Data-Driven Optimization of Plasma Electrolytic Oxidation (PEO) Coatings with Explainable Artificial Intelligence Insights

Patricia Fernández-López [1,*], Sofia A. Alves [1], Aleksey Rogov [2,3], Aleksey Yerokhin [2,3], Iban Quintana [1], Aitor Duo [4] and Aitor Aguirre-Ortuzar [4]

1   Plasma Coatings Technologies Unit, Tekniker, Parke Teknologikoa, C/Iñaki Goenaga, 5, 20600 Eibar, Spain; sofia.alves@tekniker.es (S.A.A.); iban.quintana@tekniker.es (I.Q.)
2   Department of Materials, University of Manchester, Oxford Road, Manchester M12 9PL, UK; aleksey.rogov@manchester.ac.uk (A.R.); aleksey.yerokhin@manchester.ac.uk (A.Y.)
3   Henry Royce Institute, University of Manchester, Manchester M12 9PL, UK
4   Faculty of Engineering, Mondragon Unibertsitatea, Loramendi 4, 20500 Arrasate-Mondragon, Spain; aduo@mondragon.edu (A.D.); aaguirre@mondragon.edu (A.A.-O.)
*   Correspondence: patricia.fernandez@tekniker.es

Abstract: PEO constitutes a promising surface technology for the development of protective and functional ceramic coatings on lightweight alloys. Despite its interesting advantages, including enhanced wear and corrosion resistances and eco-friendliness, the industrial implementation of PEO technology is limited by its relatively high energy consumption. This study explores the development and optimization of novel PEO processes by means of machine learning (ML) to improve the coating thickness. For this purpose, ML models random forest and XGBoost were employed to predict the thickness of the developed PEO coatings based on the key process variables (frequency, current density, and electrolyte composition). The predictive performance was significantly improved by including the composition of the used electrolyte in the models. Furthermore, Shapley values identified the pulse frequency and the TiO_2 concentration in the electrolyte as the most influential variables, with higher values leading to increased coating thickness. The residual analysis revealed a certain heteroscedasticity, which suggests the need for additional samples with high thickness to improve the accuracy of the model. This study reveals the potential of artificial intelligence (AI)-driven optimization in PEO processes, which could pave the way for more efficient and cost-effective industrial applications. The findings achieved further emphasize the significance of integrating interactions between variables, such as frequency and TiO_2 concentration, into the design of processing operations.

Keywords: plasma electrolytic oxidation (PEO); machine learning; prediction models; cast Al-Si alloys; coating thickness; process digitalization

1. Introduction

Plasma electrolytic oxidation (PEO) is an advanced surface engineering technology used to produce protective ceramic coatings on lightweight alloys such as Al, Ti, and Mg. This electrochemical technique involves the application of a high voltage between the surface of the metallic material and a cathode, both submerged in an electrolyte, resulting in the generation of microarc discharges at the interface between the material and the electrolyte [1,2]. By employing aqueous-based alkaline electrolytes, PEO processes are eco-friendly and induce to the growth of ceramic-like films with high hardness and thickness and enhanced wear and corrosion properties [3,4]. Several parameters that control this process, including electrical parameters and electrolyte composition, play a major role in critical coating features such as thickness, morphology, topography, density, porosity, hardness, crystalline structure, and chemistry, which ultimately influence properties such as wear, corrosion, and tribocorrosion behavior [5–8].

Nowadays, the implementation of PEO technology remains restricted due to the relatively high costs associated with using high current densities (up to 10–100 A·dm^{-2}) and high voltages (up to 400–1000 V). This results in a high energy consumption, which increases the cost of power supplies and cooling equipment [9–11]. Novel approaches have been increasingly employed to solve these problems, based on the application of special current regimes using bipolar pulses in the kHz range [12,13]. Pulsed bipolar polarization stands out among the existing electrical regimes since, by means of time-varying current waveforms that alternate between anodic and cathodic pulses, this regime enables the formation of coatings with enhanced performance [12,14]. More specifically, several studies have been shown that the use of pulsed bipolar current results in a better-quality coatings, particularly in terms of thickness and compactness, which strongly impact coatings properties [7,15]. However, the design of PEO processes that produce coatings with multiple functionalities while displaying low energy consumption, and therefore a feasible cost to be implemented at industrial scale, is still missing. Pulsed bipolar power supplies have increasingly shown a high potential to overcome these limitations.

Pulsed bipolar power supplies allow for a more specific process design, enabling the development of PEO coatings with a wider range of properties to meet more precise requirements [16,17]. Although the simultaneous design and optimization of all parameters to be adjusted in these power supplies is challenging, especially considering the nonlinear multiscale nature of PEO, proper control of the entire process would allow for a more comprehensive and efficient design of PEO coatings according to different requirements [18]. Therefore, the development of artificial intelligence (AI) approaches to optimize and digitalize PEO processes becomes particularly relevant. In order to carry out these actions, the collection of adequate process data is essential, together with the selection of the most relevant variables to be optimized [18]. For this aim, it is necessary to properly collect PEO process data, focusing on the most influential process variables (such as potentials, currents, and temperature), and to link the data from each process to the characterization of the corresponding coating.

Given the nonlinear characteristics of the aforementioned process, there is a body of work on modelling and explaining nonlinear processes using ML techniques. Finke et al. [19] aimed to relate the time required for the visual appearance of corrosion and the extent of corrosion to the electrochemical input parameters to the process, including three electrolyte compositions using an artificial neural network (ANN), which showed 90% accuracy in predicting coating corrosion after 200 h of accelerated salt spray testing. Tagirova et al. [18] also recently developed ML and neural network models to predict the thickness and other properties of PEO coatings. The neural network model for calculating the coating thickness was developed based on the approximation of the experimental dependencies of the coating thickness on time and the current parameters. The results of these experiments have shown that the thickness of PEO coatings was dependent on the process duration, being observed that for the first 400 s of the process, as the duration increased, the thickness increased, and then the thickness remained constant with time. However, no predictions were made between the coating thickness and the frequency applied during the PEO process.

Furthermore, Rodriguez et al. [20] highlighted the difficulties faced by the scientific community as a result of differences in coating production between different research teams, due to the complex mechanisms involved in coating formation. Therefore, they proposed the use of AI techniques for data processing to increase the knowledge of the relations between the input parameters and the tribological aspects of the coating. Similarly, other studies focus on the use of neural networks to predict coating thickness over other processes. Varol et al. [21] used ANN to simulate the milling process and were able to determine that an increase in milling time resulted in an increase in coating thickness. Thus, ML and AI methods can help to understand and optimize the coating layer formed by PEO processes. To the best of the authors' knowledge, there are no publications using ML methods for the development of PEO coatings.

One of the major drawbacks of some ML/AI algorithms is that they behave as black boxes, limiting the interpretability of the importance of individual input data. A large body of literature has grown around the theme of the use of AI explainability tools such as model-agnostic explanations (LIME) or Shapley additive explanations (SHAP). A number of studies have shown that these tools increase the transparency of predictions. Baptista et al. [22] suggested that LIME explanations are not stable with nonlinear models. Therefore, the authors of the present work employed SHAP values due to the nonlinear nature of the process. These tools allow for the understanding of the direction and magnitude of each feature's impact on the output.

The main objective of this work involved the development and evaluation of AI approaches for the optimization of the PEO process, digitalizing the information generated, and linking it to the process for future analytical use. On the other hand, this research carried out the selection of appropriate inputs and ML methods to build numerical predictive models. In addition, the predictive capability of predictive ML models was used to predict materials or properties not previously studied, as well as the use of explainable artificial intelligence (XAI) methods for the explicability of the models created, promoting the understanding of the properties of the predicted materials.

2. Materials and Methods

2.1. PEO Treatments

In the present work, PEO coatings were grown on a secondary cast Al-Si substrate (A380 | EN AC-46500) that was recycled from scrap and manufactured by high-pressure die-casting (HPDC). Cast Al-Si samples were machined into 24 and 54 mm diameter discs with surface areas of 15.08 and 22.90 cm^2, respectively. Before the PEO treatments, all the discs were rinsed with distilled water and cleaned with isopropanol, without any other pre-treatment or polishing. It should be noted that the PEO process parameters were adjusted to accommodate the different sample diameters (24 mm and 54 mm discs), ensuring consistent thickness results across varying sample sizes, both in the center and the edges.

The PEO processes were performed under pulsed bipolar polarization, applied by a two GX-series DC power supplies (Analoge & Digitale Leistungselektronik GmbH, Darmstadt, Germany) connected through a MAGPULS 1000/35/200 bp-as (Magplus Stromversogungen GmbH, Kartung, Germany). The PEO process was driven by means of symmetric bipolar pulses with a duty cycle of 40%. The treatments were carried out in 2 L and 5 L homemade electrochemical cells equipped with a steel cathode, while the immersed treated samples acted as anodes.

Several electrolytes were formulated and used in the development of the PEO processes. These electrolytes were composed of the different combinations of the following reagents: KOH, Na_2HPO_4, Na_2WO_4, $Na_2WO_4 \cdot 2H_2O$, $Na_4P_2O_7 \cdot 10H_2O$, $NaAlO_2$, TiO_2 (size particle: 100–200 nm), and α-Al_2O_3. Each electrolyte was prepared by dissolving the reagents in distilled water under magnetic stirring, and the electrolytic baths were continuously stirred during the PEO treatments.

The average thickness of each developed coating was measured by using a Dual-Scope device (Fischer Technology Inc., Waltham, MA, USA).

2.2. Data Pre-Processing and ML Model Building

2.2.1. Input Variables and Pre-Processing

For the generation of the ML models, (i) electrolyte composition; (ii) electrical parameters, including frequency, J_{anodic}, $J_{cathodic}$, R, I_{anodic}, $J_{cathodic}$, and duration of the initial ramp in anodic potential; and (iii) process parameters, time, and applied area were considered. These parameters were chosen to build the models, as they most significantly influence the properties of the resulting PEO coatings. The acronyms used, their description, and measuring units can be found in Table 1.

Table 1. Employed PEO process related variables for thickness model generation.

Variable	Description	Units
Electrolyte composition	Sum of electrolyte composition elements (e.g.: "10 g/L NaAlO$_2$ + 10 g/L Na$_2$HPO$_4$")	g·L^{-1}
Frequency	Pulse frequencies applied during the processes (i.e., 1000; 500; 200; 100; 50 Hz)	Hz
J$_{anodic}$	Anodic (+) current densities applied during the processes (i.e., 213; 175; 137.5; and 125 mA·cm^{-2})	mA·cm^{-2}
J$_{cathodic}$	Cathodic (−) current densities applied during the processes (i.e., 100; 125; 137.5; 140; 150; 175; 210; 227.5; 245; and 255.6 mA·cm^{-2})	mA·cm^{-2}
R	Relation between the cathodic and anodic current densities: R = Jc/Ja (i.e., 0.8; 1.0; 1.2; 1.3; 1.4)	-
I$_{anodic}$	Anodic (+) currents applied during the processes (i.e., 1885; 2639; 2862.5; 3148.75; 3212 mA)	mA
I$_{cathodic}$	Cathodic (−) currents applied during the processes (i.e., 1508; 1885; 2111.2; 2262; 2290; 2639; 2862.5; 3148.7; 3166.8; 3430.7;3435; 3694.6; 3854.5 mA)	mA
Time	Duration of the PEO processes (i.e., 20; 30; 36; 40; 45 min)	min
Area	Total coated areas (i.e., 15.08; 22.90 cm^2)	cm^2
Duration of the initial ramp of potential	Duration of the initial ramp of potential applied at the beginning of the processes (V+ = 300 V and V− = 50 V) (i.e., 15; 30 s)	s
Thickness	Coating layer thickness—this is the ML model output	μm

Regarding the pre-processing of the data used, it was decided to create one column per constituent element found in the electrolyte, which means having as many columns as there are elements in the dataset (also known as one-hot encoding): KOH, Na$_2$HPO$_4$, Na$_2$WO$_4$, Na$_2$WO$_4$·2H$_2$O, Na$_4$P$_2$O$_7$·10H$_2$O, NaAlO$_2$, TiO$_2$, and α-Al$_2$O$_3$. In case of the absence of one of the components in the test performed, the value of that chemical element for that specific test shall be 0. This ensures that each element is represented independently, preventing the model from interpreting any ordinal relationship between the elements.

Besides this initial pre-processing, exploratory data analysis (EDA) was employed to uncover underlying patterns, identify anomalies, and test hypotheses, setting the stage for precise model development in digitalizing the PEO process.

2.2.2. Tested Models

In this study, an ML approach is used to predict the thickness of a PEO coating. As a first step, a model selection process is applied first. Model selection involves comparing and evaluating different models created with different ML algorithms to determine which one provides the best results based on certain established criteria. As for the algorithms used, the random forest and XGBoost algorithms were selected, both in their regression versions. These two models are two of the most commonly used in classic ML projects and where more advanced techniques such as deep learning cannot be used due to lack of data:

Random forest is an ensemble learning method that constructs multiple decision trees during the training phase and outputs the mean prediction of individual trees in case of regression problems. Random forests correct for decision trees' habit of overfitting to their training set by introducing randomness in two ways: firstly, by selecting a random subset of the training data to build each tree (bootstrap sampling), and secondly, by choosing a random subset of features to consider at each split in the tree.

XGBoost is another ensemble learning algorithm. However, unlike random forests, its core functioning is based on the boosting technique. This is an ML method that combines multiple weak learning models to build a stronger model. Boosting is based on the idea of building a sequence of models, where each model tries to correct the errors of the previous model. This error or loss minimization is carried out by using advanced techniques such as gradient descent.

The training of the models for subsequent selection was carried out using cross-validation. This validation is a technique used in data analysis and ML to evaluate and validate the performance of a model. It consists of dividing the dataset into multiple subsets and performing several iterations of training and evaluation, using different combinations of subsets as training and test data. This provides a more accurate and reliable estimate of model performance, reducing the bias introduced by a single data split. This validation has been used both in the model selection phase and in the hyperparameter optimization where it has been performed.

The comparison of the models was carried out based on two regression metrics (RMSE and R2), selecting the best one for the subsequent optimization of the final model. This optimization, also known as hyperparameter optimization, is a process in which the hyperparameter values of the ML algorithms are found and adjusted in order to maximize the performance of the model based on a specific evaluation metric.

The Hyperopt (Hyperopt: Distributed Asynchronous Hyper-parameter Optimization, https://hyperopt.github.io/hyperopt (accessed on 24 July 2024)) library was used to optimize the hyperparameters of the generated models. The hyperparameters are set before training a model and determine its performance and behavior. Unlike traditional grid search, which methodically tests all possible combinations, Hyperopt leverages Bayesian optimization for a more strategic search. Bayesian optimization is a sequential optimization method that uses probabilistic models to guide the search. It automatically explores the hyperparameter space intelligently, focusing on the areas that are likely to produce the best results.

For each set of hyperparameters evaluated during the optimization process, a cross-validation with K = 5 is applied. Cross-validation consists in separating the dataset into a number K of subsets, in this case into 5 subsets of data. The algorithm will be trained on 4 of these subsets and tested on the remaining subset, and repeated K times until all subsets have been used to test the model. The objective during model optimization is to minimize the aggregation of the 5 metrics obtained during cross-validation. In this case, the metric to be minimized was the root mean square error (RMSE). It measures the average difference between values predicted by a model and the actual values, a critical indicator of regression model performance.

For training, selection and validation of the generated models, the data were separated into a training set, which constitutes 80% of the data (56 samples), while 20% of the data (14 samples) were used for the model validation. The model selection and optimization processes were performed on 80% of the data selected for training while 20% were used for final model validation.

One of the tests carried out in terms of variable selection was to evaluate the impact of introducing or not the concentrations of the electrolyte elements to generate the model. In the case of introducing a new component in the composition, the model has to be re-trained to consider this new feature.

A residual analysis was performed for final model validation. Residual analysis involves the examination of the differences between the actual values measured in the

laboratory and the predictions made by the models. The objective of the residual analysis is to detect patterns in the predictions made and to propose possible areas for improvement of the models created.

For the identification of the most influential variables or parameters on the output variables of the models, Shapley values were obtained. In the context of machine learning, Shapley [23] values are used to identify the most influential parameters or features in model prediction. By attributing a value to each feature based on its impact on model output, Shapley values help to understand which features contribute most to model performance. In the case of identifying the most influential variables or parameters for model output variables, Shapley values can be calculated to determine the relative importance of each parameter in driving the model prediction toward a more positive value or a more negative value.

All the above methods were employed to model the effects of PEO process parameters and predict coating features. The main results achieved employing these methods are described and discussed in the next section.

3. Results and Discussion

3.1. Modelling of PEO Processes

In this work, models were developed to correlate the multiple-input PEO process variables with critical features of the resulting PEO coatings: the coating thickness. Among all the critical features related to PEO coatings (e.g., thickness, roughness, hardness, morphology, microstructure), the thickness was selected to be predicted by the current model (Figure 1) after considering it as a representative key variable strongly related to the functionality of the coating, such as wear and corrosion resistance.
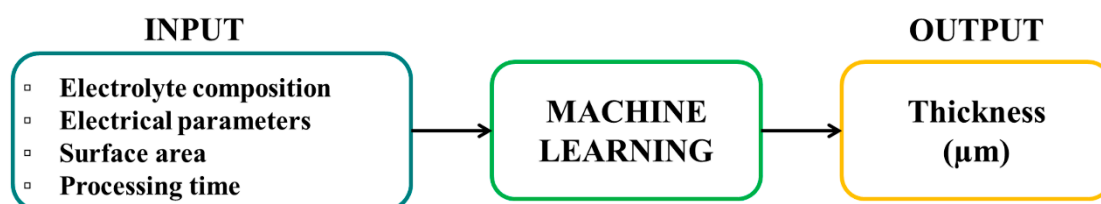

Figure 1. Schematic illustration on ML models' application to predict PEO coating thickness.

The data used to carry out this work were obtained by performing several PEO processes with different process parameters (e.g., frequency) in order to study their influence on certain coating properties, such as thickness. Figure 2 is a representative example of some of the PEO processes carried out in this study, showing the evolution of anodic and cathodic potentials and the appearance of plasma microdischarges during the PEO processes performed at frequencies of 50 Hz, 200 Hz, and 1000 Hz. These processes were carried out under a current ratio, $R = J_{cathodic}/J_{anodic}$, of 1.2.

As can be depicted from Figure 2, the anodic potential decreases earlier during the process as the frequency increases. This decrease in anodic potential during the process is known as "soft sparking" and represents a specific regime of thinner and softer plasma discharges that has been correlated with the development of coatings of higher density and smaller pore size, along with a reduction in the energy consumption of the process [24,25]. Despite the already demonstrated benefits of soft-sparking, the transition to this regime is undesirable when the potential drop is excessive, as this would cause known issues with corrosion properties [26] or even coating delamination [27]. In particular, the PEO process carried out at the highest frequency (1000 Hz, R = 1.2) revealed a significant decrease in the intensity and density of the discharges after 690 s of process. This decrease in the intensity of the discharges was also observed in the process carried out at 200 Hz and R = 1.2, although in this case, it did not appear until 1600 s into the process. In contrast, the process carried out at the lowest frequency, 50 Hz and R = 1.2, did not show this drop in potential or change in microdischarge (Figure 2). This may be related to the complex reactions that take

place during PEO processes performed at different frequencies, as discussed in previous works [28,29].
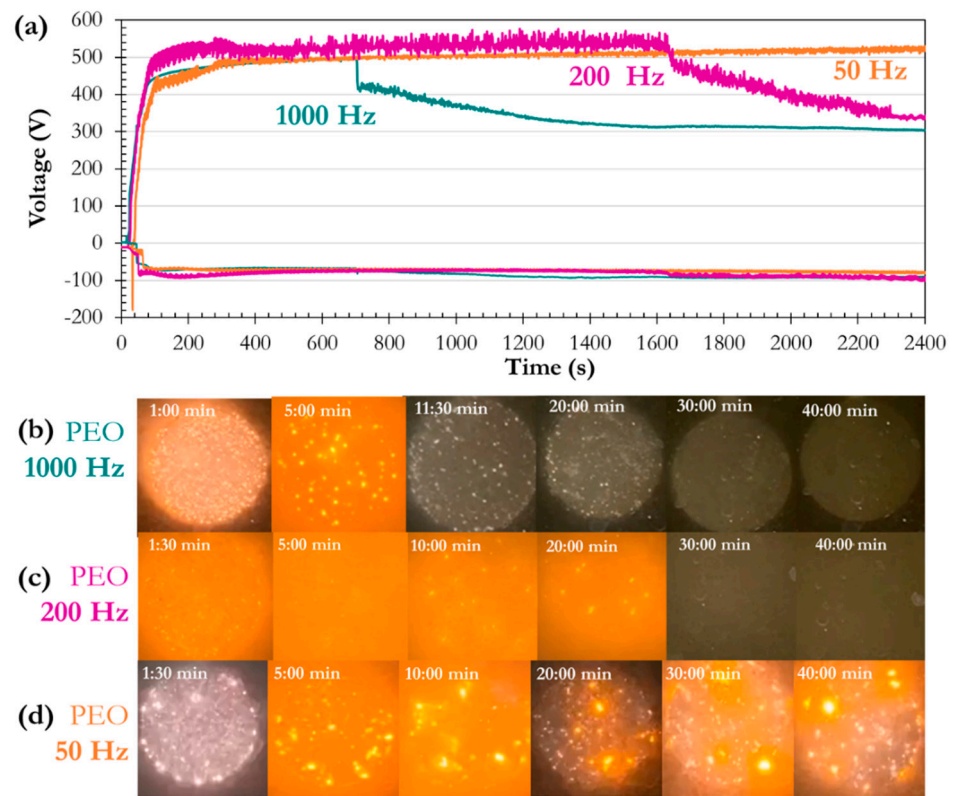
Figure 2. (**a**) Evolution of the anodic (+) and cathodic (−) potentials during the PEO processes carried out at 1000 Hz, 200 Hz, and 50 Hz, under a constant current ratio of 1.2. Appearance of the plasma microdischarges during the PEO processes carried out applying frequencies of (**b**) 1000 Hz, (**c**) 200 Hz, and (**d**) 50 Hz.

Thickness measurements of the resulting coatings developed under different PEO process conditions were performed to be used later in the ML models to predict their trend. Since each of the thickness values were measured in multiple replicates to ensure the accuracy and consistency of the data collected, no outliers or out-of-range values were considered in either the input variables or the thickness measurements.

Figure 3 shows the kernel density estimation of the variable thickness, together with the mean, first quartile, and third quartile of the mentioned variable. When analyzing the distribution of the data, it is observed that approximately 50% of the thickness measurements obtained in this study are concentrated in the range of 50 to 65 μm; this is equivalent to the interquartile range of analyzed data, which is defined as the zone between the first (Q1) and third (Q3) quartiles. In addition, approximately 25% of the thicknesses are below 50 μm (below Q1), while another 25% are above 65 μm (above Q3).

The correlations between thickness and various parameters demonstrate that TiO_2 (0.6) and frequency (−0.6) exert the most significant influences, with TiO_2 exhibiting a positive and frequency a negative effect on thickness. Moderate positive correlations are observed for Na_2HPO_4 (0.4), Na_2WO_4 (0.5), $NaAlO_2$ (0.4), and time (0.4), indicating that these parameters contribute to an increase in thickness. Conversely, moderate negative correlations are observed for J_{anodic} (−0.4), which implies a decrease in thickness. Weak correlations are seen with KOH (−0.3), $Na_2WO_4 \cdot 2H_2O$ (−0.3), and IRA (0.3), which suggests minor influences on thickness. Overall, TiO_2 and frequency are the most significant factors affecting thickness.
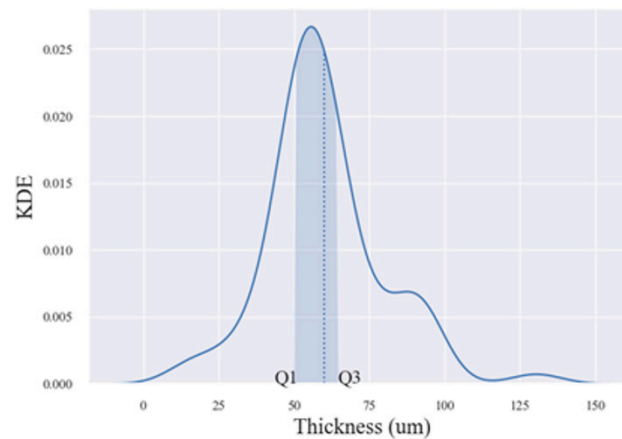
Figure 3. Kernel density estimation of target variable (thickness).

The moderate positive linear relationship of frequency suggests a relationship between lower pulse frequencies and thicker coatings during the PEO process. This insight may be very useful to select the range of frequency values to be used in a PEO process that aims to produce coatings with specific thickness values. This finding has significant practical implications, particularly in industrial applications where stringent dimensional requirements must be met.

For the selection of variables to build the models, the results returned by the random forest model were observed. Random forest is preferred for variable selection because of its simple feature importance ranking, which allows for easy identification of key predictors. For this, it was considered to introduce or not the concentrations of the electrolyte elements. Table 2 shows the results (RMSE and R2) including and excluding the electrolyte elements and without considering them.
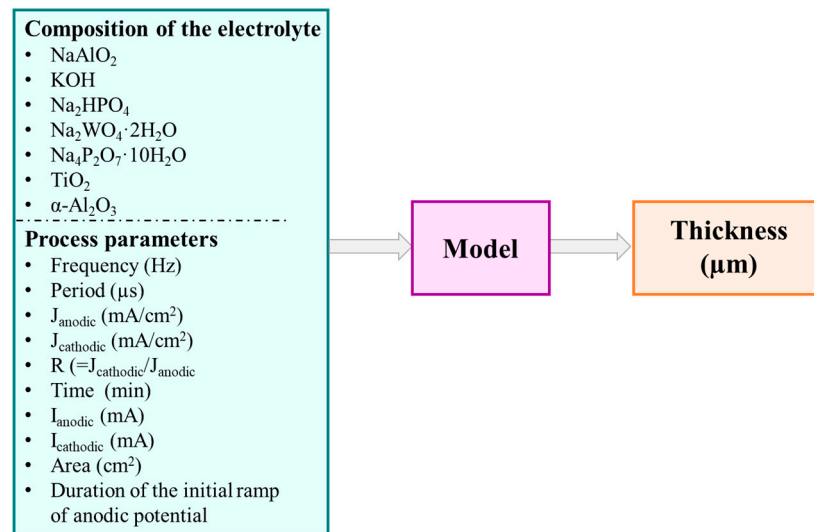
Table 2. Results considering the concentrations of the electrolyte elements and excluding them in the random Forest model.

Model	RMSE	R2
Random forest "without electrolyte"	9.27	0.74
Random forest "with electrolyte"	7.54	0.83

According to preliminary results, an improvement in the RMSE (~20% of the decrease) and R2 (~12% of the increase) coefficients is observed when the electrolyte composition is considered as a variable in the prediction model. This led to the decision to keep the elements composing the electrolyte in the analysis. It is important to note that if a new chemical element is added to the prediction, the model developed in this work must be retrained to incorporate that element. This implies adding an additional variable to the model input dataset. During this analysis, it was observed that the variables "Duration of the initial ramp of anodic potential" and "Area" had no significance in terms of the predictions made and were therefore eliminated for the creation of the models.

A generalized approach for the development of machine learning models is illustrated in Figure 4. This figure shows the input variables, which are composed of all the electrolyte elements present in the dataset, as well as the process conditions applied to the PEO process. In addition, the target variable is represented, which is, in this case, the thickness.

Composition of the electrolyte
- $NaAlO_2$
- KOH
- Na_2HPO_4
- $Na_2WO_4 \cdot 2H_2O$
- $Na_4P_2O_7 \cdot 10H_2O$
- TiO_2
- $\alpha\text{-}Al_2O_3$

Process parameters
- Frequency (Hz)
- Period (µs)
- J_{anodic} (mA/cm^2)
- $J_{cathodic}$ (mA/cm^2)
- R ($=J_{cathodic}/J_{anodic}$)
- Time (min)
- I_{anodic} (mA)
- $I_{cathodic}$ (mA)
- Area (cm^2)
- Duration of the initial ramp of anodic potential

Model

Thickness (µm)

Figure 4. Generalized model of the input and output variables of the machine learning model for thickness prediction obtained in PEO processes.

Table 3 shows the results obtained for the models considered in this study with the hyperparameters of the default algorithms and the optimized hyperparameters. The optimization of the hyperparameters has a positive impact on the final model, improving the two metrics used to evaluate the models, both when using the random forest model and when using the XGBoost algorithm. Thus, a residual analysis helps in the selection of the best model to predict the thickness of the coating on PEO processes.

Table 3. Results of the random forest and XGBoost models using the default hyperparameters and optimized hyperparameters.

Model	RMSE	R2
XGBoost "default"	11.42	0.44
XGBoost "optimized"	8.48	0.81
Random forest "default"	12.07	0.37
Random forest "optimized"	7.93	0.81

XGBoost and random forest exhibit significant improvement in performance after hyperparameter optimization. Optimized models show lower RMSE and higher R2 values compared to their default implementations. Random forest slightly outperforms in terms of RMSE, indicating a small advantage in prediction accuracy.

3.2. Data-Driven Optimization of PEO Processes Assisted by AI Methods

For the evaluation of the created model, a plot of the residuals was generated for both the training subset and the test subset, as shown in Figure 5. The residuals are the difference between the observed values and the predicted values. For 50% of the model core data (thicknesses between 50 and 65 µm), the plot shows that the model performance is accurate in terms of predictions. However, for values closer to the tails of the distribution, the model tends to have a larger dispersion.
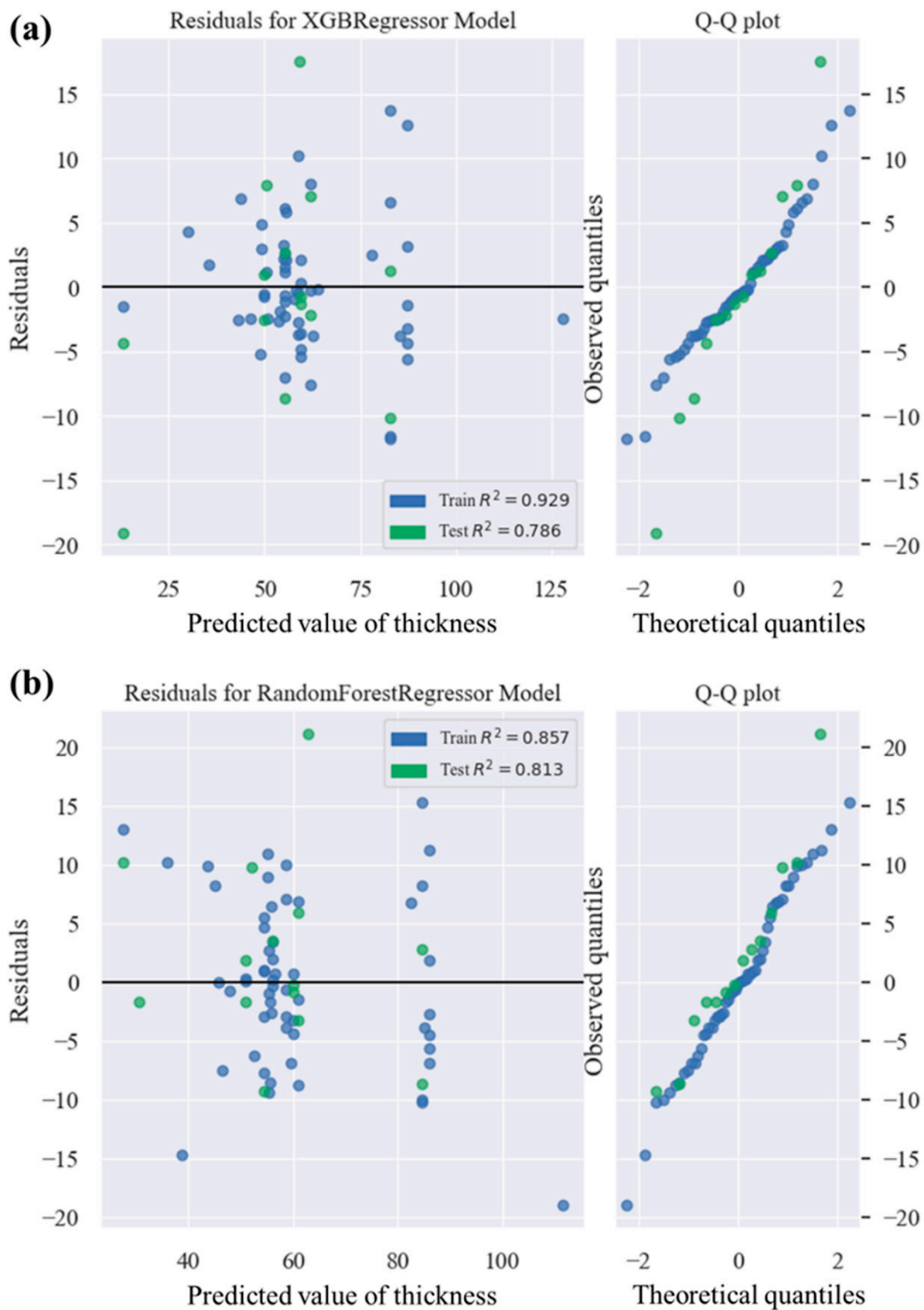
Figure 5. Residuals of the (**a**) XGBRegressor model and (**b**) random forest regressor on the training subset and the test subset for predictions of coating thickness (μm) over PEO processes.

When analyzing the residuals, a certain heteroscedasticity is observed in Figure 5a XGBRegressor, i.e., the higher the thickness values, the more error the predictions tend to have. This heteroscedasticity can have a significant impact on the accuracy and reliability of thickness predictions. To address this situation and improve the quality of predictions, it is suggested to consider increasing the number of samples from the third quartile of the

target variable. Increasing the number of samples can help to reduce heteroscedasticity and improve accuracy in these regions. The Q-Q plot shows that near the mean, the distribution follows a normal distribution, and there is good alignment between the theoretical quantiles and the observed quantiles. In Figure 5b, a more random pattern is shown, suggesting better distributed residuals in this model. Additionally, the Q-Q plot for this model illustrates a closer adherence to normality across the dataset.

Figure 6 shows the SHAP values obtained for each variable. The most influential variables are frequency and titanium dioxide concentration. As can be seen, the thickness tends to increase both at low frequencies and for higher concentrations of titanium dioxide in the electrolyte. In addition, other influential effects are identified, although of lesser relevance, such as the "I_{anodic}" variable, which tends to increase the thickness as its values increase, and the opposite effect is observed in the R variable, where the lower the value, the higher the thickness of the coating.
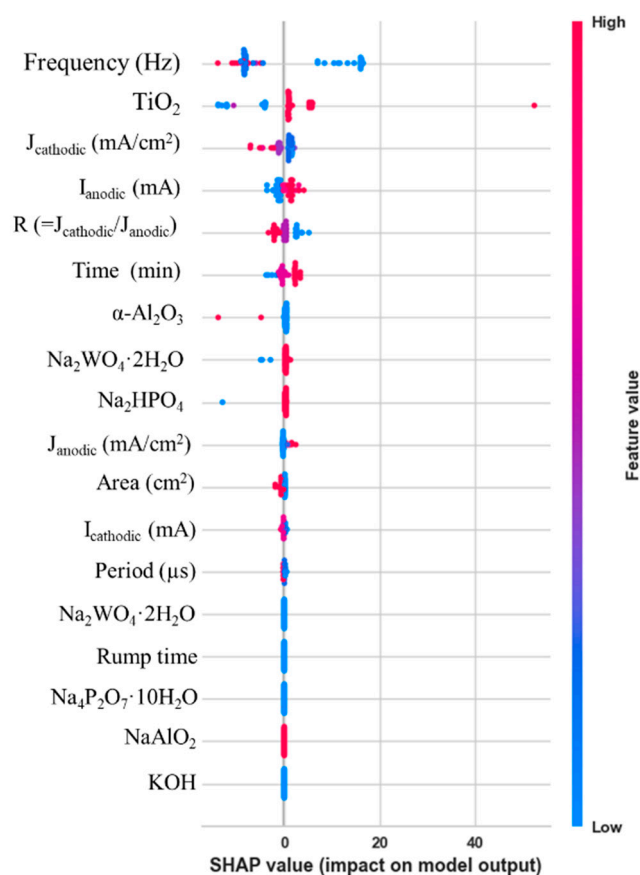
Figure 6. SHAP values obtained for each variable with XGBoost model.

Observing the effects of these variables on the prediction, it is interesting to analyze the possible interactions that could exist between them. Figure 7 shows the independent effect of each of the variables. As suggested in the EDA, the effect of frequency is confirmed, as well as titanium dioxide's effect, with values of 10.0 and 5.4, respectively. In addition, the interaction between these variables is observed with a value of 4.4. Also, to explore the possible interactions that might occur between variables on the prediction, Figure 8 provides a visual representation of the independent effect of each variable under consideration and the possible interactions between the variables. The data in this figure support the existence of a significant effect of both frequency and titanium dioxide on the final thickness prediction. However, the detection of an interaction between these variables, evidenced by a value of 4.4, suggests that the effect of one variable on prediction cannot be considered in isolation. This indicates that the relationship between frequency and titanium dioxide is

more complex than could be inferred by simply examining their individual effects. This information improves the ability to interpret and predict the results.
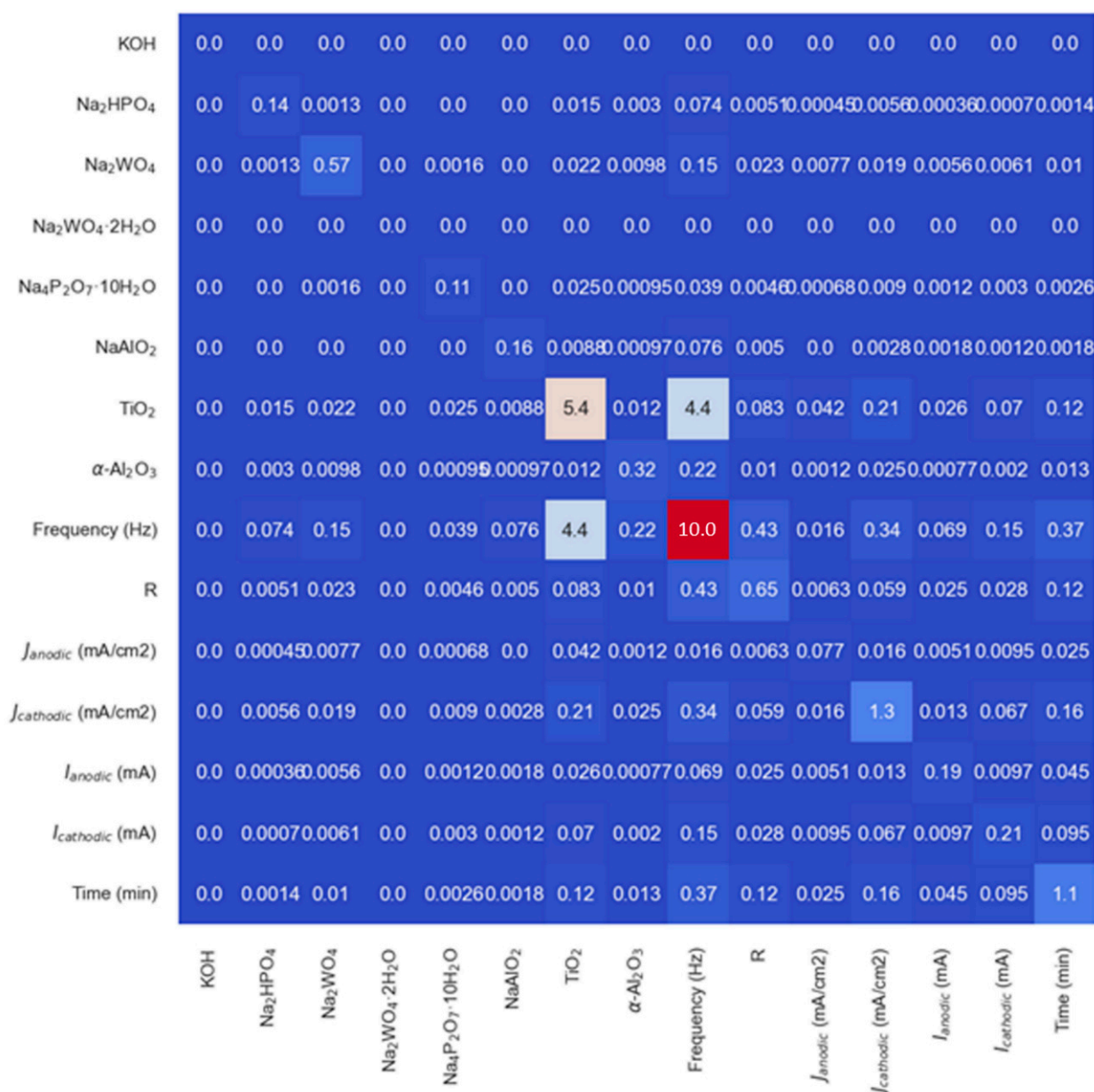
	KOH	Na$_2$HPO$_4$	Na$_2$WO$_4$	Na$_2$WO$_4$·2H$_2$O	Na$_4$P$_2$O$_7$·10H$_2$O	NaAlO$_2$	TiO$_2$	α-Al$_2$O$_3$	Frequency (Hz)	R	J$_{anodic}$ (mA/cm2)	J$_{cathodic}$ (mA/cm2)	I$_{anodic}$ (mA)	I$_{cathodic}$ (mA)	Time (min)
KOH	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Na$_2$HPO$_4$	0.0	0.14	0.0013	0.0	0.0	0.0	0.015	0.003	0.074	0.0051	0.00045	0.0056	0.00036	0.0007	0.0014
Na$_2$WO$_4$	0.0	0.0013	0.57	0.0	0.0016	0.0	0.022	0.0098	0.15	0.023	0.0077	0.019	0.0056	0.0061	0.01
Na$_2$WO$_4$·2H$_2$O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Na$_4$P$_2$O$_7$·10H$_2$O	0.0	0.0	0.0016	0.0	0.11	0.0	0.025	0.000095	0.039	0.0046	0.00068	0.009	0.0012	0.003	0.0026
NaAlO$_2$	0.0	0.0	0.0	0.0	0.0	0.16	0.0088	0.000097	0.076	0.005	0.0	0.0028	0.0018	0.0012	0.0018
TiO$_2$	0.0	0.015	0.022	0.0	0.025	0.0088	5.4	0.012	4.4	0.083	0.042	0.21	0.026	0.07	0.12
α-Al$_2$O$_3$	0.0	0.003	0.0098	0.0	0.00099	0.00097	0.012	0.32	0.22	0.01	0.0012	0.025	0.00077	0.002	0.013
Frequency (Hz)	0.0	0.074	0.15	0.0	0.039	0.076	4.4	0.22	10.0	0.43	0.016	0.34	0.069	0.15	0.37
R	0.0	0.0051	0.023	0.0	0.0046	0.005	0.083	0.01	0.43	0.65	0.0063	0.059	0.025	0.028	0.12
J$_{anodic}$ (mA/cm2)	0.0	0.00045	0.0077	0.0	0.00068	0.0	0.042	0.0012	0.016	0.0063	0.077	0.016	0.0051	0.0095	0.025
J$_{cathodic}$ (mA/cm2)	0.0	0.0056	0.019	0.0	0.009	0.0028	0.21	0.025	0.34	0.059	0.016	1.3	0.013	0.067	0.16
I$_{anodic}$ (mA)	0.0	0.00036	0.0056	0.0	0.0012	0.0018	0.026	0.000077	0.069	0.025	0.0051	0.013	0.19	0.0097	0.045
I$_{cathodic}$ (mA)	0.0	0.0007	0.0061	0.0	0.003	0.0012	0.07	0.002	0.15	0.028	0.0095	0.067	0.0097	0.21	0.095
Time (min)	0.0	0.0014	0.01	0.0	0.0026	0.0018	0.12	0.013	0.37	0.12	0.025	0.16	0.045	0.095	1.1

Figure 7. Independent effect and interactions between input variables.

It can be seen in Figure 9 that when the titanium dioxide concentration is zero, an increase in frequency can lead to an increase in the obtained thickness. Therefore, the ML model developed in this work suggests that frequency plays a significant role in thickness prediction, especially when the titanium dioxide concentration is low.
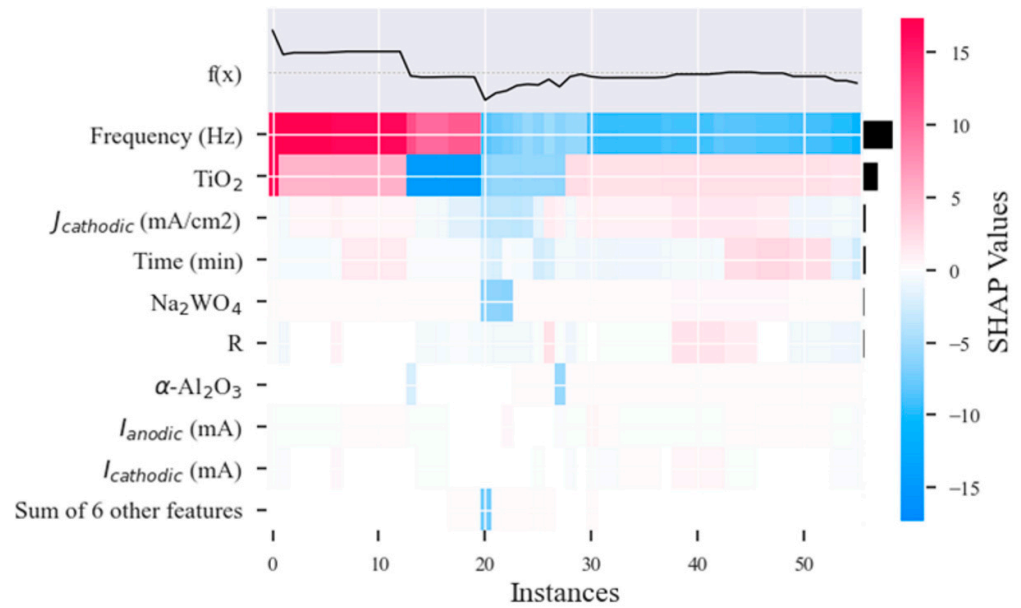
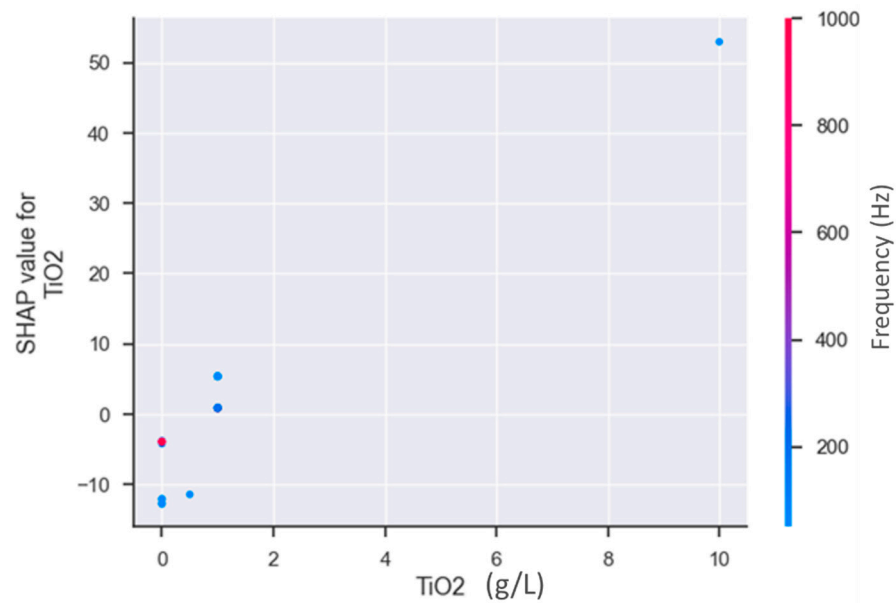**Figure 8.** SHAP values heatmap for random forest regressor.

Figure 9. Graph of dependence between the variables "Frequency" and concentration of titanium dioxide (TiO_2 added to the electrolyte).

Controversial results are still found in the literature regarding the influence of frequency on the thickness of PEO coatings [30]. On the one hand, Martin et al. [8] previously investigated the effect of electrical parameters of the PEO process on the properties of the obtained coatings by building a correlation between microdischarge characteristics and oxide layer features, including morphology, growth rate, and surface roughness. The highest coating growth rate was achieved for the combination of higher current density and the highest current pulse frequency. Ceriani et al. [31] also worked with PEO processes conducted at different electrical regimes to produce functional coatings on titanium substrates and during their experiments observed that when working at higher frequency (1000 Hz), thicker PEO coatings were produced with a uniform structure, providing improved corrosion resistance. On the other hand, Aljohani et al. [32] addressed the impact of processing parameters of bipolar waveform PEO on corrosion resistance of a magnesium alloy. The authors found that the size of micropores of the PEO coatings is interrelated with the duty

cycle and current polarities, and that higher frequencies cause thinner coatings layers, with fewer micropores (denser), and, consequently, improved corrosion resistance. These results could be explained by the fact that the application of lower frequencies induces the appearance of longer and more intense discharges that promote higher temperatures in the coating. The increased molten material generated by the higher temperatures exits more easily through the discharge channels of the coating, leaving larger pores, and promoting faster coating growth [30,33]. Otherwise, the use of higher frequencies would shorten the lifetime of the plasma discharges, decreasing the size of the discharge channels [34]. Pulsed power bipolar power supplies are regulated by multiple parameters, each influencing the properties of the final coatings in distinct ways. Developing machine learning methods to correlate PEO process parameters with the features of the resulting coatings is therefore of paramount importance to develop PEO coatings with advanced features. Machine learning algorithms, by analyzing vast datasets, identifying complex, nonlinear relationships that traditional methods might miss, will allow for optimizing the PEO process with unprecedented precision due to the complexity that all the variables involved might have on PEO coating growth mechanisms and respective features. Regardless of the frequency employed in this study, the relation between titanium dioxide concentration and thickness was studied: an increase in titanium dioxide concentration correlates with an increase in the obtained thickness (Figure 9). This indicates that titanium dioxide concentration exerts a direct and consistent influence on thickness, regardless of how the frequency varies. The importance of carefully considering the interaction between these variables in thickness analysis and prediction should be emphasized. These results provide valuable information for the understanding of the phenomenon studied and may have practical implications in the control and adjustment of thickness through the proper manipulation of frequency and titanium dioxide concentration in the system. Ceriani et al. [31] studied the effect of anatase and rutile TiO_2 microparticle incorporation in PEO coatings produced on titanium substrates. In agreement with the results achieved in the present work, the authors observed that the addition of rutile TiO_2 particles to the PEO electrolyte induced growth in the coatings, with significantly higher thickness and homogeneity.

4. Conclusions

In this work, digital models were successfully developed to establish a correlation between PEO process input variables (e.g., electrolyte composition, duty cycle, period, R factor, current density, negative/positive voltage, etc.) and the selected output variable—thickness. Frequency and TiO_2 concentration were identified as the variables most significantly influencing the thickness of the PEO coatings: the higher the frequency and TiO_2 concentration, the greater the thickness of the PEO coatings obtained. Additionally, it was found that the influence of TiO_2 concentration on thickness is independent of variations in frequency.

The relationship identified in this study between frequency, electrolyte composition, and the thickness of the PEO coatings may represent the first step towards PEO process modeling and digitalization. Developing machine learning methods to correlate PEO process parameters with the features of the resulting coatings is of paramount importance for advancing materials science and industrial applications. This approach can lead to the production of coatings with tailored properties such as enhanced thickness, hardness, corrosion resistance, and wear resistance, thereby meeting specific application requirements while optimizing processes for energy consumption. Furthermore, machine learning can significantly reduce experimental costs and time by predicting optimal parameters, facilitating more efficient and sustainable manufacturing processes. Ultimately, integrating machine learning into PEO research not only accelerates the development of high-performance coatings but also drives innovation in various industries, from aerospace to biomedical engineering.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors Patricia Fernández-López, Sofia A. Alves and Iban Quintana were employed by the company Tekniker. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kaseem, M.; Fatimah, S.; Nashrah, N.; Ko, Y.G. Recent Progress in Surface Modification of Metals Coated by Plasma Electrolytic Oxidation: Principle, Structure, and Performance. *Prog. Mater. Sci.* **2021**, *117*, 100735. [CrossRef]
2. Fernández-López, P.; Alves, S.A.; López-Ortega, A.; San José-Lombera, J.T.; Bayón, R. High Performance Tribological Coatings on a Secondary Cast Al–Si Alloy Generated by Plasma Electrolytic Oxidation. *Ceram. Int.* **2021**, *47*, 31238–31250. [CrossRef]
3. Cheng, Y.; Cao, J.; Mao, M.; Peng, Z.; Skeldon, P.; Thompson, G.E. High Growth Rate, Wear Resistant Coatings on an Al–Cu–Li Alloy by Plasma Electrolytic Oxidation in Concentrated Aluminate Electrolytes. *Surf. Coat. Technol.* **2015**, *269*, 74–82. [CrossRef]
4. Fernández-López, P.; Alves, S.A.; Azpitarte, I.; San-José, J.T.; Bayón, R. Corrosion and Tribocorrosion Protection of Novel PEO Coatings on a Secondary Cast Al-Si Alloy: Influence of Polishing and Sol-Gel Sealing. *Corros. Sci.* **2022**, *207*, 110548. [CrossRef]
5. Lu, X.; Mohedano, M.; Blawert, C.; Matykina, E.; Arrabal, R.; Kainer, K.U.; Zheludkevich, M.L. Plasma Electrolytic Oxidation Coatings with Particle Additions—A Review. *Surf. Coat. Technol.* **2016**, *307*, 1165–1182. [CrossRef]
6. Dehnavi, V.; Luan, B.L.; Shoesmith, D.W.; Liu, X.Y.; Rohani, S. Effect of Duty Cycle and Applied Current Frequency on Plasma Electrolytic Oxidation (PEO) Coating Growth Behavior. *Surf. Coat. Technol.* **2013**, *226*, 100–107. [CrossRef]
7. Sikdar, S.; Menezes, P.V.; Maccione, R.; Jacob, T.; Menezes, P.L. Plasma Electrolytic Oxidation (Peo) Process—Processing, Properties, and Applications. *Nanomaterials* **2021**, *11*, 1375. [CrossRef] [PubMed]
8. Martin, J.; Melhem, A.; Shchedrina, I.; Duchanoy, T.; Nominé, A.; Henrion, G.; Czerwiec, T.; Belmonte, T. Effects of Electrical Parameters on Plasma Electrolytic Oxidation of Aluminium. *Surf. Coat. Technol.* **2013**, *221*, 70–76. [CrossRef]
9. Mohedano, M.; Lopez, E.; Mingo, B.; Moon, S.; Matykina, E.; Arrabal, R. Energy Consumption, Wear and Corrosion of PEO Coatings on Preanodized Al Alloy: The Influence of Current and Frequency. *J. Mater. Res. Technol.* **2022**, *21*, 2061–2075. [CrossRef]
10. Matykina, E.; Arrabal, R.; Skeldon, P.; Thompson, G.E. Optimisation of the Plasma Electrolytic Oxidation Process Efficiency on Aluminium. *Surf. Interface Anal.* **2010**, *42*, 221–226. [CrossRef]
11. Zhang, X.M.; Tian, X.B.; Yang, S.Q.; Gong, C.Z.; Fu, R.K.Y.; Chu, P.K. Low Energy-Consumption Plasma Electrolytic Oxidation Based on Grid Cathode. *Rev. Sci. Instrum.* **2010**, *81*, 103504. [CrossRef] [PubMed]
12. Ntomprougkidis, V.; Martin, J.; Nominé, A.; Henrion, G. Sequential Run of the PEO Process with Various Pulsed Bipolar Current Waveforms. *Surf. Coat. Technol.* **2019**, *374*, 713–724. [CrossRef]
13. Nominé, A.; Nominé, A.V.; Braithwaite, N.S.J.; Belmonte, T.; Henrion, G. High-Frequency-Induced Cathodic Breakdown during Plasma Electrolytic Oxidation. *Phys. Rev. Appl.* **2017**, *8*, 031001. [CrossRef]
14. Rogov, A.B.; Yerokhin, A.; Matthews, A. The Role of Cathodic Current in Plasma Electrolytic Oxidation of Aluminum: Phenomenological Concepts of the "Soft Sparking" Mode. *Langmuir* **2017**, *33*, 11059–11069. [CrossRef] [PubMed]
15. Rogov, A.B.; Huang, Y.; Shore, D.; Matthews, A.; Yerokhin, A. Toward Rational Design of Ceramic Coatings Generated on Valve Metals by Plasma Electrolytic Oxidation: The Role of Cathodic Polarisation. *Ceram. Int.* **2021**, *47*, 34137–34158. [CrossRef]
16. Geebarowski, W.; Pietrzyk, S. Influence of the Cathodic Pulse on the Formation and Morphology of Oxide Coatings on Aluminium Produced by Plasma Electrolytic Oxidation. *Arch. Metall. Mater.* **2013**, *58*, 241–245. [CrossRef]
17. Godja, N.; Hansal, W.E.G.; Mann, R.; Kleber, C.; Hansal, S. Pulsed Plasma Electrolytic Oxidation Processes for Aeronautical Applications and Their Technical Application. *Trans. IMF* **2013**, *91*, 321–329. [CrossRef]

18. Tagirova, K.; Aubakirova, V.; Vulfin, A. Neural Network Control System for the Process of Plasma Electrolytic Oxidation. In *Advances in Automation V—RusAutoCon 2023*; Lecture Notes in Electrical Engineering; Springer: Cham, Switzerland, 2024; Volume 1130, pp. 321–333. [CrossRef]

19. Finke, A.; Escobar, J.; Munoz, J.; Petit, M. Prediction of Salt Spray Test Results of Micro Arc Oxidation Coatings on AA2024 Alloys by Combination of Accelerated Electrochemical Test and Artificial Neural Network. *Surf. Coat. Technol.* **2021**, *421*, 127370. [CrossRef]

20. Rodriguez, L.; Paris, J.Y.; Denape, J.; Delbé, K. Micro-Arcs Oxidation Layer Formation on Aluminium and Coatings Tribological Properties—A Review. *Coatings* **2023**, *13*, 373. [CrossRef]

21. Varol, T.; Canakci, A.; Ozsahin, S.; Erdemir, F.; Ozkaya, S. Artificial Neural Network-Based Prediction Technique for Coating Thickness in Fe-Al Coatings Fabricated by Mechanical Milling. *Part. Sci. Technol.* **2018**, *36*, 742–750. [CrossRef]

22. Baptista, M.; Mishra, M.; Henriques, E.; Prendinger, H. Using Explainable Artificial Intelligence to Interpret Remaining Useful Life Estimation with Gated Recurrent Unit. *Preprint* **2021**, 5–7. [CrossRef]

23. Hart, S. Shapley Value. In *Game Theory*; Eatwell, J., Milgate, M., Newman, P., Eds.; The New, Palgrave; Pal-grave Macmillan: London, UK, 1989. [CrossRef]

24. Tsai, D.S.; Chou, C.C. Review of the Soft Sparking Issues in Plasma Electrolytic Oxidation. *Metals* **2018**, *8*, 105. [CrossRef]

25. Cheng, Y.; Teng, F.; Cheng, Y. A Systematic Study of the Role of Cathodic Polarization and New Findings on the Soft Sparking Phenomenon from Plasma Electrolytic Oxidation of an Al-Cu-Li Alloy. *J. Electrochem. Soc.* **2022**, *169*, 071505. [CrossRef]

26. Gębarowski, W.; Pietrzyk, S. Growth Characteristics of the Oxide Layer on Aluminium in the Process of Plasma Electrolytic Oxidation. *Arch. Metall. Mater.* **2014**, *59*, 407–411. [CrossRef]

27. Martin, J.; Nominé, A.; Brochard, F.; Briançon, J.L.; Noël, C.; Belmonte, T.; Czerwiec, T.; Henrion, G. Delay in Micro-Discharges Appearance during PEO of Al: Evidence of a Mechanism of Charge Accumulation at the Electrolyte/Oxide Interface. *Appl. Surf. Sci.* **2017**, *410*, 29–41. [CrossRef]

28. Hakimizad, A.; Raeissi, K.; Golozar, M.A.; Lu, X.; Blawert, C.; Zheludkevich, M.L. Influence of Cathodic Duty Cycle on the Properties of Tungsten Containing Al2O3/TiO2 PEO Nano-Composite Coatings. *Surf. Coat. Technol.* **2018**, *340*, 210–221. [CrossRef]

29. Troughton, S.C.; Clyne, T.W. Cathodic Discharges during High Frequency Plasma Electrolytic Oxidation. *Surf. Coat. Technol.* **2018**, *352*, 591–599. [CrossRef]

30. Toulabifard, A.; Rahmati, M.; Raeissi, K.; Hakimizad, A.; Santamaria, M. The Effect of Electrolytic Solution Composition on the Structure, Corrosion, and Wear Resistance of PEO Coatings on AZ31 Magnesium Alloy. *Coatings* **2020**, *10*, 937. [CrossRef]

31. Ceriani, F.; Casanova, L.; Massimini, L.; Brenna, A.; Ormellese, M. TiO2 Microparticles Incorporation in Coatings Produced by Plasma Electrolytic Oxidation (PEO) on Titanium. *Coatings* **2023**, *13*, 1718. [CrossRef]

32. Aljohani, T.; Aljadaan, S.; Rubayan, M.; Khoshnaw, F. Impact of Processing Parameters in Plasma Electrolytic Oxidation on Corrosion Resistance of Magnesium Alloy Type AZ91. *Eng. Rep.* **2021**, *4*, e12459. [CrossRef]

33. Lv, G.; Chen, H.; Gu, W.; Li, L.; Niu, E. Effects of Current Frequency on the Structural Characteristics and Corrosion Property of Ceramic Coatings Formed on Magnesium Alloy by PEO Technology. *J. Mater. Process. Technol.* **2008**, *8*, 9–13. [CrossRef]

34. Bala Srinivasan, P.; Liang, J.; Balajee, R.G.; Blawert, C.; Störmer, M.; Dietzel, W. Effect of Pulse Frequency on the Microstructure, Phase Composition and Corrosion Performance of a Phosphate-Based Plasma Electrolytic Oxidation Coated AM50 Magnesium Alloy. *Appl. Surf. Sci.* **2010**, *256*, 3928–3935. [CrossRef]