ASME Journal of Manufacturing Science and Engineering
Online journal at:
https://asmedigitalcollection.asme.org/manufacturingscience

Samuele Piandoro
Department of Industrial Engineering,
Alma Mater Studiorum University of Bologna,
Bologna 40136, Italy
e-mail: samuele.piandoro2@unibo.it

Dexiang Zha[1]
Department of Industrial Engineering,
Alma Mater Studiorum University of Bologna,
Bologna 40136, Italy
e-mail: dexiang.zha2@unibo.it

Erica Liverani
Department of Industrial Engineering,
Alma Mater Studiorum University of Bologna,
Bologna 40136, Italy
e-mail: erica.liverani2@unibo.it

Alessandro Ascari
Department of Industrial Engineering,
Alma Mater Studiorum University of Bologna,
Bologna 40136, Italy
e-mail: a.ascari@unibo.it

Alessandro Fortunato
Department of Industrial Engineering,
Alma Mater Studiorum University of Bologna,
Bologna 40136, Italy
e-mail: alessandro.fortunato@unibo.it

A Physics-Informed Neural Network–Based Framework For Near-Real-Time Keyhole Laser Welding Simulation

Achieving stable and repeatable keyhole laser welding requires precise control of process parameters, since even small variations can lead to insufficient penetration, overpenetration, or process instabilities, particularly in high-value assemblies where a single defective weld may result in the rejection of the entire component. Numerical simulation plays a central role in process understanding and optimization; however, state-of-the-art multiphysics models are characterized by high computational costs, which prevent their direct use as predictive tools for online process control despite the growing demand for models that can be integrated within real-time monitoring and control frameworks. To overcome these limitations, a physics-informed neural network (PINN) framework is proposed as a near-real-time surrogate model for keyhole laser welding. The approach embeds the transient heat conduction equation, coupled with a double-conical volumetric heat source, directly into the neural network loss function, avoiding the need for large labelled datasets. The model is calibrated through an inverse analysis using a limited set of experiments, establishing empirical correlations between laser power, scanning speed, and heat source geometry. Validation against experimental data and high-fidelity computational fluid dynamics (CFD) simulations shows good agreement, with relative errors typically below 10% for weld depth and width. Once trained, the PINN predicts the thermal field and weld bead geometry within milliseconds, enabling rapid mapping of the process window and supporting laser welding process optimization.
[DOI: 10.1115/1.4071658]

Keywords: advanced materials and processing, laser processes, modeling and simulation, welding and joining, machine learning

1 Introduction

Laser beam welding (LBW) has emerged as a state-of-the-art technology thanks to its high precision, repeatability, and processing speed, as well as its ability to join both similar and dissimilar metals [1]. New advancements in laser sources and optics have increased the freedom and flexibility for LBW in industrial applications. Strategies like static or dynamic beam shaping have shown promising results [2,3]. In high-power-density regimes, the laser–material interaction leads to the formation of a narrow vapor cavity, commonly referred to as a keyhole, generated by intense material evaporation. The trapping of the laser beam within this cavity significantly enhances energy absorption, enabling deep penetration and high aspect-ratio welds. However, achieving stable and high-quality welds in keyhole mode requires careful fine-tuning of process parameters, including laser power, speed, beam shape, and scanning strategy. This optimization is typically performed through costly and time-consuming experimental trials or computationally intensive numerical simulations, both of which present inherent limitations.

Numerical modeling of laser-based processes generally follows two main approaches: multiphysical and thermomechanical [4]. While the former targets a comprehensive physical description of the process, including melt flow, keyhole formation, and vaporization effects, the latter prioritizes computational efficiency by adopting a purely thermal approach, which enables faster simulations and facilitates their integration into process design and optimization workflows. Multiphysical models, starting from early formulations [5,6], have progressively incorporated key phenomena such as recoil pressure, phase transitions, and laser multiple reflections [7,8]. The predictive capabilities of these models have been validated against experimental data obtained through advanced diagnostics, such as high-speed imaging and synchrotron X-ray measurements, confirming their ability to accurately reproduce melt pool and keyhole dynamics [9]. However, their complexity and computational cost make them less practical for real-time monitoring applications or process control. Even commercial software like FLOW-3D, which is based on the volume of fluid (VoF) method

[1]Corresponding author.
Manuscript received January 9, 2026; final manuscript received April 9, 2026; published online April 29, 2026. Assoc. Editor: Arif S. Malik.

and has been successfully applied to several industrially relevant cases, often require extensive parameter tuning and calibration to ensure reliable predictions [10–13]. Thermomechanical models provide a more practical alternative when the focus is on predicting thermal fields and weld dimensions, rather than capturing fluid dynamics in detail [14]. These models treat the material as remaining in the solid phase throughout the process and simulate heat transfer solely via conduction. Simplified volumetric heat sources are used to approximate the laser–material interaction, ranging from standard types, like surface-based Gaussian distributions, to combined or advanced profiles, such as Goldak's double ellipsoid and double-conical models [15,16]. While these assumptions significantly reduce computational cost, defining the heat source and the temperature-dependent material properties remains a major challenge, as both critically influence the accuracy of the thermal predictions [17,18].

Although numerical simulation remains the primary tool for analyzing manufacturing processes, recent advances in computational resources have enabled the increasing adoption of machine learning (ML) techniques [19]. ML models, inspired by the structure of the human brain, are composed of layered networks of interconnected neurons. These models learn to map input features to output responses by minimizing a cost (or loss) function, which quantifies the discrepancy between predicted and known data. A wide range of architectures has been proposed in the literature, ranging from relatively shallow convolutional models, such as AlexNet [20], to extremely deep neural networks powering large-scale models such as GPT [21]. While physics-based simulations face the challenge of high computational costs, conventional data-driven ML models require large labeled training datasets to achieve accurate predictions. In manufacturing applications, however, generating such datasets, whether through experimental measurements or high-fidelity numerical simulations, is often prohibitively expensive, time-consuming, or technically challenging. To overcome the limitations of conventional simulations and purely data-driven models, hybrid techniques have been proposed that integrate physical knowledge into the learning process. Among these, physics-informed neural networks (PINNs) have emerged as a powerful framework, as they embed governing physical laws—such as conservation of mass, momentum, or energy—directly into the training process [22]. This is achieved by incorporating the problem-specific partial differential equations (PDEs), along with initial (IC) and boundary conditions (BC), into the loss function. In addition, sparse experimental or simulated data—such as temperature measurements through high-speed thermal cameras—can be seamlessly integrated as auxiliary constraints, improving the model's generalization capabilities even with limited datasets. PINNs compute derivatives via automatic differentiation (AD), which ensures high precision and avoids the need for mesh generation or numerical discretization [23].

PINN models have already been applied to laser-based additive manufacturing processes such as laser powder bed fusion (LPBF) [24] and direct energy deposition (DED) [25], but no prior work has addressed their use in laser welding, particularly in keyhole mode. Given the similarities between laser-based additive processes and laser welding, several studies in the literature provide a solid foundation for implementing PINNs in keyhole welding applications. For example, Hosseini et al. [26] employed an unsupervised PINN to predict the temperature distribution and melt pool geometry in a single-track LPBF experiment on Hastelloy X, enabling the estimation of the temperature field and the grain size via a cellular automata model. Sharma et al. [27] incorporated the Navier–Stokes equations and used labeled data generated from high-fidelity finite element method (FEM) simulations to train a PINN capable of predicting both pressure and velocity fields within the melt pool. Peng et al. [24] demonstrated a transfer learning approach in which a PINN, trained on a known dataset, was successfully applied to predict the 3D temperature field in a DED process using only surface temperature data from a high-speed thermal camera, even for a material and laser source not included in the training set. In all these studies, PINNs proved effective at reducing computational costs while maintaining high prediction accuracy, both with and without labeled data.

In the present work, a PINN model is proposed for keyhole-mode laser welding, capable of predicting the temperature field and bead dimensions with lower computational cost and higher speed compared to traditional simulations. Compared to previous models, our proposed framework is tailored to high-power-density scenarios, where the temperature involved reaches the vaporization temperature and the resulting weld bead has the typical deep and narrow shape. The model was defined, although it is not limited to a specific experimental setup, consisting of an overlap welding of two aluminum alloy sheets of 1.5 mm thickness. The PINN architecture, activation function, and collocation point sampling strategy have been systematically optimized. The network was trained without any experimental data, using only the constitutive partial differential equations and boundary conditions that govern the problem. To simplify the model, only the heat conduction equation was directly considered, while the main multiphysics effects were implicitly accounted for through the temperature-dependent material properties. Specifically, the PINN follows a modified Fourier equation that incorporates a volumetric heat source that models the laser–material interaction. To enable the model to automatically adapt to different welding conditions, the heat source's shape and size are defined by five geometrical parameters. The training dataset consisted of a range of input parameters allowing the model to learn how to predict the physically consistent temperature field corresponding to any combination of inputs. The heat source geometric parameters are then correlated with the laser process conditions through an empirical relationship. This calibration is performed using Pearson correlation and a targeted experimental campaign of five weld bead cross sections. Once trained and calibrated, the model can predict the weld bead geometry and temperature field within milliseconds, at negligible computational cost and without manual tuning, only by the laser power and scanning speed. The model is then validated against both experimental welds and high-fidelity numerical simulation. In the end, the PINN model was employed to search for the optimal laser configuration. The model was able to scan the entire parameter range and find the minimum power needed to achieve the desired penetration.

2 Materials and Methods

2.1 General Workflow. The aim of the proposed model is to predict the temperature field and weld bead dimensions during a keyhole laser welding process. The model is presented considering the setup described in Sec. 2.2 with reference to an overlap weld of two 1.5 mm thick Al6082 sheets, but the approach can be extended to any material, dimension, and laser source. The overall workflow was implemented as illustrated in Fig. 1.

Step 1: The PINN, implemented in PYTHON using the PyTorch framework [28], consists of multiple layers of interconnected neurons that transform input variables into outputs through trainable weights and biases, and nonlinear activation functions. The problem's physics is considered through the partial differential equation, boundary and initial conditions defined within the computational domain. The specific PDE is a modified version of the transient Fourier equation for heat conduction described in Sec. 2.3. In the equation, the laser–material interaction is represented as a double-conical volumetric heat source. The model's inputs are the material properties (thermal conductivity and volumetric heat capacity), the laser source parameters (power and scanning speed), and the five geometric parameters of the volumetric heat source (three radii and two heights). In contrast to traditional supervised learning models that rely on labeled datasets, this PINN model does not require labeled data for training. Instead, a range of input parameters is used as the training dataset. The

model's outputs (temperature) at several collocation points are used to evaluate whether the governing physics is satisfied. Whenever the model's predictions deviate from the physical behavior, a residual is computed. The mean squared of these residuals across all collocation points defines the PDE component of the loss function. Similarly, the residuals from the initial condition and boundary conditions are averaged to define the other two components, which are then combined with appropriate weights to calculate the final loss function. A loss function is a mathematical measure of the model's predictive accuracy: the lower its value, the better the model can accurately estimate the temperature.

Step 2: Following the PINN formulation, a hyperparameter tuning phase was conducted to optimize the model configuration, as described in Sec. 3. These hyperparameters, distinct from the trainable parameters (weights and biases of each neuron) updated by the model during learning, include the network architecture (i.e., the number of layers and neurons per layer), activation functions, and the strategies used to sample the collocation points inside the domain. Unlike trainable parameters, hyperparameters are not learned by the model and must be selected by the user, often through trial-and-error or optimization strategies. A comprehensive investigation, considering both standard and advanced hyperparameters, was carried out to find a configuration that ensured fast and stable training.

Step 3: Considering a realistic range for each input parameter, the training phase was performed. The neural network was trained until each loss component converged below a predefined threshold. A curriculum learning strategy was employed as suggested by similar PINN applications [29]. Instead of training the model using the full training range, different subsets were generated with the increasing width. The model is then trained starting with the smallest set, the easiest to learn, and only when the training is sufficiently advanced, the range is increased in a subsequent step. This strategy acts as a numerical continuation method, preventing the optimizer from converging toward physically inconsistent local minima, which is a common challenge in high-gradient PINN applications. By gradually increasing the complexity of the task, the network leverages the knowledge gained from simpler regimes to achieve a more stable and robust convergence across the entire parameter space. After this stage the model is able to predict a physically consistent temperature field given any combination of the inputs inside the training range.

Step 4: Because no theoretical relationship exists between the laser parameters and the volumetric heat source geometry, an experimental calibration was performed (see Sec. 4). A Pearson sensitivity analysis was conducted to quantify the influence of each geometric parameter on the weld bead dimensions. Subsequently, the trained PINN was employed within an inverse analysis algorithm. Five different experimental weld bead cross sections, selected from a subset of the total training range, were used. By minimizing the discrepancy between the experimental and the PINN weld bead dimensions, a constitutive relationship was established between laser power and speed, and the volumetric heat source parameters. As a result, the model can be considered calibrated to the specific laser setup and can predict the temperature field based solely on laser power and speed.

Step 5: After the calibration, the PINN is validated against both experimental welds and numerical simulations (see Sec. 4.1). Four different laser power and scanning speed configurations were used, two inside the inverse analysis range and two outside to assess the generalization capability. The model's performance was evaluated by comparing the weld bead shape and the temperature evolution extracted from the high-fidelity numerical model described in Sec. 4.2. After that, the model can be considered validated and can be used as a surrogate of the laser welding process.

Step 6: Finally, an example of a practical application is presented. The model scans the full range of laser power and scanning speed, utilizing the constitutive relationship found in the calibration phase, to compute the weld penetration and width. The regions of interest, specifically defining the operational boundaries for overpenetration and lack of penetration, are identified.

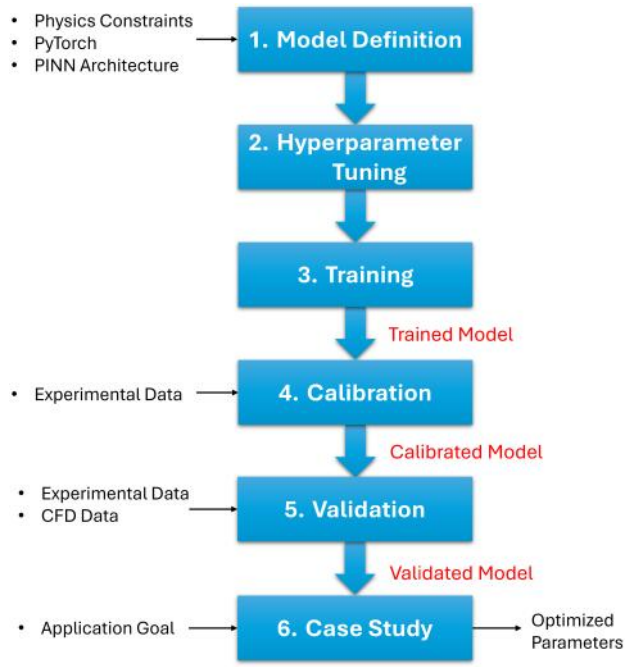

Fig. 1 Flowchart of process optimization

2.2 Experimental Setup. The experimental setup is described in Fig. 2(*a*). An nLight Alta Yb:Fiber CW multimode laser source with a wavelength of 1070 nm, a maximum output power of 3 kW, and a delivery fiber core diameter of 50 μm was employed. The beam was focused and displaced by means of a galvo scanner (Scanlab HurrySCAN 30) equipped with a Wavelength F-Theta fused silica lens. The lens, having a focal length of 160 mm, produced a theoretical spot diameter equal to 65 μm. The laser characteristics are summarized in Table 1. The laser beam directly irradiated the welding area, with the focal point positioned on the sample surface at the center of the joint. The selected configuration consists of an overlap welding of two Al6082 alloy sheets with a thickness of 1.5 mm. The overlapped samples are positioned inside a dedicated clamping system, designed to fix their position and minimize the interface gap. An experimental campaign was carried out for both calibration and validation stages (steps 4 and 5). The process parameters used are reported in Table 2, where P and v denote the laser power (W) and the linear scanning speed (m/s), respectively.

After welding, cross sections of the welded samples were prepared using electrical discharge machining and then mounted in phenol-formaldehyde resin. The samples were ground using silicon carbide abrasive papers (80–2500 grit), polished with a 1 μm alumina suspension, and finally etched with Keller's reagent (a mixture of nitric, hydrochloric, and hydrofluoric acids) to reveal the weld seam geometry (shown in Fig. 2(*b*)). Geometric measurements were carried out using a Zeiss Observer A1M optical microscope, considering three different cross sections for each weld to obtain the average depth (MD) and width (MW).

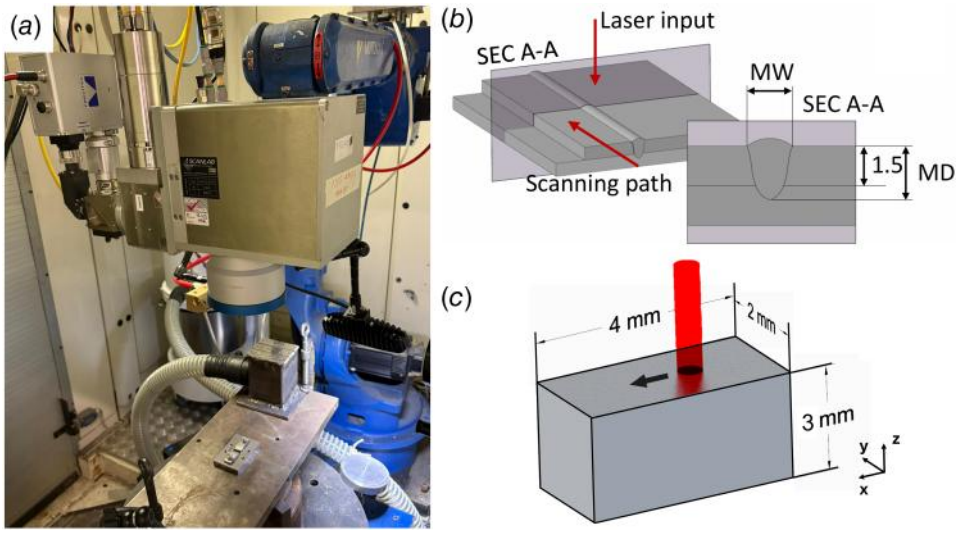

Fig. 2 (*a*) Experimental setup, (*b*) cross section with weld dimensions, and (*c*) domain discretized by the PINN

2.3 Laser Keyhole Welding Model. This section provides a detailed description of the physical model of the LBW process used as an input for the PINN. The computational domain is a simplification of the real experimental setup and is shown in Fig. 2(*c*). It extends 4 mm in the scanning direction (x) [−2,2], 2 mm in width (y) [−1,1], and 3 mm in depth (z) [−3, 0]. These dimensions were chosen to reflect the experimental configuration and to minimize boundary effects that could distort the thermal field. In particular, the domain length ensures that a representative weld of approximately 1 mm can be formed while leaving sufficient space ahead of and behind the heat source to capture realistic cooling and heating gradients. The width and depth of the domain were selected based on the characteristics of the laser system and weld configuration used in the subsequent calibration and validation steps. To allow for thermal preconditioning of the material's initial temperature, the temporal domain begins at −0.1 s, with the laser initially inactive. It ends when the laser completes the 1 mm weld segment. The total simulation time is therefore determined by the laser scanning speed and the target weld length.

To simplify the complex multiphysics phenomena involved in keyhole laser welding, while preserving the dominant thermal effects, the following assumptions were made:

- The material is treated as a continuous solid medium.
- The material is assumed to be isotropic and homogeneous, with temperature-dependent thermophysical properties.
- Fluid flow, recoil pressure, and convection are neglected.
- The laser beam is modeled as an equivalent volumetric heat source.

Under these assumptions, the temperature distribution $T(\mathbf{x}, t)$ within the domain is governed by the transient heat conduction equation described in Eq. (1).

$$\rho\tilde{C}_p \frac{\partial T(\mathbf{x}, t)}{\partial t} - \nabla \cdot (\tilde{k}\nabla T(\mathbf{x}, t)) - q_v(\mathbf{x}, t) = 0 \tag{1}$$

Table 1 Experimental equipment

Maximum power (W)	3000
Collimation length (mm)	120
Focal length (mm)	160
Fiber core diameter (μm)	50
Beam divergence (mm · mrad)	2
Focused spot diameter (μm)	65
Wavelength (nm)	1070

where $\mathbf{x}$ represents the spatial coordinates x, y, and z, t represents time, and $\rho(T)$ (kg/m^3), $\tilde{C}_p(T)$ (J/kg/K), and $\tilde{k}(T)$ (W/m/K) denote the density, apparent heat capacity, and effective thermal conductivity of the material, respectively. The density is considered constant at 2700 kg/m^3, while the other two thermophysical quantities are temperature dependent, as shown in Fig. 3. Specifically, the thermal conductivity $\tilde{k}$ is represented by a solid line, and the volumetric heat capacity $\rho\tilde{C}_p$ by a dashed line. These properties extend the material's intrinsic characteristics to account for phase change phenomena and melt pool dynamics that are not explicitly modeled by the PDE. Below the melting point (933 K), constant values from the literature were adopted. For higher temperatures, an effective thermal conductivity approach was implemented following the methodology of Lampa et al. [30]. Although the physical conductivity of aluminum typically decreases upon melting, $\tilde{k}$ was artificially increased by a factor of 2.5 in the liquid phase to account for the enhanced heat transfer induced by thermocapillary (Marangoni) flow. This simplification is essential in purely thermal formulations to accurately predict the weld bead width where convective effects are dominant. Regarding the energy storage, the latent heat of fusion (0.4 MJ/kg) and vaporization (1.08 MJ/kg) are incorporated into the effective volumetric heat capacity using a smoothed apparent heat capacity formulation [31]. This approach represents the latent heat as increase in $\rho\tilde{C}_p$ near the respective phase change temperatures (933 K and 2600 K), ensuring that the energy balance is maintained during the transition without requiring a separate interface-tracking algorithm. The term q_v represents the volumetric heat source given by the moving

Table 2 Process parameters used for the calibration and subsequential validation step

Parameter set	P (W)	v (m/s)
Calibration set		
CAL_1	1000	0.1
CAL_2	1000	0.3
CAL_3	1500	0.2
CAL_4	2000	0.1
CAL_5	2000	0.3
Validation set		
VAL_1	750	0.05
VAL_2	1250	0.2
VAL_3	1750	0.2
VAL_4	2500	0.5

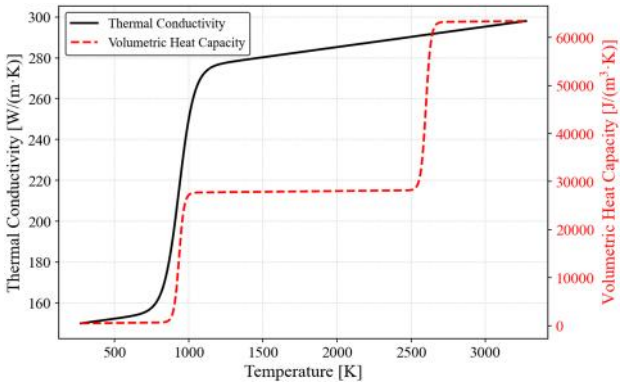

Fig. 3 Temperature-dependent properties used in the model

laser beam. To capture the characteristic morphology of keyhole-mode welds, a double-conical volumetric heat source—originally derived from the three-dimensional conical model by Wu et al. [16] and later refined by Farrokhi et al. [32]—was adopted in this study. Compared to other commonly used volumetric heat sources, notably the Goldak or the single cone, this source was designed for deep or hybrid laser keyhole welding and has been successfully applied in such applications [33]. Given the model's specific task and the specimens thickness, it was selected as the most suitable heat source. The source is shown in Fig. 4 and is composed of two separate volumetric contributions:

- q_{v_l} (see Eq. (2)) describes the lower conical region, extending from the bottom of the keyhole at z_b to the intermediate interface at z_i, over a distance H_l.
- q_{v_u} (see Eq. (3)) defines the upper conical region, extending from z_i to the top surface at z_t over a distance H_u.

$$q_{v_l} = \frac{9\eta P_l e^3}{\pi(e^3-1)(z_i-z_b)(r_i^2+r_i r_b+r_b^2)} \cdot \exp\left(-\frac{3[(x-vt)^2+y^2]}{r_1^2}\right) \tag{2}$$

$$q_{v_u} = \frac{9\eta P_u e^3}{\pi(e^3-1)(z_t-z_i)(r_i^2+r_i r_i+r_i^2)} \cdot \exp\left(-\frac{3[(x-vt)^2+y^2]}{r_2^2}\right) \tag{3}$$

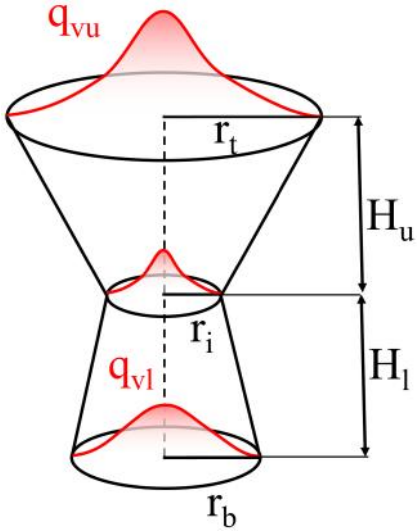

Fig. 4 Double conical heat source with the five parameters shown

In these expressions, η is the dimensionless absorption rate (with fixed value of 0.9), while the total laser power is split into the lower (P_l) and upper (P_u) components. The geometric parameters r_t, r_i, and r_b denote the radii at the top, interface, and bottom surface located at vertical coordinates z_t, z_i, and z_b, respectively. Inside the volume, the heat distribution is Gaussian and the corresponding radial profiles, r_1 and r_2, which vary linearly along the height of each cone, are defined as follows:

$$r_1 = r_i - (r_i - r_b)\frac{z_i - z}{z_i - z_b} \tag{4}$$

$$r_2 = r_t - (r_t - r_i)\frac{z_t - z}{z_t - z_i} \tag{5}$$

By adjusting the five geometric parameters, three radii (r_b, r_i, and r_t), and two lengths (H_u and H_l), it is possible to model a wide range of heat intensity distributions along the workpiece thickness, providing both high flexibility and physical realism.

To solve Eq. (1), appropriate boundary and initial conditions were imposed as detailed below:

- The initial temperature was set equal to room temperature T_0 (25°C) throughout the domain:

$$T(x, y, z, 0) = T_0(x, y, z) = 298\ \text{K} \tag{6}$$

- A Dirichlet BC with fixed temperature at the bottom surface:

$$T(\mathbf{x}, t) = 298\ \text{K}, \quad \mathbf{x} \in \partial\Omega \tag{7}$$

- An adiabatic Neumann BC at the external boundaries:

$$\left.\frac{\partial T}{\partial \hat{n}}\right|_{\partial\Omega} = 0, \quad \mathbf{x} \in \partial\Omega \tag{8}$$

where $\hat{n}$ denotes the outward unit vector to the boundary surface $\partial\Omega$.

- A Robin BC at the top surface to account for heat loss due to external convection and radiation:

$$\tilde{k}\nabla T - h(T - T_0) - \epsilon\sigma(T^4 - T_0^4) = 0 \tag{9}$$

Here, h is the convective coefficient, set to $20\ \text{W/m}^2/\text{K}$, ϵ is the material emissivity, equal to 0.9, and σ is the Stefan–Boltzmann constant, $5.67 \times 10^{-8}\ \text{W/m}^2/\text{K}^4$.

Unlike traditional numerical methods (e.g., finite element or finite volume methods) that rely on mesh-based discretization to approximate derivatives, the solution to Eq. (1) is here approached through a mesh-less framework. The partial derivatives required by the governing equations are evaluated using automatic differentiation. This technique allows for the computation of exact derivatives of the neural network output with respect to its input coordinates at any point within the domain, bypassing the truncation errors of numerical schemes. This approach will be formally integrated into the loss function of the PINN model described in Sec. 3.

3 PINN and Hyperparameters Tuning

This section presents the implementation and tuning of the proposed physics-informed neural network, developed to predict the thermal field during laser keyhole welding (steps 1 and 2 in Fig. 1). The framework combines a feedforward neural network architecture with the information provided by the governing physical law. The neural network is trained to satisfy both the energy conservation equation (Eq. (1)) and the initial and boundary conditions described in Sec. 2.3 without requiring any experimental data. Figure 5 shows the scheme of the proposed PINN model.

To approximate the solution of partial differential equations, PINNs employ standard feedforward neural networks as surrogate function approximators.

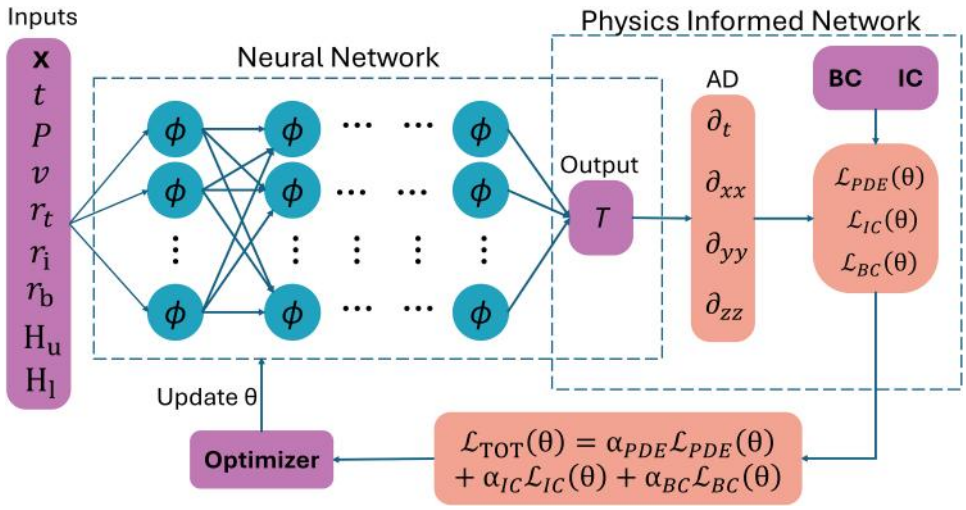

Fig. 5 Scheme of the proposed PINN model

Following the formulation proposed by Lu et al. [34], a fully connected feedforward neural network with L layers is defined as $\mathcal{F}^L(\mathbf{x}) : \mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_{\text{out}}}$, where d_{in} and d_{out} are the input and output dimensions, respectively. The network consists of an input layer, multiple hidden layers with N_l neurons connected through nonlinear activation functions, and an output layer:

$$\begin{aligned} \textit{Input layer :} \quad & \mathcal{F}^0(\mathbf{x}) = \mathbf{x} \in \mathbb{R}^{d_{\text{in}}} \\ \textit{Hidden layers :} \quad & \mathcal{F}^l(\mathbf{x}) = \phi(\mathbf{W}^l \mathcal{F}^{l-1}(\mathbf{x}) + \mathbf{b}^l) \in \mathbb{R}^{N_l} \\ \textit{Output layer :} \quad & \mathcal{F}^L(\mathbf{x}) = \mathbf{W}^L \mathcal{F}^{L-1}(\mathbf{x}) + \mathbf{b}^L \in \mathbb{R}^{d_{\text{out}}} \end{aligned}$$

where ϕ is a nonlinear activation function, $\mathbf{W}^l \in \mathbb{R}^{N_l \times N_{l-1}}$ is the weight matrix, and $\mathbf{b}^l \in \mathbb{R}^{N_l}$ is the bias vector of the lth layer. The trainable parameters of the network, denoted by $\boldsymbol{\theta}$, consist of all the weight matrices and bias vectors.

Considering a generic PDE with a differential operator $\mathcal{N}[\cdot;\lambda]$ parameterized by λ:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathcal{N}[\mathbf{u}(\mathbf{x}, t); \lambda] = 0, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d, \quad t \in [t_0, t_{\text{end}}] \tag{10}$$

with initial and boundary conditions:

$$\mathbf{u}(\mathbf{x}, t_0) = \mathbf{u}_0(\mathbf{x}), \quad \mathbf{x} \in \Omega \tag{11}$$

$$\mathcal{B}[\mathbf{u}] = 0, \quad \mathbf{x} \in \partial\Omega, \quad t \in [t_0, t_{\text{end}}] \tag{12}$$

The PINN approximates the solution $\mathbf{u}(\mathbf{x}, t)$ with $\hat{\mathbf{u}}(\mathbf{x}, t; \boldsymbol{\theta})$ by minimizing the total loss function:

$$\begin{aligned} \mathcal{L}_{\text{TOT}}(\boldsymbol{\theta}) = \alpha_{\text{PDE}} \cdot \mathcal{L}_{\text{PDE}}(\boldsymbol{\theta}) \\ + \alpha_{\text{IC}} \cdot \mathcal{L}_{\text{IC}}(\boldsymbol{\theta}) + \alpha_{\text{BC}} \cdot \mathcal{L}_{\text{BC}}(\boldsymbol{\theta}) \end{aligned} \tag{13}$$

Each term of the loss function is evaluated over a specific set of training points distributed within the domain. The total loss is a weighted sum of these components, where the balancing factors are denoted as α_{PDE}, α_{BC}, and α_{IC}. In particular, the PDE loss term $\mathcal{L}_{\text{PDE}}$ (Eq. (14)) is computed over a batch of N_{coll} collocation points sampled within the space-time domain to enforce the governing differential equation. The boundary condition loss $\mathcal{L}_{\text{BC}}$ (Eq. (15)) is computed over N_{bc} points located on the domain boundaries $\partial\Omega$, whereas the initial condition loss $\mathcal{L}_{\text{IC}}$ (Eq. (16)) is evaluated at N_{ic} points at $t = t_0$. For each loss component, the mean squared error is used to quantify the discrepancy between the model's prediction and the physical constraints. It is crucial to note that this process does not rely on precomputed or labeled datasets. Instead, training is performed on a set of unlabeled coordinates (collocation points) that consist solely of the input features. During each iteration of the optimizer, AD is performed "on the fly" to evaluate the partial derivatives of the network $\hat{\mathbf{u}}$ at the current collocation points. Specifically, AD traverses the computational graph of the neural network to compute exact derivatives (e.g., $\partial\hat{\mathbf{u}}/\partial t$, $\nabla^2\hat{\mathbf{u}}$) at every training step. While this approach avoids the need for generating expensive supervised datasets prior to training, it ensures that the physics is enforced directly and continuously throughout the optimization process, effectively acting as a self-supervised solver.

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_{\text{coll}}} \sum_{i=1}^{N_{\text{coll}}} \left\| \frac{\partial \hat{\mathbf{u}}}{\partial t} + \mathcal{N}[\hat{\mathbf{u}}(t^i_{\text{coll}}, \mathbf{x}^i_{\text{coll}}; \boldsymbol{\theta}); \lambda] \right\|_2^2 \tag{14}$$

$$\mathcal{L}_{\text{BC}} = \frac{1}{N_{\text{bc}}} \sum_{i=1}^{N_{\text{bc}}} \| \mathcal{B}[\hat{\mathbf{u}}(t^i_{\text{bc}}, \mathbf{x}^i_{\text{bc}}; \boldsymbol{\theta})] \|_2^2 \tag{15}$$

$$\mathcal{L}_{\text{IC}} = \frac{1}{N_{\text{ic}}} \sum_{i=1}^{N_{\text{ic}}} \| \hat{\mathbf{u}}(t_0, \mathbf{x}^i_{\text{ic}}; \boldsymbol{\theta}) - \mathbf{u}_0 \|_2^2 \tag{16}$$

A value of 5×10^3 was selected for both α_{BC} and α_{IC}, while α_{PDE} was defined equal to 100. A systematic hyperparameter tuning (step 2) was conducted to identify the most effective network configuration, focusing on four key aspects: the number of hidden layers (L), the number of neurons per layer (N_l), the activation function (ϕ), and the sampling strategy used to generate collocation points. The strategies evaluated are summarized in Table 3.

The objective of the hyperparameter tuning was to identify a reliable model configuration and to evaluate the influence of different parameters on the training stability, rather than conducting an exhaustive grid search. A sequential tuning strategy was adopted, investigating the network architecture first, followed by the activation function and the sampling strategy.

Table 3 Hyperparameters tested

Hyperparameter	Values tested
Hidden layers	4, 8
Neurons per layer	16, 32, 64
Activation function	ReLU, Tanh, Sin, Swish
Collocation points sampling strategy	Grid, Random, Sobol, Moving

Table 4 Ranges of model inputs and output variables

Category	Variable	Range/definition
Time		
Time	t	$[-0.1,\ 1]$
Spatial coordinates		
Scan direction	x	$[-1,\ 3]$ mm
Width	y	$[-1,\ 1]$ mm
Depth	z	$[-3,\ 0]$ mm
Material properties		
Thermal conductivity	$\tilde{k}$	Temperature dependent
Volumetric heat capacity	$\rho\tilde{C}_p$	Temperature dependent
Laser parameters		
Power	P	[500, 2500] W
Scanning speed	v	[0.05, 0.5] m/s
Heat source geometry		
Top radius	r_t	[0.1, 0.7] mm
Intermediate radius	r_i	[0.1, 0.4] mm
Bottom radius	r_b	[0.05, 0.3] mm
Upper cone height	H_u	[0.15, 0.5] mm
Lower cone height	H_l	[0.4, 2.5] mm
Output		
Temperature	T	Temperature field

To ensure a consistent comparison across all tested architectures, the total number of training points during this phase was fixed at 2^{18} (262,144). This value was selected as the maximum resolution achievable within the hardware constraints (specifically the GPU memory limit) when considering the most complex "bottleneck" configuration (8 layers and 64 neurons). These points were distributed as 75% for the PDE collocation ($\mathcal{L}_{\text{PDE}}$) and 25% partitioned between boundary ($\mathcal{L}_{\text{BC}}$) and initial ($\mathcal{L}_{\text{IC}}$) conditions. The use of a power of two for the point count was chosen to optimize the computational throughput of the PyTorch libraries on GPU hardware. As stated, during this tuning stage, training was truncated at 25,000 iterations to assess the convergence trends, while the final selected model was trained until full convergence using the maximum sampling number of points allowed by the GPU memory.

After the hyperparameters tuning, the training range of interest was defined considering the specific experimental setup and realistic values for the volumetric heat source geometrical parameters. Table 4 summarizes the inputs and the output of the model with the corresponding range of training. The maximum and minimum laser power and scanning speed were selected based on the experimental setup in Fig. 2(*a*). Due to the difficulty of training the model over the whole range, a curriculum learning strategy was adopted. The laser power and scanning speed ranges were divided into ten subsets of increasing width, and the training was performed sequentially. This approach is based on the idea of gradually increasing the complexity of the database, in our case the range of parameters. As a result, the model is first trained on simpler configurations and is subsequently exposed to more complex input, allowing it to satisfy the governing PDE for high-power regimes only after converging on lower complexity cases. It has shown promising results when applied to PINNs [35,36], and our study confirms the strength of this approach. The training was performed to minimize the total loss function Eq. (13) using the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm, as suggested by similar physics-informed studies.

It is important to emphasize that the PINN is not trained as a single simulation for a fixed configuration; instead, it is designed as a parametric surrogate model. Specifically, the network receives as input not only the space-time coordinates (x, y, z, t) but also the five geometric parameters that define the volumetric heat source ($\lambda = [r_t, r_i, r_b, H_u, H_l]$). By including these geometric features directly in the input layer ($\mathbb{R}^{d_{\text{in}}}$), the model learns a generalized mapping between the heat source shape and the resulting temperature field T. This parametric nature is fundamental for the subsequent inverse analysis (step 4), as it allows the framework to explore the geometric parameter space nearly instantaneously to find the configuration that best matches experimental observations.

4 Calibration

This section presents the calibration procedure used to define the volumetric heat source model as a function of laser power and scanning speed (step 4). To achieve accurate predictions of the melt pool geometry, it is essential to establish the correlation between the input process parameters and the geometric parameters of the double-conical heat source model (Eqs. (2) and (3)), within the PINN framework. Since no general law exists, the correlation was carried out in two successive stages, as illustrated in Fig. 6:

(1) Perform a Pearson sensitivity analysis to identify the geometric parameters of the heat source model that exhibit strong correlations with the melt pool dimensions.
(2) Perform an inverse analysis of the heat source model using the trained PINN and experimental data.

Pearson correlation analysis was used to quantify the linear relationship between each volumetric heat source parameter (r_t, r_i, r_b, H_u, and H_l) and the resulting melt pool geometry, specifically the depth and width (MD and MW). The Pearson correlation coefficient, r, is defined as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{17}$$

where X_i and Y_i denote individual samples of heat source parameters and melt pool dimensions, respectively, while $\bar{X}$ and $\bar{Y}$ represent their mean values. This statistical measure provides a normalized index of linear correlation, ranging from –1 to 1. Following the Pearson sensitivity analysis, the five welds denoted as calibration set of Table 2 were performed and the experimental MD_{exp} and MW_{exp} were measured.

To extract the melt pool dimensions from the temperature field predicted by the PINN, a differentiable formula is employed as follows:

$$\text{MD}_{\text{PINN}} = \int_{z_{\min}}^{z_{\max}} \mathcal{S}(T_\theta(x, 0, z) - T_m)\, \mathrm{d}z \tag{18}$$

$$\text{MW}_{\text{PINN}} = \int_{y_{\min}}^{y_{\max}} \mathcal{S}(T_\theta(x, y, z_{\max}) - T_m)\, \mathrm{d}y \tag{19}$$

where $\mathcal{S}$ is a smooth approximation function (e.g., a logistic sigmoid), T_m represents the material melting temperature, and T_θ corresponds to the temperature field predicted by the PINN.

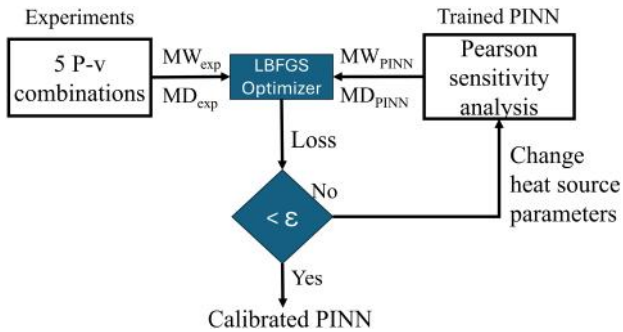

Fig. 6 Calibration flowchart

The inverse analysis for each experiment was initialized with random values for the heat source geometric parameters and performed using the L-BFGS optimizer minimizing the following loss functions:

$$\mathcal{L}_{MD} = \|MD_{PINN} - MD_{exp}\| + \mathcal{L}_{pr} \tag{20}$$

$$\mathcal{L}_{MW} = \|MW_{PINN} - MW_{exp}\| + \mathcal{L}_{pr} \tag{21}$$

The first term of each loss function ensures that the computed depth and width remain close to the experimental value. The second term, $\mathcal{L}_{pr}$, constrains the optimized model parameters to stay within the PINN training range and to be physically realistic, preventing, for example, MD from exceeding the maximum domain depth. Based on the five sets of parameters in Table 2 empirical mapping was constructed through fitting and used in the now calibrated PINN model.

4.1 Validation and Case Study. The model was validated using the four experimental welds listed in Table 2 as "Validation Set," two of them (VAL_2 and VAL_3) are inside the calibration range, while the other two (VAL_1 and VAL_4) are outside. All these configurations were both experimentally performed and numerically simulated. The weld bead dimensions and shapes were compared and, by extracting the temperature field from the numerical simulations, it was also possible to assess the model's capability of predicting the temperature dynamics. The model was then applied to search for the maximum scanning speed and minimum power that would achieve a required MD_{target} of 1.7 mm to 2.7 mm. This range prevents insufficient depth that may result in lack of fusion between the plates, reduced joint strength, and shortened service life. On the contrary, a penetration depth higher than 2.7 mm could result in overpenetration [37,38]. The entire training range was used along with the calibrated relationships found in the previous step within the model.

4.2 Computational Fluid Dynamics Model. This section outlines the numerical framework employed as a benchmark for validating the PINN-based results (step 5). The commercial multiphysics software FLOW-3D WELD (version 2025R1) was used to simulate the laser welding process. The software is based on the VoF method, treating the metal as an incompressible fluid with temperature-dependent properties. A scalar field, called volume fraction F, is computed, equal to the ratio of the volume occupied by the fluid to the total volume of the grid cell. The interface between the liquid metal and the surrounding air is accurately captured, where $0 < F < 1$, allowing the dynamics of the keyhole and the weld bead dimensions to be resolved. The main models used in the simulations are as follows: viscous flow, heat transfer and energy conservation, phase change with evaporation and solidification, volumetric thermal expansion, surface tension with Marangoni force, multiple laser reflections with Fresnel absorption, and gravity. A detailed description of the underlying models is beyond the scope of this work; interested readers are referred to Refs. [2,10] for an in-depth discussion. Material properties and numerical parameters were selected in accordance with the validated setup used in a previous study [2]. By exploiting the symmetry of the laser weld bead with respect to the y-axis, only half of the full 3D domain shown in Fig. 2(*c*) was modeled in the software to reduce computational cost. The domain was discretized using two nonuniform nested structured Cartesian meshes as shown in Fig. 7. A refined mesh with 10 μm cell size was applied in the laser interaction zone to capture fine-scale phenomena, while a coarser thermal diffusion mesh with 40 μm cell size covered the remaining volume with a total number of cells of about 4 million. A fixed atmospheric pressure boundary condition is imposed on the top surface, while the others, except for the symmetric boundary, are defined as walls.

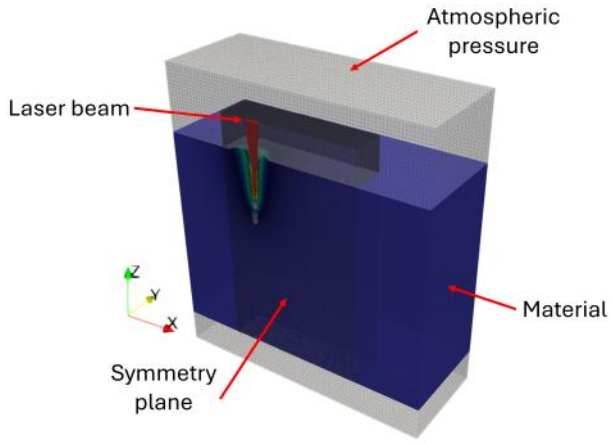

Fig. 7 Computational domain of the high-fidelity CFD simulations

5 Results

5.1 Hyperparameter Tuning. The hyperparameters listed in Table 3 were evaluated for their ability to reduce the loss function, particularly the PDE loss, while keeping a stable training. The four activation functions tested are shown in Fig. 8, along with their first (*b*) and second derivatives (*c*).

Each activation function was evaluated under identical conditions, maintaining all other hyperparameters constant for 25,000 cumulative iterations. The rectified linear unit (ReLU) function, defined as the nonnegative part of its argument, is widely used in machine learning models due to its simplicity and computational efficiency. However, training the proposed PINN model with the ReLU activation function led to numerical instabilities, and the

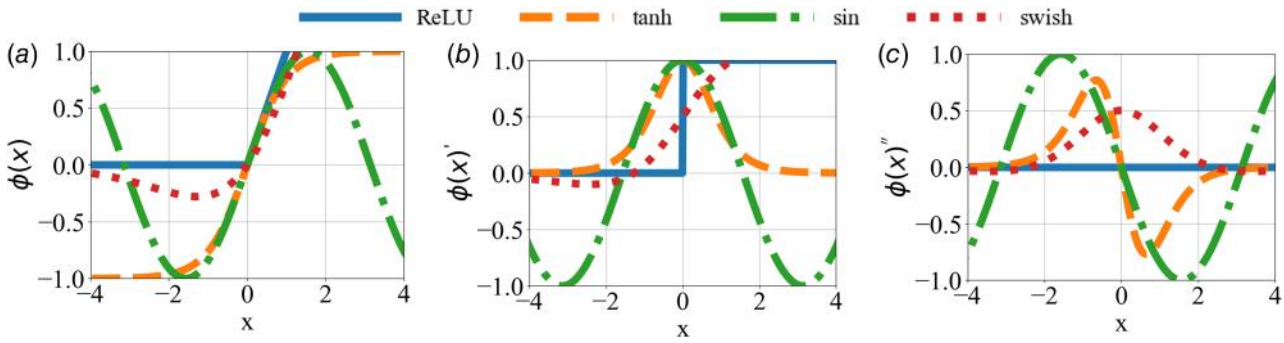

Fig. 8 (*a*) Activation functions tested with the (*b*) first and (c) second derivatives

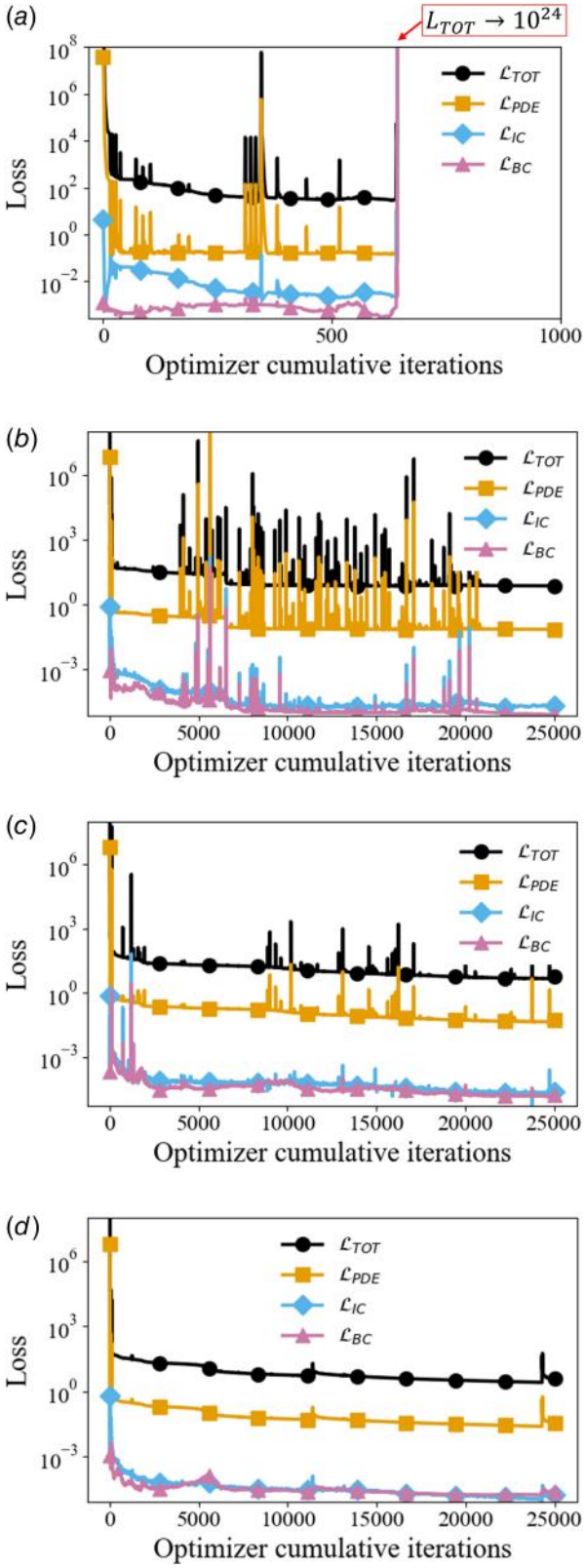

Fig. 9 Comparison of different activation functions performance: (*a*) ReLU(x), (*b*) Swish(x), (c) Tanh(x), and (*d*) Sin(x)

PDE residual loss rapidly diverged after approximately 700 iterations (see Fig. 9(*a*)). This behavior can be attributed to the lack of continuous second derivatives (being zero almost everywhere as shown in Fig. 8(*c*)), which makes ReLU unsuitable for such applications. On the contrary, the Swish activation function [39] provided a significant improvement in training stability compared to ReLU, preventing loss divergence. However, some instability persisted, indicated by the spikes in the graph (Fig. 9(*b*)). Yet this improvement came at the cost of longer training time, owing to the more complex structure of the activation function. The tanh activation function demonstrated stable and relatively fast convergence (see Fig. 9(*c*)), but the sin function (Fig. 9(*d*)) achieved even faster and more consistent training and was therefore selected for the final model.

Regarding the network's architecture, configurations with only four layers (shallow network) failed to effectively minimize the PDE loss, regardless of the number of neurons used. Specifically, during the initial tuning phase (25,000 iterations), the four-layer architectures were unable to reduce $\mathcal{L}_{\text{PDE}}$ below the 10^0 order of magnitude, often stagnating near 10^1. While these shallow networks achieved initial and boundary condition losses ($\mathcal{L}_{\text{IC}}$ and $\mathcal{L}_{\text{BC}}$) comparable to deeper models (approximately 0.7×10^{-4}), they lacked the expressive capacity to resolve the steep thermal gradients required by the governing equation. Consequently, a deeper architecture with eight hidden layers was selected as the most viable option, as it consistently achieved $\mathcal{L}_{\text{PDE}}$ values between 10^{-1} and 10^{-2} within the same iteration window, showing a clear downward trend toward convergence. On the other hand, increasing the number of neurons per layer to 64 significantly increased the training time without notable improvements in accuracy. The added complexity not only slowed convergence but also made the model more susceptible to unstable optimization, likely as a result of overparameterization. A configuration with 32 neurons per layer was found to provide a good balance, allowing for efficient training while maintaining stability and accuracy.

Finally, the comparison between the four different collocation point sampling strategies (Fig. 10) was performed. The grid strategy divides the domain in evenly spaced points, the random strategy uses a pseudorandom number generator to fill the domain, and the Sobol strategy uses a low-discrepancy generator to fill the space more evenly. The moving strategy, however, was defined considering a cylinder centered at the laser position at each time-step. The points are then distributed around this moving center using a Gaussian distribution with a standard deviation of 0.05, resulting in a high density of points under the heat source. This strategy produced a more stable and predictable convergence profile, as no sharp peaks were observed in the loss trend, although the overall convergence was slower compared to the other strategies. This can be attributed to the higher number of points located in the high-gradient zone, resulting in larger PDE losses but, at the same time, greater significance.

Ultimately, the moving sampling strategy was selected, but it was modified to incorporate 30% of sampling points distributed across the entire domain using the Sobol sequence. This adjustment ensures that information is not lost in regions far from the laser source while keeping the focus on the more complex region. For the initial and boundary condition points, a Sobol strategy sampler was selected instead. The overall model is summarized in Table 5.

The final PINN model was trained to reach a PDE loss lower than 10^{-4}, and boundary and initial losses lower than 10^{-6} with a final total loss of 10^{-2}. The training was performed using an RTX 4090 GPU with 24 GB of dedicated memory, and the total time was around 34 h.

5.2 Calibration. The results of the Pearson sensitivity analysis, presented in Fig. 11, indicate that the total cone height ($H_u + H_l$) exhibits the strongest correlation with the melt pool depth. Between the two heights, the lower cone (H_l) has the strongest effect (0.99), which is twice that of the upper cone (0.42). Additionally, the top radius R_t has a dominant influence on the melt pool width, with the intermediate radius having slightly more than half of its influence (0.48 compared to 0.88). In contrast, the bottom radius has almost no effect on the melt pool dimensions;

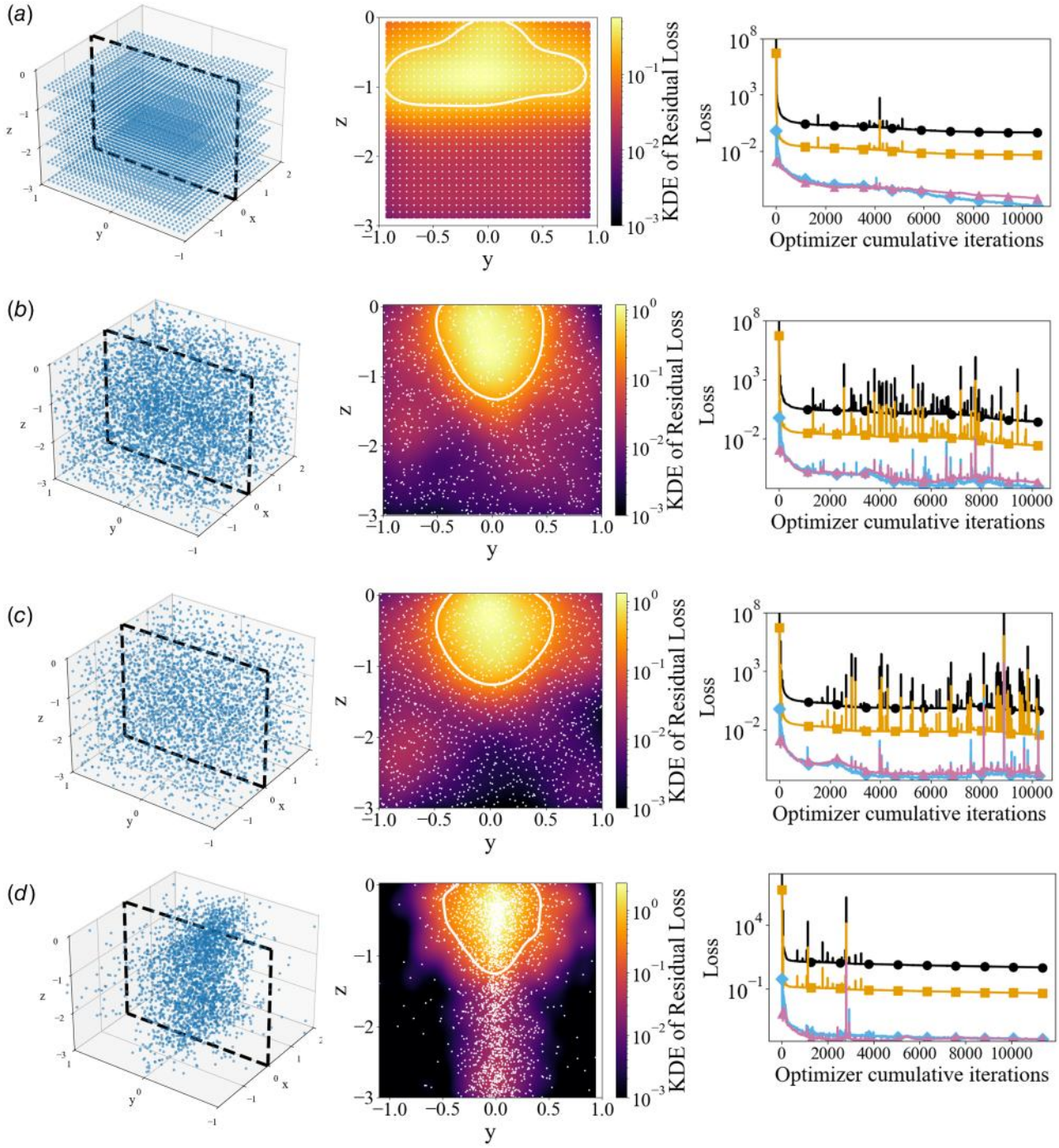

Fig. 10 Comparison of collocation point sampling strategies

however, it significantly affected the overall shape: a low value resulted in a more tapered weld bead cross section compared to higher values. The top radius also has a negative effect on the total weld depth, while the effect of the other due radii is negligible.

Based on these findings, the selection of the calibration variables for the inverse analysis was strictly driven by their statistical importance. Specifically, the parameters exhibiting the highest correlation coefficients (R_t and H_l) were selected as the variables to be optimized. Conversely, the three parameters with lower Pearson coefficients (R_i, R_b, and H_u) were held constant, thereby reducing the complexity of the optimization problem. In particular, a multicase inverse calibration was performed.

The optimal values for the top radius and lower cone height were found at different fixed values of the other parameters. The configuration that exhibited the best agreement with all five experimental results uses the following fixed parameters: $H_u = 0.5$, $R_i = 0.8 \cdot R_t$, and $R_b = 0.1$. A top cone height equal to 0.5 mm is necessary to ensure a full penetration condition (3 mm) since the maximum value for the lower cone height was set to 2.5 mm. The optimal heat source geometry parameters, $R_{t_{opt}}$ and $H_{l_{opt}}$, were determined. These optimized parameters led to a total loss lower than 10^{-5} and are listed in Table 6 for each case tested.

Based on these findings and the Pearson correlation analysis, the fitting of the geometric parameters was performed using two combined variable transformations.

Although linear relationships were initially considered, they lack physical consistency as they allow for negative predictions of geometric dimensions outside specific ranges. To address this

Table 5 Final model architecture

Parameter	Value
Hidden layers	8
Neurons per layer	32
Activation function	Sin
Sampling strategy	Custom moving
Collocation points	1,048,576

limitation and ensure strictly positive values, a power law approach was adopted. Since the total depth is directly influenced by the power and inversely by the scanning speed, the lower cone height was fitted against the ratio $P/\sqrt{v}$. Consequently, H_l approaches zero only if P is zero or as the scanning speed tends to infinity. These represent extreme conditions rather than feasible process windows, confirming the robustness of the model compared to a linear regression. The resulting power law equation is expressed as follows:

$$H_l(P, v) = m_1 \cdot \left(\frac{P}{\sqrt{v}}\right)^{n_1} \tag{22}$$

where the exponent n_1 allows the model to capture the nonlinear scaling of the depth. Conversely, the top radius exhibited an inverse correlation with both the power and the scanning speed. To maintain the constraint of positive values while capturing this inverse relationship, a power law against the product of power

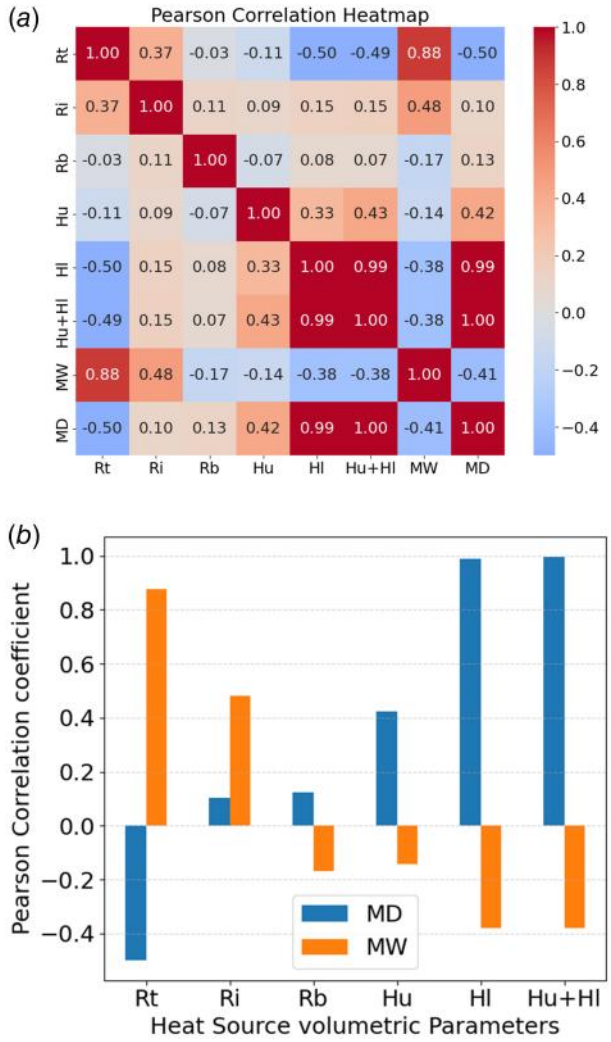

Fig. 11 Pearson correlation analysis between heat source geometric parameters and MW and MD: (*a*) Pearson correlation heatmap and (*b*) Pearson correlation coefficients

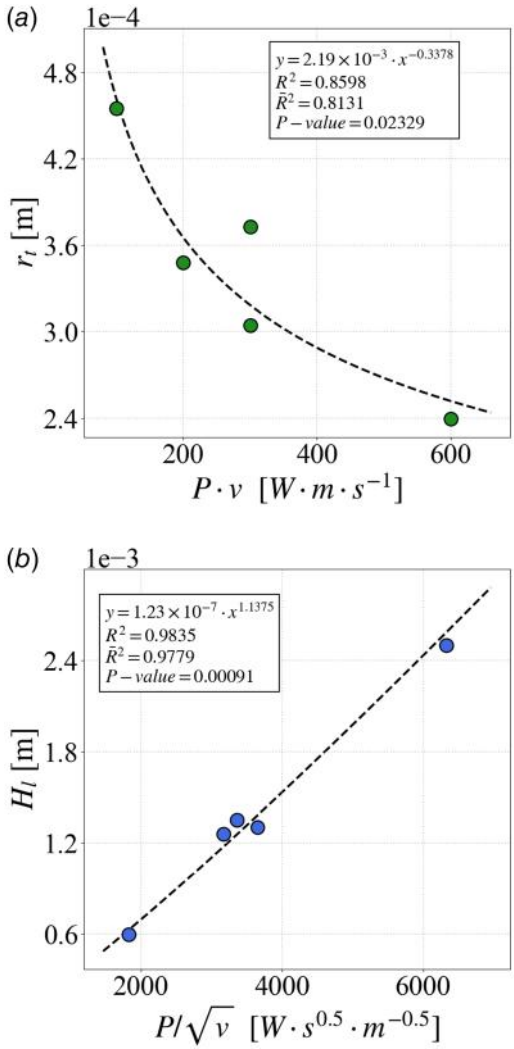

Fig. 12 Fitted parameters using the proposed assumptions: (*a*) top radius against $P \cdot v$ and (*b*) lower cone height against $P/\sqrt{v}$

Table 6 Calibrated geometric parameters of the five evaluated process conditions

ID	MD_{exp} (mm)	MW_{exp} (mm)	R_t (mm)	R_i (mm)	R_b (mm)	H_u (mm)	H_l (mm)
CAL_1	1.785	0.852	0.455	0.364	0.1	0.5	1.258
CAL_2	1.098	0.510	0.304	0.243	0.1	0.5	0.598
CAL_3	1.898	0.707	0.373	0.298	0.1	0.5	1.350
CAL_4	3.000	0.799	0.348	0.278	0.1	0.5	2.498
CAL_5	1.856	0.467	0.240	0.192	0.1	0.5	1.305

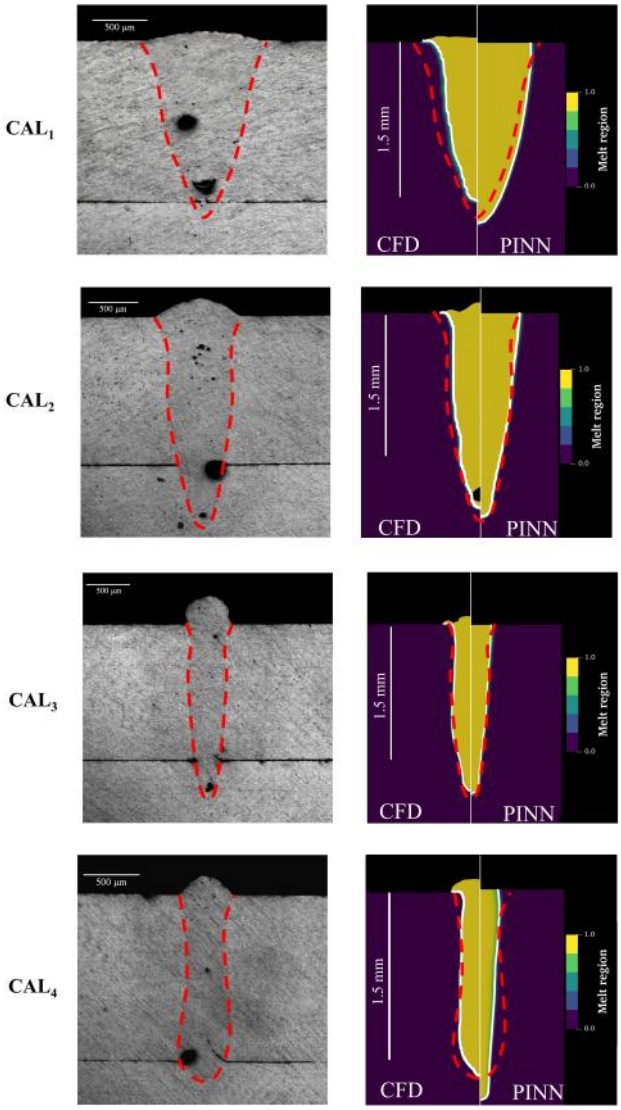

Fig. 13 Comparison between experimental cross section (left) and PINN prediction (right) at (*a*) case VAL_2 and (*b*) VAL_4

and scanning speed is proposed:

$$R_t(P, v) = m_2 \cdot (P \cdot v)^{n_2} \tag{23}$$

In this formulation, a negative exponent n_2 would mathematically describe the inverse correlation, ensuring the radius asymptotically approaches zero rather than becoming negative for extreme process values. Although this formulation presents a singularity at zero values (implying infinite width for zero power), it effectively prevents unphysical negative radius predictions within the operational process window. This confirms the mathematical suitability of the model for the investigated range of parameters. The final fitted curves are shown in Fig. 12 and the following correlations were established between the top radius R_t, the lower cone height H_l, and the laser power P and scan speed v:

$$R_t = 2.19 \cdot 10^{-3} \cdot (P \cdot v)^{-0.3378} \tag{24}$$

$$H_l = 1.23 \cdot 10^{-7} \cdot \left(\frac{P}{\sqrt{v}}\right)^{1.1375} \tag{25}$$

For both regressions, the coefficient of determination R^2, the adjusted coefficient $\bar{R}^2$, and the p-values are calculated. The fitting of the lower cone height (Fig. 12(*b*)), having an extremely low p-value of 9.1×10^{-4} and a $\bar{R}^2$ of 0.978, confirms that our assumption is correct. The top radius has a higher p-value of 0.02329 and a lower $\bar{R}^2$ of 0.8131. However, the p-value below 0.05 suggests that our assumption has a physical basis despite the lower R^2 value.

5.3 Validation. Regarding the validation cases located within the calibration range (VAL_2 and VAL_3), the model shows good agreement with the experiments, with relative error below 9 % for both MD and MW. Outside the calibration range, the PINN maintains a satisfactory predictive capability for the melt pool depth, with relative deviation less than 10%. However, a larger discrepancy is observed for the width, with VAL_4 having a relative error of approximately 14%.

Figure 13 illustrates the validation welds. The side-by-side comparison of the computational fluid dynamics (CFD) (left half) and PINN (right half) cross sections is colored according to the melt region variable. Conversely, the PINN slightly overestimates the penetration depth in most cases. The overall weld shape predicted by the PINN is consistent with the physical specimens, with the

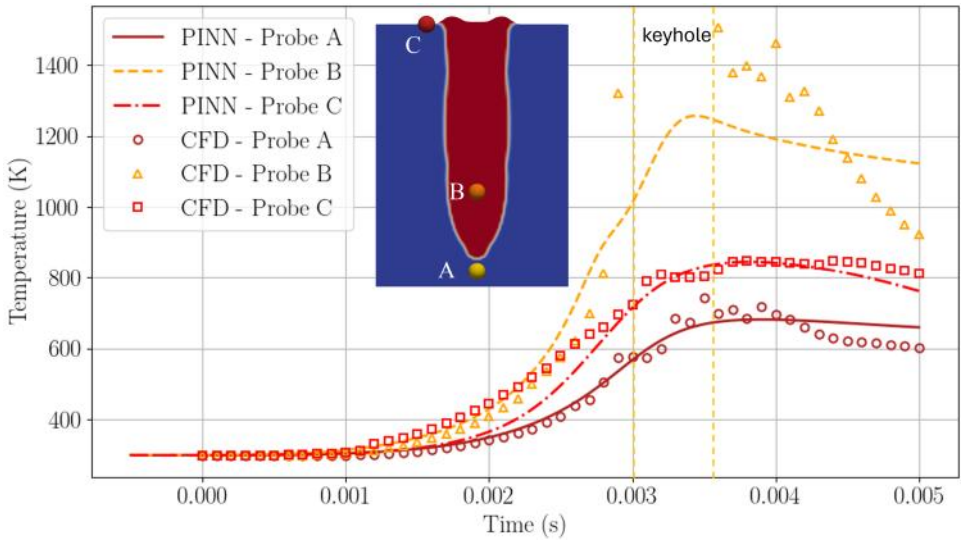

Fig. 14 Temperature evolution at three virtual probes extracted from the PINN model and numerical simulations

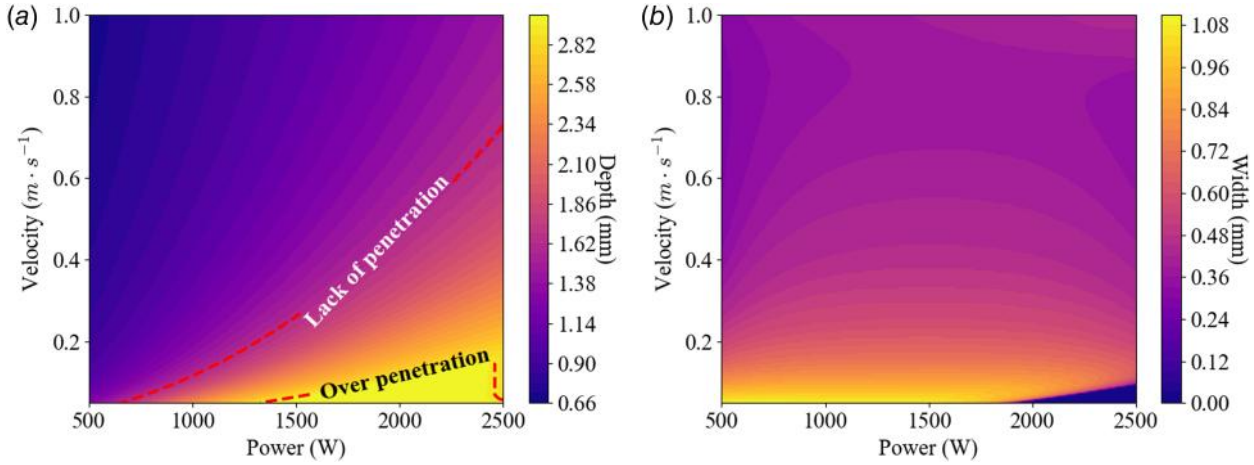

Fig. 15 Predicted melt pool (*a*) depth and (*b*) width for the complete process parameters

exception of the VAL_4 case. Finally, it is worth noting that since the PINN model solves only for thermal conduction, it cannot predict surface humping. In terms of computational cost, the duration of the high-fidelity CFD simulations was strictly dependent on the laser scanning speed ($VAL_1 > 28$ h; $VAL_4 \approx 12$ h). In contrast, the trained PINN model is capable of inferring the entire thermal field in milliseconds.

A further evaluation of the model capability is shown in Fig. 14. Three virtual probes were extracted from both the numerical simulations and the PINN and the temperature data were plotted. The comparison between the temperature predicted by the CFD model and those predicted by the calibrated PINN shows excellent agreement.

5.4 Case Study. The validated model was deployed to predict weld geometry across the entire process parameter space. A comprehensive grid of 250×250 combinations was computed, and the resulting maps for weld depth and width are presented in Fig. 15.

These predictions serve as a powerful tool for defining the optimal process window. In Fig. 15(*a*), two dashed lines delineate the boundaries to avoid: overpenetration (>2.9 mm) and lack of penetration (<1.7 mm). It is also worth noting a numerical artifact observed in the bottom-right corner of the process parameter domain (Fig. 15(*b*)). In this region, characterized by high speed and high power, the predicted width collapses to zero. However, since this area falls entirely within the overpenetration zone, this anomaly does not affect the identification of the optimal process window. Furthermore, as observed in the upper-left corner of Fig. 15(*a*), the model predicts a minimum penetration depth of 0.5 mm direct consequence of fixing H_u to 0.5 mm.

6 Conclusions

In this work, a PINN framework was developed and validated for the simulation of keyhole-mode laser welding of Al6082 aluminum alloy.

The following conclusions can be drawn from the study:

- *Hyperparameter Optimization:* A sine activation function combined with a dynamic "moving" sampling strategy proved to be the most effective configuration for ensuring training stability and convergence.
- *Calibration Strategy:* Calibrated on a small set of five experimental cross sections, the model established a robust correlation between process inputs (power and speed) and heat source dimensions via a power law regression.
- *Accuracy and Efficiency:* The calibrated PINN showed high predictive accuracy, with mean relative errors generally below 10%. Crucially, the model acts as a near-real-time surrogate, predicting thermal fields in milliseconds.
- *Model Reliability:* The study provides a dual validation of the approach: the agreement with experimental and CFD data confirms the physical soundness of the double-conical heat source, while establishing PINNs as a robust bridge between traditional analytical models and high-fidelity numerical techniques.
- *Process Optimization:* The framework demonstrated its practical utility in a case study, effectively mapping the entire process window to identify optimal laser power and speed configurations.

Future developments will focus on extending the model to include fluid dynamics effects, different joint configuration and material alloys.

Acknowledgment

The authors are grateful to Alessandro Sauna for his valuable support and assistance during this study.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Katayama, S., 2013, *Handbook of Laser Welding Technologies*, Woodhead Publishing, Cambridge, UK.
[2] Piandoro, S., Fortunato, A., Liverani, E., and Ascari, A., 2025, "Numerical Investigation of Pulsed Welding on Aluminum Alloy With a Dual-Region Laser Source: Exploring Non-Equal Pulse Parameters," Mater. Res. Proc., **54**, pp. 2130–2140.
[3] Prieto, C., Vaamonde, E., Diego-Vallejo, D., Jimenez, J., Urbach, B., Vidne, Y., and Shekel, E., 2020, "Dynamic Laser Beam Shaping For Laser Aluminium Welding in E-Mobility Applications," Procedia Cirp, **94**, pp. 596–600.
[4] Dal, M., and Fabbro, R., 2016, "[Invited] An Overview of the State of Art in Laser Welding Simulation," Opt. Laser Technol., **78**, pp. 2–14.
[5] Mazumder, J., and Steen, W. M., 1980, "Heat Transfer Model for CW Laser Material Processing," J. Appl. Phys., **51**(2), pp. 941–947.

[6] Norman, A. F., Ducharme, R., Mackwood, A., Kapadia, P., and Prangnell, P. B., 1998, "Application of Thermal Modelling to Laser Beam Welding of Aluminium Alloys," Sci. Technol. Welding Join., **3**(5), pp. 260–266.
[7] Kaplan, A., 1994, "A Model of Deep Penetration Laser Welding Based on Calculation of the Keyhole Profile," J. Phys. D: Appl. Phys., **27**(9), pp. 1805–1814.
[8] Semak, V. V., Damkroger, B., and Kempka, S., 1999, "Temporal Evolution of the Temperature Field in the Beam Interaction Zone During Laser Material Processing," J. Phys. D: Appl. Phys., **32**(15), pp. 1819–1825.
[9] Florian, T., Schricker, K., Zenz, C., Otto, A., Schmidt, L., Diegel, C., Friedmann, H., et al., 2025, "Combining In Situ Synchrotron X-Ray Imaging and Multiphysics Simulation to Reveal Pore Formation Dynamics in Laser Welding of Copper," Int. J. Mach. Tools. Manuf., **204**, p. 104224.
[10] Chianese, G., Hayat, Q., Jabar, S., Franciosa, P., Ceglarek, D., and Patalano, S., 2023, "A Multi-Physics CFD Study to Investigate the Impact of Laser Beam Shaping on Metal Mixing and Molten Pool Dynamics During Laser Welding of Copper to Steel for Battery Terminal-to-Casing Connections," J. Mater. Process. Technol., **322**, p. 118202.
[11] Huang, W., Cai, W., Rinker, T. J., Bracey, J., and Tan, W., 2023, "Effects of Laser Oscillation on Metal Mixing, Microstructure, and Mechanical Property of Aluminum–Copper Welds," Int. J. Mach. Tools. Manuf., **188**, p. 104020.
[12] Lee, K., Rinker, T. J., Pour, M. M., Cai, W., Huang, W., Tan, W., Bracey, J., and Li, J., 2023, "A Study on Cracks and Imcs in Laser Welding of Al and Cu," Manuf. Lett., **35**, pp. 221–231.
[13] Angeloni, C., Piandoro, S., Liverani, E., and Fortunato, A., 2015, "Numerical and Experimental Study of High-Speed Laser Cutting of Copper Current Collectors: Process Optimization for Quality Assessment," Adv. Mater. Technol., **10**(14), p. 2401905.
[14] Li, Z., Rostam, K., Panjehpour, A., Akbari, M., Karimipour, A., and Rostami, S., 2020, "Experimental and Numerical Study of Temperature Field and Molten Pool Dimensions in Dissimilar Thickness Laser Welding of Ti_6Al_4V Alloy," J. Manuf. Process., **49**, pp. 438–446.
[15] Goldak, J., Chakravarti, A., and Bibby, M., 1984, "A New Finite Element Model for Welding Heat Sources," Metallurgical Trans. B, **15**, pp. 299–305.
[16] Wu, C., Wang, H., and Zhang, Y., 2006, "A New Heat Source Model for Keyhole Plasma ARC Welding in FEM Analysis of the Temperature Profile," Welding J., **85**(12), pp. 284–s–291–s.
[17] Kik, T., 2020, "Heat Source Models in Numerical Simulations of Laser Welding," Materials, **13**(11), p. 2653.
[18] Murua, O., Arrizubieta, J., Lamikiz, A., and Schneider, H., 2024, "Numerical Simulation of a Laser Beam Welding Process: From a Thermomechanical Model to the Experimental Inspection and Validation," Ther. Sci. Eng. Prog., **55**, p. 102901.
[19] Jordan, M. I., and Mitchell, T. M., 2015, "Machine Learning: Trends, Perspectives, and Prospects," Science, **349**(6245), pp. 255–260.
[20] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, Vol. 25, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, eds., Curran Associates, Inc., Red Hook, NY, pp. 1097–1105.
[21] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I., 2018, "Improving Language Understanding by Generative Pre-Training," OpenAI Technical Report.
[22] Raissi, M., Perdikaris, P., and Karniadakis, G., 2018, "Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations," J. Comput. Phys., **378**, pp. 686–707.
[23] Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M., 2018, "Automatic Differentiation in Machine Learning: A Survey," J. Mach. Learn. Res., **18**(153), pp. 1–43.
[24] Peng, S., Yang, S., Gao, B., Liu, W., Wang, F., and Tang, Z., 2025, "Prediction of 3D Temperature Field Through Single 2D Temperature Data Based on Transfer Learning-Based PINN Model in Laser-Based Directed Energy Deposition," J. Manuf. Process., **138**, pp. 140–156.
[25] Farrag, A., Yang, Y., Cao, N., Won, D., and Jin, Y., 2024, "Physics-Informed Machine Learning for Metal Additive Manufacturing," Progr. Addit. Manuf., **10**(1), pp. 171–185.
[26] Tang, J., Scheel, P., Mohebbi, M. S., Leinenbach, C., De Lorenzis, L., and Hosseini, E., 2024, "On the Calibration of Thermo-Microstructural Simulation Models for Laser Powder Bed Fusion Process: Integrating Physics-Informed Neural Networks With Cellular Automata," Addit. Manuf., **96**, p. 104574.
[27] Sharma, R., Guo, Y. B., Raissi, M., and Guo, W. G., 2024, "Physics-Informed Machine Learning of Argon Gas-Driven Melt Pool Dynamics," ASME J. Manuf. Sci. Eng., **146**(8), p. 081008.
[28] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., and Killeen, T., 2019, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," *Advances in Neural Information Processing Systems*, Vol. 32, Curran Associates, Inc., Red Hook, NY, pp. 8024–8035.
[29] Krishnapriyan, A. S., Gholami, A., Zhe, S., Kirby, R. M., and Mahoney, M. W., 2021, "Characterizing Possible Failure Modes in Physics-Informed Neural Networks," 35th Conference on Neural Information Processing Systems, Sydney, Australia.
[30] Lampa, C., Kaplan, A. F., Powell, J., and Magnusson, C., 1997, "An Analytical Thermodynamic Model of Laser Welding," J. Phys. D: Appl. Phys., **30**(9), pp. 1293–1299.
[31] Proell, S. D., Wall, W. A., and Meier, C., 2020, "On Phase Change and Latent Heat Models In Metal Additive Manufacturing Process Simulation," Adv. Model. Simul. Eng. Sci., **7**(1), p. 24.
[32] Farrokhi, F., Endelt, B., and Kristiansen, M., 2019, "A Numerical Model for Full and Partial Penetration Hybrid Laser Welding of Thick-Section Steels," Opt. Laser Technol., **111**, pp. 671–686.
[33] Rong, Y., Mi, G., Xu, J., Huang, Y., and Wang, C., 2018, "Laser Penetration Welding of Ship Steel EH36: A New Heat Source and Application to Predict Residual Stress Considering Martensite Phase Transformation," Marine Struct., **61**, pp. 256–267.
[34] Lu, L., Meng, X., Mao, Z., and Karniadakis, G. E., 2021, "Deepxde: A Deep Learning Library for Solving Differential Equations," SIAM Rev., **63**(1), pp. 208–228.
[35] Monaco, S., and Apiletti, D., 2023, "Training Physics-Informed Neural Networks: One Learning to Rule Them All?," Results Eng., **18**, p. 101023.
[36] Münzer, M., and Bard, C., 2022, "A Curriculum-Training-Based Strategy for Distributing Collocation Points during Physics-Informed Neural Network Training," Machine Learning and the Physical Sciences Workshop, NeurIPS 2022, New Orleans, LA.
[37] Gu, J., Yang, S., Duan, C., Xiong, Q., and Wang, Y., 2019, "Microstructure and Mechanical Properties of Laser Welded Al–Mg–Si Alloy Joints," Mater. Trans., **60**(2), pp. 230–236.
[38] Garavaglia, M., Demir, A. G., Zarini, S., Victor, B. M., and Previtali, B., 2020, "Fiber Laser Welding of AA 5754 in the Double Lap-Joint Configuration: Process Development, Mechanical Characterization, and Monitoring," Int. J. Adv. Manuf. Technol., **111**(5–6), pp. 1643–1657.
[39] Ramachandran, P., Zoph, B., and Le, Q. V., 2017, "Searching for Activation Functions," e-print arXiv.