

Article

Not peer-reviewed version

Optimization of Rotary Friction Welding Parameters Through AI-Augmented Digital Twin Systems

[Piotr Lacki](#)^{*}, [Janina Adamus](#), [Kuba Lachs](#), [Wiktor Lacki](#)

Posted Date: 27 March 2025

doi: 10.20944/preprints202503.2105.v1

Keywords: Rotary Friction Welding; Artificial Neural Networks; Digital Twin; Finite Element Method



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Optimization of Rotary Friction Welding parameters through AI-augmented Digital Twin systems

Piotr Lacki ^{1,*}, Janina Adamus ¹, Kuba Lachs ² and Wiktor Lacki ³

¹ Faculty of Civil Engineering, Czestochowa University of Technology; J.H. Dabrowskiego 69 Str.; 42-201 Czestochowa, Poland; janina.adamus@pcz.pl

² Faculty of Computer Science and Artificial Intelligence, Czestochowa University of Technology; J.H. Dabrowskiego 69 Str.; 42-201 Czestochowa, Poland; kublac24@gmail.com

³ Faculty of Mechanical Engineering, Czestochowa University of Technology; J.H. Dabrowskiego 69 Str.; 42-201 Czestochowa, Poland wiktor.lacki07@gmail.com

* Correspondence: piotr.lacki@pcz.pl

Abstract: In this study, Artificial Neural Networks (ANN) were employed to develop a Digital Twin (DT) of the Rotary Friction Welding (RFW) process. The neural network models were trained to predict the peak temperature generated during the welding process of dissimilar Ti Grade 2/AA 5005 joints. This prediction was based on a parametric numerical model of the RFW process constructed using the Finite Element Method (FEM) within the ADINA System software. Numerical simulations enabled a detailed analysis of the temperature distribution within the weldment. Accurate temperature predictions are essential for assessing the mechanical properties and microstructural integrity of the welded materials. Artificial Intelligence (AI) models, trained on historical data and real-time inputs, dynamically adjust critical process parameters—such as rotational speed, axial force, and friction time—to maintain optimal weld quality. A key advantage of employing AI-augmented DT systems in the RFW process is the ability to conduct real-time optimization and adaptive control. By integrating a Genetic Algorithm (GA) with a DT of the RFW process, the authors have developed an effective tool for analyzing a wide range of parameters to identify optimal welding conditions to improve joint quality, minimize defects, and maximize process efficiency.

Keywords: Rotary Friction Welding; Artificial Neural Networks; Digital Twin; Finite Element Method

1. Introduction

Joining different materials using conventional welding methods causes many problems due to their varying thermophysical properties and the formation of numerous intermetallic compounds, which impart brittleness to the welded joints. Additionally, most welding processes are affected by defects (such as porosity, shrinkage, hot cracking, and segregation), which form during solidification and reduce the quality of the joint [1]. Classical welding processes also pose physical hazards (excessive noise and dangerous thermal and optical radiation) as well as dust and gas pollution, which contains numerous hazardous substances that can cause air pollution and negatively impact human health and ecosystems.

To overcome the difficulties encountered in conventional welding, RFW, a solid-state joining process, can be used. In the RFW process, the heat required to join materials comes from the direct conversion of mechanical energy into thermal energy, resulting from friction in the contact area of the welded parts. This process avoids the melting of materials, thereby preventing solidification-related defects. Additionally, the lower temperature in the welding zone limits the formation of

brittle intermetallic compounds, enabling the joining of different materials with varying physical properties that cannot be joined using conventional welding methods.

Therefore, RFW is widely used in the aerospace, automotive, oil and gas, and medical industries, where it is utilized to join parts of various shapes and sizes, such as turbine shafts, automotive parts, electrical connectors, cutting tools, and medical implants. This process allows for the joining of both similar materials [2] and dissimilar materials, such as AA7075 and AA2024 aluminum alloys [3], or aluminum and titanium alloys [4], enabling the creation of lightweight and durable structures. The RFW process is characterized by high weld quality, often comparable to or better than the strength of the base material.

With the development of AI and digitalization in production, new perspectives for creating products with predictable properties are emerging. The rapid development of AI tools, including Genetic Algorithms (GA), Artificial Neural Networks (ANN), and other Machine Learning (ML) techniques, offers new possibilities for precise control of the RFW process. Digital Twin (DT) technology enables comprehensive data recording and analysis of welding process states, as well as full visualization of the entire process. These systems integrate ML algorithms with virtual replicas of the physical welding setup to predict and adjust key process variables such as rotational speed, axial pressure, and friction time, enhancing joint quality and mechanical properties. Advanced models like ANN and GA are used to analyze large datasets, uncover complex relationships between input parameters and weld outcomes, and propose optimal settings without the need for extensive physical trials. As stated in [5], AI techniques such as ANN and GA have been successfully applied to optimize welding parameters, achieving an accuracy level of 85-99% in predicting outcomes. Hybrid methods combining AI with traditional optimization strategies, like Response Surface Methodology (RSM), have been shown to enhance prediction accuracy and optimize process parameters effectively.

Simulation tools, such as ANSYS and SIMULINK, allow for the analysis of material behavior under various welding conditions, enabling the identification of optimal parameters like friction pressure and rotational speed [6]. The use of DT systems facilitates real-time monitoring and adjustment of welding parameters, further improving the quality and efficiency of the welding process [5]. Studies highlight how these DT systems improve efficiency by minimizing defects and reducing material waste while enabling adaptive control in response to varying material behaviors. This approach not only accelerates product development cycles but also ensures greater consistency and performance in critical applications like aerospace and automotive manufacturing.

Although there are more and more successful implementations in research and industrial work, some authors [7] point out that a full evaluation of these solutions is impossible due to the lack of detailed, publicly available information. Additionally, the issue of defining a DT also remains unresolved [8], even though the DT concept has been proposed for over 20 years and is one of the main ideas associated with Industry 4.0. These challenges highlight the ongoing complexity and evolving nature of DT technology. In [8], the definitions of the DT concept in scientific literature were analyzed, tracing its origins in aviation and space to its latest interpretations in production and smart manufacturing research.

DT offers immense potential in various domains of the product engineering process. As noted by the authors of paper [9], although current approaches to utilizing DT focus only on separate disciplines, it is anticipated that the holistic use of DT models in product development and production will dominate future generations of products, enabling the creation of high-performance products competitively. The authors of [10] note that the stringent requirements for durability, reliability, and efficiency, especially for future aircraft and spacecraft, will necessitate a change in the approach to their design, testing, and maintenance. The current approach—relying on statistical material properties, heuristic design principles, physical tests, and assumed similarity between test and operational conditions—is unlikely to meet extreme requirements. A paradigm shift is needed.

The introduction of DT will integrate high-fidelity simulations with the onboard vehicle health management system, maintenance history, and all available data to accurately reflect the lifespan of

its flying twin, ensuring an unprecedented level of safety and reliability. A similar transformation is occurring in the construction industry. While Building Information Modeling (BIM) provides procedures, technologies, and data schemas for the standardized semantic representation of building components and systems, the concept of DT offers a more holistic, socio-technical, and process-oriented characterization of complex artefacts. It utilizes the synchronization of cyber-physical bidirectional data flows, as noted in [11], paving the way for the concept of construction DT.

Paper [12] discusses how DT enhances intelligent automation in various fields. This paper outlines the concept, traces the evolution and development of DT, reviews key enabling technologies, examines trends and challenges, and explores its applications. While the integration of AI in RFW optimization has shown promising results, challenges remain in standardizing these techniques across different materials and welding scenarios. Further research is needed to fully explore AI's potential in improving the adaptability and precision of welding processes.

2. Materials and Methods

This study serves as an extension and continuation of the work presented in [1,13], focusing on the prediction and optimization of RFW parameters. The numerical calculations conducted in this research were performed using the ADINA System, which facilitates the analysis of heat generation due to friction within an axisymmetric thermomechanical model. Frictional resistance was modeled using a temperature dependent friction coefficient to accurately represent material behavior during the welding process. The adopted thermo-mechanical model required precise determination of the elastic-plastic properties of the analyzed materials as a function of temperature. Friction welding simulations were conducted for varying welding parameters and time durations to assess their effects on the process. The numerical simulations enabled the prediction of temperature distribution within the weld cross-section, which is a critical factor for subsequent analyses. The resulting temperature fields provide essential data for evaluating structural transformations, hardness variations, residual stress development, and material deformation. In this study, dissimilar joints of Ti Grade 2 and AA 5005 aluminum alloy, previously described in detail in [1], were utilized to construct the DTs models. AA 5005 belongs to the 5000 series aluminum alloys, where magnesium serves as the primary alloying element. This alloy is widely used as a construction material due to its relatively high strength, excellent formability, especially in soft temper conditions superior fatigue resistance, good weldability, and robust corrosion resistance. Commercially pure titanium Grade 2 is known for its excellent corrosion resistance, relatively good formability, and moderate mechanical strength. These properties make it a suitable candidate for various industrial applications, especially when paired with lightweight aluminum alloys in dissimilar metal joining processes. The integration of numerical modeling and DT technology in this work contributes to a deeper understanding of the thermal and mechanical behaviors during RFW, ultimately aiding in the optimization of welding parameters for improved joint performance.

The effectiveness of the DT depends on incorporating as many RFW process parameters as possible into the model. **Błąd! Nie można odnaleźć źródła odwołania.** shows the key parameters of the RFW process included in the DT.

The maximum temperature at the faying surfaces during the RFW process is influenced by several key process parameters. These factors determine the amount of frictional heat generated, the efficiency of heat transfer, and the thermal and mechanical behaviors of the materials being joined. Previous studies [1,13] indicate that higher rotational speeds increase frictional heat generation due to the greater rate of sliding contact. This typically results in higher maximum temperatures at the interface. The maximum temperature generally rises with rotational speed, though an optimal range exists depending on the material properties. Higher axial forces increase frictional heat generation by enhancing frictional resistance. However, excessive axial force may lower the temperature by accelerating material deformation and shortening the frictional sliding phase. While the maximum temperature initially rises with axial force, it can decrease at very high forces due to excessive material flow. Friction time determines how long heat is generated at the interface. A longer friction time

allows the temperature to build up gradually, reaching higher maximum values. The maximum temperature increases with friction time until the system reaches thermal equilibrium.

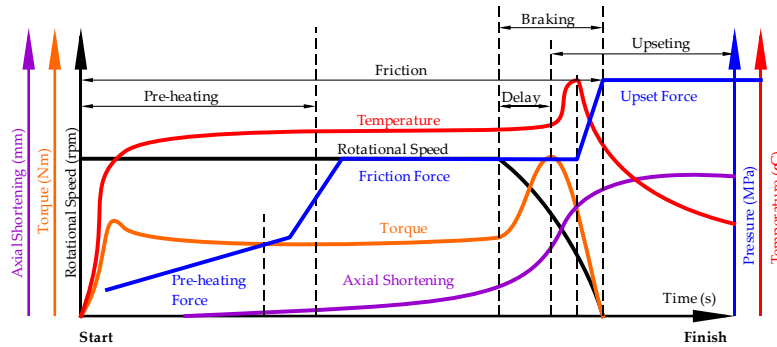


Figure 1. RFW process parameters considered in the DT framework.

Commented [M1]: Figure 1 is not cited, please check.

The interaction of these parameters is often nonlinear, and the optimal settings depend on factors such as the material pair, component geometry, and the desired joint quality. High rotational speed, combined with moderate axial force and friction time, is generally optimal for achieving sufficient heat and plastic deformation. However, excessive axial force or excessively long friction times may lead to material defects, such as voids or cracking, caused by overheating or improper material flow. In this study, numerical models and simulations, including those based on the Finite Element Method (FEM), were employed to investigate the combined effects of these parameters. Furthermore, the use of AI-augmented Digital Twin (DT) systems enhances the ability to predict and optimize maximum temperatures and welding outcomes across varying conditions.

2.1. DT Based on MES

The Finite Element Method (FEM) plays a crucial role in the creation and functionality of Digital Twin (DT) systems, particularly in the case of complex physical systems and structures. FEM enables the simulation and analysis of how physical objects respond to various forces, stresses, and environmental conditions, thereby enhancing the accuracy and predictive capabilities of DT. FEM also enables stress, fatigue, and vibration analysis, allowing DT to predict potential failures. This predictive capability supports proactive maintenance strategies, reducing downtime and extending equipment life. Advanced FEM models, integrated with real-time sensor data, enable DT to adjust simulations based on current operating conditions. This dynamic updating improves performance predictions and decision-making. FEM-driven DT systems enable virtual testing of different design variations without the need for costly physical prototypes. Engineers can optimize designs for weight reduction, structural integrity, and performance by running multiple FEM simulations. Furthermore, FEM allows DT to simulate extreme operating conditions, such as high loads, impacts, or thermal shocks, ensuring safety validation and compliance with standards and regulations.

FEM results can be used to train machine learning (ML) algorithms within Digital Twin (DT) systems, enhancing predictive accuracy. Figure 2 presents the developed parametric numerical FEM model, which is used to train the artificial neural network (ANN). The model accounts for a wide range of variable parameters, particularly for the Ti Grade 2/AA 5005 friction pair. It is proposed to adopt a temperature-dependent friction coefficient, as described in detail in [1]. The friction coefficient is a material property that determines the efficiency of converting mechanical work into heat. A higher friction coefficient leads to more effective heat generation, thereby increasing the peak temperature. The friction coefficient may vary with temperature, material condition, and surface cleanliness. Materials with higher friction coefficients typically achieve higher peak temperatures during RFW.

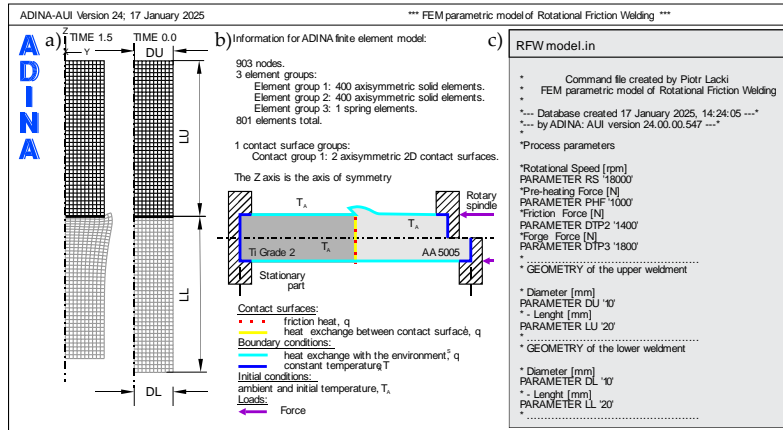


Figure 2. Parametric FEM model; (a) finite element mesh; (b) boundary and initial conditions; (c) part of the parametric model file.

Materials with low thermal conductivity (Ti Grade 2) retain heat at the interface, leading to higher maximum temperatures. In contrast, materials with high thermal conductivity (AA 5005) dissipate heat more rapidly, reducing peak temperatures. Materials with higher heat capacity require more heat input to achieve the same temperature rise, which can lower the maximum temperature for a given set of parameters. Materials with lower flow stress at elevated temperatures deform more easily, potentially reducing frictional heating as material deformation dominates over sliding. The geometry and surface roughness of the components influence the initial contact conditions, which, in turn, affect heat generation. Rougher surfaces may increase initial friction, resulting in higher temperatures, while smoother surfaces may reduce heat generation at the interface. Heat dissipation through the surrounding material and environment can limit the maximum temperature. Materials with higher thermal conductivity or external cooling systems dissipate heat more efficiently, reducing peak temperatures. By combining FEM with AI, DT can predict complex behaviors without the need for full-scale FEM simulations each time..

2.2. DT Based on ANN

DT uses ANN to learn relationships between inputs (friction force and rotational speed) and outputs (temperature) based on historical and real-time data. This enables rapid simulations without the need to solve complex physical equations. ANN can work alongside traditional models (e.g., FEM) to speed up simulations or fill in gaps where physical models are incomplete. Hybrid models combine data-driven learning with physics-based constraints to create more accurate and efficient simulations. In the work [14], an ANN-based incremental learning system was developed to predict crack growth and fatigue in the RFW joints of AA5083 and AA7075 aluminum alloys. The results demonstrate that this method can effectively accommodate input temperatures and the S-N curve, providing reliable predictions of expected fatigue. This approach can reduce labor costs and time spent on crack propagation tests, thereby enhancing production processes and reducing overall costs. Furthermore, it shows promise for extrapolating monotonic functions, paving the way for solving various predictive problems. Additionally, verification using AISI 2205 and AISI 1020 steel showed that neural networks could obtain S-N curve values for other metals with an error rate of less than 8%.

The detailed scope of the numerical data used to analyze the ANN is presented in Table A1 of Appendix A. Figure 3 graphically presents the data grid, divided into three sets: (a) Training Set, (b) Validation Set, and (c) Test Set. Before training the neural network, a representative dataset was

generated that accurately reflects the problem domain. The dataset from the FEM calculation contained 105 cases, divided into three subsets: training set (47%), validation set (36%), and test set (20%). This dataset size was optimal for network performance. The data also included scenarios that, while not practically relevant, were still useful for the analysis. All features contribute equally to the task. Care was taken to ensure that each data sample in each set had the same probability of occurrence.

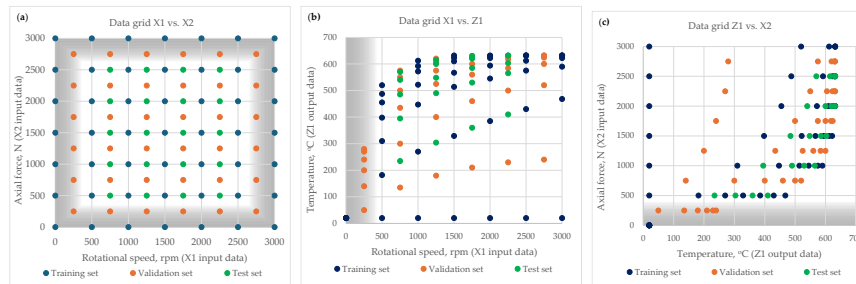


Figure 3. Scatter plot of the data grid for three subsets in different configurations; (a) rotational speed vs. axial force; (b) rotational speed vs. temperature; (c) temperature vs. axial force.

Each subset plays a crucial role in ensuring effective learning and generalization to unseen data. The data obtained from the FEM model are consistent, and no augmentation, dropout, or weight decay techniques were applied. Preparing a high-quality training set is essential for ensuring that the neural network learns effectively and generalizes well. The RFW process parameters that are not used in the analyzed case but are theoretically possible to use, and were calculated by FEM, are marked in gray in Figure 3.

Designing a neural network topology involves choosing the number of hidden layers, the number of neurons per layer, and the activation functions based on the problem's complexity. In general, the selection of the topology depends on the nature of the data and the target function. In the analyzed case, note that the network comprises two input neurons and one output neuron, with nonlinear dependencies between the input and output data. Therefore, the problem cannot be solved using a simple linear model; it requires nonlinear transformations. The input data interact in a complex manner that necessitates at least one hidden layer. There is no fixed formula for determining the number of layers or the number of neurons per hidden layer. The practical approach used in this work assumes that a single hidden layer (a 2-H-1 configuration) is often sufficient for most nonlinear problems, although two hidden layers (a 2-H1-H2-1 configuration) may be required for more complex relationships. To determine the starting number of neurons, an empirical formula was employed:

$$H = \frac{N_{in} + N_{out}}{2} + k, \quad (1)$$

where H is the number of hidden neurons, N_{in} is the number of inputs (2), N_{out} is the number of outputs (1), and k is a small integer (from 2 to 10). Example: $H = (2+1)/2 + 7 = 8$, so we start with 8 neurons in the hidden layer.

The initial 2-8-2 topology was expanded by increasing the hidden layer to 16 neurons ($H=16$), and then a second hidden layer with 16 neurons was added. Figure 4 presents the ANN topologies used in the work.

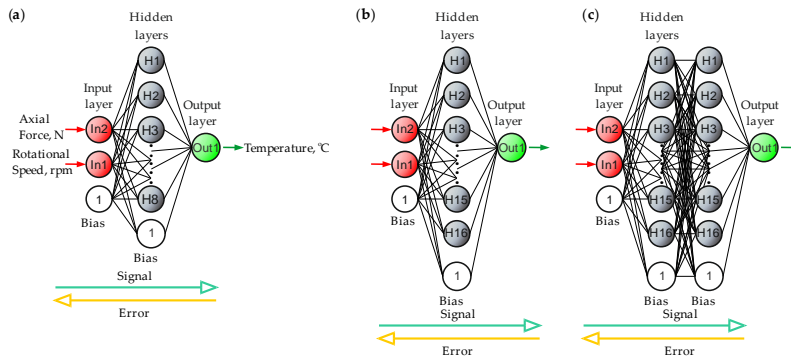


Figure 4. Applied ANN topologies; (a) topology: 2-8-1; (b) topology 2-16-1; (c) topology 2-16-16-1.

Activation functions introduce nonlinearity into neural networks, enabling them to model complex relationships in data. In this work, three types of activation functions are used: Linear, Hyperbolic Tangent (Tanh), and Sigmoid (or Logistic). The Linear activation function is applied at the output of the input layer to evenly distribute the signal to all neurons in the first hidden layer. It is the simplest and most computationally efficient, as its output is directly proportional to its input:

$$f(x) = x, \quad (2)$$

Range: $(-\infty, +\infty)$; Derivative: $f'(x)=1$ (constant)

The Tanh activation function is a scaled version of the Sigmoid function that outputs values in the range $(-1, 1)$. It is characterized by a zero-centered output, which makes training more stable than using the Sigmoid function. The Tanh function provides better nonlinearity than a linear activation and works well for hidden layers in neural networks, especially for problems where the data are centered around zero. However, due to its shape, it is computationally more expensive than the linear activation function:

$$f(x) = \frac{2}{1+e^{-2x}} - 1, \quad (3)$$

Range: $(-1, 1)$; Derivative: $f'(x)=1-\tanh^2(x)$

The Sigmoid function is a nonlinear function that squashes input values to the range $(0, 1)$. It is well-suited for binary classification problems because its output can be interpreted as a probability. While it is differentiable and smooth, for very large or very small input values, the gradients approach zero, which slows down training:

$$f(x) = \frac{1}{1+e^{-x}}, \quad (4)$$

Range: $(0, 1)$; Derivative: $f'(x)=f(x)(1-f(x))$

The learning rate α is a crucial hyperparameter in neural networks that controls the step size at which the model updates its weights during training. It determines how much the model learns in each iteration and significantly affects convergence speed, performance, and stability. In this work, depending on the network model, the learning rate was set in the range from $\alpha = 0.1$ to 0.0001 . An optimal learning coefficient was selected to balance speed and stability while enabling efficient convergence without overshooting. If α is too high, weight updates become too large and the model may never converge; conversely, if α is too low, training becomes very slow and may get stuck in local minima. In this work, one full pass through the dataset (i.e., one epoch) ranged from 100 to 10,000 iterations, depending on the network model, while the batch size was set to 1.

Statistical metrics describing individual ANN cases and the hyperparameters adjusted to obtain the optimal ANN are presented in Table 1.

Table 1. ANN model architecture for selected hyperparameters and associated statistical metrics.

ID	Topology	Activation functions	Learning rate	Epoch (All/Best)	MSE*	R ²	Training time, s
1	2-8-1	Linear-Linear-Linear	0.1	100/61	0.157/0.175/0.157	0.6	3.94
2	2-16-1	Linear-Linear-Sigmoid	0.01	100/57	0.135/0.147/0.136	0.6848	6.55
3	2-16-1	Linear-Tanh-Sigmoid	0.01	100/100	0.125/0.137/0.130	0.7329	6.61
4	2-16-1	Linear-Tanh-Sigmoid	0.01	500/485	0.071/0.057/0.086	0.9129	33.42
5	2-16-16-1	Linear-Tanh-Tanh-Sigmoid	0.01	500/210	0.055/0.065/0.063	0.9732	89.41
6	2-16-16-1	Linear-Tanh-Tanh-Sigmoid	0.001	1000/1000	0.054/0.056/0.055	0.947	178.87
7	2-16-16-1	Linear-Tanh-Tanh-Sigmoid	0.001	5000/2038	0.045/0.045/0.045	0.9661	946.12
8	2-16-16-1	Linear-Tanh-Tanh-Sigmoid	0.0001	10000/10000	0.041/0.055/0.079	0.9314	1827.32

* MSE for training set / validation set / test set.

In the applied neural network model, Mean Squared Error (MSE) was used as the loss function. MSE measures the average squared difference between the actual and predicted values and helps train the network by minimizing errors via gradient descent. A lower MSE indicates more accurate predictions, whereas a higher MSE suggests larger errors. Since MSE squares the errors, it penalizes large deviations more than small ones, making it well-suited for minimizing errors during ANN training. In addition, R-squared (R²) was used to evaluate the model's performance after training. R² is a statistical metric that assesses how well the model's predictions match the actual values by measuring the proportion of variance in the dependent variable explained by the independent variables. Unlike MSE, which measures absolute error, R² indicates how well the model generalizes, facilitating comparisons between different models. Both MSE and R² were analyzed together to ensure a well-performing regression model.

The last column of Table 1 provides the computation times for individual ANNs. The training time of an ANN depends on multiple factors, including the number of epochs, batch size, network complexity, dataset size, and available computational power. The number of epochs used (ranging from 100 to 10,000) was sufficient to ensure learning while avoiding overfitting. Early stopping was employed to determine the optimal training point. Batch size affects training stability and speed; smaller batch sizes lead to more frequent updates, whereas larger batch sizes require more memory. A fixed batch size of 1 was used. Network complexity, defined by the number of layers and neurons, increases training time but enhances model capacity, necessitating techniques such as dropout and batch normalization for efficiency. Large datasets and high-dimensional data require more processing but improve generalization, often benefiting from dimensionality reduction methods. Finally, computational resources such as GPUs and TPUs significantly accelerate training compared to CPUs.

2.3. GA-Based Optimization

GA are powerful optimization tools inspired by the principles of natural selection and evolution. They are well-suited for solving complex, nonlinear, and multi-objective optimization problems, such as those encountered in the RFW process. By integrating GA with the digital twin (DT) of the RFW process, it is possible to efficiently explore a large parameter space and identify optimal welding conditions to improve joint quality, minimize defects and maximize process efficiency. GA solves multi-criteria problems by balancing goals such as cost, performance, energy efficiency, and durability. DTs use GAs to find trade-offs between conflicting objectives. GAs assist DTs in calibrating system models by automatically tuning parameters to match real-world behavior,

resulting in more accurate and realistic simulations. They allow DTs to continuously evolve and adapt their models over time as new data are collected, thereby improving decision-making and system resilience. GAs can find global optima in complex, nonlinear, and multi-modal problem spaces. They handle both discrete and continuous variables, evaluating multiple solutions simultaneously to speed up optimization. Unlike gradient-based methods, GAs do not require differentiable models, making them ideal for black-box optimization. By integrating GAs, DTs become highly adaptive and capable of discovering innovative solutions that drive efficiency, performance, and resilience across industries.

In [15], a GA was used to optimize the parameters of the RFW process for dissimilar light alloys (AZ31B/AA7075). The optimized values were compared with experimental results, and the GA method increased the tensile strength of the joint from 88 to 180 MPa. A similar approach to optimizing the parameters of the RFW process was presented in [16]. Parameters such as spindle speed, friction pressure, upset pressure, and burn-off length were optimized to achieve maximum Ultimate Tensile Strength (UTS) and hardness for the aluminum metal-matrix composite AA7075-10 vol% SiCP-T6 using GA with a Combined Objective Function (COF). The main advantage of using GA is that it converges to optimized values for various weighting factors in a minimal number of iterations, making it suitable for solving problems with a large search space, such as friction welding. The strong correlation between the optimized and validated results (in terms of tensile strength and hardness) demonstrates the method's accuracy in achieving the objective.

The first step in applying genetic algorithms to RFW optimization is to formulate an Fitness Function (F) that quantifies the quality of the weld. The proposed optimization function includes parameters such as: maximizing joint strength (e.g., tensile strength, fatigue resistance); minimizing defects (e.g., voids, cracks, excessive flow); optimizing temperature distribution to prevent unwanted intermetallic formation in dissimilar welding; minimizing residual stresses and deformations; and reducing cost or energy consumption.

A weighted sum function was used to optimize the RFW process, combining several qualitative or quantitative criteria into a single F with appropriate weighting factors. An F with penalties was added to account for additional constraints (e.g., maximum temperature, minimum strength, limited intermetallic layer thickness). In this approach, the F is extended with components that "penalize" solutions violating these constraints. The F proposed for the optimization of the RFW process can be mathematically expressed as:

$$F = \sum_{i=1}^n \omega_i \cdot f_i(x) + \sum_{l=1}^m \lambda_l \cdot \max\{0, g_l(x)\} \quad (5)$$

where:

ω_i - weighting coefficients defining the relative importance of individual criteria,

$f_i(x)$ - individual objective functions (e.g. maximizing strength, minimizing defects, minimizing temperature differences, etc.),

λ_l - penalty coefficients that determine how much violating a constraint affects the value of the objective function,

$g_l(x)$ - constraint functions (may take positive values when the constraint is violated).

In the adopted F, each criterion is normalized to a common scale to ensure that their impacts are properly compared. The F value is calculated as the sum of the products of individual criteria and their weights. Solutions that satisfy all constraints receive no penalties, so the F value equals $f(x)$. If any constraint is violated, a penalty is added to the objective function, which deteriorates the evaluation of the solution. Optimization involves finding a set of parameters that minimizes or maximizes the function F—depending on the definition—while meeting technical and qualitative criteria. The selection of appropriate weights is subjective and requires preliminary research or experimentation. In some cases, the function may not fully reflect the trade-offs between criteria, particularly when dealing with strongly divergent scales or conflicting goals.

GA requires encoding the RFW parameters into a chromosome-like structure. The following coding scheme was adopted:

$$\text{Chromosome}=[N, F_A, t_f] \quad (6)$$

where: N = Rotational speed, rpm; F_A = Axial force, N; t_f = Friction time, s.

A population (PS) of RFW parameter sets is randomly generated, with each individual representing a different combination of parameters. Each individual is evaluated using the objective function, and the Fitness Score (FS) indicates how well its parameters fit the F. Based on the FS, the best-performing individuals are selected using a Selection Method (SM) technique. In this paper, a tournament selection strategy is employed; the Tournament Size (TS) specifies the number of individuals randomly selected to compete for selection, with larger sizes increasing selection pressure. To explore new solutions, the selected individuals undergo crossover, during which parts of their parameter sets are swapped to produce offspring with mixed traits. The Crossover Rate (CR) indicates the probability that two parental chromosomes will undergo crossover to produce offspring; a higher CR promotes the exploration of new solutions. To introduce diversity and avoid premature convergence, a Mutation Rate (MR) is applied to randomly alter some parameters. The MR represents the probability that individual genes (parameters) will be changed during mutation, which is crucial for maintaining diversity and avoiding local minima, thereby ensuring that the algorithm explores a wider search space. The Number of Generations (NG) denotes the total number of iterations the GA will perform to develop solutions. The process repeats over multiple generations, refining the parameters to reach an optimal solution. Convergence occurs when the fitness function stabilizes, meaning that further iterations produce no significant improvements. The Elitism Rate (ER) indicates the proportion of the best individuals preserved into the next generation without modification, helping to retain high-quality solutions across generations. The Termination Criteria (TC) specify the conditions under which the GA stops running (e.g., reaching the maximum number of generations or a convergence threshold), thereby preventing unnecessary computation once an optimal solution is reached. Finally, the Encoding Scheme (EC) employs a real-valued representation of decision variables, which is consistent with the nature of the problem parameters.

3. Results

The microstructure of dissimilar joints between Ti Grade 2 and AA 5005 aluminum alloy, obtained through RFW, reveals distinct regions with characteristic features depending on the magnification level. Observing the welded interface at different magnifications, as shown in Figure 5, facilitates the analysis of metallurgical changes, intermetallic phase formation, and material flow behavior. Microstructure images were captured for a sample joint produced at a rotational speed of $n = 1440$ rpm and an axial force of $F_A = 1000$ N. However, the following discussion applies to the entire scope of the study.

At low magnification Figure 5 (a), the microstructure provides a macroscopic view of the welded joint, highlighting the overall morphology of the interface between Ti Grade 2 and AA 5005. The boundary between titanium and aluminum is clearly visible, with its distinct characteristics influenced by the process parameters. Flash formation due to excessive material flow may also be evident, particularly on the aluminum side, given its lower strength and higher deformability. The heat-affected zone (HAZ) appears as a transition region near the interface, with the aluminum side typically showing a larger HAZ due to its higher thermal conductivity and lower melting point. In contrast, titanium exhibits minimal deformation due to its higher strength and lower thermal conductivity, resulting in a narrower HAZ.

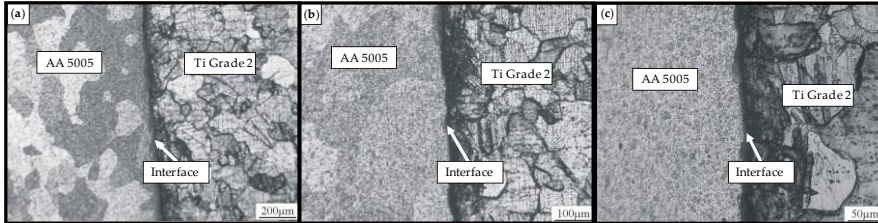


Figure 5. Welded interface at different magnifications; (a) low magnification; (b) medium magnification, (c) high magnification.

At medium magnification Figure 5 (b), the microstructure reveals more detailed features of the interface and adjacent regions. The titanium side near the interface retains a refined, equiaxed grain structure, while its small HAZ shows slight grain elongation or refinement depending on the distance from the weld interface. In the aluminum alloy, significant grain refinement occurs due to dynamic recrystallization, leading to smaller and more uniform grains in the HAZ compared to the base material. Evidence of plastic flow is visible as elongated grains near the interface, indicating material mixing and deformation during welding. A thin intermetallic compound (IMC) layer may also be present at the interface, formed due to high temperatures and chemical interaction between titanium and aluminum. Ensuring controlled plastic deformation and uniform material flow, as well as managing heat generation along the radial direction, helps minimize harmful compounds and improve the uniformity of the joint's mechanical properties. The formation of IMCs at the interface is primarily influenced by welding parameters such as friction load, burn length, and upsetting load [1].

At high magnification Figure 5 (c), finer details of the interface and intermetallic compounds become visible. IMC thickness increases with higher temperatures or prolonged friction time, potentially leading to joint embrittlement if excessive. Any microvoids, cracks, or discontinuities at the interface can be observed, often resulting from improper welding parameters or uneven material flow. The presence of oxide layers or impurities may also be detected, particularly if pre-welding surface treatment was inadequate. Small regions of aluminum may become embedded in the titanium surface due to high-pressure contact and deformation, while titanium fragments may appear in the aluminum matrix, forming a mixed zone near the interface. Grain refinement on both sides contributes to improved mechanical properties, particularly tensile strength and fatigue resistance. However, defects such as microvoids, cracks, or excessive IMC formation can compromise joint performance, necessitating process optimization to mitigate these issues. This multiscale analysis underscores the importance of carefully controlling RFW parameters to optimize microstructure and ensure a robust joint between Ti Grade 2 and AA 5005.

3.1. FEM Results

The distribution of plastic strain in the weldment volume during RFW process of a Ti Grade 2 rod and an AA 5005 aluminum alloy rod depends on the axial force F_A applied during the process. Higher axial forces lead to greater plastic deformation due to increased contact pressure, material flow, and heat generation. Figure 6 presents a quantitative analysis of plastic strain distribution for different axial force values ($F_A = 500$ N, 1000 N, and 2000 N).

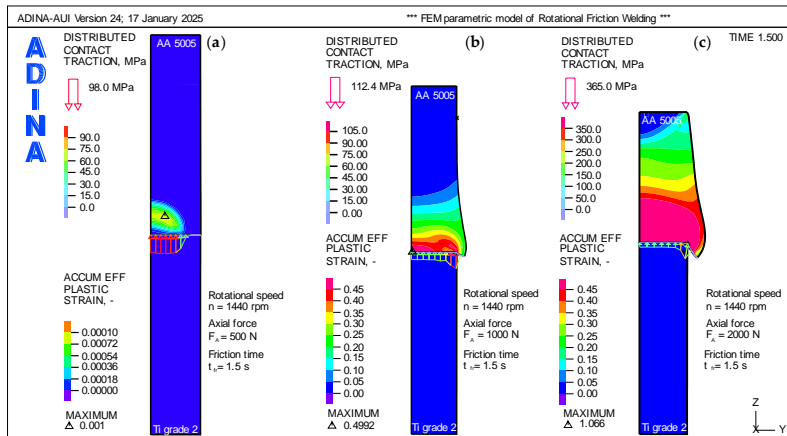


Figure 6. Plastic strains as a function of axial forces, N; (a) $F_A=500$ N; (b) $F_A=1000$ N; (c) $F_A=2000$ N.

At a low axial force Figure 6 (a); $F_A=500$ N, frictional heat generation is relatively low, resulting in limited thermal softening and material flow. The maximum plastic strain ($\epsilon_{\max} = 0.001$) is localized near the interface on the aluminum alloy (AA 5005) side, due to its lower strength and higher plasticity compared to titanium. Titanium exhibits minimal deformation due to its high yield strength and stiffness. The HAZ is small, with minimal material flow or deformation outside the immediate interface region. The aluminum side exhibits small plastic strains near the interface, whereas the titanium side remains largely unaffected. Weak material flow on the aluminum side results in less pronounced flash formation, which may lead to an incomplete joint.

At an optimal axial force Figure 6 (b); $F_A=1000$ N the plastic strain at the interface becomes more pronounced ($\epsilon_{\max} = 0.4992$), and the aluminum alloy undergoes significant plastic flow due to increased pressure and heat generation. Titanium begins to exhibit slight plastic deformation near the interface, primarily due to higher temperatures and increased contact pressure. Material flow significantly improves on the aluminum side, leading to noticeable flash formation. However, titanium still exhibits limited material flow due to its higher stiffness and lower thermal conductivity.

At a high axial force Figure 6 (c); $F_A=2000$ N plastic strain increases substantially ($\epsilon_{\max} = 1.066$) on the aluminum side. The aluminum alloy undergoes severe plastic flow, with material being extruded as flash. However, the titanium side does not undergo noticeable plastic strain near the interface despite the increased heat and pressure. Excessive axial force may lead to overheating and the formation of thick intermetallic layers (e.g., $TiAl_3$) at the interface, which can reduce joint strength due to brittleness.

The temperature distribution in the weldment volume during the RFW process of a Ti Grade 2 rod and an AA 5005 aluminum alloy rod depends on frictional heat generation at the faying surfaces, the thermal conductivity of the materials, and heat dissipation into the surrounding volume. Figure 7 illustrates how the temperature distribution varies at different rotational speeds.

At low rotational speeds Figure 7 (a); $n = 500$ rpm, the relative motion at the interface generates limited frictional heat. The peak temperature ($T_{\max} = 312$ °C) at the interface remains relatively low — insufficient to fully soften or plasticize the materials. Heat is concentrated at the immediate interface, with a steep temperature gradient extending into both materials. The aluminum alloy (AA 5005) side experiences a greater temperature rise than the titanium (Ti Grade 2) side, due to aluminum's lower thermal conductivity. The HAZ is narrow, and insufficient thermal energy may lead to weak bonding.

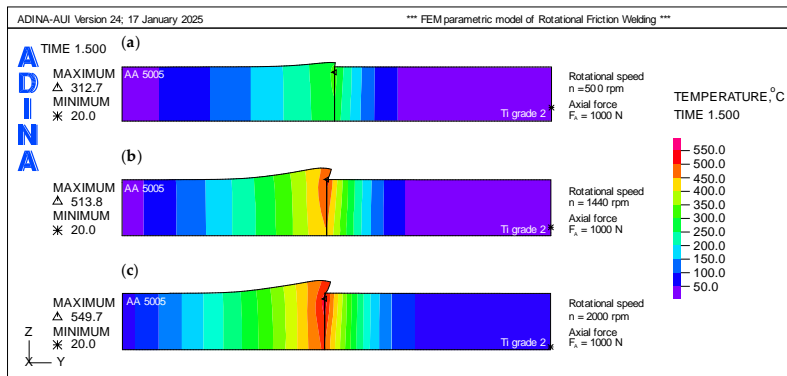


Figure 7. Temperature, °C as a function of rotational speed, rpm; (a) $n = 500$ rpm; (b) $n = 1440$ rpm; (c) $n = 2000$ rpm.

At optimal rotational speeds Figure 7 (b); $n=1440$ rpm, sufficient frictional heat is generated, raising the interface peak temperature to $T_{\max} = 513$ °C. The aluminum alloy softens and may approach its melting range, while titanium remains solid but thermally softened. The interface reaches higher temperatures, promoting material flow and enhanced mixing of the dissimilar metals. Due to its lower thermal conductivity, the aluminum alloy side experiences a greater temperature gradient and deeper heat penetration. In contrast, the titanium side undergoes a moderate temperature increase, but its thermal gradient remains steeper compared to aluminum. The heat-affected zone (HAZ) is larger than at lower speeds, leading to improved joint quality.

At high rotational speeds Figure 7 (c); $n=2000$ rpm) frictional heat generation becomes intense, raising the interface peak temperature to $T_{\max} = 549$ °C. The aluminum alloy may reach or exceed its melting point, while titanium remains in a solid state due to its higher melting temperature. However, excessive heat at high speeds may lead to undesirable effects, such as the formation of brittle intermetallic compounds (e.g., $TiAl_3$) or defects resulting from material over-softening on the aluminum side.

3.2. ANN Results

The result of an ANN is a set of predicted outputs generated from input data after training. The accuracy and reliability of these predictions depend on the quality of the training data, network architecture, and optimization process. When trained correctly, the ANN approximates complex relationships within the data, producing outputs that closely match expected values. The ANN's performance is typically evaluated using statistical metrics such as MSE and R^2 , which quantify the difference between predicted and actual values.

As part of the analysis, eight cases (ID 1–8) were examined, as shown in Table 1. The individual networks differ primarily in topology and the number of neurons, while activation functions also play a crucial role. A linear activation function was applied at the output of each input layer. The ID 1–8 networks exhibit a progressive increase in complexity. Beginning with ID 1, each subsequent network demonstrates higher accuracy. The networks evolve from a simple 2-8-1 structure to more complex 2-16-1 and 2-16-16-1 architectures, highlighting the trade-off between complexity, performance, and training time.

Figure 8 presents three plots of MSE versus epochs for representative ANN models. Model (a), ID 1 (2-8-1, Linear-Linear-Linear, $LR = 0.1$, 100 epochs), shows rapid error reduction in the initial epochs but exhibits significant fluctuations, particularly in validation MSE. While training and test MSE remain relatively stable, the validation MSE oscillates, indicating poor generalization. The model stops early (epoch 61), suggesting early plateauing or overfitting. The linear activation

functions restrict complexity, making the model less effective at capturing intricate patterns. Although the model learns quickly, it lacks expressive power, resulting in greater error variance.

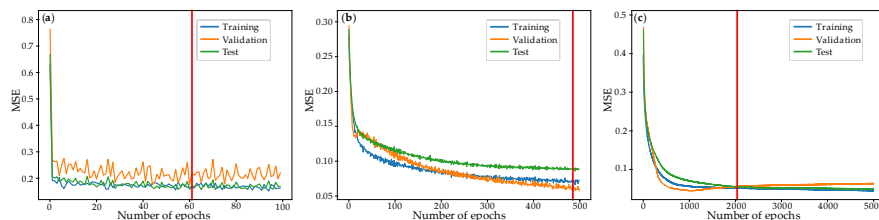


Figure 8. Variability of MSE for different ANN models as a function of the number of epochs; (a) ID 1; (b) ID 4; (c) ID 7.

The simplest model utilizes linear activation functions throughout and a high learning rate of 0.1, achieving moderate performance with $MSE = 0.157$ and $R^2 = 0.6$ in under four seconds. As the architecture expands to a 2-16-1 configuration, and activation functions are adjusted—initially using a Sigmoid output, then incorporating a Tanh hidden layer followed by a Sigmoid output—the network achieves higher accuracy with reduced MSE and increased R^2 values, albeit with longer training times. The transition to a deeper 2-16-16-1 network, incorporating two hidden layers with Tanh activation functions and a Sigmoid output, further enhances predictive capability, achieving lower error metrics and higher R^2 , particularly when the learning rate is lowered to 0.001 or 0.0001 and epochs are increased. However, this improvement comes at the cost of significantly longer training times. Overall, the transition from a simple to a more complex network illustrates how increasing neurons and layers, fine-tuning activation functions, and adjusting learning rates and epochs enhance the approximation of FEM data, while also highlighting the inherent trade-offs between accuracy and computational efficiency.

In case (b), ID 4 (2-16-1, Linear-Tanh-Sigmoid, $LR = 0.01$, 500 epochs), the MSE decreases smoothly during the initial epochs, indicating stable learning. Using the Tanh activation function in the hidden layer enhances nonlinearity, enabling better representation of complex relationships. Training proceeds until epoch 485, demonstrating continuous improvement with additional epochs. The validation and test MSE remain close to the training MSE, suggesting a good balance between learning and generalization. This network is better optimized than the previous one, exhibiting improved nonlinearity, lower MSE, and stable training.

In the best case (c), ID 7 (2-16-16-1, Linear-Tanh-Tanh-Sigmoid, $LR = 0.001$, 5000 epochs), the MSE gradually decreases due to the small learning rate (0.001). Training continues until epoch 2038, after which the validation/test MSE starts to increase slightly, indicating potential overfitting. The deeper architecture with two hidden layers utilizing Tanh activation functions enhances feature extraction, resulting in the lowest MSE. The longer training time is a result of the lower learning rate and deeper network structure. While the deep network achieves high accuracy, it requires significantly more epochs, and overtraining may lead to overfitting.

The scatter plots in Figure 9 compare FEM data with ANN predictions for different network architectures. The ANN model's quality is determined by how closely its predicted values align with FEM data along the red regression line and its corresponding R^2 value. For the shallow ID 1 network, $R^2 = 0.6$. The points deviate significantly from the regression line, particularly for low and midrange prediction temperatures. The wide spread indicates higher prediction errors. The linear activation functions limit the model's ability to capture complex relationships, reducing performance. This model demonstrates moderate accuracy but struggles with nonlinearity, leading to noticeable deviations.

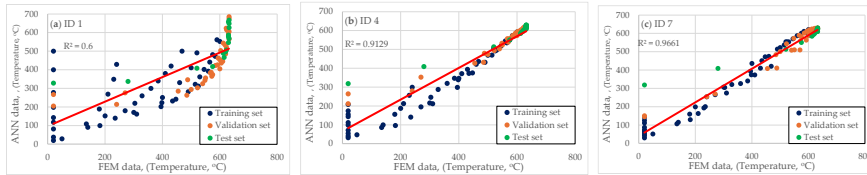


Figure 9. Scatter plots of FEM vs. ANN data for the network architectures; (a) ID 1; (b) ID 4; (c) ID 7.

The network with architecture ID 4 has an R^2 value of 0.9129. Its alignment with the regression line is significantly improved, exhibiting lower scatter and variance. The Tanh activation function enhances the model's ability to capture nonlinear patterns, improving accuracy compared to ID 1. However, some deviations persist for midrange temperature values, suggesting slight underfitting. This model achieves higher accuracy, demonstrating a more effective mapping between FEM and ANN predictions.

The deep network ID 7 achieves the highest R^2 value of 0.9661 among the analyzed architectures. The data points are positioned closer to the regression line than in ID 4, indicating high accuracy. Minimal scatter indicates strong generalization across different datasets (training, validation, and test). The increased network depth and lower learning rate enable more precise predictions. This model provides the best fit by effectively capturing FEM data patterns and generating accurate predictions.

The analysis of temperature distribution predictions for the examined network architectures is best illustrated through contour plots. Figure 10 presents contour plots depicting the temperature distribution as a function of axial force, N and rotational speed, rpm for three ANN models. The accuracy of these predictions depends on network depth, activation functions, and training quality.

For the shallow ID 1 network, the temperature distribution appears oversimplified, exhibiting linear gradient patterns. The temperature isolines appear as straight lines, indicating poor generalization. Unrealistic temperature values appear in the color bar, ranging from $-150^{\circ}C$ to $900^{\circ}C$, significantly deviating from expected physical trends. This model fails to capture nonlinear relationships, resulting in unrealistic and unreliable predictions.

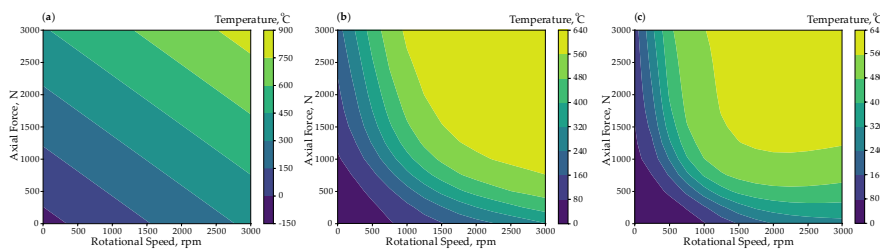


Figure 10. Temperature distribution prediction, $^{\circ}C$ by ANN; (a) ID 1; (b) ID 4; (c) ID 7.

For the ID 4 network, the temperature distribution appears more structured, with smooth and continuous transitions. The high-temperature region (yellow) is concentrated at higher rotational speeds and moderate axial forces, aligning with expected physical trends. The contour lines indicate a better representation of nonlinearities, resulting in more realistic predictions. This model performs significantly better than ID 1, handling complex relationships more effectively, although some deviations persist.

The ID 7 network exhibits the smoothest temperature distribution. The high-temperature region is well-defined and physically consistent. The model captures clearer boundary transitions, enhancing prediction accuracy. Its temperature distribution aligns most closely with expected FEM

results, suggesting better generalization and learning capability. ID 7 is the most accurate and physically consistent model, effectively predicting realistic temperature variations.

3.3. GA Result

Each fitness function variant (A–H) in Table 2 represents a unique approach to integrating RFW parameters (n , F_A , T) and temperature constraints into a single metric, minimized by the genetic algorithm. Temperature constraints, ensuring the temperature remains within specified limits, are crucial for preventing poor bonding (too low) or excessive intermetallic formation (too high). The fitness functions combine n , F_A , and T in various ways to capture different aspects of the welding process—such as energy input, mechanical stability, or final joint quality.

Table 2. Variants of the fitness functions in a genetic algorithm and their hyperparameters.

ID	F*	Constraints	Search domain		Hyperparameter values common to all FFs	ANN record selected by GA [n, F _A , T]
			common to all FFs	PS NG		
A	$f(n,T)=n+T$	$380 < T < 600$		60 10	CR = 0.6	[500.0, 1750.0, 398]
B	$f(n,T)=1/n+T$	$380 < T < 600$		65 15	MR = 0.2	[1000.0, 750.0, 389]
C	$f(n,T)=1/n+1/T$	$380 < T < 600$		70 20	SM = Tournament	[2250.0, 1000.0, 588]
D	$f(F_A,T)=F_A+T$	$400 < T < 580$	$500 < n < 2500$	80 25	TS = 4	[2500.0, 500.0, 411]
E	$f(F_A,T)=1/F_A+T$	$400 < T < 580$	$500 < F_A < 2500$	85 30	ER = 1%	[500.0, 2000.0, 409]
F	$f(F_A,T)=1/F_A+1/T$	$400 < T < 580$		90 35	TC = Max generations	[1000.0, 2500.0, 563]
G	$f(n,F_A)=n+F_A$	$470 < T < 530$		100 50	ES = Real-valued	[1000.0, 1000.0, 472]
H	$f(n,F_A)=1/(n+F_A)$	$470 < T < 530$		100		[750.0, 2500.0, 510]

* Rotational Speed: n , rpm; Axial Force: F_A , N; Temperature: T , °C. All functions are subject to minimization.

All fitness functions are subject to minimization. Typically, a lower function value indicates a more optimal set of parameters (e.g., minimal temperature deviation, minimal defect formation, or minimal energy consumption). The differences lie in how temperature is penalized or rewarded, how rotational speed and axial force are scaled or combined, and how constraints are enforced.

These differences, along with varying GA hyperparameters (population size, selection method, and mutation/crossover rates), facilitate a systematic exploration of the solution space to determine the optimal RFW settings. Variants with complex or multi-objective fitness definitions required larger populations and more generations for reliable convergence.

Ultimately, selecting a fitness function depends on optimization goals—whether focusing on defect reduction, mechanical performance, energy efficiency, or a balanced approach.

4. Discussion

ANN play a transformative role in the creation and operation of DT by enabling them to learn, adapt, and make accurate predictions based on vast datasets. ANNs complement traditional physics-based models, such as the FEM, by offering data-driven insights, improving efficiency, and enhancing real-time decision-making. They are particularly effective in modeling complex, nonlinear systems where physical modeling is challenging or computationally expensive. ANNs excel at recognizing patterns in large datasets, allowing DT to predict equipment failures or performance degradation. By analyzing sensor data, ANNs can forecast maintenance needs, reducing downtime and costs. Additionally, they enable DT to make real-time adjustments for optimized system performance by continuously learning from operational data and recommending changes to improve efficiency, energy consumption, or output.

To validate a neural network trained on FEM data, various evaluation techniques should be employed, including statistical metrics (MSE, R^2 , and error analysis), graphical comparisons (scatter plots and residuals), and sensitivity tests. An optimized ANN should exhibit low errors, high R^2 values, and close alignment with FEM results. If discrepancies persist, refining the training process, increasing data diversity, or adjusting the network architecture may be necessary.

Heatmaps in Figure 11 illustrate the temperature distribution as a function of axial force, N and rotational speed, rpm. The left heatmap represents FEM-dependent data, while the right heatmap displays ANN-predicted data. The temperature distribution in FEM-dependent data follows an orderly and physically consistent trend, with higher temperatures occurring at higher rotational speeds and axial forces. The values are well-organized, increasing smoothly from bottom-left (low force/speed) to top-right (high force/speed). The highest recorded temperature, $T_{\max} = 632^\circ\text{C}$, is in the upper-right region, indicating maximum thermal impact. This dataset serves as the ground truth reference for evaluating ANN predictions.

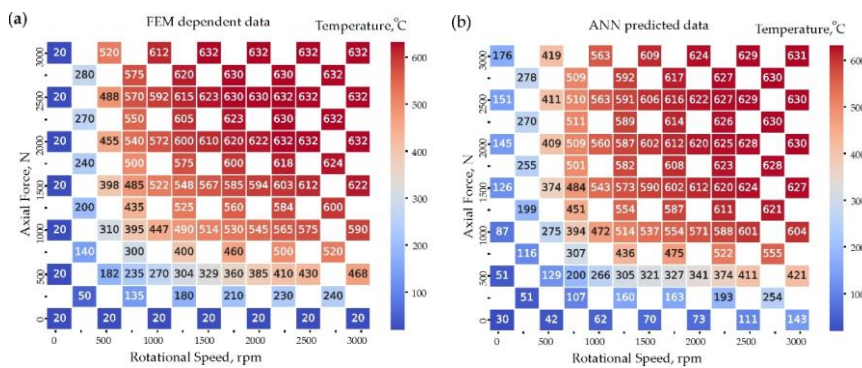


Figure 11. Comparison of annotated heatmaps for; (a) FEM-dependent data; (b) ANN-predicted data.

The ANN model captures the general trend observed in the FEM data but exhibits some deviations. In low axial force and low rotational speed regions, the ANN underestimates temperatures slightly. Conversely, in higher rotational speed and force regions, ANN predictions closely align with FEM values, demonstrating a good approximation of the physical behavior. Minor inconsistencies are present, such as slight overestimation or underestimation in midrange temperatures, highlighting areas for potential improvement. Further optimization of the network architecture, training data, or hyperparameters could enhance the ANN's performance and minimize deviations. While ANNs provide a strong but imperfect approximation of FEM results, they remain a valuable yet improvable tool for predicting temperature distributions. Figure 12 presents a comparative error analysis between FEM data and ANN predictions, using absolute error ($^\circ\text{C}$) and relative error (%). These heatmaps provide insights into how well the ANN model approximates FEM-generated temperature data across various axial force (N) and rotational speed (rpm) combinations. The absolute error represents the direct numerical difference between FEM and ANN temperature predictions. Most values remain low, indicating that the ANN performs well in approximating FEM data. However, higher errors (above 100°C) are observed at extreme values of axial force (e.g., 3000 N) and rotational speed. The majority of midrange values (1000–2500 rpm, 500–2000 N) have errors below 50°C , suggesting good ANN generalization in these regions. This analysis indicates that ANN errors tend to increase at high forces and speeds, potentially due to nonlinearities in thermal behavior.

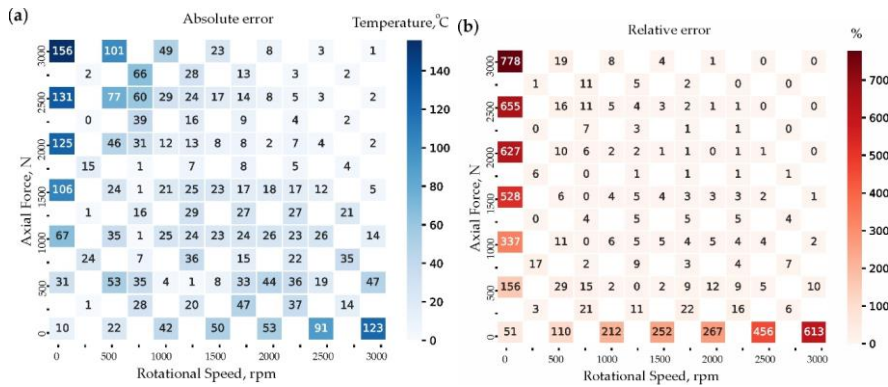


Figure 12. Statistical and graphical evaluation of the similarity between FEM and ANN predictions; (a) Absolute error, °C; (b) Relative error, %.

The relative error (%) highlights the proportional deviation between ANN predictions and FEM results. Higher relative errors occur in low-temperature regions, particularly at low axial forces (0–500 N) and low speeds, where absolute temperatures are small, making any deviation appear large in percentage terms. In high-speed/high-force conditions, the relative error significantly decreases, indicating better ANN accuracy in this range. Some extreme values (e.g., 613% at 3000 rpm, 0 N) arise due to FEM predicting very low temperatures in these regions, causing small absolute deviations to appear disproportionately large in percentage terms. The ANN closely follows FEM results, exhibiting relatively low absolute error.

Extreme values of both absolute and relative error occur at high force and low speed, as well as low force and high speed. The ANN performs well in mid-range values ($n = 1000\text{--}2500$ rpm, $F_A = 500\text{--}2000$ N) but struggles at extreme conditions. The primary source of error likely stems from the neural network's limitations in capturing nonlinear heat distribution. Future refinements, such as incorporating more training data in low-temperature regions, could further reduce prediction errors.

Figure 13 compares temperature predictions from the ANN ID 7 architecture (blue dots) with reference FEM data (red curve) in two scenarios. Case (a) varying rotational speed at a constant axial force ($F_A = 2500$ N). As the rotational speed increases, both reference FEM data and ANN predictions exhibit a rapid rise in temperature, followed by a plateau at higher speeds. The close alignment of the blue dots with the red curve suggests that the ANN effectively models the thermal behavior in this range. Small deviations appear at intermediate speeds, but overall, the ANN captures the trend well near the maximum temperature.

Case (b) varying axial force at a constant rotational speed ($n = 2250$ rpm). As the axial force increases from low to high values, the temperature rises steeply initially, then levels off once the force becomes sufficiently large. The ANN predictions closely track the FEM reference curve, demonstrating good generalization across different force values. Minor discrepancies occur at mid-range forces, but the model still provides a smooth and physically plausible temperature profile.

Overall, the ANN ID 7 network demonstrates strong agreement with reference data under both constant force and constant speed conditions. Its Tanh-Tanh-Sigmoid architecture, combined with sufficient training, enables the model to capture nonlinear thermal effects effectively.

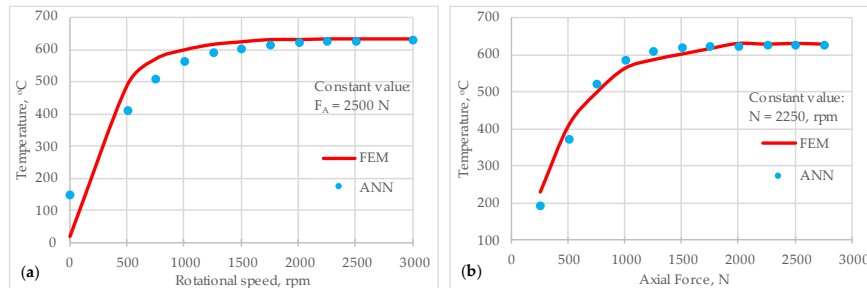


Figure 13. Temperature prediction for ANN ID 7 network architecture; (a) Rotational Speed, rpm vs. Temperature, °C at constant axial force $F_A = 2500$ N; (b) Axial Force, N vs. Temperature °C at constant rotational speed $n = 2250$ rpm.

Figure 14 presents a color-coded contour map of temperature, °C predicted by the ANN ID 7 architecture, as a function of rotational speed, rpm and axial force, N. As ID 7 is the deepest and most accurate ANN model, its predictions yield a well-defined temperature distribution. The contour map suggests that the model successfully captures nonlinear thermal effects across a broad operating range. Points A through H mark specific operating conditions identified through an optimization procedure. By examining the locations of these points on the map and their corresponding temperature values, one can assess how different combinations of rotational speed and axial force influence thermal outcomes.

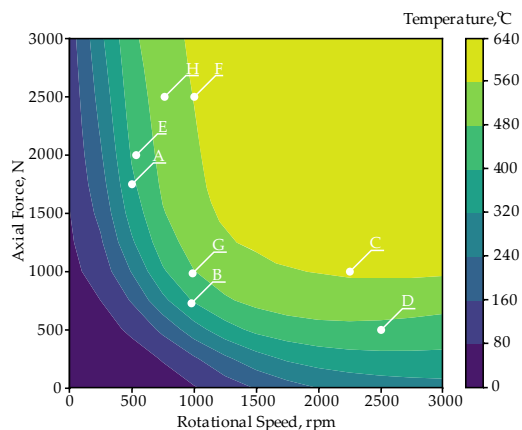


Figure 14. Optimization results (A-H) marked on the temperature distribution map for the ANN ID 7 architecture.

The optimization points (A-H) are distributed across the diagram, covering low and high rotational speeds and forces. Each labeled point represents a trade-off between process constraints and objectives—for instance, maintaining lower temperatures while ensuring sufficient force or maximizing temperature within safe operational limits. These trade-offs depend on engineering goals, such as optimizing material properties to meet specific performance requirements.

Overall, the labeled points (A-H) on the temperature distribution map demonstrate how ANN predictions can be leveraged to optimize the RFW process, revealing different operating conditions that achieve desired temperature levels.

5. Conclusions

This study investigated the integration of ANN and GA within a DT framework to optimize the RFW process and accurately predict temperature distributions and joint quality. Extensive numerical simulations FEM, experimental validations, and error analyses were employed to assess the impact of key parameters and the performance of various ANN architectures. As found in experimental and FEM studies, careful selection of axial force, along with rotational speed and welding time, is crucial for achieving an optimized plastic strain distribution. Optimal RFW process parameters will provide a balance between sufficient deformation and heat generation, leading to a solid joint with minimal defects. Considering the results of the conducted studies and their discussion, it was concluded that:

ANN Model Performance

Different ANN architectures were evaluated, revealing that shallow networks (ID 1–2) learn quickly but exhibit unstable behavior and poor generalization. In contrast, moderate-depth networks (ID 3–4) demonstrate balanced learning and stable convergence. The deep network (ID 7), employing nonlinear Tanh-Tanh-Sigmoid activation functions and a low learning rate, achieved the highest accuracy, with predictions matching FEM data. The ANN model successfully captures the general temperature distribution, though minor inaccuracies appear in the low-force, low-speed region. High-temperature regions are well predicted, confirming the ANN's ability to model complex thermal behaviors. The ANN model demonstrates strong predictive capabilities, with absolute and relative errors mostly below 50°C in key operational ranges. Relative error spikes at low force/speed conditions, suggesting that improvements in model training or architecture could enhance accuracy.

Optimization via GAs

GAs were effectively integrated with the DT to systematically explore the RFW parameter space. This approach enables the optimization of complex, multi-criteria objectives such as joint quality, defect minimization, and energy efficiency, thereby facilitating real-time adjustments and improved process performance.

Error Analysis and Predictive Capabilities

Comparative analyses using statistical metrics and graphical methods demonstrated that ANN predictions generally align well with FEM results—particularly in mid-range operational conditions—although discrepancies are more pronounced at extreme force and speed settings. These findings underscore the potential for further improvements in ANN training and network architecture. The temperature contour map highlights optimal operating conditions, helping engineers balance factors such as temperature, force, and speed. The A–H optimization points illustrate trade-offs in welding parameters, assisting in process optimization for different material and performance requirements.

Overall, the combination of ANN-based predictive models with GA optimization within a DT framework offers a powerful, computationally efficient tool for enhancing the RFW process. The approach successfully captures complex thermal behaviors and material interactions, leading to superior joint quality and operational efficiency. Future work should focus on refining network architectures and expanding training datasets, to further improve model performance across all operating conditions.

Author Contributions: Conceptualization, P.L.; methodology, P.L.; software, P.L. K.L. and W.L.; validation, P.L. K.L. and W.L.; formal analysis, P.L.; investigation, P.L. and J.A.; resources, P.L. and J.A.; data curation, P.L. and J.A.; writing—original draft preparation, P.L.; writing—review and editing, J.A.; visualization, P.L. K.L. and W.L.; supervision, P.L. and J.A.; project administration, J.A.; funding acquisition, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financed by Czestochowa University of Technology under research project BS/PB-500-301-1/23 Z.3.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANN	Artificial Neural Networks
COF	Combined Objective Function
CR	Crossover Rate
DT	Digital Twins
ER	Elitism Rate
ES	Encoding Scheme
F	Fitness Function
FEM	Finite Element Method
FF	Fitness Function
FS	Fitness Score
GA	Genetic Algorithm
HAZ	Heat Affected Zone
ML	Machine Learning
MR	Mutation Rate
MSE	Mean Square Error
NG	Number of Generations
PS	Population Size
R ²	R-squared
RFW	Rotary Friction Welding
RSM	Response Surface Methodology
SM	Selection Method
TC	Termination Criteria
TS	Tournament Size
UTS	Ultimate Tensile Strength

Appendix A

The dataset from the FEM calculation. It contains 105 cases and was divided into three subsets: training set, validation set and test set with the ratio: training (47%) / validation (36%) / test (20%). The training set was used to learn patterns. The validation set helped to tune the model and avoid overfitting. The test set was used for the final evaluation of the model performance.

Table A1. Dataset from the FEM calculation.

Training set (49%)				Validation set (34%)				Test set (19%)			
ID	X1	X2	Z1	ID	X1	X2	Z1	ID	X1	X2	Z1
1	3000	0	20	50	2750	250	240	86	2250	500	410
2	2500	0	20	51	2250	250	230	87	1750	500	360
3	2000	0	20	52	1750	250	210	88	1250	500	304
4	1500	0	20	53	1250	250	180	89	750	500	235
5	1000	0	20	54	750	250	135	90	2250	1000	565
6	500	0	20	55	250	250	50	91	1750	1000	530
7	0	0	20	56	2750	750	520	92	1250	1000	490
8	3000	500	468	57	2250	750	500	93	750	1000	395

9	2500	500	430	58	1750	750	460	94	2250	1500	603
10	2000	500	385	59	1250	750	400	95	1750	1500	585
11	1500	500	329	60	750	750	300	96	1250	1500	548
12	1000	500	270	61	250	750	140	97	750	1500	485
13	500	500	182	62	2750	1250	600	98	2250	2000	632
14	0	500	20	63	2250	1250	584	99	1750	2000	620
15	3000	1000	590	64	1750	1250	560	100	1250	2000	600
16	2500	1000	575	65	1250	1250	525	101	750	2000	540
17	2000	1000	545	66	750	1250	435	102	2250	2500	632
18	1500	1000	514	67	250	1250	200	103	1750	2500	630
19	1000	1000	447	68	2750	1750	624	104	1250	2500	615
20	500	1000	310	69	2250	1750	618	105	750	2500	570
21	0	1000	20	70	1750	1750	600				
22	3000	1500	622	71	1250	1750	575				
23	2500	1500	612	72	750	1750	500				
24	2000	1500	594	73	250	1750	240				
25	1500	1500	567	74	2750	2250	632				
26	1000	1500	522	75	2250	2250	630				
27	500	1500	398	76	1750	2250	623				
28	0	1500	20	77	1250	2250	605				
29	3000	2000	632	78	750	2250	550				
30	2500	2000	632	79	250	2250	270				
31	2000	2000	622	80	2750	2750	632				
32	1500	2000	610	81	2250	2750	630				
33	1000	2000	572	82	1750	2750	630				
34	500	2000	455	83	1250	2750	620				
35	0	2000	20	84	750	2750	575				
36	3000	2500	632	85	250	2750	280				
37	2500	2500	632								
38	2000	2500	630								
39	1500	2500	623								
40	1000	2500	592								
41	500	2500	488								
42	0	2500	20								
43	3000	3000	632								
44	2500	3000	632								
45	2000	3000	632								
46	1500	3000	632								
47	1000	3000	612								
48	500	3000	520								
49	0	3000	20								

References

1. Lacki, P.; Adamus, J.; Więckowski, W.; Motyka, M. A New Method of Predicting the Parameters of the Rotational Friction Welding Process Based on the Determination of the Frictional Heat Transfer in Ti Grade 2/AA 5005 Joints. *Materials (Basel)* **2023**, *16*, doi:10.3390/ma16134787.
2. Lakache, H.E.; May, A.; Badji, R.; Benhammouda, S.E.; Ramtani, S. The Impact of the Rotary Friction Welding Pressure on the Mechanical and Microstructural Characteristics of Friction Welds Made of the Alloys TiAl6V4 and 2024 Aluminum. *Exp Tech* **2024**, *48*, 473–484, doi:10.1007/s40799-023-00671-z.
3. Bauri, S.K.; Babu, N.K.; Ramakrishna, M.; Rehman, A.U.; Prasad, V.J.; Suryanarayana Reddy, M.R. Microstructural and Mechanical Properties of Dissimilar AA7075 and AA2024 Rotary Friction Weldments. *Crystals* **2024**, *14*, 1011, doi:10.3390/cryst14121011.
4. Hu, M.; Zhi, S.; Chen, J.; Li, R.; Liu, B.; He, L.; Yang, H.; Wang, H. Restraint of intermetallic compound and improvement of mechanical performance of Ti/Al dissimilar alloy by rotary friction welding based on laser powder bed fusion. *Journal of Manufacturing Processes* **2024**, *131*, 440–454, doi:10.1016/j.jmapro.2024.09.021.

5. Yaknesh, S.; Rajamurugu, N.; Babu, P.K.; Subramanian, S.; Khan, S.A.; Saleel, C.A.; Nur-E-Alam, M.; Soudagar, M.E.M. A technical perspective on integrating artificial intelligence to solid-state welding. *Int J Adv Manuf Technol* **2024**, *132*, 4223–4248, doi:10.1007/s00170-024-13524-9.
6. Mattie, A.A.; Ezdeen, S.Y.; Khidhir, G.I. Optimization of parameters in rotary friction welding process of dissimilar austenitic and ferritic stainless steel using finite element analysis. *Advances in Mechanical Engineering* **2023**, *15*, doi:10.1177/16878132231186015.
7. Sharma, A.; Kosasih, E.; Zhang, J.; Brintrup, A.; Calinescu, A. Digital Twins: State of the art theory and practice, challenges, and open research questions. *Journal of Industrial Information Integration* **2022**, *30*, 100383, doi:10.1016/j.jii.2022.100383.
8. Negri, E.; Fumagalli, L.; Macchi, M. A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing* **2017**, *11*, 939–948, doi:10.1016/j.promfg.2017.07.198.
9. Wagner, R.; Schleich, B.; Haefner, B.; Kuhnle, A.; Wartzack, S.; Lanza, G. Challenges and Potentials of Digital Twins and Industry 4.0 in Product Design and Production for High Performance Products. *Procedia CIRP* **2019**, *84*, 88–93, doi:10.1016/j.procir.2019.04.219.
10. Glaessgen, E.H.; Stargel, D.S. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. *Paper for the 53rd Structures, Structural Dynamics, and Materials Conference: Special Session on the Digital Twin* **2012**.
11. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction* **2020**, *114*, 103179, doi:10.1016/j.autcon.2020.103179.
12. Attaran, M.; Celik, B.G. Digital Twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal* **2023**, *6*, 100165, doi:10.1016/j.dajour.2023.100165.
13. Więckowski, W.; Lacki, P.; Derlatka, A.; Wiczorek, P. Numerical Study of Rotary Friction Welding of Automotive Components. *Adv. Sci. Technol. Res. J.* **2024**, *18*, 336–346, doi:10.12913/22998624/188742.
14. Nayif, A.M.; A. D., Y.; Al Sarraf, Z.S. Investigation of the Process Parameters in Rotary Friction Welded Dissimilar AA7075/AA5083 Aluminum Alloy Joints on Fatigue Initiation using FEA and ANN. *WSEAS TRANSACTIONS ON APPLIED AND THEORETICAL MECHANICS* **2024**, *19*, 97–112, doi:10.37394/232011.2024.19.11.
15. Winiczenko, R.; Skibicki, A.; Skoczylas, P. Optimization of Friction Welding Parameters to Maximize the Tensile Strength of Magnesium Alloy with Aluminum Alloy Dissimilar Joints Using Genetic Algorithm. *Processes* **2021**, *9*, 1550, doi:10.3390/pr9091550.
16. Sreenivasan, K.S.; Satish Kumar, S.; Katiravan, J. Genetic algorithm based optimization of friction welding process parameters on AA7075-SiC composite. *Engineering Science and Technology, an International Journal* **2019**, *22*, 1136–1148, doi:10.1016/j.jestch.2019.02.010.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.