IJSOM International Journal of Supply and Operations Management

Optimization of Key Parameters in a Welding Process Using Multivariate Statistical Analysis and Design of Experiments

Narges Delavar Pasikhani [1] ✉, **Abbas Foroozanfar** [2], **Seyed Taghi Akhavan Niaki** [3]

1. Corresponding author, Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran. E-mail: delavar.narges@ie.sharif.edu
2. Department of Industrial Engineering, Sharif university of technology, Tehran, Iran. E-mail: forouzanfar.abbas@ie.sharif.edu
3. Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran. E-mail: niaki@sharif.edu

Article Info

Article type:
Research Article

Article history:
Received July 8, 2025
Received in revised form November 12, 2025
Accepted May 29, 2026
Available online May 29, 2026

***Keywords*:**
design of experiment
welding process
2^K factorial design
principal component analysis
linear discriminant analysis

ABSTRACT

Objective: Welding quality strongly affects product performance, durability, scrap rate, and production cost. This study aims to develop an integrated data-driven framework combining Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Design of Experiments (DOE) to optimize welding parameters and improve weld quality in an industrial manufacturing process.

Methods: Historical production data from a turbine manufacturing company in Iran were analyzed. The investigated welding parameters included current intensity, welding position, welding speed, preheat temperature, and electrode diameter. A 2^k factorial design was applied to evaluate their effects on weld depth, which was maximized, and weld width, which was minimized. Multi-response optimization was conducted using response surface methodology and desirability functions, with different weighting schemes assigned to the response variables.

Results: The findings revealed that current intensity was the most influential parameter affecting weld quality. Moreover, a significant interaction effect was identified between current intensity and welding position, indicating that the joint effect of these parameters plays an important role in determining weld profile characteristics. Compared with the historical data, the optimized parameter settings improved weld depth by an average of 13.77% and enhanced weld width by an average of 3.56%. Following the implementation of the optimal parameter values in the production process, the average scrap rate decreased from 19.8% to 13.1% over a six-month period.

Conclusion: The proposed framework effectively improved weld quality and production performance. The reduction in scrap rate and enhancement of welding responses resulted in an estimated 3–4% decrease in production costs, demonstrating the practical value of the approach for industrial welding optimization.

Cite this article: Delavar Pasikhani, N., Foroozanfar, A., & Akhavan Niaki, S. T. (2026). Optimization of Key Parameters in a Welding Process Using Multivariate Statistical Analysis and Design of Experiments. *International Journal of Supply and Operations Management*, *XX*(X), pages. https://doi.org/10.22034/ijsom.2026.110844.3429

Publisher: Kharazmi University
DOI: https://doi.org/10.22034/ijsom.2026.110844.3429

1. Introduction

Welding is an important part of many industrial processes and serves as the foundation for assembling structures and components (Breitenbach et al., 2021). Many factors affect the quality and performance of welding, including the amount of heat used, the welding speed, the type of electrode used, and the welding position. Due to the complex interactions between these factors, it can be challenging to optimize the welding process for improved weld performance, including increased efficiency, reduced defects, and enhanced joint strength. Organizations must improve welding procedures through systematic optimization if they want to satisfy customers and prosper in the current competitive landscape (Sabegh et al., 2014). When designing experiments for welding, researchers often employ trial-and-error methods or full-factorial designs that encompass every possible factor. These methods have worked well in the past, but they become less valuable as the number of parameters increases (H. Zhang et al., 2024).

However, the performance of the welding process over all degrees of freedom is not only time-consuming but prohibitively expensive. To address these challenges, data-driven approaches are being increasingly developed as efficient tools for identifying and ranking significant factors in experimental design. These solutions are crucial for accuracy and minimizing chance errors in weld property prediction and process parameter optimization, which results in higher-quality welds (Mahadevan et al., 2021). The integration of intelligent systems into welding processes has led to the development of intelligent welding systems (IWS), which enhance product quality, reduce human intervention, and improve process efficiency through advanced modeling, prediction, and control of key process parameters (S. Kumar et al., 2022).

This work offers a new approach for the optimization of welding processes by associating statistical techniques (namely Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA)) and Design of Experiments DOE methods. The high-dimensional parameter space is analyzed using PCA and LDA, and the most influential factors on the response variable are identified. They allow a more focused and resource-efficient experimental design to emerge as they shrink the dimensionality of the parameter space. According to the identified influential factors, a 2^K factorial experimental design is used to experimentally study the influences and interactions of the parameters on the welding process. The 2^K design is an efficient design that reveals the main and interaction effects of K factors in just a few experimental runs.

The integration of statistical methods and experimental design addresses the limitations of conventional methods. By selecting the most significant parameters, this approach strikes a balance between the accuracy of the experiment and its cost-effectiveness. This not only leads to savings in terms of cost and time spent performing the experiments, but also increases the robustness of the results, as the parameters are chosen in a data-driven manner.

The goal of this study is to highlight the synergy between statistical techniques and experimental design elements in optimizing the welding process. The proposed approach is scalable and systematic, making it generalizable to other complex manufacturing processes. A schematic representation of the proposed approach is provided in Figure 1.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. Section 3 presents the methodology adopted in this study. Section 4 discusses the case study and the corresponding results. Section 5 evaluates the practical impact on scrap rate and cost, and Section 6 concludes the paper and offers directions for future research.

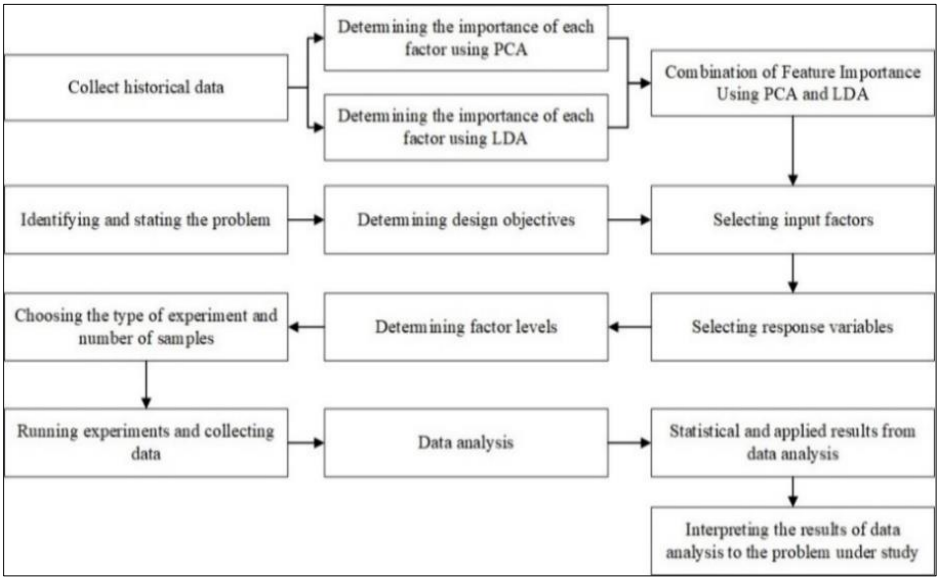

Figure 1. Graphical abstract of proposed methodology

2. Literature Review

This section reviews the application of Machine Learning (ML) techniques in optimizing welding process parameters, with a focus on experimental design methodologies, including PCA and LDA, for feature selection, followed by the implementation of 2^K factorial designs.

2.1. Implementation of Machine Learning in Design of Experiments

The optimization of welding parameters has been extensively investigated, with various techniques employed to improve weld quality and efficiency. Early approaches primarily relied on empirical methods and statistical techniques, such as the Taguchi method (Ahmed et al., 2022). The advent of ML has opened new possibilities for welding optimization. Various ML algorithms have been successfully applied to model and predict weld characteristics, leading to improved process control and reduced defect rates(Anand et al., 2015). Before applying ML models, it is often beneficial to select the most relevant process parameters for the optimization study. PCA and LDA are two powerful statistical techniques that can be used for this purpose (Guo et al., 2021).

PCA's primary role in experimental design is dimensionality reduction, which simplifies complex datasets by identifying the most informative variables (Reddy et al., 2020). Antony (2000) introduces a novel feature extraction methodology that accounts for interactions between variables, addressing a gap in diagnostic research. Utilizing a fractional factorial DOE, the authors extract features through PCA to capture variation patterns in tonnage signals. The methodology employs regression analyses for modeling relationships between features and process variables, along with hierarchical classifiers and cross-validation for root-cause determination and diagnostic performance evaluation, illustrated through a real-world example. Zhang and Edgar (2008) propose a PCA-based optimal criterion (P-optimal) for DOE in dynamic systems, enhancing parameter estimation for process system engineers. This approach integrates PCA with information matrix analysis, encompassing various widely used optimal design criteria and automatically selecting the best objective function for specific differential-algebraic (DAE) systems. Khoshnevisan et al. (2021) focus on modeling energy consumption and water content in natural gas processing plants (NGPP) using a radial basis function neural network (RBF-NN). By applying PCA on normalized data, the authors develop a reduced-dimensional dataset for their model. Babanova et al. (2014) employ DOE and PCA to enhance the performance of air-breathing

gas-diffusional enzymatic electrodes. Using DOE, the authors systematically evaluate factors affecting the composite system, leading to the development of an improved cathode. PCA is utilized to validate DOE findings and visualize the contributions of key factors, such as enzyme amount, phosphate buffer volume, and gas-diffusion layer thickness, to the cathode's performance. Myftarago et al. (2023) investigate the performance of ternary binders, focusing on Portland Cement-Calcium Aluminate Cement-Calcium Sulphate (PC-CAC-CS) compositions, mineral powders, and chemical admixtures. Using a Taguchi design with a limited number of trials, the research assesses factors like binder composition, replacement ratios, and admixture levels. Measurements on hydration kinetics, dimensional stability, and compressive strength are conducted, with PCA employed to analyze hydration curves and X-ray diffraction data.

In welding-related studies, Meng et al. (2020) investigated the optimization of ultrasonic metal welding (UMW) parameters to enhance joint quality, with a focus on peel and shear strength. They proposed a machine learning-based response surface methodology to model complex relationships between welding parameters and quality. Comparing various models, they found that Gaussian Process Regression (GPR) and Support Vector Regression (SVR) with radial basis function kernels provided the most accurate predictions. UMW is an effective method for joining thermoplastics without compromising the material's integrity. He et al. (2022) propose a novel approach to quality prediction and process parameter optimization in resistance spot welding, integrating a Chaos Game Optimization (CGO) algorithm with a Multi-output Least Squares Support Vector Regression (MLSSVR) model. This method effectively predicts weld quality indicators, such as nugget diameter and tensile shear load, using small sample data, and demonstrates improved accuracy compared to traditional models. Furthermore, the use of the Particle Swarm Optimization (PSO) algorithm enables the determination of optimal welding parameters. (Baraya et al., 2024) utilized a full factorial design combined with a Backpropagation Neural Network (BPNN) to analyze the impact of ultrasonic welding parameters on the joint strength, heat generation, and electrical properties of copper wire and non-woven PVC textiles, showcasing the efficiency of machine learning in optimizing welding processes for innovative textile applications.

Apart from their individual applications, various recent studies have explored the potential capabilities of PCA and LDA to complement each other by being integrated or combined effectively for feature extraction and classification tasks in diverse fields of science. Thus (Li et al., 2009) suggested a face recognition approach by applying PCA initially as a variance-based feature transformation followed by the implementation of LDA as a class-separated transformation in the PCA-transformed space, achieving a significantly improved recognition rate than that achieved solely by PCA. (Lee et al., 2010) described the generation of robust features by applying LDA transformation to the PCA descriptions for local features, validating the capability of their combined transformation approach to produce more consistent global features effectively in object detection applications. (Hayati et al., 2024) designed a novel chemometric model by integrating PCA, LDA, and a Support Vector Machine (SVM) classifier to predict the origin of coconuts measured by their NIR spectra, validating that the integration approach between PCA and LDA significantly enhanced the accuracy rate of discrimination even with relatively complex and large datasets. Likewise, (Ali et al., 2022) designed an intelligent PCA-LDA-based feature fusion scheme to identify plant leaf diseases, applying PCA as a feature selector providing cues to LDA classifiers, resulting in an accuracy rate over 98%. (Sharma et al., 2006), theoretically described the integrated classifier approach by uniting class-dependent PCA and LDA, validating that both transformations could be statistically combined as linear projections to enhance class discriminability efficiently, preventing singularity issues due to the involvement of a covariance matrix. Together, all mentioned studies profoundly validate the idea that PCA transformation with a variance orientation and another LDA transformation with a class separability orientation are actually two relatively complementary linear projections and can be effectively combined or integrated either sequentially or concurrently without violating any assumptions related to consistency, hence make this integration study meaningful and acceptable by both theory and practice.

2.2. Factorial Design

After feature selection, the next step in the optimization process involves designing appropriate experiments. 2^K factorial designs are commonly used in experimental design for their efficiency in exploring the effects of multiple factors and their interactions (Camarena-Martinez et al., 2023).

Ramasamy et al. (2002) evaluate arc stud welding strength on uncoated and galvanized sheets, examining factors such as stud coating, electrode wear, polarity, power supply, stud design, substrate thickness, surface oils, and steel coating condition. Weld quality was assessed through various tests and metallographic analysis. A 3^2-experiment design analyzed eight factors and four quality criteria. Results indicated that larger studs and thicker sheets had a significant impact on weld performance, despite an increase in internal porosity. Haribalaji et al. (2022) investigate Friction Stir Welding (FSW) of dissimilar AA2014 and AA7075 aluminum alloys using a Taguchi design of experiments. Key input parameters include tool rotation speed, welding speed, axial force, and tool tilt angle, while output parameters are tensile strength and microhardness. Surface topography was analyzed via Scanning Electron Microscopy (SEM), and Analysis of variance (ANOVA) identified significant factors. The study by Kumar et al. (2023) highlights the significance of optimizing welding parameters in the FSW process for Aluminum Alloy 6061, using methods such as the Taguchi approach and ANOVA to determine the influence of parameters like tool geometry, rotational speed, and welding speed on mechanical properties, which serves as a foundation for integrating machine learning and DOE in further enhancing the welding process efficiency and accuracy (R. R. Kumar et al., 2023).

Jagtap et al. (2017) investigated the effects of welding current, feed rate, and the number of welding passes on the tensile strength and yield strength of IS 2062 E250 steel through experimental methods and mathematical modeling. Multiple linear regression models were employed to establish correlations between welding parameters and performance metrics. ANOVA was utilized to predict the impact of welding parameters on response variables. The findings indicate that tensile strength and yield strength are significantly influenced by higher levels of welding current and the number of passes. Optimal welding parameters were determined using the Response Surface Methodology (RSM). Das et al. (2025) investigate the impact of welding parameters, including welding current, arc distance, electrode vertex angle, and welding speed, on the geometry of TIG welds in Zircaloy-2 cladding, to develop mathematical models to optimize weld bead depth and width. Rathi (2021) utilized a fractional factorial design to optimize submerged arc welding (SAW) parameters, demonstrating the significant influence of process variables, such as voltage, trolley speed, and wire feed rate, on the weld bead geometry and hardness in mild steel.

Choudhary et al. (2024) employed a two-level half-factorial design to investigate the effects of welding current, arc voltage, welding speed, and nozzle-to-plate distance on key bead geometry parameters, including weld penetration shape factor (WPSF), weld reinforcement form factor (WRFF), and dilution percentage. The results indicate significant interactions between these process variables, with welding current and voltage showing pronounced effects on WPSF and WRFF, while welding speed and nozzle-to-plate distance influence the dilution percentage and bead volume. Mounika et al. (2024) propose an optimization approach integrating Taguchi's L27 orthogonal array, RSM, and fuzzy logic to optimize friction welding parameters, minimizing metal loss and maximizing tensile strength, with results validated through comparison to experimental data. (Boriwal et al., 2021) investigate the effect of input process parameters on the strength of spot-welded joints of dissimilar materials, utilizing full factorial design to create an experimental matrix and develop a mathematical model. The model's validity is confirmed through ANOVA and experimental testing.

Richmire et al. (2018) examined the identification of process parameters and their effects on FSW outcomes using a statistical approach, specifically the Taguchi method. This research focused on the impact of tool rotation speed (RPM) and traverse speed on hardness distribution in the stir zone for the magnesium alloy AM60. Experiments were conducted at three different rotation speeds and three traverse speeds, utilizing a 3^2 factorial design for data analysis. The results indicate that both parameters significantly affect the hardness of the stir zone, and interactions between

these two factors were also statistically evaluated. Recent studies have explored the optimization of welding process parameters through various methods, including the Taguchi design method, to enhance welding quality. For instance, Tesfaw et al. (2022) employed the Taguchi method to optimize the Metal Active Gas (MAG) welding parameters, achieving significant improvements in welding hardness through a systematic analysis of key factors, including welding voltage, current, wire speed, and gas flow rate.

(Madrid et al., 2019) present an innovative approach for optimizing the design of welded aircraft components by integrating experimental design techniques with welding simulation. Traditional methods often overlook the complexities of the welding process, resulting in higher costs due to redesign. This study introduces a virtual experimental design method to model and optimize the interactions between design and welding parameters during the early design phase. A case study focuses on optimizing weld penetration and distortion, using initial physical welds to develop a new heat source model. An experimental matrix is created to effectively test various parameter combinations, facilitating robust analyses of weld manufacturability. Kuo et al. (2022) investigated the inefficiencies of trial-and-error methods in optimizing ultrasonic welding (UW) parameters, which can result in suboptimal outcomes. They utilized Taguchi methods to identify optimal parameters for achieving high weld strength in plastic chargers. Key factors affecting weld strength include amplitude, welding pressure, hold time, and trigger position.

In the reviewed literature, it has been observed that there is a lack of research on the integration of PCA and LDA for dimensionality reduction and the identification of key factors in experimental design. This study innovatively utilizes existing historical data within the system to design experiments and identify significant factors, offering a comprehensive approach to the welding process. Furthermore, the research leverages the advantages of both PCA and LDA methods, introducing an innovative approach that combines the importance scores from both techniques. This approach enables experts to select relevant factors as final determinants, enhancing the effectiveness of the experimental design process.

3. Methodology

3.1. PCA

PCA is a dimensionality-reduction technique that identifies the directions (principal components) corresponding to linear combinations of the original variables along which the projected data variance is maximized. The step-by-step mathematical explanation and implementation are outlined below:

Step 1. Problem setup: A dataset $\mathbf{X}$ with n samples and p features ($\mathbf{X} \in \mathbb{R}^{n\times p}$) is given. First, the dataset needs to be standardized, which involves calculating the mean and standard deviation for each feature (column) independently. Each feature is then scaled to have a mean of zero and a standard deviation of one.

$$\hat{\mathbf{x}}_i = \frac{\mathrm{x}_i - \mu_i}{\sigma_i} \tag{1}$$

where $\boldsymbol{\mu_i}$ is the mean of each feature ($\boldsymbol{\mu} \epsilon \mathbb{R}^p$) and σ_i is the standard deviation of the i-th feature (column) .

Step 2. Compute the correlation matrix: Since the data are standardized, the correlation matrix is used to capture the relationships between features, rather than the covariance matrix. The correlation matrix is calculated by normalizing the covariance matrix, where each feature is scaled by its standard deviation. This ensures that each feature contributes equally to the analysis, and the relationship between standardized features is appropriately captured.

$$\mathbf{R} = \frac{1}{n-1}\sum_{m=1}^{n}\left(\frac{p_{mi}-\bar{p}_m}{\sigma_i}\right)\left(\frac{p_{zi}-\bar{p}_z}{\sigma_z}\right) \tag{2}$$

The correlation matrix $\mathbf{R}$ is calculated by summing over all n samples, where p_{mi} and p_{zi} represent the values of the i-th and z-th features for the m-th sample, $\bar{p}_m$and $\bar{p}_z$ are the mean values of the i-th and z-th features across all samples, and σ_i and σ_z are the standard deviations of the i-th and z-th features, respectively.

Step 3. Solve the Eigen equation: Find the eigenvalues (λ) and eigenvectors ($\boldsymbol{\omega}$) of the correlation matrix **R**.

$$\mathbf{R}\boldsymbol{\omega} = \lambda\boldsymbol{\omega} \tag{3}$$

The eigenvalues λ represent the variance explained by each principal component and the eigenvectors $\boldsymbol{\omega}$ represent the directions of the principal components.

Step 4. Sort eigenvalues and eigenvectors:

1. Arrange the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$ in descending order.

2. Reorder the eigenvectors $\boldsymbol{\omega_1}, \boldsymbol{\omega_2}, \dots, \boldsymbol{\omega_p}$ accordingly.

Step 5. Select top components: Select the top **k** eigenvectors corresponding to the k largest eigenvalues. These define the principal components that capture the most variance in the data.

$$\mathbf{w}_{pca} = [\boldsymbol{\omega_1}, \boldsymbol{\omega_2}, \dots, \boldsymbol{\omega_k}] \tag{4}$$

Step 6. Transform the data: Project the original data $\widehat{\mathbf{x}_i}$ onto the principal components.

$$\mathbf{X}_{pca} = \widehat{\mathbf{x}_i}\mathbf{w}_{pca} \tag{5}$$

$\mathbf{X}_{pca} \in \mathbb{R}^{n\times k}$ is the reduced dataset with k principal components.

Step 7. Compute feature importance: The importance of each feature in the original dataset can be derived from the magnitudes of the entries in the eigenvectors.

$$IoF_j^{PCA} = \sum_{i=1}^{k} \boldsymbol{\omega}_{ij}^2 \tag{6}$$

Where $\boldsymbol{\omega}_{ij}$ is the j-th element of the i-th eigenvector ($j = 1, \dots, n$).

Final outcome:

1. Reduced dataset: The dataset is reduced to k-dimensions while retaining most of the variation.

2. Feature importance: Each feature has an importance score derived from its corresponding eigenvector.

3.2. LDA

LDA is a dimensionality reduction technique and a classifier. Below a step-by-step implementation for LDA, focusing on its mathematical foundation and how it assigns importance to each factor in the input dataset is explained.

Step 1. Problem setup: A dataset **X** with n samples and p features ($\mathbf{X} \in \mathbb{R}^{n\times p}$) and a target vector **y** with k unique classes ($\mathbf{y} \in \{1, \dots, k\}$) are given.

Step 2. Compute class-wise statistics: For each class c ($c = 1, \dots, k$):

1. Compute the mean vector:

$$\boldsymbol{\mu}_c = \frac{1}{n_c}\sum_{i:y_i=c} \mathbf{X}_i \tag{7}$$

where n_c is the number of samples in class c, and $\mathbf{X}_i$ is the i-th sample.

2. Compute the overall mean:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i \tag{8}$$

Step 3. Compute scatter matrices:

1. Within-class scatter matrix:

$$\mathbf{S}_W = \sum_{c=1}^{k} \sum_{i:y_i=c} (\mathbf{X}_i - \boldsymbol{\mu}_c)(\mathbf{X}_i - \boldsymbol{\mu}_c)^T \quad (9)$$

2. Between-class scatter matrix:

$$\mathbf{S}_B = \sum_{c=1}^{k} n_c (\boldsymbol{\mu}_c - \mu)(\boldsymbol{\mu}_c - \mu)^T \quad (10)$$

Step 4. Solve the generalized Eigen equation: Find the eigenvalues (λ) and eigenvectors ($\boldsymbol{\omega}$) of the matrix equation.

$$\mathbf{S}_W^{-1}\mathbf{S}_B\boldsymbol{\omega} = \lambda\boldsymbol{\omega} \quad (11)$$

The eigenvalues λ represent the variance explained by each discriminant component and the eigenvectors $\boldsymbol{\omega}$ define the direction of the new axes (discriminant components).

Step 5. Sort and select components:

1. Sort the eigenvectors $\boldsymbol{\omega}$ by their corresponding eigenvalues λ in descending order.

2. Select the top $\mathbf{k} - 1$ eigenvectors (as there are $k - 1$ discriminant components in k-class problems).

Step 6. Transform the data: Transform the original dataset $\mathbf{X}$ onto the new axes.

$$\mathbf{X}_{lda} = \mathbf{XW} \quad (12)$$

where $\mathbf{W}$ is the matrix of selected eigenvectors.

Step 7. Assign importance to features: The significance of each original feature can be measured by the magnitude of its corresponding entry in the eigenvectors $\boldsymbol{\omega}$. For example, the importance of the feature j can be represented as:

$$IoF_j^{LDA} = \|\boldsymbol{\omega}_j\| \quad (13)$$

After following these steps, each factor in the input dataset $\mathbf{X}$ has a numeric importance, derived from the eigenvectors. Additionally, the dataset is transformed into a lower-dimensional space for further analysis or classification.

3.3. Combination of Feature Importance Using PCA and LDA

In this section, the significance of features previously obtained through two methods, PCA and LDA, is explored. These importance metrics are combined linearly to derive a final assessment of each feature's significance. This linear combination not only allows us to leverage the advantages of both methods in evaluating feature importance but also mitigates any biases that may arise from relying solely on one approach. In other words, this combination provides a more comprehensive and accurate representation of feature significance in data analysis.

Consequently, this approach enables us to make more informed decisions regarding the selection and evaluation of features, facilitating necessary optimizations to enhance the performance of analytical models.

$$IoF_j^{combined} = \alpha IoF_j^{PCA} + (1 - \alpha) IoF_j^{LDA} \quad (14)$$

where α represents the weight of importance given to each method. Finally, after determining the final importance of each factor, decision-makers and experts will decide how many factors to include in the subsequent stage—the experimental design—based on the top K highest values derived from this final importance assessment.

Although PCA and LDA are traditionally derived from different statistical premises—PCA assuming that observations originate from a single population with covariance Σ, and LDA assuming k class-conditional populations with a common Σ—these assumptions merely describe idealized conditions for optimality rather than strict prerequisites for applying the methods. In practical data analysis, particularly in complex industrial or spectroscopic

systems where neither assumption is perfectly satisfied, PCA and LDA are frequently employed together because they capture complementary structures: PCA reveals the dominant variance directions in the overall data cloud, while LDA emphasizes directions that best separate response categories. Numerous studies have demonstrated that combining these projections can improve stability and discrimination without logical inconsistency. For instance, (Sharma et al., 2006) formulated a unified class-dependent PCA - LDA classifier and showed mathematically that integrating the two linear transforms enhances discrimination and mitigates singularity issues in ($\mathbf{S}_W$). (Ali et al., 2022) proposed a feature-fusion PCA–LDA model for plant-leaf disease identification, where PCA features and LDA discriminants were jointly optimized to achieve higher recognition accuracy. Similarly, (Hayati et al., 2024) applied a PCA–LDA–SVM hybrid in chemometric spectroscopy, demonstrating that the combined approach substantially increased class separability in high-dimensional NIR data. These examples confirm that PCA (variance-oriented) and LDA (class-separation-oriented) can be used concurrently as complementary linear projections even when their underlying Gaussian assumptions are only approximate, thereby justifying the present study's use of a parallel PCA–LDA weighting framework for factor importance evaluation.

3.4. Comparative Analysis and Complementarity of PCA and LDA in the Present Study

While PCA and LDA are both multivariate statistical techniques, their analytical objectives differ substantially. PCA is an unsupervised method that transforms correlated input variables into a smaller number of orthogonal principal components that retain the maximum possible variance of the original dataset. This property allows PCA to reveal the intrinsic structure and dominant sources of variability among the process parameters without reference to any output response. In contrast, LDA is a supervised method that determines linear combinations of predictors that best discriminate among predefined classes of the response variable, thereby emphasizing variables that are most influential in distinguishing weld quality levels. In the context of this research, the available production data include nine recorded welding parameters and a corresponding response variable related to weld quality. From an exploratory viewpoint, PCA enables the identification of parameters that contribute most strongly to overall variation in the welding process—such as current, position, speed, and electrode characteristics—thereby highlighting variables that drive general process dynamics. However, when the goal shifts toward understanding how these parameters influence weld quality outcomes, LDA becomes essential, as it explicitly considers the relationship between predictors and categorized response levels.

By combining the two approaches, the analysis benefits from both perspectives: PCA ensures that the selection of key factors is not biased by the specific labeling or discretization of the response variable, while LDA validates and refines this selection by incorporating the response-driven discriminative information. This integrated PCA–LDA framework thus provides a balanced and comprehensive basis for selecting the most influential process parameters before performing the DOE. It enhances the robustness of factor screening and ensures that the subsequent experimental optimization is grounded in both statistical variation and performance relevance.

3.5. 2^K Factorial Design

The following outlines the steps for implementing a 2^K experiment transitioning from a full model to a reduced model.

1. Full model representation

A 2^k factorial experiment explores the effects of K factors, each at two levels (low and high), on a response variable. When blocking is introduced, the response is adjusted for block effects to reduce variability. Including center points allows the detection of curvature in the response surface.

Full model for the experiment, incorporating main effects, interaction effects, and block effects, is expressed as:

$$Y_{ijk} = \mu + \sum_{i=1}^{K} \beta_i X_i + \sum_{i<j} \beta_{ij} X_i X_j + \cdots + \delta_{b(k)} + \epsilon_{ijk} \quad (15)$$

Where Y_{ijk} Observed response for the i-th run in the block k, μ Overall mean, β_i Main effect of the factor i, β_{ij} Interaction effect between factors i and j, $\delta_{b(k)}$ Effect of the block k, and ϵ_{ijk} Random error ($\epsilon_{ijk} \sim N(0, \sigma^2)$).

2. Inclusion of center points

Center points provide a check for nonlinearity or curvature in the model. For k factors, the center point corresponds to the mid-level of each factor. The response at the center is modeled as:

$$Y_{center} = \mu + \delta_{b(k)} + \tau_c + \epsilon \tag{16}$$

where τ_c represents the curvature effect.

3. Reduce the full model

After fitting the full model, terms are iteratively removed to achieve a reduced model without sacrificing predictive power. The steps are:

1. Perform significance tests: Use ANOVA or regression analysis to test the significance of each term.

2. Refine the model: Sequentially remove nonsignificant terms (main effects and interactions) while maintaining the hierarchical order.

3. Final reduced model: The resulting reduced model, retaining only significant terms, can be expressed as:

$$Y_{ijk} = \mu + \sum_{i \in I} \beta_i X_i + \sum_{(i,j) \in J} \beta_{ij} X_i X_j + \delta_{b(k)} + \epsilon_{ijk} \tag{17}$$

where I and J represent the sets of retained main effects and interactions, respectively.

4. Case Study

In this case study, a turbine manufacturing company in Iran is examined, with a particular focus on a specific component, NiCr23Co12Mo, designated for welding processes. An image of welding lines related to the designed experiment is presented in Figure 2. This component is utilized within the mixing chamber of the turbine's combustion chamber, and the welding performed on it is among the most sensitive welding operations employed in turbine manufacturing.

Prior to the design of experiments, production operators meticulously recorded welding settings on the component across nine distinct factors, while simultaneously measuring the response variable of weld depth. The nine factors identified include: current, voltage, speed, position, environment temperature, electrode diameter, heat input, arc length, and preheat temperature. Although other factors associated with machine wear and environmental conditions could potentially be considered, the parameters selected for this study were chosen due to their significance, continuity, and measurability, which are critical for the scope of the research. Utilizing this data, the primary factors influencing weld quality were identified. Based on insights from industry experts, it was determined that conducting experiments with five factors would be more rational, as the costs and time associated with experiments involving all nine factors proved to be impractical. Following the identification of these five effective factors, the design of experiments proceeded to include an additional response variable, known as weld width. Notably, the objective is to maximize weld depth and minimize weld width, reflecting the dual goals of enhancing penetration while reducing excess material spread. Furthermore, during the experimentation phase, the suppliers of the welding electrodes were considered as blocks to account for their effects on the response variables. It is worth mentioning that there are two suppliers of welding electrodes. Ultimately, using the response optimizer in Minitab version 22, optimal settings for the five specified parameters were achieved, maximizing weld depth and minimizing weld width. The subsequent steps of this study will be detailed in the following sections.

Figure 2. A sample welding image related to the conducted designed experiments

4.1. Factor Elimination

In this study, a dataset containing nine influencing factors and 100 samples was previously recorded by operators. Additionally, the response variable, representing the depth of the weld, was documented and required transformation into linguistic variables for application in the LDA method. To facilitate this transformation, seven linguistic variables were utilized: very low, low, low average, average, high average, high, and very high. Ultimately, the significance of the criteria was assessed using both PCA and LDA methods, with the results combined under an equal weighting scheme. Based on the derived importance values, the factors were ranked, leading to the identification of the five most critical factors: current, welding speed, welding position, preheat temperature, and electrode diameter, as illustrated in Table 1. The interval ranges corresponding to each linguistic variable, based on expert judgment, are presented in Table 3 displays the high and low levels of each factor that were determined as a result of the analysis. Thus, the depth of the weld and the width of the weld were the two response variables. In addition to the five selected factors that contributed towards the modelling accuracy, five central points were included in each block to determine the linearity of the model, resulting in a total of 74 experiments. The experiments were performed in random sequence by MINITAB software in the order indicated in the Run Order column. The two response variables were methodically recorded in the relevant columns after completing all experiments and observational data.

Table 2.

Table 1. Combination of factor importance

Factor	PCA importance	LDA importance	Final importance	Rank
Welding speed	0.108936	0.045253	0.077094	1
Current	0.101854	0.036694	0.069274	2
Position	0.121335	0.016355	0.068845	3
Preheat temperature	0.113398	0.019175	0.066286	4
Electrode diameter	0.11695	0.013457	0.065204	5
Heat input	0.117746	0.006719	0.062232	6
Voltage	0.098936	0.024431	0.061683	7
Environment temperature	0.112431	0.010101	0.061266	8

Arc length	0.108414	0.010273	0.059344	9

In combining the PCA- and LDA-based importance scores for factor ranking, equal weighting was applied to avoid bias toward either statistical perspective. PCA and LDA operate under different assumptions—unsupervised variance maximization and supervised class discrimination, respectively—and therefore capture complementary aspects of the data structure. Assigning equal weights ensures that neither purely statistical variance nor response-based discrimination dominates the selection process. This balanced treatment reflects expert consensus within the project team and was judged appropriate for preserving objectivity in factor prioritization. Similar combinations of PCA and LDA have been applied effectively in other domains where both dimensionality reduction and class separation were desirable, such as chromatographic data analysis for gasoline adulteration detection (Skrobot et al., 2007) and image-based classification of agricultural samples (Soleimanipour et al., 2018).

4.2. 2^K Factorial Design

Using the methodology outlined in the previous section, five key factors—current (A), position (B), speed (C), electrode diameter (D), and preheat temperature (E)—were identified as influential variables.

Table 3 displays the high and low levels of each factor that were determined as a result of the analysis. Thus, the depth of the weld and the width of the weld were the two response variables. In addition to the five selected factors that contributed towards the modelling accuracy, five central points were included in each block to determine the linearity of the model, resulting in a total of 74 experiments. The experiments were performed in random sequence by MINITAB software in the order indicated in the Run Order column. The two response variables were methodically recorded in the relevant columns after completing all experiments and observational data.

Table 2. Converting the response variable to a linguistic format for executing the LDA method

Level	Response variable range
very low	[8-10)
low	[10 -12)
low average	[12-14)
average	[14-16)
high average	[16-18)
high	[18-20)
very high	≥ 20

Table 3. Factor levels

Factor	High and low level	Centre point	Measurement unit
Current	80-160	120	Amper
Position	0-45	22.5	Degree
Welding speed	20-40	30	Millimetre/Second
Electrode diameter	2-4	3	Millimetre
Preheat temperature	25-100	62.5	Celsius degree

4.2.1. First Response Variable (Depth)

The results for the first response variable, which represents the weld depth, are systematically summarized and presented in Table 4, providing a detailed overview of the observed data and their corresponding analysis.

Table 4. Codded coefficients (y_1)

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		16.061	0.137	117.20	0.000	
Blocks						
1		0.254	0.127	1.99	0.053	1.00
Current	3.734	1.867	0.137	13.63	0.000	1.00
Position	-0.953	-0.477	0.137	-3.48	0.001	1.00
Speed	-0.672	-0.336	0.137	-2.45	0.019	1.00
Diameter	0.222	0.111	0.137	0.81	0.423	1.00
Temperature	0.347	0.173	0.137	1.27	0.213	1.00
Current*Position	-1.041	-0.520	0.137	-3.80	0.000	1.00
Current*Speed	-0.184	-0.092	0.137	-0.67	0.505	1.00
Current*Diameter	-0.103	-0.052	0.137	-0.38	0.709	1.00
Current*Temperature	-0.091	-0.045	0.137	-0.33	0.743	1.00
Position*Speed	0.441	0.220	0.137	1.61	0.116	1.00
Position*Diameter	0.384	0.192	0.137	1.40	0.168	1.00
Position*Temperature	-0.516	-0.258	0.137	-1.88	0.067	1.00
Speed*Diameter	-0.097	-0.048	0.137	-0.35	0.726	1.00
Speed*Temperature	0.166	0.083	0.137	0.60	0.549	1.00
Diameter*Temperature	0.147	0.073	0.137	0.54	0.595	1.00
Current*Position*Speed	0.228	0.114	0.137	0.83	0.410	1.00
Current*Position*Diameter	-0.241	-0.120	0.137	-0.88	0.385	1.00
Current*Position*Temperature	0.447	0.223	0.137	1.63	0.111	1.00
Current*Speed*Diameter	0.028	0.014	0.137	0.10	0.919	1.00
Current*Speed*Temperature	-0.022	-0.011	0.137	-0.08	0.937	1.00
Current*Diameter*Temperature	-0.278	-0.139	0.137	-1.01	0.316	1.00
Position*Speed*Diameter	-0.022	-0.011	0.137	-0.08	0.937	1.00
Position*Speed*Temperature	0.266	0.133	0.137	0.97	0.338	1.00
Position*Diameter*Temperature	-0.128	-0.064	0.137	-0.47	0.643	1.00
Speed*Diameter*Temperature	-0.059	-0.030	0.137	-0.22	0.830	1.00
Current*Position*Speed*Diameter	0.103	0.052	0.137	0.38	0.709	1.00
Current*Position*Speed*Temperature	-0.372	-0.186	0.137	-1.36	0.182	1.00
Current*Position*Diameter*Temperature	0.147	0.073	0.137	0.54	0.595	1.00
Current*Speed*Diameter*Temperature	-0.159	-0.080	0.137	-0.58	0.564	1.00
Position*Speed*Diameter*Temperature	0.003	0.002	0.137	0.01	0.991	1.00

Current*Position*Speed*Diameter*Temperature	-0.047	-0.023	0.137	-0.17	0.865	1.00
Ct Pt		-0.111	0.373	-0.30	0.768	1.00

The $p - value_{Ct-Pt} = 0.768$ for the test of second-degree curvature indicates the absence of significant curvature in the response variable. This result supports the assumption of linearity in the model, allowing for the removal of higher-order terms in the reduced version. The adequacy of the linear model was verified through a curvature test using replicated center points, as outlined by (Montgomery, 2017). The statistical analysis revealed that the difference between the average responses at factorial and center points was statistically insignificant ($p > 0.05$), indicating no evidence of curvature. Consequently, the first-order model was deemed sufficient to describe the relationship between factors and the response within the studied region. Given that the curvature test provides a formal test for nonlinearity, the inclusion of higher-order or kernel-based models was not justified for this dataset. To simplify the model, it is essential to identify the significant and impactful factors while eliminating the non-essential ones. It is clear from Table 4 that factors A, B, C, as well as the two-way interaction AB should stay in the model. The central points would be removed from the reduced model, as the resulting p-value was not statistically significant. However, regarding the block, as the p-value for the block is 0.053 (very close to 0.05, which is the threshold for removal), it was preferred to include the block factor in this reduced model. This choice is made to alleviate the variability due to the electrode supplier changing. The ANOVA table after pruning the model is given in **Error! Not a valid bookmark self-reference.**.

Table 5. Analysis of Variance for reduced model (y_1)

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	5	266.989	53.398	49.58	0.000
Blocks	1	4.776	4.776	4.43	0.039
Linear	3	244.887	81.629	75.79	0.000
Current	1	223.129	223.129	207.18	0.000
Position	1	14.535	14.535	13.50	0.000
Speed	1	7.223	7.223	6.71	0.012
2-Way Interactions	1	17.326	17.326	16.09	0.000
Current*Position	1	17.326	17.326	16.09	0.000
Error	68	73.234	1.077		
Curvature	1	0.106	0.106	0.10	0.756
Lack-of-Fit	59	64.972	1.101	1.08	0.498
Pure Error	8	8.156	1.019		
Total	73	340.224			

In the reduced model obtained, all the considered terms, including the block effect, are statistically significant. This indicates that, in addition to factors A, B, C, and the interaction term AB, the selection of the electrode supplier also influences the achieved weld depth. To verify whether any significant effect has been inadvertently excluded from the initial model, the p-value for the lack-of-fit test is examined. Given the obtained value of 0.498, it can be concluded

that the hypothesis suggesting the omission of a significant effect in the reduced model is rejected. To further assess the adequacy of the model, the residual plots were then examined.

A detailed analysis of the main effects plot (Figure 3) for weld depth (y_1) reveals that current is the dominant factor influencing the response variable, showing a clear and consistent increase in depth as current increases. This indicates a direct relationship between current and weld depth, where higher current values result in a deeper weld. In contrast, position and speed exhibit minimal effects, suggesting that varying these factors within the experimental range does not significantly influence the weld depth. Furthermore, the interaction plot (Figure 4) for weld depth (y_1) shows a pronounced interaction between current and position. Weld depth increases significantly with rising current, particularly at higher amperage levels. This suggests that optimizing current and position together is crucial for achieving the desired weld depth, with the maximum depth reached at the highest current level and the lowest position level. The analysis emphasizes the critical role of current in achieving the optimal weld depth, while position has an important but secondary interaction effect.

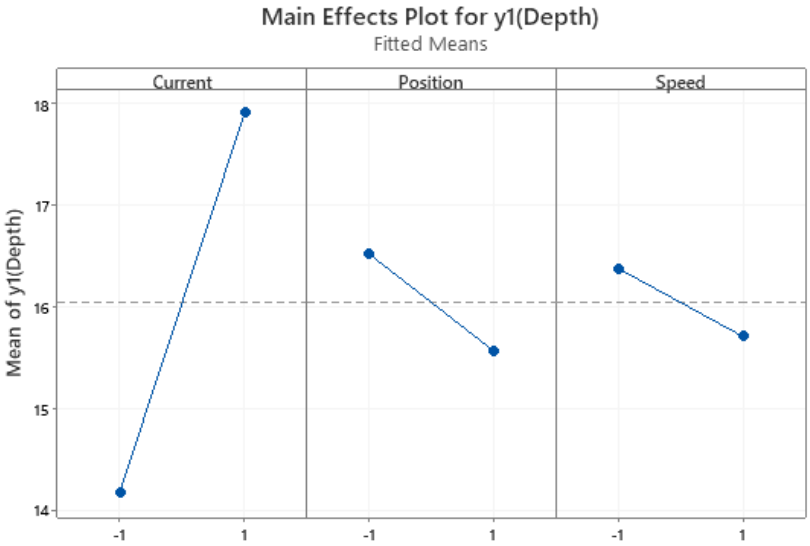

Figure 3. Main effect plot (y_1)

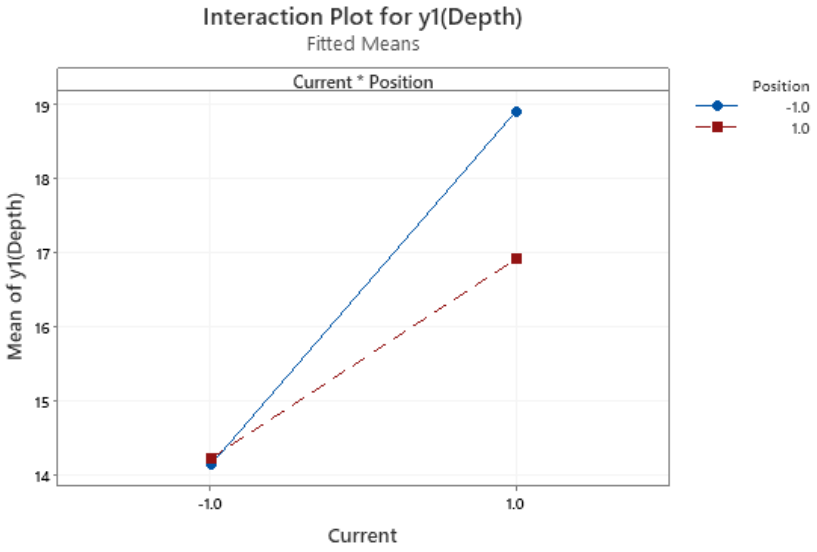

Figure 4. Interaction effect plot (y_1)

Following the lack-of-fit assessment, residual diagnostic plots were examined to further evaluate the adequacy of the fitted model. As shown in Figure 5, the residual plots for y_1 (Depth) provide a comprehensive assessment of the residuals' distribution, including the normal probability plot, histogram, and plots of residuals versus fitted values and observation order. These plots suggest that the residuals are generally normally distributed, with only a few minor deviations. To formally verify the normality assumption, the Kolmogorov-Smirnov (K-S) test (Massey Jr, 1951) was also conducted on the residuals, and the results indicated that the residuals are indeed normally distributed. Furthermore, Figure 6 displays the synthetic Cumulative Distribution Function (CDF) of the residuals for y_1 (Depth). The empirical CDF (blue bars) aligns closely with the normal CDF (red line), offering strong evidence that the residuals follow a normal distribution. While individual data points such as observations 3, 7, and 61 show slight deviations from the normal curve, these points remain within acceptable limits. The residuals versus fits plot, shown in Figure 5, further supports the independence of the residuals, although a slight increase in residual variance between predicted values of 14 and 15 suggests a minor inconsistency in the spread of the residuals. Despite this, the residuals' overall behaviour confirms that the model is adequate and valid, satisfying the normality assumption.

Furthermore, Figure 6 displays the synthetic cumulative distribution function (CDF) of the residuals for y_1 (Depth). The empirical CDF (blue bars) aligns closely with the normal CDF (red line), offering strong evidence that the residuals follow a normal distribution. While individual data points such as observations 3, 7, and 61 show slight deviations from the normal curve, these points remain within acceptable limits. The residuals versus fits plot, shown in Figure 5, further supports the independence of the residuals, although a slight increase in residual variance between predicted values of 14 and 15 suggests a minor inconsistency in the spread of the residuals. Despite this, the residuals' overall behaviour confirms that the model is adequate and valid, satisfying the normality assumption.

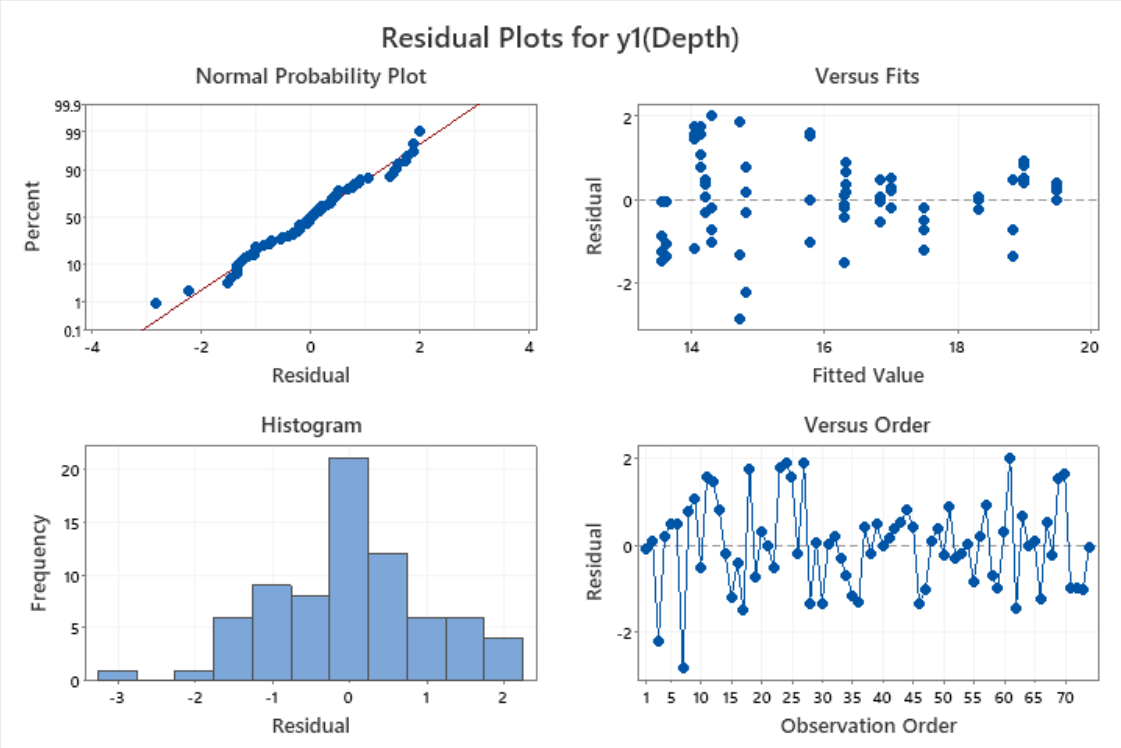

Figure 5. Residual for reduced model (y_1)

Finally, the cube plot for the weld depth response variable, as shown in Figure 7, is investigated to understand under which conditions the weld depth reaches its highest value. According to the fitted equation, the maximum weld depth occurs when the current is at its high level of 160 amps. Due to the negative coefficients of speed and welding position in the equation, they should be kept at their respective low levels to increase the weld depth. In other words, higher depth values are attained at a welding speed of 20 millimetres per second with a welding position of zero degree.

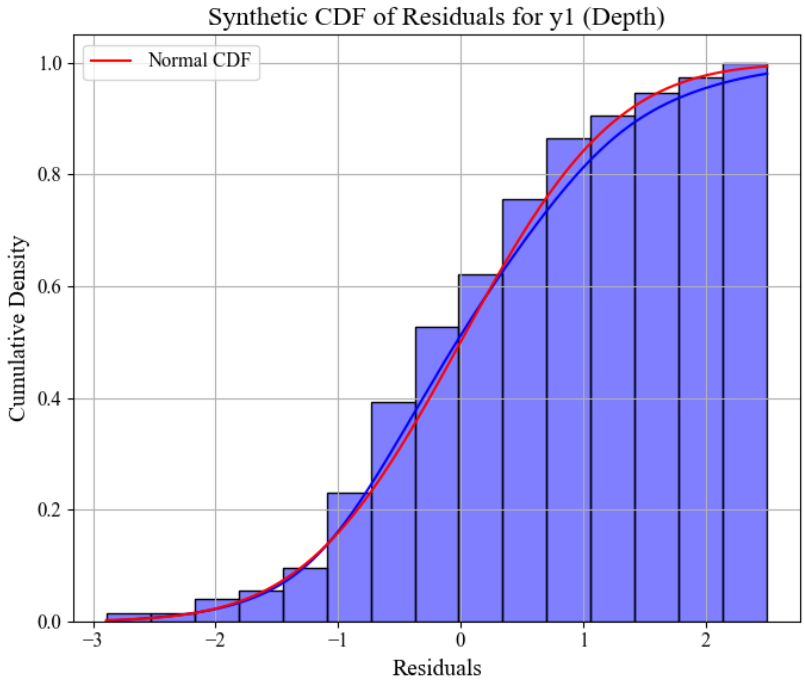

Figure 6. CDF of residuals for y_1

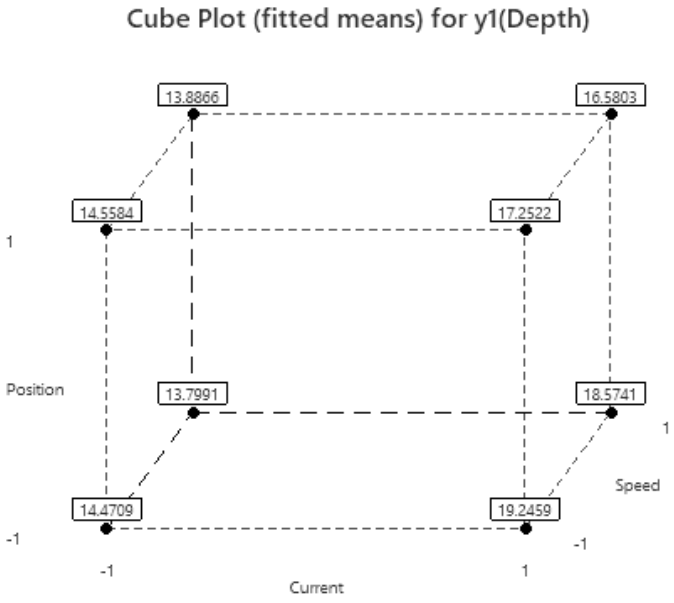

Figure 7. Cube plot for reduced model (y_1)

4.2.2. Second Response Variable (Width)

Regarding the second response variable, which is weld width, by considering the block and central points and taking all terms into account, the obtained results are as shown in Table 6.

Table 6. Codded coefficients table (y_2)

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		8.966	0.113	79.04	0.000	
Blocks						
1		0.191	0.105	1.81	0.078	1.00
Current	3.363	1.681	0.113	14.82	0.000	1.00
Position	-0.912	-0.456	0.113	-4.02	0.000	1.00
Speed	-0.687	-0.344	0.113	-3.03	0.004	1.00
Diameter	0.156	0.078	0.113	0.69	0.495	1.00
Temperature	0.312	0.156	0.113	1.38	0.176	1.00
Current*Position	-0.919	-0.459	0.113	-4.05	0.000	1.00
Current*Speed	-0.219	-0.109	0.113	-0.96	0.341	1.00
Current*Diameter	-0.138	-0.069	0.113	-0.61	0.548	1.00
Current*Temperature	0.019	0.009	0.113	0.08	0.935	1.00
Position*Speed	0.494	0.247	0.113	2.18	0.035	1.00
Position*Diameter	0.275	0.137	0.113	1.21	0.233	1.00
Position*Temperature	-0.456	-0.228	0.113	-2.01	0.051	1.00
Speed*Diameter	-0.138	-0.069	0.113	-0.61	0.548	1.00
Speed*Temperature	0.106	0.053	0.113	0.47	0.642	1.00
Diameter*Temperature	0.012	0.006	0.113	0.06	0.956	1.00
Current*Position*Speed	0.225	0.113	0.113	0.99	0.327	1.00
Current*Position*Diameter	-0.106	-0.053	0.113	-0.47	0.642	1.00
Current*Position*Temperature	0.313	0.156	0.113	1.38	0.176	1.00
Current*Speed*Diameter	0.081	0.041	0.113	0.36	0.722	1.00
Current*Speed*Temperature	-0.025	-0.013	0.113	-0.11	0.913	1.00
Current*Diameter*Temperature	-0.181	-0.091	0.113	-0.80	0.429	1.00
Position*Speed*Diameter	0.081	0.041	0.113	0.36	0.722	1.00
Position*Speed*Temperature	0.213	0.106	0.113	0.94	0.355	1.00
Position*Diameter*Temperature	-0.019	-0.009	0.113	-0.08	0.935	1.00
Speed*Diameter*Temperature	-0.019	-0.009	0.113	-0.08	0.935	1.00
Current*Position*Speed*Diameter	0.012	0.006	0.113	0.06	0.956	1.00
Current*Position*Speed*Temperature	-0.306	-0.153	0.113	-1.35	0.185	1.00
Current*Position*Diameter*Temperature	0.050	0.025	0.113	0.22	0.827	1.00
Current*Speed*Diameter*Temperature	-0.225	-0.113	0.113	-0.99	0.327	1.00
Position*Speed*Diameter*Temperature	0.050	0.025	0.113	0.22	0.827	1.00
Current*Position*Speed*Diameter*Temperature	-0.044	-0.022	0.113	-0.19	0.848	1.00
Ct Pt		-0.176	0.309	-0.57	0.572	1.00

The p-value analysis indicates that the response variable of weld width, similar to the previously examined response variable of weld depth, does not exhibit curvature. Consequently, the assumption of linearity in the model is not rejected. Based on the results obtained in the reduced model, the central points are removed. A notable distinction between this response variable and the previous case is the insignificant p-value for the block effect, suggesting that the change in electrode supplier does not have a statistically significant impact on weld width. In contrast, for the weld depth response variable, this value was very close to the significance threshold and was ultimately retained in the reduced model. Therefore, it can be inferred that while variations in electrode supplier influence weld depth, they do not significantly affect weld width, implying that supplier changes can essentially be considered inconsequential in determining weld width variations. As illustrated in Table 6, factors A, B, and C, along with the two-way interaction terms BC and AB, should be retained in the model. As mentioned earlier, the central points will not be included in the reduced model due to their high p-value. Additionally, given the p-value for the block effect (p-value $_{\text{Block}} = 0.078$) and its considerable distance from the significance threshold, it will also not be included in the reduced model, based on the reasoning outlined earlier. The ANOVA table following model reduction is presented in Table 7.

Table 7. Analysis of Variance for reduced model (y_2)

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	5	219.194	43.839	59.14	0.000
Linear	3	201.788	67.263	90.74	0.000
Current	1	180.903	180.903	244.04	0.000
Position	1	13.322	13.322	17.97	0.000
Speed	1	7.563	7.563	10.20	0.002
2-Way Interactions	2	17.406	8.703	11.74	0.000
Current*Position	1	13.506	13.506	18.22	0.000
Position*Speed	1	3.901	3.901	5.26	0.025
Error	68	50.406	0.741		
Curvature	1	0.267	0.267	0.36	0.552
Lack-of-Fit	59	41.632	0.706	0.66	0.826
Pure Error	8	8.508	1.063		
Total	73	269.600			

For weld width (y_2), the main effects plot is shown in Figure 8 clearly indicates that current is the most influential parameter, as increasing current significantly increases weld width. Since the objective is to minimize weld width, lower current settings are preferred. Welding position also shows a moderate effect—weld width decreases slightly as the position shifts from low (-1) to high (1), suggesting that optimal positioning can help reduce weld width. Welding speed exhibits a minor influence, with higher speeds slightly reducing width, though its overall impact is less pronounced. These findings suggest that effective control of weld width can be primarily achieved by minimizing current, while position serves a secondary role in fine-tuning the result. Speed has the least effect and may be adjusted based on other process considerations.

The interaction plot for weld width (y_2), shown in Figure 9, reveals important relationships between process parameters. The most notable interaction occurs between current and position, where increasing current results in a substantial increase in weld width, particularly at the lower position setting (-1). However, at the higher position setting (1), the increase in width due to the current is less pronounced. This suggests that operating at a higher welding position can help mitigate the width-increasing effect of high current, making it an effective strategy for controlling weld width. The current-speed interaction shows minimal variation, suggesting that speed does not significantly influence the effect of current on width. Likewise, the position-speed interaction reveals a slight reduction in width at higher levels of both parameters, but the effect is relatively minor. Overall, the interaction effects suggest that to minimize weld width, the most effective approach is to use lower current settings and higher welding positions, while welding speed plays a secondary, limited role.

Figure 8. Main effect plot (y_2)

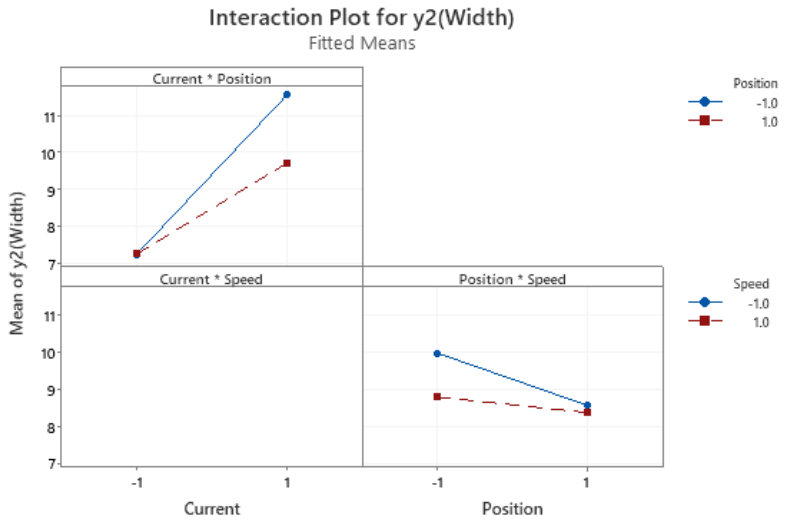

Figure 9. Interaction effect plot (y_2)

Subsequently, the residuals were analysed to evaluate the validity of the model. As shown in Figure 10, the residual plots for y_2 (Width) provide a clear indication that the residuals generally follow a normal distribution, with no major discrepancies. To formally verify this assumption, Figure 11 displays the synthetic CDF of the residuals for y_2 (Width). The empirical CDF (green bars) closely aligns with the normal CDF (red line), confirming that the residuals are normally distributed. However, four specific data points—observations 18, 24, 25, and 61—appear as outliers, particularly observation 61, which is an outlier for both the weld width and weld depth response variables. Given this, it is recommended to repeat the experiment for observation 61 and obtain new response values for verification. The residuals versus fits plot, depicted in Figure 10, shows no apparent pattern, further confirming the independence of the residuals. It is, however, worth noting that the variance of the residuals is slightly higher around the predicted values of 6.6 to 9 compared to other regions, suggesting a minor tendency toward heteroscedasticity. Despite this slight indication of unequal variance, the overall analysis of the residuals supports the model's suitability and validity, with no major issues that would invalidate its use.

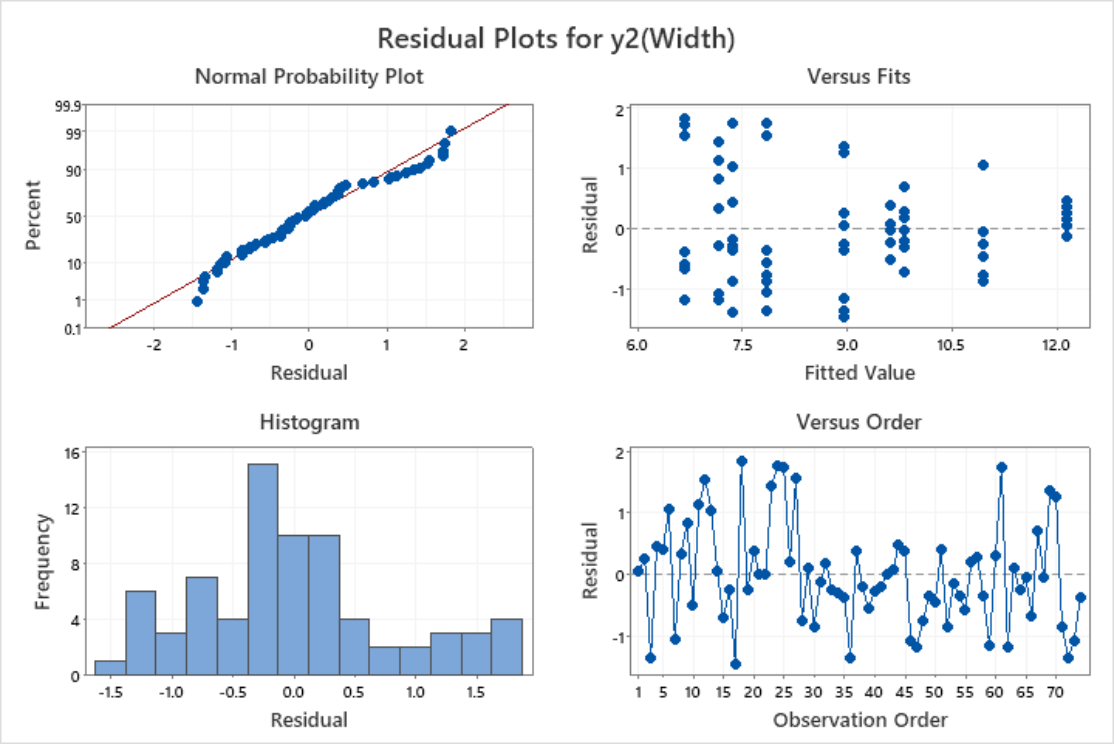

Figure 10. Residual for reduced model (y_2)

Finally, the cube plot for the weld width response variable, as shown in Figure 12, the process parameters that result in the minimum weld width were examined. The analysis indicates that the narrowest weld width is achieved when the current intensity is set to its lowest value of 80 amperes. Furthermore, to minimize the weld width, the welding speed and position should be maintained at their highest levels, specifically at a speed of 40 mm/s and a welding angle of 45 degrees.

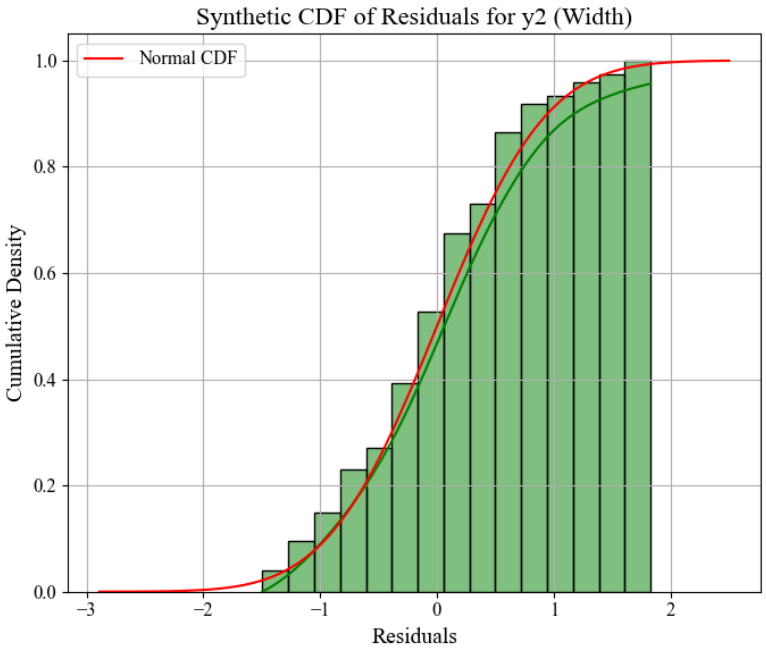

Figure 11. CDF of residuals for y_2

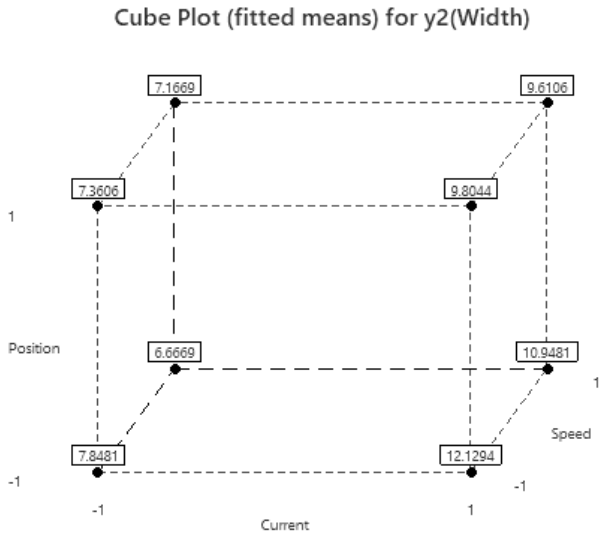

Figure 12. Cube plot for reduced model (y_2)

4.3. Response optimizer

Multi-response optimization was achieved to reach optimal welding performance in terms of many quality characteristics, using the response optimizer in Minitab. The analysis focused on two major responses: weld depth, which was to be maximized, and weld width, which was to be minimized. After the statistical screening, only three parameters were found to be statistically significant among the input factors and hence subjected to optimization: current, position, and welding speed. Thus, all variables were coded to standardized levels for interpretability and consistency in the analysis. An interesting application feature of the response optimizer is that it allows the user to assign weights to each response to create trade-off strategies depending upon process objectives. To ensure whether or not the optimization results were sensitive to the weight assignments, a systematic sensitivity analysis was conducted

by varying the weights (w_1 for weld depth, and w_2 for weld width) over a series of nine different combinations. For each combination, optimal values were set for the three process parameters. The fitted values of the two responses under each weighting scenario are presented in Figure 13, illustrating how changes in the relative importance of each objective affect the predicted weld characteristics. Weld depth (blue line) shows fairly stable, albeit oscillatory, behavior and achieves maximum values at w_1=3, w_2=1; this means that with an increased priority given to depth, it can be considerably improved. Weld width (red line) shows opposite trends-it has its lower values as its weight increases, especially at w_1=1, w_2=3- thus confirming the trade-off expected between the two objectives. These trends highlight the inherent conflict: improving one response typically leads to the deterioration of the other.

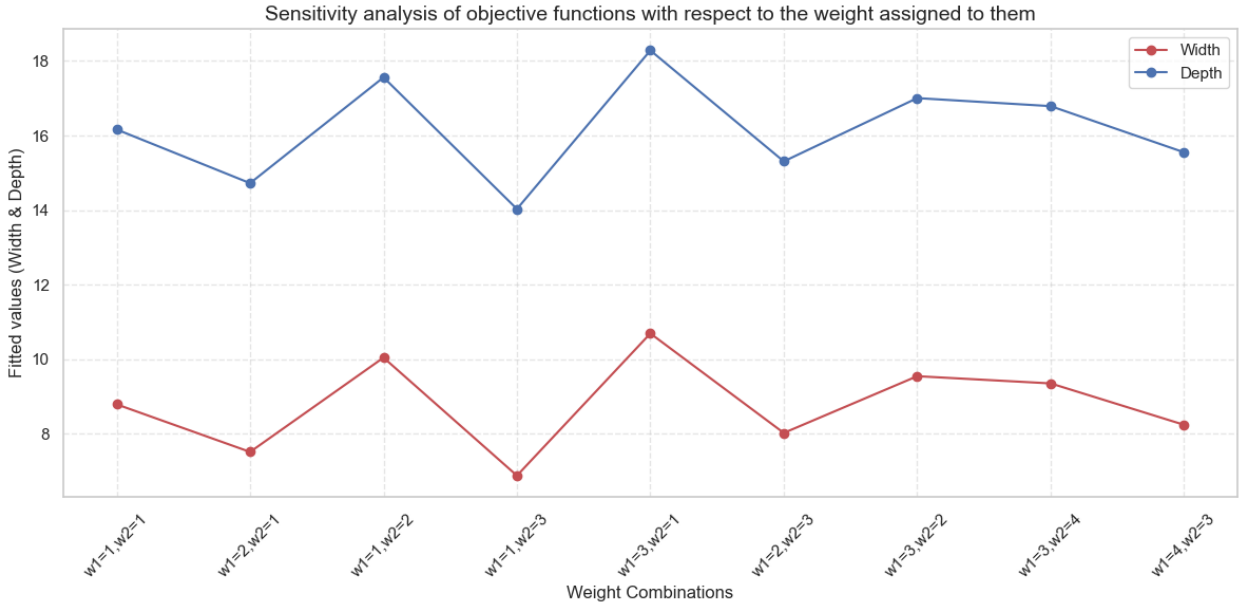

Figure 13. Sensitivity analysis of process optimization

To further elucidate the influence of factor settings under each weight scenario, the corresponding optimization profiles generated by the response optimizer are shown in Figure 14. Each subpanel presents the predicted response values and associated desirability scores for weld depth and width, as well as the optimal levels of the three process factors. The profiles reveal precise adjustments in the parameter settings as the optimization priorities shift. Notably, position and speed tend to converge toward their extreme levels (e.g., position at -1 and speed at +1) in configurations favoring weld depth, suggesting a more consistent role in influencing penetration. Meanwhile, current appears to adjust more flexibly across scenarios, serving as a balancing lever between the conflicting objectives. In situations where both response variables hold equal significance, the current should be adjusted to 119.596 amps, the welding angle should be set to zero degree, and the welding speed should be maintained at 40 mm/s. This configuration will yield a weld depth of 16.1625 mm and a weld width of 8.7858 mm. The composite desirability index (D), which combines the individual desirability scores into a single optimization metric, varies from 0.1158 to 0.5530 across the weight scenarios. This wide range indicates the substantial impact of weight assignment on the overall quality of optimization. Intermediate weight combinations tend to yield more balanced composite scores, while extreme prioritization of a single response significantly reduces the composite desirability.

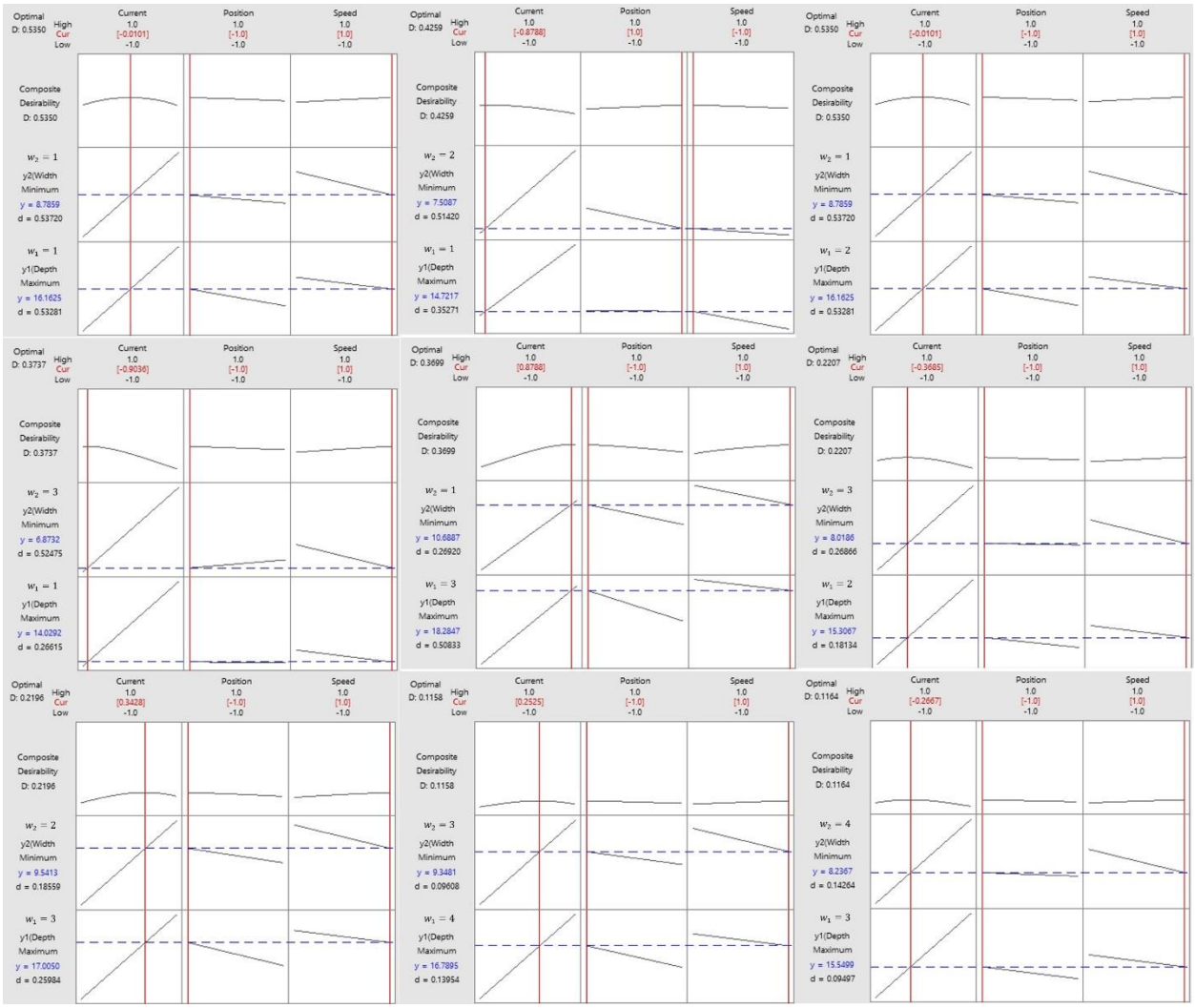

Figure 14. Response optimizer for different weight combinations

4.4. Validation of the optimization results

Historical production data from the case study company were used to demonstrate the effectiveness of the proposed optimization methodology. This data set consisted of measured values of weld depth and weld width under varying operational settings that had previously been used in the actual manufacturing process. The average values derived from this historical data served as a baseline for performance comparison. Three distinct optimization scenarios—generated using the response optimizer—were evaluated: (1) equal weighting of both responses ($w_1 = 1$, $w_2 = 1$), (2) depth-prioritized optimization ($w_1 = 3$, $w_2 = 1$), and (3) width-prioritized optimization ($w_1 = 1$, $w_2 = 3$). Each scenario produced a set of optimal factor levels and corresponding fitted values for the two responses. These predicted values were compared against the historical averages, and the percentage improvement in weld depth and reduction in weld width were calculated. The results obtained, which are presented in **Error! Reference source not found.**, verify the fact that a meaningful improvement was attained with the three scenarios, with respect to the historical achievements. The equal-weight scenario improved both quality parameters equally. The depth and width prioritized scenarios achieved their respective optimization goals, that is, to maximize penetration and minimize weld spread. These results

substantiate the validity and usefulness of the developed multi-response optimization framework in enhancing welding performance under various production priorities.

Table 8. Comparison of historical data with optimized results under different weighting scenarios

Scenario	Weight (w_1, w_2)	Weld Depth (mm)	% Depth Improvement	Weld Width (mm)	% Width Reduction
Historical Average	—	14.20	—	9.10	—
Equal Weights	(1, 1)	16.16	+13.82%	8.78	−3.52%
Depth-Prioritized	(3, 1)	18.28	+28.77%	10.68	+17.38%
Width-Prioritized	(1, 3)	14.02	-1.27%	6.87	−24.55%
Average	-	-	+13.77%	-	-3.56%

5. Practical Assessment of Scrap Rate and Cost Impact

To evaluate the practical effects of the optimized welding parameters, production performance was examined over a six-month period, consisting of three months before and three months after implementation. Approximately twenty components of this type are welded each month, and each defective part incurs substantial material, labor, and inspection costs.

As shown in **Error! Reference source not found.**, the monthly average scrap rate decreased from about 19.8% in the baseline stage to approximately 13.1% after optimization, representing a relative reduction of nearly 6%. This improvement led to an estimated 3–4 % reduction in total production cost, reflecting the high unit value of the welded component. The trend of cost improvement presented in

Figure **16** demonstrates a steady decline in losses following adoption of the optimized parameters. These findings confirm that the proposed optimization not only enhanced weld geometry but also contributed directly to improved manufacturing efficiency and cost effectiveness at the plant level.

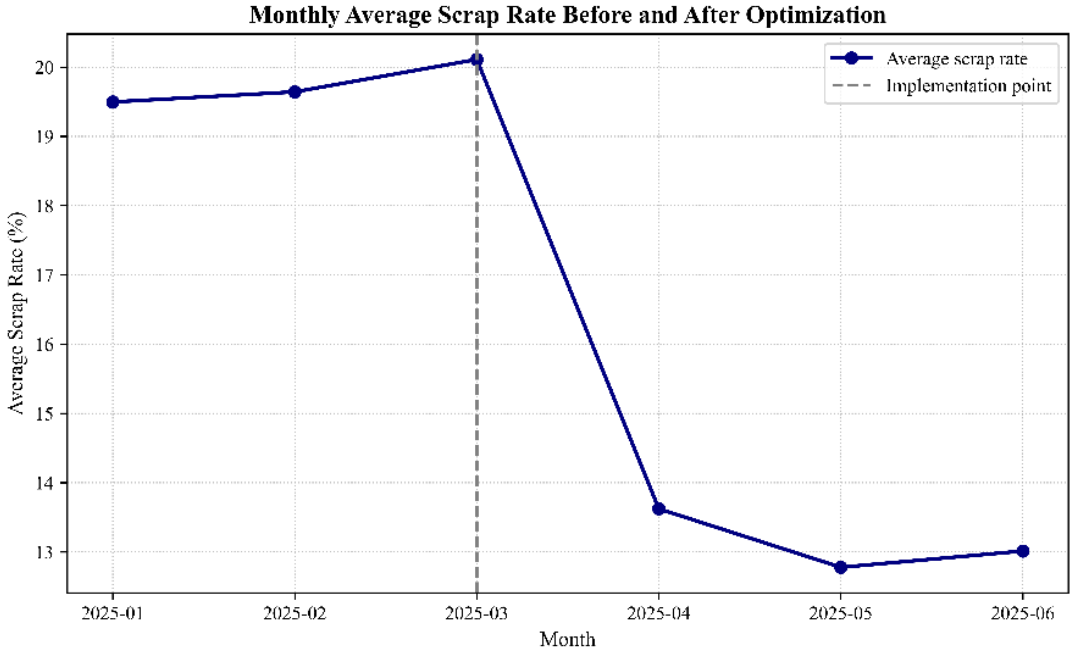

Figure 15. Monthly average scrap rate before and after optimization

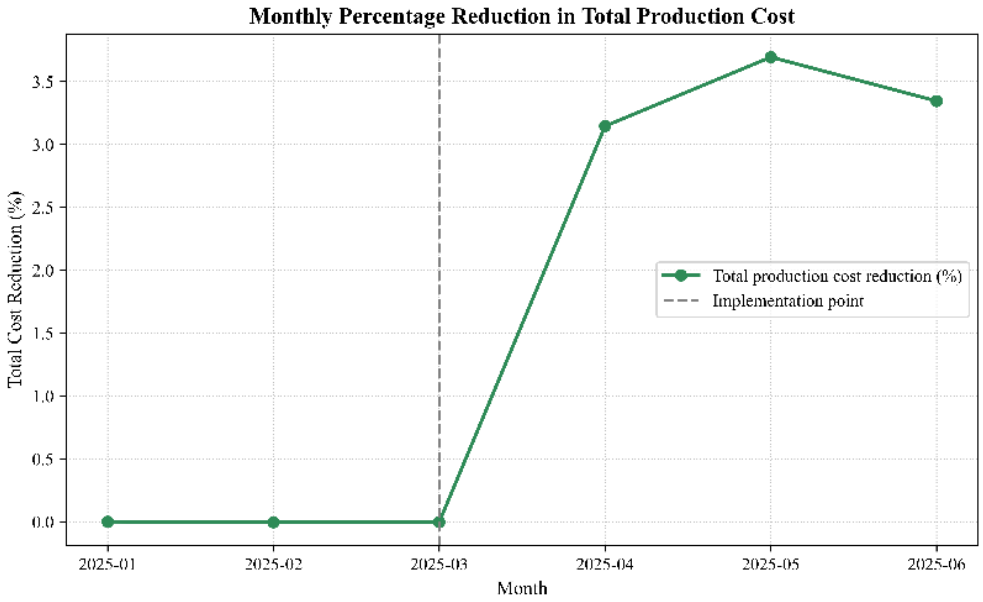

Figure 15. Cumulative cost savings following implementation of optimized settings

6. Conclusion

This study establishes a robust methodology by integrating statistical techniques—specifically PCA and LDA—with a structured 2^K factorial design of experiments to optimize welding parameters. The hybrid approach effectively identified current, position, and speed as dominant factors governing weld quality, with qualitative analysis revealing that current serves as a primary adjustable lever for balancing the inherent trade-off between depth maximization and width minimization. Validation demonstrated significant improvements: under industry-standard balanced weighting (w_1=w_2=1), weld depth increased by 13.8% while width decreased by 3.5%; for depth-prioritized weighting (w_1=3, w_2=1 – selected as the extreme achievable depth from sensitivity analysis), depth surged by 28.8%; and under width-prioritized weighting (w_1=1, w_2=3 – representing the extreme achievable width reduction), width diminished by 24.55% with minimal depth trade-off. These scenarios collectively map the pareto frontier of achievable outcomes, with the composite desirability range (0.1158–0.5530) quantifying the framework's adaptability to diverse production goals. Future work should extend this methodology to real-time adaptive welding systems and applications involving dissimilar materials.

Although this study was conducted on a specific industrial case involving the NiCr23Co12Mo alloy in turbine manufacturing, the proposed data-driven and experimental workflow is not limited to this material or context. Given that design of experiments can be costly in other manufacturing, the same approach could be effectively applied to other alloys and welding processes, such as laser beam welding, gas metal arc welding, friction stir welding, or resistance spot welding. Extending and validating this framework across different materials and joining methods represents an interesting direction for future research and industrial application.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Ethical considerations

The authors avoided data fabrication, falsification, and plagiarism, and any form of misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

References

Ahmed, M. M. Z., Touileb, K., El-Sayed Seleman, M. M., Albaijan, I., & Habba, M. I. A. (2022). Bobbin tool friction stir welding of aluminum: parameters optimization using taguchi experimental design. *Materials*, *15*(8), 2771. https://doi.org/10.3390/ma15082771

Ali, S., Hassan, M., Kim, J. Y., Farid, M. I., Sanaullah, M., & Mufti, H. (2022). FF-PCA-LDA: intelligent feature fusion based PCA-LDA classification system for plant leaf diseases. *Applied Sciences*, *12*(7), 3514. https://doi.org/10.3390/app12073514

Anand, K., Barik, B. K., Tamilmannan, K., & Sathiya, P. (2015). Artificial neural network modeling studies to predict the friction welding process parameters of Incoloy 800H joints. *Engineering Science and Technology, an International Journal*, *18*(3), 394–407.

Antony, J. (2000). Multi-response optimization in industrial experiments using Taguchi's quality loss function and principal component analysis. *Quality and Reliability Engineering International*, *16*(1), 3–8.

Babanova, S., Artyushkova, K., Ulyanova, Y., Singhal, S., & Atanassov, P. (2014). Design of experiments and principal component analysis as approaches for enhancing performance of gas-diffusional air-breathing bilirubin oxidase cathode. *Journal of Power Sources*, *245*, 389–397. https://doi.org/10.1016/j.jpowsour.2013.06.031

Baraya, M., El-Asfoury, M. S., Fadel, O. O., & Abass, A. (2024). Experimental Analyses and Predictive Modelling of Ultrasonic Welding Parameters for Enhancing Smart Textile Fabrication. *Sensors*, *24*(5), 1488. https://doi.org/10.3390/s24051488

Boriwal, L., Sarviya, R. M., & Mahapatra, M. M. (2021). Process analysis and regression modelling of resistance spot welded joints of austenitic stainless steel 304L and low carbon steel sheets by using surface response methodology. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, *235*(1), 24–33. https://doi.org/10.1177/0954408920940888

Breitenbach, J., Dauser, T., Illenberger, H., Traub, M., & Buettner, R. (2021). A Systematic Literature Review on Machine Learning Approaches for Quality Monitoring and Control Systems for Welding Processes. *2021 IEEE International Conference on Big Data (Big Data)*, 2019–2025.

Camarena-Martinez, R., Baeza-Serrato, R., & Lizarraga-Morales, R. A. (2023). Optimization of Welding Process of Geomembranes in Biodigesters Using Design of Factorial Experiments. *Energies*, *16*(18), 6583. https://doi.org/10.3390/en16186583

Choudhary, D. K. C. D., Pandey, R. M., Prakash, C., & Kumar, V. (2024). Mathematical Modeling of Shape relationship including% Dilution and Total Bead Volume in SAW Process by using Two Level Half Factorial Technique. *Journal of Manufacturing Engineering*, *19*(4), 108–125. https://doi.org/10.37255/jme.v19i4pp108-125

Das, S., AnanthaKrishna, K. V, Rajkumar, J. V, Kumar, B., Arora, U. K., & Tewari, R. (2025). A Study on the Effects of Welding Parameters on Weld Bead Geometry in TIG Welding Process of Zircaloy Fuel Pins

Using Response Surface Methodology. *Transactions of the Indian Institute of Metals*, *78*(2), 40. https://doi.org/10.1007/s12666-024-03484-9

Guo, H., Lin, L., Lv, Y., Liu, J., & Tong, C. (2021). Machine Learning for Determining Key Parameters in Welding Process of Underground Engineering Equipment. *2021 IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, 33–41. https://doi.org/10.1109/SDPC52933.2021.9563365

Haribalaji, V., Boopathi, S., & Asif, M. M. (2022). Optimization of friction stir welding process to join dissimilar AA2014 and AA7075 aluminum alloys. *Materials Today: Proceedings*, *50*, 2227–2234. https://doi.org/10.1016/j.matpr.2021.09.499

Hayati, R., Munawar, A. A., Lukitaningsih, E., Earlia, N., Karma, T., & Idroes, R. (2024). Combination of PCA with LDA and SVM classifiers: A model for determining the geographical origin of coconut in the coastal plantation, Aceh Province, Indonesia. *Case Studies in Chemical and Environmental Engineering*, *9*, 100552. https://doi.org/10.1016/j.cscee.2023.100552

He, Y., Yang, K., Wang, X., Huang, H., & Chen, J. (2022). Quality prediction and parameter optimisation of resistance spot welding using machine learning. *Applied Sciences*, *12*(19), 9625. https://doi.org/10.3390/app12199625

Jagtap, K. R., Rojekar, M. S., Dravid, S. V, & Deshpande, A. R. (2017). Effect of welding parameters on tensile & yield strength of is 2062 grade steel using design of experiment approach. *Materials Today: Proceedings*, *4*(8), 7875–7883. https://doi.org/10.1016/j.matpr.2017.07.123

Khoshnevisan, L., Hourfar, F., Alhameli, F., & Elkamel, A. (2021). Combining design of experiments, machine learning, and principal component analysis for predicting energy consumption and product quality of a natural gas processing plant. *International Journal of Energy Research*, *45*(4), 5974–5987. https://doi.org/10.1002/er.6217

Kumar, R. R., Singh, A., Kumar, A., Ansu, A. kumar, Kumar, A., Kumar, S., Kumar, D., Goyal, A., Oza, A. D., & Singh, D. (2023). Enhancement of friction stir welding characteristics of alloy AA6061 by design of experiment methodology. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *17*(5), 2659–2671. https://doi.org/10.1007/s12008-022-01106-6

Kumar, S., Gaur, V., & Wu, C. (2022). Machine learning for intelligent welding and manufacturing systems: research progress and perspective review. *The International Journal of Advanced Manufacturing Technology*, *123*(11), 3737–3765. https://doi.org/10.1007/s00170-022-10403-z

Kuo, C.-C., Tsai, Q.-Z., Li, D.-Y., Lin, Y.-X., & Chen, W.-X. (2022). Optimization of Ultrasonic Welding Process Parameters to Enhance Weld Strength of 3C Power Cases Using a Design of Experiments Approach. *Polymers*, *14*(12), 2388. https://doi.org/10.3390/polym14122388

Lee, W., Jung, S., Noh, H.-S., & Park, M. (2010). Reliable feature extraction using linear discriminant analysis. *SCIS & ISIS SCIS & ISIS 2010*, 1233–1236.

Li, J., Zhao, B., & Zhang, H. (2009). Face recognition based on PCA and LDA combination feature extraction. *2009 First International Conference on Information Science and Engineering*, 1240–1243. https://doi.org/10.1109/ICISE.2009.581

Madrid, J., Lorin, S., Söderberg, R., Hammersberg, P., Wärmefjord, K., & Lööf, J. (2019). A virtual design of experiments method to evaluate the effect of design and welding parameters on weld quality in aerospace applications. *Aerospace*, *6*(6), 74. https://doi.org/10.3390/aerospace6060074

Mahadevan, R., Jagan, A., Pavithran, L., Shrivastava, A., & Selvaraj, S. K. (2021). Intelligent welding by using machine learning techniques. *Materials Today: Proceedings*, *46*, 7402–7410. https://doi.org/10.1016/j.matpr.2020.12.1149

Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*(253), 68–78.

Meng, Y., Rajagopal, M., Kuntumalla, G., Toro, R., Zhao, H., Chang, H. C., Sundar, S., Salapaka, S., Miljkovic, N., & Ferreira, P. (2020). Multi-objective optimization of peel and shear strengths in ultrasonic metal welding using machine learning-based response surface methodology. *Mathematical Biosciences and Engineering*, *17*(6). https://doi.org/10.3934/mbe.2020379

Montgomery, D. C. (2017). *Design and analysis of experiments*. John wiley & sons.

Mounika, G., Rajyalakshmi, K., Rajkumar, G. V. S., & Sravani, D. (2024). Prediction and optimization of process parameters using design of experiments and fuzzy logic. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *18*(4), 2333–2343. https://doi.org/10.1007/s12008-023-01446-x

Myftarago, A., Bier, T. A., Qoku, E., Aliti, R., & Zogaj, M. (2023). Multi-Response Optimization on Hydrated Calcium Aluminate Rich Ternary Binders Using Taguchi Design of Experiments and Principal Component Analysis. *Buildings*, *13*(10), 2494. https://doi.org/10.3390/buildings13102494

Ramasamy, S. R., Gould, J., & Workman, D. (2002). Design-of-experiments study to examine the effect of polarity on stud welding. *WELDING JOURNAL-NEW YORK-*, *81*(2), 19-S.

Rathi, A. K. (2021). To study the effect of submerged arc welding parameters on bead geometry and hardness for mild steel (IS-2062A) using fractional factorial design. *Materials Today: Proceedings*, *34*, 525–530.

Reddy, G. T., Reddy, M. P. K., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). Analysis of dimensionality reduction techniques on big data. *Ieee Access*, *8*, 54776–54788. https://doi.org/10.1109/ACCESS.2020.2980942

Richmire, S., Hall, K., & Haghshenas, M. (2018). Design of experiment study on hardness variations in friction stir welding of AM60 Mg alloy. *Journal of Magnesium and Alloys*, *6*(3), 215–228. https://doi.org/10.1016/j.jma.2018.07.002

Sabegh, M. H. Z., Mirzazadeh, A., Salehian, S., & Weber, G.-W. (2014). A literature review on the fuzzy control chart; classifications & analysis. *International Journal of Supply and Operations Management*, *1*(2), 167.

Sharma, A., Paliwal, K. K., & Onwubolu, G. C. (2006). Class-dependent PCA, MDC and LDA: A combined classifier for pattern classification. *Pattern Recognition*, *39*(7), 1215–1229. https://doi.org/10.1016/j.patcog.2006.02.001

Skrobot, V. L., Castro, E. V. R., Pereira, R. C. C., Pasa, V. M. D., & Fortes, I. C. P. (2007). Use of principal component analysis (PCA) and linear discriminant analysis (LDA) in gas chromatographic (GC) data in the investigation of gasoline adulteration. *Energy & Fuels*, *21*(6), 3394–3400. https://doi.org/10.1021/ef0701337

Soleimanipour, A., Chegini, G. R., & Massah, J. (2018). Classification of Anthurium flowers using combination of PCA, LDA and support vector machine. *Agricultural Engineering International: CIGR Journal*, *20*(1).

Tesfaw, T., Singh, A. P., Gezahegn, A. M., & Garedew, B. T. (2022). Optimization of MAG welding process parameters using Taguchi design method on dead mild steel used in automotive industry. *International Journal of System Assurance Engineering and Management*, *13*(1), 79–89. https://doi.org/10.1007/s13198-021-01107-w

Zhang, H., Bai, X., Dong, H., & Zhang, H. (2024). Modelling and Prediction of Process Parameters with Low Energy Consumption in Wire Arc Additive Manufacturing Based on Machine Learning. *Metals*, *14*(5), 567. https://doi.org/10.3390/met14050567

Zhang, Y., & Edgar, T. F. (2008). PCA combined model-based design of experiments (DOE) criteria for differential and algebraic system parameter estimation. *Industrial & Engineering Chemistry Research*, *47*(20), 7772–7783. https://doi.org/10.1021/ie071206c