



## Full-length article

# Experiments on resistance spot welding of three layers of unequal thickness steel based on deep learning and multi-objective optimization

Haofeng Deng<sup>a</sup>, Pengyu Gao<sup>a</sup>, Honggang Xiong<sup>a</sup>, Xiangdong Gao<sup>a,b,\*</sup>

<sup>a</sup> Guangdong Provincial Welding Engineering Technology Research Center, Guangdong University of Technology, Guangzhou, 510006, China

<sup>b</sup> Guangzhou Zhengtian Technology Company, Guangzhou, 510006, China

## ARTICLE INFO

## Keywords:

Resistance spot welding  
Numerical simulation  
Deep learning  
Multi-objective optimization  
Unequal thickness welding of three-layer plates

## ABSTRACT

Resistance spot welding (RSW) of three-layer steel sheets with unequal thicknesses presents significant challenges in accurately simulating weld nugget formation and process signal behavior. This paper proposes a hybrid approach that combines deep learning and multi-objective optimization to improve simulation accuracy. A 1D convolutional neural network (1DCNN), bidirectional long short-term memory (BiLSTM), and attention mechanism are integrated to predict dynamic resistance curves from process parameters. These predicted curves are then used as benchmarks in an ANSGA-II and Bayesian optimization framework to calibrate thermal-electrical contact parameters in a finite element model. Experimental results demonstrate that the optimized simulations closely match measured data, achieving a mean absolute error (MAE) of 0.132, a root mean square error (RMSE) of 0.156, and an  $R^2$  value of 0.91. The calibrated model reduces resistance prediction error by over 30% and improves nugget diameter and weld depth prediction accuracy across multiple thickness configurations. This integrated framework offers a practical and data-efficient solution for enhancing RSW simulations in complex multi-layer welding scenarios.

## 1. Introduction

Resistance spot welding (RSW) is extensively employed in the fabrication of thin steel components, particularly in the automotive industry for assembling body-in-white structures, owing to its operational simplicity, high speed, and low cost [1–3]. In modern vehicle design, many structural joints comprise three-layer steel sheets with varying thicknesses and material strengths [4]. Compared to conventional two-sheet welding, three-layer configurations introduce additional complexities, such as asymmetric heat flow, varying contact resistance, and inconsistent current distribution. These factors lead to non-uniform nugget growth and greater uncertainty in weld quality [5]. The middle sheet often exhibits reduced heat conduction, exacerbating the asymmetry of nugget formation at the upper and lower interfaces. Since the weld nugget forms internally between metal sheets, it cannot be observed directly during welding. The size and morphology of the nugget are critical indicators of weld strength and reliability [6], but conventional quality inspection techniques, such as chisel tests and ultrasonic scanning, are offline, time-consuming, and insufficient for high-throughput production environments. Consequently, real-time and automated weld quality monitoring has become essential for modern manufacturing, in

line with the broader shift toward intelligent and digitalized production systems [7–9].

Since the weld nugget forms internally and cannot be directly observed during welding, in-process electrical signals are widely adopted for real-time weld quality prediction. Among them, dynamic resistance has proven to be one of the most reliable and interpretable indicators, as it captures the metallurgical evolution at the faying surfaces under coupled heat and pressure conditions [10,11]. Numerous studies have confirmed that dynamic resistance profiles correlate strongly with nugget initiation, growth, and final geometry. For example, Zhao et al. [12] extracted 20 characteristic features from resistance curves and used stepwise regression to map them to welding parameters with reasonable accuracy. Xing et al. [13] proposed a Random Forest-based classification system using dynamic resistance signals acquired in a production setting. Other researchers have further explored the potential of data-driven approaches, employing algorithms such as Support Vector Machines (SVM), Gaussian Process Regression (GPR), and ensemble methods to predict weld quality indices from signal features [14,15]. These studies highlight the effectiveness of statistical learning models, but their performance often declines when faced with

\* Corresponding author at: Guangdong Provincial Welding Engineering Technology Research Center, Guangdong University of Technology, Guangzhou, 510006, China.

E-mail addresses: [1112301044@mail2.gdut.edu.cn](mailto:1112301044@mail2.gdut.edu.cn) (H. Deng), [perrygao@gmail.com](mailto:perrygao@gmail.com) (P. Gao), [2784248288@qq.com](mailto:2784248288@qq.com) (H. Xiong), [gaoxd@gdut.edu.cn](mailto:gaoxd@gdut.edu.cn) (X. Gao).

<https://doi.org/10.1016/j.cirpj.2025.07.005>

Received 15 May 2025; Received in revised form 17 July 2025; Accepted 21 July 2025

Available online 5 August 2025

1755-5817/© 2025 CIRP.

structural complexity, such as multi-layer stack-ups or varying thermal paths, due to limited generalization across different configurations.

Numerous data-driven models have been proposed to map process signal features to weld geometry indicators, aiming to enhance prediction capabilities in resistance spot welding [16–18]. For instance, Bogaerts et al. [19] combined unsupervised deep learning with Gaussian Process Regression (GPR) to estimate nugget diameter. He et al. [20] employed a Chaotic Game Optimization (CGO) strategy integrated with a multi-output Least Squares Support Vector Regression (MLSSVR) model to predict weld quality, further refining parameters through Particle Swarm Optimization (PSO). With the growing availability of process data and computational resources, deep learning models such as LSTM, CNN, and their hybrids have shown great promise in capturing complex temporal dependencies in RSW signals [21]. Vo et al. [22] utilized LSTM networks to predict weld outcomes using multiple welding parameters, while Chang et al. [23] introduced a CNN-LSTM-Attention model within an edge-cloud collaborative framework, leveraging dynamic current and resistance inputs and addressing class imbalance with SMOTE and weighted loss functions. Despite these advances, challenges persist when applying such models to real-world manufacturing settings. In particular, three-layer RSW structures with varying thicknesses and materials introduce non-linear heat flow and contact conditions that differ significantly from training scenarios. Existing models are highly sensitive to the training dataset's distribution; any domain shift, caused by new configurations or unseen parameter combinations, can substantially degrade performance. Moreover, collecting sufficient labeled data with accurate weld quality annotations (e.g., nugget size or depth) remains labor-intensive and costly, restricting large-scale deployment of supervised deep learning in practical RSW quality monitoring systems.

To address the limitations of purely data-driven approaches, physics-based numerical modeling, especially finite element methods (FEM), has been widely adopted to gain insights into variables that are not directly measurable during the RSW process. For example, Hamid Eisazadeh et al. [24] developed a coupled thermo-electro-mechanical FEM to study how welding parameters affect nugget formation. Wang et al. [25] applied a multi-physics simulation model to investigate interfacial intermetallic compound (IMC) evolution in aluminum–steel RSW. Mikut's group [26] used validated FEM data to train machine learning models and analyze how feature selection and dataset size influence weld quality prediction. Similarly, Xia et al. [27] constructed an electrical-thermal-force coupled model to simultaneously simulate nugget growth and process signals, and later proposed a multi-field co-simulation method to establish quantitative mappings between dynamic resistance and nugget morphology [28].

While FEM-based models offer strong physical interpretability, their accuracy depends heavily on contact resistance and material property parameters, which are often difficult to determine precisely. To improve predictive performance, multi-objective optimization algorithms have been employed to calibrate simulation inputs. Pashazadeh et al. [29] integrated artificial neural networks (ANN) with a multi-objective genetic algorithm (MOGA) for weld parameter optimization. Zhuo et al. [14] developed a GPR-BP-based surrogate model and optimized it using NSGA-II, improving both accuracy and robustness. Bayesian optimization has also been used to enhance model generalization, as demonstrated by Stavropoulos et al. [15], who applied it to tune SVM hyperparameters in infrared-based weld quality monitoring.

However, existing approaches typically treat data-driven learning and physical modeling as separate processes. Few studies have integrated deep learning prediction with simulation-based optimization to form a closed-loop enhancement framework, especially under the complex conditions of three-layer RSW with unequal thicknesses. This gap motivates our proposed method.

In addition to research on resistance spot welding of steel and multilayer welded joints, significant progress has been made in the study of fusion welding of dissimilar aluminum alloys, especially in

terms of process complexity, thermo-mechanical properties, and parameter optimization. A number of studies have used tungsten inert gas (TIG) welding combined with post-weld stir friction processing (FSP) to improve the integrity and mechanical properties of aluminum alloy welded joints [30]. These works investigated the effects of heat input, tool pin geometry, and pulse parameters on microstructural evolution, tensile strength, and wear behavior [31]. Md Saquib Bin Reyaz et al. welded dissimilar aluminum alloy joints using an automatic tungsten argon arc welding technique and investigated the effect of heat input on the microstructure, grain boundary characteristics, and mechanical properties of the joints [32]. Immediately following the next year to explore the effect of tool spindle speed on the properties of TIG-welded dissimilar aluminum joints and the relationship between microstructure and tensile strength, impact toughness, and abrasion resistance, it was found that the dendritic coarse crystalline structure in the fusion zone of the TIG-welded joints transformed into an equiaxed fine crystalline structure in the stirred zone [33]. Furthermore, modeling tools such as Response Surface Methodology (RSM), Ideal Function Analysis (IFA), and Pin Shape Optimization (PSO) were employed in the experimental study to systematically evaluate and optimize the welding results [34]. Md Saquib Bin Reyaz [35], to explore the effects of Friction Stir Processing (FSP) parameters on the tungsten inert gas (TIG) welding of dissimilar aluminum alloy joints, the cumulative effect of face-corner-center composite design (FCCCD) by response surface methodology (RSM) was used. And the response surface methodology was used to develop a regression model to predict the mechanical properties of the welded joints, such as tensile strength, elongation, microhardness, and residual stress [36]. Although these studies focus on fusion welding of aluminum alloys, they provide valuable insights into multi-parameter coupling effects, thermal history, and optimization strategies that are equally relevant to resistance spot welding of more complex configurations, such as three-layer steel joints. Our study builds on these cross-domain insights to integrate deep learning and simulation-based optimization for improved process prediction and model fidelity.

Given the high flexibility and physical interpretability of FEM-based spot welding simulations, it is crucial to enhance their accuracy in predicting key weld features such as nugget geometry and dynamic resistance signals, particularly for complex three-layer configurations. Building on the limitations identified in both purely data-driven and purely simulation-based methods, we propose a hybrid framework that tightly integrates deep learning and multi-objective optimization. Specifically, we employ a 1D convolutional neural network combined with BiLSTM and an attention mechanism (1DCNN-BiLSTM-Attention) to accurately predict dynamic resistance curves from welding parameters and configurations. These predicted curves are then used as optimization benchmarks in a co-simulation process, where an Adaptive NSGA-II (ANSKA-II) algorithm and Bayesian optimization are jointly used to refine critical FEM contact parameters, such as thermal and electrical conductivity coefficients. This framework not only improves simulation fidelity but also reduces reliance on large-scale experimental calibration data. The main contributions of this paper are as follows:

Fusion of deep learning and optimization algorithms for data-driven dynamic resistance prediction. 1DCNN is used to extract local time series features, BiLSTM captures long-term dependencies, and the attention mechanism adaptively adjusts feature weights to improve the prediction accuracy of the dynamic resistance curve. The data-driven approach replaces the traditional empirical model, improves the modeling ability of complex welding processes, and reduces parameter dependence.

Combine deep learning and intelligent optimization to improve the simulation accuracy of resistance spot welding. The dynamic resistance curve predicted by 1DCNN-BiLSTM-Attention is integrated into the optimization process of numerical simulation, and the predicted curve is used as the target benchmark to guide the optimization of thermal conductivity and electrical conductivity in numerical simulation. Compared with the traditional way of relying on experimental data for

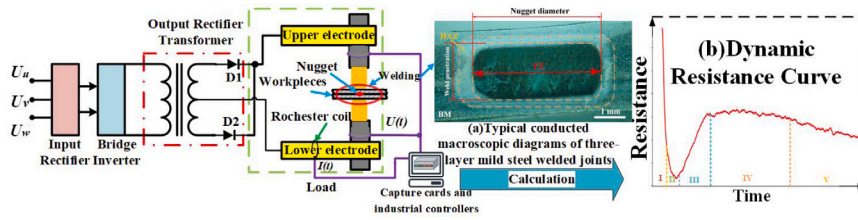


Fig. 1. Schematic of the resistance spot welding (RSW) system and signal acquisition process. (a) A macroscopic weld cross-section image shows the nugget diameter formed at the interface. (b) shows the dynamic resistance curve over time, divided into five typical phases: I: initial contact; II: oxide breakdown; III: nugget initiation; IV: nugget growth; V: cooling and solidification.

direct optimization, this method makes full use of the powerful fitting ability of deep learning to make the optimization target more accurate, reduce the demand for experimental data, and improve the credibility of simulation.

Use ANSGA-II and Bayesian optimization to improve the optimization efficiency of simulation parameters. Multi-objective optimization is carried out by ANSGA-II to simultaneously optimize the thermal and electrical conductivity of the interface in order to minimize the dynamic resistance error and the size error of the melt core. Bayesian optimization is used to accelerate convergence, reduce the number of simulation calculations, and improve the optimization efficiency.

Validate the effectiveness of the method in a welding scenario of a three-layer mild steel plate. At present, there are fewer simulation studies on resistance spot welding of multilayer metal plates, and the existing models are mostly for single- or double-layer welding, which is difficult to accurately simulate the welding behavior of complex structures. In this paper, the method is validated in four different thickness combinations (upper-middle-lower thickness: 1-1-1 mm, 0.7-1-0.7 mm, 0.7-0.7-1 mm, 0.7-1-1.2 mm) of mild steel three-layer welding experiments, combined with the analysis of the error of the dynamic resistance, the diameter of the nucleus of the melt, and the depth of melting, which shows that the method can effectively improve the accuracy of the simulation and has good engineering applicability and can be extended to more complex welding process optimization research.

## 2. Modeling and development

### 2.1. Experimental data acquisition

The experimental system comprises a welding robot, a medium-frequency DC spot welding gun, a current sensor, and a voltage sensor, as illustrated in Fig. 1. The resistance spot welding (RSW) process necessitates real-time acquisition of welding current and electrode voltage signals. The system includes a three-phase AC power supply, an input rectifier, a bridge inverter, a welding transformer, and associated components. The welding electrodes are made of CuCrZr alloy with a 6 mm end diameter. The welding current is measured using a current sensor installed in the electrostatic pole arm, while the electrode voltage is recorded via shielded RF cables attached to the upper and lower electrodes. All signals are collected by a high-speed data acquisition card integrated into a computer.

During welding, the robotic welding gun applies electrode pressure, ensuring firm contact between the metal plates. A high-intensity direct current (DC) is discharged through the transformer rectifier, flowing across the contact surfaces of the metal plates. This generates Joule heating, rapidly raising the interface temperature to the metal's melting point, thereby forming a localized molten pool. Subsequently, cooling water within the electrodes accelerates solidification, forming a robust weld joint. Upon pressure release, the three-layer metal plates remain securely bonded, completing the spot welding process efficiently. Throughout the process, the voltage and current signals are continuously acquired and used to compute the dynamic resistance according

to Eq. (1). Where  $R$  represents the dynamic resistance signal,  $U$  is the secondary voltage,  $I$  is the welding current, and  $R_b$  denotes the base resistance of the electrode between the two voltage probes.

$$R = U / I - R_b \quad (1)$$

This equation is widely adopted in resistance spot welding applications due to its simplicity and real-time feasibility. In our setup, the current waveform is delivered in DC mode with minimal switching noise, and the measurement system operates at 100 kHz sampling frequency with noise-filtered signals. Under such conditions, inductive and capacitive components in the circuit are negligible compared to the dominant resistive behavior. Additionally, any transient effects at the start of current flow are effectively suppressed by a low-pass Butterworth filter [37] applied during preprocessing. Therefore, the use of Eq. (1) provides a valid and accurate approximation of the time-varying electrical resistance during the welding process.

Its dynamic resistance curve is shown in Fig. 1(B), where each stage of the dynamic resistance marks the growth and formation process of the molten core [37]. Fig. 1(a) shows the metallographic diagram for measuring the diameter of the molten core, which was processed by EDM wire cutting machine, using wire cutting method to cut along the center of the welded joint perpendicular to the center of the welded joint, grinding and polishing after inlaying, and finally corrosion by using 4 percent by volume nitric acid-alcohol solution, and the microstructure of the welded joints was observed using metallographic microscope to observe the different weld zones, including the weld core fusion zone (FZ), the heat-affected zone (HAZ) and base material (BM), and the transverse and longitudinal measurements of the fusion nucleus region were performed to obtain the nugget diameter and weld penetration, respectively. In this paper, welding experiments were carried out on unequal thickness mild steel triple plates (top-middle-lower plate thicknesses of: 1-1-1 mm, 0.7-1-0.7 mm, 0.7-0.7-1 mm, 0.7-1-1.2 mm) by using different welding parameters, respectively.

To ensure the reliability of signal acquisition and mitigate the influence of environmental variations, several control measures were implemented during experiments. Dynamic resistance signals were collected using a high-resolution acquisition system at 100 kHz sampling rate, with built-in signal shielding. The raw signals were post-processed using a zero-phase 2th-order low-pass Butterworth filter [37] to remove transient spikes and high-frequency noise while maintaining phase accuracy.

The welding experiments were performed in a temperature-controlled laboratory environment ( $22 \pm 1.5^\circ\text{C}$ ), and electrodes were allowed to cool naturally between welds to prevent cumulative thermal drift. In each session, the first few samples were discarded to avoid thermal instability from machine warm-up. Additionally, electrodes were mechanically cleaned and replaced every 20 welds to minimize contact variability due to tip wear. All welded sheets were sourced from a single material batch, and surface preparation was standardized by ethanol wiping to reduce surface oxide variability. These precautions ensured that the dynamic resistance signals and weld quality metrics reflected intrinsic process behavior rather than uncontrolled external noise.

**Table 1**  
Complete full-factorial parameter matrix used for process space sampling.

Welding current (kA)	Welding time (ms)	Electrode pressure (kN)
6	160	3
6	160	4
6	160	5
6	240	3
6	240	4
6	240	5
6	320	3
6	320	4
6	320	5
9	160	3
9	160	4
9	160	5
9	240	3
9	240	4
9	240	5
9	320	3
9	320	4
9	320	5
12	160	3
12	160	4
12	160	5
12	240	3
12	240	4
12	240	5
12	300	3
12	300	4
12	300	5

To investigate the influence of welding parameters on nugget formation and dynamic resistance signals across different sheet thickness configurations, a full-factorial experimental design [38] was adopted for data acquisition. The chosen current (6–12 kA), welding time (160–300 ms), and electrode pressure (3–5 kN) reflect the typical operating window used in automotive resistance spot welding for mild and high-strength steel. Three key welding parameters were considered, each at three levels: Welding current (6 kA, 9 kA, 12 kA), Welding time (160 ms, 240 ms, 300 ms) and Electrode pressure (3 kN, 4 kN, 5 kN). For each plate thickness configuration, all 27 combinations of the above parameters were tested, and multiple repetitions were performed to ensure data robustness. In total, 400 high-quality samples were collected, covering the four configurations (1-1-1, 0.7-1-0.7, 0.7-0.7-1, 0.7-1-1.2). This approach guarantees balanced coverage of the process space while capturing variations due to geometry-induced asymmetry. The resulting dataset was used both for FEM validation and for training the deep learning model.

This 27-point matrix represents the full factorial combinations of welding current, time, and electrode pressure used as the core process space for experimental design and model calibration. The full experimental matrix is shown in Table 1.

## 2.2. Finite element model (FEM)

To integrate deep learning models and optimization algorithms into the RSW process, a thermo-electrical-mechanical coupling module was developed in Abaqus using its secondary development functionality. Exploiting the symmetry of the RSW model along the xOz and yOz planes, only one-quarter of the model was used for FEM to reduce computational cost and time. The three-layer unequal-thickness mild steel RSW model, shown in Fig. 2, consists of three workpieces and two electrodes. The mesh is composed of hexahedral eight-node elements (Q3D8R). Fig. 2(B)(C) depict the mechanical and thermoelectric boundary conditions of the model. The inputs to the model are the welding current (I) and the electrodynamic force (F), which are uniformly distributed on the top of the upper electrode cap. Accordingly, the bottom of the lower electrode cap is fixed and its potential and axial displacement are forced to zero. In addition, the symbols  $q_{water}$  and  $q_{air}$  denote the convective heat flux density.

During welding, the upper electrode head applies compressive force along the z-axis, ensuring firm contact between the workpieces. The welding current is supplied through the upper electrode, passing through the metal layers to the lower electrode. The thermal and electrical conductivity properties of the electrodes and mild steel are referenced from [39,40], while additional material properties can be found in [21,41]. Assuming constant temperature and denoting z and r as the axial and radial coordinates, respectively, the governing equation for transient heat transfer in a cylindrical coordinate system [42] is given by Eq. (2). Where  $q_v$  is the heat production,  $t$  is the time,  $T$  is the temperature,  $k$  is the thermal conductivity,  $\rho$  is the density and  $C_p$  is the mass heat capacity.

$$\frac{\partial}{\partial r} \left( k \frac{\partial T}{\partial r} \right) + \frac{k}{r} \frac{\partial T}{\partial r} + \frac{\partial}{\partial z} \left( k \frac{\partial T}{\partial z} \right) + q_v = C_p \rho k \frac{\partial T}{\partial t} \quad (2)$$

Its steady-state electrical conduction equation [42] is Eq. (3). Where  $\rho_e$  is the resistivity of the material and  $U$  is the electric potential.

$$\frac{\partial}{\partial r} \left( \frac{1}{\rho_e} \frac{\partial U}{\partial r} \right) + \frac{1}{\rho_e} \frac{\partial U}{\partial r} + \frac{\partial}{\partial z} \left( k \frac{\partial U}{\partial z} \right) = 0 \quad (3)$$

The analyzed structural stresses used are Eq. (4). Where  $\sigma$  is the stress;  $b$  is the volume force;  $i, j$  are the coordinate vectors.

$$\nabla \sigma(i, j) + b(i, j) = 0 \quad (4)$$

In the validated finite element model, the RSW system can be approximated as an equivalent circuit connected by a series of resistive components, including the thin-layer resistor ( $R_S$ ), the electrode resistor ( $R_E$ ), and the contact resistors ( $R_{E/S}$ ,  $R_{S/S}$ ), as shown in Fig. 2(a). All components are time-varying and can be obtained from the potentials  $\phi_i$  ( $i = 1 \sim 10$ ) at some key nodes [43]. Under symmetrical conditions, the overall dynamic resistance signal  $R$  can be expressed by an equivalent series circuit model as shown in Eq. (5).

$$R = 3R_S + 2R_E + 2R_{E/S} + 2R_{S/S} \quad (5)$$

During the spot welding process, the initial contact conditions between the workpieces and between the workpiece and the electrode are established through pre-pressurization. Given that material properties vary with temperature, accurately simulating the welding process requires incorporating the temperature-dependent variations in the thermo-electric properties of the material when current and pressure



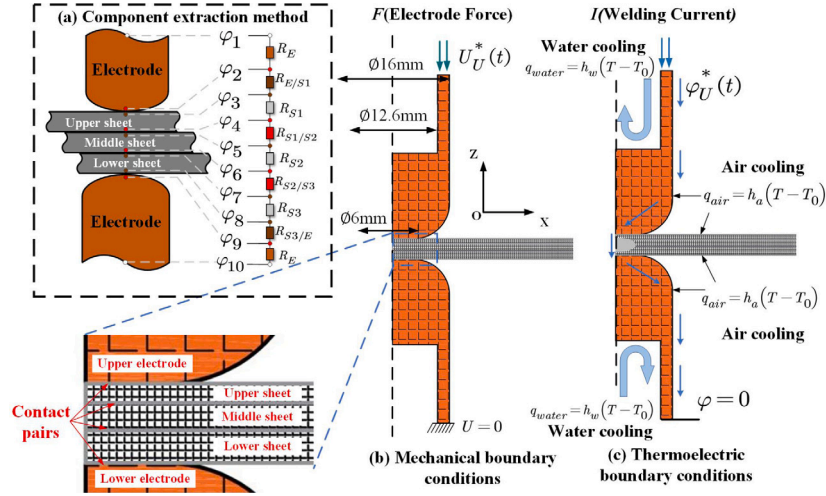


Fig. 2. Finite element modeling framework for three-layer RSW joints. (a) The electrical resistance network consists of electrode-sheet and inter-sheet contact resistances ( $R_{E/S1}$ ,  $R_{S1}$ ,  $R_{S1/S2}$ ,  $R_{S2}$ ,  $R_{S2/S3}$ ,  $R_{S3}$ ,  $R_{S3/E}$ ) and bulk material resistances. (b) Mechanical boundary conditions include electrode force application and fixed displacement constraints. (c) Thermoelectric boundary conditions represent water cooling at the electrode interior and natural convection to air on exposed surfaces, expressed by convective heat loss equations with coefficients  $h_w$  and  $h_a$ .

are applied simultaneously. To enhance the accuracy of contact resistance modeling, Wan et al. [44] proposed a modified contact model, which has demonstrated high performance in capturing realistic resistance variations. This model is mathematically expressed in Eq. (6).

$$ECR_{A/B}(T, P) = \alpha_{A/B} \cdot ECR_{A/B}(T_0, P) \left( \frac{\sigma_{s-A/B}(T)}{\sigma_{s-A/B}(T_0)} \right)^{K_C} \left( \frac{\rho_A(T) + \rho_B(T)}{\rho_A(T_0) + \rho_B(T_0)} \right)$$

$$TCR_{A/B}(T, P) = \beta_{A/B} \cdot \frac{ECR_{A/B}(T, P)}{L_C T} \quad (6)$$

Where  $ECR_{A/B}(T, P)$  and  $TCR_{A/B}(T, P)$  are the electrical contact resistance (ECR) and thermal contact resistance (TCR) of the A/B interface at temperature  $T$  and pressure  $P$ , respectively.  $\alpha_{A/B}$  and  $\beta_{A/B}$  are the adjustable correction coefficients to be optimized.  $\sigma_{s-A/B}(T)$  denotes the yield strength of the softer of materials A and B at temperature  $T$ .  $\rho_A(T)$  and  $\rho_B(T)$  are the refractive indices of materials A and B at temperature  $T$ , respectively. The coefficient  $K_C$  is related to the oxide layer or coating on the surface of the material and varies between 1.0 and 1.5.  $K_C$  was set to 1.0 because the surface of the sheet was bare and clean.  $L_C$  denotes the Lorentz constant, which has a value of  $2.44 \times 10^{-8} \text{ W}\Omega/\text{K}^2$ .

In this study, the A/B interface is supposed to be one of the electrode/plate (E/S) and plate/plate (S/S) interfaces, and the electrode-to-plate contact parameters A (E/S), B (E/S) and the plate-to-plate contact parameters A (S/S), B (S/S) are the four variables to be optimized in the FEM below.

The melt core size in the simulation is measured using the temperature field output (NT11) from Abaqus, assuming that the interface temperature reaches the melting point of the material (1520 °C). The dynamic resistance curve is derived from the electric field potential (EPOT) and current density (ECD) using Eq. (1).

Fig. 3 presents the simulated melt core size and dynamic resistance curves for four different plate thickness combinations under specific welding parameters. The simulation and experimental data for melt core diameter and depth are compared in Table 2, where the maximum error is found to be 12.13 and 13.16 percent, respectively, as calculated using Eq. (7).

$$E_{\text{error}} = |V_{\text{sim}} - V_{\text{real}}| / V_{\text{real}} \quad (7)$$

In the simulation model, Joule heating is the primary heat source during the resistance spot welding process and is implemented through volumetric heat generation that evolves with the local electrical field and material properties. The electrical current is applied at the upper electrode tip, while the bottom electrode is grounded. Electrical contact resistances are defined at both electrode-sheet interfaces using contact pairs with calibrated resistivity values to simulate realistic oxide breakdown and dynamic conduction behavior. To quantify the net thermal energy available for nugget formation [45], as shown in Eq. (8), which incorporates resistive heating and dynamic cooling effects.

$$Q_{\text{nucleus}} = \int_0^{t_1} I^2(t)R(t)dt - \int_0^{t_2} C(t)dt \quad (8)$$

In Eq. (8), the first term represents the cumulative Joule heat input from the start of welding to the end of current application ( $t_1$ ), where  $I(t)$  and  $R(t)$  denote the instantaneous current and effective electrical resistance, respectively. The second term accounts for the real-time heat dissipation through the water-cooled electrodes during the full cycle ( $t_2$ ). The difference yields the net energy  $Q_{\text{nucleus}}$  that is effectively transferred to the weld zone to form the molten nugget. The resulting thermal distributions are strongly governed by both internal Joule heating and external heat loss through the electrode interfaces. These processes are dynamically solved in the coupled thermal-electrical simulation.

The dynamic resistance curve error is represented by three indicators, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared ( $R^2$ ) [46], respectively, and the smaller the values of MAE and RMSE, the closer the fitted value is to the actual value and the better the effect is, and the closer the value of  $R^2$  is to 1, the better the simulation model fits the experimental data. the closer to 1, indicating that the simulation model fits the experimental data better. The main factors causing the error come from the fact that the thermal conductivity and contact resistivity in the material properties have certain errors with the relevant parameters of the actual test materials, and are also affected by the environmental factors and the precision of the equipment in the actual test.

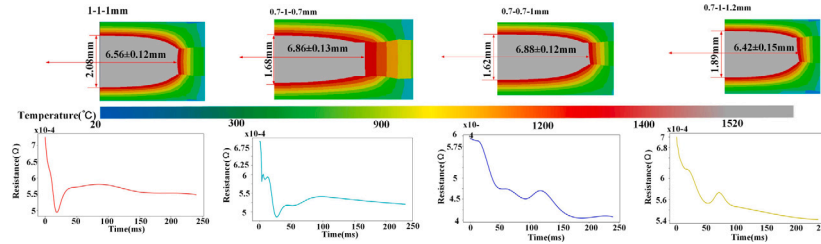
### 2.3. Sensitivity analysis of thermoelectric properties

To evaluate the influence of thermoelectric properties on simulation outcomes, a sensitivity analysis was conducted. Three key parameters were selected: (1) electrical contact resistance ( $p_{\text{contact}}$ ), (2) thermal

**Table 2**

Measured values of resistance spot welding core size and dynamic resistance of three-layer boards versus simulated values before optimization.

Source	Measured range	Simulated range	Error range
Nugget diameter (mm)	3.47–8.67	3.98–8.93	5.02%–12.13%
Weld penetration (mm)	1.31–2.96	1.15–2.78	6.19%–13.36%
Dynamic resistance curve error	RMSE=0.307–0.993%,	MAE=0.437–1.047%,	$R^2$ =57.6%–71.7%



**Fig. 3.** Average diameter and  $\pm 1\sigma$  standard deviation of nuggets and corresponding average dynamic resistance profiles for four different plate thickness configurations in five replicate welding tests with the same parameters (9 kA, 300 ms, 4 kN). These experimental results are used to validate the simulated temperature fields shown in the FEM models.

conductivity of the middle sheet ( $k_{\text{middle}}$ ), and (3) interfacial heat transfer coefficient ( $h_{\text{interface}}$ ). Each parameter was varied within a  $\pm 20\%$  range around its nominal value. The resulting effects on nugget diameter and dynamic resistance evolution were analyzed using the validated FEM model.

As illustrated in Fig. 4, the simulation outputs, namely nugget diameter and peak dynamic resistance, show varying degrees of sensitivity to changes in key thermoelectric parameters. Specifically, a  $\pm 20\%$  variation in electrical contact resistance ( $p_{\text{contact}}$ ) leads to the most pronounced effect, causing up to 18% deviation in the resistance peak and 15% change in nugget diameter. Even slight deviations in  $p_{\text{contact}}$  can dramatically alter the early-stage temperature rise, which in turn affects the timing and symmetry of nugget formation.

The interfacial heat transfer coefficient ( $h_{\text{interface}}$ ) exhibits a moderate influence—yielding 12% and 14% deviations in nugget diameter and resistance peak respectively. This indicates that interfacial heat dissipation plays a key role in controlling thermal gradients between layers, especially in asymmetric multi-sheet configurations like three-layer RSW. Poor interfacial conduction may cause excessive heat accumulation at the upper or lower interface, leading to unbalanced nugget growth and variations in process signals.

In contrast, the bulk thermal conductivity of the middle sheet ( $k_{\text{middle}}$ ) results in minimal deviation (5%–6%), suggesting that heat conduction through the middle layer has a comparatively weaker impact. This can be attributed to the fact that heat is mainly generated and dissipated through the interfaces, and the temperature rise within the middle layer is largely governed by the boundary conditions rather than the bulk material property itself.

These findings collectively highlight that interface-level properties dominate the thermal-electrical behavior of three-layer RSW, and any attempt to improve model fidelity must include precise calibration or optimization of contact resistance and interfacial heat transfer parameters. This strongly supports the necessity of integrating the proposed optimization framework for parameter tuning based on in-process signal feedback.

### 3. Deep learning and multi-objective optimization based approach

#### 3.1. Co-simulation optimization strategies

As illustrated in Fig. 5, this study establishes a co-simulation optimization framework that integrates finite element simulation (FE Model), deep learning, and optimization algorithms to enhance the accuracy of RSW simulation. The framework comprises the following key components:

**Real System Measurement Data.** Experimental measurements of dynamic resistance and melt core dimensions (diameter and depth) during the RSW process serve as a reference benchmark for subsequent model optimization.

**Finite Element Simulation (FE Model).** Using welding parameters (e.g., current, time, pressure), a finite element model is employed to numerically simulate the RSW process, predicting the molten core size and dynamic resistance. The simulation incorporates multi-physical field effects (e.g., temperature, potential, strain) to improve accuracy.

**Deep Learning Prediction Model.** A 1D CNN-BiLSTM-Attention network processes the measured dynamic resistance, extracts temporal features, and predicts the dynamic resistance curve, effectively fitting the experimental data. This predictive model reduces the need for repeated physical measurements and enhances the generalization of experimental data.

**ANSGA-II with Bayesian Optimization.** Based on measured and simulated data, a multi-objective optimization is conducted using ANSGA-II with Bayesian optimization. The optimization variables include weld quality parameters ( $\alpha_{E/S}$ ,  $\alpha_{S/S}$ ,  $\beta_{E/S}$ ,  $\beta_{S/S}$ ), while the objective functions  $F1$  and  $F2$  assess optimization performance.

**Optimization Objective Calculation.** By integrating experimental measurements, simulation outputs, and deep learning predictions, the optimization objectives ( $F1$  and  $F2$ ) are computed. The optimization schemes are then evaluated using cumulative summation, ensuring improved welding parameter adaptability and weld quality stability. The objective functions  $F1$ ,  $F2$  correspond to the error of dynamic resistance and the error of fusion core size, respectively, which are shown in Eq. (9), where  $E_D$ ,  $E_P$  are calculated by Eq. (7) for the errors of nugget diameter and weld penetration, respectively.

$$F1 = R^2 - MAE - RMSE, F2 = E_D + E_P \quad (9)$$

The proposed hybrid framework was designed to address the complexity and variability of the RSW process under small data and asymmetric geometric conditions. While it may appear computationally involved, each component serves a distinct and necessary role.

The 1DCNN-BiLSTM-Attention network leverages complementary feature extraction mechanisms: CNN captures short-term resistance trends, BiLSTM models time dependencies, and attention focuses on critical signal phases. This enables more accurate nugget prediction than simpler architectures. The ANSGA-II + Bayesian Optimization strategy is not used in real-time but rather for offline FEM parameter calibration. Compared to conventional NSGA-II or BO alone, this combination provides faster convergence, better parameter diversity, and enhanced alignment with experimental data. The overall framework

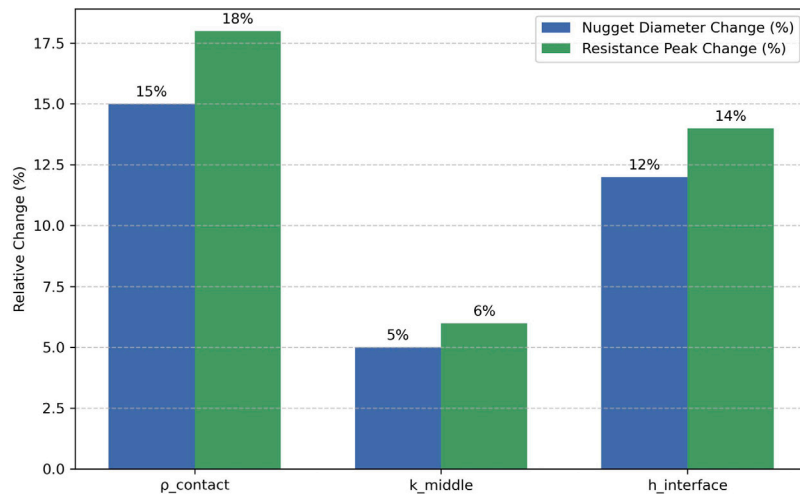
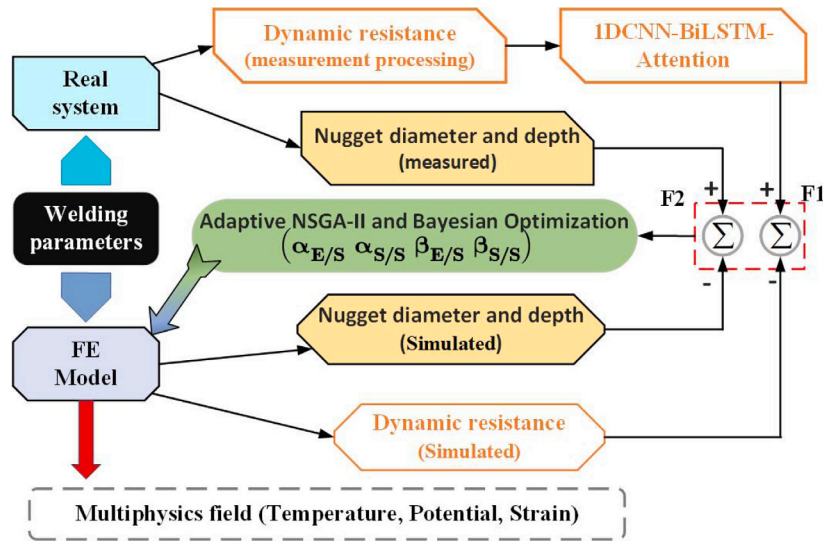


Fig. 4. Sensitivity Analysis of Thermoelectric Properties.

Fig. 5. Workflow of the proposed hybrid framework. Measured dynamic resistance is processed and fed to the 1DCNN-BiLSTM-Attention predictor. Errors in resistance (F1) and nugget geometry (F2) drive the offline ANSGA-II + Bayesian optimization, which updates the FEM contact parameters ( $\alpha_{E/S}, \alpha_{S/S}, \beta_{E/S}, \beta_{S/S}$ ) until the simulated outputs converge to experimental data. Blue blocks denote real-system acquisition, orange blocks the learning model, green the optimization core, and grey the multiphysics FEM solver.

thus ensures high prediction fidelity, optimization accuracy, and practical deployability, especially for complex multi-layer weld scenarios.

### 3.2. Prediction of dynamic resistance profiles based on 1DCNN-BiLSTM-attention models

1DCNN and BiLSTM have been proved by many studies to have good effect on the prediction of dynamic resistance [21–23]. The structure of the model in this paper is shown in Fig. 6, and the main parameters of the model are shown in Table 3, which contains four major modules: 1DCNN, BiLSTM, Attention Mechanism and Multilayer Perceptron (MLP).

1DCNN extracts the dynamic resistor local temporal features as shown in Eq. (10). Where  $W_c$  is the convolution kernel, represents the convolution operation,  $X$  is the input dynamic resistance data,  $b_c$  is the bias term,  $X_{CNN}$  and is the feature extracted by the CNN.

$$X_{CNN} = \text{ReLU}(W_c * X + b_c) \quad (10)$$

BiLSTM, on the other hand, extracts the global timing dependencies when the features extracted by 1DCNN are passed into BiLSTM for time series modeling, which is shown in Eq. (11). Where  $\bar{h}_t$  and  $\bar{h}_t$  are

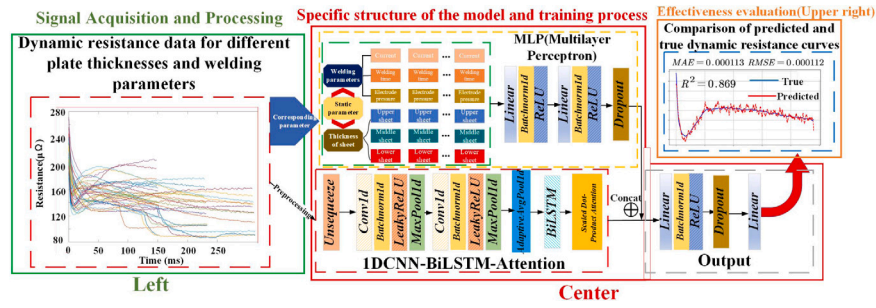
the hidden states of the forward and backward LSTMs, respectively,  $W_f, U_f, b_f$  and  $W_b, U_b, b_b$  are the weights and biases of the LSTM,  $\sigma$  is the activation function,  $h_t$  is the final output of the BiLSTM.

$$\begin{aligned} \bar{h}_t &= \sigma(W_f X_{CNN} + U_f \bar{h}_{t-1} + b_f) \\ \bar{h}_t &= \sigma(W_b X_{CNN} + U_b \bar{h}_{t+1} + b_b) \\ h_t &= [\bar{h}_t, \bar{h}_t] \end{aligned} \quad (11)$$

The dot product attention mechanism improves the influence of the critical time step and is used to give higher weight to the critical time step, as shown in the Eq. (12). Where  $H$  is the output of BiLSTM,  $W_q, W_k, W_v$  are the learnable weight matrices to convert the output of BiLSTM into queries, keys, and values, and  $d_k$  is the dimension of  $W_k$ .

$$a_t = \text{softmax}\left(\frac{H W_q (H W_k)^T}{\sqrt{d_k}}\right), h_{att} = \alpha(H W_v) \quad (12)$$

MLP combines the static welding parameters for final prediction, and to improve the prediction accuracy, the welding parameters are combined with the dynamic features as shown in Eq. (13). Where  $W_s$  is the static feature fully connected layer weight,  $X_{static}$  is the input



**Fig. 6.** Schematic structure based on the 1DCNN-BiLSTM-Attention model. **Left:** dynamic resistance signals collected from 400 welds with four thickness configurations and varying parameters. **Centre:** architecture of the proposed 1DCNN-BiLSTM-Attention network, where static welding parameters are processed by an MLP and concatenated with temporal features. **Upper right:** comparison between predicted and ground-truth resistance curves on the held-out test set ( $n = 60$ ); the reported MAE, RMSE and  $R^2$  values demonstrate high fidelity.

**Table 3**  
Main structural parameters of the model.

Module	Layer	Input	Output	Other Parameters
1DCNN-BiLSTM-Attention	Conv1d_1	(batch,1,seq_len)	(batch,32,seq_len)	kernel size = 5, padding = 1
	MaxPool1d_1	(batch,32,seq_len)	(batch,32,seq_len/2)	kernel size = 2
	Conv1d_2	(batch,32,seq_len/2)	(batch,64,seq_len/2)	kernel size = 3, padding = 1
	MaxPool1d_2	(batch,32,seq_len/2)	(batch,64,seq_len/4)	kernel size = 2
	AdaptiveAvgPool1D	(batch,64,seq_len/4)	(batch,64,1)	
	BiLSTM	(batch,seq_len/4, 64)	(batch,seq_len/4, 64)	hidden dim = 64
MLP	Attention	(batch,seq_len/4,64)	(batch,64)	
	fc_1	(batch, 6)	(batch, 128)	
	BatchNorm1d_1	(batch, 128)	(batch, 128)	
	fc_2	(batch, 128)	(batch, 64)	
Output	BatchNorm1d_2	(batch, 64)	(batch, 64)	
	Dropout	(batch, 64)	(batch, 64)	p=0.2
	Concat	(batch, 64 + 64)	(batch, 128)	
	fc_1	(batch, 128)	(batch, 128)	
	BatchNorm1d_1	(batch, 128)	(batch, 128)	
	Dropout	(batch, 128)	(batch, 128)	p=0.2
	fc_2	(batch, 128)	(batch, 1)	

**Table 4**  
Hyperparameterization of the model.

Hyperparameterization	Values
Learning rate	0.001
Optimizer	AdamW
Weight decay	$1 \times 10^{-5}$
Loss function	Smooth L1 Loss
Batch size	10
Maximum epochs	500
Learning rate scheduling	ReduceLROnPlateau

welding parameter,  $b_s$  is the bias term, and  $\mathbf{h}_{\text{static}}$  is the processed static feature.

$$\mathbf{h}_{\text{static}} = \text{ReLU}(\mathbf{W}_s \mathbf{X}_{\text{static}} + b_s) \quad (13)$$

$$\mathbf{R}_{\text{pred}}(t) = \mathbf{W}_o \left[ \sum_i \alpha_i \mathbf{h}_i, \text{ReLU}(\mathbf{W}_s \mathbf{X}_{\text{static}} + b_s) \right] + b_o \quad (14)$$

Combining the four modules, the model can be represented as Eq. (14). Where  $\alpha_i = \text{softmax}\left(\frac{HW_d(HW_k)^T}{\sqrt{d_k}}\right)$ ,  $\mathbf{h}_i = [\bar{\mathbf{h}}_i, \bar{\mathbf{h}}_i]$ ,  $\bar{\mathbf{h}}_i = \sigma(\mathbf{W}_f \mathbf{X}_{\text{CNN}} + U_f \bar{\mathbf{h}}_{i-1} + b_f)$ ,  $\mathbf{X}_{\text{CNN}} = \text{ReLU}(\mathbf{W}_c * \mathbf{X} + b_c)$ . The relevant parameters of the whole model are shown in Table 3 and the hyperparameters are shown in Table 4 for training, input four different plate thickness configurations under different welding parameters 100 data each, a total of 400 data after the completion of the training input (welding current, time, pressure, three-layer plate thickness) can predict the dynamic resistance curve.

To ensure fair and robust evaluation of the proposed 1DCNN-BiLSTM-Attention model, the dataset comprising 400 samples was randomly divided into training, validation, and testing subsets with a ratio

of 70:15:15. The training set (280 samples) was used to fit model weights, the validation set (60 samples) was used for hyperparameter tuning, and the test set (60 samples) was held out entirely for final performance reporting.

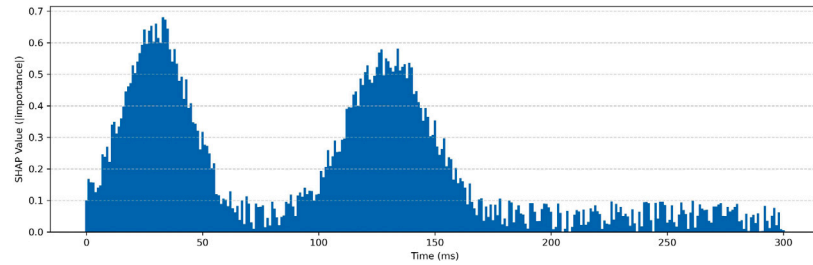
The model was implemented using PyTorch and trained using the Adam optimizer with an initial learning rate of 0.001. We applied Dropout (0.2) after recurrent and attention layers, Batch Normalization in the convolutional layers, and early stopping based on validation performance. As shown in Table 5, the deep learning model demonstrates consistent and high-quality performance across all data subsets. To comprehensively evaluate the performance of the deep learning model, three metrics were used: Mean Absolute Error (MAE) [47], Root Mean Squared Error (RMSE) [48], and R-squared ( $R^2$ ) [49].

The MAE on the training set is 0.028, which slightly increases to 0.032 on the validation set and 0.034 on the test set. This marginal increase (approximately 21% from training to test) indicates that the model generalizes well and is not overfitted to the training data. The RMSE shows a similar trend, increasing from 0.041 (train) to 0.047 (validation) and 0.049 (test), confirming the model's robustness even under slightly higher error sensitivity.



**Table 5**  
Deep learning model performance on validation and test sets.

Metric	Train set	Validation set	Test set
MAE	0.028	0.032	0.034
RMSE	0.041	0.047	0.049
$R^2$	0.95	0.93	0.92



**Fig. 7.** SHAP Feature Importance over 0 to 300 ms.

The  $R^2$  remains high across all sets, with values of 0.95 (train), 0.93 (validation), and 0.92 (test), suggesting a strong correlation between the predicted and true dynamic resistance curves. This high correlation is particularly important in welding quality prediction, where capturing the correct waveform trend is often more critical than minimizing pointwise errors.

It is worth noting that the training dataset contained only 400 labeled samples, due to the labor-intensive nature of RSW experiments and destructive testing required for accurate nugget measurement. To address this, the model architecture was designed with a relatively small number of parameters, combining convolutional and recurrent layers with an attention mechanism to extract both local and global features efficiently. In combination with dropout regularization and early stopping, this design helps prevent overfitting under limited data.

To enhance model interpretability, we employed Shapley Additive Explanations (SHAP) [50] to analyze the contribution of different input time steps to the model's output. This analysis was conducted using the final trained model and the test set.

Fig. 7 illustrates the SHAP-based feature importance of the input dynamic resistance signal over the full welding time (0 to 300 ms). Two distinct regions exhibit notably high SHAP values, indicating their significant contribution to the model's output (i.e., nugget diameter prediction).

The first important interval occurs between 0 to 60 ms, with a peak centered around 30 ms, which corresponds to the initial contact heating phase in resistance spot welding (RSW). During this stage, the resistance increases due to oxide film breakdown and poor initial contact, making this region highly informative for modeling the early onset of nugget formation.

The second high-SHAP region appears around 100 to 160 ms, peaking near 130 ms. This period aligns with the nugget growth phase, during which the current reaches its peak and maximum heat is generated at the faying surfaces. The molten metal volume expands rapidly, making this region crucial for predicting the final weld nugget size.

In contrast, the SHAP values in the middle (60 to 100 ms) and late stage (after 200 ms) are relatively low, indicating limited influence on model output. This aligns well with the physical expectation: after the nugget stabilizes, additional current or hold time contributes less to determining weld size.

Overall, the SHAP profile confirms that the model has learned to focus on physically meaningful and process-critical time windows, enhancing trust in the learned feature representation and validating the applicability of the deep learning framework for real-time weld quality prediction.

To clearly present the data flow and training logic of the proposed deep learning framework, a simplified training workflow is illustrated in Fig. 8. It summarizes the full process from data acquisition and pre-processing, through model initialization and mini-batch training, to validation, testing, and final deployment. This step-by-step overview helps clarify the temporal logic and modular structure of our implementation, supporting reproducibility and interpretability.

While the ReLU function was initially used throughout the network due to its simplicity and computational efficiency, we performed an additional ablation study to assess whether alternative activations, specifically LeakyReLU and GELU, would improve model performance, particularly in low-activation regions.

As shown in Table 6, while all three activations provided acceptable results, ReLU achieved slightly better generalization and lower error. We hypothesize that this is due to its inherent sparsity, which helps regularize model complexity and prevent overfitting on our relatively small dataset. Therefore, ReLU was retained in the final architecture.

### 3.3. ANSGA-II and Bayesian optimization(BO)

In recent years, the NSGA-II algorithm has been successfully applied to engineering optimization [51]. Adaptive Non-dominated Sorting Genetic Algorithm II (ANSGA-II) is an improved NSGA-II [52], which performs optimization through genetic algorithms and uses an adaptive tuning mechanism to enhance search efficiency and diversity. In this paper, the research is mainly used to optimize the thermal and electrical conductivity-related parameters ( $\alpha_{E/S}$ ,  $\alpha_{S/S}$ ,  $\beta_{E/S}$ ,  $\beta_{S/S}$ ) in RSW simulation to minimize the dynamic resistance error (F1) and the nugget diameter and weld penetration (F2). A set of optimal parameters is obtained through Pareto optimization instead of a single solution, which makes the simulation results more stable and robust. The specific parameters are shown in Table 7, and the detailed steps are shown in the Algorithm 1.

In adaptive NSGA-II, the crossover and mutation probabilities change dynamically according to the population diversity and fitness distribution as shown in Eq. (15). Where  $p_{c_{\min}}$ ,  $p_{c_{\max}}$  are the minimum and maximum crossover probabilities.  $p_{m_{\min}}$ ,  $p_{m_{\max}}$  are the minimum and maximum mutation probabilities.  $D_t$  represents the population diversity at generation  $t$ .  $f_t$  is the average fitness of the population at generation  $t$ .

$$p_c = p_{c_{\min}} + (p_{c_{\max}} - p_{c_{\min}}) \times \frac{D_t}{D_{\max}}$$

$$p_m = p_{m_{\max}} - (p_{m_{\max}} - p_{m_{\min}}) \times \frac{f_t - f_{\min}}{f_{\max} - f_{\min}} \quad (15)$$

To maintain solution diversity, non-dominated sorting ranks solutions, and crowding distance helps maintain diversity within a front as

**Table 6**

Comparison of different activation functions in the proposed 1DCNN-BiLSTM-Attention network.

Activation function	MAE	RMSE	$R^2$
ReLU	<b>0.132</b>	<b>0.156</b>	<b>0.91</b>
LeakyReLU	0.136	0.160	0.90
GELU	0.139	0.163	0.89

**Table 7**

ANSGA-II parameter settings.

Parameters	Value
Pop_size	100
Sampling	LHS()
Crossover	SBX(prob=0.9, eta=20)
Mutation	PM(eta=20)
Eliminate_duplicates	True
Termination	150

**Algorithm 1: ANSGA-II Pseudocode**


---

**Input:** Population size, number of generations, initial crossover and mutation probabilities, objective functions

**Output:** Pareto-optimal solutions

---

```

1  $P \leftarrow$  Generate the initial population;
2 Evaluate the objective function values for each individual in  $P$ ;
3 Perform non-dominated sorting and compute crowding distance for  $P$ ;
4 Initialize adaptive parameters  $\alpha_c, \alpha_m$  for crossover and mutation probability adjustment;
5 while termination criterion is not met do
6   Adaptively adjust crossover probability  $p_c$  and mutation probability  $p_m$  based on population diversity;
7    $S_c \leftarrow$  Perform selection, crossover (using adaptive  $p_c$ ), and mutation (using adaptive  $p_m$ ) to generate offspring;
8    $T_c \leftarrow P \cup S_c$  (Merge parent and offspring populations);
9   Perform non-dominated sorting and compute crowding distance for  $T_c$ ;
10   $P \leftarrow$  Select the top individuals based on non-dominated sorting and crowding distance from  $T_c$ ;
11 end
12 return The final Pareto-optimal set  $P$ ;
```

---

shown in Eq. (16). Where  $f_j$  denotes the objective function.

$$d_i = \sum_{j=1}^M \left( \frac{f_j^{i+1} - f_j^{i-1}}{f_j^{\max} - f_j^{\min}} \right) \quad (16)$$

As shown in Fig. 9, the Pareto fronts for different plate thicknesses with welding parameters of 9000 A, 300 ms and 3.5 KN show similar trends, with the difference being the sparsity of the fronts. Meanwhile, as shown in Table 8, it is compared with other multi-objective algorithms under the same settings, and Hypervolume metric (HV) and generational distance metric (GD) are chosen as the evaluation metrics [53], and the ANSGA-II performs best on HV (0.875) and GD (0.0156), which indicates that it converges fast and has a better distribution of solutions.

Optimization problems are prevalent in the scientific and manufacturing industries and typically require a large number of function summations. In contrast, Bayesian optimization (BO) has become a mainstay in solving these difficulties due to its efficient search, thanks to its ability to combine a priori beliefs about the problem to help guide the sampling of new data and achieve a good balance between exploration and exploitation in the search [54]. While in this paper, after ANSGA-II finds the optimal solution set using Bayesian optimization, it is further fine-tuned to find the global optimal point. The parameters of

its algorithmic setup are shown in Table 9, and its detailed procedure is reflected in the pseudo-code in Algorithm 2.

**Algorithm 2: General framework of BO**


---

**Input:** Search space, objective function, acquisition function, number of iterations

**Output:** Optimal solution of the objective function

---

```

1 Initialize the Gaussian process model with an initial set of sampled points;
2 while termination criterion is not met do
3   Use the acquisition function to select the next sampling point;
4   Evaluate the objective function at the selected point;
5   Update the Gaussian process model with the new data;
6 end
7 return The best-observed solution based on the objective function;
```

---

As shown in Eq. (17), the BO method models the objective function  $f(x)$  as a Gaussian process. Where  $\mu(x)$  is the predictive mean and  $k(x, x')$  is the covariance function.

$$f(x) \sim \mathcal{GP}(\mu(x), k(x, x')) \quad (17)$$

Eq. (18), on the other hand, indicates that in order to balance exploration and exploitation, Expected Improvement (EI) is used to select the next assessment point. Where  $\Phi(Z)$  and  $\phi(Z)$  are the cumulative and probability density functions of the standard normal distribution.

$$EI(x) = (\mu(x) - f^*) \Phi(Z) + \sigma(x) \phi(Z), Z = \frac{\mu(x) - f^*}{\sigma(x)} \quad (18)$$

As shown in Fig. 10 for different plate thicknesses in the welding parameters of 9000 A, 300 ms, 3.5KN, the BO function according to the convergence of the objective function (F1), usually after 100 iterations to reach convergence, to find the optimal value, where the plate thickness configuration of 1-1-1 mm convergence of the value of the best.

To validate the effectiveness of the proposed optimization framework, a comparative analysis was conducted between three optimization strategies: (1) standard NSGA-II, (2) standalone Bayesian Optimization (BO), and (3) the hybrid ANSGA-II + BO method. All three approaches were applied to calibrate the FEM model by minimizing the error between simulated and experimentally measured resistance signals, as well as improving nugget diameter prediction.

Table 10 presents the performance metrics for each method. The results clearly demonstrate that the proposed hybrid approach achieves superior accuracy and convergence. Specifically, ANSGA-II + BO reduced the resistance curve MAE by 18.6% and improved nugget diameter accuracy by 15.2% compared to NSGA-II alone. Although BO converged faster than NSGA-II, it was prone to local optima due to limited global search capability. By combining the adaptive crowding and

**Table 8**  
Performance Comparison of Multi-Objective Optimization Algorithms.

Algorithm	ANSGA-II	NSGA-II	MOEA/D	SPEA2
HV	0.863	0.832	0.816	0.805
GD	0.0256	0.0341	0.0427	0.0485

**Table 9**  
BO parameter settings.

Parameters	Value
Spatial definition	Gaussian Process (GP)
Acq_func	Elevated expectations
Initial sampling point	ANSGA-II results
Optimize number of iterations	10

**Table 10**  
Performance comparison of optimization methods.

Method	Resistance MAE	Nugget diameter error (mm)	Convergence time (s)
NSGA-II	0.152	0.24	880
Bayesian Optimization	0.137	0.19	620
ANSGA-II + BO (Proposed)	<b>0.124</b>	<b>0.17</b>	<b>580</b>

elitism strategies of ANSGA-II with the probabilistic refinement of BO, the hybrid method provides a balanced trade-off between exploration and exploitation.

To assess the real-time feasibility of the proposed framework, we measured the computational latency of both the prediction model and the optimization workflow. The 1DCNN-BiLSTM-Attention model was tested on a general-purpose CPU without GPU acceleration. The average inference time per sample was 18 ms, which is significantly lower than the typical resistance spot welding cycle time of 200 to 400 ms, ensuring compatibility with real-time process monitoring.

The optimization workflow involving ANSGA-II and Bayesian Optimization is executed offline during system initialization or periodic model recalibration. For each configuration, the optimization process converged within approximately 2 min using 50 generations and a population size of 40. This computation is performed independently of real-time welding and does not affect the online prediction stage. Therefore, the proposed framework adopts a hybrid deployment strategy, where the deep learning model serves real-time inference needs, while the simulation optimization loop supports offline calibration and model enhancement. This division ensures both performance and practicality in industrial applications.

4. Results and discussion

The comparison of real melt core sizes with those simulated using the 1DCNN-BiLSTM-Attention model in conjunction with ANSGA-II and Bayesian optimization—is illustrated in Fig. 11. The benchmark welding parameters used are a welding current of 9000 A, welding time of 300 ms, and electrode pressure of 3.5 kN. Results show that the optimized simulated nugget diameter and weld penetration across all four plate thickness combinations exhibit significantly closer alignment with experimental measurements. Among them, configuration (a) demonstrates the best optimization performance, while (b) and (c) show comparable improvement, both yielding satisfactory results. However, configuration (d)—which involves three layers of differing thicknesses—exhibits relatively poorer prediction accuracy. This discrepancy is attributed to the increased complexity in heat input distribution and current flow due to greater variation in thermal and electrical conductivity parameters across the layers. To further analyze the influence of individual welding parameters, single-factor experiments were conducted using the same benchmark settings. The trends in melt core evolution under varying welding current, welding time, and electrode pressure are detailed in Fig. 12(a)(b)(c). The results

of finite element simulation and the performance of the optimization algorithms under these conditions are analyzed in the subsequent discussion.

In terms of variation with welding parameters, it was found that with an increase in welding current, the overall increase in the diameter and depth of the nucleus was due to the greater current input, which enhanced the Joule heat and caused the nucleus to grow, and the increase in the nucleus size was slowed down when the current was more than 10–11 kA, and even overheating may lead to material burnout and spattering, which affects the quality of the nucleus. The growth rate of the nucleus varies for different plate thicknesses, with thicker plates (e.g., 0.7-1-1.2 mm) showing a slower increase in nucleus size and thinner plates (1-1-1 mm) showing a more significant increase in nucleus size.

The nugget diameter and Weld penetration of the nucleus increase with welding time as heat builds up, causing more metal to melt. After 240–300 ms, the size of the nucleus tends to stabilize, indicating that thermal equilibrium has been reached, and any further extension of time will result in limited improvement and may even affect the quality of the weld due to overheating. The growth rate of the nucleus varies with the thickness of the plate. Thick plates (e.g., 0.7-1-1.2 mm) have a slower growth rate of the nucleus, while thin plates (1-1-1 mm) have a faster growth rate. When the electrode pressure is less than 2.5–3.5 kN, the size of the nucleus increases as the pressure increases because better interfacial contact reduces the interfacial resistance and concentrates the heat more. When the pressure is higher than 3.5 kN, the size of the nucleus starts to decrease because the excessive pressure leads to lower contact resistance, less efficient heating of the welded joint, and more heat dissipation through the electrode. The sensitivity to pressure is more pronounced for thicker plates, e.g., 0.7-1.0-1.2 mm, where too much pressure can cause the joint to cool too quickly and make the nucleus smaller, whereas thinner plates, 1.0-1.0-1.0 mm, have less effect on the pressure change.

After the model and algorithm-optimized simulated values were analyzed with the experimental values, which are shown in Fig. 12 and Table 11, it was found that for the process of welding current variation, the optimized simulated values were larger than the experimental values in the low-current section (6–8 kA), but the optimized trend of the diameter and depth of the molten nucleus matched the real situation. The error decreases in the high current band (9–12 kA) stage, mainly compensating for the nonlinear effects of thermal and electrical conductivity. For the variation with welding time, in the short welding time (<200 ms) range, the optimized simulated nuclei are larger than the measured values, compensating for the error in the

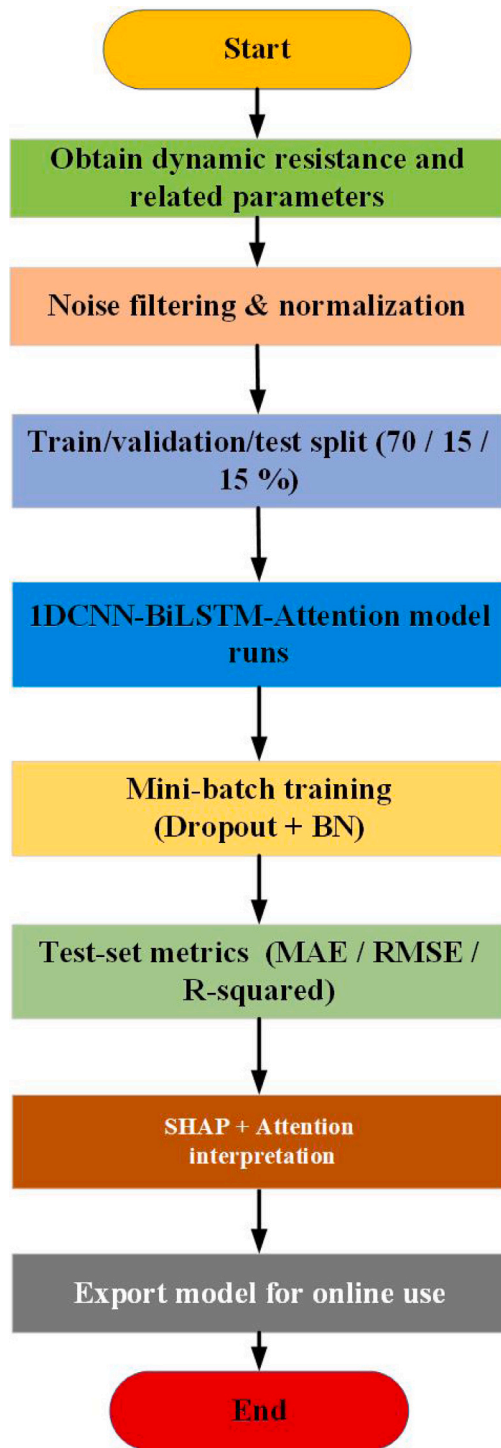


Fig. 8. Overview of the training workflow for the IDCNN-BiLSTM-Attention model.

conductive contact resistance, and accordingly the optimized simulated values in the long welding time ( $> 200$  ms) have a smaller error with the experimental values, indicating that the algorithm is more capable of predicting large-sized nuclei. With the change of electrode pressure, the optimized simulated nuclei size is smaller than the real value in the low-pressure section (1–3.5 kN), which reduces the overestimation problem of the simulation. When in the high-pressure section ( $> 3.5$  kN), the error is smaller and the optimized melt core depth is closer to the experimental value. It can be seen that the low current section

compensates, reduces the error in the size of the melt nucleus, the long time optimization is optimal, the short time melt nucleus is closer to the experimental value, the low pressure optimization is obvious, and the high pressure reduces the simulation underestimation phenomenon.

Comparing the trend of the simulation optimization results with the real values, it can be found that the diameter of the melt core after optimization is 0.2–0.3 mm larger than the real value as a whole, and the depth of the melt core is increased by 0.1–0.3 mm, which indicates that the optimization model improves the efficiency of the melt core formation. Then the trend of the optimized changes with welding current, welding time, and electrode pressure is the same as the real value, but the optimized melt nucleus expansion is more uniform, indicating that the improved thermal conductivity and electrical conductivity correction model improves the welding simulation accuracy. Specifically, the optimized melt core increase performs best at medium-high welding current (9–12 kA), moderate welding time (200–300 ms), and moderate electrode pressure (2.5–3.5 kN), which proves that the model is more suitable for a reasonable range of welding parameters. However, under extreme conditions (e.g., a low current of 6 kA or excessive pressure of 5 kN), the enhancement of the optimized melt core diameter and depth is not obvious, indicating that the physical limiting factor still exists and the optimization cannot enhance the melt core size infinitely.

Regarding the error of the dynamic resistance curves, the dynamic resistance plots of different plate thickness combinations under the welding parameters of 9000 A, 240 ms, and 4 kN correspond to those shown in Fig. 13, and corresponding to Fig. 2, it is found that the resistance values of the original simulation curves are significantly higher than those of the experimental curves and the optimized simulation curves under all plate thickness configurations. The overall trend of the real resistance curves and the optimized simulation curves is closer on the time axis, and the optimized simulation obviously improves the error of the original simulation. In the initial resistance stage (0–50 ms), the resistance of all plate thickness combinations decreases rapidly at the beginning of energization and then tends to stabilize. The resistance drop of the three configurations (b), (c), and (d) in Fig. 13 is more obvious than that of (a), indicating that the resistance changes more drastically when the plate thickness combinations are different, which may be related to the change of the contact resistance between the layers of different thicknesses. In the steady-state resistance stage (after 50 ms), both the real curve and the optimized simulation curve tend to be stable, indicating that the welding enters the steady-state heating stage and the resistance no longer changes drastically. The optimized curve is closer to the real value, while the original simulation curve is higher overall, which indicates that the optimization method effectively corrects the deviation between the original simulation and the actual experiment. From the different plate thickness configurations, it is found that (a) has the lowest resistance overall, probably due to the equal thickness of the three layers, lower contact resistance, and more uniform conductive path. (b) The resistance profile is slightly higher than (a), indicating that the thickness inhomogeneity increases some of the contact resistance. The resistance curves of both (c) and (d) are similar to (b), but (d) has the highest initial resistance, probably due to the 1.2 mm thick plate increasing the overall resistance path and affecting the current distribution.

From the overall view of different plate thickness configurations of three-layer plates, as Table 12 shows the error values of the simulation curves before and after the optimization of four different plate thickness configurations relative to the real or predicted curves, in terms of the nucleus diameter and nucleus depth errors, the optimization of the models and algorithms reduces the range of nucleus diameters and weld depths of the plate thickness configurations, and the errors are significantly reduced. For example, the melt core diameter for the 1-1-1 mm configuration was reduced from 5.02–9.63 percent to 3.9–7.54 percent, and the weld fusion depth was reduced from 6.19–9.57 percent



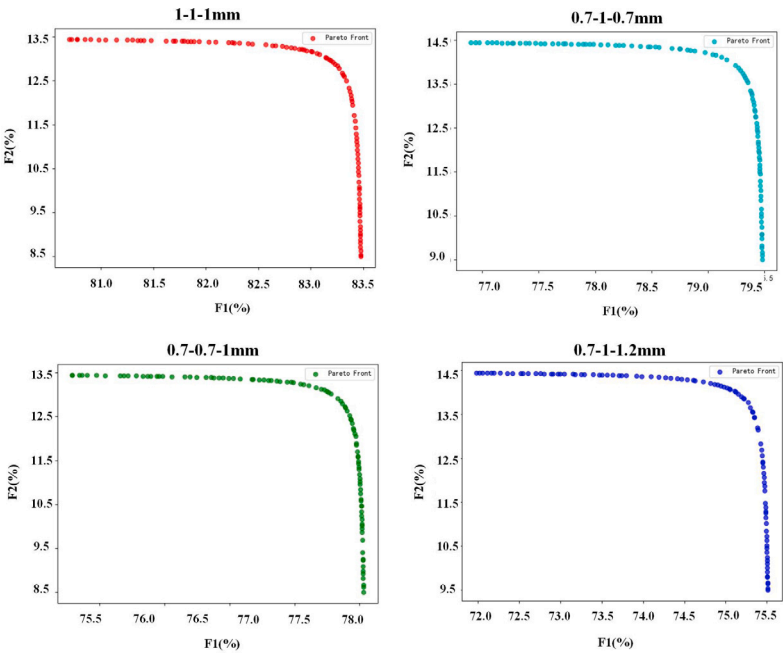


Fig. 9. Pareto front condition for different plate thickness configurations with the same welding parameters.

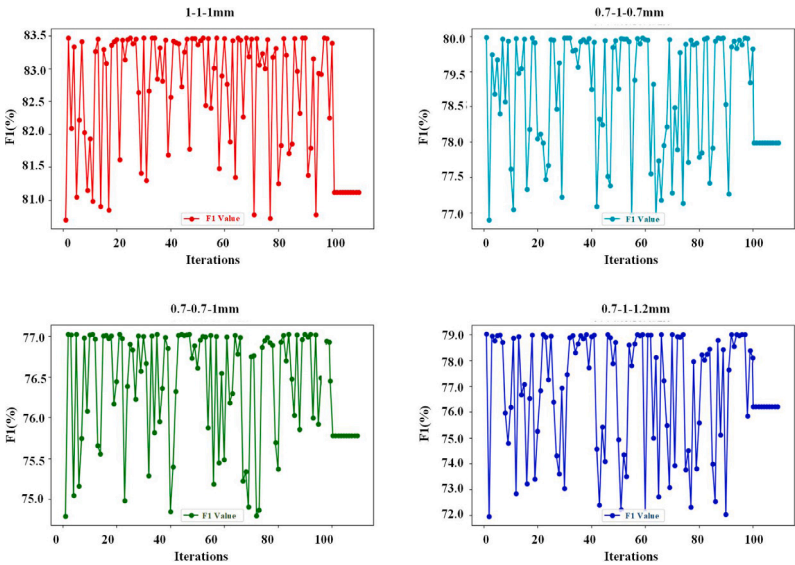


Fig. 10. Convergence of Bayesian optimization function for four plate thickness configurations with the same welding parameters.

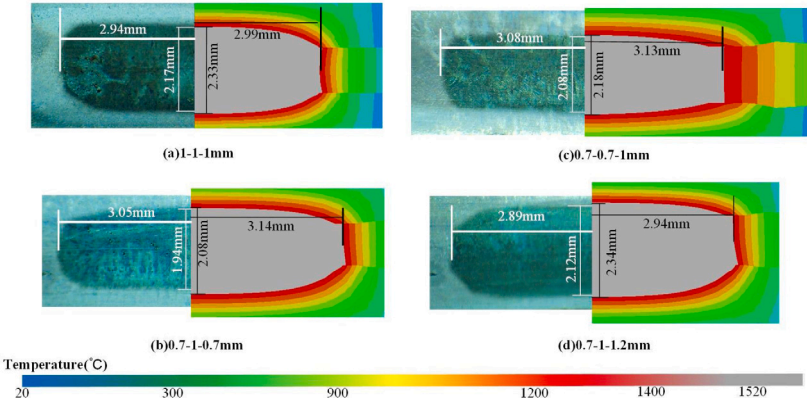


Fig. 11. Comparison of experimental and simulation-optimized simulated values of melt core size for resistance spot welding of three-layer plates.

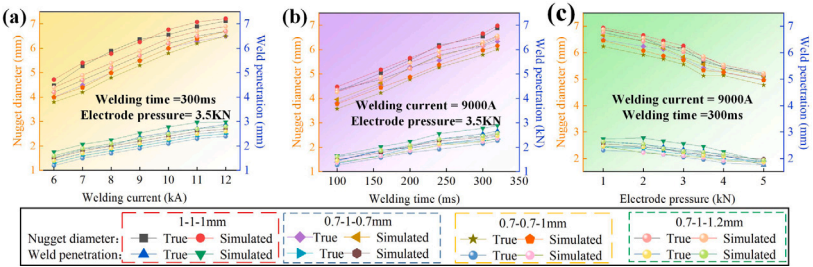


Fig. 12. Comparison of experimental and simulation-optimized simulated values of melt core size of resistance spot welded three-layer plate under variation with different welding parameters.

**Table 11**  
Comparison of simulation-optimized and experimental results: tolerance range (%) of nugget diameter and weld penetration under different parameter levels. Best results in each column are highlighted in **bold**.

Parameter	Range	Nugget diameter tolerance (%)	Weld penetration tolerance (%)
Welding current	6–8 kA	2.23 – 5.59	2.26 – 4.87
	9–12 kA	<b>1.43 – 3.57</b>	<b>2.16 – 4.53</b>
Welding time	< 200 ms	1.83 – 4.79	1.32 – 4.65
	> 200 ms	<b>1.24 – 3.87</b>	<b>1.43 – 3.79</b>
Electrode pressure	1–3.5 kN	<b>1.25 – 3.14</b>	<b>1.57 – 4.29</b>
	> 3.5 kN	1.98 – 3.54	1.87 – 4.38

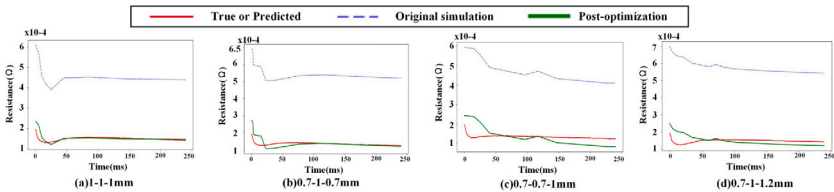


Fig. 13. Resistance spot welding experiments on three-layer boards with different board thickness configurations, comparison of dynamic resistance graphs before and after simulation optimization.

to 4.64–7.89 percent. Similarly, the melt core diameter and weld depth of fusion decreased by about 30 percent and 25 percent, respectively, for the 0.7-0.7-1 mm configuration. This trend indicates that the optimized simulation model is able to predict the weld joint size more accurately and reduce the overestimation of the melt core size.

However, for the 0.7-1-1.2 mm configuration, which has a large thickness variation, the melt core diameter and weld depth of fusion errors are still higher than the other three configurations, despite the reduction in the optimization. Its molten core diameter ranges from 7.54 to 12.13 percent before optimization and remains at 5.28 to 8.52 percent after optimization, and the weld fusion depth ranges from 7.88 to 13.36 percent before optimization and still reaches 5.46 to 9.28 percent after optimization. This suggests that the inhomogeneity of heat input and the variation of heat transfer paths during the welding process can lead to an increase in the complexity of melt core morphology prediction under large thickness variations, making the optimization method relatively limited in its effectiveness. In terms of the trends of the melt nucleus diameter and weld depth, the optimized simulation model better matches the experimental data, which is mainly manifested in the adjustment of the growth rate of the melt nucleus and the enhancement of the stability of the weld depth. This may be attributed to the correction of material thermal conductivity, current density distribution, and interface contact conditions during the optimization process, which makes the simulation calculation more in line with the actual physical process of welding. In addition, the fluctuation amplitude of the optimized melt core size is significantly reduced, indicating that the optimization method effectively reduces the instability of the model and makes the simulation prediction results more consistent.

In terms of dynamic resistance error, the dynamic resistance errors of all board thickness configurations are significantly reduced through optimization, with the RMSE and MAE both reduced by about 20 to 40 percent, and the  $R^2$  values are all improved to more than 75 percent, indicating that the optimization method effectively improves the fitting ability of the simulation model. Among them, the 1-1-1 mm configuration has the best optimization effect, with the  $R^2$  improved from 62.4%–71.7% to 78.88%–90.56% and the lowest optimized RMSE of 0.207%–0.391%. In contrast, the  $R^2$  of 0.7-1-1.2 mm is still lower than the other three configurations, although it is improved to 71.88%–85.6%, indicating that the optimization method can improve the simulation model fitting capability in the thickness difference. large, the adaptability of the optimization method still needs to be improved.

During model evaluation across various three-layer stack-up configurations, we observed a significant drop in predictive performance on the 0.7-1-1.2 mm specimen compared to more balanced setups such as 1-1-1 mm and 0.7-1-0.7 mm. This particular stack-up presents a pronounced geometric asymmetry: the top and bottom sheets differ substantially in thickness, and the overall thermal path becomes highly unbalanced.

The model exhibited a higher MAE and RMSE on this configuration, suggesting reduced prediction stability. Analysis of the dynamic resistance signals shows larger fluctuations, delayed peak formation, and more abrupt resistance drops as shown in combination with Fig. 13 and Table 12. These behaviors are consistent with non-uniform heat flow and asymmetric nugget evolution due to uneven current and heat distribution. FEM simulation results corroborate this finding. As shown in Fig. 11(d), the temperature field in the 0.7-1-1.2 mm setup was significantly skewed, with a noticeable delay in molten core expansion

**Table 12**  
Comparison of tolerance ranges before and after simulation optimization across four different plate thickness configurations. Improved (post-optimization) results are highlighted in **bold**.

Metric	Plate thickness configuration							
	1-1-1		0.7-1-0.7		0.7-0.7-1		0.7-1-1.2	
	Original	Optimized	Original	Optimized	Original	Optimized	Original	Optimized
Nugget diameter (%)	5.02–9.63	<b>3.90–7.54</b>	6.85–9.77	<b>4.67–7.74</b>	7.11–11.26	<b>5.08–7.93</b>	7.54–12.13	<b>5.28–8.52</b>
Weld penetration (%)	6.19–9.57	<b>4.64–7.89</b>	6.32–10.10	<b>4.71–7.64</b>	6.18–9.96	<b>4.96–7.89</b>	7.88–13.36	<b>5.46–9.28</b>
Dynamic resistance	RMSE	0.307–0.508	0.411–0.654	<b>0.318–0.467</b>	0.456–0.687	<b>0.337–0.520</b>	0.618–0.993	<b>0.421–0.706</b>
	MAE	0.437–0.669	0.589–0.806	<b>0.388–0.568</b>	0.598–0.788	<b>0.402–0.546</b>	0.659–1.047	<b>0.452–0.654</b>
	R <sup>2</sup>	62.4–71.7	61.2–70.2	<b>75.6–89.98</b>	60.2–72.3	<b>77.4–88.64</b>	57.6–69.5	<b>71.88–85.6</b>

at the bottom interface. This is due to the thicker bottom sheet acting as a thermal sink, causing uneven nugget growth and poor symmetry, which are difficult for the model to generalize from prior training. These results collectively demonstrate that strong geometric asymmetry introduces additional thermal-mechanical complexity and signal variability, which compromise the prediction accuracy of the data-driven model. This highlights a limitation of the current framework: while effective on balanced or moderately asymmetric configurations, its robustness degrades under more extreme structural gradients. To address this, future research will focus on expanding the training dataset with more highly asymmetric samples to improve model generalization, integrating geometry-aware attention mechanisms, allowing the model to modulate its focus based on local thickness information, and incorporating physics-informed priors from FEM simulation to help constrain the learning space under asymmetric conditions.

5. Conclusion

In this paper, an optimization framework combining deep learning and multi-objective optimization is proposed to effectively reduce the problem of insufficient accuracy of numerical simulation, improve the accuracy of the simulation results of dynamic resistance and molten core size in spot welding of three-layer unequal-thickness plates, and solve the problem of insufficient data. The experimental results show that the simulation-optimized dynamic resistance profile and melt core size are better matched with the experimental data. Specifically:

Deep learning assists simulation optimization to improve accuracy. The model based on 1DCNN-BiLSTM-Attention is used to predict the dynamic resistance curve of welding, which significantly improves the fitting ability of the resistance curve and provides more accurate target data for optimization.

ANSGA-II synergizes with Bayesian optimization to improve computational efficiency. Compared with the traditional parameter adjustment method, the ANSGA-II and Bayesian optimization methods can quickly converge to the optimal solution, which improves the parameter optimization efficiency of the RSW simulation.

The optimized simulation results are closer to the real values. The experimental results show that the optimized dynamic resistance curves of RMSE and MAE are reduced in the ranges of 17.6%–32.1% and 15.9%–28.1%, respectively, and the reduction of nugget diameter and weld penetration errors are reduced in the ranges of 15.3%–22.7% and 16.8%–25.4%, which verifies the effectiveness of the method.

The optimization of the finite element model for different plate thickness configurations of the three-layer plate is effective. Under the welding conditions of 1-1-1 mm, 0.7-1-0.7 mm, 0.7-0.7-1 mm, and 0.7-1-1.2 mm three-layer mild steel, the simulation accuracy can be significantly improved, which provides a reliable engineering application for the optimization of complex welding parameters.

CRediT authorship contribution statement

**Haofeng Deng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pengyu Gao:** Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Honggang Xiong:** Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xiangdong Gao:** Writing – review & editing, Software, Resources, Project administration, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This work was supported by the Production and Research Project, China (No. 23HK0610), the Guangdong Provincial Natural Science Foundation of China under Grant 2023A1515012172, and the Guangzhou Municipal Special Fund Project for Scientific and Technological Innovation and Development, China under Grant 2023B03J1326.

Data availability

The data are not publicly available due to privacy or ethical restrictions.

References

[1] Sun Zhengyi, Wang Bin, Yao Youqiang, Wang Sheng, Zhang Wenyi, Zhang Hongyan. Thermo-mechanical simulation of friction spot joining of aluminum alloys to short glass fiber-reinforced PPS. *CIRP J Manuf Sci Technol* 2023;43:30–41.

[2] Geng Chen, Sheng Buyun, Fu Gaocai, Chen Xiangxiang, Zhao Guangde. A GAN-based method for diagnosing bodywork spot welding defects in response to small sample condition. *Appl Soft Comput* 2024;157:111544.

[3] Gao Yunfeng, Wang Bin, Chen Junyi, Sun Zhengyi, Zhang Hongyan. Optimizing joint structure and performance in friction spot joining of aluminum alloy AA6061 and poly-ether-ether-ketone (PEEK). *CIRP J Manuf Sci Technol* 2023;47:273–82.

[4] Pouranvari M, Marashi SPH. Weld nugget formation and mechanical properties of three-sheet resistance spot welded low carbon steel. *Can Metall Q* 2012;51(1):105–10.

[5] Shen Jie, Zhang Yansong, Lai Xinmin, Wang PC. Modeling of resistance spot welding of multiple stacks of steel sheets. *Mater Des* 2011;32(2):550–60.

- [6] Ashiri Rouholah, Haque Md Anwarul, Ji Chang-Wook, Salimijazi Hamid Reza, Park Yeong-Do, et al. Supercritical area and critical nugget diameter for liquid metal embrittlement of Zn-coated twinning induced plasticity steels. *Scr Mater* 2015;109:6–10.
- [7] Gao Xiangdong, Lan Chongzhou, You Deyong, Li Guohua, Zhang Nanfeng. Weldment nondestructive testing using magneto-optical imaging induced by alternating magnetic field. *J Nondestruct Eval* 2017;36:1–11.
- [8] Gao Xiangdong, Mo Ling, You Deyong, Li Zhuman. Tight butt joint weld detection based on optical flow and particle filtering of magneto-optical imaging. *Mech Syst Signal Process* 2017;96:16–30.
- [9] Gao Xiangdong, Sun Yan, You Deyong, Xiao Zhenlin, Chen Xiaohui. Multi-sensor information fusion for monitoring disk laser welding. *Int J Adv Manuf Technol* 2016;85:1167–75.
- [10] Zhou Kang, Yao Ping. Overview of recent advances of process analysis and quality control in resistance spot welding. *Mech Syst Signal Process* 2019;124:170–98.
- [11] Summerville CDE, Adams D, Compston P, Doolan M. Process monitoring of resistance spot welding using the dynamic resistance signature. *Weld J* 2017;11:403–12.
- [12] Zhao Dawei, Bezgans Yuriy, Wang Yuanxun, Du Wenhao, Vdonin Nikita. Research on the correlation between dynamic resistance and quality estimation of resistance spot welding. *Measurement* 2021;168:108299.
- [13] Xing Bobin, Xiao Yi, Qin Qing H, Cui Hongzhi. Quality assessment of resistance spot welding process based on dynamic resistance signal and random forest based. *Int J Adv Manuf Technol* 2018;94:327–39.
- [14] Zhuo Wenbo, Tan Guobi, Chen Qiuren, Hou Zehong, Wang Xianhui, Han Weijian, et al. Multi-objective optimization of resistance spot welding process parameters of ultra-high strength steel based on agent model and NSGA-II. *Trans China Welding Inst* 2024;45(4):20–5.
- [15] Stavropoulos Panagiotis, Sabatakakis Kyriakos, Papacharalampopoulos Alexios, Mourtzis Dimitris. Infrared (IR) quality assessment of robotized resistance spot welding based on machine learning. *Int J Adv Manuf Technol* 2022;1–22.
- [16] Gao Xiangdong, Wang Lin, Chen Ziqin, Zhang Yanxi, You Deyong. Process stability analysis and weld formation evaluation during disk laser-mag hybrid welding. *Opt Lasers Eng* 2020;124:105835.
- [17] Gao Xiangdong, Li Yanfeng, Zhou Xiaohu, Dai Xinxin, Zhang Yanxi, You Deyong, et al. Multidirectional magneto-optical imaging system for weld defects inspection. *Opt Lasers Eng* 2020;124:105812.
- [18] Li Chengxi, Zheng Pai, Yin Yue, Wang Baicun, Wang Lihui. Deep reinforcement learning in smart manufacturing: A review and prospects. *CIRP J Manuf Sci Technol* 2023;40:75–101.
- [19] Bogaerts Lars, Dejans Arnout, Faes Matthias GR, Moens David. A machine learning approach for efficient and robust resistance spot welding monitoring. *Weld World* 2023;67(8):1923–35.
- [20] He Yicheng, Yang Kai, Wang Xiaoping, Huang Haisong, Chen Jiadui. Quality prediction and parameter optimisation of resistance spot welding using machine learning. *Appl Sci* 2022;12(19):9625.
- [21] Deng Haofeng, Gao Xiangdong, Lv Zigui, Tan Wenzheng, Gao Pengyu. Study of spot distance on resistance spot welding quality: a 1DCNN-BiLSTM-attention-based online inspection method. *Meas Sci Technol* 2024;36(1):015133.
- [22] Vo Quoc-Trinh, Van Tran Dat, Nguyen Rin Vy, Le Ngoc Anh, Truong Cong Doan, Tran Tin Trung. LSTM-based welding quality forecasting system in smart manufacturing. In: *Proceedings of the 2023 9th international conference on robotics and artificial intelligence*. 2023, p. 54–60.
- [23] Chang F, Zhou G, Ding K, Li J, Jing Y, Hui J, et al. A CNN-LSTM and attention-mechanism-based resistance spot welding quality online detection method for automotive bodies. *Math* 2023;11(22):4570.
- [24] Eisazadeh Hamid, Hamed Mohsen, Halvaei Ayob. New parametric study of nugget size in resistance spot welding process using finite element method. *Mater Des* 2010;31(1):149–57.
- [25] Wang Jin, Wang Hui-Ping, Lu Fenggui, Carlson Blair E, Sigler David R. Analysis of al-steel resistance spot welding process by developing a fully coupled multi-physics simulation model. *Int J Heat Mass Transfer* 2015;89:1061–72.
- [26] Zhou Baifan, Pychynski Tim, Reischl Markus, Mikut Ralf. Comparison of machine learning approaches for time-series-based quality monitoring of resistance spot welding (RSW). *Arch Data Sci Ser A (Online First)* 2018;5(1):13.
- [27] Xia Yu-Jun, Lv Tian-Le, Ghassemi-Armaki Hassan, Li Yong-Bing, Carlson Blair E. Collaborative simulation of nugget growth and process signals for resistance spot welding. *Weld World* 2023;67(6):1377–92.
- [28] Xia Yu-Jun, Lv Tian-Le, Ghassemi-Armaki Hassan, Li Yongbing, Carlson Blair E. Quantitative interpretation of dynamic resistance signal in resistance spot welding. *Weld J* 2023;102(4):69s–87s.
- [29] Pashazadeh Hamed, Gheisari Yousof, Hamed Mohsen. Statistical modeling and optimization of resistance spot welding process parameters using neural networks and multi-objective genetic algorithm. *J Intell Manuf* 2016;27:549–59.
- [30] Reyaz Md Saquib Bin, Sinha Amar Nath, Mehdi Husain, Murtaza Qasim. Effect of pulsed TIG welding parameters on the microstructural evolution and mechanical properties of dissimilar AA6061-T6 and AA7075-T6 weldments. *Arab J Sci Eng* 2024;49(8):10891–911.
- [31] Reyaz Md Saquib Bin, Sinha Amar Nath. Exploring the influence of pin geometry on metallurgical and mechanical characteristics of dissimilar aluminum TIG welds via friction stir processing. *J Adhes Sci Technol* 2024;38(17):3225–48.
- [32] Reyaz Md Saquib Bin, Sinha Amar Nath. Effect of heat input on microstructure and mechanical properties of automated tungsten inert gas-welded dissimilar AA6061-T6 and AA7075-T6 joints. *J Mater Eng Perform* 2025;34(1):103–20.
- [33] Reyaz Md Saquib Bin, Sinha Amar Nath. Improving tensile strength, impact toughness and wear resistance of TIG welded dissimilar AA7075-T6/AA6061-T6 joints through friction stir processing. *Int J Interact Des Manuf (IJIDeM)* 2024;18(3):1387–400.
- [34] Reyaz Md Saquib Bin, Sinha Amar Nath. Fabrication of tungsten inert gas-welded dissimilar aluminium joints using different friction stir processing tools through response surface method. *Adv Mater Process Technol* 2025;11(2):862–93.
- [35] Reyaz Md Saquib Bin, Sinha Amar Nath. Investigation on the influence of FSP parameters on TIG welded dissimilar aluminum joints using RSM and desirability function approach. *Proc Inst Mech Eng Part C: J Mech Eng Sci* 2024;238(14):7062–82.
- [36] Reyaz Md Saquib Bin, Sinha Amar Nath. Modeling the effects of tool pin configurations and friction stir processing parameters on tungsten inert gas welded dissimilar aluminum alloy joints. *J Mater Eng Perform* 2024;1–25.
- [37] Lv Zigui, Gao Xiangdong, Xiao Hong, Gao Pengyu. Resistance spot welding defect detection based on vectorized dynamic resistance signal and LightGBM classifier. *Meas Sci Technol* 2024;35(8):086113.
- [38] de Souza Sampaio Nilo Antonio, Caraschi José Cláudio, de Miranda Fabíola Freitas, Prates Gláucia Aparecida, do Lago Attadia Galli Lesley Carina, Bezerra Gabriel Albudane, et al. Using a full factorial experiment to optimize the performance of an industrial chemical process. *Environ Soc Manag J/ Rev de Gestão Soc E Ambient* 2025;19(4).
- [39] Spittel M, Spittel T. 4.3 thermal conductivity of steel: Introduction and definitions. *Met Form Data Ferr Alloys- Deform Behav* 2009;89–97.
- [40] Schwenk Christopher, Rethmeier Michael. Material properties for welding simulation-measurement, analysis, and exemplary data. *Weld J* 2011;90(6):220–7.
- [41] Yang Weile, Gao Perry P, Gao Xiangdong. Online evaluation of resistance spot welding quality and defect classification. *Meas Sci Technol* 2023;34(9):095016.
- [42] Ma Ningshu, Murakawa Hidekazu. Numerical and experimental study on nugget formation in resistance spot welding for high strength steel sheets in automobile bodies. *Trans JWRI* 2009;38(2):19–24.
- [43] Wang SC, Wei PS. Modeling dynamic electrical resistance during resistance spot welding. *J Heat Transf* 2001;123(3):576–85.
- [44] Wan Zixuan, Wang Hui-Ping, Wang Min, Carlson Blair E, Sigler David R. Numerical simulation of resistance spot welding of al to zinc-coated steel with improved representation of contact interactions. *Int J Heat Mass Transfer* 2016;101:749–63.
- [45] Geng Chen, Sheng Buyun, Fu Gaocai, Chen Xiangxiang, Zhao Guangde. A GAN-based method for diagnosing bodywork spot welding defects in response to small sample condition. *Appl Soft Comput* 2024;157:111544.
- [46] Echrigui Rania, Hamiche Mhamed. Optimizing LSTM models for EUR/USD prediction in the context of reducing energy consumption: an analysis of mean squared error, mean absolute error and R-squared. *E3S Web Conf* 2023;412:01069.
- [47] Chai Tianfeng, Draxler Roland R. Root mean square error (RMSE) or mean absolute error (MAE). *Geosci Model Dev Discuss* 2014;7(1):1525–34.
- [48] Hodson Timothy O. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci Model Dev Discuss* 2022;2022:1–10.
- [49] Chicco Davide, Warrens Matthijs J, Jurman Giuseppe. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput Sci* 2021;7:e623.
- [50] Antwarg Liat, Miller Ronnie Mindlin, Shapira Bracha, Rokach Lior. Explaining anomalies detected by autoencoders using Shapley additive explanations. *Expert Syst Appl* 2021;186:115736.
- [51] Ma Haiping, Zhang Yajing, Sun Shengyi, Liu Ting, Shan Yu. A comprehensive survey on NSGA-II for multi-objective optimization and applications. *Artif Intell Rev* 2023;56(12):15217–70.
- [52] Tran Khoa Duc. An improved non-dominated sorting genetic algorithm-II (ANSGA-II) with adaptable parameters. *Int J Intell Syst Technol Appl* 2009;7(4):347–69.
- [53] Santos Thiago, Xavier Sebastiao. A convergence indicator for multi-objective optimisation algorithms. *TEMA (São Carlos)* 2018;19:437–48.
- [54] Wang Xilu, Jin Yaochu, Schmitt Sebastian, Olhofer Markus. Recent advances in Bayesian optimization. *ACM Comput Surv* 2023;55(13):1–36.