machines

MDPI

Article

Composition-Aware SDAS Prediction in Recycled Aluminum Alloys via Physics-Informed Machine Learning Guided by Analytical Solidification Physics

Hamed Rezvanpour, Alberto Vergnano *, Paolo Veronesi and Francesco Leali

Enzo Ferrari Department of Engineering, University of Modena and Reggio Emilia, Via P. Vivarelli 10, 41125 Modena, Italy; hamed.rezvanpour@unimore.it (H.R.); paolo.veronesi@unimore.it (P.V.); francesco.leali@unimore.it (F.L.)

* Correspondence: alberto.vergnano@unimore.it

Abstract

The mechanical performance of secondary aluminum alloys depends on Secondary Dendrite Arm Spacing (SDAS). Commercial casting simulations accurately predict local thermal history but typically neglect the influence of compositional variability on SDAS by using fixed material constants. This study introduces a physics-informed machine learning framework to bridge macroscopic process simulation and microscopic solidification physics. A computational Design of Experiments covering 500 AlSi7 alloy variants was generated, and a theoretical SDAS ground truth was calculated using an analytical model incorporating the growth restriction factor. A Gradient Boosting Regressor surrogate was trained to predict the physics-informed SDAS from thermal and chemical inputs. The analysis reveals a solute sensitivity gap, where standard simulations misestimate SDAS by up to 20% for high-impurity batches. The surrogate model captures this variance ($R^2 = 0.95$, MAE = 0.24 µm), enabling rapid, composition-specific microstructural prediction without additional simulation cost. This approach supports the reliable simulation of casting with secondary alloys, where the composition can be hardly considered constant.

Keywords: secondary aluminum alloys; SDAS; physics-informed machine learning; growth restriction factor; process simulation; surrogate modeling

1. Introduction

Structural parts in the automotive and aerospace industries must achieve a high strength-to-weight ratio and strong resistance to corrosion. Additionally, these properties must be delivered in large quantities and complex geometries [1,2]. Considering this, aluminum alloys are the best candidates for lightweight structures as they provide the required strength-to-weight ratio together with corrosion resistance, and aluminum casting is one of the most widely used production methods for delivering large, complex, nearly net-shaped components [3,4]. However, producing primary aluminum is an energy-intensive procedure, accounting for approximately 2% of all greenhouse gas emissions in recent years [5]. On the other hand, recycling of aluminum scraps consumes only 5% of the energy compared to the production of primary aluminum [6]. Therefore, the use of recycled aluminum to produce casting parts is essential to reduce the carbon footprint of the industry, moving toward a greener future to meet the EU Green Deal objectives [7]. However, the increasing use of secondary alloys introduced new challenges in maintaining

Academic Editor: Dimitrios Manolakos

Received: 5 February 2026
Revised: 2 March 2026
Accepted: 6 March 2026
Published: 10 March 2026

uniform material quality. In cast aluminum parts, both global and local mechanical properties are inherently nonuniform and are heavily influenced by the alloy microstructure, which itself depends on the solidification dynamics and the heat transfer map during casting. In addition, maintaining consistent mechanical properties becomes more challenging while using recycled aluminum, which introduces variation in alloy composition. Although conventional simulation tools are available for predicting mechanical properties during the design phase [8], they are often calibrated for nominal alloy composition. Consequently, performing new simulations for every possible compositional variant becomes computationally expensive and time-consuming. This motivates a microstructure-centric approach in which the solidification history is summarized by key descriptors that control mechanical response.

From a microstructural control perspective, the SDAS can be expressed with well-established scaling relations to the local thermal path, typically as:

$$SDAS = K \cdot \dot{T}^{-n} \quad (1)$$

where $\dot{T}$ is the cooling rate and K is a material dependent constant. In aluminum alloys, exponents commonly cluster around $n \approx 0.33$–0.5 [9]. This relation motivates a physics-guided strategy in which thermal parameters are treated as first-order drivers, and composition is analyzed as a bounded secondary effect. In this context, SDAS serves as a measurable microstructural target that links the thermal history imposed by the process to the mechanical behavior of the casting. Therefore, SDAS prediction is a fundamental requirement for controlling casting quality. Recent investigations further indicated that variations in alloying composition, such as Mg, Si, Sr, and Fe, can affect the SDAS by changing the solidification behavior and dendrite growth mechanism. For example, in a combined experimental–numerical study on cylindrical A356 coupons, Bayoğlu et al. showed that increasing the Fe content leads to a reduction in SDAS, but simultaneously degrades overall tensile performance due to the higher fraction of brittle β-AlFeSi intermetallics [10]. Fracchia et al. [11] investigated the role of Si (alongside Fe, Mg, and Cu) SDAS in Al–Si cast alloys, reporting a parabolic dependence for unmodified alloys between SDAS and the averaged alloying content. Furthermore, after addition of Sr, SDAS decreases across all alloys, with the strongest reduction observed for low-Si compositions. Zhang et al. [12] provided evidence that SDAS follows the classical coarsening form:

$$SDAS = A\,(M\,t_f)^b \quad (2)$$

Moreover, under a rotating magnetic field, the ripening exponent increases to $b = 1/2$ for AlSi7; in the Fe-bearing alloy AlSi7Fe1, however, the formation of β-Al_5FeSi suppresses interdendritic flow and the coarsening law remains at $b = 1/3$. Additionally, several studies have demonstrated that SDAS directly affects mechanical properties, such as tensile strength, ductility, and fatigue resistance, proving that a finer SDAS usually increases mechanical strength and ductility by refining the microstructure, whereas larger SDAS often results in decreased mechanical performance [13–18].

Building on prior evidence, SDAS can be treated as the operative microstructural length scale in a Hall–Petch-type framework, where strength increases as the dendrite size decreases. In cast Al–Si alloys, the classic Hall–Petch relation for size strengthening written as $\sigma = \sigma_0 + k\,d^{-1/2}$. Ghassemali et al. studied an Al-10Si cast alloy with Sr-modified Si and Al-5Ti-1B grain refiner under different cooling rates, producing SDAS $\approx$ 6–35 µm and grain sizes $\approx$ 138–300 µm. They also performed conventional tensile tests plus in situ SEM-EBSD to track cracking and deformation. It was concluded that SDAS (not grain size) is the physically meaningful length scale in a Hall–Petch relation for flow stress, and that observed cracks occurred mainly at dendrite/eutectic boundaries with trans-granular

propagation [19]. This SDAS-based Hall–Petch framing aligns with broader size-effect mechanics in which mechanical response scales with the inverse square root of a controlling microstructural length [20].

Beyond thermal control, composition can be compactly summarized by the growth-restriction factor:

$$Q = \sum_i m_i \, (k_i - 1) \, C_{0,i} \tag{3}$$

which captures how solutes impede dendrite/grain growth through liquidus slope m_i and partitioning k_i at concentration $C_{0,i}$ [21–23]. Published tables and reviews for Al alloys report element-wise Q contributions (e.g., Ti, Si, Mg, Fe), enabling a practical, physics-based chemistry descriptor that complements SDAS and the thermal path [24,25]. In this work, Q (alongside composition flags) is used as a bounded chemistry feature together with the thermal SDAS law to quantify how realistic composition variability can shift SDAS within the operating window.

Together with physics-based relations, data-driven machine learning (ML) models are also used in recent studies to capture residual and non-linear material behavior in casting industry. Lan et al., for example, used neural networks to estimate the strength and elongation of A380 from composition, cooling rate, and porosity [1]. To accurately predict the yield strength and ultimate tensile strength of various aluminum alloys, Merayo et al. used a feedforward ANN model [26]. Additionally, Srivastava et al. developed an ANN model to predict the mechanical properties of Al–MWCNT–RHA composites, achieving a regression accuracy greater than 99 % [27]. Jimenez-Martinez et al. used neural network models to assist with the design of secondary aluminum alloys with optimized fatigue life, highlighting the capability of ML to support material design purposes [28]. Shabani and Mazahery used ANN-FEM models to predict A356 alloy mechanical behavior, focusing on macroscopic properties like yield and tensile strength, without addressing dendritic features such as SDAS [29]. Similarly, Oh et al. utilized neural networks to optimize mechanical properties in titanium alloys, demonstrating the broader applicability of ML in alloy systems, yet without addressing microstructural attributes like SDAS [30]. More recently, Dorbane et al. used neural networks and Gaussian process regression to model stress–strain curves of Al6061 across temperatures, focusing on mechanical response rather than microstructural features like SDAS [31].

Although the effect of SDAS on the final mechanical properties is widely explored and ML models are capable of predicting the mechanical properties of casting parts, few models are capable of predicting SDAS in secondary aluminum alloys while cleanly separating thermal effects from composition changes. Unlike prior ML studies, which focus primarily on predicting macroscopic mechanical properties, the present work introduces a physics-informed machine learning (PIML) framework to target SDAS as the primary microstructural descriptor. By integrating standard Finite Volume Method (FVM) thermal simulations with analytical solute-coarsening theories, this study establishes a computational surrogate model that explicitly accounts for realistic composition variability in secondary AlSi7 casting, bridging the gap between macroscopic process simulation and microscopic physical theory.

2. Materials and Methods

This study adopts a PIML method to enhance the predictive capabilities of standard casting simulations. By integrating FVM simulations with analytical solidification theories, the proposed framework explicitly captures the influence of alloy composition on microstructural evolution. The overall workflow, illustrated in Figure 1, replaces traditional experimental validation with a robust ground-truth dataset derived from solute-dependent

coarsening physics. The following subsections detail the computational DoE, the analytical framework, and the ML architecture utilized to construct this predictive surrogate model.

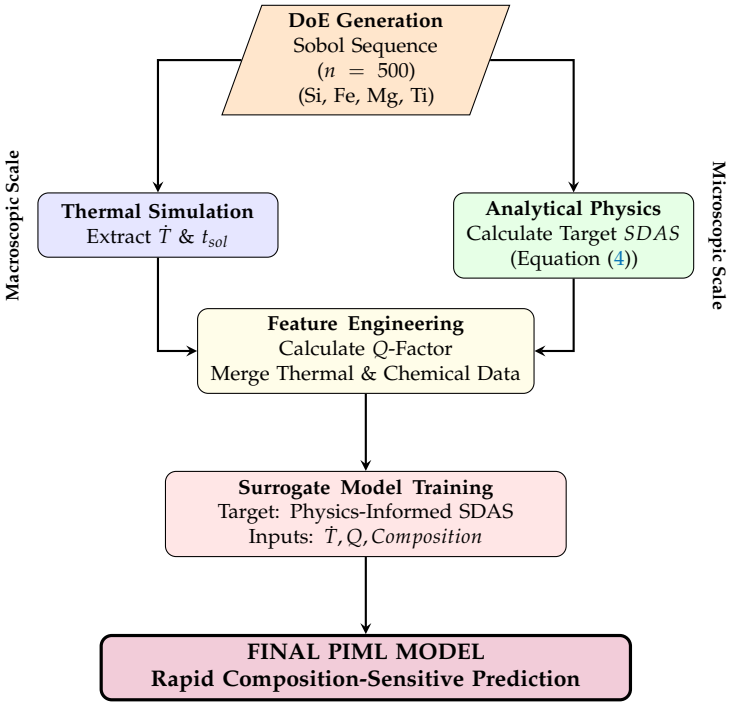
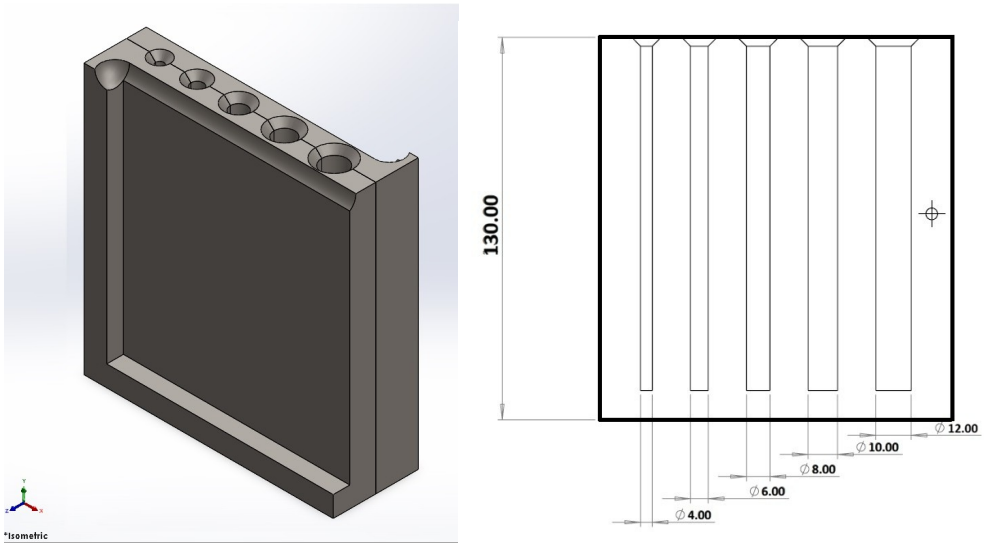


Figure 1. Flowchart of the PIML framework. The methodology bridges the gap between macroscopic thermal simulations (MAGMA) and microscopic solute physics by training an ML surrogate model on theoretically derived ground-truth data.

2.1. Computational Design of Experiments (DoE)

To decouple the effects of thermal history and composition on microstructural evolution, a comprehensive DoE was generated using the commercial casting simulation software Magmasoft® (Version 6). A virtual dataset consisting of 500 unique simulations was conducted to map the solidification behavior of secondary AlSi7 alloys using a cylindrical mold design with different cylinder diameters to simulate different solidification times, as shown in Figure 2.



(**a**) (**b**)

Figure 2. Mold design: (**a**) isometric and (**b**) front view.

To ensure a uniform and unbiased coverage of the multi-dimensional compositional space, the input variables (Si, Fe, Mg, Ti) were sampled using a Sobol quasi-random sequence. Unlike standard Monte Carlo sampling, the Sobol sequence minimizes the clustering of data points, ensuring that the boundaries of the operational window are explored as thoroughly as the center. The range of compositional variations investigated in the DoE is detailed in Table 1. A fine compositional resolution was selected to explicitly resolve the steep, non-linear gradients associated with titanium-driven growth restriction, which would otherwise be obscured by coarser sampling strategies. To achieve this high-fidelity coverage across the multi-dimensional solute space without computationally prohibitive costs, a Sobol low-discrepancy sequence was utilized.

Table 1. Compositional variation ranges used in the computational DoE.

Element	Min (wt%)	Max (wt%)	Step Size
Si	6.50	7.50	0.1
Fe	0.1	0.60	0.1
Mg	0.25	0.45	0.1
Ti	0.1	0.5	0.05

The simulations utilized a stepped casting geometry (cylindrical moduli of 4, 6, 8, 10 and 12 mm) to generate a wide range of cooling rates. Thermal history data ($\dot{T}$, t_{sol}) were extracted from the geometric center of each cylinder to serve as the macroscopic input for the predictive model.

2.2. Physics-Based Hybrid Modeling Framework

To overcome the solute-sensitivity limitations of the commercial FVM solver, this study establishes a hybrid ground truth that integrates macroscopic thermal history with microscopic solute kinetics. Although the commercial solver accurately computes the thermal field ($\dot{T}$), it utilizes a simplified power law ($SDAS = K \cdot \dot{T}^{-n}$) that neglects complex solute interactions. To correct this, we calculate the target SDAS (λ_{phy}) using a rigorous semi-empirical adaptation of the solute-dependent coarsening model.

It is critical to distinguish the proposed model from the classical framework of Easton and StJohn [21]. Although Easton and StJohn identified the growth restriction factor (Q) as the governing thermodynamic parameter, their original model ($d_{gs} \propto 1/Q$) describes nucleation-driven grain refinement. In contrast, SDAS is governed by Ostwald Ripening (diffusion-driven coarsening). Consequently, we integrated the Q-factor into the cubic coarsening kinetics established by Roósz and Exner [32], yielding an adapted constitutive law that respects both the chemical thermodynamics and the diffusive physics:

$$SDAS = K_{cal} \cdot \left(\frac{1}{Q_{total} \cdot \dot{T}} \right)^{1/3} \tag{4}$$

where $\dot{T}$ is the local cooling rate extracted from the FVM simulation, and Q_{total} is the summation of the growth restriction effects of all alloying elements mentioned in Equation (3). The solute parameters m and k for the Al-Si-Fe-Mg-Ti system are summarized in Table 2.

Crucially, the coarsening coefficient K_{cal} is not an arbitrary fitting parameter. To ensure the hybrid model remains consistent with established industrial standards, K_{cal} was determined via a Baseline Anchoring Strategy. The coefficient was mathematically derived by equating Equation (4) to the commercial solver's output specifically for the standard binary alloy (Al-Si7Mg with $Ti < 0.01$ wt%). This calibration ensures that the hybrid model reproduces the trusted commercial baseline for standard alloys (Residual ≈ 0), isolating the ML task to predicting the discrepancy caused by potent grain refiners like titanium.

Table 2. Solute growth restriction parameters used for the calculation of the Q-factor [21,23,24,33].

Element	Liquidus Slope m [K/wt%]	Partition Coeff. k	Growth Factor $m(k-1)$ [K/wt%]
Si	−6.6	0.11	5.9
Fe	−3.0	0.02	2.9
Mg	−6.2	0.51	3.0
Ti	33.3	≈9.0	220.0

2.3. Physics-Informed Machine Learning Architecture

To address the solute sensitivity limitations of the commercial finite element solver, this study employs a PIML framework. Unlike traditional black-box machine learning which maps inputs directly to outputs without physical constraints, or standard Physics-Informed Neural Networks (PINNs) which embed physical laws as soft penalties, the proposed architecture utilizes a hard-constraint hybrid approach. This method leverages the FVM simulation ($\mathcal{M}_{FVM}$) as a baseline for thermal history and geometry, while employing an analytical physics model ($\mathcal{P}_{phy}$) to serve as the ground truth for microstructural kinetics. The machine learning model is strictly tasked with learning the *discrepancy* (or residual) between the thermal-only simulation and the solute-aware physics. While the target dataset in this study is generated via an analytical relationship (Equation (4)), the introduction of a ML surrogate serves a dual purpose. First, it ensures computational efficiency: rather than extracting node-by-node thermal data from the 3D FVM mesh to manually compute Equation (4) for every new scrap composition, the ML model learns a universal relative correction factor that can be instantly applied to the static FVM baseline. Second, this residual learning architecture future-proofs the framework. Analytical models are idealized. by establishing a surrogate model capable of handling synthetic noise (representing measurement uncertainty), the architecture is primed to be fine-tuned with empirical, noisy industrial data in future iterations, capturing complex thermodynamic interactions that strict analytical equations may oversimplify.

The distinction between the standard PINN approach and the residual learning approach used in this work is mathematically fundamental. A standard PINN operates by minimizing a composite loss function where the physics is treated as a regularization term (soft constraint) [34]:

$$\mathcal{L}_{PINN} = \mathcal{L}_{data} + \lambda_{phys} \cdot ||\mathcal{N}(u) - f||^2 \tag{5}$$

where $\mathcal{N}(u)$ represents the differential equation governing the system. While effective for solving partial differential equations with sparse data, PINNs are computationally inefficient for this application because the analytical solution for growth restriction is already known. In contrast, our PIML framework formally defines the final predicted SDAS as the sum of the simulation baseline and a learned correction term [35]:

$$SDAS = \mathcal{M}_{FVM}(\mathbf{x}) + \delta_{ML}(\mathbf{x}) \tag{6}$$

where $\mathbf{x}$ represents the input vector of alloying elements (Si, Fe, Mg, Ti) and local cooling conditions. Input data for training the PIML algorithm is listed as follows:

- t_{sol}
- Q_{total}
- $(Fe, Si, Mg, Ti)wt\%$

The target residual δ for training is derived directly from the discrepancy between the analytical physics model (Equation (4)) and the simulation output:

$$\delta_{target} = \mathcal{P}_{phy}(Q, \dot{T}) - \mathcal{M}_{FVM}(\dot{T}) \tag{7}$$

By training the gradient boosting regressor to approximate this specific residual (δ_{target}), the model effectively "injects" the missing solute physics (specifically the growth restriction factor, Q) into the 3D simulation results. This architecture ensures that the final prediction retains the macroscopic accuracy of the commercial solver (e.g., heat transfer in complex geometries) while enforcing the correct microstructural response to potent grain refiners like titanium, which are otherwise ignored by the baseline solver.

To implement the regression function $\delta_{ML}(\mathbf{x})$, this study utilized a Gradient Boosting Regressor (GBR). Unlike deep neural networks which often require massive datasets to converge, ensemble tree-based methods are well-suited for tabular metallurgical data where the interaction between features (e.g., cooling rate vs. solute content) is highly non-linear but the dataset size is finite. The model was constructed using a sequential additive training process, where each subsequent decision tree corrects the residual errors of its predecessor.

To ensure the ML model captures the underlying physical trends rather than simply memorizing the training data (overfitting), its settings were optimized using a grid search with 5-fold cross-validation. A range of values was systematically tested for four key parameters: the number of trees (100, 300, and 500), the maximum depth of each tree (3, 4, and 5), the learning rate (0.01, 0.05, and 0.1), and the minimum number of samples required to split a node (2, 5, and 10). All these combinations were evaluated to identify the configuration that produced the lowest Mean Absolute Error (MAE). For the standard dataset, the optimal settings were determined to be 300 trees, a maximum depth of 4, a minimum split of 2, and a learning rate of 0.05. Furthermore, to evaluate how the model handles real-world experimental uncertainty, random Gaussian noise (representing a typical measurement error of about 1.5%) was artificially added to the training data. When exposed to this noisy dataset, the grid search automatically adapted by selecting more restricted settings: 500 trees, a shallower maximum depth of 3, a minimum split of 5, and a slower learning rate of 0.01. This practical test demonstrates that the grid search process successfully restricts the model's flexibility, forcing it to learn the actual thermodynamic relationships instead of just copying random experimental errors.

For validation, the dataset was partitioned using a stratified 80/20 train–test split. To rigorously test the robustness of the PIML framework, synthetic Gaussian noise ($\sigma \approx 0.25$ µm) was injected into the test set target variables. This procedure simulates the inherent measurement uncertainty of optical microscopy, forcing the model to learn the underlying solute-refinement signal rather than fitting the clean analytical curve.

3. Results

3.1. Thermal History and Nominal Simulation Baseline

The casting process simulation was initialized using the standard gravity die casting parameters summarized in Table 3, representing the nominal processing window for Al-Si7Mg alloys.

Table 3. Process parameters and boundary conditions used for the simulations (MAGMAsoft).

Parameter	Value
Pouring Temperature (T_{pour})	720 °C
Mold Pre-heat Temperature (T_{mold})	250 °C
Die Material	H13 Tool Steel
Heat Transfer Coefficient (HTC)	Temperature Dependent (Standard)

The computational DoE successfully mapped the fluid dynamics and subsequent solidification conditions across the varying mold geometries. Figure 3 illustrates the filling

sequence, confirming a quiescent bottom-filling pattern that minimizes turbulence and air entrapment.

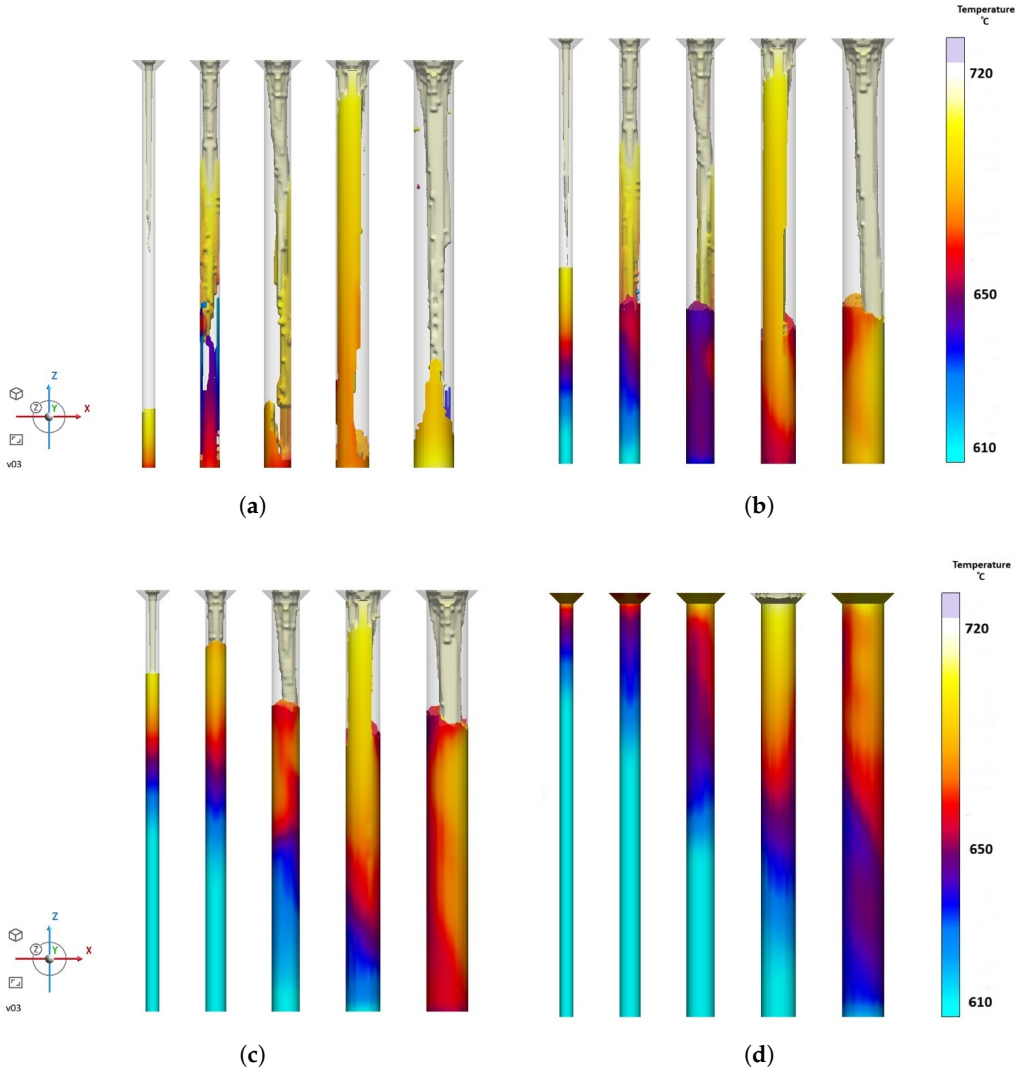


Figure 3. Simulation of the mold filling sequence: (**a**) 25% filled; (**b**) 50 % filled; (**c**) 75% filled; (**d**) 100% filled.

Following the filling analysis, the thermal history was characterized to define the local solidification conditions. To ensure a consistent quantitative baseline, the local cooling rate ($\dot{T}$) was derived specifically from the simulated cooling curves at the geometric center of each step. Figure 4 illustrates the typical thermal profile observed in the simulations. The solidification interval (ΔT) was determined by the difference between the liquidus temperature ($T_{liq} \approx 613$ °C) and the solidus temperature ($T_{sol} \approx 565$ °C), as shown in Figure 4, yielding a characteristic interval of ($\Delta T \approx 48$ K) for the Al-Si7Mg alloy system. The local cooling rate is thus defined as

$$\dot{T} = \frac{T_{liq} - T_{sol}}{t_{sol}} = \frac{48\ \text{K}}{t_{local}} \tag{8}$$

where t_{local} is the time required for the specific section to pass through the solidification range. This definition allows for a direct correlation between the thermal simulation outputs and the microstructural measurements. To ensure data consistency, all quantitative measurements (solidification time and SDAS) were virtually probed at the geometric center of each step section (ϕ4, ϕ6, ϕ8, ϕ10, and ϕ12 mm). This specific sampling location was

selected to isolate the bulk cooling behavior and eliminate edge effects or chill-zone artifacts near the mold walls.

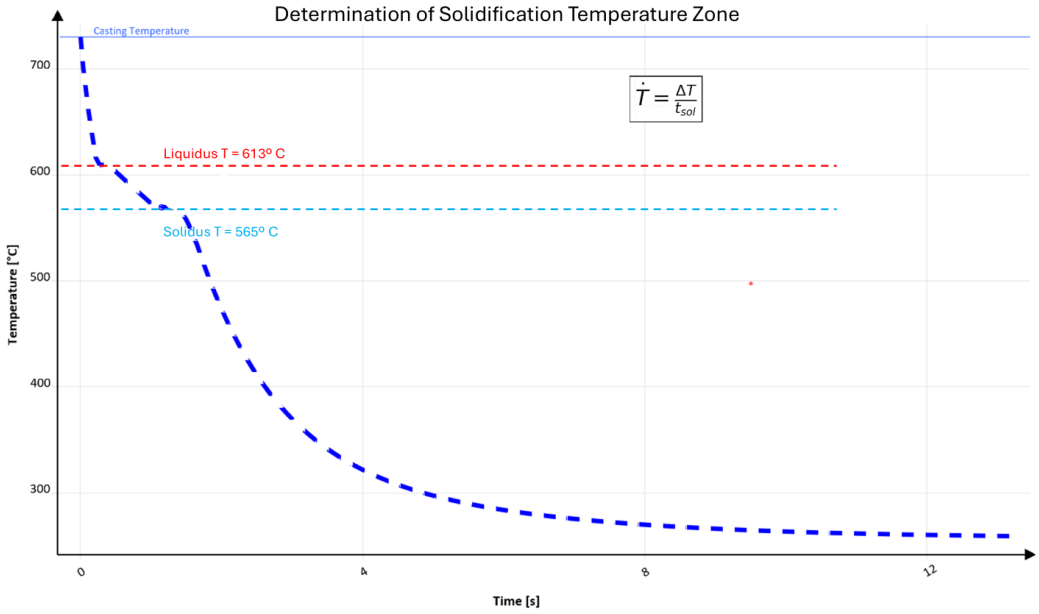
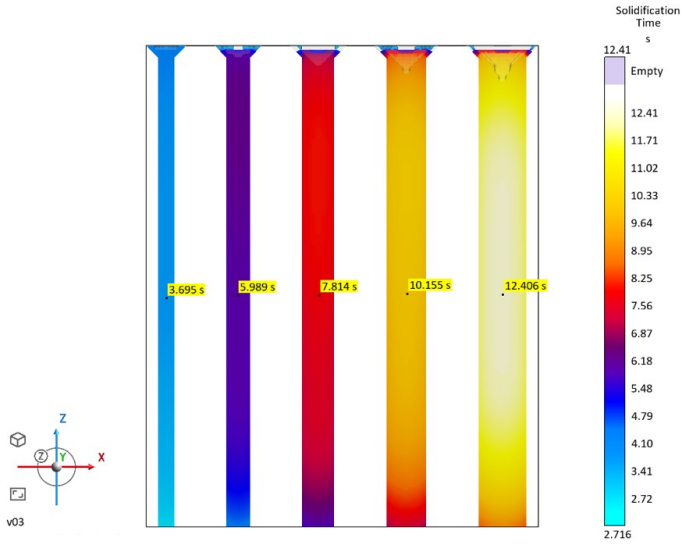


Figure 4. Methodology for extracting the cooling rate ($\dot{T}$) from the simulated thermal history. The solidification interval ($\Delta T \approx 48$ K) is defined between the liquidus and solidus temperatures.

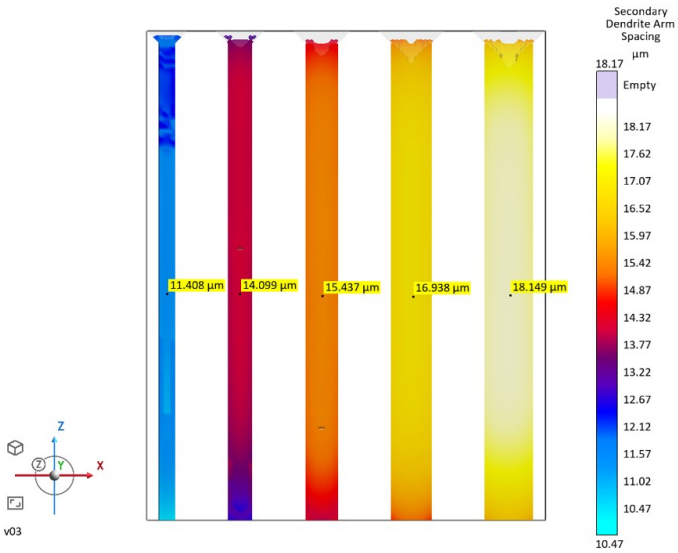
3.2. Finite Volume Method Analysis

The step-mold configuration generated a broad spectrum of local cooling rates ($\dot{T}$), ranging from approximately 4.1 K/s in the thickest 12 mm section to over 20.9 K/s in the thinnest 4 mm section. Figure 5 visualizes the fraction solid evolution and the resulting microstructural variation across the geometry.



(**a**)

Figure 5. *Cont.*



(**b**)

Figure 5. Contour maps from the MAGMAsoft FVM simulation of the step-mold casting: (**a**) local final solidification time, t_{sol}, highlighting the increase in solidification time with section thickness; (**b**) corresponding SDAS field predicted by the commercial solver, showing a coarser dendritic spacing in slow-cooling (thick) regions and refinement in fast-cooling (thin) sections. Quantitative values used in the analysis were extracted at the geometric centers of each step.

Figure 6 presents the aggregated SDAS prediction extracted directly from the MAGMAsoft at these center points. The software efficiently calculates the microstructural trend using a standard calibration for the AlSi7 alloy class. As expected for a macroscopic process simulation, the predicted SDAS follows a consistent power-law curve dependent on the local cooling rate:

$$SDAS_{FVM} \propto \dot{T}^{-1/3} \tag{9}$$

This baseline accurately represents the behavior of a nominal binary composition, providing a reference point for the general solidification trend. However, strictly relying on thermal history implies that the software is insensitive to variations in solute-driven growth restriction.

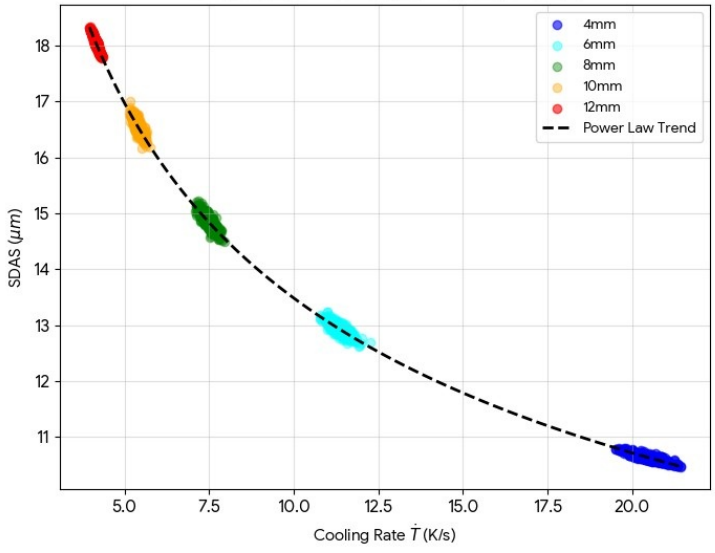


Figure 6. Nominal simulation output: The software provides a consistent baseline prediction based on the cooling rate, representing the behavior of the standard AlSi7 alloy class.

3.3. Sensitivity to Solute-Driven Refinement

While the commercial solver robustly captures the macroscopic thermal trends (Figure 6), its standard calibration is designed to represent a nominal alloy composition. Figure 7 isolates the specific influence of titanium variation on the predicted microstructure for the 12 mm section ($\dot{T} \approx 4.1$ K/s).

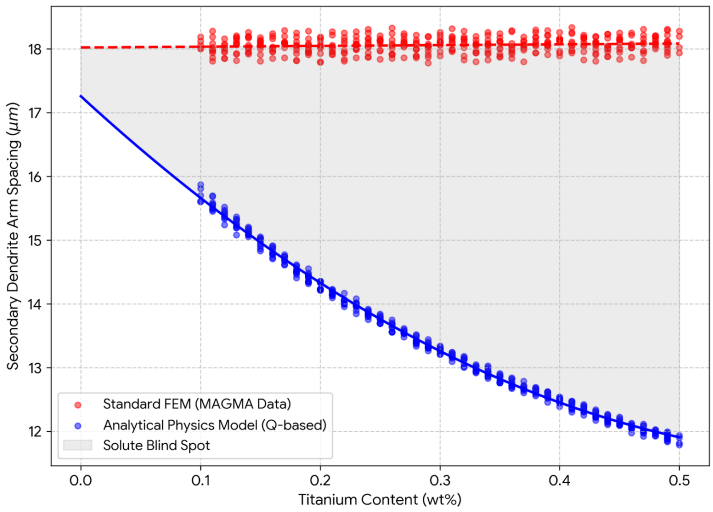


Figure 7. Comparison of SDAS prediction between the standard FVM solver (red dashed) and the PIML (blue solid) as a function of titanium content. The shaded region represents the divergence in predictive accuracy due to solute blindness.

The FVM prediction (red dashed line) remains statistically invariant (SDAS $\approx$ 18.0 µm) across the analyzed compositional range (0.10–0.50 wt% Ti). This behavior reflects the standard formulation of macroscopic process simulations. Consequently, the model relies primarily on local cooling conditions to predict microstructural scale, assuming a constant chemical baseline.

In contrast, the PIML model (blue curve) accounts for the thermodynamic impact of solute segregation. As the titanium content increases from 0.10% to 0.50%, the growth restriction factor rises significantly, driving a corresponding reduction in SDAS from 18.0 µm down to ≈13.5 µm. The divergence between the two curves highlights the scope for enhanced precision in secondary alloy modeling: while the standard thermal approach is sufficient for stable primary alloys, capturing the fluctuations inherent in recycled batches requires accounting for the non-linear thermodynamic effects of trace elements.

In this work, we refer to the solute sensitivity gap as the systematic difference between SDAS predicted by the commercial FVM solver and SDAS predicted by a solute-aware model when the thermal history is the same. The FVM solver uses a calibrated SDAS power-law that depends mainly on cooling conditions and assumes a fixed chemical baseline, so it is largely insensitive to batch-to-batch composition changes. In contrast, the PIML model includes the effect of solute segregation through the growth restriction factor (Q). Therefore, when trace elements change (especially titanium in our alloy window), Q changes and the predicted SDAS shifts, while the FVM prediction remains nearly constant. Figure 7 illustrates this directly: the shaded region represents the solute sensitivity gap (i.e., the SDAS mismatch caused by solute blindness).

It is important to note that the experimental range of titanium (0.10–0.50 wt%) was selected to extend beyond typical industrial limits (<0.15 wt%) intentionally. In secondary aluminum processing, scrap variability can lead to unexpected spikes in trace element concentrations. By expanding the DoE to include these higher-than-usual levels, the study effectively "scales up" the microstructural response. This allows for a clear mathematical

resolution of the non-linear saturation point of the growth restriction factor (Q), ensuring that the model captures the full asymptotic behavior of the alloy system rather than just a linear approximation near the baseline.

3.4. Physics-Informed Machine Learning Performance

To bridge the gap between the static commercial solver and the dynamic physical reality, the PIML model was trained to predict the Relative Refinement Factor (δ_{solute}). This target variable isolates the chemical contribution from the thermal scale, allowing the model to learn a universal correction rule independent of the local solidification time:

$$SDAS_{corrected} = SDAS_{FVM} \cdot (1 + \delta_{solute}) \quad (10)$$

The model performance was evaluated on a held-out test set (20% of the data) with synthetic Gaussian noise ($\sigma \approx 0.25$) added to represent measurement-like variability. Figure 8 presents the parity plot comparing the PIML-corrected predictions against the ground truth. The model achieves a coefficient of determination (R^2) of 0.95, indicating that the chemical correction factor can be predicted with high precision using composition as inputs.

To understand the value of the PIML framework, it is helpful to consider the limitations of standard machine learning. Traditional data-driven models attempt to map inputs directly to outputs without any knowledge of underlying physical laws. Because they rely entirely on data, these models typically require massive datasets to train effectively and often struggle to make accurate predictions when the alloy composition falls outside the normal training range. In contrast, the PIML approach uses known thermodynamic equations as a reliable baseline. Instead of learning the entire process from scratch, the algorithm is only tasked with learning the residual difference between the thermal simulation and the actual physical behavior. This greatly reduces the amount of training data needed and ensures that the final predictions remain physically realistic, even when dealing with the highly variable compositions typical of recycled aluminum alloys.

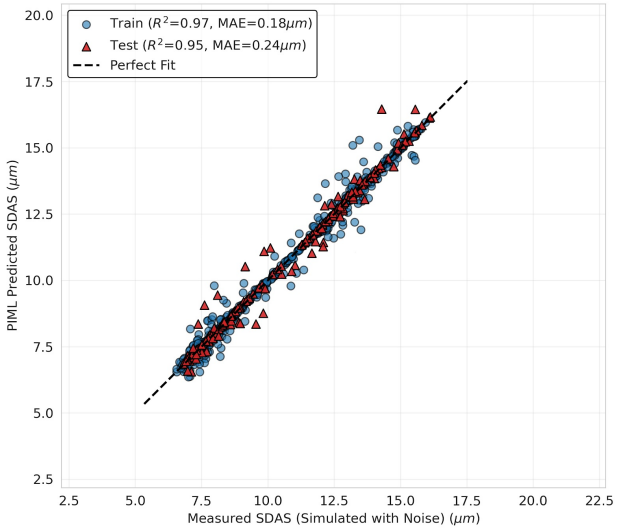
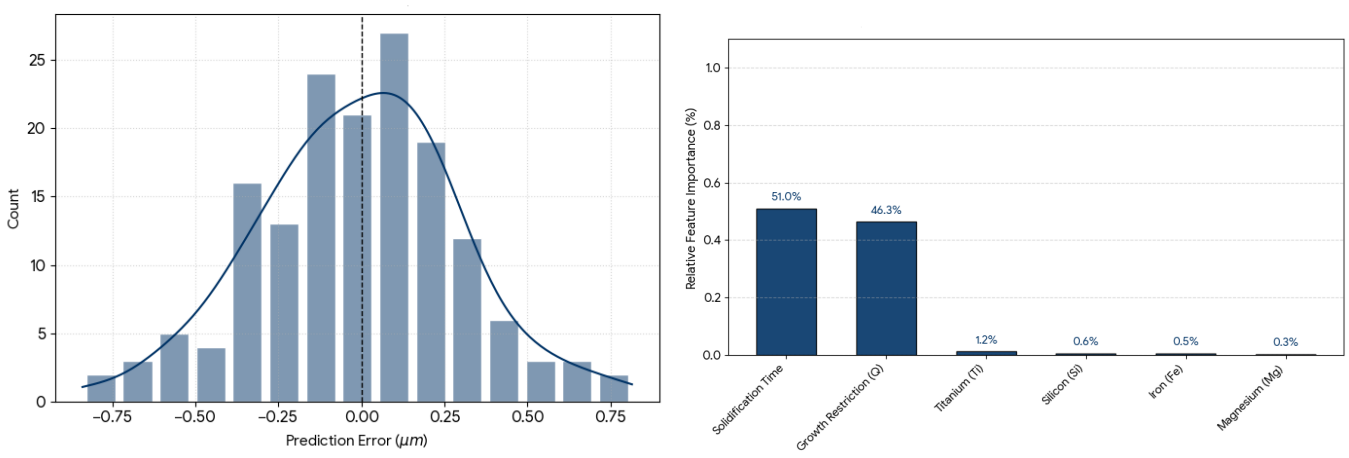


Figure 8. Parity plot comparing the PIML-predicted SDAS against the physical ground truth (simulated with experimental noise). The model's performance is evaluated on both the 80% training dataset (blue circles) and the 20% unseen test dataset (red triangles). The model demonstrates high predictive accuracy across both sets ($R^2_{train} = 0.97$, $MAE_{train} = 0.18$ µm; $R^2_{test} = 0.95$, $MAE_{test} = 0.24$ µm). The minimal generalization gap between the training and testing metrics confirms that the gradient boosting architecture effectively captures the underlying thermodynamic scaling laws without overfitting to the stochastic experimental noise.

The error analysis is further detailed in Figure 9, which shows the distribution of prediction residuals. The errors follow a Gaussian distribution centered at zero, supporting that the GBR has captured key physical trends with limited systematic bias, even in the presence of noise. The robustness of the model is validated through the feature importance analysis (Figure 9). The feature importance analysis (Figure 9b) shows that the correction is mainly controlled by the solidification time and the growth restriction factor Q. Among the individual composition flags, Ti is the most influential element, consistent with its high growth-restriction potency, but most of its effect is already summarized in Q. This confirms that the PIML algorithm is showing limited signs of overfitting considering the fact that the grid search method has been implemented to the hyperparameter tuning.



(**a**) (**b**)

Figure 9. (**a**) Distribution of prediction errors ($SDAS_{Pred} - SDAS_{True}$), showing a normal distribution consistent with random experimental error. (**b**) Feature importance was computed over the GBR inputs (t_{sol}, Q_{total}, $(Fe, Si, Mg, Ti)wt\%$), where Q_{total} is calculated from (3).

3.5. Sensitivity and Robustness Analysis

To address the robustness of the solute sensitivity gap, a systematic sensitivity analysis was performed to verify that the PIML model specifically targets potent grain refiners without introducing spurious corrections for standard alloying elements. Figure 10 presents the model's predicted refinement impact (relative to the FVM baseline) as the concentration of each key element (Si, Fe, Mg, Ti) is swept across its entire experimental range (normalized from 0 to 1).

The analysis reveals a stark contrast in model behavior. For titanium (red solid line), the model predicts a steep refinement trajectory, reducing the SDAS by up to 25% as the concentration increases from 0.0% to 0.5%. This response is physically consistent with the high growth restriction factor of Ti ($m(k-1) \approx 220$ K/wt%). Conversely, for the baseline alloying elements—silicon (6.5–7.5%), iron (0.1–0.6%), and magnesium (0.2–0.5%)—the model's response is effectively flat (near 0% impact). When sweeping Si/Fe/Mg, Ti is held at its reference value (e.g., dataset mean 0.30 wt%), so the predicted correction includes the baseline Ti-driven refinement; hence, the curves are offset but show negligible sensitivity to Si/Fe/Mg. This result provides critical validation of the model's robustness: the PIML framework shows low sensitivity to the noise of standard specification variations, isolating and correcting only the thermodynamically potent impurity fluctuations that drive the discrepancy between simulation and reality.

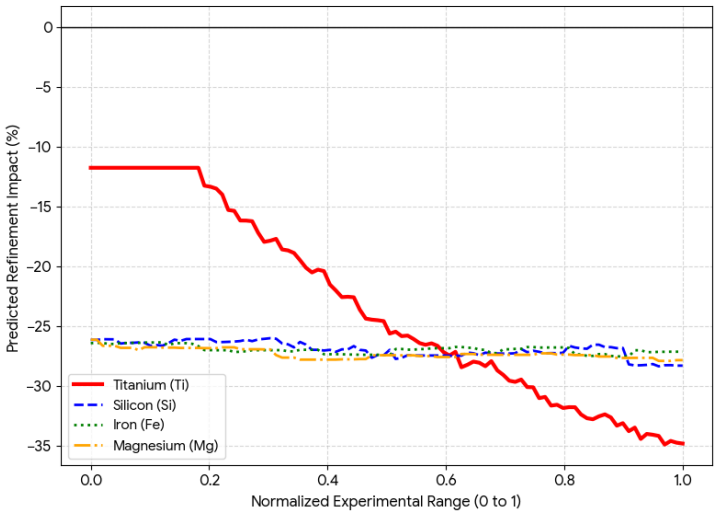


Figure 10. Sensitivity analysis demonstrating the robustness of the solute correction. The model predicts a significant, non-linear refinement of up to 25% for titanium (red solid line), while remaining statistically invariant to fluctuations in silicon (blue dashed), iron (green dotted), and magnesium (orange dash-dot). This confirms the "solute sensitivity gap" is exclusively driven by thermodynamic potency.

4. Discussion

4.1. Decoupling Thermal Scaling from Solute Drivers

The feature importance analysis (Figure 9) provides a critical insight into the limitations of standard casting simulations. The PIML model attributes the predictive weight almost evenly between solidification time (51%) and the growth restriction factor (46%). This split perfectly reflects the physical governing equation of dendritic coarsening, where the microstructural scale is a product of the local cooling condition and the alloy's thermodynamic state.

The dominance of solidification time represents the scaling effect: macroscopic cooling conditions determine the baseline magnitude of the dendrite spacing. Consequently, predictive errors are naturally larger in slower-cooling regions (thick sections) compared to rapid-cooling zones. However, time alone cannot predict the deviation from the baseline.

The second half of the variance is controlled by the chemical driver—specifically the Q factor. This parameter dictates the constitutional undercooling capacity of the melt. Standard FVM solvers effectively treat this term as a constant ($C_{calibrated}$), rendering them blind to batch-to-batch chemistry changes. Even though Q_{total} is calculated from the elemental composition, using both as inputs helps the model learn more reliably. The GBR model uses Q as a general physical baseline. At the same time, including the individual element concentrations helps the model catch complex behaviors or interactions that the simple addition in the growth restriction formula might miss. The PIML model's reliance on Q confirms that accurate prediction requires decoupling these two factors: the thermal history sets the scale, but the solute chemistry sets the slope of the coarsening behavior.

4.2. The Threshold Effect of Titanium

To fully appreciate the microstructural impact of titanium, it is necessary to examine its specific influence on the SDA beyond its traditional role in primary grain boundary refinement. Recent studies demonstrate that titanium is exceptionally potent at restricting the growth of secondary dendrites. For instance, Sun et al. [36] observed that introducing just 0.1 wt.% Ti to an Al-Si-Mg-Cu-Ni alloy reduced SDAS by 30%, dropping the spacing from 9.06 μm to 6.33 μm. This profound reduction is driven by the high Q factor of

titanium relative to its solubility. Because solute titanium effectively inhibits the growth of the advancing solid–liquid interface, it strictly restricts dendritic expansion and enhances the overall uniformity of the microstructure. This mechanism is further supported by Lee et al. [37], who confirmed that even in the absence of primary Ti-based inoculant particles, significant refinement originates directly from the growth restriction exerted by solute Ti. Consequently, the reduction in SDAS fundamentally alters the distribution of subsequent phases. The restricted dendritic growth increases the available grain boundary area, which significantly decreases the degree of secondary phase agglomeration during the final stages of solidification. Furthermore, titanium actively modifies the composition and size of nanoscale features, such as reducing the lattice misfit of $L1_2$ precipitates to refine their size, and drives the formation of complex $Al_{18}Mg_3Ti_2$ intermetallics. However, the efficacy of this SDAS refinement is strictly bounded by the solubility limit of titanium in aluminum, which is approximately 0.1 wt.%. When Ti additions exceed this threshold, SDAS exhibits a reversal, increasing as the refinement effect degrades. This saturation occurs because excess titanium reacts with aluminum and silicon to form coarse, elongated $(Al, Si)_3Ti$ ternary intermetallic phases. The precipitation of these massive phases depletes the surrounding matrix of solute titanium, thereby weakening the constitutional growth restriction effect that is responsible for maintaining a fine SDAS [36,37].

A deeper examination of the chemical driver reveals the disproportionate impact of titanium. As shown in the "Solute-Only" analysis (Figure 7), titanium fluctuations account for over 97% of the chemical variance, effectively masking the contributions of silicon, iron, and magnesium.

This dominance is rooted in the extreme growth restriction potency of titanium ($m(k-1) \approx 220$ K/wt%) compared to standard alloying elements like silicon (≈5.9 K/wt%). In the context of secondary aluminum alloys, where scrap sourcing introduces unavoidable variability in trace elements, this creates a threshold effect. Below a critical titanium level (<0.10 wt%), the physics aligns with the standard calibration. However, as Ti content rises above this threshold—a common occurrence in recycled batches—the thermodynamic driver Q increases non-linearly, causing the titanium anomaly observed in Figure 7. Standard quality control models that monitor only Si and Mg levels will fail to detect this refinement, leading to inconsistent mechanical performance despite in-spec primary chemistry.

5. Conclusions

This study successfully developed and validated a PIML framework designed to enhance the predictive accuracy of commercial casting simulations for secondary AlSi7 alloys. By integrating macroscopic FVM thermal data with microscopic analytical solidification physics, the proposed model effectively bridges the gap between process simulation and solute-driven microstructural evolution.

A key finding of this research is the identification and quantification of the solute sensitivity gap. Standard thermal-only casting simulations do not inherently account for batch-to-batch chemical variations, leading to a potential miscalculation of SDAS in highly alloyed or recycled batches. To address this, the trained Gradient Boosting Regressor demonstrated high predictive capability. The model achieved a coefficient of determination (R^2) of 0.95 and an MAE of 0.24 µm on unseen test data featuring synthetic noise, showing its robustness against typical measurement uncertainty.

Furthermore, feature importance analysis confirmed that microstructural evolution can be decoupled into a thermal scaling factor and a chemical driving factor. While the local solidification time establishes the baseline SDAS scale, accounting for 51.0% of the model's predictive weight, the thermodynamic growth restriction factor (Q) dictates the non-linear refinement slope, accounting for 46.3%. The study particularly highlights the notable

influence of titanium trace elements in recycled aluminum. Variations in titanium can drive a non-linear reduction in SDAS by up to 25%, a phenomenon that traditional quality control and simulation tools may not fully capture when the element exceeds standard nominal thresholds.

By learning a universal relative correction factor, the PIML architecture allows for rapid, composition-specific microstructural predictions. This reduces the need to perform computationally expensive 3D FVM recalculations for every new secondary aluminum composition variant. Ultimately, this hybrid modeling approach provides a scalable tool for the aluminum casting industry, supporting the broader adoption of secondary alloys by ensuring reliable microstructural prediction despite inherent scrap variability. Future work will focus on the empirical validation of these synthetic findings using real-world industrial casting data, alongside extending the framework to predict macroscopic mechanical properties based on the resolved microstructure.

Author Contributions: Conceptualization, A.V. and H.R.; methodology, H.R.; software, H.R. and A.V.; validation, A.V. and P.V.; writing—original draft preparation, H.R.; writing—review and editing, H.R. and A.V.; visualization, H.R.; supervision, A.V. and P.V. and F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Union with the project Recycling Technologies for Circular Aluminium (RecAL), grant agreement ID 101138747, https://recal-project.eu/ (accessed on 1 January 2024).

Data Availability Statement: All data that supported this study can be made available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SDAS	Secondary Dendrite Arm Spacing
PIML	Physics-Informed Machine Learning
FVM	Finite Volume Method
ML	Machine Learning
MAE	Mean Absolute Error
HTC	Heat Transfer Coefficient
GBR	Gradient Boosting Regressor
PINN	Physics-Informed Neural Network

References

1. Lan, Q.; Wang, X.; Sun, J.; Chang, Z.; Deng, Q.; Sun, Q.; Liu, Z.; Yuan, L.; Wang, J.; Wu, Y.; et al. Artificial neural network approach for mechanical properties prediction of as-cast A380 aluminum alloy. *Mater. Today Commun.* **2022**, *31*, 103301. [CrossRef]
2. Wu, X.; Zhang, H.; Cui, H.; Ma, Z.; Song, W.; Yang, W.; Jia, L.; Zhang, H. Quantitative Relationship Analysis of Mechanical Properties with Mg Content and Heat Treatment Parameters in Al–7Si Alloys Using Artificial Neural Network. *Materials* **2019**, *12*, 718. [CrossRef]
3. Javidani, M.; Larouche, D. Application of cast Al–Si alloys in internal combustion engine components. *Int. Mater. Rev.* **2014**, *59*, 132–158. [CrossRef]
4. Li, D.; Slater, C.; Cai, H.; Hou, X.; Li, Y.; Wang, Q. Joining Technologies for Aluminium Castings—A Review. *Coatings* **2023**, *13*, 958. [CrossRef]
5. Sáez-Guinoa, J.; García-Franco, E.; Lera-Sastresa, E.; Romeo, L.M. The effects of energy consumption of alumina production in the environmental impacts using life cycle assessment. *Int. J. Life Cycle Assess.* **2024**, *29*, 380–393. [CrossRef]
6. Quinkertz, R.; Rombach, G.; Liebig, D. A scenario to optimise the energy demand of aluminium production depending on the recycling quota. *Resour. Conserv. Recycl.* **2001**, *33*, 217–234. [CrossRef]

7. European Commission. The European Green Deal Striving to Be the First Climate-Neutral Continent. 2019. Available online: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal (accessed on 11 October 2025).
8. Vergnano, A.; Rezvanpour, H.; Spessotto, M.; Leali, F. Robustness Assessment in the Optimization of Low-Pressure Die Casting Subject to Variations in Secondary Alloy Composition. *Int. J. Met.* **2025**, *19*, 2864–2880. [CrossRef]
9. Jegede, O.E.; Haque, N.; Mullis, A.M.; Cochrane, R.F. Relationship between cooling rate and SDAS in liquid phase separated metastable Cu–Co alloys. *J. Alloys Compd.* **2021**, *883*, 160823. [CrossRef]
10. Bayoğlu, S.; Tunçay, T. The Effect of Iron Content on Microstructure and Mechanical Properties of A356 Cast Alloy. *Metall. Mater. Trans. B* **2017**, *48*, 794–804. [CrossRef]
11. Fracchia, E.; Gobber, F.S.; Rosso, M. Effect of Alloying Elements on the Sr Modification of Al-Si Cast Alloys. *Metals* **2021**, *11*, 342. [CrossRef]
12. Zhang, H.; Wu, M.; Rodrigues, C.M.; Ludwig, A.; Kharicha, A. Directional Solidification of AlSi7Fe1 Alloy Under Forced Flow Conditions: Effect of Intermetallic Phase Precipitation and Dendrite Coarsening. *Metall. Mater. Trans. A* **2021**, *52*, 3007–3022. [CrossRef]
13. Yajjala, R.K.; Inampudi, N.M.; Jinugu, B.R. Correlation between SDAS and mechanical properties of Al–Si alloy made in Sand and Slag moulds. *J. Mater. Res. Technol.* **2020**, *9*, 6257–6267. [CrossRef]
14. Ceschini, L.; Morri, A.; Morri, A.; Gamberini, A.; Messieri, S. Correlation between ultimate tensile strength and solidification microstructure for the sand cast A357 aluminium alloy. *Mater. Des.* **2009**, *30*, 4525–4531. [CrossRef]
15. Wang, Q.; Praud, M.; Needleman, A.; Kim, K.; Griffiths, J.; Davidson, C.; Cáceres, C.; Benzerga, A. Size effects in aluminium alloy castings. *Acta Mater.* **2010**, *58*, 3006–3013. [CrossRef]
16. Wang, Q.; Apelian, D.; Lados, D. Fatigue behavior of A356-T6 aluminum cast alloys. Part I. Effect of casting defects. *J. Light Met.* **2001**, *1*, 73–84. [CrossRef]
17. Li, Y.; Liu, J.; Huang, W.; Zhang, S. Microstructure related analysis of tensile and fatigue properties for sand casting aluminum alloy cylinder head. *Eng. Fail. Anal.* **2022**, *136*, 106210. [CrossRef]
18. Goulart, P.R.; Spinelli, J.E.; Osório, W.R.; Garcia, A. Mechanical properties as a function of microstructure and solidification thermal variables of Al–Si castings. *Mater. Sci. Eng. A* **2006**, *421*, 245–253. [CrossRef]
19. Ghassemali, E.; Riestra, M.; Bogdanoff, T.; Kumar, B.S.; Seifeddine, S. Hall-Petch equation in a hypoeutectic Al-Si cast alloy: Grain size vs. secondary dendrite arm spacing. *Procedia Eng.* **2017**, *207*, 19–24. [CrossRef]
20. Arzt, E. Size effects in materials due to microstructural and dimensional constraints: A comparative review. *Acta Mater.* **1998**, *46*, 5611–5626. [CrossRef]
21. Easton, M.; StJohn, D. An analysis of the relationship between grain size, solute content, and the potency and number density of nucleant particles. *Metall. Mater. Trans. A* **2005**, *36*, 1911–1920. [CrossRef]
22. Quested, T.; Dinsdale, A.; Greer, A. Thermodynamic modelling of growth-restriction effects in aluminium alloys. *Acta Mater.* **2005**, *53*, 1323–1334. [CrossRef]
23. Xu, H.; Xu, L.; Zhang, S.; Han, Q. Effect of the alloy composition on the grain refinement of aluminum alloys. *Scr. Mater.* **2006**, *54*, 2191–2196. [CrossRef]
24. Lei, Z.; Wen, S.; Huang, H.; Wei, W.; Nie, Z. Grain Refinement of Aluminum and Aluminum Alloys by Sc and Zr. *Metals* **2023**, *13*, 751. [CrossRef]
25. Cziegler, A.; Geraseva, O.; Schumacher, P. Numerical and Experimental Investigation of the Influence of Growth Restriction on Grain Size in Binary Cu Alloys. *Metals* **2017**, *7*, 383. [CrossRef]
26. Merayo, D.; Rodríguez-Prieto, A.; Camacho, A.M. Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys. *Materials* **2020**, *13*, 5227. [CrossRef]
27. Srivastava, N.; Singh, L.K.; Yadav, M.K. Utilization of ANN for the Prediction of Mechanical Properties in AlP0507-MWCNT-RHA Composites. *Met. Mater. Int.* **2023**, *30*, 1106–1122. [CrossRef]
28. Jimenez-Martinez, M.; Alfaro-Ponce, M.; Muñoz-Ibañez, C. Design of an Aluminum Alloy Using a Neural Network-Based Model. *Metals* **2022**, *12*, 1587. [CrossRef]
29. Shabani, M.O.; Mazahery, A. The ANN application in FEM modeling of mechanical properties of Al–Si alloy. *Appl. Math. Model.* **2011**, *35*, 5707–5713. [CrossRef]
30. Oh, J.M.; Narayana, P.; Hong, J.K.; Yeom, J.T.; Reddy, N.; Kang, N.; Park, C.H. Property optimization of TRIP Ti alloys based on artificial neural network. *J. Alloys Compd.* **2021**, *884*, 161029. [CrossRef]
31. Dorbane, A.; Harrou, F.; Anghel, D.C.; Sun, Y. Machine Learning Prediction of Aluminum Alloy Stress–Strain Curves at Variable Temperatures with Failure Analysis. *J. Fail. Anal. Prev.* **2023**, *24*, 229–244. [CrossRef]
32. Roósz, A.; Exner, H. Numerical modelling of dendritic solidification in aluminium-rich Al-Cu-Mg alloys. *Acta Metall. Mater.* **1990**, *38*, 375–380. [CrossRef]
33. Han, Q. The role of solutes in grain refinement of hypoeutectic magnesium and aluminum alloys. *J. Magnes. Alloys* **2022**, *10*, 1846–1856. [CrossRef]

34. Raissi, M.; Perdikaris, P.; Karniadakis, G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
35. Parish, E.J.; Duraisamy, K. A paradigm for data-driven predictive modeling using field inversion and machine learning. *J. Comput. Phys.* **2016**, *305*, 758–774. [CrossRef]
36. Sun, X.; Hou, P.B.X.; Liu, S.; Feng, Z.; Cao, G.; Cao, Y. Effect of Ti Content on Microstructure and Mechanical Properties of Al-Si-Mg-Cu-Ni Alloys. *JOM* **2026**, *78*, 2772–2786. [CrossRef]
37. Lee, S.H.; Jung, J.G.; Baik, S.I.; Park, S.H.; Kim, M.S.; Lee, Y.K.; Euh, K. Effects of Ti addition on the microstructure and mechanical properties of Al–Zn–Mg–Cu–Zr alloy. *Mater. Sci. Eng. A* **2021**, *801*, 140437. [CrossRef]