Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Advancing Circular and Sustainable Surface Engineering through Machine Learning-Assisted Plasma Electrolytic Polishing of TC4 Alloy: Reusability, Stirring and Immersion Properties

Abdulkadir Çebi

abdulkadir.cebi@samsun.edu.tr

Samsun Üniversitesi: Samsun Universitesi https://orcid.org/0000-0002-3074-6554

Hasan Demirtaş

Çağın Bolat

Furkancan Demircan

Research Article

Keywords: Plasma electrolytic polishing, TC4, Electrolyte stirring, Immersion depth, Machine learning, SHAP analysis

Posted Date: November 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11590729/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Advancing Circular and Sustainable Surface Engineering through Machine Learning-Assisted Plasma Electrolytic Polishing of TC4 Alloy: Reusability, Stirring and Immersion Properties

Abdulkadir Çebi[1,*], Hasan Demirtaş[1], Çağın Bolat[1], Furkancan Demircan[2]

[1]Department of Mechanical Engineering, Engineering and Natural Science Faculty, Samsun University, 55420 Ondokuzmayıs, Samsun, Türkiye

[2]Department of Software Engineering, Computer and Information Science Faculty, Samsun University, 55420 Ondokuzmayıs, Samsun, Türkiye

**Corresponding author: abdulkadir.cebi@samsun.edu.tr*

ABSTRACT

Plasma electrolytic polishing (PeP) is an anodic finishing process in which surface asperities are dissolved beneath a vapor–gaseous envelope (VGE). Previous studies have mainly addressed voltage and electrolyte properties, whereas immersion conditions and stirring remain less explored. This study investigated the effects of immersion depth, electrolyte reusability, immersion cycle, and stirring on the PeP of TC4 alloy using a 27-run full factorial design. Surface roughness, gloss, morphology, and elemental composition were evaluated. After PeP, R_a decreased from 13.18 μm to 1.214-3.173 μm, reaching its minimum at 50 mm depth, 600 rpm stirring, and single cycle, owing to a thinner VGE at greater depths and electrolyte renewal through agitation. Fluorine residues decreased from 9.8 to 0.8 wt.% with increasing depth and stirring, while Al and V contents remained nearly unchanged. Edge erosion cavities at 10 mm without stirring were suppressed by deeper immersion and stirring. Twenty-three regression models and a target-specific hybrid approach were compared for predicting R_a and gloss using stirring-stratified nine-fold cross-validation. The selected configurations achieved pooled out-of-fold R^2 values of 0.7159 (R_a) and 0.9009 (gloss). SHAP analysis identified stirring as the most influential predictor for both responses, followed by depth for Ra and cycle for gloss.

Keywords: Plasma electrolytic polishing; TC4; Electrolyte stirring; Immersion depth; Machine learning; SHAP analysis.

1. INTRODUCTION

TC4 is an α+β titanium alloy widely preferred in the aerospace, biomedical, automotive, marine, and chemical industries due to its low density, high specific strength, excellent corrosion resistance, and biocompatibility [1–5]. Components fabricated from this alloy are typically subjected to demanding service conditions involving elevated temperatures, high pressures, and cyclic loading [6]. In such harsh environments, even minor machining-induced surface defects can act as stress concentrators and crack initiation sites, ultimately leading to fatigue failure. Consequently, removing these defects through polishing is crucial for ensuring the reliability and service life of these components. However, polishing TC4 by using conventional methods is particularly challenging due to its low thermal conductivity, high chemical reactivity, and low elastic modulus [7]. In addition, conventional polishing techniques suffer from several drawbacks, including poor adaptability to complex geometries, insufficient accuracy, low efficiency, inconsistent surface quality, and environmental pollution [8–11].

In recent years, unconventional polishing methods, including magnetic abrasive finishing [12,13], laser polishing [14,15], chemical-mechanical polishing [16,17], and plasma electrolytic polishing (PeP) [18,19], have become increasingly attractive for surface finishing TC4 components. PeP is an unconventional surface treatment technique that achieves high surface quality on metals with complex geometries by using eco-friendly and sustainable electrolytes [11]. The material removal mechanism of PeP is based on anodic dissolution, governed by the combined action of chemical, thermal, electrochemical, and plasma-chemical reactions [11]. In this multiple finishing process, the workpiece serves as the anode and is immersed in an environmentally compatible aqueous electrolyte under a high applied voltage. Once the voltage exceeds a critical threshold, a stable vapor gaseous envelope (VGE) forms around the anode surface. Afterward, the plasma discharges within this envelope promote the controlled, preferential dissolution of surface asperities. Through this mechanism, smooth and lustrous surfaces can be obtained with minimal material loss.

The PeP process involves numerous variables, such as voltage, electrolyte temperature, electrolyte concentration, electrolyte composition, polishing time and immersion speed. Many studies focusing on the optimization of these parameters can be found in the literature. For instance, Wang et al. [20] investigated the effect of voltage values on the polishing of TC4 alloy using the PeP method and found that the average surface roughness was reduced to 1/24 of the initial value at 300 V. Likewise, Tan et al. [21] observed that the Al content increased while the Ti content decreased after polishing at 350 V for the TC4 alloy. Regarding electrolyte temperature parameter, Liu et al. [22] examined the relationship between electrolyte

temperature and the VGE and reported that increasing the temperature shifted the main reaction product from $TiOF_2$ to TiO_2 owing to a stronger oxidation effect. With a further rise in temperature, the influence of fluorine became predominant, and TiF_4 ultimately emerged as the main reaction product. Regarding the polishing time variable, Grigoriev et al. [23] analysed the polishing of TC4 alloy using an ammonium fluoride-based electrolyte with the addition of complexing agents at various polishing times. The findings indicate that surface roughness values were significantly improved at a polishing time of 20 minutes. In another study, Zou et al. [24] determined that a polishing duration of 30 minutes yielded the best surface quality for TC4 alloys produced via additive manufacturing (AM). In the similar way, Ghezri et al. [25] reported that the surface roughness of TC4 alloys produced via AM decreased from 9.2 μm to 0.5 μm after a 20-minute polishing period. For works conducted regarding multiple parameter optimization, Wang et al. [26] selected four different variables for PeP parameter optimization and identified voltage, polishing time, depth, and electrolyte temperature as the parameters influencing surface roughness. Similarly, Lv et al. [27] investigated the parameters influencing the polishing of TC4 alloy and listed as polishing time, electrolyte temperature, voltage, electrolyte concentration, and salt ratio as variables that affect surface roughness, respectively.

Surface finishing requires predictive models that relate processing conditions to both surface topography and optical appearance. Machine learning (ML) offers a means of estimating these responses from experimental observations, including when the relationships between process inputs and surface quality are nonlinear. For instance, in belt polishing, Qi et al. [28] developed a backpropagation artificial neural network using the maximum cutting depth of abrasive grains, belt rotational speed, and workpiece feed rate as inputs. Their approach combined a mechanistic description of abrasive contact with data-driven prediction of surface roughness, illustrating how physically relevant input variables can support the modelling of polishing outcomes. Neural networks have also been applied to successive finishing operations. Soler et al. [29] modelled the arithmetic mean roughness of selectively laser-melted TC4 parts after blasting and electropolishing. Their network incorporated manufacturing and finishing parameters and was evaluated against 429 experimental results, after which it was used to investigate processing conditions associated with reduced roughness. Deep learning extends this approach through multiple hidden layers that represent nonlinear relationships between processing variables and surface response. Pan et al. [30] developed a deep neural network for predicting surface roughness during chemical mechanical polishing of lithium niobate and combined the trained model with a genetic algorithm to search for processing parameters

corresponding to a specified roughness target. These studies demonstrate distinct uses of learning-based models, from predicting final surface quality to supporting the selection of finishing conditions. Gloss prediction introduces an additional response that should be modelled explicitly alongside roughness. Thasana and Wetchakama [31] used an artificial neural network to predict the gloss of injection-moulded plastic parts from three-dimensional surface roughness obtained through virtual machining and from injection-process parameters. Although that work concerned moulded plastics, it showed that gloss can be predicted from surface texture and process parameters. These studies show that learning-based models can relate finishing conditions to both topography and optical appearance, provided that their predictive performance is validated for the dataset at hand [32].

Numerous optimization studies on voltage, electrolyte temperature, electrolyte composition, and polishing time have been reported in the literature. However, the efficiency and effectiveness of PeP are not governed by these variables alone. Comprehensive reviews [8,33] have identified initial surface roughness [25,34,35], surface cleaning [18], and immersion depth [36,37] as additional factors. Çebi et al. [11] also highlighted the immersion cycle [38] as a parameter influencing PeP performance. Nevertheless, electrolyte stirring has rarely been considered, and its combined effect with immersion depth and immersion cycle remains unexplored.

This study aims to clarify the effects of electrolyte stirring, immersion depth, and immersion cycle on the PeP of TC4 alloy. Unlike most previous studies, the specimens were polished directly in the as-WEDM condition, without mechanical pretreatment. The three parameters were varied in a 27-run full factorial design, and the reusability of the electrolyte was evaluated across successive polishing series. Changes in surface quality were characterized in terms of arithmetic mean roughness (R_a), root-mean-square roughness (R_q), maximum roughness depth (R_z), material removal rate (MRR), and surface gloss. In addition, 23 statistical, machine learning, and ensemble regression methods and a target-specific hybrid approach were used to predict R_a and gloss. Regularized linear, kernel-based, tree-based, and ensemble models were compared using pooled out-of-fold R^2, RMSE, and MAE under nine-fold cross-validation stratified by stirring level. SHAP analysis was then applied to determine the contribution of each parameter to the predictions. To the best of our knowledge, these parameters have not previously been investigated together for the PeP of as-WEDM TC4 alloy. In addition, no earlier PeP study has modelled surface roughness and gloss as separate responses.

2. MATERIALS and METHODS

The experimental workflow followed in this study is illustrated in Figure 1. Initially, TC4 alloy samples were cut to the required dimensions by wire electrical discharge machining (WEDM), and the resulting surfaces were subjected to PeP without any prior mechanical pretreatment. Each sample was weighed before and after polishing, and the mass loss was used to calculate the MRR. The polished surfaces were then characterized in terms of surface roughness (R_a, R_q, and R_z) and surface gloss. Surface morphology was examined by SEM, while the elemental composition and phase constitution of the polished surfaces were determined by EDX and XRD analyses, respectively. Finally, the experimental results obtained from the 27 experimental conditions were used to train and evaluate 23 statistical, machine learning, and ensemble regression models, together with a target-specific hybrid approach, for predicting R_a and surface gloss.

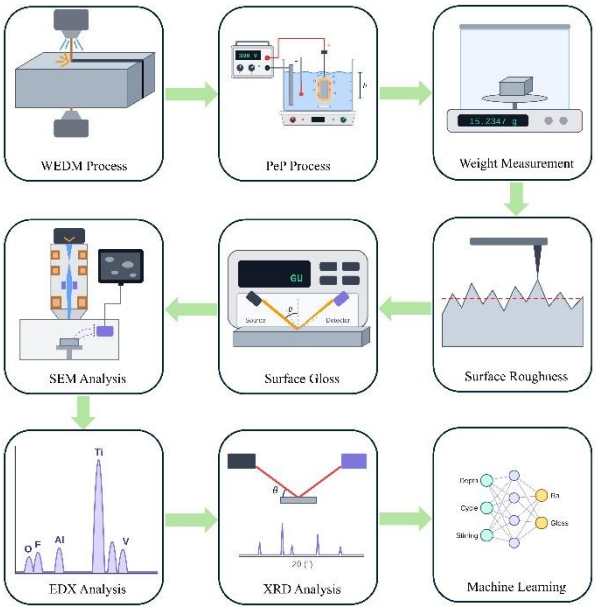

Figure 1. Process flow diagram for experimental conditions.

2.1. Sample Preparation

Samples with dimensions of 25×8.8×2 mm^3 were sectioned from a titanium plate by WEDM bench (DK7740, Hightech CNC, Jiangsu, China). The dimensions of the samples were then checked with a digital calliper (500-181-30, Mitutoyo Corporation, Kawasaki, Japan). No pretreatment, such as sanding or grinding, was applied to the samples. In the as-WEDM condition, the surface had R_a, R_q, and R_z values of 13.18 µm, 15.98 µm and 73.379 µm respectively, and a gloss of 12.6 GU.

2.2.Experimental Setup

The experiments were conducted using a modified three-axis CNC machine. The setup consists of a DC power supply (2060H, Owon Technology, Zhangzhou, China), a magnetic hotplate stirrer (MS-H280-PRO, DLAB Scientific, Beijing, China), a thermometer (059.02.002, IsoLab, İstanbul, Türkiye), a 1 L beaker (025.01.901 model, IsoLab, İstanbul, Türkiye), and connectors. In addition, an SS304 plate was used as the cathode and mounted inside the beaker. The cathode plate measured 250×25×1.5 mm^3 and was positioned 80 mm from the specimen surface, parallel to the polished face. The electrolyte was heated to 80°C, and its temperature was monitored with the thermometer. The hotplate controller maintained the bulk temperature at 80°C (±1°C); at the end of each 240 s polishing time, the electrolyte temperature had risen to 88°C, and the next run started only after it had returned to 80°C. As soon as this temperature was reached, the workpiece was immersed to various depths according to the experimental conditions. Meanwhile, the voltage was set to 300 V using the power supply software (PC software for OWP-H series, Owon Technology, Zhangzhou, China), and current time data were recorded throughout the process. The experimental setup is demonstrated in Figure 2, which presents both an overview and a close-up view.

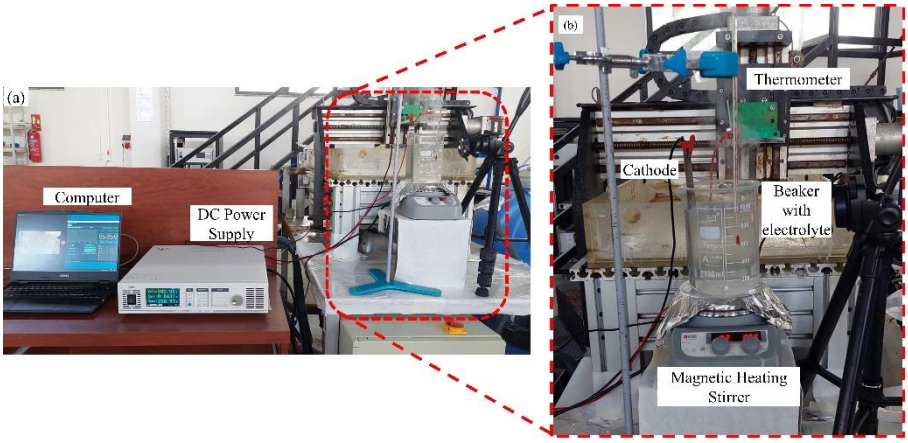

Figure 2. Experimental setup and equipments a) Overview and b) Close-up view.

Based on preliminary experiments, the electrolyte consisted of 2.813% (wt.) sodium fluoride (TekKim Kimya, Bursa, Türkiye), 1.687% (wt.) ammonium fluoride (Sigma-Aldrich, Darmstadt, Germany), and 0.5% (wt.) hydrochloric acid (080231.01001 code, TekKim Kimya, Bursa, Türkiye) in deionized water, giving a total solute content of 5% (wt.). Each batch contained 1000 mL of electrolyte, and its initial pH and electrical conductivity at 80 °C were 6.2 and 70.8 mS cm^{-1}, respectively. Because acidified fluoride solutions can release hydrogen fluoride, all experiments were conducted with acid-resistant personal protective equipment.

Before polishing, the electrolyte was stirred with the magnetic hotplate stirrer for 5 min to ensure homogeneous mixing.

Immersion depth was defined as the vertical distance between the free electrolyte surface, measured before stirring, and the upper edge of the specimen, so that the entire specimen was submerged at all three levels of 10, 30, and 50 mm (Figure 2c). The corresponding hydrostatic overpressure at the sample ranges from approximately 0.10 kPa to 0.74 kPa, which is less than 1% of atmospheric pressure. Depth was therefore expected to act mainly through the amount of vapor and gas rising along the specimen toward the free surface [58], rather than through pressure itself. These depths were combined with immersion cycling, and single, quadruple, and octuple immersions were performed with 240 s, 60 s, and 30 s per immersion, respectively. In the quadruple and octuple modes, the specimen was raised 10 mm above the electrolyte surface at 10 mm/s by the Z-axis of the CNC system, held in air for 1 s with the voltage applied, and reimmersed at the same speed. The total immersion time was 240 s in all modes. In addition, to the best of the authors' knowledge, few studies have addressed the effect of electrolyte stirring. Therefore, stirring speeds of 0 (no stirring), 300 rpm, and 600 rpm were applied using the magnetic hotplate stirrer.

To evaluate electrolyte reusability, a single electrolyte batch was used for all three immersion depths in succession. The fresh electrolyte was first used for the 10 mm series, and the same batch was then reused for the 30 mm and 50 mm depth groups. As a result, each reuse level corresponds to a single immersion depth. The surface quality obtained at a given depth therefore reflects the combined effect of immersion depth and electrolyte usage. Accordingly, the results related to reusability are interpreted as the combined effect of depth and reuse rather than as the effect of reuse alone.

2.3.Material Removal Rate and Surface Analysis Procedure

MRR is widely regarded as a key indicator of polishing performance. Accordingly, the mass of each sample was measured before and after PeP using an electronic balance (602.31.002, IsoLab, İstanbul, Türkiye) with a readability of 0.001 g. Afterward, the difference between the two values was taken as the mass loss resulting from polishing. With mass losses of approximately 32 mg and 75 mg, the balance readability of 1 mg corresponds to an uncertainty of about 1.3% to 3.1% in MRR.

Surface roughness was measured with a surface roughness tester (HDS-G30, Mitutoyo Corporation, Kawasaki, Japan) in accordance with ISO 4287:1997, using a cutoff wavelength

(λc) of 2.5 mm, two sampling lengths (evaluation length of 5.0 mm), and a traverse speed of 0.25 mm/s. All surfaces, including the as-WEDM surface, were measured under the same conditions, so that the values remain directly comparable. Profiles were recorded perpendicular to the WEDM wire traces at the centre of the polished face. For each condition, three specimens were polished independently and each was measured [three] times at different positions; the values reported in Table 1 are the averages of these measurements. The surface morphology of the samples before and after PeP was examined by scanning electron microscopy (SEM; SU1510, Hitachi High-Tech Corporation, Tokyo, Japan), and their elemental composition was analysed with an integrated energy-dispersive X-ray spectroscopy (EDX) detector. The phase composition was investigated by X-ray diffraction (XRD) using a SmartLab diffractometer (Rigaku Corporation, Hokuto, Japan), with data collected between 2θ interval 5° and 90° at 3°/min. Surface gloss was measured directly on the polished face of each sample with an SL268 gloss meter (SerLab Service Laboratory Systems, İstanbul, Türkiye) at a 60° geometry in accordance with ISO 2813. The instrument was calibrated against its reference tile before each session, and five readings were averaged for each sample, oriented parallel to the WEDM traces.

2.4. Machine Learning and Hybrid Models

This study evaluated 23 statistical, machine learning, and ensemble regression models together with a target-specific hybrid approach to predict surface roughness and gloss from experimentally controlled process parameters. The dataset comprised 27 experimental observations obtained from a full 3×3×3 factorial design involving Depth, Cycle, and Stirring. Depth was investigated at levels of 10, 30, and 50, Cycle at levels of 1, 4, and 8, and stirring at levels of 0, 300, and 600. These factor combinations produced 27 unique experimental conditions, with one observation recorded for each condition, as presented in Table 1. The predictive analysis focused on arithmetic mean surface roughness, R_a, and Gloss, with Depth, Cycle, and Stirring included as predictors for both responses. Reusability was excluded from the predictor set because it represented a one-to-one relabelling of Depth and therefore provided no additional independent information. Separate models were developed for R_a and Gloss to characterize their respective relationships with the processing conditions, and SHAP analysis was subsequently applied to examine the contributions of the three predictors to the model predictions.

Table 1. Original dataset of process parameters and measured outcomes.

Exp. No	Depth (mm)	Cycle	Stirring (rpm)	Reusability	R_a (µm)	R_q (µm)	R_z (µm)	Gloss (GU)
1	10	1	0	1	3,173	5,835	31,277	42,9
2	10	1	300	1	2,96	5,527	24,116	49,48
3	10	1	600	1	2,603	4,439	20,035	52,32
4	10	4	0	1	2,948	5,117	26,987	45,5
5	10	4	300	1	2,812	4,84	21,633	51,2
6	10	4	600	1	2,347	4,167	18,354	53,66
7	10	8	0	1	2,576	4,561	20,964	47,86
8	10	8	300	1	2,751	4,322	26,957	50,94
9	10	8	600	1	2,092	3,685	20,981	54,12
10	30	1	0	2	3,08	5,413	28,527	42,38
11	30	1	300	2	2,804	4,586	25,583	47,74
12	30	1	600	2	1,815	3,439	15,698	50,58
13	30	4	0	2	3,003	4,268	25,688	47,58
14	30	4	300	2	2,365	5,02	23,428	51,64
15	30	4	600	2	2,189	4,094	19,989	52,98
16	30	8	0	2	2,614	4,628	20,411	49,48
17	30	8	300	2	2,128	3,879	20,903	50,28
18	30	8	600	2	1,563	3,144	14,203	53,86
19	50	1	0	3	2,437	4,353	20,66	41,94
20	50	1	300	3	1,917	4,901	19,601	48,54
21	50	1	600	3	1,214	2,795	12,981	53,8
22	50	4	0	3	2,189	4,168	22,764	48,16
23	50	4	300	3	2,04	3,997	20,148	51,8
24	50	4	600	3	1,887	3,757	18,489	52,82
25	50	8	0	3	2,04	4,471	17,243	51,84
26	50	8	300	3	2,398	3,463	21,651	51,88
27	50	8	600	3	1,894	3,087	21,527	53,9

The predictive analysis included 23 statistical, machine learning, and ensemble regression methods together with an additional target-specific hybrid model, resulting in 24 approaches in the final comparison. The evaluated methods comprised Linear Regression, Ridge Regression, Lasso Regression, Bayesian Ridge Regression, Partial Least Squares Regression, K-Nearest Neighbors Regression, Support Vector Regression, Decision Tree Regression, Random Forest Regression, Extra Trees Regression, Gradient Boosting Regression, AdaBoost Regression, Bagging Regression, Stacking Regression, XGBoost Regression, Gaussian Process Regression, second-degree Polynomial Ridge Regression, cross-validation-tuned Ridge Regression, cross-validation-tuned Elastic Net Regression, Huber Regression, tuned Kernel Ridge Regression, LightGBM Regression, and Voting Regression. Separate models were developed for R_a and Gloss using Depth, Cycle, and Stirring as predictors for both responses. Reusability was excluded from the predictor set because it represented a one-to-one

relabelling of Depth and provided no additional independent information. Predictor variables were standardized within each training fold, while R_a was logarithmically transformed and Gloss was standardized before model fitting. Predictions were subsequently returned to their original measurement scales for performance evaluation. The validation analysis compared shuffled three-fold, five-fold, and nine-fold cross-validation, three-fold and nine-fold cross-validation stratified by Stirring level, and leave-one-out cross-validation. Each randomized validation design was evaluated using ten random seeds, whereas leave-one-out cross-validation was performed once. Following comparison of predictive performance and variability across seeds, the final analysis used nine-fold cross-validation stratified by Stirring level with a random seed of 0. Each validation fold contained three observations, comprising one observation from each Stirring level, while the remaining 24 observations were used for training. The resulting out-of-fold predictions were used to calculate R^2, RMSE, and MAE for each response and to construct the target-specific hybrid models by combining Ridge predictions with those of a selected complementary model. Because the validation design, final seed, hybrid partner, and blending weight were selected using the same dataset, the final performance estimates describe the selected modelling configuration and do not constitute an independent assessment of its generalization capability.

2.4.1. Correlation Analysis

The Pearson correlation matrix illustrates the strength and direction of the linear relationships among Depth, Cycle, Stirring, R_a, and Gloss. Stirring exhibited a strong positive correlation with Gloss and a negative correlation with R_a, indicating that higher stirring levels were associated with greater gloss and lower arithmetic mean surface roughness within the investigated conditions. Depth was negatively correlated with R_a but showed only a weak positive association with Gloss. Cycle exhibited a moderate positive correlation with Gloss and a weak negative correlation with R_a, suggesting a stronger linear association with optical appearance than with surface roughness. R_a and Gloss were negatively correlated, indicating that smoother surfaces generally exhibited higher gloss. The correlations among Depth, Cycle, and Stirring were zero, consistent with the balanced full factorial experimental design. Because Pearson coefficients capture only linear association, the significance and relative contribution of each factor were evaluated by ANOVA (Section 3.1).

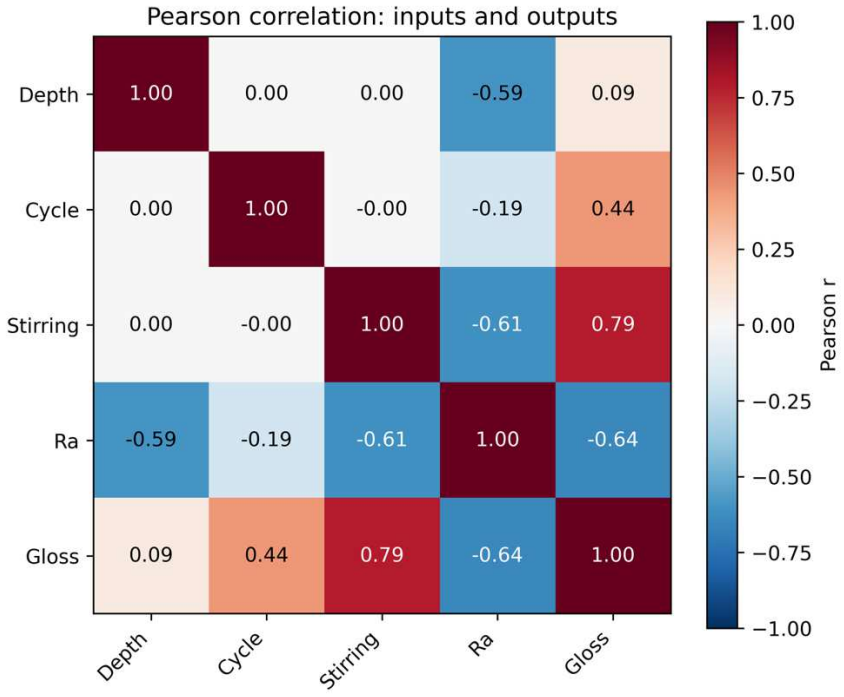

Figure 3. Pearson correlation matrix of the process parameters and measured response variables.

A broad range of statistical, machine learning, ensemble, and kernel-based regression approaches was employed to characterize the linear and nonlinear relationships between Depth, Cycle, and Stirring and the surface quality responses R_a and Gloss. The model set provided complementary predictive structures incorporating linearity, regularization, local similarity, nonlinear interactions, robustness, and prediction aggregation. A total of 23 base regression models were evaluated independently for each response, followed by the construction of a target-specific hybrid model using weighted combinations of out-of-fold predictions from Ridge Regression and a selected complementary model, resulting in 24 approaches in the final comparison.

2.4.2. Regularization Methods

The regularization methods followed the formulations of Ridge Regression [39], Lasso [40], and Elastic Net [41]. Tree ensembles were based on the Random Forest framework [42] and sequential gradient boosting [43], while the kernel-based approaches included Support Vector Regression [44]. The XGBoost implementation follows the regularized tree-boosting framework described by Chen and Guestrin [45]. These methodological foundations supported the evaluation of 23 base regression models and an additional target-specific hybrid approach for predicting R_a and Gloss from Depth, Cycle, and Stirring. Each model was fitted separately for the two responses using standardized predictors, with logarithmic transformation applied to

R_a and standardization applied to Gloss. Predictions were subsequently transformed back to their original measurement scales before performance evaluation and hybrid construction.

Linear Regression represented each response as an additive linear combination of the predictors, with coefficients estimated by minimizing the sum of squared residuals [46]. Ridge Regression incorporated an L2 penalty to constrain coefficient magnitudes [39], with the regularization parameter fixed at $\alpha = 1.0$. Lasso Regression employed an L1 penalty capable of shrinking coefficients to zero, using $\alpha = 0.01$ and a maximum of 20,000 optimization iterations. Bayesian Ridge Regression assigned probability distributions to the regression parameters and updated them using the observed data, providing regularized estimates and a probabilistic representation of parameter uncertainty [47]. Partial Least Squares Regression constructed latent components that captured covariance between the predictors and the response [48], with two components used in the present implementation. K-Nearest Neighbours Regression estimated each response from nearby observations in the standardized predictor space [46]. The implemented model used three neighbours and inverse-distance weighting, giving greater influence on observations closer to the prediction point. Support Vector Regression employed an epsilon-insensitive loss and a radial basis function kernel to represent nonlinear relationships [44], with $C = 10.0$ and $\varepsilon = 0.05$. Decision Tree Regression recursively partitioned the predictor space into regions according to splitting rules that reduced prediction error [46], with the maximum tree depth restricted to four. Random Forest Regression combined predictions from trees constructed using randomized sampling and predictor selection [42], with 300 trees included in the ensemble. Extra Trees Regression introduced additional randomness through the selection of candidate splitting thresholds [49] and also employed 300 trees. Gradient Boosting Regression sequentially added learners to reduce the loss of the existing ensemble [43]. The implemented model used 200 estimators, a maximum tree depth of two, and a learning rate of 0.05. AdaBoost Regression adjusted the influence of training observations according to their prediction errors and aggregated the resulting estimators [50], using 150 estimators. Bagging Regression fitted estimators to bootstrap samples and averaged their predictions to reduce the variability associated with individual learners [51], with 150 estimators included in the implementation. Stacking Regression integrated the predictions of multiple base learners through a secondary regression model [52]. Random Forest and Gradient Boosting models containing 150 estimators each served as the base learners, while Ridge Regression with $\alpha = 1.0$ served as the final estimator. Three-fold internal cross-validation generated the base-model predictions used to train the final estimator. XGBoost Regression constructed an additive tree

ensemble within a regularized optimization framework [45], using 300 trees, a maximum depth of two, and a learning rate of 0.05. Gaussian Process Regression represented the response through a probability distribution over functions, with a covariance kernel describing relationships between observations [53]. The implemented covariance function comprised the product of a constant kernel and a radial basis function kernel, together with an additive white-noise component. Three optimizer restarts and internal response normalization were applied during model estimation. Second-Degree Polynomial Ridge Regression expanded the predictors into linear terms, squared terms, and pairwise interactions before applying L2-regularized regression [39,46], with $\alpha = 1.0$. Cross-Validation-Tuned Ridge Regression selected its regularization strength from 25 logarithmically spaced α values between 10^{-3} and 10^{3} using the cross-validation procedure provided by the RidgeCV estimator [32,39]. Cross-Validation-Tuned Elastic Net Regression combined L1 and L2 penalties and evaluated L1 ratios of 0.1, 0.5, 0.9, and 1.0 alongside 25 logarithmically spaced α values between 10^{-3} and 10^{1}. Parameter selection used three-fold internal cross-validation, with a maximum of 20,000 optimization iterations [32,41]. Huber Regression employed a robust loss that was quadratic for small residuals and linear for sufficiently large residuals, limiting the influence of large response errors [54]. The maximum number of optimization iterations was set to 5,000. Tuned Kernel Ridge Regression combined L2 regularization with a radial basis function kernel [39,46]. Its α and γ parameters were selected through three-fold grid-search cross-validation using negative mean squared error as the scoring criterion. The search evaluated five logarithmically spaced α values between 10^{-2} and 10^{2} and five logarithmically spaced γ values between 10^{-2} and 10^{1}. LightGBM Regression employed computationally efficient gradient-boosted decision trees [55]. The implemented model contained 150 estimators, with a maximum tree depth of three, a maximum of seven leaves, a minimum child-sample setting of three, and a learning rate of 0.05. Voting Regression averaged the predictions of separately fitted models according to the general principle of prediction combination [56]. The implemented ensemble combined Ridge Regression with $\alpha = 1.0$ and Gradient Boosting Regression containing 150 estimators. The Gradient Boosting estimators within the Voting and Stacking ensembles retained their library-default depth and learning-rate settings, which differed from those specified for the standalone Gradient Boosting model.

Target-Specific Hybrid Regression combined out-of-fold predictions using a response-specific partner-selection and weighting procedure consistent with the general principle of prediction combination [56]. Ridge Regression served as the reference predictor for both R_a and

Gloss, and each alternative base model was evaluated as a potential partner. The weight assigned to the partner model ranged from 0.0 to 1.0 in increments of 0.1, with the complementary weight assigned to Ridge Regression. The combination yielding the highest pooled out-of-fold coefficient of determination (R^2) was retained independently for each response. In the final analysis, the selected R_a prediction comprised 30% Ridge Regression and 70% Support Vector Regression. For Gloss, the selected weight was 100% Stacking Regression, making the resulting hybrid prediction identical to that of the Stacking model. The internal cross-validation procedures used for parameter tuning and stacking were distinct from the final nine-fold evaluation stratified by Stirring level. However, the validation design, final random seed, hybrid partner, and blending weight were selected using performance estimates obtained from the same dataset. Consequently, the reported results represent the performance of the selected configuration and may be optimistic because model selection was not separated from independent evaluation [57]. An additional outer validation procedure or an independent test dataset would be required to assess the generalization performance of the complete selection process.

2.4.3. Performance Metrics

Predictive performance was evaluated using the R^2, root mean squared error (RMSE) and mean absolute error (MAE). These metrics were calculated independently for R_a and Gloss using pooled out-of-fold predictions from all 27 experimental observations. The final evaluation employed nine-fold cross-validation stratified by Stirring level with a random seed of 0, with each validation fold containing three observations and the remaining 24 observations used for model training. Fold-specific performance metrics were calculated separately to assess variation among validation subsets. Predictions for R_a and Gloss were transformed back from their logarithmic and standardized representations, respectively, before performance evaluation, ensuring that prediction errors were expressed on the original measurement scales. Let $y_{i,t}$ denote the experimentally measured value of response t for observation i, let $\hat{y}_{i,t}$ denote its corresponding out-of-fold prediction, and let $\bar{y}_t$ represent the mean measured value across all $n = 27$ observations, where $t \in \{R_a, \text{Gloss}\}$. The coefficient of determination, presented in Equation (1), compares the sum of squared prediction errors with the total sum of squares around the observed response mean. Higher R^2 values indicate closer agreement between measured and predicted responses, whereas negative values indicate poorer performance than the corresponding mean-response reference. Pooled R^2 was calculated using the overall

response mean, while fold-specific R^2 used the mean within each validation subset, meaning that pooled R^2 does not equal the arithmetic mean of the fold-specific values.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{t,i} - \hat{y}t, i\right)^2}{\sum_{i=1}^{n}\left(y_{t,i} - \bar{y}_t\right)^2} \quad (1)$$

An (R_t^2) value approaching 1 indicates strong predictive agreement, while a value of 0 indicates performance equivalent to predicting the overall mean response. A negative value indicates that the model performs worse than the mean-based reference prediction. The root mean squared error, defined in Equation (2), quantifies the typical magnitude of prediction errors while imposing a greater penalty on larger deviations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{t,i} - \hat{y}_{t,i}\right)^2} \quad (2)$$

The mean absolute error, defined in Equation (3), quantifies the average magnitude of the differences between the measured and predicted responses, irrespective of whether the individual errors are positive or negative.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{t,i} - \hat{y}_{t,i}\right| \quad (3)$$

Lower RMSE and MAE values indicate greater predictive accuracy, with both metrics calculated on the original measurement scales of R_a and Gloss. RMSE assigns greater influence on large prediction errors through the squared residuals, whereas MAE measures the average absolute deviation and is less sensitive to occasional extreme errors. The pooled coefficient of determination was calculated using the mean of all 27 observed values for each response, while each fold-specific value used the mean of the three observations within the corresponding validation subset. Consequently, pooled R^2 is not the arithmetic mean of the nine-fold-specific R^2 values. The final nine-fold cross-validation design stratified by Stirring level supported the calculation of R^2 within each fold, although the small validation subsets could produce substantial variation in this metric, particularly when the observed responses within a fold exhibited limited variability.

3. RESULTS and DISCUSSIONS

3.1.Surface Characterization

The XRD patterns of the raw sample and the sample polished (Exp. No 21) are presented in Figure 4. Both patterns exhibit the characteristic reflections of hexagonal close-packed α-Ti, together with a weak β-Ti (1 1 0) peak at approximately 39.5°, consistent with the α+β microstructure of TC4. In the raw sample, the reflections were weak and broad, which can be attributed to the WEDM based recast layer containing fine-grained or partially amorphous features, residual stresses, and microcracks, as well as to the high initial roughness. After PeP, the peak intensities increased markedly and the reflections became sharper, indicating that the recast layer was largely removed. In addition, the (0 0 2) reflection became dominant over the (1 0 1) reflection, suggesting a preferred basal orientation, most likely originating from the rolling texture of the plate. No peaks attributable to TiO_2, $TiOF_2$, or TiF_4 were detected, implying that the reaction products formed during PeP did not accumulate on the surface in detectable quantities.

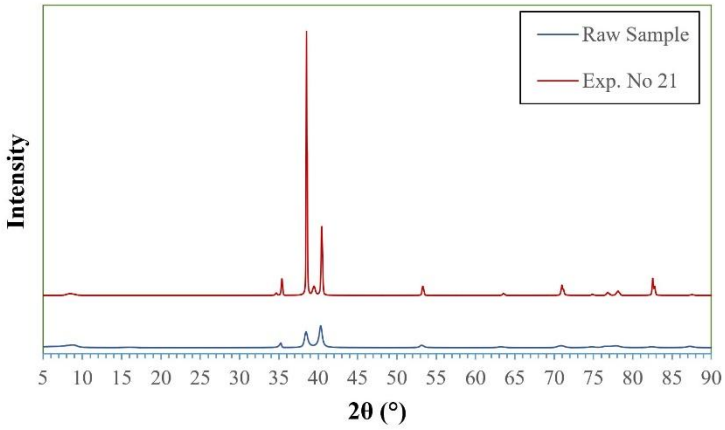

Figure 4. XRD analysis of the samples before and after polishing.

In order to evaluate the effect of the process parameters on surface morphology after PeP process, the samples exhibiting the lowest and highest roughness values were examined by SEM device. SEM images of the raw sample are also included for comparison. All images were acquired from the same region of each sample. Figure 5 presents SEM images of the raw sample (Figure 5a) and of the samples exhibiting the highest (Figure 5b) and lowest (Figure 5c) roughness values after PeP process. The raw sample surface displays the typical morphology of a WEDM-machined surface, consisting of overlapping discharge craters, resolidified droplets, globular debris, and irregular ridges, together with micropores formed during the rapid melting and resolidification of the recast layer. After PeP process under the conditions (10 mm depth,

single cycle, no stirring) that yielded the highest roughness, these features were substantially attenuated, and the surface became markedly more uniform. Nevertheless, residual ridges and pits remained locally, appearing with bright edges owing to enhanced secondary electron emission at topographic peaks. This indicates that material removal under these conditions was insufficient to eliminate the deepest WEDM-based surface shapes. In contrast, the sample with the lowest roughness exhibited a homogeneous and nearly featureless surface, with no remnants of discharge craters or resolidified parts. This observation is consistent with the selective dissolution of surface asperities within the VGE and demonstrates that, under appropriate process conditions, PeP can completely remove the WEDM recast layer without any prior mechanical pretreatment.

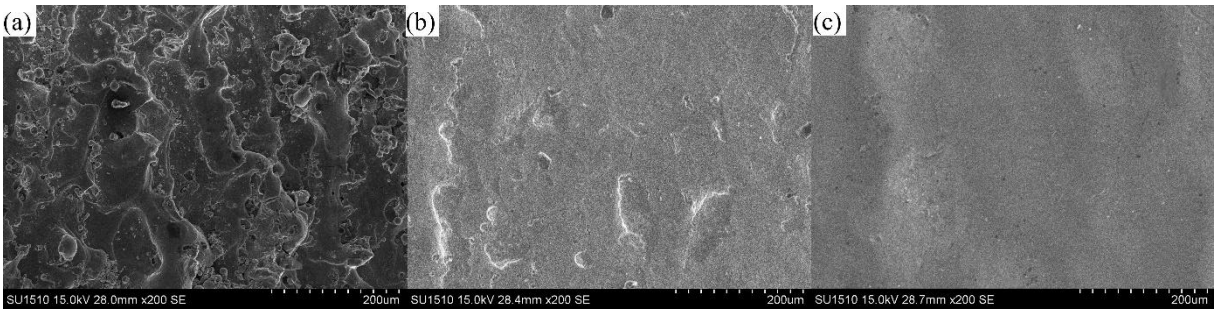

Figure 5. SEM images of the TC4 alloys for a) Raw, b) Exp. No 1 and c) Exp. No 21.

Figure 6 presents the SEM images of the polished samples at 10 mm immersion depth without stirring (Figure 6a to Figure 6c), 30 mm immersion depth with stirring at 300 rpm (Figure 6d to Figure 6f), and 50 mm immersion depth with stirring at 600 rpm (Figure 6g to Figure 6i), each under single, quadruple, and octuple immersion cycle, respectively. As also shown in Figure 5, the discharge craters and resolidified droplets of the WEDM recast layer were largely removed in all samples. In addition, R_a values decreased from an initial value of 13.18 µm to the between 1.214 µm and 3.173 µm, corresponding to reductions of approximately 75.9% and 90.8%. Nevertheless, the degree of surface homogenization differed markedly among the conditions.

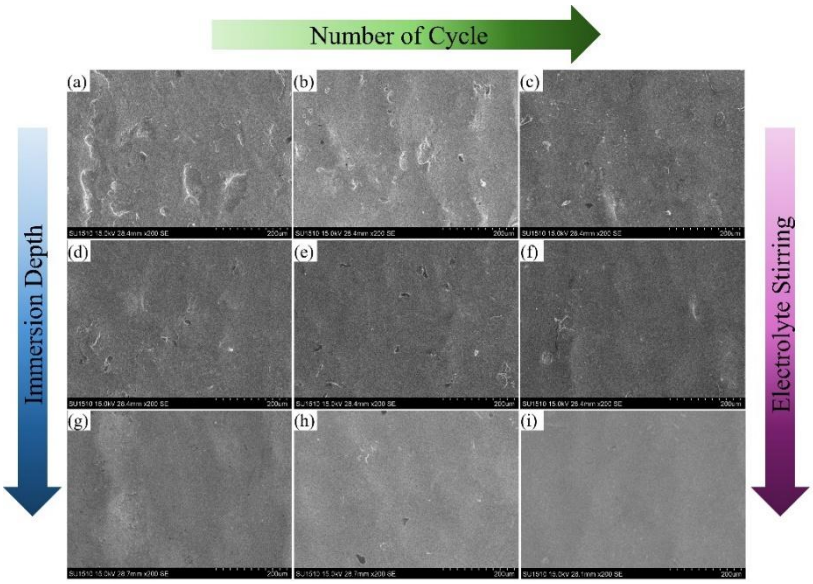

Figure 6. SEM images of the polished samples accordance to a) Exp. No 1, b) Exp. No 4, c) Exp. No 7, d) Exp. No 11, e) Exp. No 14, f) Exp. No 17, g) Exp. No 21, h) Exp. No 24 and i) Exp. No 27.

The combined increase in immersion depth and stirring speed generally improved the surface quality. At 10 mm depth and no stirring condition, residual ridges and pits remained visible, and the highest roughness values were obtained, with R_a, R_q, and R_z values reaching 3.173 µm, 5.835 µm, and 31.277 µm, respectively. In contrast, for the TC4 alloys in Figures 4d to 4f, R_a values decreased to between 2.128 mm and 2.804 µm and small, scattered pits and debris were observed. The lowest roughness values were obtained at 50 mm depth, 600 rpm stirring and single cycle (Figure 4g), and the R_a, R_q, and R_z values were 1.214 mm (improvement rate 90.79%), 2.795 mm (improvement rate 82.51%), and 12.981 µm (improvement rate 82.31%), respectively. These results are related to the nonuniform distribution of the VGE along the immersed sample. The vapor and gas generated at the anode rise toward the free surface of the electrolyte, so the VGE becomes thicker in the regions close to the electrolyte and air interface [59]. Since nearly the entire applied voltage drops across the VGE, regions covered by a thinner layer experience a higher current density [59,60]. Because PeP behaves predominantly as an electrochemical process, this higher current density accelerates the preferential removal of surface peaks. Accordance with this phenomenon, lower roughness values have been reported for deeper immersion [61]. On the other hand, this case is related to electrolyte stirring. During stirring process, the resulting convective flow continuously supplies fresh electrolyte to the anode surface, including the areas where polishing is most intense. Thus, the refreshed electrolyte can enhance PeP performance, leading to improved surface quality.

The electrolyte reusability was also evaluated, since the same composition was used in successive polishing conditions. In the experimental sequence, the electrolyte was utilized for the first time in the 10 mm depth, for the second time in the 30 mm depth, and for the third time in the 50 mm depth. As shown in Figure 6, surface quality did not deteriorate with repeated usage of electrolyte. On the contrary, the maximum R_a value decreased from 3.173 µm in the first usage to 2.804 µm in the second usage and 1.894 µm in the third usage. The corresponding minimum values were 2.576 µm, 2.128 µm, and 1.214 µm, respectively. The SEM images followed the same trend as residual ridges and pits were visible after the first usage, only scattered pits remained after the second usage, and a nearly featureless surface was obtained after the third usage. In addition, each reusage level coincided with a different immersion depth, and the samples in Figure 6 also differed in stirring speed, these improvements cannot be attributed to electrolyte reusage alone. Nevertheless, no deterioration in surface quality or removal rate was observed over three consecutive usages, suggesting that the electrolyte can be reused without compromising PeP performance within the investigated range. This finding is relevant for reducing chemical consumption and waste in practical applications.

The elemental compositions of the polished surfaces obtained by EDX analysis are demonstrated in Figure 7. In all samples, the contents of Ti (72.8 wt.% to 82.0 wt.%), Al (4.6 wt.% to 5.6 wt.%), and V (2.5 wt.% to 3.6 wt.%) remained within a narrow range, indicating that PeP did not cause significant selective dissolution of the alloying elements under the investigated conditions. In contrast, the fluorine content varied considerably among the samples. The highest F concentrations were detected under the 10 mm depth and no stirring condition with single and quadruple immersion (9.8 wt.% and 7.9 wt.%, respectively), accompanied by the highest Na contents (1.6 wt.% and 1.5 wt.%). These samples also exhibited the highest roughness values, and their SEM images revealed residual pits and ridges. This suggests that, in the absence of stirring, fluoride-containing reaction products and electrolyte residues were not effectively removed from the surface and were partly retained within the remaining surface irregularities. When stirring was applied, the F content generally decreased, reaching values as low as 0.9 wt.% for quadruple immersion at 30 mm depth with 300 rpm stirring and 0.8 wt.% for single immersion at 50 mm depth with 600 rpm stirring. Notably, the latter sample also exhibited the lowest roughness. Across all conditions, a moderate positive correlation was found between F content and R_a, indicating that rougher surfaces tend to retain more fluoride residues. Although no fluoride phases such as $TiOF_2$ or TiF_4 were detected by XRD, the presence of F in the EDX spectra indicates that these reaction products may remain

on the surface as a thin or amorphous layer below the detection limit of XRD. For the 10 mm depth with no stirring, the O content increased from 2.9 wt.% to 8.0 wt.% as the number of immersions increased, while the F content decreased from 9.8 wt.% to 1.0 wt.%. However, this trend was not observed under the other conditions. The C content (3.8 wt.% to 7.0 wt.%) is most likely associated with adventitious carbon contamination, while Cl, originating from HCl in the electrolyte, remained negligible (<0.2 wt.%) in all samples.

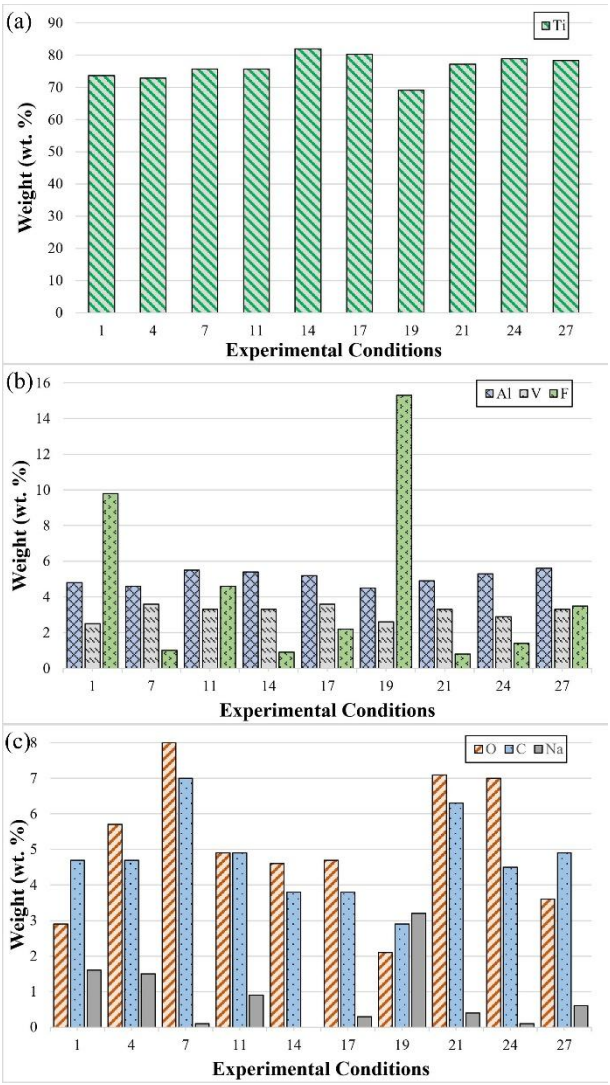

Figure 7. EDX analysis results for Exp. No 1, No 4, No 7, No 11, No 14, No 17, No 19, No 21, No 24 and No 27 accordance with a) Ti element, b) Al, V and F elements, and c) O, C, Na elements.

3.2.Corner Region Analysis

SEM images for the corner regions of polished samples as Exp. No 1, No 4, No 7, No 11, No 14, No 17, No 19, No 21, No 24 and No 27 are demonstrated in Figure 8. In all samples, the sharp WEDM corners were rounded, which is a characteristic outcome of PeP. This rounding can be attributed to the dependence of the current density distribution on surface

geometry, whereby protruding features experience higher current densities and are therefore removed faster [59,60]. For the 10 mm depth and no stirring condition (Figure 8a to Figure 8c), large and irregular erosion cavities with a porous morphology formed along the edges. These defects were most extensive at octuple immersion (Figure 8c), even though this sample exhibited the lowest roughness within its group (Figure 8a to Figure 8c). This indicates that R_a values measured on the flat face do not fully reflect the surface condition at the edges. The formation of these cavities can be related to localized discharges at the corners, where the electric field is concentrated. Such discharges may intensify during the initiation stage of each immersion, since PeP is prone to high current peaks while the VGE is being formed [62]. In the absence of stirring, the accumulation of gas near the edges may further destabilize the VGE. For the 30 mm depth and 300 rpm stirring condition (Figure 8d to Figure 8f), the erosion cavities became smaller and fewer, and only scattered small pits remained at quadruple and octuple immersion. On the other hand, under the 50 mm depth and 600 rpm stirring condition (Figure 8g to Figure 8i), the corners were smoothly rounded and free of erosion cavities, and the sample polished with single immersion (Figure 8g) exhibited the most uniform corner morphology. These results show that increasing the immersion depth and stirring speed not only reduces surface roughness but also suppresses localized damage at the corners.

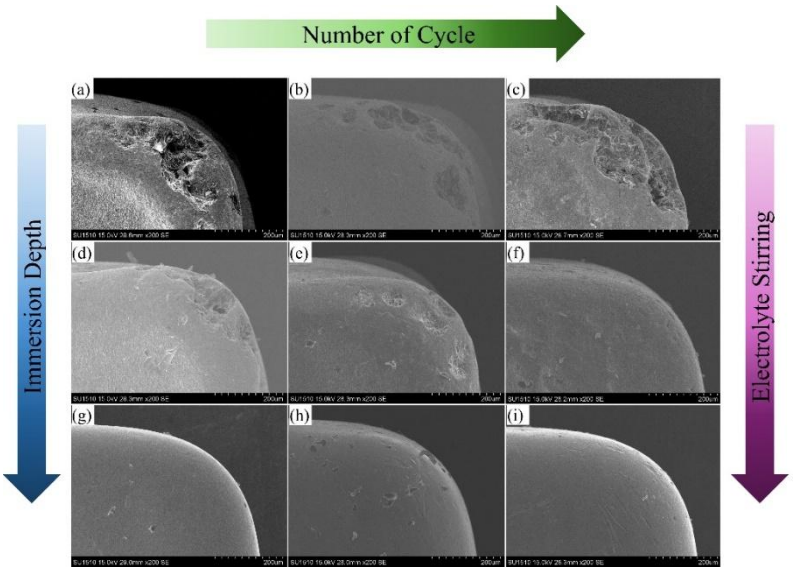

Figure 8. SEM images for the corner regions of the polished samples a) Exp. No 1, b) Exp. No 4, c) Exp. No 7, d) Exp. No 11, e) Exp. No 14, f) Exp. No 17, g) Exp. No 21, h) Exp. No 24 and i) Exp. No 27.

The corner regions in Figure 8 exhibited a similar trend with repeated usage of the electrolyte. After the first usage, large, irregular erosion cavities with a porous morphology formed along the edges, reaching their greatest extent after octuple immersion cycles (Figure 8c). After the second usage, these cavities became smaller and shallower. Only isolated pits

with porous rims remained, while the corner processed with octuple cycles (Figure 8f) was largely free of erosion damage. After the third usage, all corners were smoothly rounded without erosion cavities, apart from a few scattered shallow pits in Figure 8h. Repeated usage of therefore did not promote localized damage at the edges, which are the regions most susceptible to discharge concentration. As noted above, this improvement mainly reflects the concurrent changes in immersion depth and stirring speed rather than the reusage itself. Nevertheless, the absence of edge degradation after three consecutive uses supports the conclusion that the electrolyte retains its polishing performance under the investigated conditions.

In order to evaluate the elemental composition of the corner regions, EDX analyses were performed on the corners of the samples polished for Exp. No 1, No 11 and No 21 (Figure 9). Compared with the corresponding flat surfaces, the corner regions exhibited markedly lower Ti contents, together with higher F and Na concentrations. The increase in F was particularly pronounced at 30 mm depth with 300 rpm stirring and 50 mm depth with 600 rpm stirring, where the corner values (19.9 wt.% and 6.6 wt.%) were approximately four and eight times higher than those of the flat surfaces (4.6 wt.% and 0.8 wt.%). By contrast, the Al/Ti ratio at the corners (0.062 to 0.066) was comparable to that measured on the flat surfaces (0.063 to 0.073). The lower absolute Al contents at the corners therefore stem from dilution by surface elements rather than from selective dissolution of Al. These results suggest that fluoride-containing reaction products preferentially accumulated at the corners. In PeP, the local current density depends strongly on surface geometry and is concentrated at peaks features such as corners. Therefore, the corners are exposed to more intense electrochemical reactions, which likely promotes the local formation of fluoride and oxyfluoride products. For the Exp. No 1, the corner also exhibited high O and C contents (11.2 wt.% and 14.9 wt.%), which can be associated with the porous erosion cavities observed in Figure 8a that retain reaction products and contaminants. Notably, although the corner of the sample polished at Exp. No 21 displayed a smooth morphology without erosion cavities (Figure 8g), it still contained elevated O and F levels (14.7 wt.% and 6.6 wt.%). This indicates that a thin oxide or oxyfluoride film may form at the corners even when no morphological damage is visible.

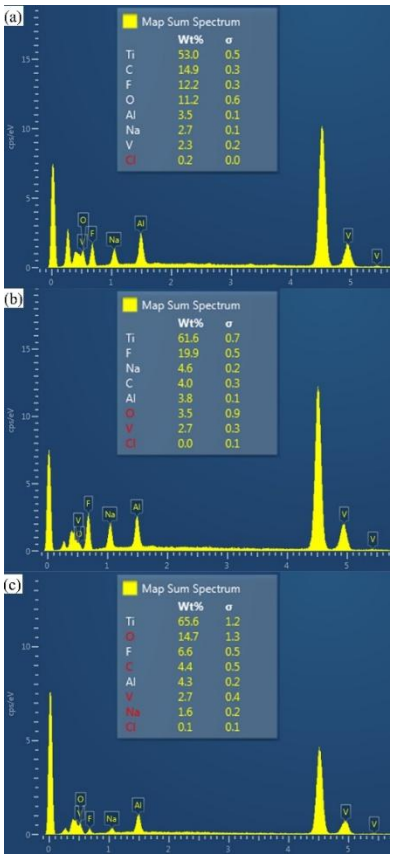

Figure 9. EDX analysis result for the corners area of polished samples a) Exp. No 1, b) Exp. No 11 and c) Exp. No 21.

3.3. Material Removal Rate

Figure 10 shows the MRR values obtained under the 27 experimental conditions. MRR ranged from 0.133 mg/s to 0.313 mg/s. However, 26 of the 27 runs fell within a relatively narrow band of 0.133 mg/s to 0.200 mg/s, and the average MRR was 0.170 mg/s. The only exception was Exp. No 21 (50 mm depth, single cycle, 600 rpm stirring), which yielded 0.313 mg/s, nearly twice the average MRR. Immersion depth had the most consistent influence on removal: the average MRR increased from 0.157 mg/s at 10 mm depth to 0.165 mg/s and 0.188 mg/s at 30 mm and 50 mm depth, respectively. This case can be attributed to the non-uniformity of the VGE along the immersion depth. Electrolyte stirring followed a similar trend, with the average MRR increasing from 0.160 mg/s in the immobile electrolyte (no stirring) to 0.186 mg/s at 600 rpm stirring. This increase can be associated with enhanced convective transport, which removes dissolution products from the vicinity of the specimen surface more effectively. In contrast, the immersion cycle had no systematic effect on MRR, with mean values of 0.177 mg/s, 0.161 mg/s, and 0.173 mg/s for quadruple, and octuple cycles, respectively.

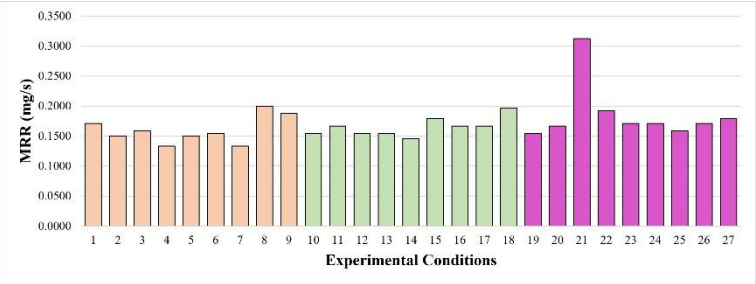

Figure 10. MRR values for all experimental conditions.

3.4. Current Time Characterization

Figure 11 shows the current-time curves recorded during PeP for single, quadruple, and octuple immersion cycles. In all experiments, the current rose sharply within the first 3 s to 5 s and reached peak values between 2.3 and 4.2 A. This initial transient corresponds to the stage preceding the formation of a continuous VGE, during which the sample surface is in direct contact with the electrolyte. Afterward, the current decreased gradually over approximately 100 s to 120 s before stabilizing at about 1 A, indicating the establishment of a stable polishing regime. For immersion process, the gradual decline can be attributed to the stabilization of the VGE, the progressive heating of the electrolyte, and the smoothing of surface irregularities. For quadruple and octuple immersion (Figure 11b and Figure 11c), the current dropped to zero during each up to Z-axis and was reestablished upon reimmersion. As the process progressed, this effect weakened, and the current converged to about 1.0 A to 1.3 A. The repeated current transients associated with VGE re-formation are consistent with the extensive edge erosion observed at octuple immersion. In addition, the highest current values under octuple immersion were recorded at 50 mm depth (Exp. No 9, No 18, and No 27), which agrees with the higher MRR obtained at greater immersion depths.

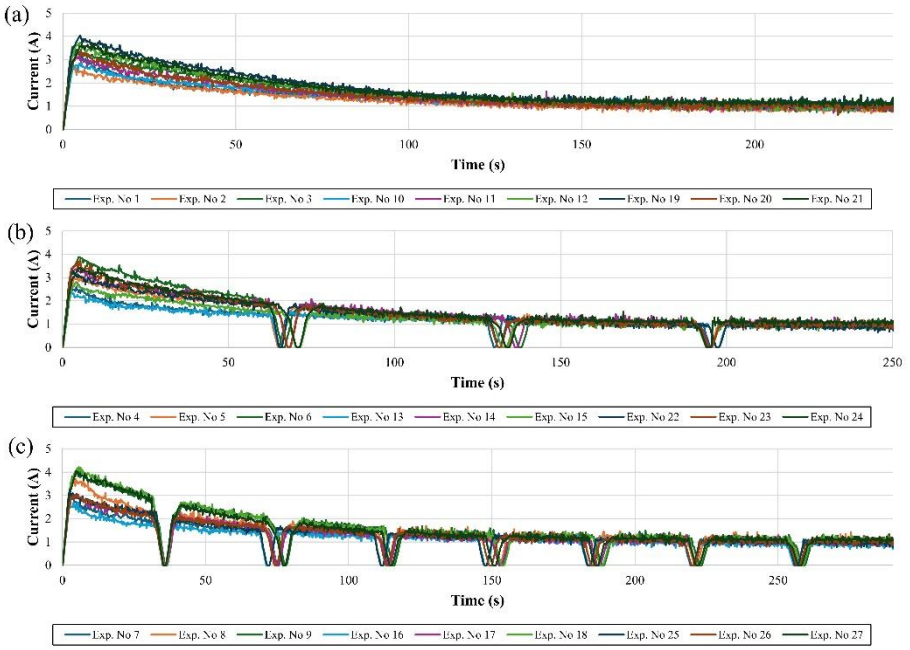

Figure 11. Current time graphs accordance with a) Single immersion, b) Quadruple immersion and c) Octuple immersion.

3.5.Machine Learning and Hybrid Model Analysis

The computational analyses were performed on a high-performance desktop workstation equipped with an Intel Core i9-13900K processor using Python 3.12 and conventional CPU-based processing. Statistical regression, machine learning, ensemble learning, Gaussian Process Regression, preprocessing, and cross-validation procedures were implemented primarily using the scikit-learn library, while XGBoost and LightGBM models were implemented using their respective Python packages. The SHAP package was used to examine the contributions of the predictor variables to model predictions. The computational framework comprised 23 base regression models and an additional target-specific hybrid approach, resulting in 24 predictive methods in the final comparison. Depth, Cycle, and Stirring were used as predictors for both R_a and Gloss, with each response modeled independently. Predictor standardization was fitted exclusively to the training observations within each evaluation fold. R_a was logarithmically transformed before model fitting, whereas Gloss was standardized using the corresponding training data. Predictions were subsequently returned to their original measurement scales before performance evaluation and hybrid construction.

The validation analysis used 27 experimental observations from a full 3×3×3 factorial design and compared shuffled three-fold, five-fold, and nine-fold cross-validation, three-fold and nine-fold cross-validation stratified by Stirring level, and leave-one-out cross-validation. Each randomized design was evaluated using ten random seeds, while leave-one-out cross-

validation was performed once. Following comparison of predictive performance and variability across seeds, nine-fold cross-validation stratified by Stirring level with a random seed of 0 was selected for the final analysis. Each validation fold contained three observations, including one from each Stirring level, while the remaining 24 observations were used for training. Model performance was evaluated using pooled out-of-fold predictions and fold-specific metrics for R_a and Gloss. The model configurations are summarized in Table 2. Internal cross-validation used for estimator tuning and stacking was distinct from the final nine-fold evaluation. Because the validation design, final seed, hybrid partner, and blending weight were selected using results from the same dataset, the reported performance does not constitute an independent evaluation of the complete selection procedure.

Table 2. Regression models and parameter configurations reported for predicting R_a and Gloss.

Models	Parameters
Linear Regression	Library default settings
Ridge Regression	alpha = 1.0
Lasso Regression	alpha = 0.01, max_iter = 20000
Bayesian Ridge Regression	Library default settings
Partial Least Squares Regression	n_components = min(2, n_features), corresponding to two components for the three predictors
K-Nearest Neighbors Regression	n_neighbors = 3, weights = "distance"
Support Vector Regression	kernel = "rbf", C = 10.0, epsilon = 0.05
Decision Tree Regression	max_depth = 4, random_state = 0
Random Forest Regression	n_estimators = 300, random_state = 0
Extra Trees Regression	n_estimators = 300, random_state = 0
Gradient Boosting Regression	n_estimators = 200, max_depth = 2, learning_rate = 0.05, random_state = 0
AdaBoost Regression	n_estimators = 150, random_state = 0
Bagging Regression	n_estimators = 150, random_state = 0
Stacking Regression	Base Random Forest with n_estimators = 150 and random_state = 0. Base Gradient Boosting with n_estimators = 150 and random_state = 0. Final estimator comprising Ridge Regression with alpha = 1.0. Internal stacking cross-validation with cv = 3
XGBoost Regression	n_estimators = 300, max_depth = 2, learning_rate = 0.05, random_state = 0, verbosity = 0
Gaussian Process Regression	Kernel defined as ConstantKernel(1.0) × RBF(length_scale = [1.0, 1.0, 1.0]) + WhiteKernel(noise_level = 0.1), with n_restarts_optimizer = 3, normalize_y = True, and random_state = 0
Second-Degree Polynomial Ridge Regression	PolynomialFeatures(degree = 2, include_bias = False) followed by Ridge Regression with alpha = 1.0
Cross-Validation-Tuned Ridge Regression	alphas = logspace(-3, 3, 25), corresponding to 25 logarithmically spaced values between 10^{-3} and 10^{3}. Other RidgeCV settings retained at their library defaults
Cross-Validation-Tuned Elastic Net Regression	l1_ratio = [0.1, 0.5, 0.9, 1.0], alphas = logspace(-3, 1, 25), cv = 3, max_iter = 20000, random_state = 0
Huber Regression	max_iter = 5000
Tuned Kernel Ridge Regression	kernel = "rbf" with grid-search optimization over alpha = logspace(-2, 2, 5) and gamma = logspace(-2, 1, 5), using cv = 3 and scoring = "neg_mean_squared_error"

The predictive models were developed independently for R_a and Gloss to characterize their relationships with Depth, Cycle, and Stirring. Model performance was assessed separately for each response across 23 base regression models and an additional target-specific hybrid

approach. Figure 12 and Figure 13 present the comparative predictive performance of the evaluated models using pooled out-of-fold R^2 values obtained through nine-fold cross-validation stratified by Stirring level with a random seed of 0. The figures illustrate differences in predictive accuracy among the statistical, machine learning, and ensemble regression approaches, allowing the highest-performing configurations to be identified for each surface quality response under the selected validation procedure. The results also demonstrate that model rankings differed between R_a and Gloss, supporting the evaluation of modelling strategies separately for surface roughness and optical appearance.

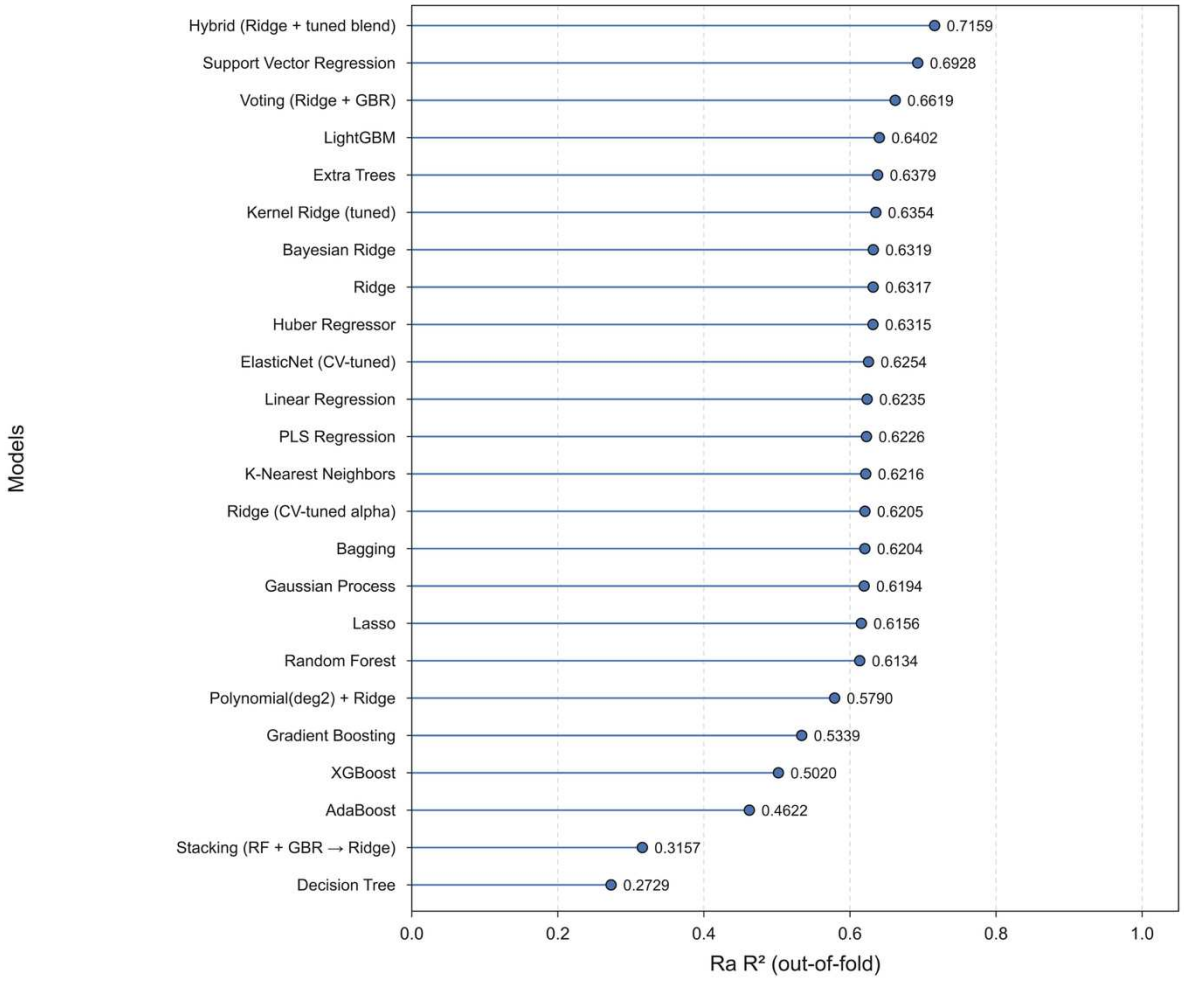

Figure 12. Comparative model performance for predicting R_a values based on the evaluated regression approaches.

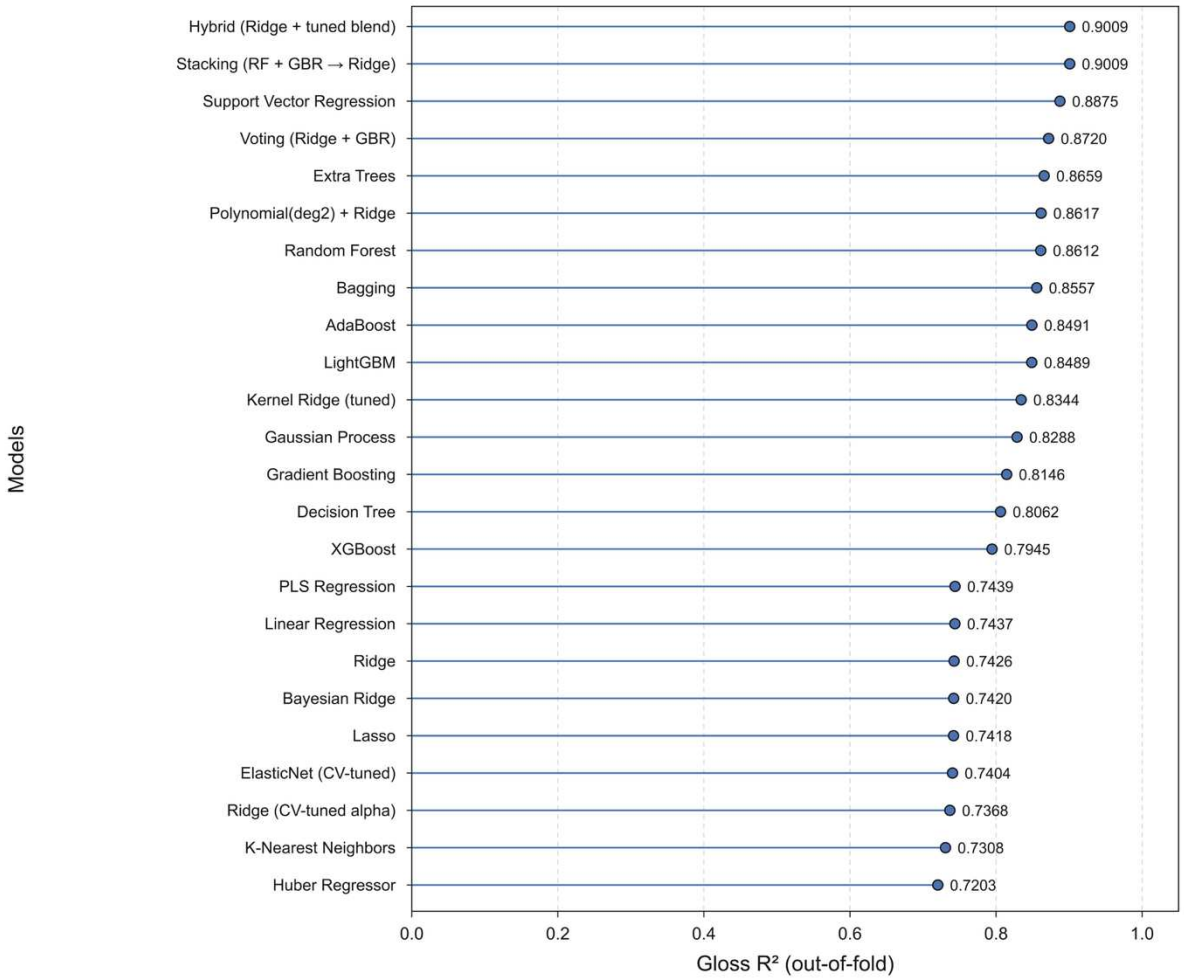

Figure 13. Comparative model performance for predicting the Gloss based on the evaluated regression approaches.

Table 3 and Table 4 present the fold-specific predictive performance of representative highest, intermediate, and lowest-ranked regression models for R_a and Gloss, respectively, using R^2, RMSE, and MAE. The models were selected according to their pooled out-of-fold R^2 rankings across the 24 evaluated approaches, and their performance was examined over nine validation folds stratified by Stirring level. For R_a, Hybrid Regression, K-Nearest Neighbors Regression, and Decision Tree Regression represented the highest, intermediate, and lowest-ranked approaches, with pooled R^2 values of 0.7159, 0.6216, and 0.2729, respectively. The hybrid model achieved fold-specific R^2 values ranging from −0.6237 to 0.9775, RMSE values between 0.0566 and 0.3828, and MAE values between 0.0501 and 0.3484. Although it achieved R^2 values above 0.95 in the third, fourth, and fifth folds, negative values occurred in the first and sixth folds, demonstrating sensitivity to the validation subset. K-Nearest Neighbors Regression also produced negative R^2 values in these two folds but exceeded the hybrid model in the second fold and marginally in the eighth fold. Decision Tree Regression exhibited weaker

overall performance, with negative R^2 values in four folds and a minimum of −3.7498 in the sixth fold.

For Gloss, Hybrid Regression, Gradient Boosting Regression, and Huber Regression represented the highest, intermediate, and lowest-ranked approaches, with pooled R^2 values of 0.9009, 0.8146, and 0.7203, respectively. The hybrid model achieved R^2 values above 0.87 in seven folds and reached 0.9724 in the eighth fold, while its RMSE and MAE ranged from 0.6550 to 1.8406 and from 0.4896 to 1.3515, respectively. However, all three models produced markedly negative R^2 values in the sixth fold, reaching −44.5713 for Hybrid Regression, −85.1310 for Gradient Boosting Regression, and −79.9063 for Huber Regression. The measured Gloss values in this fold were closely distributed at 52.32, 51.84, and 51.88, resulting in a small reference variance against which prediction errors were evaluated. Consequently, these extreme negative R^2 values did not correspond to the largest RMSE values observed across the folds. This finding illustrates the sensitivity of fold-specific R^2 to limited response variability within small validation subsets. Negative fold-specific values indicate poorer performance than prediction using the corresponding fold mean and can coexist with positive pooled R^2 because the pooled calculation uses the mean and variability of all 27 observations. The results therefore support interpreting R^2 alongside RMSE and MAE rather than using fold-specific R^2 alone to assess predictive performance.

Table 3. Fold-specific predictive performance of the selected regression models for the R_a values.

Folds		Models		
		Highest Hybrid (Ridge + tuned blend)	**Intermediate** K-Nearest Neighbours	**Lowest** Decision Tree
Fold 1	**R²**	−0.3418	−1.1758	−1.9815
	RMSE	0.3265	0.4157	0.4866
	MAE	0.3016	0.4054	0.4603
Fold 2	**R²**	0.1161	0.7913	−1.2015
	RMSE	0.2414	0.1173	0.3811
	MAE	0.2048	0.1005	0.3052
Fold 3	**R²**	0.9775	0.9066	0.6937
	RMSE	0.0566	0.1152	0.2087
	MAE	0.0550	0.0979	0.2009
Fold 4	**R²**	0.9749	0.9286	−0.1267
	RMSE	0.0566	0.0953	0.3788
	MAE	0.0501	0.0936	0.3611
Fold 5	**R²**	0.9518	0.7101	0.3589
	RMSE	0.1260	0.3089	0.4594
	MAE	0.1065	0.2659	0.3871
Fold 6	**R²**	−0.6237	−0.8455	−3.7498
	RMSE	0.2965	0.3161	0.5071
	MAE	0.2749	0.3121	0.4317
Fold 7	**R²**	0.5292	0.2935	0.2263
	RMSE	0.2689	0.3294	0.3447
	MAE	0.2055	0.2760	0.3070
Fold 8	**R²**	0.6414	0.6440	0.4113
	RMSE	0.3205	0.3193	0.4107

	MAE	0.3066	0.2816	0.3627
	R²	0.7326	0.6715	0.6302
Fold 9	**RMSE**	0.3828	0.4243	0.4502
	MAE	0.3484	0.3693	0.3855

Table 4. Fold-specific predictive performance of the selected regression models for the gloss values.

Folds		**Models**		
		Highest Hybrid (Ridge + tuned blend)	**Intermediate** Gradient Boosting	**Lowest** Huber Regression
	R²	0.8729	0.8777	0.5131
Fold 1	**RMSE**	0.8387	0.8225	1.6414
	MAE	0.6426	0.8132	1.6329
	R²	0.9047	0.8298	0.7384
Fold 2	**RMSE**	0.6550	0.8756	1.0855
	MAE	0.6266	0.5742	0.6766
	R²	0.9083	0.8475	0.8505
Fold 3	**RMSE**	1.0577	1.3636	1.3503
	MAE	0.8971	1.1016	1.3206
	R²	0.9717	0.8357	0.7311
Fold 4	**RMSE**	0.8451	2.0345	2.6029
	MAE	0.7414	1.7029	1.9089
	R²	0.9158	0.9139	0.4569
Fold 5	**RMSE**	1.0116	1.0229	2.5687
	MAE	0.8196	0.6809	2.2955
	R²	−44.5713	−85.1310	−79.9063
Fold 6	**RMSE**	1.4680	2.0182	1.9560
	MAE	1.1689	1.4269	1.4109
	R²	0.8831	0.7453	0.7521
Fold 7	**RMSE**	0.8866	1.3087	1.2911
	MAE	0.7569	1.2017	1.2399
	R²	0.9724	0.9305	0.8669
Fold 8	**RMSE**	0.7153	1.1343	1.5693
	MAE	0.4896	0.9162	1.2316
	R²	0.4905	0.2911	0.4658
Fold 9	**RMSE**	1.8406	2.1710	1.8845
	MAE	1.3515	1.7845	1.8309

Figure 14 illustrates the agreement between measured and predicted R_a values for the selected Hybrid Regression, K-Nearest Neighbors Regression, and Decision Tree Regression models. The hybrid predictions followed the overall experimental trend more closely, although overprediction remained evident at the lowest measured R_a value and underprediction occurred for several observations within the intermediate response range. K-Nearest Neighbors Regression exhibited a wider distribution around the line of equality, particularly where similar measured values corresponded to noticeably different predictions. Decision Tree Regression produced repeated prediction levels and substantial dispersion, reflecting its piecewise representation of the response. These patterns indicate that the hybrid approach provided a closer approximation of the measured roughness variation, while localized prediction errors remained across the investigated range.

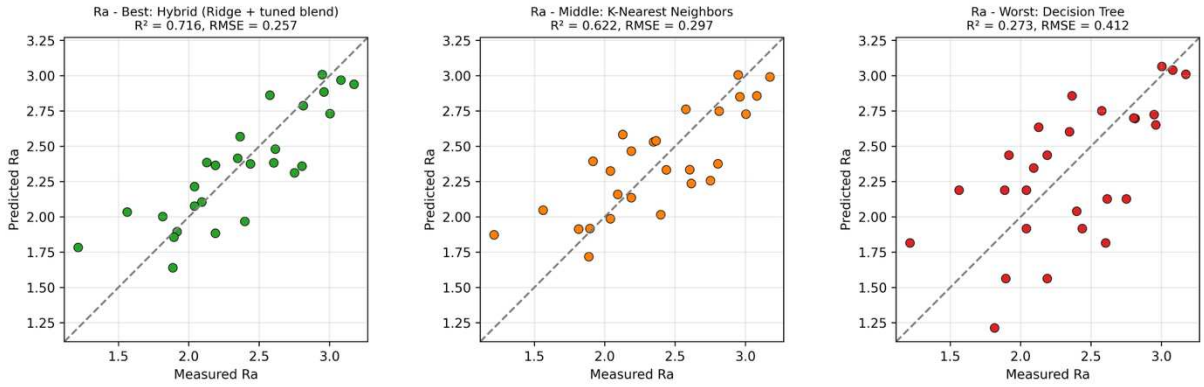

Figure 14. Measured and predicted R_a values obtained using the Highest-ranked, Intermediate-ranked, and Lowest-ranked models in the plotted comparison.

Figure 15 shows the corresponding measured and predicted Gloss values. The selected hybrid configuration, which produced predictions identical to Stacking Regression, closely followed the line of equality across much of the observed range, including the lowest Gloss values. Nevertheless, several observations at the upper end of the range were underpredicted, indicating that the model did not reproduce all experimental variations equally well. Gradient Boosting Regression captured the increasing response trend but displayed greater dispersion at intermediate and higher measured Gloss values, with both overprediction and underprediction evident. Huber Regression showed a distinct tendency to overpredict the lowest measured values and a broader distribution around the line of equality elsewhere. The visual comparisons therefore complement the fold-specific analysis by identifying where prediction deviations occurred within the measured response ranges.

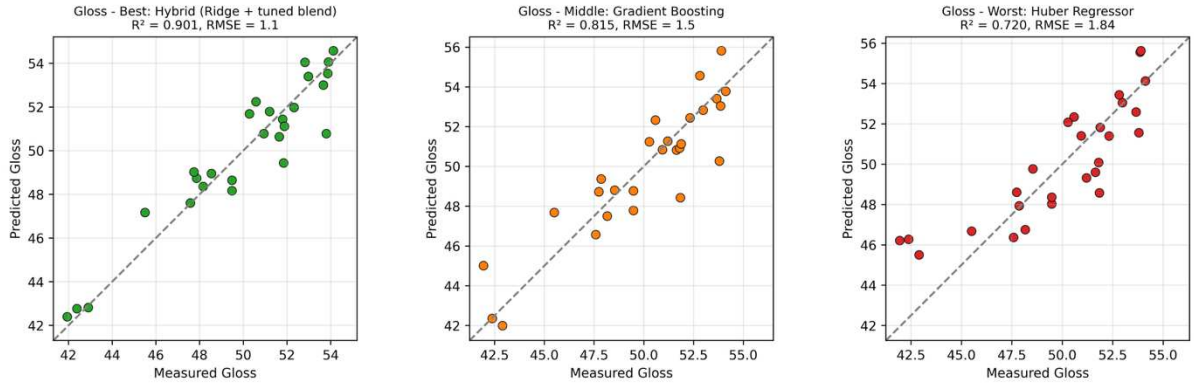

Figure 15. Measured and predicted gloss values obtained using the Highest-ranked, Intermediate-ranked, and Lowest-ranked models in the plotted comparison.

Figure 16 illustrates the contributions of the three process parameters to the predicted R_a values. Each point represents an experimental observation, with its horizontal position indicating the contribution of a predictor relative to the model's baseline prediction and its color

representing the corresponding predictor value. Positive SHAP values increase the predicted response, whereas negative values decrease it. Stirring exhibited the greatest mean absolute contribution, followed by Depth and Cycle. Higher Stirring and Depth values consistently produced negative contributions, while lower values produced positive contributions, indicating that the fitted model associated increased levels of these parameters with reduced surface roughness. Intermediate Depth values generally contributed less strongly, with SHAP values distributed near zero. Cycle exhibited a narrower distribution and a less uniform directional pattern, although higher Cycle values generally contributed to lower predicted R_a. These results indicate that Stirring and Depth accounted for the principal variations in the model predictions, while Cycle provided a smaller contribution within the investigated experimental domain.

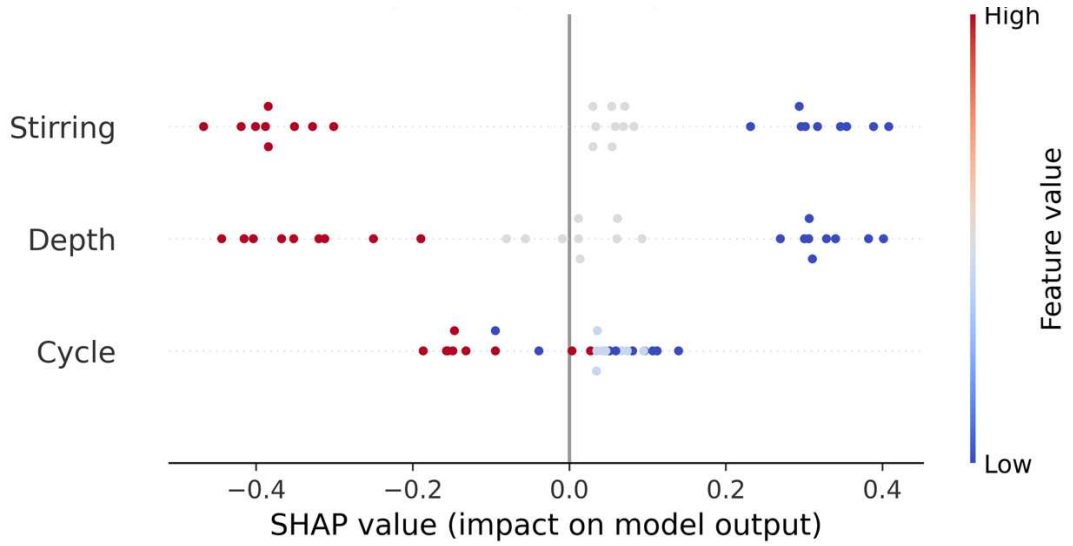

Figure 16. SHAP summary of the contributions of Depth, Cycle, and Stirring to R_a predictions from the selected hybrid regression model.

Figure 17 presents the corresponding SHAP summary for Gloss, with Stirring ranked as the most influential predictor, followed by Cycle and Depth. Higher Stirring values consistently increased predicted Gloss, whereas lower values reduced it, producing the widest distribution of SHAP contributions among the three predictors. Cycle displayed a similar directional pattern, with low values associated with negative contributions and higher values generally associated with positive contributions. Depth contributed comparatively little, as most of its SHAP values were concentrated near zero, although higher Depth values generally produced small positive contributions. The contrasting contribution patterns for R_a and Gloss indicate that the fitted models associated increased Stirring with both lower roughness and greater gloss, while the relative importance of Depth and Cycle differed between the responses. For Gloss, the selected hybrid configuration assigned its entire weight to Stacking Regression, so the displayed

attributions describe that model. These explanations were obtained after refitting the selected models to all 27 observations and characterize their learned prediction patterns rather than establishing independent causal effects.

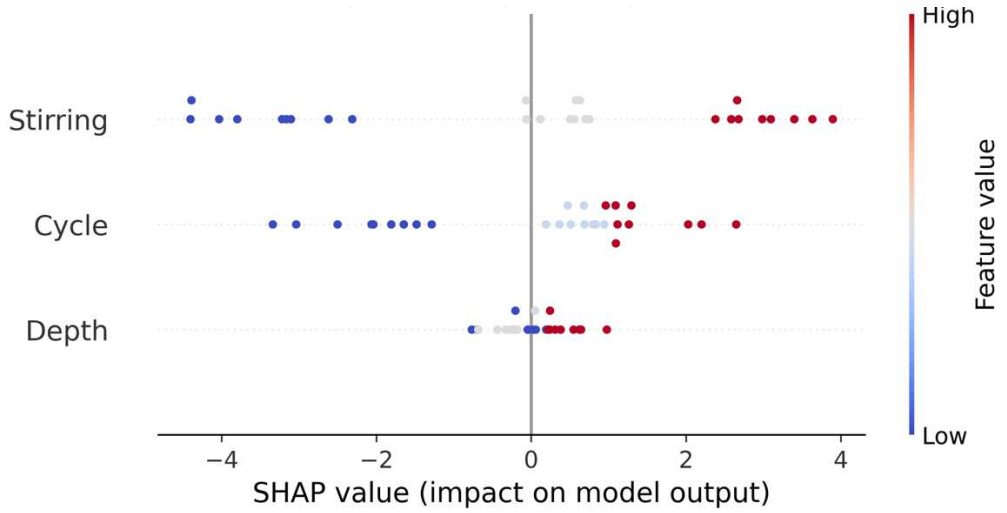

Figure 17. SHAP summary of the contributions of Depth, Cycle, and Stirring to Gloss predictions from the selected hybrid configuration.

4. CONCLUSIONS

This study examined the effects of electrolyte stirring, immersion depth, and immersion cycle on the PEP of as-WEDM Ti-6Al-4V without mechanical pretreatment, using a 27-run full factorial design and machine learning. The main conclusions are as follows.

PeP reduced R_a from 13.18 µm to between 1.214 µm and 3.173 µm. In addition, largely removed the WEDM surface irregularities under all experimental conditions. The lowest roughness (1.214 µm for R_a, 12.981 µm for R_z) was achieved at 50 mm depth, 600 rpm stirring, and a single cycle, whereas residual ridges and pits persisted at 10 mm depth without stirring. This improvement is attributed to the thinner VGE and higher local current density at greater depths, together with the continuous renewal of electrolyte at the anode surface provided by stirring.

EDX analysis showed no significant selective dissolution of the alloying elements, as the Al and V contents remained nearly constant in all samples. Fluorine residues reached 9.8 wt.% at 10 mm depth without stirring but decreased to 0.8 wt.% at 50 mm depth and 600 rpm stirring, indicating that stirring limits the retention of fluoride-containing reaction products, which tend to accumulate on rougher surfaces.

PeP rounded the sharp WEDM corners in all samples, owing to the higher current density at protruding features. At 10 mm depth without stirring, large erosion cavities formed along the edges and became most extensive at octuple immersion, despite this sample having the lowest R_a within its group. Increasing the immersion depth and stirring speed suppressed this localized damage, and smoothly rounded, no cavity corners were obtained at 50 mm depth and 600 rpm stirring.

Increased Stirring was associated with lower Ra and higher Gloss, while greater Depth was primarily associated with reduced R_a and higher Cycle values were more strongly associated with increased Gloss. SHAP analysis supported these response-specific patterns, identifying Stirring as the most influential predictor for both responses, followed by Depth for R_a and Cycle for Gloss. The selected hybrid configuration achieved pooled out-of-fold R^2 values of 0.7159 for R_a and 0.9009 for Gloss. For R_a, the configuration combined Ridge Regression and Support Vector Regression, whereas the Gloss configuration assigned its entire weight to Stacking Regression and therefore provided no additional improvement over that model.

Gloss exhibited stronger overall predictive agreement than R_a, while differences in model rankings demonstrated that increased complexity did not consistently improve accuracy. Substantial variation in fold-specific R^2, particularly where measured responses exhibited limited variability, emphasized the importance of interpreting this metric alongside RMSE and MAE. The limited dataset and the selection of the validation configuration and hybrid components using the same observations warrant independent evaluation before the models are applied to broader process optimization.

Acknowledgments: This study was funded by Samsun University Rectorate Scientific Research Projects Coordination Office (Grant No: BAP.MÜF.5502.2024.001).

Funding: This study was funded by Samsun University Rectorate Scientific Research Projects Coordination Office (Grant No: BAP.MÜF.5502.2024.001).

Author Contributions: A.Ç.: Conceptualization, Methodology, Investigation, Data curation, Visualization, Writing - original draft, Writing - review & editing. H.D.: Conceptualization, Supervision, Funding acquisition, Writing - review & editing. Ç.B.: Conceptualization, Supervision, Writing - review & editing. F.D.: Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing.

All authors have read and agreed to the published version of the manuscript.

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

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, the collection, analyses, or interpretation of data, the writing of the manuscript, or the decision to publish the results.

Declaration of Generative AI and AI-Assisted Technologies in the Manuscript Preparation Process: During the preparation of this work, the authors used Claude (Anthropic) to improve the language and readability of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

[1] Anil Kumar V, Gupta RK, Prasad MJN V, Narayana Murty SVS. Recent advances in processing of titanium alloys and titanium aluminides for space applications: A review. J Mater Res 2021;36:689–716. https://doi.org/10.1557/s43578-021-00104-w.

[2] Hao Y-L, Li S-J, Yang R. Biomedical titanium alloys and their additive manufacturing. Rare Met 2016;35:661–71. https://doi.org/https://doi.org/10.1007/s12598-016-0793-5.

[3] Gurrappa I. Characterization of titanium alloy Ti-6Al-4V for chemical, marine and industrial applications. Mater Charact 2003;51:131–9. https://doi.org/https://doi.org/10.1016/j.matchar.2003.10.006.

[4] Wang X, Xu S, Yu J, Chen G, Dong W, Shen W. Study on impact damage and failure mechanism of marine Ti-6Al-4V thin plate welded structure. Eng Fail Anal 2025;182:110114. https://doi.org/https://doi.org/10.1016/j.engfailanal.2025.110114.

[5] Avelar-Batista JC, Spain E, Fuentes GG, Sola A, Rodriguez R, Housden J. Triode plasma nitriding and PVD coating: A successful pre-treatment combination to improve the wear resistance of DLC coatings on Ti6Al4V alloy. Surf Coatings Technol 2006;201:4335–40. https://doi.org/https://doi.org/10.1016/j.surfcoat.2006.08.070.

[6] Yang H, Lu M, Zhao J, Han C, Li D, Sun B. Study on the Microstructure and Properties of TC4 Alloy Based on Water-Jet-Guided Laser Technology. Metals (Basel) 2025;15:1204. https://doi.org/10.3390/met15111204.

[7] Pramanik A, Littlefair G. Machining of Titanium Alloy(Ti-6Al-4V)—Theory to Application. Mach Sci Technol 2015;19:1–49. https://doi.org/10.1080/10910344.2014.991031.

[8] Belkin PN, Kusmanov SA, Parfenov E V. Mechanism and technological opportunity of plasma electrolytic polishing of metals and alloys surfaces. Appl Surf Sci Adv 2020;1:100016. https://doi.org/https://doi.org/10.1016/j.apsadv.2020.100016.

[9] Mohammad AEK, Wang D. Electrochemical mechanical polishing technology: recent developments and future research and industrial needs. Int J Adv Manuf Technol 2016;86:1909–24. https://doi.org/10.1007/s00170-015-8119-6.

[10] Ke X, Wu W, Wang C, Yu Y, Zhong B, Wang Z, et al. Material Removal and Surface Integrity Analysis of Ti6Al4V Alloy after Polishing by Flexible Tools with Different Rigidity. Materials (Basel) 2022;15:1642. https://doi.org/10.3390/ma15051642.

[11] Çebi A, Demirtaş H, Bolat Ç. On the brink of mass production and simultaneous multi-component surface modifications for ultra-fast polishing: Plasma electrolytic polishing process. J Manuf Process 2026;164:68–97. https://doi.org/https://doi.org/10.1016/j.jmapro.2026.02.047.

[12] Tian Y, Shi C, Fan Z, Zhou Q. Experimental investigations on magnetic abrasive finishing of Ti-6Al-4V using a multiple pole-tip finishing tool. Int J Adv Manuf Technol 2020;106:3071–80. https://doi.org/10.1007/s00170-019-04871-z.

[13] Fan Z, Tian Y, Zhou Q, Shi C. Enhanced magnetic abrasive finishing of Ti–6Al–4V using shear thickening fluids additives. Precis Eng 2020;64:300–6. https://doi.org/https://doi.org/10.1016/j.precisioneng.2020.05.001.

[14] Lu H, Wang D, Wu S, Pan Z, Wang G, Guo G, et al. A review of laser polishing on Ti6Al4V based on energy density. J Mater Process Technol 2024;331:118520. https://doi.org/https://doi.org/10.1016/j.jmatprotec.2024.118520.

[15] Zhou J, Liao C, Shen H, Ding X. Surface and property characterization of laser polished Ti6Al4V. Surf

Coatings Technol 2019;380:125016. https://doi.org/https://doi.org/10.1016/j.surfcoat.2019.125016.
[16] Deng C, Jiang L, Qian L. High-Efficiency Chemical Mechanical Polishing of Ti-6Al-4V Alloy via the Synergistic Action of H2O2 and K+ Under Alkaline Conditions. ECS J Solid State Sci Technol 2022;11:24005. https://doi.org/10.1149/2162-8777/ac495e.
[17] Deng C, Jiang L, Qin N, Qian L. Effects of pH and H2O2 on the chemical mechanical polishing of titanium alloys. J Mater Process Technol 2021;295:117204. https://doi.org/https://doi.org/10.1016/j.jmatprotec.2021.117204.
[18] Çebi A, Demirtaş H, Bolat Ç. Plasma Electrolytic Polishing of TC4 Alloys in Hydrochloric Acid Modified Novel Electrolyte: Effects of Electrolyte Reusability and Polishing Time on Surface Morphology. Coatings 2026;16:1125. https://doi.org/10.3390/coatings16091125.
[19] Yang D, Sun H, Wang J, Ji G, Duan H, Xiang Y, et al. The formation and stripping mechanism of oxide film on Ti6Al4V alloy surface during electrolytic plasma polishing. Surf Coatings Technol 2024;478:130469. https://doi.org/https://doi.org/10.1016/j.surfcoat.2024.130469.
[20] Wang J, Zhang L, Cheng J, Liu J, Zou Y, Wei D, et al. A precise machining surface obtained electrolyte plasma polishing on Ti–6Al–4V alloy: Microstructure and in-situ mechanical behaviors. J Manuf Process 2025;133:979–91. https://doi.org/10.1016/j.jmapro.2024.11.100.
[21] Tan XY, Ng Y, Meng TL, Sun CN, Huang Z, Ngo ACY, et al. Plasma electrolytic polishing-induced surface chemical and structural evolutions of additively manufactured Ti-6Al-4V coupons. Surf Coatings Technol 2024;494. https://doi.org/10.1016/j.surfcoat.2024.131557.
[22] Liu J, Liu L, Liu D, Xia G, Zhang H, Zhang P, et al. The influence of VGE characteristics on the plasma-electrochemical reaction pathways in plasma electrochemical polishing of TA4. J Manuf Process 2026;168:18–30. https://doi.org/https://doi.org/10.1016/j.jmapro.2026.04.006.
[23] Grigoriev SN, Mukhacheva TL, Tambovskiy I V, Kusmanova IA, Zaletova AA, Gaponov VA, et al. Electrolytic-Plasma Polishing of Titanium Alloy Ti6Al4V Additive Manufacturing Products. Surf Eng Appl Electrochem 2026;62:1–9. https://doi.org/10.3103/S1068375525701066.
[24] Zou Y, Sun J, Wang J, Xuan J, Wang S, Chen G, et al. Electrolyte plasma polishing: A strategy for surface and mechanical enhancement of additively manufactured Ti-6Al-4 V porous materials. Surf Coatings Technol 2026;521:133105. https://doi.org/https://doi.org/10.1016/j.surfcoat.2025.133105.
[25] Ghezri A, Nelis T, Burger J, Bessire C. Surface Finishing of Additive Manufactured Titanium Alloy by Plasma Electrolytic Polishing Without Pretreatments. Materials (Basel) 2025;18:4719. https://doi.org/10.3390/ma18204719.
[26] Wang J, Zhang L, Yang R, Huang Q, Li Y, Tian H, et al. Surface finishing of 3D-printed Ti–6Al–4 V alloy by electrolyte plasma polishing: Process optimization and microstructure. Mater Charact 2024;217:114442. https://doi.org/https://doi.org/10.1016/j.matchar.2024.114442.
[27] Lv Z, Wang L, Hu X, Bu Z, Li Y. Leveling mechanism of plasma electrolytic polishing of titanium alloy with aqueous fluoride-type solution. Int J Electrochem Sci 2024;19:100749. https://doi.org/10.1016/j.ijoes.2024.100749.
[28] Qi Junde, Zhang Dinghua, Li Shan, Chen Bing. Modeling and prediction of surface roughness in belt polishing based on artificial neural network. Proc Inst Mech Eng Part B J Eng Manuf 2018;232:2154–63. https://doi.org/10.1177/0954405416683737.
[29] Soler D, Telleria M, García-Blanco MB, Espinosa E, Cuesta M, Arrazola PJ. Prediction of Surface Roughness of SLM Built Parts after Finishing Processes Using an Artificial Neural Network. J Manuf Mater Process 2022;6:82. https://doi.org/10.3390/jmmp6040082.
[30] Pan Bo, He Zengxu, Guo Jiang, Cheung Benny, Wang Dongzhou. Modelling and optimization of surface roughness in chemical mechanical polishing based on DNN-GA. Proc Inst Mech Eng Part C J Mech Eng Sci 2023;237:3198–209. https://doi.org/10.1177/09544062221147132.
[31] Thasana W, Wetchakama W. Prediction of Gloss in Plastic Injection Parts Based on 3D Surface Roughness from Virtual Machining with Artificial Neural Networks. Int J Autom Technol 2022;16:138–48. https://doi.org/10.20965/ijat.2022.p0138.
[32] Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Stat Surv 2010;4:40–79. https://doi.org/10.1214/09-SS054.
[33] Huang Y, Wang C, Ding F, Yang Y, Zhang T, He X, et al. Principle, process, and application of metal plasma electrolytic polishing: a review. Int J Adv Manuf Technol 2021;114:1893–912. https://doi.org/10.1007/s00170-021-07012-7.
[34] Sushil K, Raj A, Jha SK, Ramkumar J, Chandraprakash C. Correlation of roughness with surface properties of SS 304 and Inconel 718 substrates processed by plasma electrolytic polishing. Surf Coatings Technol 2025;497:131770. https://doi.org/https://doi.org/10.1016/j.surfcoat.2025.131770.
[35] Belkin PN, Silkin SA, D'yakov IG, Burov S V, Kusmanov SA. Influence of Plasma Electrolytic Polishing Conditions on Surface Roughness of Steel. Surf Eng Appl Electrochem 2020;56:55–62. https://doi.org/10.3103/S1068375520010032.

[36] Navickaitė K, Roßmann K, Nestler K, Böttger-Hiller F, Penzel M, Grund T, et al. Plasma Electrolytic Polishing of Porous Nitinol Structures. Plasma 2022;5:555–68. https://doi.org/10.3390/plasma5040039.
[37] Wang C, Ding F, Li Y, Zhang T, He X, Huang Y, et al. Plasma electrolytic polishing for improving the surface quality of zirconium-based bulk metallic glass. Int J Adv Manuf Technol 2023;124:2079–93. https://doi.org/10.1007/s00170-022-10588-3.
[38] Ng Y, Tan XY, Meng TL, Sun C-N, Huang Z, Ngo ACY, et al. Material removal and surface finishing of additively manufactured Ti-6Al-4V coupons by cyclic process of plasma electrolytic polishing. Surf Coatings Technol 2025;498:131872. https://doi.org/https://doi.org/10.1016/j.surfcoat.2025.131872.
[39] Hoerl AE, Kennard RW. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 1970;12:55–67. https://doi.org/10.1080/00401706.1970.10488634.
[40] Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Ser B 1996;58:267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
[41] Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. J R Stat Soc Ser B Stat Methodol 2005;67:301–20. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
[42] Breiman L. Random Forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
[43] Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. Ann Stat 2001;29:1189–232.
[44] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004;14:199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.
[45] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., New York, NY, USA: Association for Computing Machinery; 2016, p. 785–794. https://doi.org/10.1145/2939672.2939785.
[46] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer; 2009. https://doi.org/10.1007/978-0-387-84858-7.
[47] MacKay DJC. Bayesian Interpolation. Neural Comput 1992;4:415–47. https://doi.org/10.1162/neco.1992.4.3.415.
[48] Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 2001;58:109–30. https://doi.org/https://doi.org/10.1016/S0169-7439(01)00155-1.
[49] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006;63:3–42. https://doi.org/10.1007/s10994-006-6226-1.
[50] Drucker H. Improving Regressors using Boosting Techniques. Proc. Fourteenth Int. Conf. Mach. Learn., San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 1997, p. 107–115.
[51] Breiman L. Bagging predictors. Mach Learn 1996;24:123–40. https://doi.org/10.1007/BF00058655.
[52] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6. https://doi.org/10.1038/323533a0.
[53] Wolpert DH. Stacked generalization. Neural Networks 1992;5:241–59. https://doi.org/https://doi.org/10.1016/S0893-6080(05)80023-1.
[54] Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning 2005. https://doi.org/10.7551/mitpress/3206.001.0001.
[55] Huber PJ. Robust Regression: Asymptotics, Conjectures and Monte Carlo. Ann Stat 1973;1:799–821.
[56] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. Proc. 31st Int. Conf. Neural Inf. Process. Syst., Red Hook, NY, USA: Curran Associates Inc.; 2017, p. 3149–3157.
[57] Bates JM, Granger CWJ. The Combination of Forecasts. OR 1969;20:451–68. https://doi.org/10.2307/3008764.
[58] Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics 2006;7:91. https://doi.org/10.1186/1471-2105-7-91.
[59] Danilov I, Hackert-Oschätzchen M, Zinecker M, Meichsner G, Edelmann J, Schubert A. Process Understanding of Plasma Electrolytic Polishing through Multiphysics Simulation and Inline Metrology. Micromachines 2019;10:214. https://doi.org/10.3390/mi10030214.
[60] Vana D, Podhorsky S, Hurajt M, Hanzen V. Surface Properties of the Stainless Steel X10 CrNi 18/10 after Aplication of Plasma Polishing in Electrolyte . Int J Mod Eng Res 2013;3:788–92.
[61] Wang J, Zhang L, Cheng J, Liu J, Zou Y, Zhang C, et al. Optimization of electrolytic plasma polishing process and surface performance analysis for 6061 aluminum alloy. Mater Today Commun 2024;40:110162. https://doi.org/https://doi.org/10.1016/j.mtcomm.2024.110162.
[62] Valentinčič J, Koroth JE, Zeidler H. Advancements in surface finish for additive manufacturing of metal parts: a comprehensive review of plasma electrolytic polishing (PEP). Virtual Phys Prototyp 2024;19:1–23. https://doi.org/10.1080/17452759.2024.2364222.