prediction. Fig. 8 shows some modelling results of distribution of defects such as temperature at fill time, oxides indicator and air entrainment. It could be seen that for each HPDC component, the distribution of defects varies from each other, which is a key result of elongation variability.

3.2. ML model selection and feature analysis

Table 3 summarizes the performance of various ML models in predicting the elongation of HPDC Al–Si alloys. By comparing three key metrics including R^2, RMSE, and MAE, the differences in prediction accuracy and stability among the models were systematically analyzed. The results indicate that the XGBoost delivered the best overall performance, achieving an exceptional R^2 value of 0.8719 alongside the lowest RMSE (0.3800) and MAE (0.2571), underscoring its capability to capture complex nonlinear relationships inherent in the HPDC process parameters. Gradient Boosting and RF also demonstrated strong predictive

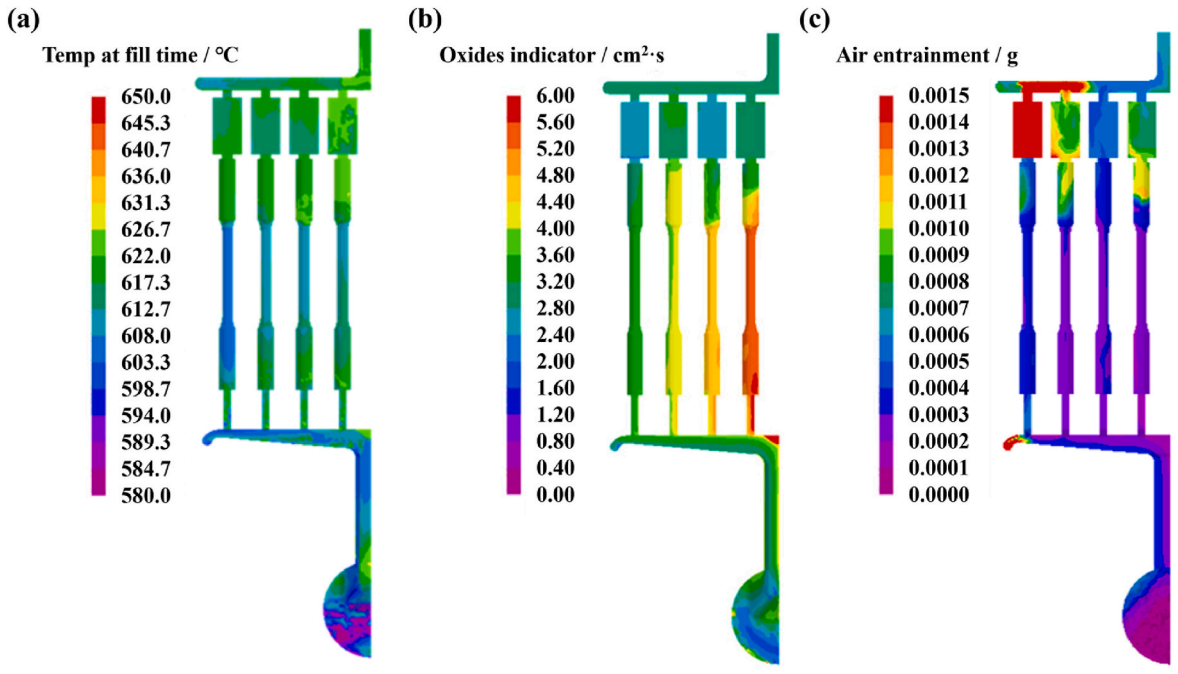

Fig. 8. Modelling results of (a) temperature at fill time, (b) oxides indicator and (c) air entrainment.

Table 3
The prediction results of El for HPDC aluminum-silicon alloy.

Model	Test set		
	R^2/1	RMSE/%	MAE/%
Linear	0.3663	0.8452	0.6224
RF	0.7819	0.4959	0.3088
ANN	0.5581	0.7058	0.5557
RidgeCV	0.3601	0.8493	0.6234
SGD	0.3647	0.8463	0.6215
KNN	0.1231	0.9943	0.7633
SVR	0.0735	1.0220	0.6977
Decision Tree	0.2653	0.9101	0.4858
CatBoost	0.6658	0.6139	0.4690
XGBoost	0.8719	0.3800	0.2571
Gradient Boost	0.8307	0.4368	0.2600
LightGBM	0.4455	0.7906	0.5371

capabilities, with R^2 values of 0.8307 and 0.7819, respectively. Conversely, simpler linear models such as Linear, RidgeCV, and SGD exhibited limited performance ($R^2 < 0.37$), due to their inability to model the nonlinear dependencies between process variables and mechanical properties. Notably, KNN and SVR underperformed significantly, with R^2 values of 0.1231 and 0.0735, respectively, suggesting challenges in generalizing to the dataset.

As shown in Fig. 9, the Taylor diagram further highlights the dominance of tree-based models in predicting elongation after optimizing input parameters. The correlation coefficient is represented by the angle, with the top position corresponding to 1.0 and decreasing clockwise. The standard deviation is indicated by radial distance, while the RMSE is denoted by a blue dashed line. Compared to other models, the XGBoost and GB models are positioned closer to the black reference point in the diagram, exhibiting higher correlation coefficients, standard deviations closer to the true values, and lower RMSE values. The XGBoost model achieves the highest correlation coefficient, whereas the GB model demonstrates a higher standard deviation than XGBoost, indicating that the GB model exhibits superior adaptability to data fluctuations. Consequently, both XGBoost and GB models were selected for subsequent hyperparameter optimization to enhance their generalization capabilities.

The importance of hyperparameter optimization in machine learning cannot be ignored, as it directly impacts model performance, generalization capability, and training efficiency. Systematic hyperparameter optimization to identify optimal hyperparameter configurations serves as an effective approach to significantly enhance model performance. In

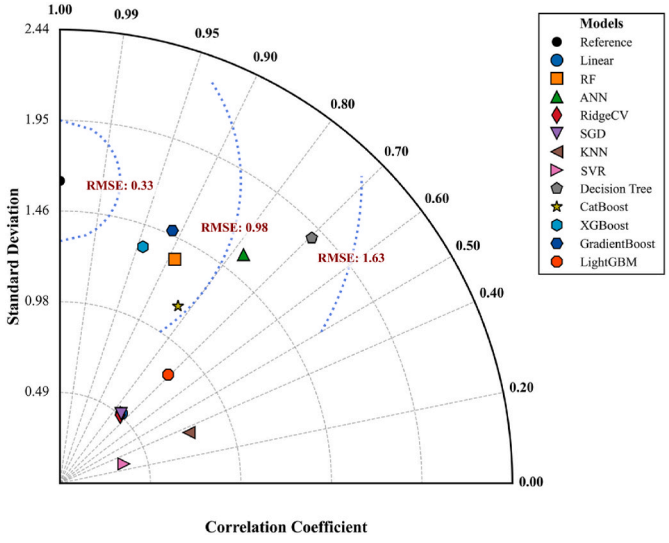

Fig. 9. The Taylor diagram of the predicted results of different ML models.

this study, the employed optimization algorithms evaluated base models on the training set using 3-fold cross-validation, with cross-validation average scores guiding the hyperparameter optimization process to ensure stable and reliable performance assessment. The test set was exclusively reserved for final model evaluation and remained entirely independent of model selection or parameter tuning. To guarantee reproducibility, a random seed was fixed at 42 throughout all experiments. Hyperparameter optimization for both CPO and BO was performed with the objectives of minimizing RMSE, MAE and maximizing R^2. Table 4 presents the hyperparameter search ranges and optimal values for XGBoost and GB models obtained through CPO and BO optimization.

Fig. 10 shows the comparison of the predictive performance of elongation by four different models. A red dashed line indicates the ideal fit and gray points represent sample data. All four models demonstrate high predictive accuracy, with R^2 values exceeding 0.86, indicating strong alignment between predictions and experimental results. Among them, the CPO-XGBoost model achieves the best performance, with R^2, RMSE, and MAE values of 0.882, 0.559 %, and 0.383 %, respectively, slightly outperforming the other three models. Additionally, XGBoost-based models generally surpass GB in elongation prediction tasks, and the CPO optimization method further enhances model performance. CPO required approximately 45 % fewer average iterations than BO when achieving similar performance. Furthermore, the model optimized by CPO attained a higher R^2 and a lower RMSE on the test set compared to the model optimized by BO. However, since the CPO algorithm performs 3-fold cross-validation for each evaluation, its execution time was 4.22 times longer than that of the BO algorithm.

To assess the novelty and effectiveness of the CPO-XGBoost model, some research was conducted and compared in Table 5, those on the prediction of the El of Al-based alloy. The CPO-XGBoost model has superior performance in predicting the El rate of HPDC Al–Si alloy.

Table 4
The optimization of hyperparameters for XGBoost and GB models by CPO and BO algorithm.

Moldel	Hyperparameter optimization algorithm	Hyperparameters and range	Optimized hyperparameters
XGBoost	BO	n_estimators [200, 1200]	971
		max_depth [1,10]	6
		learning_rate [0.001, 0.1]	0.02
		subsample [0.1, 1.0]	0.854
		colsample_bytree [0.1, 1.0]	1.0
	CPO	n_estimators [200, 1200]	1141
		max_depth [1,10]	6
		learning_rate [0.001, 0.1]	0.018
		subsample [0.1, 1.0]	0.822
		colsample_bytree [0.1, 1.0]	0.863
GB	BO	n_estimators [200, 1200]	937
		max_depth [1,10]	7
		learning_rate [0.001, 0.1]	0.008
		subsample [0.1, 1.0]	0.643
		max_features [0.1, 1.0]	0.806
	CPO	n_estimators [200, 1200]	1067
		max_depth [1,10]	8
		learning_rate [0.001, 0.1]	0.011
		subsample [0.1, 1.0]	0.709
		max_features [0.1, 1.0]	1

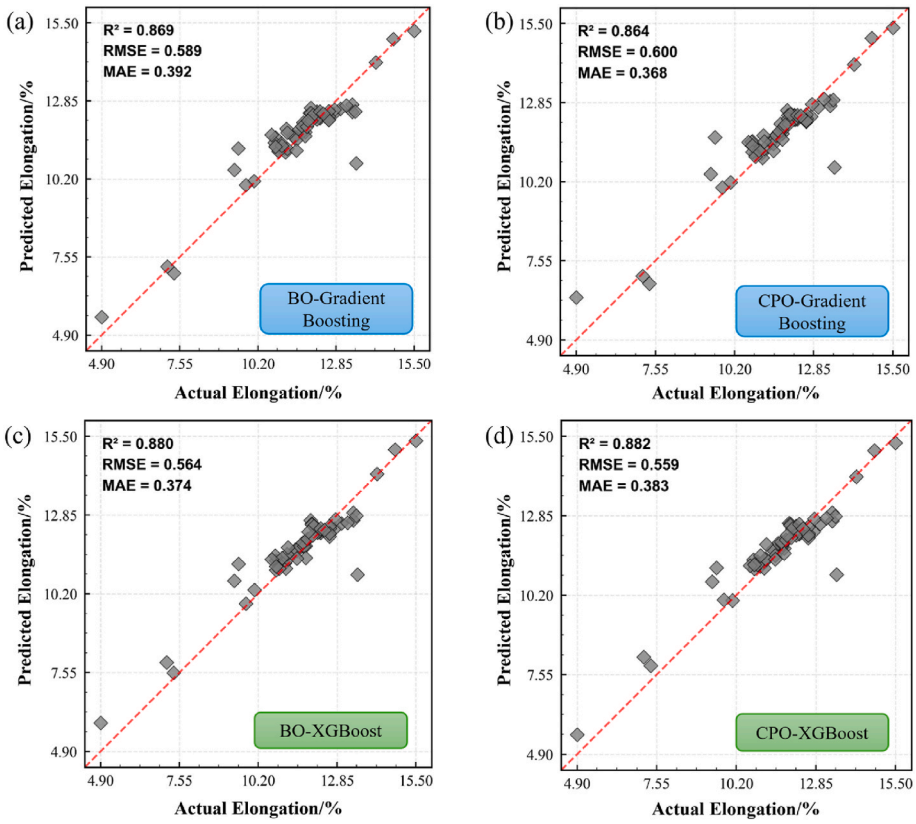

Fig. 10. The prediction performance of the models after hyperparameter optimization.

Table 5
The comparison with previous studies related to El prediction in Al–Si alloys.

Alloy	Model	R^2 of El	Ref.
Thixoformed Al–Si–Cu alloy	RF	0.87	[34]
Heat treatment-free Al–Si–Mg alloy	AdaBoost	0.84	[35]
HPDC Al7Si0.2 Mg alloy	3D CNN	0.786	[36]
HPDC Al–Si alloy	CPO-XGBoost	0.882	This work

However, due to the small number of samples in the model, there is still a lot of room for optimization.

The residuals represent the discrepancies between actual values and the regression line. The goodness of fit of the model is quantified by the sum of squared errors. By analyzing the distribution and trends of residuals, the predictive model's alignment with experimental data can be evaluated. As shown in Fig. 11, the residuals exhibit a random, unstructured pattern following a normal distribution, with a mean of −0.1135 and a standard deviation of 0.5477. These statistical characteristics indicate that the CPO-XGBoost model achieves a robust fit to the data (see Fig. 12).

Based on 5-fold cross-validation, the average absolute SHAP value of each fold model on the validation set was calculated. Fig. 12 (a) presents the box plot of the SHAP values for the features. Total shrinkage volume emerges as the most critical parameter influencing elongation, with a SHAP value of 0.475, significantly surpassing other features. Elongation reflects the material plastic deformation capability during tensile loading. Shrinkage cavities as internal defects, act as stress concentration points, promoting crack initiation and propagation under load,

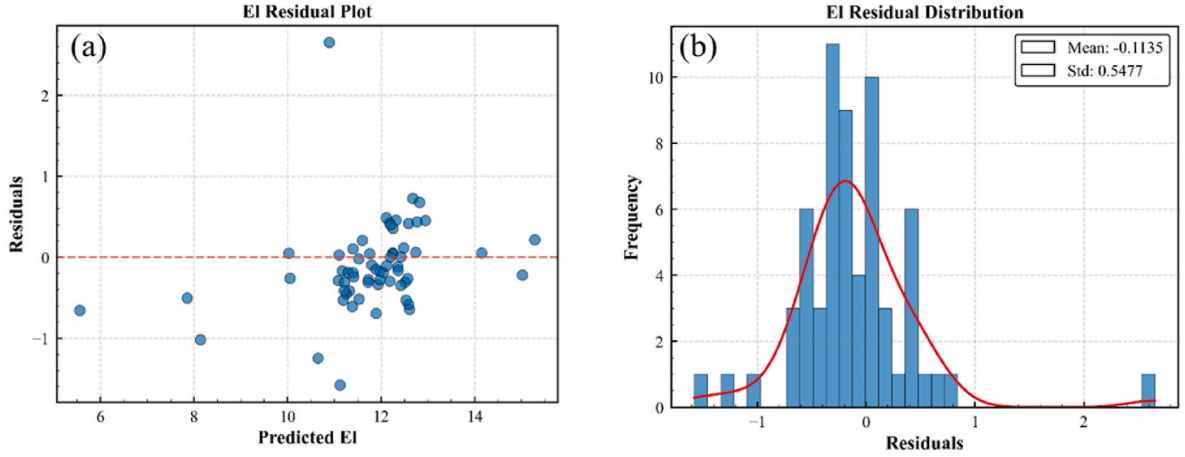

Fig. 11. (a) Residual plot and (b) residual distribution of the elongation of HPDC Al–Si alloy.

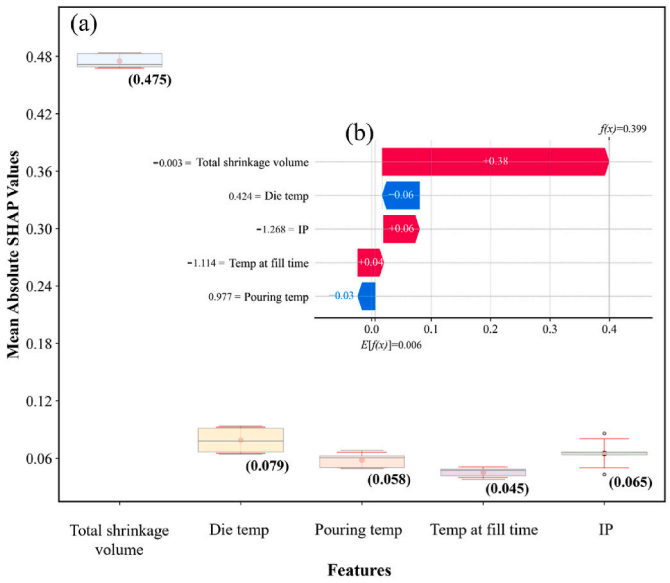

Fig. 12. (a) The box plot of the SHAP values for the features and (b) waterfall plot for an individual sample.

thereby drastically reducing elongation. Consequently, larger shrinkage volumes correlate with lower elongation, aligning with the SHAP analysis.

Fig. 12 (b) illustrates the directional influence and magnitude of the features on the prediction of model for a single sample. In this case, the relatively low total shrinkage volume exerts a positive contribution to the elongation prediction. Likewise, lower values of IP and temperature at fill time also provide a little positive contribution to the elongation prediction. In contrast, elevated die temperature and pouring temperature contribute negatively, with SHAP values of −0.006 and −0.003, respectively.

Total shrinkage volume is identified as the most critical factor in reducing elongation [37–39], as it arises from the cooling and solidification process. However, because shrinkage volume cannot be directly controlled, it is essential to optimize other process parameters to minimize this defect. Moderately reducing the piston velocity helps suppress cavitation and improve feeding during solidification [40], increasing the IP enhances die filling and feeding efficiency, compensating for volumetric contraction during solidification [41]. Temperature variations can lead to antagonistic effects [42,43]. As the temperature increases, the viscosity of the molten metal decreases, thereby improving fluidity and reducing solidification shrinkage and porosity. However, excessively high pouring temperatures may lead to undesirable grain growth, which can degrade the mechanical strength of the casting. These insights suggest that process optimization should prioritize simulation driven and inspection-based strategies to minimize shrinkage defects, followed by control adjustments of temperature and pressure parameters within optimal ranges.

3.3. Validation of ML model

An extra comprehensive evaluation of the model's generalization performance was performed through ten sets of experimental and numerical simulations with different process parameter combinations. Table 6 shows the test values and predicted values under different process parameters. These datasets, containing the same feature columns as those used during training, were imported into the model. The predicted elongation values were then compared with experimental measurements, and a comparative scatter plot was generated, with error bars representing a relative error range of 5 % for the experimental data. As shown in Fig. 13, while slight deviations exist between predicted and measured elongation values, most predictions fall within the experimental error margins. When accounting for the error range comprehensively, these results confirm the high reliability of the prediction of CPO-XGBoost model.

3.4. Future work

While the proposed model demonstrates high predictive accuracy for elongation within the process parameter range, several limitations warrant consideration and suggest avenues for future research. Firstly, the current model is trained on 313 samples generated from a specific

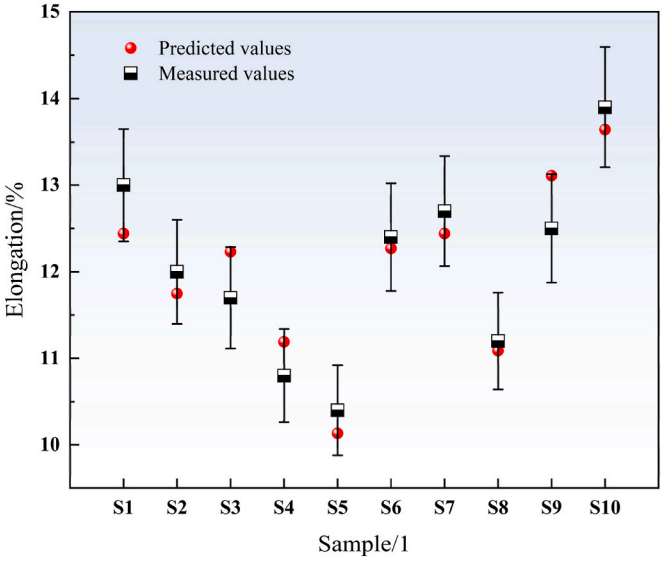

Fig. 13. The comparison between the measured and the predicted values of elongation.

Table 6
The test values and predicted values under different process parameters.

Sample	Features					Predicted values	Test values
	Total shrinkage volume	Die temperature	IP	Pouring temperature	Temperature at fill time		
1	0.23531	150.0	320	672.6	623.3	12.4412	13.0
2	0.22256	140.6	170	677.1	619.3	11.7467	12.0
3	0.24125	166.6	170	671.7	621.5	12.2301	11.7
4	0.28417	144.2	180	670.3	623.6	11.1863	10.8
5	0.28548	165.8	230	674.8	598.7	10.1262	10.4
6	0.23356	129.5	320	676.2	618.8	12.2684	12.4
7	0.23531	150.0	320	672.6	623.3	12.4412	12.7
8	0.24351	149.3	200	677.3	604.6	11.0884	11.2
9	0.23811	127.6	180	675.0	598.2	13.1075	12.5
10	0.18245	155.7	230	672.8	620.8	13.6437	13.9

Al–Si alloy, die geometry, and process parameter ranges. Although sufficient for establishing a robust model within these bounds, expanding the dataset to encompass broader operational windows and material variations would enhance model robustness. Secondly, exploring faster surrogate models for these physics-based simulations or investigating the feasibility of models relying on measurable process parameters could improve practical deployment efficiency. Future work will focus on: (1) Expanding the dataset to include more diverse alloys and process conditions; (2) Extending the ML framework to predict other mechanical properties or defect metrics; (3) Investigating transfer learning techniques to adapt the model to new die geometries or alloys with reduced data requirements; (4) Exploring real-time deployment strategies considering computational trade-offs.

4. Conclusion

1) This study further addressed the limitations of traditional quality assessment methods in HPDC Al–Si alloys by establishing a hybrid framework that integrates numerical simulations with experimental datasets. Solidification and defect formation are studied and validated to obtain reliable datasets are obtained from FEM modelling and HPDC experiments. Datasets were preprocessed through standardization and feature selection, resulting in the screening of five important features. Twelve ML models were evaluated, among which the XGBoost and GB models had excellent prediction performance.
2) Hyperparameter optimization was performed on the XGBoost and GB models using BO and CPO, respectively. The CPO-XGBoost model achieved a superior R^2 of 0.882. The SHAP analysis revealed the key influencing factors, and its stability analysis under different data divisions further confirmed the reliability of the conclusion that the total shrinkage volume is the dominant factor limiting the elongation. Additionally, the die temperature and pouring temperature exhibited negative correlations.
3) The performance of the CPO-XGBoost model was tested using ten independent experimental datasets, with the predicted values falling within the 5 % error margin of measured elongation values. Future work can continuously supplement experimental data while further modifying model parameters to improve the prediction accuracy of elongation. Furthermore, the model should be considered for application in predicting other key mechanical properties and defect formation metric, and it should be expanded for use in the production of various types and complex processes of alloys.

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

The financial supports from the Key Research and Development Program of Xiangjiang Laboratory (22XJ01002), the National Science Foundation of China (52304360) and the National Key Research and Development Program of China (No. 2023YFB3710202) are greatly acknowledged.

References

[1] Dargusch MS, Dour G, Schauer N, et al. The influence of pressure during solidification of high pressure die cast aluminium telecommunications components. J Mater Process Tech 2006;180:37–43. https://doi.org/10.1016/j.jmatprotec.2006.05.001.
[2] Dargusch MS, Hamasaiid A, Dour G, et al. The Influence of In-Cavity Pressure on Heat Transfer and Porosity Formation During High-Pressure Die Casting of A380 Alloy. JOM 2020;72:3798–805. https://doi.org/10.1007/s11837-020-04341-y.
[3] Guo Z, Xiong S, Liu B, et al. Determination of the heat transfer coefficient at metal–die interface of high pressure die casting process of AM50 alloy. Int J Heat Mass Tran 2008;51:6032–8. https://doi.org/10.1016/j.ijheatmasstransfer.2008.04.029.
[4] Gertsman VY, Li J, Xu S. Microstructure and second-phase particles in Low- and high-pressure die-cast magnesium alloy AM50. Metall Mater Trans A 2005;36A:1989–98. https://doi.org/10.1007/s11661-005-0319-5.
[5] Zhang Y, Lordan E, Dou K, et al. Influence of porosity characteristics on the variability in mechanical properties of high pressure die casting (HPDC) AlSi7MgMn alloys. J Manuf Process 2020;56:500–9. https://doi.org/10.1016/j.jmapro.2020.04.071.
[6] Lee SG, Patel GR, Gokhale AM. Variability in the tensile ductility of high-pressure die-cast AM50 Mg-alloy. Scripta Mater 2005;53:851–6. https://doi.org/10.1016/j.scriptamat.2005.06.002.
[7] Lee SG, Patel GR, Gokhale AM, et al. Quantitative fractographic analysis of variability in the tensile ductility of high-pressure die-cast AE44 Mg-alloy. Mat Sci Eng A 2006;427:255–62. https://doi.org/10.1016/j.msea.2006.04.108.
[8] Szalva P, Orbulov IN. Fatigue Testing and Non-destructive Characterization of AlSi9Cu3(Fe) Die Cast Specimens by Computer Tomography, Fatigue Fract. Eng M 2020;43:1949–58. https://doi.org/10.1111/ffe.13249.
[9] Bowles A, Nogita K, Dargusch M, et al. Grain Size Measurements in Mg-Al High Pressure Die Castings Using Electron Back-Scattered Diffraction (EBSD). Mater Trans 2004;45:3114–9. https://doi.org/10.2320/matertrans.45.3114.
[10] Hamasaiid A, Dour G, Loulou T, et al. A predictive model for the evolution of the thermal conductance at the casting–die interfaces in high pressure die casting. Int J Therm Sci 2010;49:365–72. https://doi.org/10.1016/j.ijthermalsci.2009.07.014.
[11] Law M, Hulme-Smith C, Matsushita T, et al. Assessment of Mechanisms for Particle Migration in Semi-Solid High Pressure Die Cast Aluminium-Silicon Alloys. J Manuf Mater Process 2020;4:1–18. https://doi.org/10.3390/jmmp4020051.
[12] Fiorese E, Bonollo F. Simultaneous effect of plunger motion profile, pressure, and temperature on the quality of high-pressure die-cast aluminum alloys. Metall Mater Trans A 2005;47:6453–65. https://doi.org/10.1007/s11661-016-3732-z.
[13] Otarawanna S, Gourlay CM, Laukli HI, et al. Microstructure Formation in High Pressure Die Casting. Trans Indian I Metals 2009;62:499–503. https://doi.org/10.1007/s12666-009-0081-2.
[14] Chiang K, Liu N, Tsai T. Modeling and Analysis of The Effects of Processing Parameters on The Performance Characteristics in The High Pressure Die Casting Process of Al–SI Alloys. Int J Adv Manuf Tech 2009;41:1076–84. https://doi.org/10.1007/s00170-008-1559-5.
[15] Kopper A, Karkare R, Paffenroth RC, et al. Model selection and evaluation for machine learning: deep learning in materials processing. Integr Mater Manuf I 2020;9:287–300. https://doi.org/10.1007/s40192-020-00185-1.
[16] Hart GL, Mueller T, Toher C, et al. Machine learning for alloys. Nat Rev Mater 2021;6:730–55. https://doi.org/10.1038/s41578-021-00340-w.
[17] Rai JK, Lajimi AM, Xirouchakis P. An intelligent system for predicting HPDC process variables in interactive environment. J Mater Process Tech 2008;203:72–9. https://doi.org/10.1016/j.jmatprotec.2007.10.011.
[18] Yarlagadda PKDV, W Chiang EC. A neural network system for the prediction of process parameters in pressure die casting. J Mater Process Tech 1999;89:583–90. https://doi.org/10.1016/S0924-0136(99)00071-0.
[19] Soundararajan R, Ramesh A, Sivasankaran S, et al. Modeling and analysis of mechanical properties of aluminium alloy (A413) reinforced with boron Carbide (B_4C) processed through squeeze casting process using artificial neural network model and statistical technique. Mater Today Proceed 2017;4:2008–30. https://doi.org/10.1016/j.matpr.2017.02.047.
[20] Zheng J, Wang Q, Zhao P, et al. Optimization of high-pressure die-casting process parameters using artificial neural network. Int J Adv Manuf Tech 2009;44:667–74. https://doi.org/10.1007/s00170-008-1886-6.
[21] Dou K, Zhang Y, Lordan E, et al. Understanding the Initial Solidification Behavior for Al–Si Alloy in Cold Chamber High-Pressure Die Casting (CC-HPDC) Process Combining Experimental and Modeling Approach. Metall Mater Trans A 2022;53A:3110–24. https://doi.org/10.1007/s11661-022-06731-0.
[22] Dou K, Lordan E, Zhang Y, et al. A novel approach to optimize mechanical properties for aluminium alloy in High pressure die casting (HPDC) process combining experiment and modelling. J Mater Process Tech 2021;296:117193. https://doi.org/10.1016/j.jmatprotec.2021.117193.
[23] Dou K, Lordan E, Zhang Y, et al. A complete computer aided engineering (CAE) modelling and optimization of high pressure die casting (HPDC) process. J Manuf Process 2020;60:435–46. https://doi.org/10.1016/j.jmapro.2020.10.062.
[24] Lordan Ewan, Zhang Y, Dou K, et al. High-Pressure Die Casting: a Review of Progress from the EPSRCFuture LiME Hub. Metals 2022;10:1575–91. https://doi.org/10.3390/met12101575.
[25] Ileri K. Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks. Int J Mach Learn Cyber 2025. https://doi.org/10.1007/s13042-025-02654-5.
[26] Sarothi SZ, Ahmed KS, Khan NI, et al. Machine learning-based failure mode identification of double shear bolted connections in structural steel. Eng Fail Anal 2022;239:106471. https://doi.org/10.1016/j.engfailanal.2022.106471.
[27] Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 1992;46:175–85. https://doi.org/10.1080/00031305.1992.10475879.
[28] Dorota W, Krzysztof R, Grzegorz G. Comparative analysis of the properties of the nodular cast iron with carbides and the austempered ductile iron with use of the machine learning and the support vector machine. Int J Adv Manuf Tech 2016;87:1077–93. https://doi.org/10.1007/s00170-016-8510-y.

[29] Conduit BD, Jones NG, Stone HJ, et al. Probabilistic design of a molybdenum- base alloy using a neural network. Scripta Mater 2018;146:82–6. https://doi.org/10.1016/j.scriptamat.2017.11.008.
[30] Titus T, Soorya PK, David PC, et al. Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen. Int J Hydrogen Energ 2017;42:28612–21. https://doi.org/10.1016/j.ijhydene.2017.09.149.
[31] Abdel-Basset M, Mohamed R, Abouhawwash M. Crested Porcupine optimizer: a new nature-inspired metaheuristic. Knowl-Based Syst 2024;284:111257. https://doi.org/10.1016/j.knosys.2023.111257.
[32] Lundberg SM, Lee S. A unified approach to interpreting model predictions. Proceedings of the 31st annual conference on neural information processing systems (NIPS). USA: Long Beach; 2017. https://doi.org/10.48550/arXiv.1705.07874.
[33] ProCAST. 0 - user Guide (Visual-Cast 16.0 User's Guide). https://myesi.esi-group.com/downloads/software-documentation/procast-2020.0-user-guide-visual-cast-16.0-users-guide-online-online-online-online-online; 2020.
[34] Beil WL, Veitía BDR, Fals HDC, et al. Machine learning algorithms model applied to predict mechanical properties of thixoformed Al-6wt.%Si-2.5wt.%Cu alloy. J Mater Eng Perform 2025. https://doi.org/10.1007/s11665-025-11174-w.
[35] Dong X, Liu Q, Han W, et al. Intelligent development of high strength and ductile heat treatment-free Al-Si-Mg alloys for integrated die casting through the machine learning of experimental big data. J Alloy Compd 2025;1021:179769. https://doi.org/10.1016/j.jallcom.2025.179769.
[36] Yang Y, Qiu Z, Zheng Z, et al. Al-enabled properties distribution prediction for high-pressure die casting Al-Si alloy. Adv Manuf 2024;12:591–602. https://doi.org/10.1007/s40436-024-00485-1.
[37] Zhang Y, Lordan E, Dou K, et al. Influence of porosity characteristics on the variability in mechanical properties of high pressure die casting (HPDC) AlSi7MgMn alloys. J Manuf Process 2020;56:500–9. https://doi.org/10.1016/j.jmapro.2020.04.071.
[38] Chen JK, Hung HY, Wang CF, et al. Effects of casting and heat treatment processes on the thermal conductivity of an al-si-cu-fe-zn alloy. Int J Heat Mass Tran 2017;105:189–95. https://doi.org/10.1016/j.ijheatmasstransfer.2016.09.090.
[39] Zhu H, Xia C, Zhang H, et al. Design of Non-Heat Treatable High Pressure Die Casting Al Alloys: a Review. J Mater Eng Perform 2024;33:8601–26. https://doi.org/10.1007/s11665-024-09477-5.
[40] Wang T, Huang J, Fu H, et al. Influence of process parameters on filling and feeding capacity during high-pressure die-casting process. Appl Sci 2022;12:4757. https://doi.org/10.3390/app12094757.
[41] Adamane AR, Arnberg L, Fiorese E, et al. Influence of Injection Parameters on the Porosity and Tensile Properties of High-Pressure Die Cast Al-Si Alloys: a Review. Int J Metalcast 2015;9:43–53. https://doi.org/10.1007/BF03355601.
[42] Mo W, Zhang L, Wu G, et al. Effects of processing parameters on microstructure and mechanical properties of squeeze-cast Mg-12Zn-4Al-0.5Ca alloy. Mater Design 2014;63:729–37. https://doi.org/10.1016/j.matdes.2014.07.005.
[43] Santos SL, Antunes RA, Santos SF. Influence of injection temperature and pressure on the microstructure, mechanical and corrosion properties of a AlSiCu alloy processed by HPDC. Mater Design 2015;88:1071–81. https://doi.org/10.1016/j.matdes.2015.09.095.